TEL AVIV UNIVERSITY אוניברסיטת תל-אביב

SACKLER FACULTY OF EXACT SCIENCES

SCHOOL OF COMPUTER SCIENCE

AUTOMATIC ANALYSIS OF VIBRO-ACOUSTIC HEART SIGNALS: COMBINING SIGNAL PROCESSING AND COMPUTATIONAL LEARNING TECHNIQUES FOR NON-INVASIVE EVALUATION AND MONITORING OF CARDIAC FUNCTION

THESIS SUBMITTED FOR THE DEGREE
'DOCTOR OF PHILOSOPHY'

BY

GUY AMIT

THIS WORK WAS CARRIED OUT UNDER THE SUPERVISION OF

PROF. NATHAN INTRATOR AND PROF. NOAM GAVRIELY

SUBMITTED TO THE SENATE OF TEL-AVIV UNIVERSITY

APRIL 2009

This work is dedicated to my parents, Rachel and Arie Amit,
who gave me their rock-solid support in every choice I ever made.

In memory of my beloved grandparents Yekutiel and Lea Braud.

The intellect of man is forced to choose
Perfection of the life, or of the work,
And if it take the second must refuse
A heavenly mansion, raging in the dark.

When all that story's finished, what's the news?
In luck or out the toil has left its mark:
That old perplexity an empty purse,
Or the day's vanity, the night's remorse.

W.B. Yeats, 'The Choice'

עַל רוּחַ הָאָדָם נִגְזַר לִבְחֹר
בֵּין יְצִירָה וּבֵין חַיִּים בְּלִי דֹּפִי,
וְאִם בְּיִצִירָה יִבְחַר, לִבְעֹט
בְּנְאוֹת שָׁמַיִים, חוֹרֵק שִׁנָּיו בָּאֹפֶל.

וּבְסוֹף כָּל הַסִּפּוּר, מַה נִּשְׁתַּנָּה?
זֵעַת-אַפָּיו תַּשְׁאִיר חוֹתָם מִמֵּילָא:
אוֹתוֹ אָבְדַן-עֵצוֹת – הַכִּיס הָרֵיק,
אוֹתוֹ רַהַב בַּיּוֹם, נֹחַם בַּלַּיְלָה.

ו.ב. ייטס, 'הבחירה'
תרגום שמעון זנדבנק

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisors, Prof. Nathan Intrator and Prof. Noam Gavrierly, who showed me the road into the fascinating world of physiological signal processing and guided me in my scientific exploration of this road.

I would also like to thank my collaborators, Dr. Jonathan Lessick and Dr. Khuloud Shukha for their kind assistance in designing and executing the clinical experiments.

I owe many thanks to my colleagues from the Neural Computation and Signal Processing lab and to my dear friends from BSP, for many fruitful academic discussions and numerous wise and helpful advices.

During the period of my PhD studies, my wife Raya and I formed our own family nest, giving birth to our children Tamar, Ella and Ophir, an endless source of happiness and wonder. I would like to thank Raya for her invaluable love, friendship and wisdom in our long joint journey.

The Burma road, linking Burma with China.

ABSTRACT

The human heart is an electro-mechanical system that cyclically pumps blood throughout the body. The mechanical processes of the cardiac cycle generate low-frequency vibrations that are received on the chest wall as vibro-acoustic signals. These sound and infra-sound signals carry significant information about the underlying functioning of the cardiovascular system and its interaction with other body systems, such as the pulmonary system. The purpose of this work was to develop signal analysis and pattern recognition techniques for automatic extraction of meaningful physiological information from vibro-acoustic heart signals. These techniques were used to gain new insights about the relationship between vibro-acoustic signals and the mechanical function of the heart, and to assess the potential benefit for non-invasive monitoring of cardiac function.

We describe and study the properties of a wide range of analysis methods of vibro-acosutic heart signals, including preprocessing, event and component segmentation, time-frequency signal representation, multi-cycle alignment, robust feature extraction, morphological clustering and classification. In particular, we introduce a computational signal analysis framework for identifying distinct morphologies of heart sounds and classifying them into physiological states. We applied the analysis methods on heart sound datasets, acquired during controlled alternations of the physiological conditions. Analysis of heart sound signals acquired during pharmacological stress test was used to characterize the pattern of spectral changes of S1, which was correlated with 'gold-standard' echocardiographic indices of cardiac contractility. The analysis framework was able to identify this pattern and to accurately predict the stress level from the morphology of S1. Analysis of heart sound signals acquired during alternations of the respiratory activity revealed the temporal and morphological changes induced in S1 and S2 by the pulmonary system, and showed that the respiratory phase, the resistive-load and the instantaneous pressure can be estimated from the morphology of the heart sounds.

The results suggest that quantitative automated heart sound analysis using modern computational techniques may provide a new non-invasive technology for continuous cardiopulmonary monitoring and improved detection of mechanical dysfunctions caused by cardiovascular and cardiopulmonary diseases.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACG	-	Apexcardiogram
AR	-	Auto-Regressive
ASI	-	Acoustic Spectral Index
AV	-	Atrio-Ventricular
AVI	-	Acoustic Variability Index
CC	-	Correct Classification
CO	-	Cardiac Output
COPD	-	Chronic Obstructive Pulmonary Disease
CP	-	Carotid Pulse
CT	-	Computerized Tomography
CW	-	Continuous Wave (Doppler)
CWD	-	Choi-Williams Distribution
DA	-	Discriminant Analysis
DN	-	Dicrotic Notch
DSE	-	Dobutamine Stress Echocardiography
DTW	-	Dynamic Time Warping
ECG	-	Electrocardiogram
EDV	-	End-Diastolic Volume
ET	-	Ejection Time
FDR	-	False Discovery Rate
FFT	-	Fast Fourier Transform
FT	-	Filling Time
HF	-	Heart Failure
HMM	-	Hidden Markov Model
HR	-	Heart Rate
HSPRS	-	Heart Signals with Respiratory Pressure (dataset)
HSDSE	-	Heart Signals during Dobutamine Stress Echocardiography (dataset)
ICG	-	Impedance Cardiography
IIR	-	Infinite Impulse Response (filter)

IVRT - Isovolumetric Relaxation Time
KNN - K-Nearest Neighbor
LCS - Longest Common Subsequence
LV - Left Ventricle
LVET - Left-Ventricular Ejection Time
MRI - Magnetic Resonance Imaging
PAC - Pulmonary Artery Catheterization
PCA - Principal Component Analysis
PEP - Pre-Ejection Period
PSA - Phase-Shift Averaging
PSSR - Peak Systolic Strain Rate
PSV - Peak Systolic Velocity
PWV - Pulse-Wave Velocity
RID - Reduced Interference Distribution
RV - Right Ventricle
SD - Standard Deviation
SNR - Signal to Noise Ratio
ST - S-Transform
STFT - Short-Time Fourier Transform
SV - Stroke Volume
SVM - Support Vector Machine
TDI - Tissue-Doppler Imaging
TFR - Time-Frequency Representation
WVD - Wigner-Ville Distribution

CHAPTER 1
INTRODUCTION TO VIBRO-ACOUSTIC HEART SIGNALS

1.1 Physiological and clinical background

1.1.1 The heart and circulation

The human heart is a mechanical system whose primary function is to pump blood throughout the body in order to provide adequate perfusion of organs. The anatomical structure of the heart consists of two atria for collecting blood from the veins and two ventricles for pumping the blood out to the arteries (Figure 1a). The left side of the heart is responsible for pumping oxygenated blood through the aorta into the systemic blood circulation. The deoxygenated blood flows back from the peripheral body tissues through the veins, and is collected in the right atrium. The right side of the heart pushes deoxygenated blood through the pulmonary artery into the pulmonary circulation and towards the lungs. The returning oxygen-rich blood is collected in the left atrium. The blood flow from the atria to the ventricles is regulated by a pair of atrio-ventricular (AV) valves, named mitral and tricuspid valves. A second pair of semilunar valves, named aortic and pulmonary valves, controls the flow from the ventricles to the arteries [1].

1.1.2 The mechanical cardiac cycle

The periodical activity of the heart is controlled by an electrical conducting system, triggering electrical action potentials that excite the muscle cells and cause the mechanical contraction of the heart chambers (Figure 1b). The contraction phase of the heart is called systole. The systole begins with the contraction of the atria, shortly followed by the isovolumetric contraction of the ventricles. During isovolumetric contraction, the ventricular pressure rises without change in the blood volume. The high ventricular pressure causes the atrio-ventricular valves to be closed. When ventricular pressure exceeds the arterial pressure, the semilunar valves are opened, and blood is ejected from the ventricles to the arterial circulation. At the end of the ejection phase, with the fall of ventricular pressure, the semilunar valves are closed, indicating the end of systole and beginning of diastole. Early diastole is an isovolumetric relaxation period, in which ventricular pressure drops with blood volume changes. As soon as atrial pressure

exceeds ventricular pressure, the atrio-ventricular valves are opened to start the filling phase. Filling is a passive process of blood flow from the atria to the ventricles. The flow rate is initially rapid, and is reduced in late diastole. Finally, atrial contraction contributes an additional small volume of blood to the ventricular filling [2].

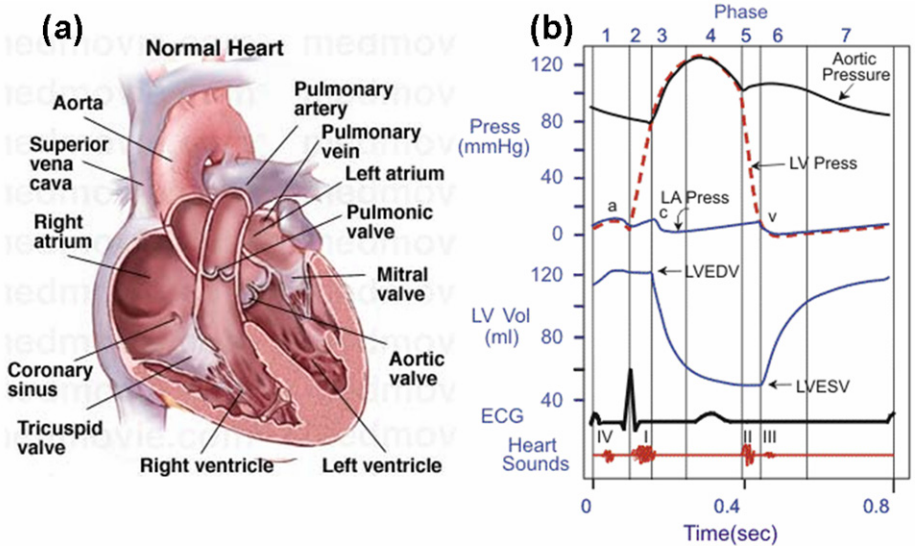

Figure 1: The anatomy of a normal heart (a) and the dynamics of aortic, left-ventrcular (LV), left-atrial (LA) pressures, LV volume, ECG and heart sound signals throughout the systolic (1-4) and diastolic (4-7) phases of the cardiac cycle (b).

1.1.3 Assessment of cardiac function

The mechanical operation of the cardiovascular system is governed by a complex interplay between pressure gradients, determined by the contraction force of the myocardial cells, the dynamics of blood flow and the compliance of cardiac chambers and blood vessels. One way to assess systolic function is to measure the ventricular stroke volume (SV), the volume of blood that was ejected during systole. SV is regulated by three primary mechanisms, namely preload, inotropy and afterload [2]. Preload is the initial stretching of the cardiac muscle cells (myocytes) prior to contraction. Increased myocyte stretching causes an increase in force generation during contraction. This principle, called 'The Frank-Starling mechanism', enforces a direct relation between the stroke volume and the extent of venous return, or ventricular filling. Preload is therefore often estimated by the ventricular end-diastolic volume. Cardiac muscle has a unique ability to alter its inotropy, or contractility, which is a preload-independent change in the rate of force or pressure generation. Inotropy, regulated by the autonomic nervous system, is directly related to the stroke volume ejected by the contraction force. Finally,

afterload is the load against which the heart ejects blood, which is closely related to the aortic pressure. When the aortic pressure is increased, the stroke volume is consequently decreased.

1.1.4 Cardio-pulmonary interaction

The pulmonary system plays an important part in modulating the cardiovascular mechanical activity [3, 4]. During respiration, the pressure between the lungs and the chest wall, named pleural pressure, is altered, causing changes in the arterial pressure and venous return. During inspiration, the pressure gradient from the extra-thoracic regions to the right atrium increases due to the lowered pleural pressure, causing an increased blood filling of the right ventricle (RV). The increased right-ventricular end diastolic volume (EDV) leads to an increased RV stroke volume by the Frank-Starling mechanism. The distended RV causes a leftward motion of the inter-ventricular septum, which reduces the compliance of the left ventricle (LV) by physical compression (ventricular interdependence) and leads to a reduced LV filling. At the same time, the distending lung and its circulatory volume tend to reduce the pressure gradient and flow from the pulmonary veins to the left ventricle. Furthermore, the transmural diastolic aortic pressure, which is the LV afterload, increases. These additive effects result in a decrease in LV stroke volume. The opposed process occurs during expiration, in which RV-SV decreases and LV-SV increases [5]. Respiration has also been shown to modulate the duration of the systolic and diastolic time intervals of the cardiac cycle [6]. The effects of the respiratory cycle and changes in intra-thoracic pressure on cardiac function are well known clinically in the form of “pulsus paradoxus” as a sign of asthma severity and in assessing the need for fluids transfusion in critically ill patients [7].

1.1.5 Clinical aspects – heart disease

The mechanical functioning of the heart may be impaired by cardiovascular and cardiopulmonary diseases. Heart failure (HF) is a clinical syndrome defined as the inability of the ventricle to fill with or eject enough blood in order to deliver adequate oxygen supply to peripheral tissues and organs [8]. HF is a major public health problem worldwide, with estimated 15 million new cases each year. HF is especially common in older populations, being the leading cause of hospitalization in patients over 65 years old.

HF is usually caused by coronary artery disease, chronic hypertension, cardiomyopathy (a primary disease of heart muscle cells) or valvular heart disease. The changes in cardiac function associated with heart failure results in a decreased stroke volume due to systolic or diastolic dysfunction. Systolic dysfunction occurs as a result of loss of intrinsic contractility or reduced viability of contracting muscle following acute myocardial infarction. Diastolic dysfunction occurs when the ventricle becomes more stiff and less compliant, which impairs its diastolic blood filling. The body has compensatory mechanisms, aiming to augment stroke volume by increasing the ventricular preload. The higher preload is achieved by an increase in the ventricular end-diastolic pressure and the blood volume. However, increased blood volume is often deleterious, because it raises venous return and can lead to accumulation of fluid in the lungs, a condition termed pulmonary congestion. HF patients with pulmonary congestion suffer from shortness of breath initially during exertion, and as disease progresses – in rest conditions as well.

1.1.6 Cardiac monitoring

Electrocardiography, the recording of the surface electrical potentials of the heart, remains the most common technology for continuous monitoring of cardiac activity, more than 100 years after its introduction by Willem Einthoven. The electrocardiogram (ECG) signals provide reliable indications for electrical dysfunctions related to the heart's pacing and conduction system, as well as for conditions of myocardial ischemia, in which lack of oxygen causes changes in the cellular bioenergetics and electrical conduction of the myocytes. Mechanical dysfunctions which are not accompanied by electrical changes are not reflected in the electrocardiogram. In addition, patients with chronic heart disease such as heart failure often have enduring ECG abnormalities [9], which reduce the efficacy of ECG monitoring in detecting disease exacerbation.

Existing medical technologies for assessing mechanical cardiac function include invasive catheterization, echocardiography, and impedance cardiography [10]. Pulmonary artery catheterization (PAC), considered as the gold-standard of hemodynamic measurements, is utilized for critically-ill patients in intensive care units. A balloon-tipped catheter is inserted through a peripheral vein into the right atrium, right ventricle and pulmonary artery. Balloon inflation occludes a branch of the pulmonary artery, and the measured pressure provide a good estimate of the left-atrial pressure, which is also the end-diastolic

left-ventricle pressure, or ventricular preload. The routine usage of this highly-invasive procedure is controversial, due it its potential complications [11].

Echocardiography, cardiac ultrasound imaging, is the most widely-used technology for evaluation of mechanical cardiac function by qualitative evaluation of cardiac wall motion, blood flow and valve operation. Hemodynamic parameters such as stroke volume, ejection fraction and left-ventricular pressure can be also quantitatively estimated by transthoracic 2D and Doppler echocardiography [12]. More accurate measurements of stroke volume and ventricular preload can be obtained by an invasive technique of transesophageal echocardiography. However, these measurements cannot be obtained continuously for monitoring purposes, as they require manual placement and adjustments of the transducer by an expert operator.

Impedance cardiography (ICG) is an emerging technology, based on measuring the resistance to transmission of a small electrical current throughout the chest area. The measured changes in the impedance are converted into measurements of blood volume changes during the cardiac cycle. Using ICG, measurements of stroke volume, vascular resistance and systolic time intervals can be obtained continuously and non-invasively. ICG has been shown to be beneficial in a variety of clinical conditions [13].

Respiratory functions can be continuously monitored by tracking thoracic and abdominal motion, tidal volume, respiratory frequency and oxygenation level of arterial and venous blood [14]. Diseases that involve airway obstruction, such as asthma and COPD, are monitored by spirometry, which measures the volume of inspired and expired air, or by a peak-flow meter device, which measures the peak-expiratory air flow rate.

1.2 Vibro-acoustic heart signals

1.2.1 History of heart auscultation and phonocardiography

The medical practice of applying the ear to the chest for direct heart auscultation is very ancient, appearing already in Hippocratic writings from 400 BC [15]. The invention of the first auscultation device, the stethoscope, is referred to the French physician René Laennec, who constructed a wooden tube monaural device in 1816 (Figure 2a). In the middle of the 19th century, George Cammann and Arthur Leared introduced the binaural stethoscope, which is the basic model of stethoscope still in use today. In 1907 Willem

Einthoven reported on the use of the string galvanometer for graphic recording of heart sounds, thus laying the grounds of phonocardiography (Figure 2b). Phonocardiography was later developed in various aspects including low-frequency phonocardiogram by Luisada [16] and spectral phonocardiography by McKusick [17]. With the vast developments in microelectronics and computing devices, electronic and later digital stethoscopes emerged. These devices provide more sophisticated control over the aural and visual properties of the acquired heart sound signals, with some automatic diagnostic capabilities [18]. However, the utilization of these devices in the common medical practice is still quite limited [19].

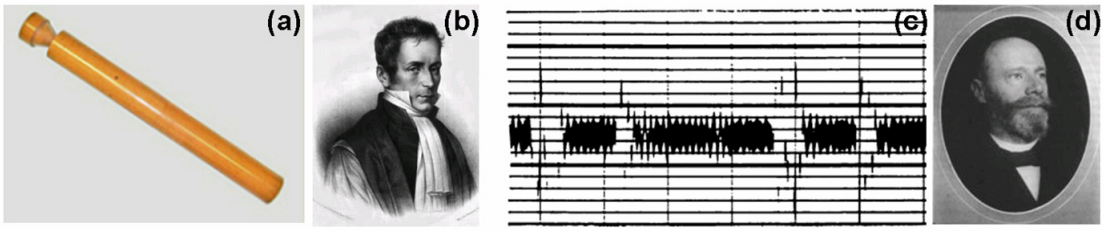

Figure 2: Rene Laennec (b), inventor of the first stethoscope (a) and Willem Einthoven (d), pioneer of phnocardiography. Displayed phonocardiogram (c) was recorded with original Einthoven string galvanometer at Massachusetts General Hospital [20].

1.2.2 Heart sounds – origin and characteristics

Heart sounds are produced by the vibrations of the cardiohemic system, composed of the blood, heart walls and valves. The vibrations are triggered by the acceleration and deceleration of blood due to abrupt mechanical events of the cardiac cycle [21, 22]. The portion of these vibrations which is within the audible frequency range is termed 'sound' while lower-frequency energy (infra-sound) is received as heart 'palpitations' (Figure 3). The two major audible sounds in a normal cardiac cycle are the first and second heart sounds, S1 and S2 (Figure 4a). S1 occurs at the onset of ventricular contraction, while S2 is heard at the end of systole, following the closure of the semilunar valves. In some cases, a third low-frequency sound, S3, may be heard at the beginning of the diastole, during the rapid filling of the ventricles. A fourth heart sound, S4, may be heard in late diastole during atrial contraction. Another type of heart sound is murmurs. These sounds are high-frequency, noise-like sounds that are heard between the two major heart sounds during systole or diastole. Existence of audible S3, S4 or murmurs may be an indication of cardiovascular dysfunction.

S1 and S2 are both low-frequency, multi-component transient signals. S1 is usually the longest and loudest heart sound. It is considered to have four sub-components (Figure 4b): the initial vibrations occur when the first contraction of the ventricle moves blood towards the atria, closing the AV-valves. The second component is caused by the abrupt tension of the closed AV-valves, decelerating the blood. The third component involves oscillation of blood between the root of the aorta and the ventricular walls, and the fourth component represents the vibrations caused by turbulence in the ejected blood flowing to the arteries. S2 usually has a shorter duration and higher frequency than S1. It has aortic and pulmonary sub-components (A2 and P2), associated with the closure of each of the two semilunar valves.

The complex interplay between pressure gradients in atria, ventricles and arteries affects the timing, magnitude and morphology of the produced heart sounds. The amplitude of S1 has been shown to be related to the degree of separation of the mitral valve leaflets, determined by the relative timing of the left atrial and ventricular systole (P-R interval of the ECG). Left-ventricular contractility was also shown to be an independent factor determining the amplitude of S1 [23, 24]. The amplitude of the aortic component of S2 has been shown to be closely related to the peak rate of development of the aortic-to-left ventricular differential pressure gradient [25]. The dyssynchrony between the dynamics of the left and right sides of the heart, which is normally associated with respiratory activity, has well-established effects of widening the delay between the sound components, thus producing a split morphology of either S1 or S2 [26, 27].

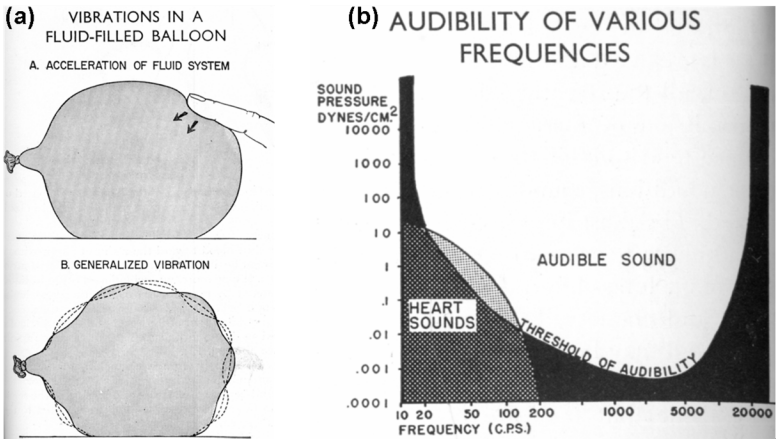

Figure 3: Cardiohemic theory of heart sound generation, described by Rushmer [21] as the vibrations in a fluid-filled balloon (a), and the frequency range of heart sounds, which only partially overlap with the range of audible sounds (b).

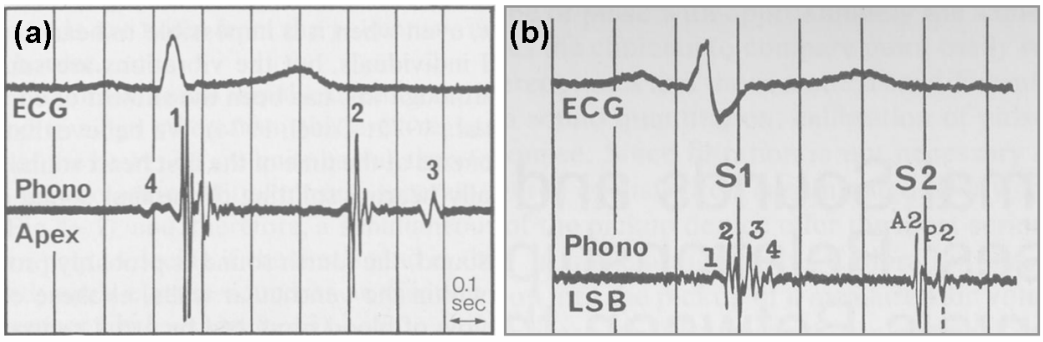

Figure 4: A phonocardiogram recording of a single heart beat, showing the two major heart sounds S1 and S2, as well as S3 and S4 (a), and a detailed description of the inner structure of S1 and S2 (b), showing their subcomponents [28].

1.2.3 Infrasound signals

The vibratory energy produced by the cardiovascular system, which is below audibility threshold, can be registered by the appropriate transducer as infrasound signals. These vibrations are of large amplitude, and can be appreciated to some extent by the tactile sense. Outward movements of the heart or large blood vessels toward the transducer are recorded as upward deflections of the signal, and retractions are registered as downward deflections. The carotid pulse signal reflects small volume changes of the artery with each heartbeat. The signal normally begins to rise abruptly with aortic ejection and reaches an initial peak (percussion wave) when the ejection is probably at its maximum (Figure 5a). There usually follows a plateau or secondary wave (tidal wave) late in systole. This wave was suggested to represent primarily the reflected pulse wave returning from the upper body. Next, the wave smoothly falls to a point termed the 'dicrotic notch', a notch produced by abrupt completion of the aortic valve closure. In early diastole, there is often a small, positive wave, designated the 'dicrotic wave', which is probably an effect of the reflected pulse from the lower body [28]. The apical pulse signal, termed 'apexcardiogram', reflects the low-frequency vibrations transmitted by the heart and great vessels to the chest wall at the point of maximal impulse. The signal is thought to mirror events caused by the left ventricle [28]. The normal apexcardiogram (Figure 5b) registers a small wave, termed the 'a' wave, at the time of left-atrial contraction, and then rises abruptly at the onset of the isovolumetric ventricular contraction ('c' point). With the onset of blood ejection from the ventricle into the aorta, this initial rise reverses abruptly and forms a peak termed the 'e' point. Following the

onset of ejection, there is a sudden down-sloping that continues to its nadir, termed the 'o' point, which is registered approximately at the time of mitral valve opening. Aortic valve closure can sometimes be observed as a brief notch on the down-sloping segment. With early rapid ventricular filling, the curve rises steeply in an ascent called the ‘rapid filling wave’, followed by a gentler incline called the ‘slow filling wave’. If ventricular diastole is sufficiently long, there follows a late diastolic plateau, called ‘stasis’ [28, 29].

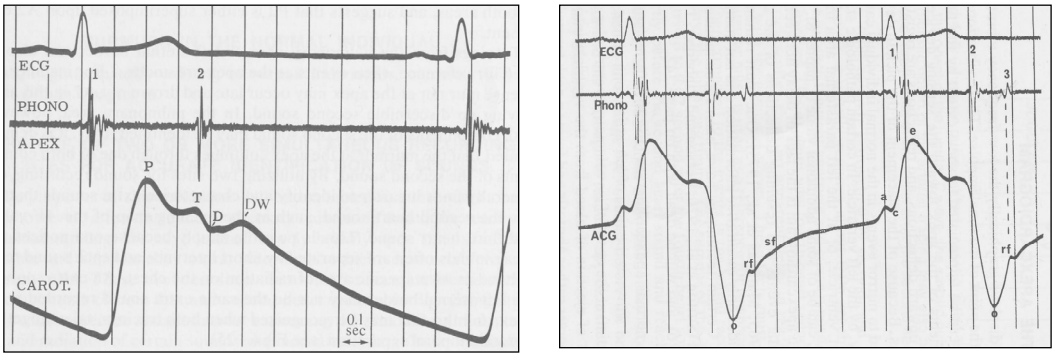

Figure 5: Infrasound heart signals of the carotid pulse (a) and the apexcardiogram (b), with simultaneous heart sounds (phono) and electrocardiogram (ECG) [28].

1.3 Research motivation and goals

Vibro-acoustic heart signals are a direct manifestation of the mechanical processes within the cardiovascular system. These signals therefore carry valuable information about events of the cardiac cycle and the mechanical functioning of the heart. Physicians have long been utilizing this information by means of heart auscultation and palpitation. However, due to the limitations of the human sensual and perceptual capabilities in receiving and analyzing these signals, their clinical use has been mostly qualitative. The development of quantitative analysis techniques for vibro-acoustic heart signals has been pushed aside in recent decades as a result of the immense advancements in cardiac imaging technologies such as echocardiography, computerized tomography (CT) and magnetic resonance imaging (MRI). As much as these imaging technologies are valuable, they require complex equipment and expert operators and interpreters. Furthermore, they can not be used continuously or outside of the hospital environment, and are therefore inadequate for applications of cardiac monitoring. Recent advancements in sensor technology, miniaturization of high-performance computing devices and wireless communication enable us to revisit quantitative analysis of vibro-acoustic heart signals

from a computational perspective. The goal of the research presented here was to develop signal analysis and pattern recognition techniques for automatic extraction of meaningful physiological and clinical information from vibro-acoustic heart signals. We have asked whether the application of computational methods, including time-frequency segmentation, robust feature extraction, morphological clustering and classification, to vibro-acoustic signal analysis, can improve our ability to automatically infer quantitative information about the mechanical functioning of the heart. Moreover, we have used these techniques to gain new insights about the underlying physiological processes and to assess the potential benefit for non-invasive monitoring of cardiac function and early diagnosis of cardiovascular and cardiopulmonary dysfunctions.

1.4 Previous research

The characteristics of the vibro-acoustic heart signals and methods for their digital analysis have been studied extensively [30, 31]. The following sections briefly review past literature of signal processing and pattern recognition techniques applied to mechanical heart signals, and applications of these techniques for deriving measures of cardiac function.

1.4.1 Signal processing techniques

Early techniques for extracting the frequency content of the phonocardiogram signal included zero-crossing analysis [32] and bandpass filter banks. McKusick et al. studied the applicability of the sound spectrograph for analysis of heart sounds and murmurs [17]. The spectrograph, originally developed for speech analysis, used a bank of bandpass filters to determine the power of the signal at each frequency band of interest. With the advancement in digital computing, Fast Fourier Transform (FFT) has been implemented and applied to analyze the frequency spectrum of S1 and S2 [33, 34]. Time-domain processing methods included synchronized envelope averaging [35, 36], computing of average power in contiguous segments [37], envelogram estimation by Hilbert transform [38] or normalized average Shannon energy [39] and complexity-based segmentation [40]. Modern signal analysis techniques can be generally partitioned into parametric and non-parametric methods. With parametric methods, the investigated signal is first modeled by an equivalent system, and the parameters of the model are then estimated

from measurements of the signal made over a limited period of time. Parametric modeling of heart sounds was studied using autoregressive modeling [41], iterative pole-zero modeling [42], and adaptive spectrum analysis with matching pursuit method [43, 44]. Durand et al. [45] evaluated the performance of parametric methods (all-pole and pole-zero modeling) and of FFT-based methods. They concluded that the basic periodogram (FFT) appears to be the best compromise for estimating both the spectral distribution and dominant frequency peaks of the closing sounds of artificial heart valves. Akay et al. [46] compared the ability of FFT and parametric methods to identify coronary artery disease from diastolic heart sounds, and pointed out the eigenvector methods that provided the best diagnostic performance. A different type of signal modeling is probabilistic hidden Markov models (HMM), applied to heart sounds for identification of S1 and S2 [47]. Non-parametric decomposition of heart sounds was used to explore the inner structure of the signals. Homomorphic filtering was used to deconvolve heart sounds into a sum of slowly varying and fast varying components, providing a smooth and robust envelogram [48]. Joint time-frequency analysis is a favorable method for decomposition and representation of heart sounds, due to the multi-component and non-stationary nature of these signals [49]. Wood et al. used time-frequency representations (TFR) to characterize the first heart sound in dogs [50]. They compared short-time Fourier transform (STFT), Wigner-Ville distribution (WVD), and reduced interference distributions (RID), including Choi-Williams distribution (CWD) and binomial transforms. Binomial transform was highlighted as an efficient algorithm with good cross-term suppression properties [51]. Other studies pointed out additional techniques, including cone-kernel distribution [52] continuous wavelet transform [53] and S-transform [54] as suitable choices for representation and analysis of S1 and S2.

1.4.2 Computational learning techniques

Signal processing methods enable to uncover the structure of the signal and to extract features which compactly represent the properties of the raw data. Feature extraction is usually a preceding step of a classification or regression task, which is used to infer application-specific information from the input signals. Feature selection can utilize domain-specific knowledge for choosing features with physical meaning (e.g. frequency, time interval, energy), or alternatively, may be automatic, relying on general properties of

the signals. Domain-specific features used for heart sound analysis include, for example, dominant frequencies of the signal's spectrum [55], bandwidth of the dominant frequencies, mean and total spectral energy, and intensity ratio of S1 and S2 [56]. In model-based signal analysis, the parameters of the model are a natural set of features. For example, Guo et al. used the coefficients of 12-order all-pole system as features [57]. Bentley et al. used a search scheme to select an optimal subset of features from a larger set, extracted using discrete wavelet transform [58]. They showed that this feature set provide better performance than morphological features extracted from CWD in classifying the condition of native and artificial heart valves. The application of identifying degenerated artificial valves from features of their closure sounds was addressed using different types of classifiers, including K-nearest neighbor (KNN), Gaussian-Bayes and neural networks. Reported results, which seem to be optimistic, indicate high accuracy of 89%-98% in detecting degeneration of different types of artificial heart valves.

Unsupervised learning using morphological cluster analysis has been applied before to biomedical imaging modalities such as magnetic resonance imaging [59], as well as to electrocardiogram signals [60]. To the best of our knowledge, there are no previous studies that have applied cluster analysis on the morphology of heart sounds.

1.4.3 Applications of heart sound analysis

The relation between the properties of the first heart sound and the ventricular pressure was demonstrated by Sakamoto et al. [24], who studied the relationship between the amplitude of the first heart sound and the maximal rate of left ventricular pressure rise in anesthetized dogs. Their results showed a strong correlation in various experimentally-induced cardiac conditions. Similar experiments with the amplitude of the second heart sound [25] showed that the amplitude of the aortic component of the second heart sound is not directly related to the aortic pressure. However, it was shown that there is a consistent relation between the peak rate of development of the aortic-to-left ventricular differential pressure and the amplitude of the aortic component of S2. Later it was shown that myocardial infarction in humans caused a shift of the maximum energy of S1 to a lower frequency range [61], and that a reduction in the spectral energy of S1 correlated well with the presence of significant coronary artery disease [62]. More recently, Chen et

al. showed a strong cross-correlation between the instantaneous frequency of S1 and dP/dt of dogs in various contractile states [63]. They suggested that the resonant frequency of S1 is proportional to the fractional power of the tension of the left-ventricular myocardium during contraction, which relates to the left ventricular pressure gradient by Laplaces's law. Heckman et al. [64] used simultaneous recordings of internal phonocardiogram and pressure waves in the left ventricle and the aorta to show that it is possible to derive close facsimiles of the phonocardiogram by double differentiation of the corresponding pressure pulse, and conversely to derive the pressure pulse by double integration of the phonocardiogram. Several studies suggested methods for estimating pulmonary artery pressure by spectral or time-frequency analysis of the second heart sound, showing high correlation with invasive measurements [65-67]. Phonocardiography during exercise test, utilizing the increase in the amplitude of S1 as an index of cardiac contractility reserve, was suggested by Xiao et al. [68]. Emergence of diastolic heart sounds, S3 and S4, during handgrip static exercise has been shown to be a useful clinical adjunct in the diagnosis of coronary artery disease [69]. Another clinical application of heart sound analysis, addressed in many studies, is automatic detection of murmurs and classification of innocent and pathologic murmurs [70-72]. Recently, Syed et al. proposed an analytical framework for analysis of cardiac sounds, which may be useful for teaching cardiac auscultation and for computer-assisted diagnosis of valvular disease [73]. Finally, Tanaka et al. [74] described coupled analysis of respiration periods and heart rate from phonocardiographic sensor, which can be used for home monitoring.

1.4.4 Applications of infrasound signal analysis

The techniques of apexcardiography became popular with the work of Benchimol and Dimond, who demonstrated the clinical usefulness of the apexcardiogram signal for timing systolic and diastolic events and assessing left ventricular function [75]. Apexcardiography-derived indexes, such as the ratio between the height of the 'a' wave and the total vertical deflection of the trace, were found to signify elevated left-ventricular end-diastolic pressure due to various conditions [28]. Manolas et al. showed that the systolic upstroke time of the apexcardiogram was in close correlation with internally measured indexes of myocardial performance: isovolumetric contraction time, maximal dP/dt and peak velocity of shortening of the contractile elements [76]. A method

for estimating the ejection fraction, and the mean velocity of the shortening of the contractile elements, based on the apexcardiogram deflection at the beginning and at the end of systole, was suggested by Antani et al. [77]. Both indexes were shown to highly correlate with measures taken by invasive angiography procedure. The diastolic function of the left ventricle was assessed using apexcardiography by Manolas and Rutishauser [78]. They showed correlations between apexcardiographic time-amplitude features, and invasively-derived indexes of ventricular relaxation and compliance (minimal dP/dt, velocity of lengthening of the contractile elements). More recently, Manolas et al. [79] showed that static handgrip apexcardiographic test is a useful method for detecting left ventricular diastolic abnormalities in patients with coronary artery disease, hypertension and cardiomyopathy.

The carotid pulse signal was commonly used for indirect measurements of systolic time intervals [28]. The duration of the pre-ejection period (PEP) was measured from the ECG Q-wave to the onset of the steep carotid upstroke, corrected for the delay in the pulse-wave transmission to the neck. Left-ventricular ejection time (LVET) was measured from the carotid upstroke to the dicrotic notch. The ratio PEP/LVET was suggested as a useful index for systolic performance. Garrard et al. [80] showed a high correlation between PEP/LVET and ejection fraction in patients with cardiac disease. However, on a large series of patients with ischemic heart disease, Swatzell et al. [81] concluded that although the PEP/LVET ratio was found to be related to the ejection fraction, the correlation was too low and the scatter of the data was too large for reliable clinical use.

Another application of pulse wave signals is measuring the pulse wave velocity (PWV) as an index of arterial distensibility [82]. The pulse wave velocity can be measured from the time delay between two pulse waves (for example, carotid and radial pulses), given the distance between the recording sites. PWV is inversely related to the arterial wall distensibility, and increased PWV was suggested to be an early indicator of atherosclerosis development.

CHAPTER 2
RESEARCH METHODOLOGY

2.1 Principles of experimental methodology

The experimental methodology used in this work was based on the following principles:

1. Acquisition of vibro-acoustic heart signals from multiple recording locations.
2. Simultaneous acquisition of supplementary signals and data including electrocardiogram, breathing pressure and echocardiographic information.
3. Controlled modulation of the hemodynamic conditions by either alternation of the respiratory activity or induction of stress response.
4. Computational analysis of the changes induced in the vibro-acoustic heart signals by the hemodynamic modulation.
5. Analysis of the relation between the information extracted from the vibro-acoustic signals and 'gold-standard' reference information obtained from the supplementary echocardiographic or respiratory data.

2.2 Principles of signal acquisition

Accurate and reliable acquisition of low-frequency mechanical heart signals from the chest wall is technically challenging and requires a proper combination of transducer, amplification and sampling hardware. The signals are picked-up using a piezoelectric contact transducer. Piezoelectric crystals are elements that acquire an electrical charge when compressed or deformed. The transducer is composed of a flexible metallic membrane coupled to a piezoelectric element. The lightweight and deformable membrane is distorted relative to the heavier housing when force is applied to its surface. The deformation of the membrane relative to the housing generates an electric output from the piezoelectric element (Figure 6). Piezoelectric sensors generate low voltage and current, and have high electric impedance of 1MΩ [83]. In order to amplify the signal from the range of 1mV to the typical range of analog-to-digital converters (±1V), a suitable pre-amplifier with matching impedance and low noise figure must be used. As the frequencies of the recorded vibrations are as low as 1Hz, the amplifier should have linear

frequency response from DC up to the upper limit of heart sounds frequencies (400Hz). Finally, the analog signals are sampled by a 16-bit analog-to-digital converter, ensuring adequate resolution and accuracy. Sampling frequency should be at least 1KHz to acquire the full range of the signal while avoiding aliasing problems.

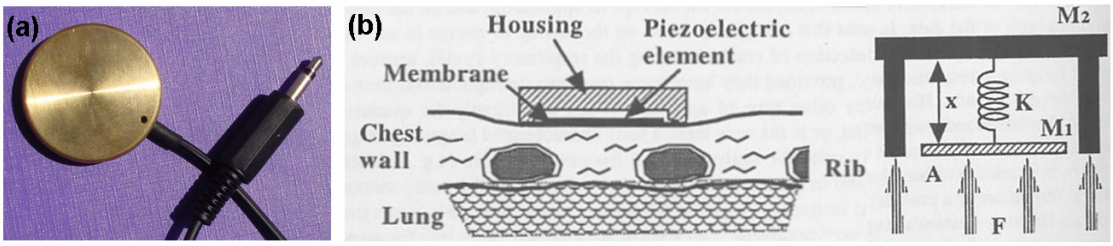

Figure 6: Piezoelectric contact transducer (a) and its principle of operation (b). When force F is applied to the surface A of the membrane, it is deformed relative to the heavier housing. The deformation x of the membrane relative to the housing generates an electric output from the piezoelectric element [83].

2.3 Modalities of echocardiography

Echocardiography was used this work for measuring parameters of cardiovascular function, and was taken as a 'gold-standard' reference. This section provides a brief overview of the principles and the terminology of the echocardiography modalities that were used. Echocardiography is a non-invasive imaging technique based on transmission of ultrasound waves and detection of their echoes returning from the heart. In two-dimensional (2D) echo, an ultrasound beam rapidly scans a selected cardiac section, producing a tomographic image that shows the spatial structure of the heart during the cardiac cycle. Doppler echocardiography is a modality that utilizes the Doppler principle (2.1) to measure the direction and velocity of moving blood within the heart. In short, transmitted sound waves returning from a moving object are received with a frequency shift f_d. This shift is directly related to the velocity (V) and angle (θ) of the movement, and the relation depends on the transmitted frequency f_0 and the velocity of sound c. Measuring the Doppler shift therefore enables to calculate the velocity of the returning object, which may be, for example, blood ejected to the aorta or filling the ventricle. More recently, similar technology has been applied for measuring the velocity of cardiac tissue. This modality, denoted tissue Doppler imaging (TDI), provides quantitative information about the dynamics of cardiac wall motion, which can be used to identify motion abnormalities [84].

$$f_d = \frac{2f_0}{c} V \cos\theta \tag{2.1}$$

Strain echocardiography is a new technique for quantitative assessment of left-ventricular systolic and diastolic function, based on post-processing of tissue Doppler imaging data [85]. Tissue-Doppler assessment of ventricular function is based on measurements of regional velocities of the cardiac walls. However, as tissue velocity measurement is affected by the motion of the entire heart during contraction, it may not provide a reliable estimation of cardiac contractility. Instead, the regional deformation of the myocardial tissue and the rate of this deformation, denoted in equation (2.2) as strain (*S*) and strain-rate (*SR*), have been shown to quantify the changes in the myocardial function during stress echocardiography [86]. Specifically, peak systolic strain rate was found to be strongly correlated to invasive gold-standard indices of left ventricular contractility [87].

$$S = \frac{L - L_0}{L_0} = \frac{\Delta L}{L_0} \quad , \quad SR = \frac{dS}{dt} \cong \frac{\Delta V}{L_0} \tag{2.2}$$

(L_0: initial length; L: instantaneous length; ΔL: change of length; ΔV: change of velocity)

2.4 Description of data sets

2.4.1 Dobutamine stress echo data set

The data set of heart signals acquired during Dobutamine stress echo test (HSDSE) was designed for studying the relationship between heart sound characteristics and left-ventricular global systolic function. To induce dynamic, yet controllable, hemodynamic changes, we used clinical settings of a routine echocardiography pharmacological stress test. In this test, the heart's contractility and rate are gradually augmented by administration of a pharmacological agent called Dobutamine, while the segmental wall motion of the heart is evaluated by echocardiography.

Patients: Data was acquired from eleven male subjects of ages 36-79 (mean 60±14), referred to a routine Dobutamine stress echo test (DSE) for assessment of ischemic heart disease (Table 1). Female subjects were not included in this study as chest-wall recording in women might have introduced technical and procedural difficulties. The referral indications included positive ergometry stress test, atypical chest pain and chest pain during physical activity. Two of the subjects had a history of coronary artery disease. These two subjects were diagnosed as positive for myocardial ischemia in the DSE test.

The remaining nine patients were diagnosed as negative for ischemic heart disease. All patients have signed an informed consent prior to their enrollment in the study.

Acquisition system: The acquisition system (Figure 7a) consisted of 4 piezoelectric contact transducers (PPG Sensor Model 3, OHK Medical Devices, Haifa, Israel), an ECG sensor (EKG-BTA, Vernier Software & Technology, Beaverton, OR), a preamplifier with high input impedance and a linear frequency range of 1Hz – 4KHz (A.S. ZLIL, Bnei-Brak, Israel), a 16-bit analog-to-digital converter (PMD-1608FS, Measurement Computing Corp., Norton, MA), and a designated signal recording software running on a portable personal computer. Echocardiography images were acquired using a GE Vivid 7 ultrasound machine (General Electric Healthcare, Wauwatosa, WI).

Signals: Four vibro-acoustic transducers were placed at the apex area, the aortic and pulmonary areas (2nd intercostal space, right and left sternal border) and at the right carotid artery. The transducers were firmly attached using either elastic straps or adhesive bands. The patients were monitored while lying on their left side (Figure 7b). Vibro-acoustic and ECG signals were continuously recorded during the stress test (30-45 minutes long) at a sample rate of 4KHz. Two-dimensional echo cine loops of a single heart beat were captured before the beginning of the stress test (baseline), during each stage of the test and following the test (recovery), from three apical views (4-chamber, 2-chamber and apical long axis) at a high frame rate of 70-100 frames per second.

Protocol: The standard Dobutamine stress echo protocol consisted of four 3-minute stages of increasing Dobutamine dosage, from 10 to 40μg/kg/min. If the target heart rate, defined as 0.85 * (220 – Age), was not achieved at the end of the final stage, 0.25 mg boluses of atropine were given at 1-min intervals, up to a maximum of 1 mg.

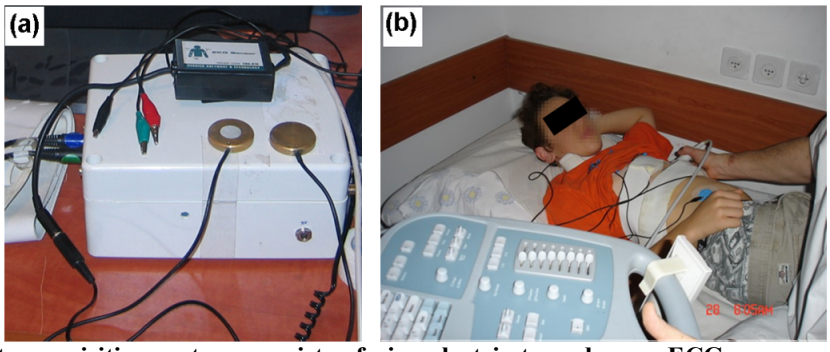

Figure 7: Data acquisition system, consists of piezoelectric transducers, ECG sensor, amplifier and analog-to-digital converter (a), and its application on a subject recorded during echocardiographic examination (b).

Table 1: Characteristics of the patients included in the HSDSE dataset

Num	ID	Age	Gender	Ref. reason	History	Test results
01	HP	54	M	atypical chest pain	None	Negative
02	AK	55	M	chest pain	None	Negative
03	SH	71	M	suspected ischemia in exercise	None	Negative
04	GM	52	M	positive ergometry	None	Negative
05	HS	61	M	positive ergometry	None	Negative
06	BM	77	M	suspected ischemia in exercise	Hypertension	Negative
07	AL	52	M	chest pain	None	Negative
08	EZ	36	M	positive ergometry	None	Negative
09	AB	48	M	atypical chest pain	COPD	Negative
10	VM	78	M	typical chest pain	MI, PCI	Positive
11	RY	79	M	chest pain	PCI, CABG	Positive
Average		60.3				
Stdev		14.2				

2.4.2 Respiratory pressure data set

The data set of heart sounds and respiratory pressure (HSPRS) was designed to assess the relationship between heart sounds and respiratory condition, by acquiring signals during controlled changes of the breathing phase - namely inspiration, expiration, and breath-holding (‘apnea’), and of the respiratory resistive load, achieved by breathing through resistive pipes.

Patients: Data was acquired from 12 healthy volunteers (age 29±12, 8 men, Table 2).

Acquisition system: The acquisition system consisted of two piezoelectric contact transducers (PPG Sensor Model 3, OHK Medical Devices, Haifa, Israel), a breathing pressure transducer (Validyne, Northridge, CA), an ECG recording system (Atlas Researchers, Hod-Hasharon, Israel), a pre-amplifier (Alpha-Omega, Nazareth, Israel), a 16-bit analog-to-digital converter (National Instruments, Austin, TX) and a designated signal recording software running on a portable personal computer (Figure 8).

Signals: Two channels of heart sounds, breathing pressure at the mouth and a single lead ECG were simultaneously acquired at sampling rate of 11.025KHz. During data recording, the subjects were sitting upright, with the heart sound transducers firmly attached by an elastic strap on the left and right para-sternal lines at the 4^{th} intercostal spaces.

Protocol: The data was recorded while the subjects were breathing through a mouthpiece that was side-connected to the pressure transducer and serially attached to plastic pipes (0.5 cm OD) of varying lengths, used for altering the respiratory resistive loads. Five levels of resistance were used: at level 0, no resistive pipe was attached, and at levels 1-4 the lengths of the resistive pipes were 8.5cm, 22cm, 66cm and 200cm, respectively. The signals were recorded in duplicates with each resistance level during 40 seconds of normal breathing, and during alternates between 15 seconds of normal breathing and 15 seconds of breath-hold (apnea).

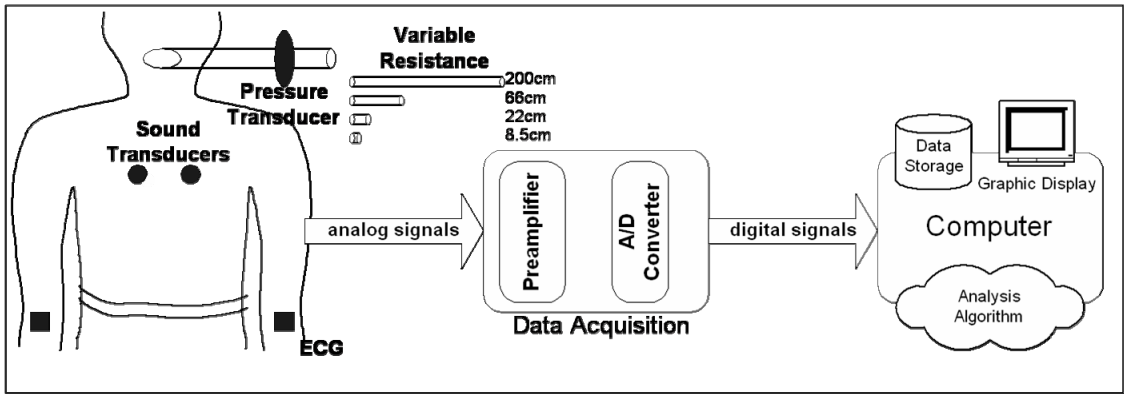

Figure 8: Experimental setup of HSPRS data set. Two channels of heart sounds, ECG and airway-opening pressure are simultaneously acquired while the subject is breathing against resistive tubes with variable length. The signals are amplified, digitally-sampled and saved for further computational analysis.

Table 2: Characteristics of the subjects included in the HSPRS dataset

Num	ID	Age	Gender	Additional information
01	GA1	31	M	
02	ND1	21	F	
03	NM1	24	M	
04	NG1	53	M	
05	NM2	25	M	
06	ND2	19	F	
07	NM3	20	F	
08	OG1	24	M	smoker
09	ST1	22	M	athlete
10	ZM1	20	F	
11	RS1	54	M	athlete
12	SS1	37	M	athlete
Average		29.2		
Stdev		12.5		

2.4.3 Echo-Doppler data set

The data set of vibro-acoustic and echo Doppler signals was acquired as a pilot for assessing the technological feasibility of signal acquisition and extraction of temporal features of the cardiac cycles, which were compared against simultaneously-acquired echo-Doppler measurements. The echo-Doppler audio signals are synthetic acoustic signals, produced by the echocardiograph to provide an aural representation of the measured velocities. The measured Doppler frequency shifts are used to synthesize a sum of sinusoid waveforms. One signal encodes the positive frequency shifts and a second signal encodes the negative frequency shifts. These acoustic signals are transmitted in two separate channels. The acquisition system described in section 2.4.1 was used to simultaneously record vibro-acoustic signals from the apex, right carotid artery, aortic area and pulmonary area, along with one-lead ECG and echo-Doppler audio signals, at a sampling rate of 8KHz. The subjects were recorded while lying on their left side, during normal quiet breathing and during breath-hold. The transducers were firmly attached using elastic straps. Doppler echocardiography was acquired from the apical window with an Acuson Sequoia C256 ultrasound machine (Siemens Healthcare, Malvern, PA).

2.5 Data preprocessing

The acquired vibro-acoustic signals were first preprocessed in order to retain the required frequency bands and prepare the data for further analysis. Heart sounds were digitally filtered by a Chebishev type-I IIR band-pass filter of order 6, with a passband between 20Hz and 250Hz. Infra-sound signals were filtered between 0.8Hz and 30Hz.

Partitioning of the continuous raw signal into cardiac cycles was done using ECG-based segmentation. The peaks of the QRS complexes (R-waves) were identified [88] and the beginning of each cardiac cycle was defined 50ms before the R-wave. In each cardiac cycle, peaks of S1 and S2 were identified from the energy envelogram, computed by applying a low-pass filter on the instantaneous amplitude of the analytic signal, given in equation (3.4). Segments of S1 were defined as 200ms slices, starting at the beginning of the cycle. Segments of S2 were defined as 200ms slices centered at the detected energy peak (Figure 9).

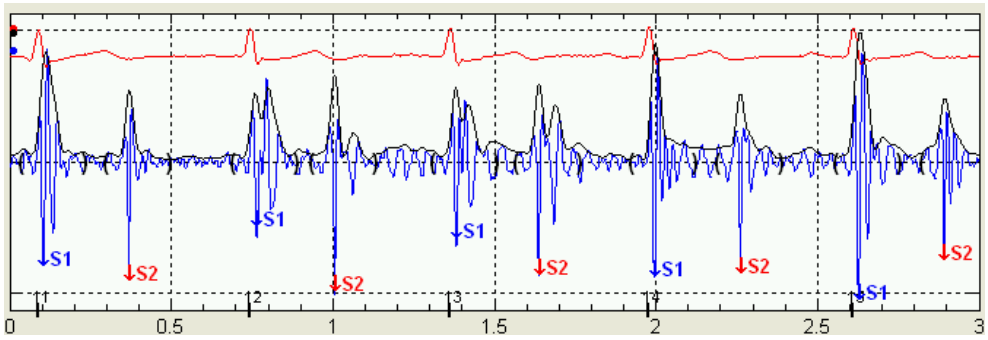

Figure 9: Basic segmentation of heart sounds. QRS peaks of the ECG are used to partition the continuous heart sound signal into cardiac cycles. Energy envelogram is used to identify S1 and S2.

CHAPTER 3
COMPUTATIONAL ANALYSIS TECHNIQUES OF HEART SIGNALS

An analysis framework of biomedical signals in general and vibro-acoustic heart signals in particular consists of the following building blocks:

1. Preprocessing – digital filtering of the acquired signals in order to retain only the required frequency bands, and remove noise and artifacts.
2. Segmentation – partitioning of the signal into segments that represent distinct components or events in the cardiac cycle. Heart signals are naturally separated into heart beats or cycles, and each cycle is further segmented to identify its components.
3. Representation – transformation of the signal into a feature domain most suitable for its characterization. Time-domain, spectral and joint time-frequency representations are examples of possible representations for vibro-acoustic heart signals.
4. Feature extraction – selection and calculation of a compact set of features, which reliably represent the properties of the signal.
5. Pattern recognition – analysis of signal's morphology in order to identify and characterize distinct morphological patterns. Techniques include signal alignment and clustering.
6. Classification – association between features and patterns of the signal and physiological or clinical conditions from the application domain.

The following sections describe the fundamentals of some of the methods used in this work for signal representation, feature extraction, clustering and classification.

3.1 Analysis techniques – theoretical background

3.1.1 Time-frequency representations

Time representation of a signal, as a series of signal values sampled in time, is probably the most natural form of examining a physical signal. Signals that have periodic components, or that are accompanied by noise, can be better analyzed in the frequency domain, by applying the Fourier transform, defined as:

$$S(f) = \int_{-\infty}^{\infty} s(t) e^{-i2\pi ft} dt \tag{3.1}$$

The Fourier spectrum is an expansion of the signal into an infinite sum of sine and cosine functions. It provides the amplitudes and phases of frequencies contained in the signal, without any localization in time. One way to describe the signal simultaneously in both time and frequency is to consider its instantaneous frequency.

$$f(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt} \tag{3.2}$$

Where $\phi(t) = \arg\{s_a(t)\}$ is the instantaneous phase of the analytic signal:

$$s_a(t) = s(t) + iHs(t) \tag{3.3}$$

The analytic signal is a complex signal, whose real part is the original signal, and imaginary part is the Hilbert transform of the signal. The Hilbert transform is an operator that shifts the phases of all frequency components in the spectrum of *s(t)* by 90º. In addition to instantaneous frequency and phase, the analytic signal can be also used to represent the instantaneous amplitude, defined as:

$$A(t) = |s_a(t)| \tag{3.4}$$

For stationary signals, whose spectral properties are constant in time, the Fourier spectrum or the instantaneous frequency may be a sufficient description. However, biomedical signals are often non-stationary, and may consist of multiple components that are distinct in time and frequency. Joint time-frequency representations (TFR) provide a solution for describing and extracting information from this type of signals. The short-time Fourier transform (STFT), defined in equation (3.5), is a commonly-used linear transform, based on the principle of multiplying the signal by a window function *W* centered at time *t*, and computing the local Fourier spectrum of the windowed signal for each time instant [89].

$$S(t, f) = \int_{-\infty}^{\infty} s(\tau) W(\tau - t) e^{-i2\pi f\tau} d\tau \tag{3.5}$$

This method introduces the inherent tradeoff between the temporal and spectral resolution of the representation: good time resolution requires a short window, whereas good frequency resolution requires a long window. This limitation, which is a

consequence of the well-known Heisenberg uncertainty principle, is partially removed in a second type of linear transform, the wavelet transform. The wavelet transform (WT) projects the signal on a set of basis functions, which are deduced from an elementary function (the mother wavelet) by operations of translation and dilation. With this representation, the time-frequency resolution is variable: good time resolution is obtained at high frequencies, while good frequency resolution is obtained at low frequencies [90]. S-transform (ST), defined in equation (3.6), is an extension of the continuous wavelet transform, which provides frequency-dependent resolution while maintaining a direct relationship with the Fourier spectrum [91].

$$S(t,f) = \int_{-\infty}^{\infty} s(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(t-\tau)^2 f^2}{2}} e^{-i2\pi f\tau} d\tau \tag{3.6}$$

An alternative to the linear atomic decomposition of STFT and WT is to examine the distribution of the signal's energy along the two variables of time and frequency. Such energy distributions are quadratic transforms of the signal, and can be expressed by a general representation (3.7), proposed by Cohen [92].

$$C_x(t,f;\phi) = \frac{1}{4\pi^2} \iiint e^{i\theta u - i2\pi f\tau - i\theta t} \phi(\theta,\tau) s(u+\frac{\tau}{2}) s^*(u-\frac{\tau}{2}) du d\tau d\theta \tag{3.7}$$

Where $s(u)$ is the time signal, $s^*(u)$ is the signal's complex conjugate, θ and τ are the frequency and time lag, and $\phi(\theta,\tau)$ is a kernel defining the specific distribution.

The Wigner-Ville distribution (WVD), given in equation (3.8) is a widely-used quadratic transform of the Cohen's class. It provides high resolution in both time and frequency and has some desirable mathematical properties. However, due to its quadratic nature, WVD produces cross terms, or interferences, which can obscure the interpretation of multi-component signals.

$$S(t,f) = \int_{-\infty}^{\infty} s(t+\frac{\tau}{2}) s^*(t-\frac{\tau}{2}) e^{-i2\pi f\tau} d\tau \tag{3.8}$$

Reduced-interference distributions are transforms that use different kernels in order to suppress cross terms. For example, Choi-Williams distribution (CWD), defined by equation (3.9), is using an exponential kernel with a single parameter σ, which controls interference suppression and frequency resolution.

$$S(t,f)=\int_{-\infty}^{\infty} e^{-i2\pi ft}\int_{-\infty}^{\infty}\sqrt{\frac{\sigma}{4\pi\tau^2}}e^{-\frac{\sigma(\mu-t)^2}{4\tau^2}}s(\mu+\frac{\tau}{2})s^*(\mu-\frac{\tau}{2})d\mu d\tau \tag{3.9}$$

3.1.2 Principal Component Analysis

Principal component analysis (PCA) is a well-known statistical technique for dimensionality reduction [93]. The principle of PCA is to project the data on a new orthogonal basis, so that the variances of the linearly transformed data are sorted in descending order along the coordinates, with the maximal variance on the first coordinate (first principal component), the second largest variance on the second coordinate, and so on. The projection of the original data on the first few principal components provides a low-dimensional representation of the data, which emphasizes the significant features (in terms of statistical variability) in the data. The choice of the significant principal components is done by examining their associated eigenvalues.

Mathematically, the PCA algorithm finds the new basis by diagonalizing the covariance matrix C of the zero-mean column vectors $x_i \in \mathbb{R}^N, i=1,...,m$ (equation (3.10)), and finding its eigenvalues λ and eigenvectors v (principal components), such that $\lambda v = Cv$.

$$C=\frac{1}{m}\sum_{j=1}^{m}x_j x_j^T \tag{3.10}$$

3.1.3 Dynamic Time Warping

Dynamic time warping (DTW) is a non-linear alignment procedure of time series, which has been extensively used in applications of speech recognition [94]. The principle of DTW is to stretch or shrink ('warp') the signals along the time axis such that the distance between them (most commonly, Euclidean distance) is minimized. Given two sequences $X=x_1,x_2,\ldots,x_n$, $Y=y_1,y_2,\ldots,y_m$, the algorithm constructs an n-by-m matrix W, such that $W(i,j)=Dist(x_i,y_j)=\left\|x_i-y_j\right\|^2$, and then retrieves the path through the matrix that minimizes the total cumulative distance between the aligned sequences (Figure 16b). The warping path is found by dynamic programming in time complexity of $O(n^2)$, or $O(nw)$, in case the warping windows is constrained to a window of width w around the diagonal of the matrix. DTW provides both full mapping between the elements of the aligned sequences, as well as a measure of the distance between the sequences.

3.1.4 Hierarchical clustering

Clustering is an unsupervised learning technique whose purpose is to partition a dataset into disjoint sub-sets (clusters), such that data elements within the same cluster share some sort of similarity. Similarity is often measured using a distance metric that is suitable for the nature of the analyzed data. One class of such metrics is the Minkowski metric, defined in (3.11) for time series *s* and *r*. The Euclidean distance is obtained for $q = 2$, while $q = 1$ yields the city-block distance.

$$d(s,r) = \left(\sum_{t=1}^{n} |s_t - r_t|^q \right)^{1/q} \tag{3.11}$$

The clustering task is defined as an optimization of a criterion function that measures the clustering quality of any data partitioning. The most widely used criterion function is the sum of squared error, defined in (3.12),

$$J_e = \sum_{i=1}^{N} \sum_{s \in C_i} \left\| s - m_i \right\|^2 \quad , \quad m_i = \frac{1}{n_i} \sum_{s \in C_i} s \tag{3.12}$$

where $C_1,\ldots,C_N$ are clusters, n_i is the size of cluster C_i and m_i is the mean of cluster C_i.

In many cases, the structure of the data is hierarchical, meaning that clusters may have subclusters. Agglomerative hierarchical clustering [95] is a procedure that initially partitions a set of *n* data elements into *n* clusters, each containing one data element, and then iteratively merges the two most similar clusters, until the entire dataset forms a single cluster. The bottom of the created hierarchical tree, often represented by a dendrogram (Figure 10), can next be pruned so that the required number of *N* clusters is obtained. Alternatively, in case there are significant changes in the quality of the clustering at different levels of the tree, pruning can be done at the hierarchy level that best reflects the 'natural' number of clusters in the data. Data elements below each cut are assigned to a single cluster, creating the output data partitioning to clusters $\{C_1,...,C_N\}$.

Agglomerative clustering requires a measure for the distance between clusters, which is used to select the next two clusters to be merged. Common choices for this measure are the single-linkage and complete-linkage algorithms (3.13), using the minimal or the maximal distance between pairs of data elements from the two clusters, respectively.

$$D_{\min}(C_i, C_j) = \min_{s \in C_i, r \in C_j} D(s,r) \quad , \quad D_{\max}(C_i, C_j) = \max_{s \in C_i, r \in C_j} D(s,r) \tag{3.13}$$

Another type of cluster distance is a step-wise optimal criterion, which chooses the clusters such that the increase in the overall sum-of-squared error after the merge is minimal. This type of linkage, suggested by Ward [96], is defined by:

$$D_{ward}(C_i, C_j) = \sqrt{\frac{n_i n_j}{n_i + n_j}} \left\| m_i - m_j \right\| \tag{3.14}$$

where n_i, n_j are respectively the sizes of clusters C_i, C_j, and m_i, m_j are their means.

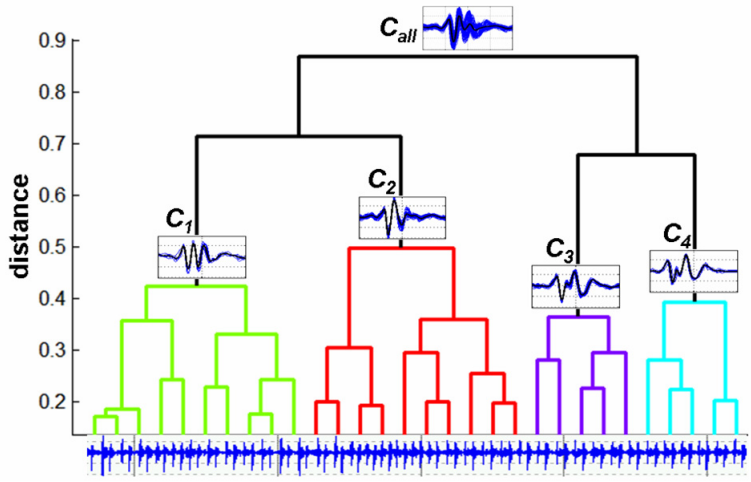

Figure 10: Hierarchical clustering tree. A dendrogram illustrating the process of iteratively merging similar clusters, following by pruning of the hierarchical tree to obtain four clusters (C_1,..,C_4). The processed signals were extracted from a time series partly displayed at the bottom of the figure.

3.1.5 Classification

Among the numerous methods of supervised learning, we describe two classical classification algorithms used in this work: K-nearest neighbor and linear discriminant analysis. K-Nearest-Neighbor (KNN) is a non-parametric method that classifies a test data element by a majority vote of the closest data elements in the training set [93]. Given a labeled training set B and a test data element d, KNN algorithm classifies d by choosing K train data elements $\{b_1,...,b_K\} \subset B$ that are the closest neighbors of d under a distance metric D: $D(d,b_1) \le D(d,b_2) \le ... \le D(d,b_K) \le D(d,b_j)\ \forall b_j \in B, j \notin \{1,...,K\}$. Then, given that l_i is the label of train data element b_i, d will be classified as the statistical mode of $\{l_1,...,l_K\}$. In the regression case, when the training data elements are associated with

real values v_i rather than discrete class labels, the value associated with d is estimated as a weighted average: $v(d) = \frac{\sum_{i=1}^{K} D(d, b_i) v_i}{\sum_{i=1}^{K} D(d, b_i)}$.

Discriminant Analysis (DA) finds a linear transform that maximizes the separation between classes in the training set [97]. The maximized objective function is: $J(w) = \frac{w^t S_B w}{w^t S_W w}$, where $S_B = \sum_{c=1}^{m} n_c (\bar{d}_c - \bar{d})(\bar{d}_c - \bar{d})^T$, $S_W = \sum_{c=1}^{m} \sum_{i \in c} (d^i - \bar{d}_c)(d^i - \bar{d}_c)^T$ are the between-classes scatter matrix and the within-class scatter matrix, respectively, n_c is the number of data elements in class c, $\bar{d}_c = \frac{1}{n_c} \sum_{i \in c} d^i$ is the mean of class c and $\bar{d} = \frac{1}{N} \sum_i d^i = \frac{1}{N} \sum_{c=1}^{m} n_c \bar{d}_c$ is the mean of the entire training set. Once the transformation w is found, by solving an eigenvalue problem, a test data element d can be classified to $\arg\min_c D(dw, \bar{d}_c w)$, the class whose center is closest to d, under a distance metric D.

The distance metric chosen to be used by the classification algorithm is the mahalanobis distance, defined by (3.15), where V is the covariance matrix of vectors $\vec{d}^i$ and $\vec{d}^j$.

$$D(\vec{d}^i, \vec{d}^j) = (\vec{d}^i - \vec{d}^j) V^{-1} (\vec{d}^i - \vec{d}^j)^T \tag{3.15}$$

3.2 Cluster analysis and classification of heart sounds

3.2.1 Problem definition

Periodic biomedical signals, such as vibro-acoustic heart signals, exhibit considerable morphological variations from beat to beat, resulting from the inherent variability of their generating processes. For instance, both S1 and S2 exhibit noticeable changes in amplitude and morphology during normal respiration (Figure 11a), and are nearly invariable during breath-holding (Figure 11b). By exploring the relations between the physiological processes and the beat-to-beat changes of the sound signals, we can

associate a physiological meaning to the morphology of the signal. Given a variable cyclical heart sound signal, we would therefore like to identify the distinct repetitious beat morphologies, and analyze the causality of the signal variations by attempting to predict the underlying physiological state.

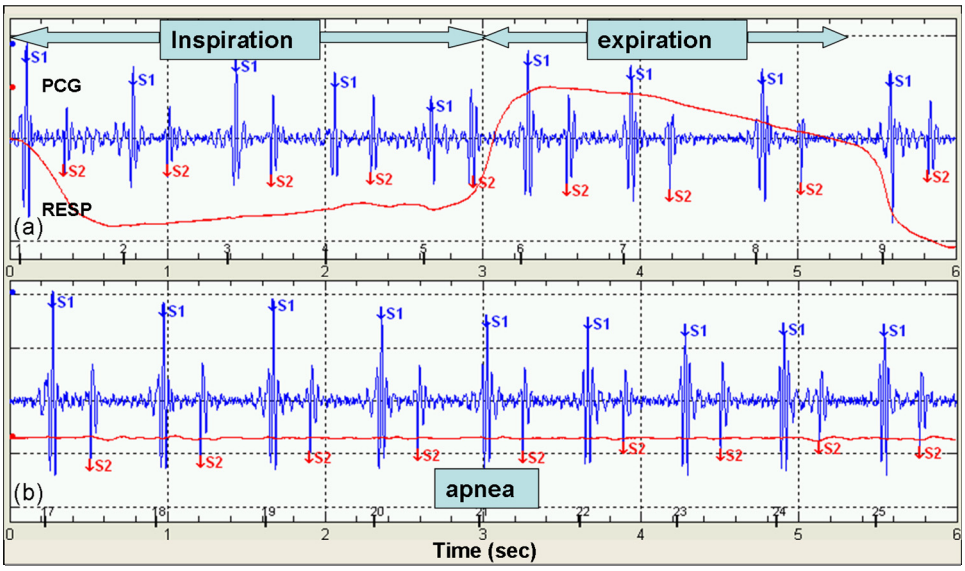

Figure 11: Morphological variability of heart sounds. Heart sound signal (PCG) acquired during a single respiration cycle (RESP) exhibits considerable morphological variability of both S1 and S2 (a). During breath-hold (apnea), there are nearly no beat-to-beat changes in S1 or S2 (b).

We describe a computational analysis framework for identifying distinct morphologies of heart sounds and classifying them into physiological states, thus inferring physiological parameters from the morphology of the signal (Figure 12) [98]. Heart sound signals are first identified and extracted from the acquired data, as described in section 2.5, and then transformed to a raw feature space in either time, frequency, or time-frequency plane. Hierarchical clustering is applied to the signals, and the obtained clusters form a compact representation of the data in the feature space of distances from the centers of the clusters. In this feature space, classification or regression algorithms are used to test whether the different signal morphologies represent different physiological states. The following sections describe the analysis framework in detail, and discuss the choices of raw signal representation, distance metrics and classification algorithms.

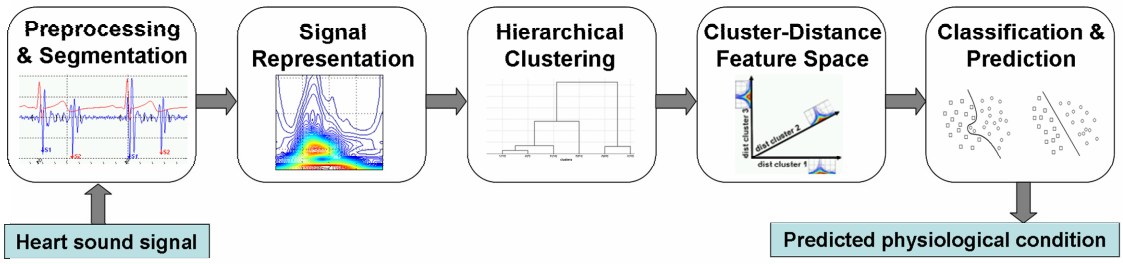

Figure 12: Signal analysis framework. Input heart sound signals are extracted, transformed to a raw feature space and clustered according to their morphologies. Distances from the centers of the clusters are used to compactly represent the data and classify it into physiological states.

3.2.2 Clustering and classification procedure

Hierarchical clustering, applied to heart sound signals, included the following steps:

1. Preprocessing. The signals are normalized to have zero-mean and unit-variance, and phase-shifted to be aligned in the time-domain.
2. Calculation of similarity between data elements. The similarity between every pair of signal cycles s_t and r_t is calculated using either Euclidean distance or correlation distance, defined in equations (3.16) and (3.17), respectively.
3. Grouping of data elements into a binary cluster tree. Pairs of similar data elements are linked together to form larger clusters, which are grouped again until a hierarchical cluster tree is formed. Ward's linkage function, defined in equation (3.14), was used to measure distance between clusters.
4. Cluster determination by pruning the hierarchical tree. Given the required number of clusters, the bottom of the hierarchical tree is pruned, and data elements below each cut are assigned to a single cluster, creating data partitioning.

$$D_{sr} = \left\| s_t - r_t \right\|_2^2 = \sum_t \left(s_t - r_t \right)^2 \tag{3.16}$$

$$D_{sr} = 1 - \frac{\sum_t (s_t - \bar{s})(r_t - \bar{r})}{\sqrt{\sum_t (s_t - \bar{s})^2} \sqrt{\sum_t (r_t - \bar{r})^2}} \quad , \quad \bar{s} = \frac{1}{n} \sum_{t=1}^{n} s_t \quad , \quad \bar{r} = \frac{1}{n} \sum_{t=1}^{n} r_t \tag{3.17}$$

While cluster analysis identified distinct signal morphologies in the data, the classification procedure is aimed to uncover the relation between these morphologies and the alternating physiological conditions. This aim is achieved by evaluating the ability of a classifier to accurately predict the label of each heart beat, given only its morphological representation. The input of the clustering-classification framework is a data set of n heart sound cycles, $B = \{(b_1, l_1), (b_2, l_2), ..., (b_n, l_n)\}$, where b_i is a representation of a heart sound

component (e.g. S1) during a single cardiac cycle, and l_i is its associated class label $l_i \in \{L_1,...,L_k\}$. The cluster analysis procedure assigns a cluster identifier to each signal cycle, producing a clustered data set $C = \{(b_1,c_1),(b_2,c_2),...,(b_n,c_n)\}$, where $c_i \in \{1,...,N\}$. Using these notations, a cluster C_j is the set of signal cycles with cluster identifier c_j: $C_j = \{i \mid (b_i,c_j) \in C\}$. Clusters that contain a minimal portion of the data, i.e. $|C_j| \geq \beta n$, are denoted as *significant* clusters (β was set by experiment to 0.05). The center of a cluster C_j is a weighted average of the cluster's elements, in which each signal cycle is weighted by its similarity to the cluster's arithmetic mean: $\bar{C}_j = \sum_{i \in C_j} \omega_i b_i$, $\omega_i = 1 - D(b_i, (\sum_{i \in C_j} b_i)/|C_j|)$, where D is a distance function.

The centers of the significant clusters provide a compact representation of the morphological variability in the entire data set. Furthermore, a signal cycle b_i can be efficiently characterized by the vector of its distances from the centers of the significant clusters $\vec{d}^{\,i} = \left(d_1^i, d_2^i, ..., d_{\hat{N}}^i\right)$, $d_k^i = D(b_i, \bar{C}_k)$. The classification algorithm is applied in this new feature-space of cluster distances.

The outline of the classification procedure is as follows:

1. Classification is applied separately on the data set B of each subject.
2. Data is partitioned into a training B^{train} set and a testing set B^{test}.
3. Hierarchical clustering is applied on the training set, producing clustered data C^{train}.
4. The centers of the significant training clusters $\bar{C}_1^{train},...,\bar{C}_{\hat{N}}^{train}$ are calculated.
5. Each beat $b_i \in B^{test} \cup B^{train}$ is characterized in the cluster distance space by the vector $\vec{d}^{\,i} = \left(d_1^i, d_2^i, ..., d_{\hat{N}}^i\right)$ of its distances from the centers of the significant training clusters.
6. A classifier F is constructed from the distance-space representation of the training set. For beat $b_i \in B^{train}$, $F(b_i) = F(d_1^i, d_2^i, ..., d_{\hat{N}}^i) = \tilde{l}_i$, $\tilde{l}_i \in \{L_1,...,L_k\}$.
7. The classification accuracy is evaluated on distance-space representation of the testing set B^{test}.

The two classification methods of K-nearest-neighbor (KNN) and discriminant analysis (DA), were used, and their performances with different signal representations and distance metrics were evaluated and compared.

Since the class labels in the studies datasets represent a continuum of physiological changes, rather than dichotomic classes, the classification accuracy CC_m was defined as the fraction of data elements (in the range $[0,1]$) classified within a certain integer range m of their actual label:

$$CC_m = \left| \left\{ b_i \in B^{test} \mid \left| \tilde{l}_i - l_i \right| \le m \right\} \right| \Big/ \left| B^{test} \right| \tag{3.18}$$

Where $l_i, \tilde{l}_i$ are discrete integers representing class labels. In most cases, m=1 was used.

3.2.3 Signal representation - simulation

The choice of the representation method of S1 and S2 signals, used in cluster analysis, is a fundamental layer of the analysis framework. Different representations highlight different features of the signals, and may lead to different clustering results. Figure 13 shows an example for S1 and S2 signals, represented in time, frequency and by different joint time-frequency distributions.

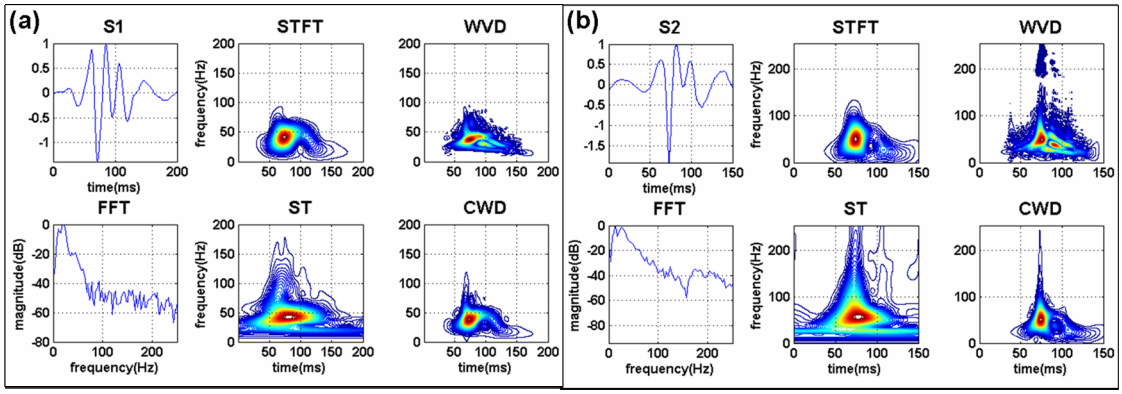

Figure 13: Representation of S1 (a) and S2 (b) in the time-domain, frequency-domain (FFT) and by various joint time-frequency transforms: short-time Fourier transform (STFT), S-transform (ST), Wigner-Ville distribution (WVD) and Choi-Williams distribution (CWD). STFT has fixed resolution, while ST has frequency-dependent resolution. WVD has higher resolution, but its quadratic nature creates cross-terms, which are suppressed in the reduced-interference CWD representation.

A simple simulation was conducted in order to realize the differences between signal representations. The baseline simulated signal was a 30Hz sinus wave with duration of 300ms and a Gaussian amplitude modulation. Random noise with Gaussian distribution was added to the signal, with initial signal-to-noise ratio set to 7dB. Three types of signal transformations were simulated: (i) time shift between -200ms to 200ms, (ii) frequency change between 20Hz to 40Hz and (iii) SNR change between -6dB and 10dB. For each of the transformations, the correlation distance between the baseline signal and the

transformed signal was calculated (Figure 14). Spectral signal representation was obviously insensitive to time shifts. Time representation, on the other hand, was over sensitive, as the distance in this case is the autocorrelation function, which fluctuates between high positive and negative values. Time-frequency representations were more robust to temporal shifts, providing a smooth change of the distance (Figure 14a). The sensitivity to changes in the signal's frequency was higher for WVD, spectral and time-domain representations, compared to ST and STFT, which have lower frequency resolution (Figure 14b). Finally, lower signal-to-noise ratio affected ST and WVD much more than STFT and spectral representations, with intermediate noise sensitivity of the time-domain representation (Figure 14c).

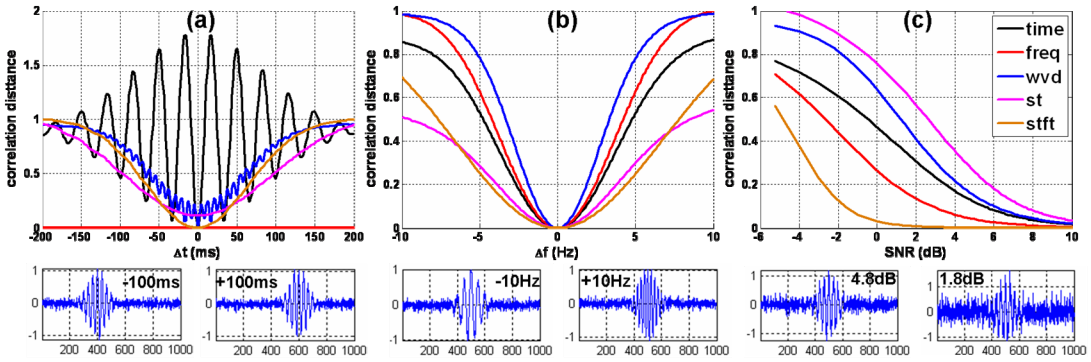

Figure 14: The sensitivity of the correlation distance under different signal representations to simulative changes of the temporal location (a), the frequency content (b) and the signal-to-noise ratio (c). The baseline simulated signal is a 300ms, 30Hz sinus with a Gaussian amplitude modulation and additive Gaussian white noise of 7dB. Bottom panels present examples of the simulated signals. Time-frequency representations are relatively robust to temporal shifts, compared to over-sensitivity of time-domain representation and insensitivity of spectral representation (a). WVD, spectral and time-domain representations are more sensitive to frequency changes than ST and STFT (b). WVD and ST are more sensitive than STFT to changes in the signal-to-noise ratio (c).

The simulation results showed that the time-frequency representations are more robust to small alignment differences between the signals, and identified significant differences in the sensitivity to changes of the frequency and the noise level. The choice of signal representation is therefore tightly related to the nature of the variability in the data. In cases where the data exhibits large variability between classes and small variability within each class, highly-sensitive representations would provide more accurate results, whereas when the changes in the data are more gradual and there is small between-class variability or large within-class variability, a representation that is less sensitive but more robust should be preferred. In order to evaluate the different representations on real data,

as well as to study the effects of the chosen distance metric and classification algorithm, the analysis framework was used to analyze the HSPRS and HSDSE dataset. This analysis is described in detail in sections 4.3, 5.2 and 5.3.

3.3 Segmentation and alignment of heart signals

3.3.1 Problem definition

The inherent variability of vibro-acoustic heart signals is realized in slight beat-to-beat changes of the components comprising the signals. Consequently, characterization of these components in time and frequency becomes a non-trivial task. The purpose of segmentation is to partition the signal into segments that represent distinct components or events in the cardiac cycle. Formally, a segmentation of a signal *s(t)* is a function $g : R \to N$, which associates a segment label $L_k \in N$ to each time sample of the signal: $g(s(t_k)) = L_k$. In order for the segmentation task to be a well-defined optimization problem, we should have a quantitative measure of the agreement between the segmentation and the 'actual' structure of the signal. In the supervised case, the signal is manually or semi-automatically annotated, and the comparison between the reference annotations and the automatic segmentation is straightforward. However, in most cases such reference annotations are not available, and the quality of the segmentation should be evaluated in an unsupervised manner. We suggest that the cyclical nature of heart signals can be beneficial for segmentation, as repetitious components that are similar in multiple signal cycles can be identified. Comparison of components between cycles requires the compared signal to be aligned. Segmentation and alignment are interdependent problems [99]. If the signals are perfectly-aligned, signal averaging will highlight the repetitious components while suppressing irregular random noise between components, simplifying the segmentation. However, in practice, alignment of physiological signals is non-trivial: beat-to-beat morphological changes can take many forms including temporal shifts, baseline drift, amplitude scaling, time scaling and additive noise. In fact, alignment can benefit from prior identification of the major components to be aligned, or in other words – prior knowledge of the segmentation will simplify signal alignment. In the following sections we describe several techniques for alignment and segmentation of vibro-acoustic heart signals, and discuss their properties.

3.3.2 Signal alignment by phase-shift averaging

Phase shift averaging is an iterative procedure of multiple signal alignment [100]. The steps of the procedure in the i^{th} iteration are as follows:

1. Compute the ensemble average of the signals: $\hat{S}^i(t) = \frac{1}{N}\sum_{k=1}^{N} s_k^i$
2. For each signal s_k^i, compute the phase shift that maximizes the cross-correlation between the signal and the average: $\tau_k^i = \arg\max_{\tau}\left\{s_k^i(t) * S^i(t+\tau)\right\}$
3. Shift each signal by the computed phase: $s_k^{i+1}(t) = s_k^i(t+\tau_k^i)$
4. Repeat stages 1-3 until a convergence criterion is met. Such criterion might require small or zero phase shifts of all signal, small standard deviation, high average correlation with the average signal, or maximal number of iterations (Figure 15).

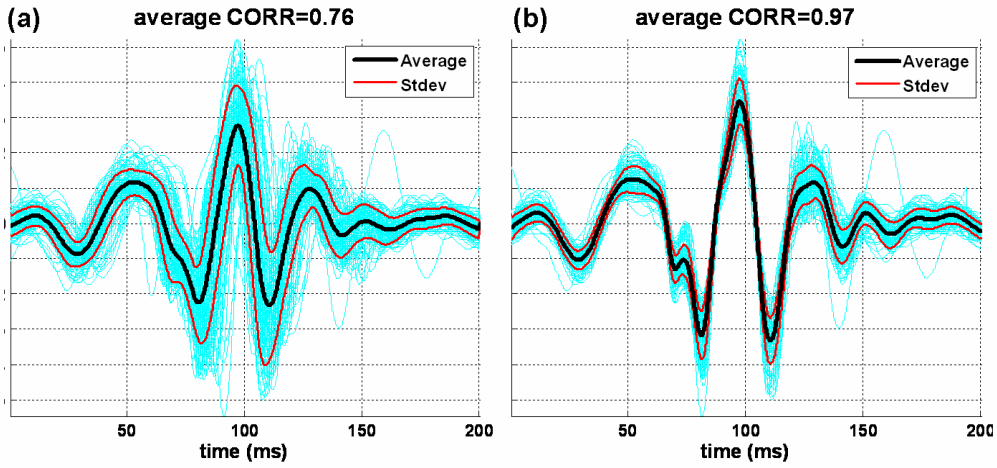

Figure 15: Phase-shift averaging of S1 signals. 109 cycles of S1, initially aligned by the peak of their energy envelope, exhibit large variability (a). Following one iteration of PSA, the standard deviation is significantly reduced, and the average correlation with the average signal increases from 0.76 to 0.97. The phase-shift averaged signal is therefore a more accurate representation of the signal's morphology.

PSA is a suitable alignment procedure when the transformation between signal cycles is of the form: $s_j(t) = s_i(t+\tau_{ij})$, i.e. there is a constant time delay between pairs of cycles. However, in case different components of the signal exhibit different temporal delays (Figure 16a), PSA will not be able to produce correct alignment for all components, resulting in inaccurate averaging of the weaker components.

3.3.3 Signal alignment by dynamic time warping

An alternative to the constant time shift of PSA alignment is to consider a non-linear time scaling transformation between signals by applying dynamic time warping (DTW). DTW aligns two time series by recursively 'stretching' them in time so that the cumulative distance between the series is minimal. DTW was a popular technique in speech recognition [94], and has been adopted for other applications of data mining and pattern recognition [101, 102]. DTW has been previously applied on biomedical signals, such as the electrocardiogram [103] and left-ventricular volume waveforms [104], showing its aptness for alignment of multi-cycle physiological signals.

This technique can successfully align multi-component signals, such as cycles of S1 and S2, where each component has a different temporal delay (Figure 16). Furthermore, the warping path provides a complete mapping between the samples of the two signals, so in case the segmentation of one signal is known, the segmentation of the other signal can be easily derived.

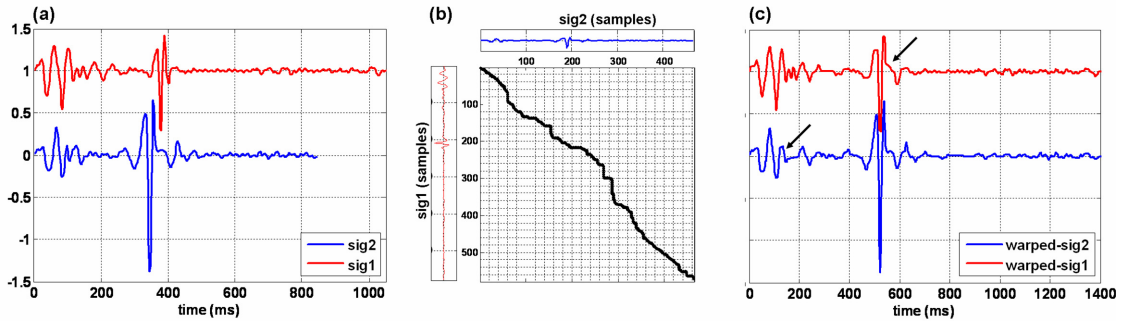

Figure 16: Dynamic time warping of heart sound signals. Two cycles of heart sounds that differ in both temporal locations and morphologies of S1 and S2 (a) were aligned using DTW, by finding the warping path with minimal cumulative error (b). Although the warped signals (c) exhibit good temporal alignment, they also exhibit some morphological distortions (black arrows) compared to the original signals. These distortions blur the physiological meaning of the signal's morphology.

Conversely, conventional DTW has some drawbacks as an alignment procedure for biomedical signals. As a general algorithm for time-series alignment, it does not necessarily preserve the morphology of the warped signals. For biomedical signals, the morphology of the components has a physiological meaning, and the transformation imposed by the warping process might be physically implausible. The signals could be mathematically aligned, but the distortions of their morphologies might cause this alignment to be physiologically meaningless (Figure 16c). In addition, DTW does not

account for partial matching, where one of the signals has a component that is missing from the other. Partial matching is a common situation in vibro-acoustic heart signals, whose inner-structure can change from beat to beat, as a result of modulating processes such as the respiratory activity, or irregular occurrences of pathologic sounds such as S3 and S4.

3.3.4 Multi-scale event-based segmentation

In order to constraint the signal alignment transformations to operations that are physiologically realizable, the 'natural' morphology of the signal, i.e. the features that best define the morphology, should be identified and preserved in the alignment process. In low-frequency pulse signals such features may be minima, maxima and inflection points of the signal. We use a multi-scale extrema detection procedure [105] to detect extrema points and associate each detected point with a scale value. The scale indicates whether the point is a global or a local extremum of the signal, and accordingly scores the point's relative importance in the description of the signal's morphology. The steps of the procedure are as follows (Figure 17):

1. The signal is smoothed by a convolution with a Gaussian kernel with an increasing standard deviation parameter, constructing a scale space representation.
2. At each scale, all maxima and minima points are detected by computing the numeric derivative of the smoothed signal.
3. Each detected point is tracked down the scale axis, from coarse to fine scale, in order to localize it in time.
4. Points with high scale are selected as candidate segmentation points.

Following the detection of morphologically important points in each cycle of the signal, segmentation can be carried out by grouping together similar points from multiple cycles, and then choosing the groups of points that have the highest repeatability and lowest inter-cycle variability [106]. Point alignment is achieved by applying DTW between the series of detected point in each cycle and in a reference cycle, allowing each point also to be 'unmatched'. Selection of the best segmentation points is done by grading each point according to the following parameters: repeatability, namely the portion of cycles in which the point was detected, standard deviation in time and in amplitude, and maximal

scale of detection. Examples of applying this algorithm on heart pulse signals are given in Figure 18.

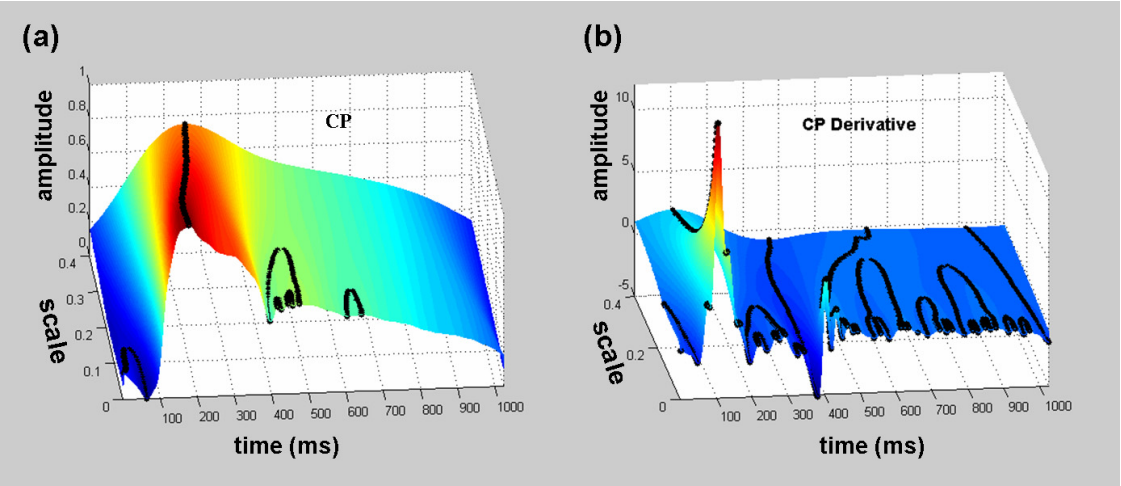

Figure 17: Scale-space extrema detection. A carotid pulse signal (a) and its derivative (b) are smoothed by a Gaussian filter with increasing standard deviations, creating a scale-space representation. Minima and maxima points are detected in the signal for each scale, and tracked from coarse to fine scale in order to identify their precise location in time.

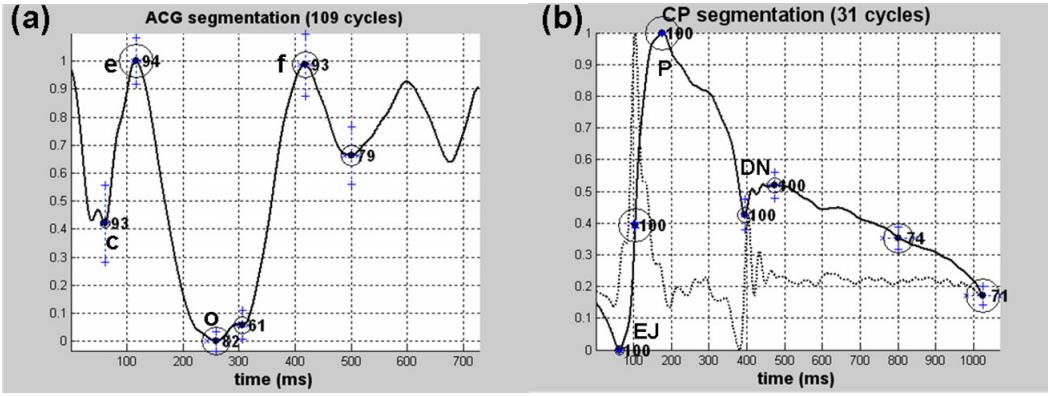

Figure 18: Event-based segmentation of 109 apexcardiogram (ACG) signal cycles (a) and of 31 carotid pulse (CP) cycles (b). The horizontal and vertical error bars indicate temporal and amplitude variability. The empty circles indicate scale and the numbers indicate repeatability percentage.

3.3.5 Time-frequency component-based segmentation

A second approach of 'component-oriented' segmentation is to first identify the major components in the signal's time-frequency energy distribution, then to use these components for alignment of multiple cycles and recognition of the repetitious and coincidental patterns in the signal. A method for the segmentation of time-frequency distribution using statistical features of the STFT [107] was applied to heart sound

signals. This method is based on the principle that STFT coefficients that contain only noise have zero-mean Gaussian distribution, while coefficients that also contain some of the signal's energy do not have zero-mean Gaussian distribution, and their second-order moment is higher. The segmentation algorithm is iterative, with each iteration consisting of the following steps:

1. Noise estimation on the non-segmented regions.
2. Local second-order moment estimation of all time-frequency locations.
3. Selection of candidate locations of signal component, using a threshold value calculated from the estimated noise and a parameter of false detection probability.
4. Region growing around the candidate components: starting at the candidate with the highest variance, neighbor candidates are assigned a joint label, creating spectral patterns of each label, until most of the candidates are segmented.

The algorithm stops when the distribution of the remaining non-segmented regions is close enough to Gaussian distribution. Gaussianity is measured by estimating the excess kurtosis, which is the fourth standardized moment, minus the kurtosis value of the normal distribution, as defined in equation (3.19)

$$K = \frac{\mu_4}{\sigma^4} - 3 \tag{3.19}$$

The application of the time-frequency segmentation algorithm on multi-cycle heart sound signal with S1, S2 and S3 components [108] is demonstrated in Figure 19. On a STFT calculated with a Hanning window of length 31 samples (28 ms), window overlap of 29 samples and 1024 frequency bins, two iterations of the time-frequency segmentation algorithm successfully identified all occurrences of the heart sounds.

Following component segmentation, the identified signal segments should be classified to distinct groups, in order to recognize repeating patterns and to point out irregularities. For this unsupervised classification task we propose two alternative approaches: component clustering and dynamic component warping. In component clustering, hierarchical clustering is applied to either STFT-derived components or segments of the original signal, defined by the boundaries of the STFT components. Since the components are non-aligned and differ in size, the clustering procedure first aligns each pair of compared components by maximizing their cross-correlation, then measures the distance between

them. Representative results of component clustering are shown in Figure 20, where segmented S1, S2 and S3 were clustered into separate clusters.

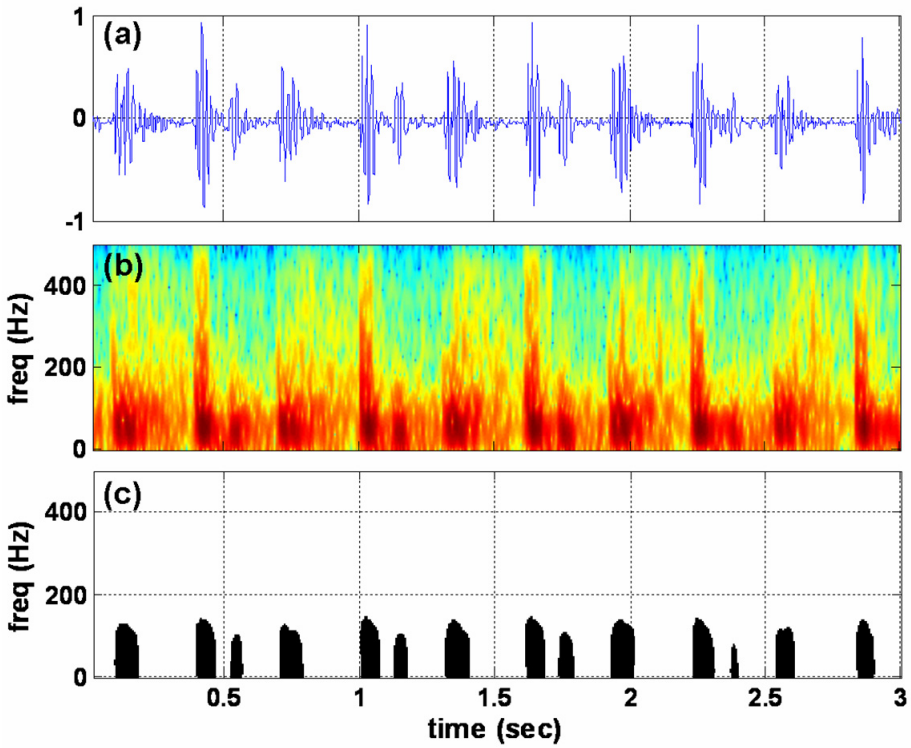

Figure 19: Time-frequency component segmentation. A continuous heart sound signal with S1, S2 and S3 (a) was segmented by calculating its STFT (b) and identifying the major components in the time-frequency representation (c).

The combination of time-frequency segmentation and clustering provides an automatic method of heart sound segmentation without ECG-based cycle partitioning and without using any heuristic information about the expected structure of the signal. In fact, this method can be applied to any cyclical signal with multiple components in the time-frequency plane.

In a scenario where the external partitioning to cardiac cycles is given, component classification can be obtained using dynamic cycle alignment of the identified components. This component-warping procedure aligns a pair of signal cycles by finding the component mapping that minimizes the overall correlation distance. The temporal order of components must be retained, but some of the components are allowed to remain unmatched, and multiple small components can be mapped to non-overlapping regions of a larger component. This technique enables subtle matching of subcomponents, as shown

in Figure 21: the first cycle has an additional component of S1, while the second cycle has an additional component of S2. The alignment procedure identified the similarity between the main components of S1 and S2 in the two cycles, and correctly matched these components, allowing the additional components to remain unmatched.

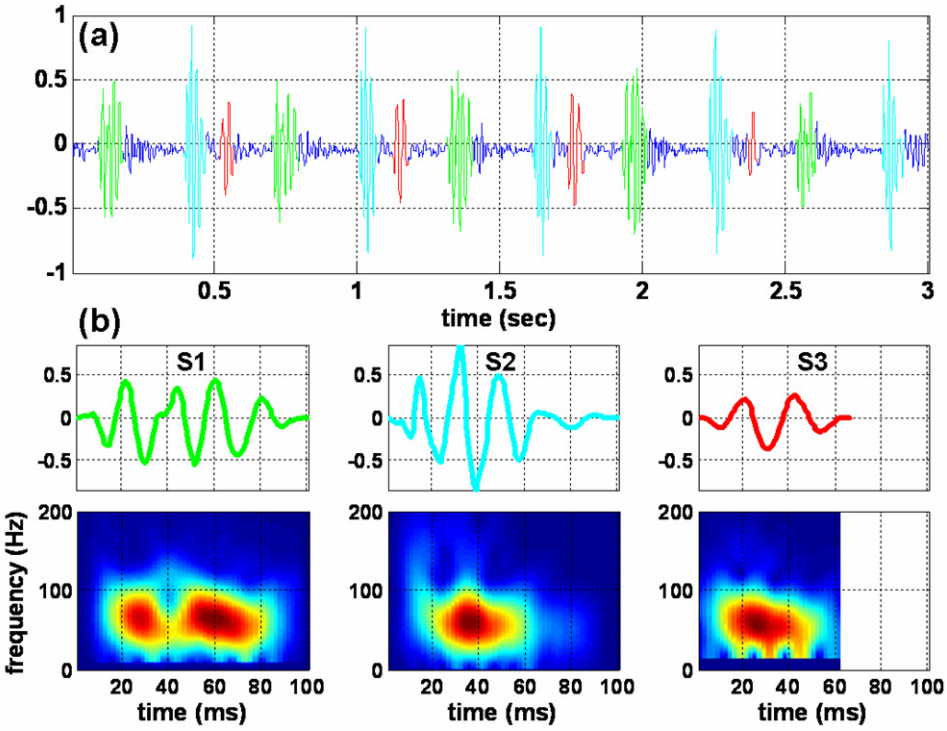

Figure 20: Time-frequency component clustering. A signal section (a) from a continuous 27-second heart sound signal that was segmented by time-frequency segmentation (Figure 19), and clustered into three morphological clusters using STFT representation (b). Segments of S1 (green), S2 (cyan) and S3 (red) were assigned to distinct clusters.

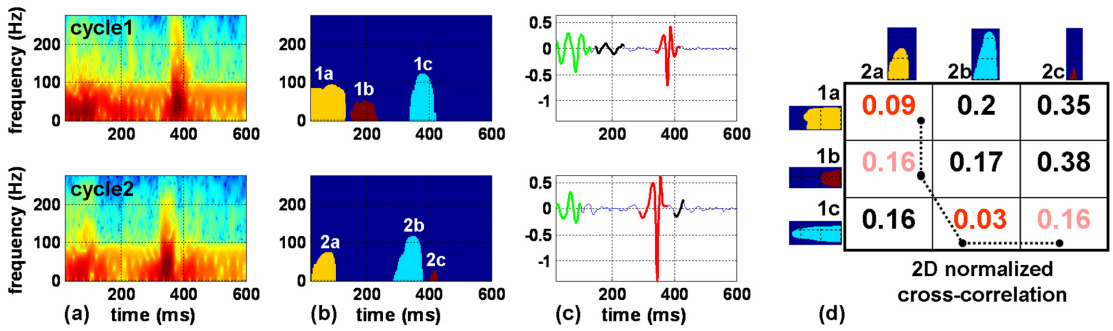

Figure 21: Time-frequency component warping. Two heart sound cycles with different morphologies of S1 and S2 were segmented using STFT (a). In the first cycle, two components of S1 were detected, and in the second cycle two components of S2 were identified (b). Component warping using two-dimensional cross-correlation distance matrix (d) matches the major components of S1 (green) and S2 (red) and marks unmatched components (black), thus providing signal segmentation.

Unlike the DTW alignment of the same pair of signals (Figure 16c), which distorts the morphology of the original signal, component-warping provides signal segmentation and alignment which preserves the physiological meaning of the subcomponents.

In summary, segmentation and alignment of heart signals can be significantly improved by adjusting the processing techniques to the morphological properties of the signal and by exploiting its cyclical nature. Accurate identification of the key events or the major components in the signal, either in the time-domain or in the time-frequency representation, can be used as a preceding step for a cycle-by-cycle alignment, which does not distort the morphology of the signal and maintains the physiological meaning of its inner-structure. The cyclicality of heart signals allows to discriminate between repetitious and coincidental patterns, making the proposed segmentation techniques more robust to noise and artifacts that are common in physiological signals.

CHAPTER 4
EXTRACTION OF CARDIAC FUNCTION INDICES FROM HEART SIGNALS

The various analysis techniques described in the previous chapter were used for studying the relations between cardiac function and vibro-acoustic heart signals and for extraction of vibro-acoustic information that has physiological significance and can be used to assess the mechanical function of the heart.

4.1 Temporal location of cardiac events

In order to assess the technical feasibility of extracting physiological features from vibro-acoustic heart signals, pilot signal recordings have been obtained from a single subject simultaneously with continuous-wave Echo-Doppler. The following sections describe the developed methods for heuristic segmentation of low-frequency pulse signals, identification of cardiac cycle events, and extraction of temporal features estimating the systolic and diastolic time intervals of the cardiac cycle.

4.1.1 Doppler-audio processing

Spectral analysis of the recorded Doppler-audio (Figure 22a) with short-time Fourier transform was used to reconstruct the Doppler sonograms and estimate reference values for the temporal location of events in the cardiac cycle. The spectral coefficients at each time instant, representing the measured Doppler frequency shifts, were converted into velocities using the Doppler equation (2.1), with parameters $\theta = 0^{\circ}$ (assuming the transducer is parallel to the direction of blood flow), $c = 1540\, m/\mathrm{s}$ (the speed of sound in tissue) and $f_0 = 2MHz$ (transmission frequency). The positive velocities, representing the movement of blood towards the transducer, quantify the blood flow through the mitral valve during diastolic ventricular filling. The negative velocities quantify the flow of blood through the aortic valve during systolic ejection. The reconstructed sonogram was preprocessed by filtering-out velocities with amplitude below the estimated noise level. The instantaneous flow through the mitral and aortic valves was estimated by integrating the intensities (I) of the positive and negative velocities (v), respectively, at each time instance, according to equation (4.1), followed by smoothing with a moving-average

filter (Figure 22b). Since the Doppler audio processing of the echocardiograph imposes a constant time delay of about 150ms between the time of the cardiac event and the time of the transmitted synthesized audio, the resulting profiles were shifted backwards in time, to compensate for this delay. The temporal location of the aortic ejection wave and the mitral filling waves were manually annotated in the instantaneous flow profiles. These annotations were used to estimate the reference systolic and diastolic time intervals and the ejection and filling amplitudes.

$$F(t) = \sum_{v} v(t) I(v,t) \tag{4.1}$$

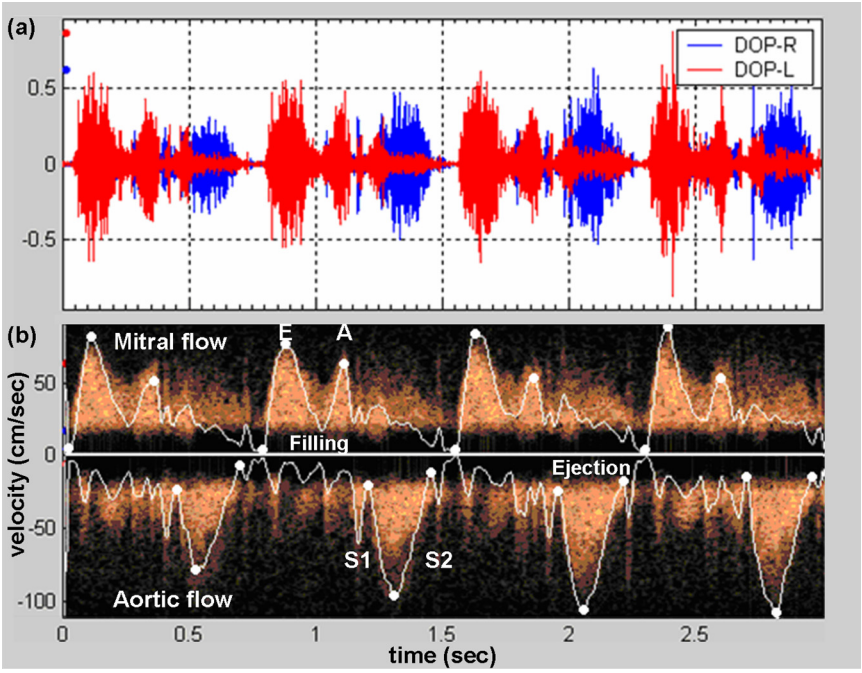

Figure 22: Doppler-audio processing. Acquired waveforms of CW-Doppler audio channels (a) are processed by short-time Fourier transform to reconstruct the sonogram of mitral and aortic flow (b). In the profile of the instantaneous mitral flow the two diastolic filling waves (E and A) can be observed, while in the profile of aortic flow the ejection wave is identified, delineated by traces of mitral valve closure (S1) and aortic valve closure (S2).

4.1.2 Vibro-acoustic signal analysis

The low-frequency pulse signals, acquired at the carotid artery (CP) and at the apex (ACG) were segmented by a heuristic algorithm, which uses the temporal location of the first and second heart sounds, and a-priory knowledge about the expected morphology of the signal, to detect the location of points of interest. Each pulse signal was pre-processed and partitioned into cardiac cycles, as describe in section 2.5. The frequency band of the

heart sounds (20-205Hz) was segmented to identify the energy peaks of S1 and S2. PSA algorithm, described in section 3.3.2, was used for adjusting the temporal location of the two major heart sounds. In the CP signal, the algorithm detected the points indicating the beginning of ejection (E), the peak of ejection (P) and the dicrotic notch (DN), using the following heuristics (Figure 23a):

- Point '*E'* was identified as a local minimum in the proximity of S1 and the maxima of the second derivative of the CP signal.
- Point 'P' was identified as a local maximum between S1 and S2.
- Point 'DN' was identified as a local minima in the proximity of S2.

In the ACG signal, the algorithm detected the points indicating phases of contraction (*c*), ejection (*e*), relaxation (*o*), filling (*f*) and atrial contraction (*a*), using the following heuristics (Figure 23b):

- Point '*e'* was identified as a local maximum in the proximity of S1.
- Point '*c'* was identified as a local minimum preceding point *e*.
- Point '*o'* was identified as local minimum in the proximity of S2.
- Point '*f*' was identified as local maximum following point *o*.
- Point '*a*' was identified as a local maxima preceding the ECG's R-wave.

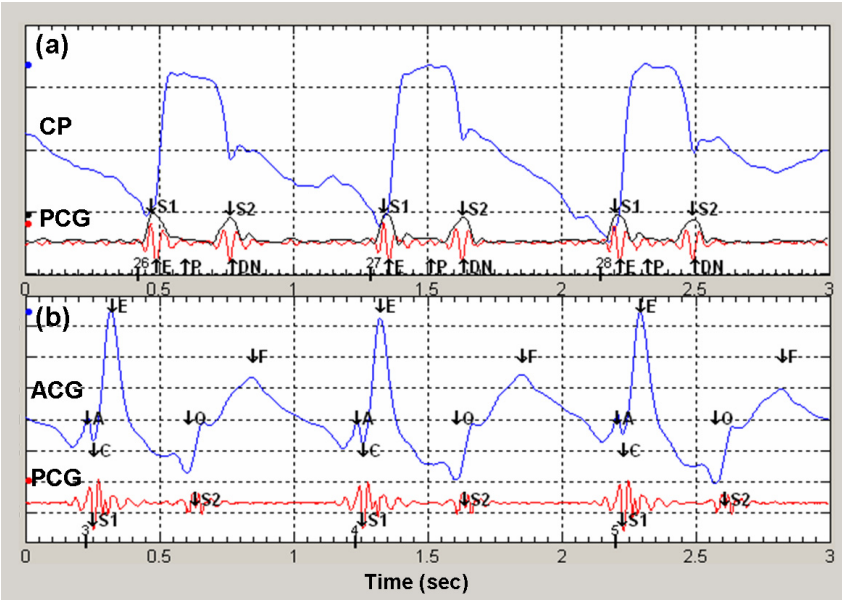

Figure 23: Identification of events in the carotid pulse (a) and apical pulse (b) signals. Cardiac cycles are defined by ECG. In each cycle, S1 and S2 are identified in the envelogram of the phnocardiogram (PCG) signal. Their location is used to identify significant points in the CP and ACG signals.

Using the identified temporal locations of cardiac cycle events, the following systolic and diastolic time intervals were extracted from the CP and ACG signals:

- Pre-ejection period (PEP).
- Ejection time (ET).
- Isovolumetric relaxation time (IVRT).
- Filling time (FT).

4.1.3 The relationship between vibro-acoustic signals and the Doppler profile

The temporal relationship between the carotid pulse signal and the pattern of aortic blood flow, obtained by Continuous-Wave Doppler, is illustrated in Figure 24. The rapid upstroke in the CP signal, indicating the beginning of rapid ejection, coincides with the beginning of the Doppler ejection wave. The dicrotic notch coincides with the closure of the aortic valve (S2) and the end of the Doppler ejection wave. Figure 25 demonstrates the temporal relationship between the apexcardiogram signal and the pattern of lateral ventricular wall movement, obtained by tissue-Doppler imaging (TDI). The low-frequency content of the apex pulse during diastole matches the two negative Doppler waves of early ventricular relaxation and late atrial contraction. Thus, information fusion from both pulse signals, as well as from sound signals, enables to locate events of the cardiac cycle and to measure the duration of the cardiac phases with high accuracy.

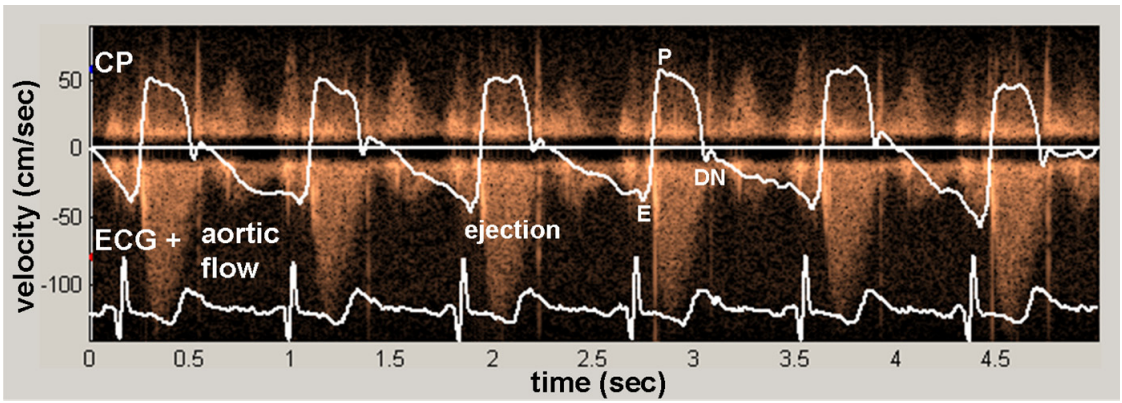

Figure 24: The relationship between simultaneously-acquired carotid pulse (CP) signal and reconstructed CW-Doppler sonogram of aortic blood flow. Point 'E' in the CP signal coincides with the beginning of the negative ejection wave, and point 'DN' coincides with the end of ejection and the closure of the aortic valve.

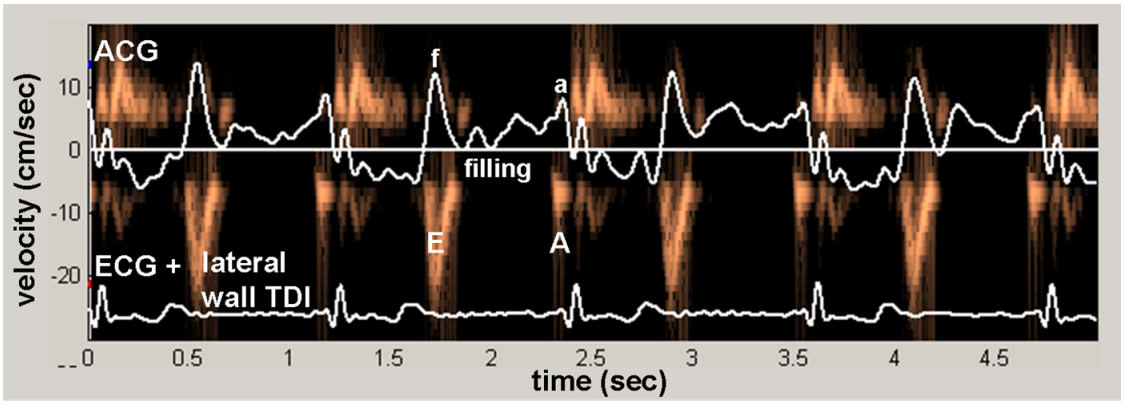

Figure 25: The relationship between simultaneously-acquired apex pulse (ACG) and the velocity profile of the lateral wall, measured by TDI. During diastole the ACG signal consists of 'f' and 'a' waves, which correspond with the negative waves (E and A) of diastolic wall movement during rapid and late ventricular filling.

The average values of the systolic and diastolic time intervals, estimated from the pulse signals and the reference CW-Doppler profile of a healthy subject are given in Table 3. A good agreement was observed between the time intervals measured by both methods in terms of average and standard deviation values. The beat-to-beat filling time, derived automatically from the apexcardiogram signal, showed a strong statistical correlation with the Doppler-derived filling time ($r=0.93$, $p<1e\text{-}40$). The average difference between the instantaneous filling times measured by the two methods was 0.77±11.6 ms (Figure 26). Statistically-significant correlations were also observed for the ejection time and ejection amplitude, derived automatically from the carotid pulse signal. However, as the physiological variations of these measures are very low in rest conditions, the inherent measurement errors of both methods mask the small beat-to-beat physiological changes, thus weakening the correlation between the instantaneous measurements. Dynamic modification of cardiac function is therefore required to elucidate the correlation of systolic features. Such conditions can be achieved during dynamic exercising or pharmacological stress test.

Table 3: Systolic and diastolic time intervals (ms) derived from apexcardiogram (ACG), carotid pulse (CP) and CW-Doppler of 109 heart cycles recorded from a healthy subject

Time interval	ACG (ms)	CP (ms)	Doppler (ms)
PEP	59.4 ± 1.8	62.4 ± 5.7	66.4 ± 5.9
ET	268.2 ± 4.1	262.2 ± 7.3	262.1 ± 7.1
IVRT	88.9 ± 10.1	-	90.1 ± 9.3
FT	313.1 ± 29.7	-	312.3 ± 30.6

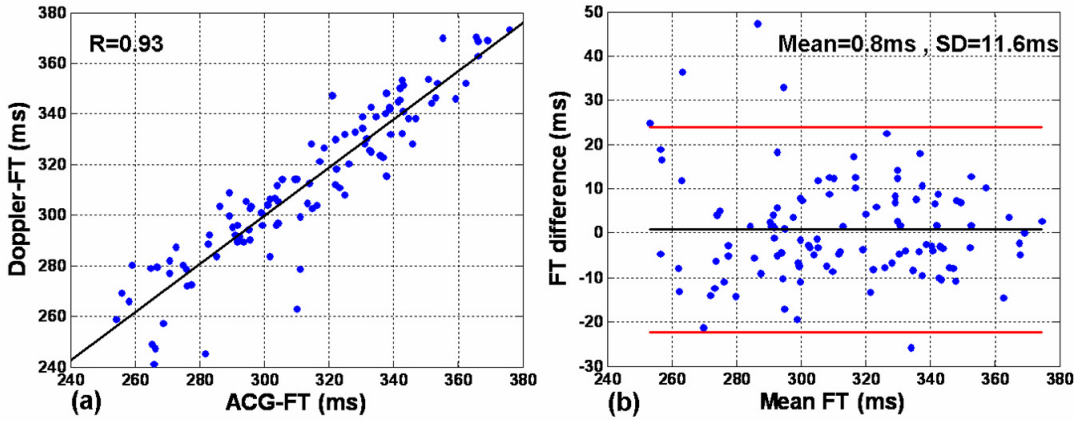

Figure 26: Correlation (a) and statistical agreement (b) between filling time derived from apexcardiogram and CW-Doppler of 109 heart cycles recorded from a healthy subject.

4.2 Left-ventricular systolic function

During a pharmacologically-induced stress response, there are marked changes in left-ventricular systolic function due to the increased cardiac contractility. The HSDSE data set was analyzed in order to identify and characterize spectral features of S1 which can be associated with the increased cardiac contractility. The physiological association was assessed by comparing these acoustic indices to reference strain-echocardiography indices that are related to left-ventricular systolic function [109].

4.2.1 Echocardiography data processing

The HSDSE data set contain two-dimensional echo cine loops of a single heart beat, captured at each stage of the stress test. The echo data was post-processed using EchoPAC Dimension ’06 software (GE Healthcare Wauwatosa, WI) in order to calculate peak systolic velocity (PSV) and peak systolic strain rate (PSSR). The calculation of strain indices was done using 2D strain analysis, based on speckle tracking technique (Figure 27). This technique allows objective analysis of the entire myocardial motion throughout the heart cycle by tracking natural acoustic markers in the image. It was shown to provide accurate strain measurements [110]. PSV and PSSR indices were first calculated separately for each cardiac wall (septal, lateral, inferior, anterior, posterior, and anteroseptal) and for three segments per wall (basal, middle and apical), and then averaged to obtain an index of global systolic function (Figure 28). Strain indices were

successfully calculated for 10 patients. One patient was excluded due to inadequate quality of the captured echo images.

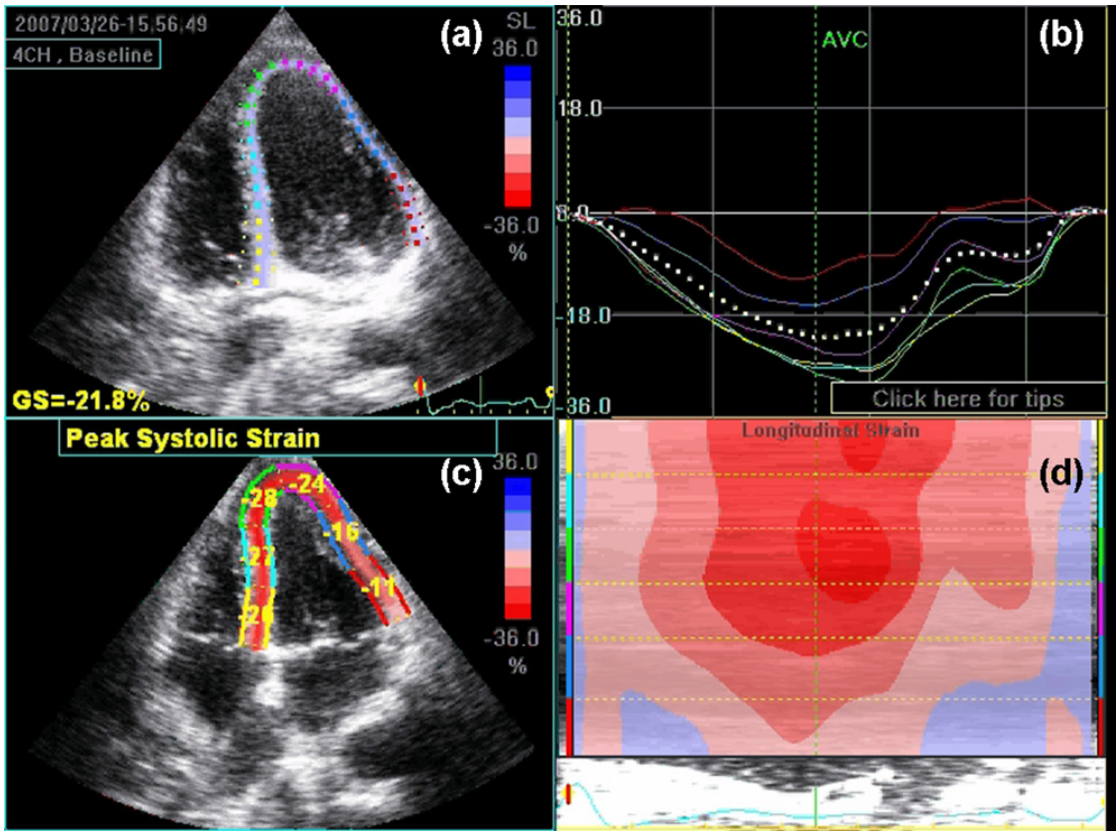

Figure 27: Strain-echocardiography analysis. The contour of the left ventricle is semi-manually defined on a 2D echo image (a). Speckle tracking is used to track basal, middle and apical segments of the lateral wall and septum during a single cardiac cycle, and to compute instantaneous strain for each tracked point (d) and for each segment (b), as well as average global strain (dotted line). The point of end systole (AVC) is used to calculate the peak systolic strain in each segment (c).

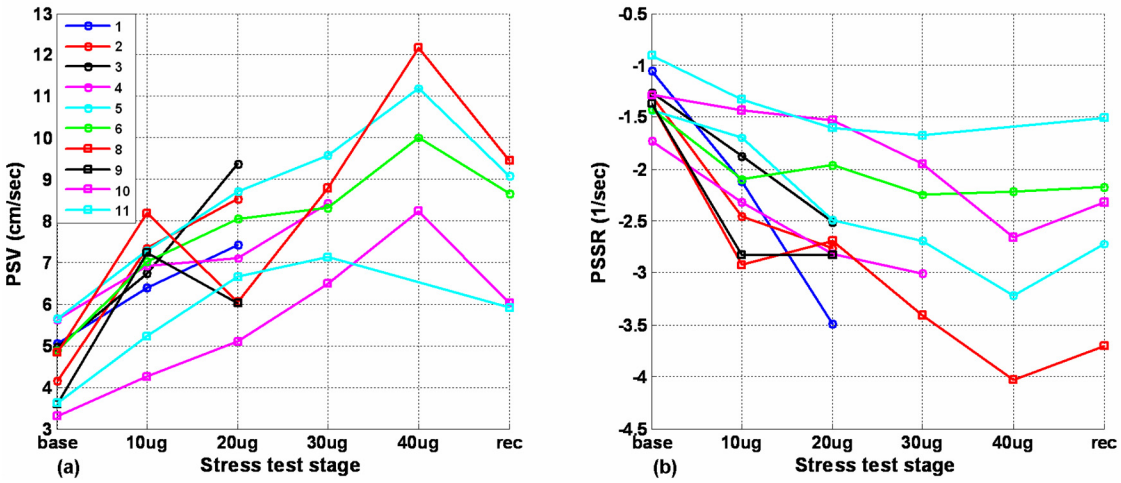

Figure 28: Strain-echocardiography indices. Global peak systolic velocity (a) and peak systolic strain rate (b) of 10 patients during different stages of Dobutamine stress test.

4.2.2 Acoustic signal processing

Each of the four recorded heart sound channels was first pre-processed and partitioned into cardiac cycles as described in section 2.5 (Figure 29a). The signal cycles were aligned by their starting points and their amplitudes were color-coded to create a two-dimensional signal map, showing the time-domain dynamics of the first and second heart sounds throughout the stress test (Figure 29b). S1 segments were extracted, and fast Fourier transform (FFT) was applied to each cycle of S1. The logarithm of the power spectrums was color-coded to generate a spectral map of S1 throughout the recording (Figure 29c). In order to characterize the joint time-frequency energy distribution of S1, S-transform was applied to each cycle of S1, and the resulting time-frequency representations were grouped by the stages of the stress test and averaged to produce a small number of representative time-frequency maps (Figure 30).

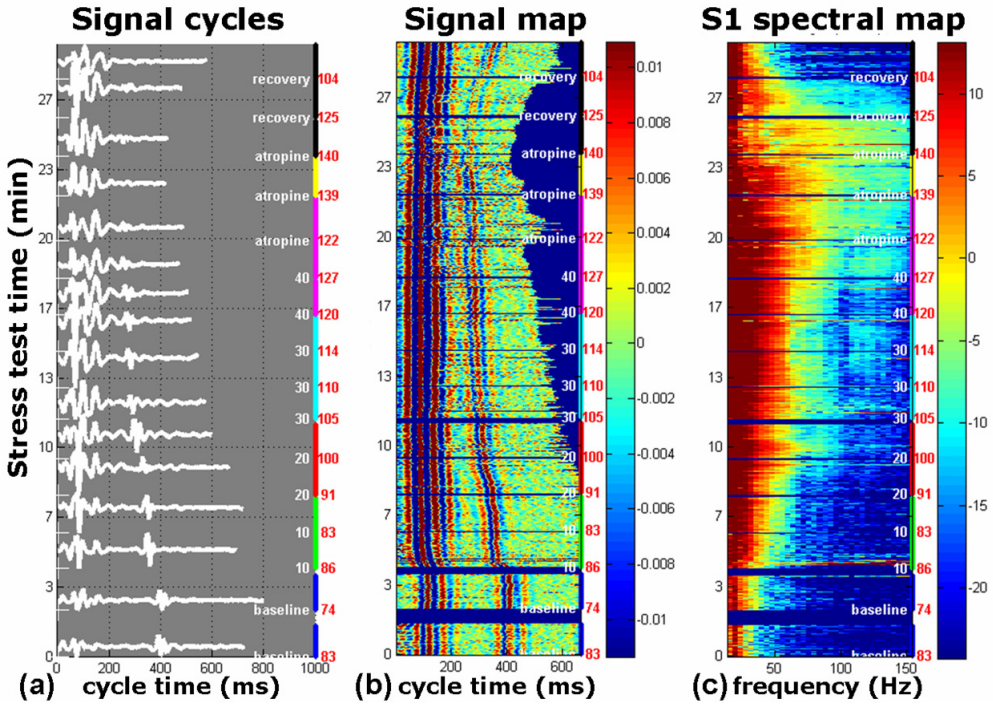

Figure 29: Heart sounds during stress test. (a) Aligned average heart sound cycles of each stage of the stress test, with heart rate (red labels) and test stages (white labels and colored segments). (b) Continuous color-coded map of all heart sound signals. (c) Continuous color-coded power spectrum of the first heart sound (S1).

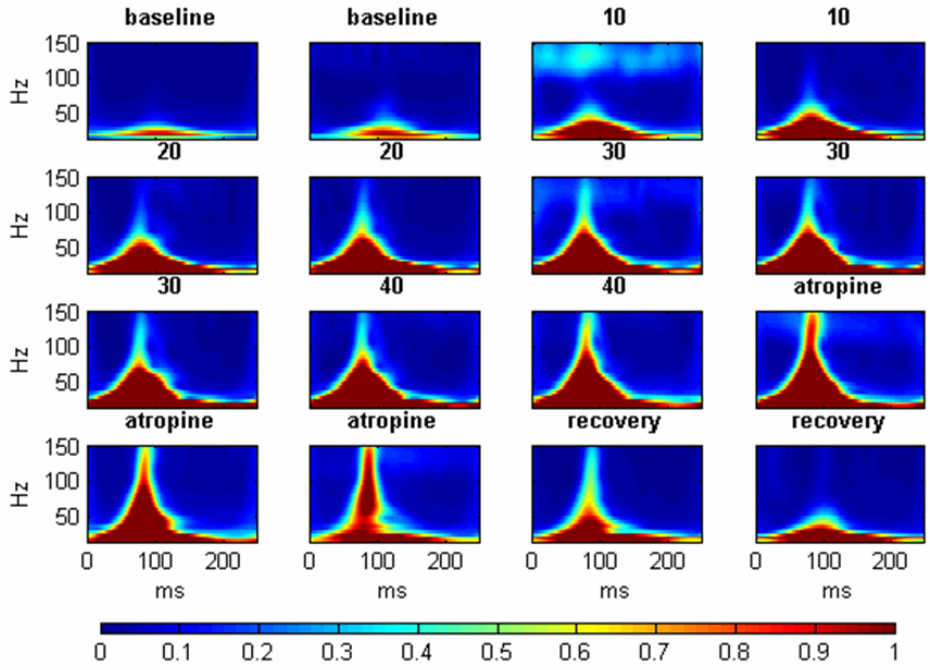

Figure 30: S-transform time-frequency representation of S1 acoustic signal obtained in a representative healthy subject during the stages of the stress test. Each plot represents an average of the S-transform of all S1 cycles over a specified period of the test.

PCA was applied on the aggregation of segmented S1 signals. The analysis was performed on the time-frequency representations produced by the S-transform (Figure 30), vectorized by concatenating adjacent columns. The most significant principal components, having eigenvalues greater than 10% of the first eigenvalue, were selected and weighted by their relative eigenvalues. The projection of the data on this weighted combination of the significant principal components was chosen as a one-dimensional feature representing the dynamic characteristics of the acoustic signal during the stress test. To obtain an interpretable trend line, this feature was normalized by the median value of the baseline stage and smoothed by a moving average filter. The resulting index, which was denoted 'acoustic variability index' (AVI), is interpreted as the trend of relative change in the spectral energy distribution of S1 (Figure 31).

A second feature extracted from the spectrum of each cycle of S1 was the frequency bandwidth of the signal, defined by the highest frequency with significant energy content. Prior to calculating this feature, signal cycles with a high wide-band energy content, compared to their local environment, were classified as noise and excluded from further processing. The bandwidth feature was calculated for each cycle by searching the

spectrum for the first frequency whose energy is at least 10dB below the maximal energy. The feature trend line obtained from all cycles was normalized by the median value of the baseline stage, and denoted 'Acosutic Spectral Index' (ASI, Figure 32).

4.2.3 Characterization of acoustic changes during stress

The color-coded signal map in Figure 29b illustrates the time-domain characteristics of the heart sound signal during the stress test. As expected, there are noticeable changes in the duration of ventricular systole and diastole, as the heart rate increases in exercise and decreases in recovery. However, there are no apparent morphological changes in the signal that can be associated with the stress response. Fourier analysis uncovers a pattern of an ascent in the spectral energy of the first heart sound as the Dobutamine dose is increased, and a descent back to baseline levels during recovery (Figure 29c). In addition to the overall energy rise, there is also an increase in the frequency bandwidth of S1, as higher frequency components in the range of 50-150Hz emerge and strengthen.

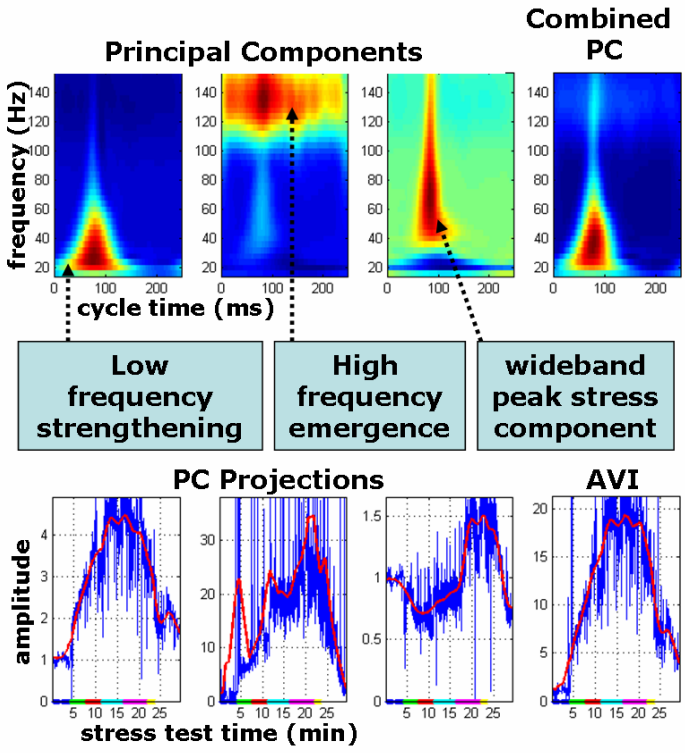

Figure 31: AVI index of S1 during stress test. The top plots show the coefficients of the first 3 principal components (PC), and their linear combination, weighted by the eigenvalues. The bottom plots show the AVI index during the entire stress test, obtained by projecting the data on the respective PC. The red lines are the result of smoothing the projected data with a moving-average filter.

The time-frequency representation, obtained by S-transform, enables localization of these spectral changes in time (Figure 30): the high-frequency components are centered about 80ms after the beginning of the signal (30ms after the peak of ECG's QRS complex), growing up to 150Hz in the highest Dobutamine dose, then falling back to the baseline upper-limit frequency of 50Hz in the recovery phase. There is no apparent time shift of the signal's energy distribution throughout the test. PCA, applied to the vectorized time-frequency distributions of S1, was able to identify the major frequency bands that contribute to the data variability and to point out the temporal location of these frequency bands. Figure 31 shows a representative example of the coefficients of the first three principal components (PC), and the projection of the time-frequency data on these principal components. The first PC, representing the axis with the largest data variability, captures the pattern already observed qualitatively in the time-frequency distributions in Figure 30: it varies from 30ms to 120ms relative to the beginning of the cycle, and from frequency of 20Hz to 70Hz, thus showing the strengthening of the signal's low-frequency components. The second PC captures the variability of the high frequency components between 110 to 150Hz for the entire duration of the S1 signal. The third PC shows a wide-band variability of frequency ranging from 40Hz to 150Hz, localized in time around 80ms from the beginning of the cycle. This component strengthens during peak stress.

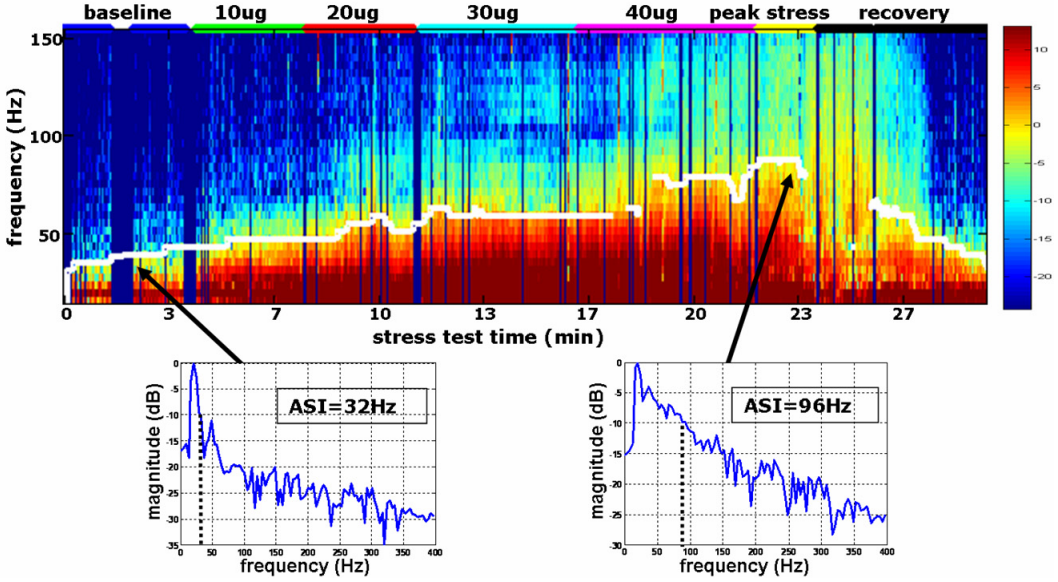

Figure 32: ASI index of S1 during stress test. Spectral coefficients are computed by FFT for each cycle of S1 throughout the stress test. ASI (white line) is defined for each cycle as the first frequency whose energy is at least 10dB below the maximal energy.

4.2.4 The relation between acoustic indices and systolic function

The AVI trend lines, extracted from each of the four transducers and averaged, are plotted in Figure 34 for four subjects, along with the stages of the stress test, the heart rate and blood pressure trends and the relative change in the echocardiographic indices of peak systolic velocity (PSV) and peak systolic strain rate (PSSR). While the AVI provides a continuous line with one point per cardiac cycle, the reference echocardiographic indices are available only at discrete time points of each stage in the stress test. Nevertheless, there are strong correlations between the two indices: the correlation coefficients between the echo indices PSV and PSSR and the corresponding AVI values, averaged over all transducers varied from 0.83 to 0.97 ($p< 0.05$ in all cases). Both paired and unpaired t-test showed that the absolute values of the acoustic spectral index (ASI) at the end of low-dose Dobutamine induction were significantly higher than the baseline values ($p<0.04$ for the 10ug stage, $p<0.003$ for the 20ug stage, Figure 33). The correspondence between the ASI and the echocardiographic indices in all of the subjects was tested by comparing the values of the relative index change at the end of the low-dose Dobutamine stages. These points were selected since the inotropic effect is more prominent at the early stages of the test. In addition, the higher heart rates at later stages of the test reduce the reliability of the tissue tracking procedure used to extract the reference echocardiographic indices. As shown in Figure 35, a good linear correlation ($r=0.78$, $p<0.01$) was observed between ASI calculated from the apex signal and the relative PSSR at the end of the 20ug stage. At the end of the 10ug stage the correlation coefficient between the two indices was 0.68 ($p< 0.03$).

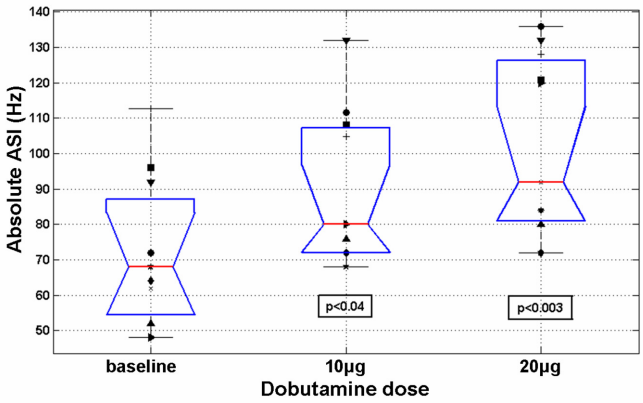

Figure 33: Absolute ASI values of all subjects at baseline and after low-dose Dobutamine induction. The box plot displays the median, lower quartile, upper quartile and data extent. Each marker symbol represents a different subject.

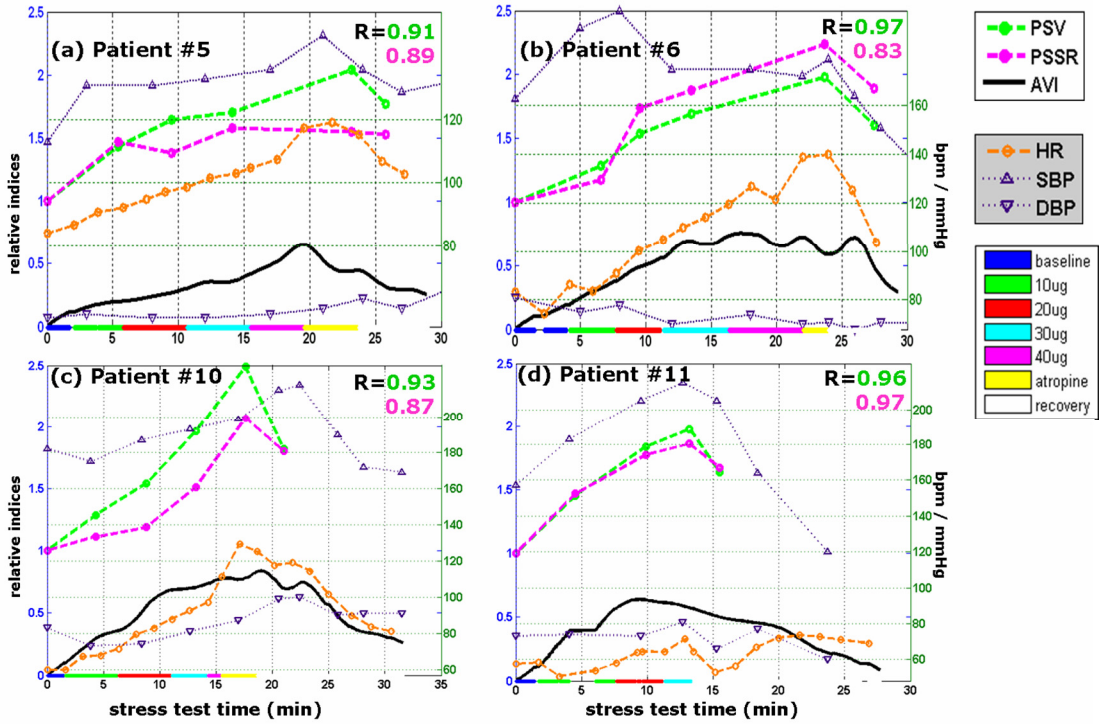

Figure 34: AVI indices of patients 5, 6, 10 and 11. Each plot displays the average trend lines of AVI from four transducers, along with the relative echo indices PSV and PSSR, trend lines of heart rate and blood pressure, and color-coded stages of the stress test.

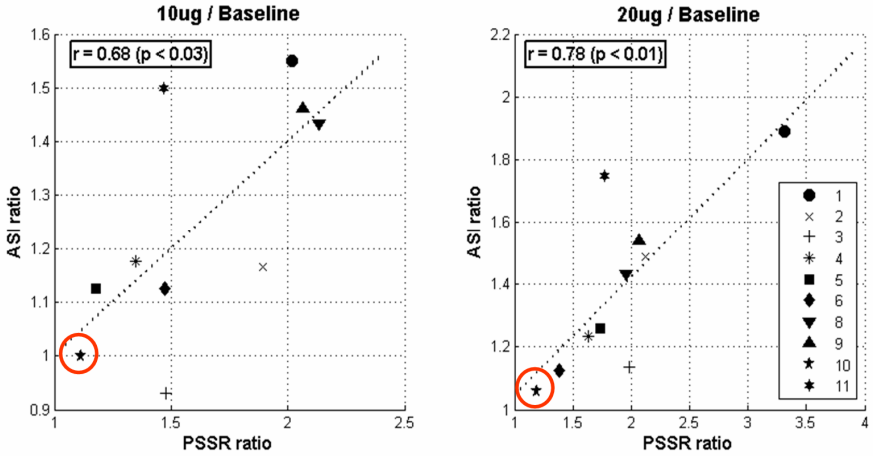

Figure 35: The correlation and regression line between relative PSSR index and relative ASI at the end of first (10ug) and second (20ug) low-dose Dobutamine induction. Each marker symbol represents a different subject.

4.3 Unsupervised analysis of heart sound morphology

The clustering and classification framework, described in section 3.2, was applied to the HSDSE data in order to evaluate its ability to predict the level of stress from the morphology of S1 [111]. For each stage of the stress test in the HSDSE dataset, the middle 1/3 of the data was used for testing, and the remaining 2/3 of the data were used for training. Each beat was associated with the label of the corresponding stress-test stage. The number of classes varied from 5 to 7 between subjects. All subjects had a 'baseline' and 'recovery' classes, and a varying number of stress stages. Analysis was performed separately on S1 signals from each of the four heart sound channels. Classification results from all four channels were combined by a majority vote scheme. Noisy test beats that were assigned to non-significant clusters in more than two channels were excluded (mean 31±13 test beats, 3.7% of the test set). Cluster analysis was applied on the training data with the required number of clusters set to 16. Before clustering, the signals were aligned by shifting each cycle to maximize the cross-correlation with an arbitrary reference cycle. Significant clusters were defined as clusters containing at least 5% of the data. Label classification was done using either KNN with K=5 and mahalanobis distance or DA with mahalanobis distance, defined in (3.15). Classification performance was evaluated by computing CC1.

The average number of processed heart beats per subject was 2549±759. The number of significant clusters identified by the cluster analysis procedure varied from 4 to 10 per subject (mean 7±2). A considerable association was observed between the clusters and stages of the stress test, where each stage was dominated by 2-3 clusters (Figure 36). The same clusters were associated with the baseline and the recovery stages, indicating that the observed morphological changes were indeed induced by the stress response. Examining the average morphology of the detected clusters revealed a pattern of increase in the spectral energy and bandwidth, directly related to the stress level. This pattern, obtained by an unsupervised learning technique, is consistent with the findings about stress-induced changes of S1 described in section 4.2. Representation of beats in the feature space of cluster distances provided a good separation between beats from different test stages (Figure 37). The observed change in the cluster-distance representation of S1 was gradual and smooth, becoming more profound at higher stress

stages, and returning back to the baseline morphology at the later stage of recovery. The average rates of correct classification (CC_1) of S1, achieved by different combinations of signal representations, distance metrics and classifiers on all 11 subjects, varied from 77% to 86% (Table 4). Correlation distance performed better than Euclidean distance, and DA classifier was slightly superior to KNN classifier. The best average classification performance of 86±7% was achieved by the DA classifier on signals represented by the S-transform and clustered using correlation distance. Time-domain representation, with correlation distance and DA classifier, provided equivalently good performance, with correct classification of 85±8%. Frequency-domain representation was inferior, compared to time-domain or joint time-frequency representations. No significant differences were observed between STFT, WVD and CWD.

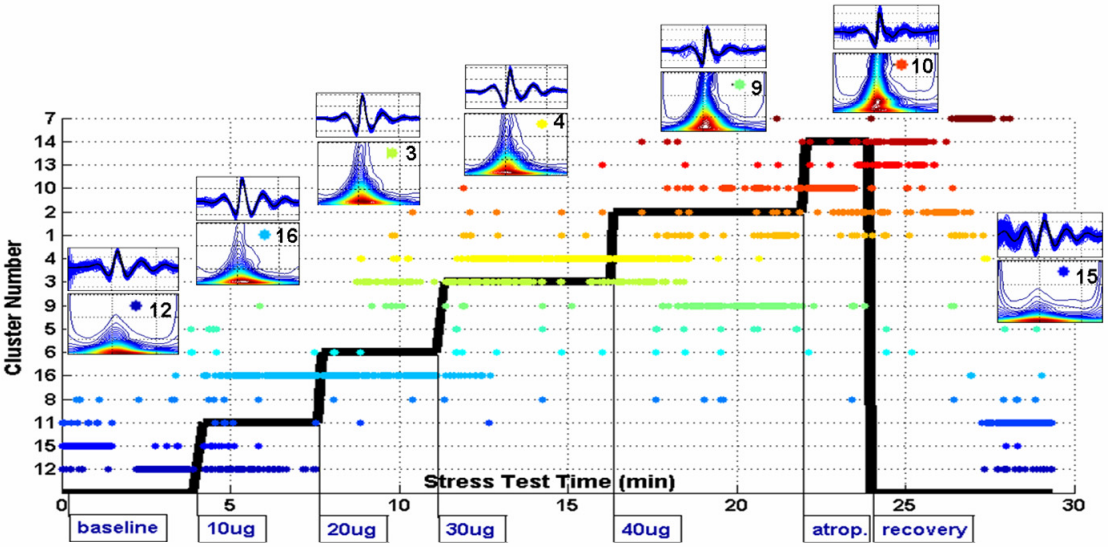

Figure 36: Clustering results of 2725 beats of S1 acquired from a single subject during 29 minutes of Dobutamine stress test. Clusters are marked by different colors and by number labels on the y-axis. The stress level is represented by the bold black line, labeled with the test stages. The time-domain and S-transform representations of the significant clusters exhibit substantial morphological changes, strongly associated with stages of the stress test, with a return to the baseline morphology during recovery.

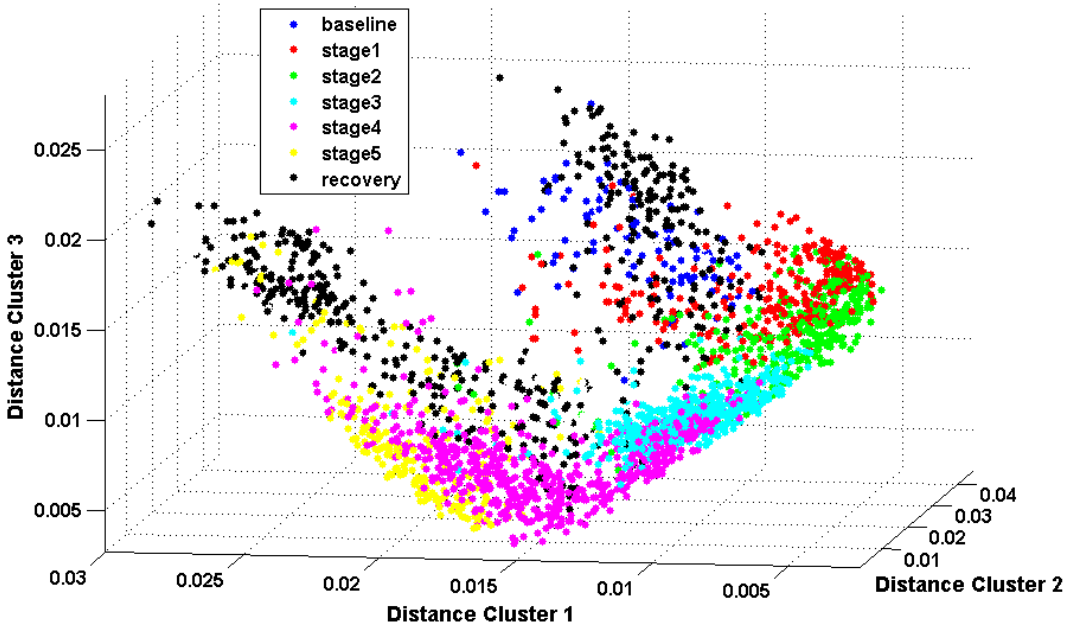

Figure 37: Cluster-distance representation of 2725 beats of S1 from a single subject, plotted by their distances from the centers of the three largest clusters. The marker colors indicate the stage of the beat in the stress test (baseline, 5 ascending stress levels and recovery). The morphology of S1 seems to vary smoothly along the stages of the test, with a distinct separation between beats of consecutive stages and a return to the baseline morphology towards the end of recovery.

Table 4: Classification performance on S1 signals from HSDSE dataset. Mean and standard deviation of correct classification measure (CC_1) of all subjects, using different configurations of signal representation, distance metric and classification algorithm (KNN=K-Nearest Neighbor, DA=Discriminant Analysis). Best results, obtained by ST and time-domain representations, are indicated by boldface.

Signal representation	**Distance metric**	**KNN CC_1(%)**	**DA CC_1(%)**
Time	**Correlation**	81±8	**85±8**
	Euclidean	81±9	82±8
Frequency	Correlation	77±9	77±11
	Euclidean	75±8	77±9
Short-Time Fourier Transform (STFT)	Correlation	80±8	80±8
	Euclidean	78±10	80±9
S-Transform (ST)	**Correlation**	**85±7**	**86±7**
	Euclidean	84±0	84±8
Wigner-Ville Distribution (WVD)	Correlation	80±9	82±9
	Euclidean	79±9	80±9
Choi-Williams Distibution (CWD)	Correlation	80±9	82±9
	Euclidean	79±9	80±9

CHAPTER 5
EVALUATION OF RESPIRATORY CONDITION BY HEART SOUND ANALYSIS

The signal analysis framework, described in section 3.2, was applied to the HSPRS dataset in order to study the effects of the respiration phase and the respiration load on the temporal and morphological properties of the first and second heart sounds, and to evaluate the ability to predict parameters of the respiratory activity from the morphology of the heart sounds [111, 112]. Analyzed data of all 12 subjects included 120 recordings of a total of 6373 heart beats acquired during normal respiration (mean±SD 531±74/subject) and additional 6275 heart beats acquired during alternations between respiration and apnea (mean±SD 523±73/subject). Cluster analysis, applied on the normal respiration recordings identified, on average, 5.5±1.6 significant clusters of S1 and 6.5±0.9 significant clusters of S2, containing 96% of the recorded beats.

5.1 Periodic morphological variability of heart sounds

5.1.1 Data analysis

Following cluster analysis, the relation between the clusters produced and the respiratory phase was first determined by assessing the morphological variability of S1 and S2 during breathing and apnea. The breathing pressure signal was automatically segmented to identify breathing activity and apnea segments. The median pressure value of the apnea segment in each file was defined as the zero-pressure. Pressure values above the zero-pressure were considered as 'expiration', and pressure values below it were considered 'inspiration'. The correlation distance between each beat and a template beat, chosen as the average of the largest cluster, was computed. The morphological variability was defined as the standard deviation of this distance, and it was computed for 15-second segments of breathing or apnea. Student's t-test was used to compare the morphological variability of S1 and S2 during respiration and during apnea across all subjects.

The periodicity of the morphological changes of S1 and S2 was evaluated by applying a robust periodicity detection algorithm [113] on the vectors of cluster-center distances. Given m beats, the vector of distances from the center of cluster k is given by:

$\vec{d}_k = \left(d_k^1, d_k^2, ..., d_k^m\right)$. $\vec{d}_k$ is non-uniformly sampled, due to the beat-to-beat variability of the heart rate, and it may contain outlier beats, due to noise interferences. The periodicity analysis is using a robust power spectral estimate, followed by Fisher's *g*-test [114], which computes the *p*-value of the null hypothesis that the time series is a Gaussian noise against the alternative hypothesis that the signal contains an added deterministic periodic component of unspecified frequency. Multiple test corrections for the *p*-value's cutoff were done using the false discovery rate (FDR) method [115]. The cluster center that provided the smallest *p*-value was selected as a template, and the identified period was compared to the average period of the breathing pressure signal.

5.1.2 Analysis results

During normal respiration, both S1 and S2 exhibited marked beat-to-beat variability, which nearly disappeared during apnea (Figure 38a). The heart sound variability was periodic and apparently synchronized with the respiratory cycle. The average morphological variability of S1 was 0.1±0.07 during respiration and 0.03±0.03 during apnea (Figure 38b). For S2, the average variability was 0.14±0.09 during respiration and 0.06±0.07 during apnea (Figure 38c). Both paired and unpaired t-tests showed that the variability of S1 and S2 during respiration was significantly higher than during apnea ($p<10^{-9}$ for all tests).

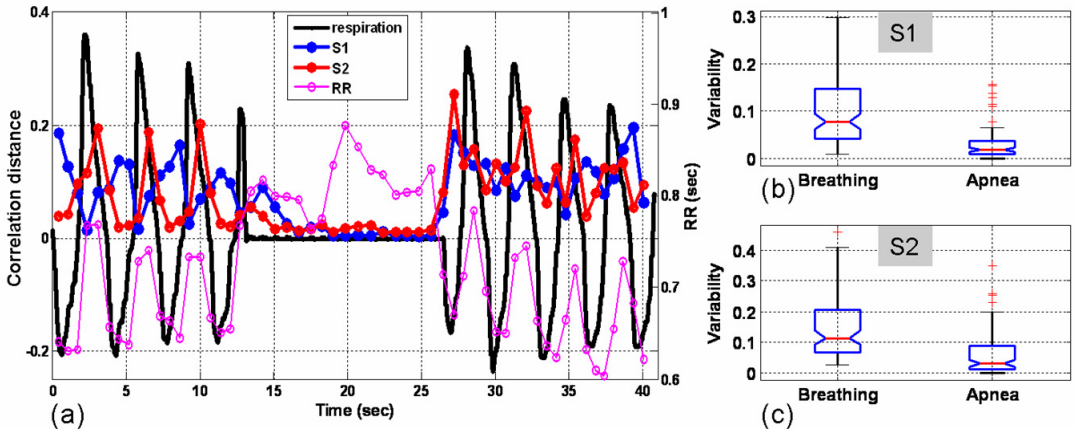

Figure 38: Respiration-induced variability. (a) Respiration pressure signal from a single recording (in arbitrary units) with beat variability of S1, S2 and heart rhythm (RR interval) during normal respiration against high resistance and during apnea. Heart sound variability is represented by the correlation distance between each beat and a fixed template. Both S1 and S2 exhibit periodic morphological changes during respiration that diminishes during apnea. The standard deviation of the correlation distance during respiration in all subjects is significantly higher than during apnea ($p<10^{-9}$) for both S1 (b) and S2 (c). The box plots display the median, lower and upper quartiles, data extent and outliers.

Cluster analysis identified distinct morphologies of S1 and S2 in all subjects. Although the heart sound morphology varied considerably between subjects, some general observations could be made about the intra-subject morphological changes. A typical instance of S1 clusters is shown in Figure 39a. The major component of S1, prominent in all clusters, is a large higher-frequency vibration, which reaches its energy peaks about 40ms after the R-wave of the ECG (90 ms from the beginning of S1 segment). While in the average of the unclustered signals the data that follows the main component is non-informative due to the high inter-beat variability, in some of the clusters (e.g. the inspiratory clusters 1,2,5,6) a peak of a secondary low-frequency component is clearly recognized 50-60 ms after the peak of the main component. This 'split' of S1 is absent from other significant clusters (e.g. clusters 3,5,8). A similar 'split' could be observed in the clustered time-frequency representation of S2, which is demonstrated in Figure 40. In this example, the clustering procedure identified a gradual emergence of a small low-frequency component, peaking 75 ms after the larger, high-frequency major component. This second component is blurred in the unclustered average of the S2 segments.

Statistical analysis of the periodicity of these apparent morphological changes was performed on 96 recordings from all 12 subjects, breathing against 4 levels of breathing resistance (2 recordings per resistance level per subject). Thresholds for significant *p*-values were determined by setting the false discovery rate (FDR) to 0.01. For S1 signals, a significant periodic component (corrected $p<0.007$) was identified in 81 of the recordings (84%). For 8 subjects periodicity was identified in all recordings, while for all subjects periodicity was identified in at least two different recordings. The measured period of S1 morphological changes was in high correlation (R=0.96) and good agreement (mean difference 0.02±0.3 sec) with the average period of the respiration cycle, measured from the breathing pressure signal (Figure 41). For S2 signals, significant periodicity (corrected $p<0.006$) was identified in 63 of the recordings (66%). All subjects had at least two recordings with periodic S2 morphology, with a strong correlation (R=0.87) and good agreement (mean difference 0.08±0.5 sec) between the measured period and the actual respiratory period.

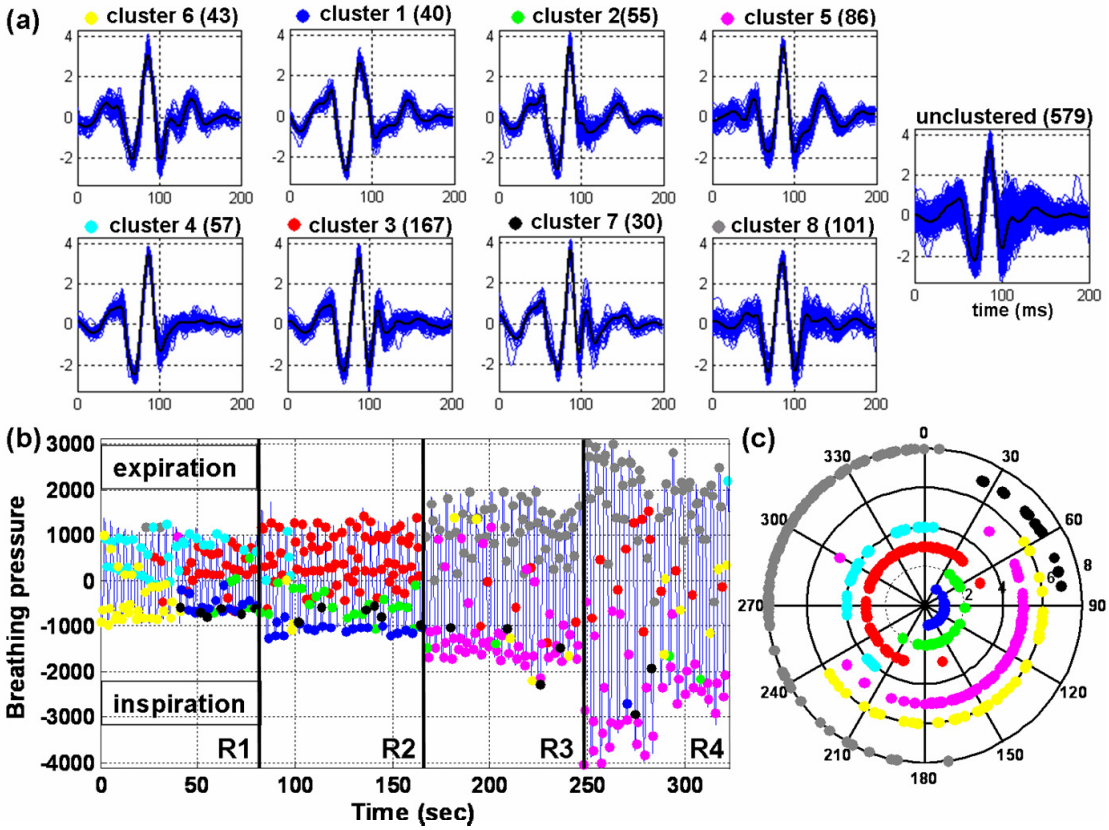

Figure 39: Clustering results of 579 beats of S1 acquired from a single subject (NM2) while breathing against variable resistance levels R1 to R4 (a). For each of the 8 clusters, the number of beats in the cluster is indicated, and the beats are plotted with the cluster's average. The morphological variability of the clustered signals is significantly lower than the variability of the unclustered data, in which subtle changes of the morphology are smeared. The relation between the morphological clusters and the respiratory activity is revealed by plotting the color-coded temporal location of the clustered beats along with the breathing pressure (b) and by a polar display of the phase in the respiratory cycle associated with each beat. A marked separation exists between inspiratory and expiratory clusters and between low and high breathing resistance levels. Note the secondary peak of energy at about 140 ms in the inspiratory clusters' morphology (yellow, blue, green and magenta clusters) that is missing in the expiratory clusters (cyan, red, gray).

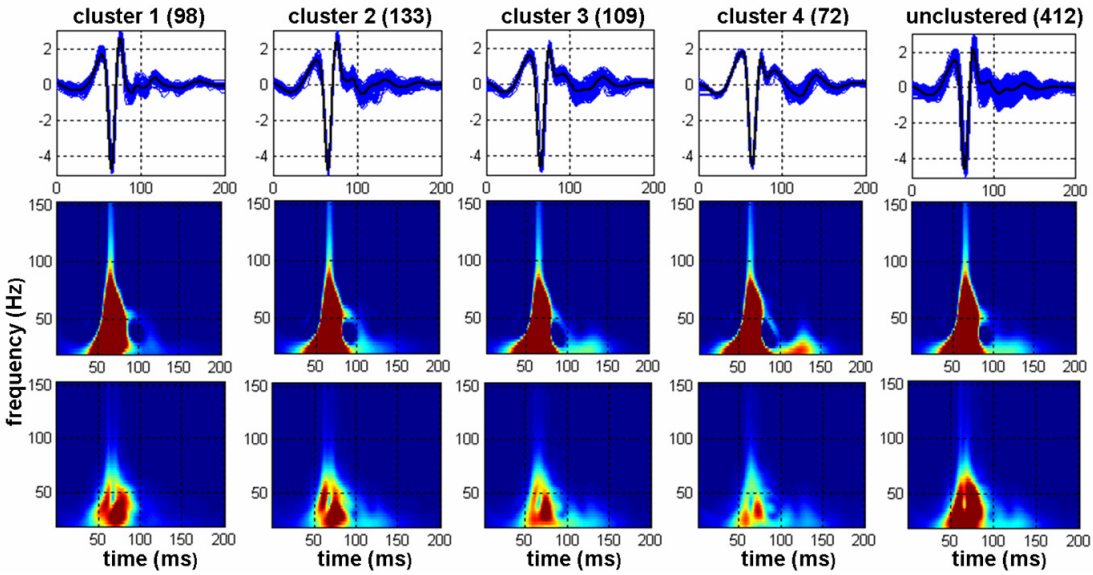

Figure 40: Clustering results of 412 beats of S2 from a single subject (ST1). For each of the 4 significant clusters, as well as for the unclustered data, the number of beats in the cluster is indicated, and the beats are plotted with the cluster's average (top row). The centers of the clusters, viewed by a time-frequency representation (middle row), emphasizes the emergence of a low-frequency late component in clusters 3 and 4. The standard deviation of the time-frequency representations (bottom row) demonstrates the larger morphological variability of the unclustered data.

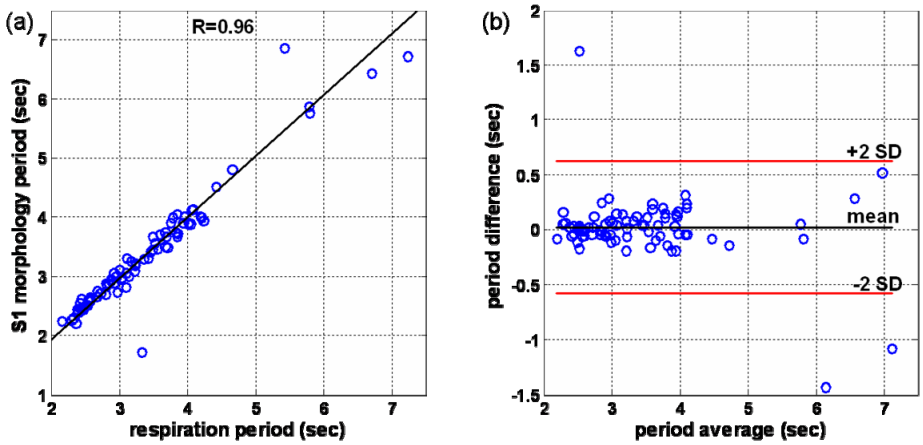

Figure 41: A linear regression plot (a) and a Bland-Altman plot (b) showing the strong correlation and the good statistical agreement between the period of the morphological changes of S1 and the actual respiration period.

5.2 Modulation of heart sounds by the respiratory phase

5.2.1 Data Analysis

To test whether there is a morphological separation between beats that occur during different phases of the respiratory cycle, each respiration cycle was mapped into the polar phase range 0-360º, where 90º is the peak of inspiration (maximal negative pressure) and

270º is the peak of expiration (maximal positive pressure). Each beat of S1 and S2 was associated with the corresponding value of the instantaneous respiratory phase (0-360º), and with the distance from the chosen cluster center. A two-tailed student's t-test was used to compare the distance values distribution of the beats occurring during inspiration (respiration phase value in the range 45º-135º) and the beats occurring during expiration (respiration phase value in the range 225º-315º). Significant *p*-value cutoff was determined by FDR method.

The ability of the computational analysis framework to predict the respiratory phase from the morphology of S1 or S2 was evaluated separately for each subject. A KNN classifier (K=5) was trained on half of the beats, and its performance was tested on the rest of the beats, by evaluating the accuracy of classifying beats into the correct half of the respiratory cycle.

5.2.2 Analysis results

The morphological difference between inspiratory and expiratory beats of S1, measured by comparing the distributions of the distances from a template beat, was found to be statistically significant (corrected $p<0.008$) in 83 of the recordings (86%), indicating that at least some of the variability in the signal's morphology is related to the respiratory phase. To visualize the effects of the respiratory phase on the heart sounds, S1 and S2 beats were sorted by their time of occurrence in the respiratory cycle (0-360º), and plotted as two-dimensional color-coded maps (Figure 42). The following observations were made regarding the variability of the heart sounds during the respiratory cycle:

Energy content of S1: In 11 of 12 subjects there was a statistically significant difference ($p<0.001$) between the energy content of S1 beats occurring in proximity to peak inspiration (phase range 45º-135º) and beats occurring in proximity to peak expiration (phase range 225º-315º). In 9 of these subjects S1 was attenuated during inspiration (phase 0-180º) and accentuated during expiration (phase 180º-360º). In the remaining two subjects the opposite relation was observed.

Timing of S1: S1 was slightly delayed during inspiration in all 12 subjects. The temporal delay from the R-wave of the ECG to the peak energy point of S1 was 4 to 20 ms longer

in inspiratory beats, compared to expiratory beats (mean 12±6 ms). This difference was statistically significant ($p<10^{-6}$) in 10 of the subjects.

Split of S1: In 6 subjects, a low-frequency second component was clearly identified in S1 signals occurring during inspiration or early expiration. The peak of this component was typically 50-60 ms after the peak of the major, high-frequency component.

Energy content of S2: In all of the subjects the energy content of S2 was significantly higher ($p<0.001$) during late inspiration and early expiration (phase range 135º-225º), compared to late expiration and early inspiration (phase range 315º-45º).

Timing of S2: S2 occurred earlier during late inspiration and early expiration in 11 of the subjects. The peak energy of S2 during this respiratory phase occurred 6 to 28 ms earlier, compared to late expiration and early inspiration beats ($p<0.001$).

Split of S2: The changes in the timing of S2 during late inspiration and early expiration were often due to the earlier occurrence of the first, aortic component of S2, while the second, pulmonary component did not change or was slightly delayed, producing noticeable split-S2 morphology in 9 of the subjects.

The ability of the cluster analysis framework to automatically identify the relations between the morphology of the first heart sounds and the respiratory phase is demonstrated in Figure 39b. There is a marked separation between clusters on the breathing pressure axis: some clusters (e.g.1,5,6) contain beats that occur in proximity to the peak of inspiration (maximal negative pressure), while other clusters (e.g. 3,4,8) are dominated by beats that occur during expiration (positive pressure). This separation is even more apparent in Figure 39c, showing the distribution of each cluster along the phase of the respiratory cycle. Beats of either S1 or S2 that are associated with inspiration are characterized by the 'split' morphology, wherein a second low-frequency component follows the major higher-frequency component, as described in the previous section. The clusters without this low-frequency component are typically associated with the expiratory or transition phases of the respiration cycle. The accuracy of the respiratory-phase classification from the heart sound morphology of all subjects is given in Table 5. The cluster-distance representation of S1 morphologies provided a good separation between beats associated with different halves of the respiratory cycle. Best accuracy was

achieved for partitioning the respiratory cycle at the points of phase 30º and 210º, allowing small error tolerance during transitions between inspiration and expiration. The accuracy of the phase classification rate varied between patients from 79% to 97% (average 87±7%). Phase classification using the morphology of S2 was much less accurate than S1, with an average correct classification of 69±8%, indicating that the morphological changes in S2 during respiration are less predictable than the changes in S1.

Table 5: Cluster analysis and classification of the respiration status from S1 and S2. Results include the number of significant clusters, correct classification (*CC*) rate of the respiratory phase and correct classification rate with maximal one level error (CC_1) of respiratory resistance from the morphology of S1 and S2.

#	ID	#beats	S1			S2		
			#Clusters	Phase-CC%	Resist-CC_1%	#Clusters	Phase-CC%	Resist-CC_1%
1	GA1	528	5	92	74	7	64	57
2	ND1	652	4	96	81	6	71	76
3	NM1	534	8	83	85	6	65	88
4	NG1	479	5	94	81	8	79	61
5	NM2	579	5	93	87	6	72	77
6	ND2	544	4	75	65	5	57	75
7	NM3	562	7	97	74	6	65	63
8	OG1	631	4	84	85	6	79	79
9	ST1	442	7	89	84	7	82	90
10	ZM1	557	4	83	83	6	60	63
11	RS1	455	8	81	90	7	59	79
12	SS1	410	5	79	89	8	72	71
Average		531.1	5.5	87.1	81.6	6.5	68.8	73.3
Stdev		73.7	1.6	7.2	7.3	0.9	8.4	10.6

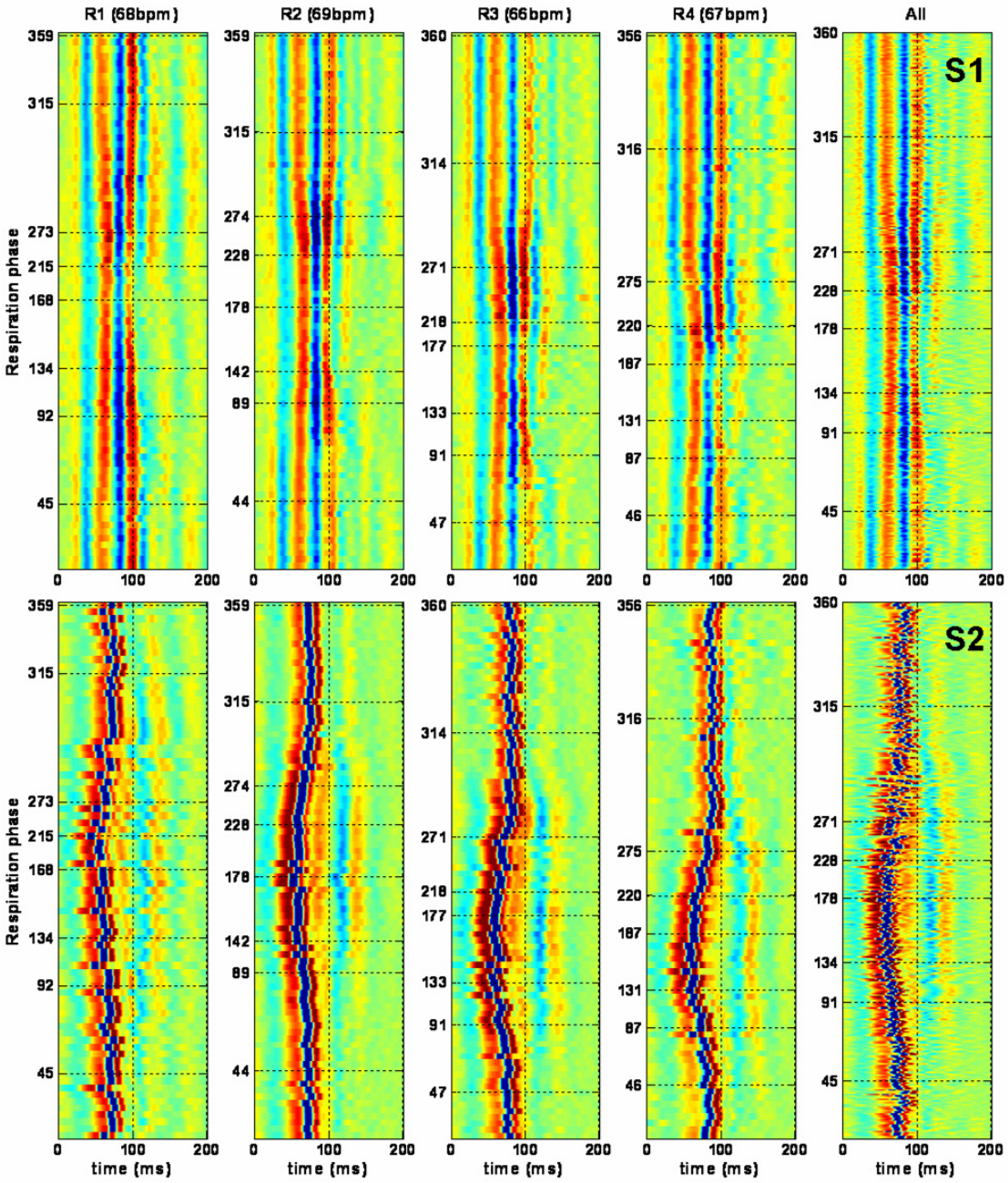

Figure 42: Morphological and temporal changes of S1 (top) and S2 (bottom), induced by the respiratory phase and load (subject ST1). S1 and S2 beats of each separate resistance level (R1-R4) and of the entire recording set (All) were sorted by the phase of their occurrence in the respiratory cycle (0-360º with inspiration occurring around 90º and expiration around 270º) and plotted as color-coded maps (red indicating positive deflection of the signal). Note that the respiration phase axis in each plot is slightly different, due to the arbitrary occurrence times of heart beats during respiration. S1 is delayed and attenuated during late inspiration. S2 occurs earlier and exhibits split morphology during late inspiration and early expiration. As the breathing load (resistance) is higher, these changes become more prominent and occur earlier in the respiration cycle.

5.3 Modulation of heart sounds by the respiratory resistive load

5.3.1 Data Analysis

The relations between the morphological changes of the heart sounds, and the variations in the breathing resistance, were examined by evaluating the performance of a classifier in predicting the breathing resistance and the instantaneous respiratory pressure from the signal's morphology. For this classification task, each beat was labeled by the level of breathing resistance used while it was acquired (R0 – R4). For each subject, A KNN classifier (K=5) was trained on half of the beats and the accuracy of resistance classification was evaluated on the other half by computing CC_1. In addition to measuring the correct classification rate per beat, the ability to correctly classify the resistance level of the entire recording, based on the classification of the majority of beats, was evaluated. Pressure estimation was also done using KNN. The estimated pressure was the weighted average of the nearest neighbors. The mean pressure estimation error, relative to the peak-to-peak amplitude of pressure variation was computed by:

$$EE = \frac{1}{\left|B^{test}\right|} \sum_{b_i \in B^{test}} \frac{\left|\tilde{p}(i) - p(i)\right|}{A_{l_i}} \tag{5.1}$$

where $p(i)$ and $\tilde{p}(i)$ are the reference and the estimated instantaneous pressure values of beat i, and A_{l_i} is the peak-to-peak variation of pressure amplitude when breathing against resistance $l_i \in \{R_1, R_2, R_3, R_4\}$. The linear correlation between $\tilde{p}(i)$ and $p(i)$ was also calculated, as well as the correlation between the peak-to-peak breathing amplitude A_{l_i} and its estimation $\tilde{A}_r$, defined by:

$$\tilde{A}_r = \max\left\{\tilde{p}(i)\middle|l_i = r\right\} - \min\left\{\tilde{p}(i)\middle|l_i = r\right\} \quad , r \in \{R_1, R_2, R_3, R_4\} \tag{5.2}$$

The reference value of the instantaneous breathing pressure associated with beat i was defined as $v_i = P(t_i\text{-}\tau)$, where t_i is the reference time point of beat i (usually, the beginning of the cycle), and τ is a constant delay parameter ($0 \le \tau \le 800$ms). The delay parameter τ was chosen to provide the maximal separation between the pressure values of the significant clusters, in terms of Fisher's separation criterion [97], defined by:

$$FC = \sum_{j=1}^{N} p_j (m_j - \bar{m})^2 \Big/ \sum_{j=1}^{N} p_j S_j \tag{5.3}$$

where $m_j = \frac{1}{|C_j|} \sum_{i \in C_j} v_i$, $S_j = \frac{1}{|C_j| - 1} \sum_{i \in C_j} (v_i - m_j)^2$, $p_j = \frac{|C_j|}{\sum_k |C_k|}$, $\bar{m} = \frac{1}{\sum_k |C_k|} \sum_{i \in \{C_1, \ldots, C_{\hat{N}}\}} v_i$

5.3.2 Analysis results

In addition to the cyclic morphological changes induced to the heart sounds by the respiratory phase, there are also changes induced by the extent of the respiratory resistive load (Figure 42). The changes in the temporal location of S1 and S2 are more prominent when the breathing load is higher. The delay of S1 during inspiration becomes longer and the delay of S2 during late inspiration/early expiration becomes shorter in high breathing resistances, compared to low breathing resistances. In some of the subjects, the magnitude of the changes in the energy and morphology of the heart sounds was also related to the level of breathing resistance. Furthermore, the resistance level affected the occurrence time of the aforementioned changes in the respiratory cycle: as the breathing load was higher, the respiration-induced changes of the heart sounds occurred earlier in inspiration. This phenomenon was observed in 10 of the subjects for temporal, morphological or energy-related changes of S1 and S2.

Cluster analysis was able to recognize resistance-induced changes, as shown in Figure 39b: while breathing against high resistance levels (R3 and R4), distinct clusters of S1 were identified for both inspiratory and expiratory phases. A representative separation between beats of different resistance levels is visualized in Figure 43. The ability of the clustering and classification framework to correctly identify the breathing resistance from the beat's morphology was quantified by the classification results, given in Table 5. Using time-domain signal representation, the accuracy of resistance classification with a maximal one-level error (CC_1) varied from 65% to 90% (mean 82±7%) using S1, and from 57% to 90% (mean 73±11%) using S2. With S1-based classification, 51% of the beats in the entire test set were classified to their exact resistance level (CC_0), and 93% were classified with a maximal two-level error (CC_2), indicating that there is a good separation between low-resistance and high-resistance beats. Since for practical applications, resistance classification may be needed for a series of beats rather than for a single beat, the classification performance per test recording was also evaluated. Automatic classification of the resistance level of the entire recording, using the majority classification of the recording's beats, was exact (CC_0) in 45 of the 60 test recordings (75%), and correct with maximal one-level error (CC_1) in 55 of the recordings (92%). Performance comparison between different parameters of the analysis framework

pointed-out time-domain and S-transform representations with correlation distance and KNN classifier as preferred configurations, with an average correct classification rate of 82±7% (Table 6). The performance differences between the methods were not significant: correlation distance was somewhat better than Euclidean distance, and the KNN classifier was slightly better than DA. All representation methods achieved low estimation errors, with the best result of 19±6% achieved by WVD. A high correlation was obtained between the breathing pressure estimated from the morphology of S1 and the pressure value associated with each beat. The correlation coefficient was 0.76 for the 2057 test beats of all 12 subjects (Figure 44A). In addition to instantaneous pressure estimation, the peak-to-peak amplitude of the estimated pressure in each breathing resistance level was strongly correlated with the actual pressure variation, or the actual breathing effort (R=0.92, Figure 44B). To ascertain that these relations are indeed a consequence of the morphological differences between beats, correctly derived by the analysis framework, the selection of the *K* nearest neighbors in the cluster-distance space was replaced by a random selection of *K* training beats that were used for classification and pressure estimation. Using this random classification, the results were significantly worse ($p<10^{-5}$) with average CC_1 of 52%, and average *EE* of 35%. There was no correlation whatsoever between the randomly-estimated and the actual pressure per beat.

Table 6: Classification performance on S1 signals from HSPRS dataset. Mean and standard deviation of correct classification (CC_1) and relative estimation error (EE) of all subjects, using different configurations of signal representation, distance metric and classification algorithm (KNN=K-Nearest Neighbor, DA=Discriminant Analysis). Best results, obtained by ST, WVD and time-domain representations, are indicated by boldface.

Signal representation	Distance metric	KNN CC_1(%)	DA CC_1(%)	*EE*(%)
Time	**Correlation**	**82±7**	76±11	20±7
	Euclidean	82±7	80±9	21±6
Frequency	Correlation	77±6	72±9	20±4
	Euclidean	75±7	74±5	23±5
Short-Time Fourier Transform (STFT)	Correlation	78±10	73±10	21±5
	Euclidean	78±8	76±8	23±6
S-Transform (ST)	**Correlation**	**82±7**	76±10	20±7
	Euclidean	78±8	80±9	22±5
Wigner-Ville Distribution (WVD)	**Correlation**	81±7	77±8	**19±6**
	Euclidean	80±7	79±8	20±6
Choi-Williams Distibution (CWD)	Correlation	78±8	73±9	20±6
	Euclidean	78±8	76±9	22±7

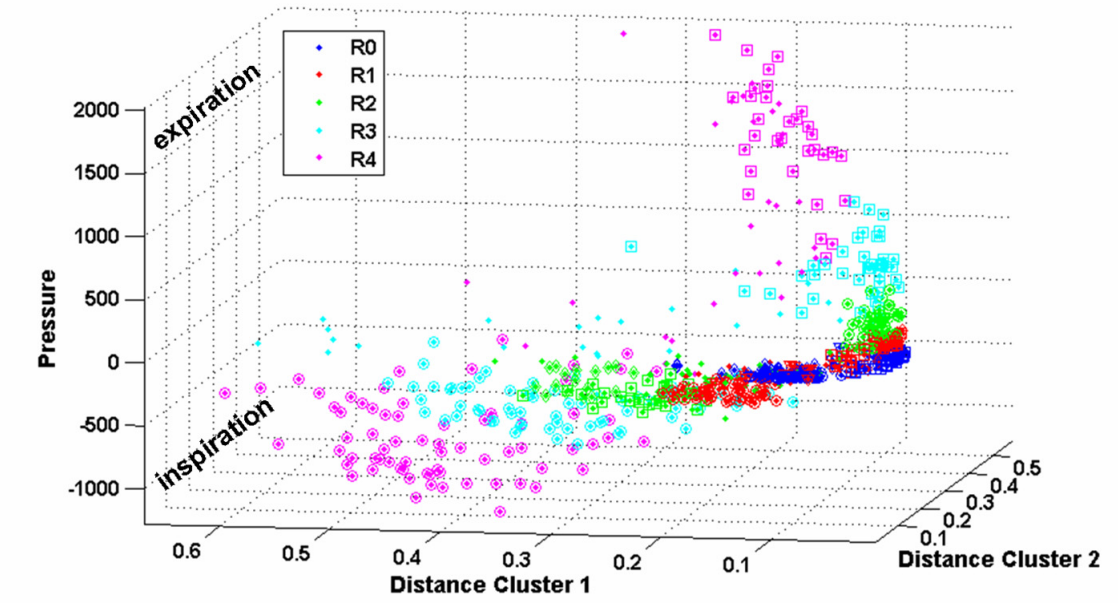

Figure 43: Cluster-distance representation of 652 beats of S1 from a single subject (ND1). Each beat is plotted by its distances from the centers of the two largest clusters (x-axis and y-axis) and by its associated breathing pressure (z-axis). The marker colors show the breathing resistance levels and the marker symbols designate the significant clusters of each level. There is a marked separation between beats of different resistance levels, and within each level, between beats associated with different respiratory phases

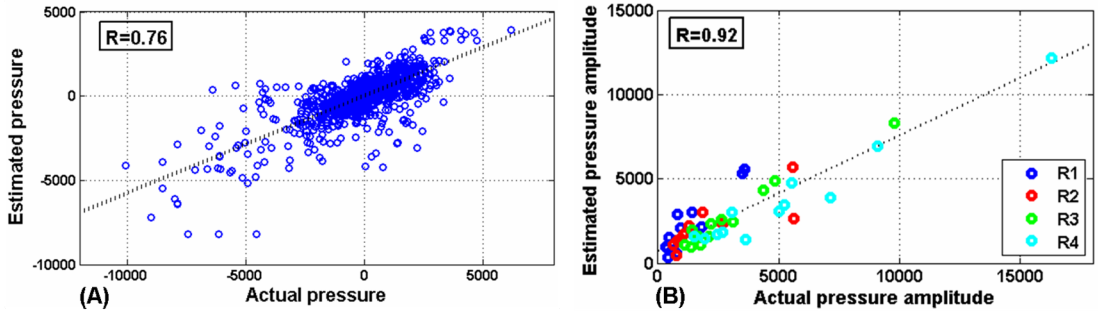

Figure 44: (A) Estimated breathing pressure of 2057 test beats of S1 from all 12 subjects, plotted against the actual breathing pressure associated with the beat. The correlation coefficient is 0.76. The absolute pressure differences were normalized by the peak-to-peak amplitude of pressure variation to obtain the reported average estimation error. (B) Estimated peak-to-peak amplitude of breathing pressure of 12 subjects (4 resistance levels per subject, indicated by marker colors), plotted against the measured amplitude of breathing pressure. The correlation coefficient is 0.92. Pressure values are specified in arbitrary non-calibrated transducer units.

CHAPTER 6
DISCUSSION

The relations between the physiological processes producing the heart's sounds and infra-sound vibrations and the morphology of the externally acquired signals are highly complex. The mechanical interplay between myocardial contraction, blood flow and valve activity is continuously regulated by the autonomous nervous system, and is affected by hormonal and pulmonary activities. The filtering effects of the thoracic cavity and the skin conducting the vibrations considerably alter the morphology of the signal [116]. Nevertheless, vibro-acoustic heart signals, being a direct manifestation of the mechanical cardiac cycle, bear valuable information about the functioning of the cardiovascular system. The methods and results described in this work for extracting this information may provide new means of continuous monitoring and assessment of cardiovascular mechanical function. The signal analysis techniques and their utilization on heart sound datasets provide computational insights about the choice of appropriate methods and parameters for processing vibro-acoustic heart signals and extracting their information. In addition, the results provide physiological insights about the nature of the morphological variability of heart sounds and its relation to changes of ventricular function and respiration activity.

6.1 Computational analysis techniques

6.1.1 Signal representation

Successful predictions of the physiological condition from the morphology of the signal require suiting of the analysis techniques to the properties of the analyzed signals. Signal representation is one of the fundamental choices within the proposed analysis framework. When considering the problem of accurate decomposition of the signal into its subcomponents, there are considerable differences between methods. Simple STFT is limited by its fixed resolution, which imposes a tradeoff between temporal and spectral resolutions. One way to avoid the resolution tradeoff is by using linear transforms with frequency-dependent resolution, such as the wavelet transform and S-transform. Alternatively, quadratic transforms, such as WVD and its reduced-interference

derivatives like CWD, can be used. There is no consensus in the literature regarding the most suitable time-frequency representation of S1 and S2. Different studies point out different techniques, such as the binomial transform [51], cone-kernel distribution [52] and continuous wavelet transform [53] as preferred choices. In our analysis framework, the preservation of the relative morphological similarity between signals, under a certain representation method, is of greater importance than the absolute accuracy of the signal's decomposition. The optimal signal representation and distance metric should have the right balance between sensitivity and robustness. Sensitivity is important for detecting minute differences between beats, and robustness is essential to reject noise-related differences. The classification results obtained by simple time-domain representation were comparable in most cases with the results obtained by time-frequency representation (TFR). Both types of representations hold the same amount of information about the signal. In TFR, this information is represented in two dimensions, with the cost of either sub-optimal time/frequency resolution, or interference of artifactual cross-terms, resulting from the non-linearity of the energy distribution transforms. The simulation results (Figure 14) demonstrated the differences in the sensitivity of various representations to changes in time, frequency and noise level, and emphasized the tight relation between the preferred representation and nature of the data: In cases where the data exhibits large variability between classes and small variability within each class, highly-sensitive representations would provide more accurate results, whereas when the changes in the data are more gradual and there is small between-class variability or large within-class variability, a representation that is less sensitive but more robust should be superior. On both data sets of real heart sounds S-transform representation provided better average classification performance than other time-frequency representations. Its advantage was evident in the HSDSE dataset, in which the changes between classes are mostly spectral. Surprisingly, time-domain representation achieved equivalently good classification results. These results can be explained by the fact that the beats were aligned in the time-domain prior to their clustering, and this initial alignment helped to mitigate most of the temporal mismatch between them. Once the starting points of all beats were correctly aligned, intra-beat temporal variations were relatively minute, enabling time-domain clustering to perform well.

6.1.2 Feature extraction

Regardless of the chosen representation method, in order to classify heart sound signals they must be transformed to a lower-dimension feature space. Typical S1 or S2 signals are represented in the time-domain by a couple of hundred samples (200ms segments sampled at ~1000Hz). Time-frequency representations add another order of magnitude (30-40 frequency bins per time point), resulting in a very high dimensionality of the raw data, which is inappropriate for classification. Previous studies on heart sound classification used domain-specific features, such as dominant frequencies, spectral bandwidth and signal intensities [56]. More general feature extraction techniques used either model estimation [57] or search-based feature selection [58]. While the domain-specific features have physical meaning and can therefore be easily interpreted, they need to be specifically determined for every type of signal and for every data set. Automatic feature extraction and selection methods provide a more systematic solution, but their loose relation with the underlying physiological processes make the classification results less traceable. We have examined two approaches for dimensionality reduction and feature extraction of heart sounds. The proposed methods for extracting acoustic features of cardiac function, described in section 4.2.3, include domain-specific features such as the bandwidth of S1, as well as a unique utilization of PCA on the time-frequency representation of the signal. PCA, which is commonly used for dimensionality reduction, enables to characterize and localize the variability of the data. The data projection on a new orthogonal basis of the principal components accentuates physiologically meaningful patterns, while diminishing noise-related components. The derived one-dimensional index is therefore more robust to noise and artifacts than straight-forward features of spectral energy.

An alterative approach is the unsupervised representation of the data in the feature space of cluster distances, introduced in section 3.2. The centers of the significant clusters constitute a concise description of the most prominent signal morphologies, and the vector of distances of each beat from these centers is extremely informative, yet computationally sensible. For the studied heart sound datasets, typical distance-space representation had 7-10 dimensions, enabling efficient feature extraction and classification. The key advantage of this clustering-based method is that it is entirely

unsupervised, not requiring any prior knowledge on the processed signals. Cluster belonging alone, as a one-dimensional feature, is too coarse to reliably classify a data element. As there is also considerable within-cluster morphological variability, and since boundaries between similar clusters might be arbitrary, multi-dimensional characterization of each beat relatively to the different major morphologies in the data is more powerful. Another benefit of the method is its inherent robustness to irregular noise interferences. The clustering process distinguishes between repetitious morphologies, which are assigned to significant clusters, and coincidental patterns, which usually reflect artifacts and are assigned to insignificant clusters. The cluster-distance feature space is constructed using only significant clusters, and enables to easily identify noisy beats, which are considerably distant from all major morphologies. Cluster-distance representation is a general approach, applicable to different types of periodic signals, and at the same time it has a simple physical interpretation of morphological signal similarity, which can be partially visualized using three-dimensional plots (Figure 37, Figure 43).

6.1.3 Signal matching

Measuring the similarity or distance between pairs or clusters of signals is another building block of the proposed analysis framework. The simple Euclidean and correlation distances used for heart sound analysis provided clear visual separation between distinct signal morphologies, as well as good overall classification results. On both datasets, average classification results were slightly better when correlation distance was used. This result is primarily due to the sensitivity of Euclidean distance to scale differences between the compared signals, while the correlation distance provide more accurate comparison of the morphologies, regardless of the absolute scales. Many time-series comparison techniques have been proposed in the pattern recognition and data mining literature [117]. These include methods for signal normalization and transformation, longest common subsequence (LCS) measures and piecewise-linear representation [118]. Dynamic time warping has also been studied, with different variants developed to improve its computational efficiency [119] and its ability to preserve the ‘natural’ morphology of the signals [120]. The landmark model [121] aimed to be consistent with the human perception of similarity, by selecting important points (landmarks) and smoothing the compared signals between them. Our approach of event-based

segmentation (section 3.3.4) extends this idea by identifying the extrema points in a multi-scale manner and using their maximal scale and multi-cycle variability as measures of their importance. Prior identification of the significant events or components in the signal, i.e. signal segmentation, enables to accurately align multi-component signals and to measure their similarity component-by-component. The technique of combined time-frequency segmentation and component-wise warping, demonstrated in section 3.3.5, provides a signal similarity measure that is suitable for heart sounds, especially when there is large variability in the inner component structure of the compared signals. The potential contribution of this method to the further performance improvement of the clustering and classification framework is left for future work.

6.1.4 Classification

The classifiers used in this work, K-nearest neighbor, and linear discriminant analysis represent two common types of data separation: linear separation (DA) and radial separation (KNN). On the HSPRS dataset, KNN classification provides better results than DA. Opposite results were obtained for the HSDSE dataset, where DA was consistently better. This apparent discrepancy can be explained by the different structures of the cluster spaces created for the two datasets. In the HSDSE data, the Dobutamine-induced stress causes substantial spectral changes of S1 that are reflected as extremely distinct clusters, separable by a linear projection. The respiratory-induced variations in this dataset are masked by the dominant stress-induced changes, and have a negligible influence on the clustering results. On the other hand, the morphological changes of S1 in the HSPRS data are more subtle. Most of the variation is related to the respiratory phase, while the variation caused by changes in the breathing resistance is smaller. Beats of different breathing resistances are therefore intermixed in the cluster-distance space, and non-linear separation of KNN achieves better results than linear DA. In recent years, support vector machines (SVM) have been established as an advanced classification method, which perform better than classical methods in a wide range of applications [122]. We have also applied SVM on the datasets of heart sounds, obtaining equivalent performance to KNN and DA. We therefore concluded that for the task of multi-class separation of the analyzed heart sound data, SVMs may offer only a modest performance improvement, which does not justify its large computational overhead. The utilization of

multiple heart sound channels, acquired simultaneously from different locations, has an important contribution to the high classification accuracy. On the HSDSE dataset, each separate channel provided a lower classification rate of about 80%, while the fusion of the four classifiers made the analysis more robust and more accurate with 86% correct classification. This approach can be extended to combining classifiers that use different signal representations, distance metrics or classification algorithms to further improve the accuracy.

6.2 Assessment of cardiac function

The effects of hemodynamic changes on heart sounds have been studied for many years. The amplitude of S1 has been shown to be directly related to left-ventricular contractility by Sakamoto et al., who reported a nearly linear relationship between the amplitude of S1 and the maximum of the time derivative of the left ventricular systolic pressure (dP/dt) in dogs [24]. Their experiments included different variations of cardiac dynamics by drug administration, obstruction of great vessels and induced myocardial infarction. More recently, Chen et al. studied the properties of intracardiac and thoracic S1 in dogs in normal, increased and decreased contractile states [63]. They analyzed S1 using cone-kernel time-frequency distribution and identified the dominant components of the signal in a band around 50Hz. They also reported a good cross-correlation between the instantaneous frequency of S1 and dP/dt, thus connecting the resonant frequency of S1 to the left ventricular pressure gradient. Clinical studies in humans showed that myocardial infarction caused a shift in the maximum energy of S1 to a lower frequency range [61], and that a reduction in the spectral energy of S1 correlated well with the presence of significant coronary artery disease [62]. Our analysis of the changes in S1 during Dobutamine-induced stress response, using principal component analysis of the time-frequency distributions of S1, provided results that are consistent with these previous studies. Our analysis pointed out a significant increase in the energy of the main low-frequency component, concentrated in the band of 20-70Hz, along with the emergence of high-frequency components in the band of 110-150Hz. These general patterns were consistently observed in multiple recording locations in all of the subjects, despite the significant inter-subject differences in the signal's morphological and spectral

characteristics. Normalizing the extracted acoustic features by their baseline value enabled to continuously track the relative changes of S1 in each patient. The correlation of these features with the strain-echocardiography indices supports the hypothesis that these features of S1 reflect changes in the left ventricular systolic function.

Dobutamine affects cardiac function by stimulation of β_1-adernergic receptors in the heart, causing a significant increase in myocardial contractility and a mild increase in heart rate (HR). As a result, both stroke volume (SV) and cardiac output (CO=SV*HR) are augmented. With low concentrations of Dobutamine, the dominant cause of the increased CO is the improved ventricular contractility, whereas with high concentrations of Dobutamine there is no further change in SV, and the dominant factor is increased HR [123]. For this reason, the correlation between the acoustic indices extracted from S1 and the strain-echocardiography indices of contractility was examined at the end of the low-dose Dobutamine administration (Figure 35). Although in most of the patients the pattern of heart rate changes resembled the trend line of the acoustic indices, there was a single patient (#11, Figure 34d) with very small changes in heart rate but significant changes in both echocardiographic and acoustic contractility indices, showing that the acoustic changes are independent of the heart rate. The characteristics of heart sound indices during pathologic reduction of cardiac function could not be addressed quantitatively in this work, due to the small number of subjects and the fact that the great majority of the subjects had normal cardiac function. However, the single subject that was diagnosed in the echocardiography examination with a reduced segmental wall motion during stress, due to myocardial ischemia (subject #10) had the lowest values of absolute and relative ASI, as well as the lowest values of PSSR, suggesting that the compromised wall motion might result in a frequency reduction of the first heart sound. This clinical correlation should clearly be further addressed in future work.

The clustering and classification analysis framework, applied to the HSDSE dataset in section 4.3, offers a more general approach to assessing the cardiac function based on the morphology of heart sounds, without explicit characterization of the monitored features. Clustering of the time-frequency representations of S1 provided unsupervised confirmation to the aforementioned patterns of spectral changes in S1 during stress response. The clusters that were automatically generated, without using any prior

knowledge about the underlying physiological condition (Figure 36) showed evident resemblance to the averaged representations of the corresponding test stage (Figure 30). This visual impression was quantitatively validated by the ability to predict the stress-level from the cluster-distance representation of S1 with a high average accuracy of 86%. This framework can therefore assess cardiac function by morphological classification of test beats into a set of physiological states, on which it was trained beforehand. Alternatively, it can identify any recurring abnormalities that reflect morphological deviations from the training set. It can also adapt to new data within normal morphology by continuously updating its internal cluster-based representation of signal morphologies.

6.3 Assessment of respiratory function

The cyclic respiratory activity modulates the mechanical function of the left and right heart through changes in the pleural pressure (Figure 45) and pulmonary blood flow. The lowered pleural pressure during inspiration causes enhanced venous return to the right atrium, and increased preload and stroke-volume of the right ventricle. The preload and stroke volume of the left ventricle are decreased due to ventricular interdependence and increased afterload [5, 124]. The left ventricle contracts with a decreased force, against a higher arterial resistance, and S1 is attenuated. The increased difference between aortic and left-ventricular pressure causes S2 to be accentuated. In addition, the aortic component of S2 occurs earlier, while the pulmonary component is delayed as the right-ventricle pressure is high. These temporal changes result in a wider split of S2.

The presented analysis of the respiratory-induced changes of heart sounds confirms and reinforces this physiological model using modern computational tools. The clustering of heart sounds, along with the compact representation of morphology in the feature space of cluster distances, enabled to quantitatively analyze the complex relationship between heart sounds and respiratory activity. Both S1 and S2 exhibited strong morphological variability during respiration, and nearly no variability during apnea. The morphological variability of heart sounds was found to be periodic, and the estimated period was in good agreement with the measured duration of the respiration cycle. This apparent relation between the respiration phase and the characteristics of heart sounds was confirmed by identifying statistically-significant differences in the template-distance, energy content,

and time of occurrence between beats of S1 and S2 acquired during different phases of the respiration cycle. The common dynamics in most subjects was attenuation of S1 during inspiration, accompanied by a small temporal delay, and accentuation of S2 during late inspiration and early expiration, with earlier occurrence of the aortic component, and wider split morphology. Intensity changes induced to the heart sounds by the respiration cycle has been described by Ishikawa et al., who compared heart beats occurring in proximity to peak inspiration and peak expiration [125]. They reported an increased intensity of both S1 and S2 during expiration. The results of the current study are consistent with these previous findings, and provide a more extensive and precise analysis of the relations between heart sounds and respiration, owing to the utilization of computerized signal analysis. For some of the studied subjects, only few of these respiration-induced changes were observed, and there was a large inter-subject variability in the exact characteristics of the heart sound changes. However, as the group of subjects was relatively small, it was impractical to analyze the differences between subjects. We did not identify specific clinical characteristics that could explain this variability in a post-hoc evaluation. The hemodynamic changes induced by inspiration to the atrial and arterial pressures are exaggerated during loaded inspiration [124, 126]. We have used a simple experimental model of variable breathing resistances to obtain higher fluctuations of pleural pressure. In most of the studied subjects, respiration-induced changes in the timing and morphology of S1 and S2 were indeed more prominent in high-resistance respiration. As the amplitude of heart sound was measured in uncalibrated units, and was sensitive to slight movements of the transducers or the subject, comparison of absolute energy content in different recordings was unreliable.

A new physiological insight from our current analysis is the relation between the breathing resistance and the relative temporal occurrence of the morphological changes in the respiratory cycle: with higher resistance, the heart sounds change earlier in inspiration. This observation is consistent with the hypothesis that the lowered pleural pressure induces the sound changes, since in high breathing effort the pleural pressure becomes low enough to affect the cardiovascular homodynamics earlier in the respiratory cycle. Automatic classification of respiratory phase and resistance level from the cluster-distance representation of S1 morphology achieved good average accuracy of 87±7% and

82±7%, respectively. The instantaneous breathing pressure of a single beat could be estimated with an average error of less than 20%. These results provide additional credence to the relation between the respiratory function and the heart sounds.

Although the statistical analysis showed that S2 is undergoing significant morphological changes during respiration, classification of the respiratory condition using S2 signals consistently achieved inferior performance compared to S1. A possible reason is that S2 has shorter duration and lower amplitude than S1, and its morphological changes are more subtle. In addition, the multi-cycle alignment of S2 signals is less accurate as it is performed without external ECG reference. Consequently, S2-based analysis is more sensitive to noise interferences and signal misalignment, and may require finer methods of signal alignment and distance measure to achieve more accurate clustering.

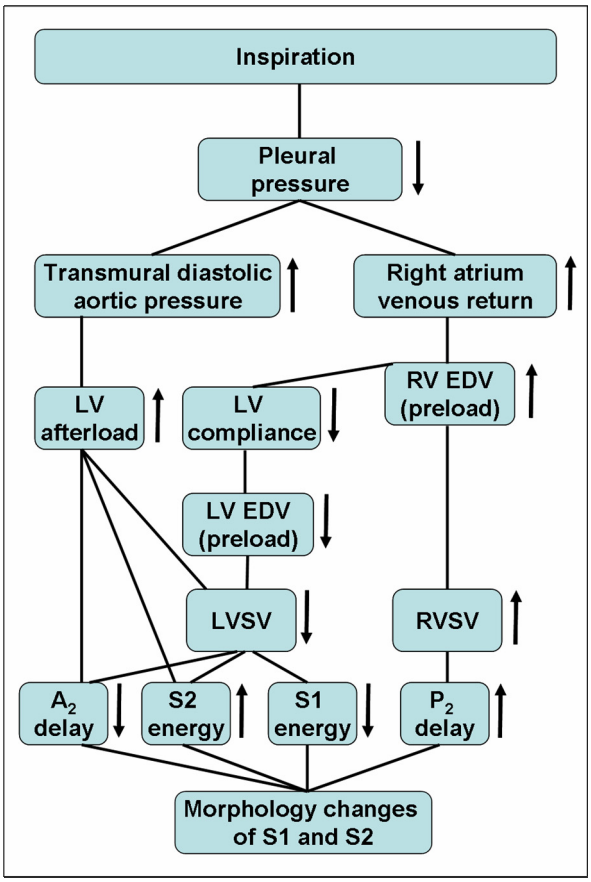

Figure 45: Physiological factors affecting the morphology of heart sounds. During inspiration there is an increase in the end-diastolic volume (EDV), or preload, of the right ventricle (RV) and a decrease in the preload of the left ventricle (LV). The latter causes a reduced contraction of the LV and attenuated S1. In addition, the increase in the LV afterload results in an earlier and accentuated aortic component of S2 (A2). The delay in the pulmonary component of S2 (P2) due to the larger and stronger RV stroke volume (SV) contributes to the split morphology of S2.

6.4 Applications for cardiopulmonary monitoring

Continuous monitoring of cardiovascular and cardiopulmonary functions may be beneficial in a wide range of clinical conditions such as critically-ill patients with hemodynamic instabilities, patient after cardiac surgery with possible postoperative complications, patients after myocardial infarction and heart failure patients with systolic or diastolic dysfunction. In all of these conditions, the mechanical function of the heart may be deteriorated by changes of ventricular contractility and compliance, or alterations of arterial, venous and intra-cardiac pressures, which can be reflected by changes in the vibro-acoustic signals. In the case of heart failure, the compensatory increase of blood volume can lead to accumulation of fluid in the lungs and complications of pulmonary congestion. The resulting increased respiratory load can then be reflected in the changes induced to the heart sounds. Other respiratory dysfunctions that may affect heart sounds include chronic obstructive pulmonary disease (COPD), asthma and mechanical ventilation. Continuous vibro-acoustic monitoring can be carried out by online extraction of specific signal features, such as time intervals, spectral energy and frequency bandwidth. In addition, a more general monitoring technique is offered by the clustering and classification framework, which can monitor morphological changes of heart sounds, and classify heart beats into physiological states. This technique requires prior training per subject, which may be done under controlled conditions such as a stress test, controlled changes of respiratory resistance, or controlled induction of medication. Alternatively, in case such multi-class training data is not available, baseline signals of a stable condition can be used for training, and the framework can be used to identify significant variations from baseline morphologies. Recurrent beats with morphological dissimilarity to the training beats can be alerted as a suspected abnormality. The monitor can also continuously adapt its training set to new signal morphologies that were confirmed by a human inspector. Using heart sound monitoring, gradual disease exacerbation may be identified before there are noticeable clinical symptoms or electrocardiographic abnormalities. Another potential use is assessment of the patient's response to pharmaceutic treatment and adjustment of medication doses according to monitored effects on the vibro-acoustic signals. Vibro-acoustic monitoring is a non-invasive method, which passively acquire the physiological signals. It is simple to use,

low-cost and portable, suitable for employment as a wearable home-monitoring device. Alternatively, it can also be performed in an out-patient setup, requiring only minimal equipment and technical support.

6.5 Limitations and future work

The described study of automatic heart signal analysis has some limitations. The number of subjects in each of the analyzed datasets was relatively small while the inter-subject variability of the signal's morphology was naturally large. As a result, the analysis was limited to relative changes in signal features per subject. Each subject should be separately trained with its own baseline signal, as it is not feasible to construct a generalized classifier for multiple subjects. We were also unable to characterize differences between sub-groups of subjects, and specifically between 'normal' and 'abnormal' subjects.

The analyzed vibro-acoustic signals, recorded on the surface of the chest wall, are prone to be distorted by different noise sources. Possible interferences include body movements, environmental noise and other physiological sounds originating from muscles, lungs or stomach. Some of these interferences produce irregular signal morphologies, which can be detected and excluded by the analysis framework. However, non-cardiac interferences that occur in a regular, periodic manner should be carefully identified. The filtering effects of the thorax, lungs and skin, for instance, cannot be easily distinguished from the cardiopulmonary-induced modulation of the signals. However, the fact that opposite effects were consistently observed for S1 and S2 during inspiration (S1 was attenuated and delayed while S2 was accentuated and occurred earlier) indicates that the contribution of the conducting medium is not a major determinant in the detected morphological changes of the signals. In order to isolate the effects of the conducting medium, intrathoracic or transesophageal heart sound signals should be acquired as well, which naturally requires a much more invasive research protocol.

Invasive measures of intra-cardiac and intrathoracic pressures can also improve the accuracy and reliability of the 'gold-standard' measures of cardiac function, to which the extracted vibro-acoustic information is compared. Although the strain-echocardiography

indices used for reference assessment of ventricular function have been shown to correlate with invasive measures of contractility, they merely remain an estimation. An additional drawback of these echocardiography indices is that they are given as discrete values along a few time points rather than a continuous beat-to-beat trend line. Similarly, the breathing pressure data, used to assess the relation between heart sounds and respiratory activity, is an indirect estimate of the intrathoracic pressure.

The utilization of ECG for cycle segmentation and temporal location of S1 was a simplifying choice, taken to ensure reliable, straightforward signal segmentation. However, the methods for heart sound segmentation without ECG, proposed in section 3.3.4, can be incorporated in the analysis framework in order to remove this limitation. Further improvement of the clustering and classification accuracy can be gained by using component-based signal alignment as a similarity measure suited for heart sound signals.

In summary, our work has laid the grounds for more extensive clinical studies by demonstrating some of the principle relations between heart sounds and cardiopulmonary physiology, and the computational methods for exploring these relations. Additional research with large groups of patients is now required in order to characterize and generalize the changes of vibro-acoustic heart signals in specific clinico-pathological conditions.

6.6 Conclusions

In this work, we introduced a variety of signal processing and pattern recognition techniques for automatic analysis of vibro-acoustic heart signals. In particular, we have studied methods of preprocessing, segmentation, alignment, time-frequency data representation, robust feature extraction, morphological clustering and classification of sound and infra-sound signals. We have acquired vibro-acoustic heart signals from human subjects during experimentally-controlled modulations of the physiological conditions, and used computational analysis methods to characterize the relations between heart sounds and cardiorespiratory function. We found that temporal, spectral and morphological features extracted from heart sounds can be used to estimate and monitor indices of left-ventricular systolic function. Using a computational framework of morphological clustering and classification of heart sounds we have shown that the

cardiac stress level, the respiratory phase, the respiratory resistive load and the instantaneous breathing pressure can be predicted from the signal's morphology with high accuracy and good noise robustness.

We demonstrated that automatic analysis of vibro-acoustic heart signals can be applied to continuous non-invasive monitoring of cardiac and respiratory functions, thus providing a promising technology for detection and diagnosis of mechanical dysfunctions caused by cardiovascular and cardiopulmonary diseases. The cost-effectiveness of the technology and its suitability for home-monitoring justify the revisit of these long-known signals with an armory of novel analysis algorithms, which can fulfill an actual clinical necessity.

REFERENCES

1. Berne, R.M. and M.N. Levy, *Cardiovascular physiology*. 6th ed. 1992: St. Louis : Mosby-Year Book.
2. Klabunde, R.E., *Cardiovascular Physiology Concepts*. 2004, Lippincott Williams & Wilkins.
3. Pinsky, M.R., *Cardiovascular Issues in Respiratory Care.* Chest, 2005. **128**: p. 592-597.
4. Bernardi, L., et al., *Modulatory effects of respiration.* Autonomic Neuroscience: Basic and Clinical, 2001. **90**: p. 47-56.
5. Bromberger-Barnea, B., *Mechanical effects of inspiration on heart functions : a review* Federation Proc., 1981. **40**: p. 2172-2177.
6. Van Leeuwen, P. and H.C. Kuemmell, *Respiratory modulation of cardiac time intervals.* Br Heart J, 1987. **58**(2): p. 129-35.
7. Knowles, G.K. and T.J. Clark, *Pulsus paradoxus as a valuable sign indicating severity of asthma.* Lancet, 1973. **2**(7842): p. 1356-9.
8. Nagendran, T., *The Syndrome of Heart Failure.* Hospital Physician, 2001. **37**(4): p. 46-57.
9. Davie, A.P., et al., *Value of the electrocardiogram in identifying heart failure due to left ventricular systolic dysfunction.* Bmj, 1996. **312**(7025): p. 222.
10. Luisada, A.A., A. Singhal, and F. Portaluppi, *Assessment of left ventricular function by noninvasive methods.* Adv Cardiol, 1985. **32**: p. 111-41.
11. Frazier, S.K. and G.J. Skinner, *Pulmonary artery catheters: state of the controversy.* J Cardiovasc Nurs, 2008. **23**(2): p. 113-21.
12. Quinones, M.A., et al., *Recommendations for quantification of Doppler echocardiography: a report from the Doppler Quantification Task Force of the Nomenclature and Standards Committee of the American Society of Echocardiography.* J Am Soc Echocardiogr, 2002. **15**(2): p. 167-84.
13. Albert, N.M., *Bioimpedance cardiography measurements of cardiac output and other cardiovascular parameters.* Crit Care Nurs Clin North Am, 2006. **18**(2): p. 195-202.
14. Tobin, M.J., *Respiratory monitoring in the intensive care unit.* Am Rev Respir Dis. , 1988. **138**(6): p. 1625-42.
15. Selig, M.B., *Stethoscopic and phonoaudio devices: historical and future perspectives.* Am Heart J, 1993. **126**(1): p. 262-8.
16. Rosa, L.M. and A.A. Luisada, *Low frequency tracings of precordial displacement and acceleration: technical comparison of various systems.* Am J Cardiol, 1959. **4**: p. 669-74.
17. McKusick, V.A., et al., *On cardiovascular sound: further observations by means of spectral phonocardiography.* Circulation, 1955. **11**(6): p. 849-70.
18. Myint, W. and D. B., *An Electronic Stethoscope with Diagnosis Capability*, in *33rd Southeastern Symposium on System Theory*. 2001. p. 133-137.
19. Tavel, M.E., *Cardiac auscultation: a glorious past--and it does have a future!* Circulation, 2006. **113**(9): p. 1255-9.
20. Sprague, H.B. and P.A. Ongley, *The clinical value of phonocardiography.* Circulation, 1954. **9**(1): p. 127-34.

21. Rushmer, R.F., *Cardiovascular Dynamics*. 4th ed. 1978, Philadelphia: WB Saunders Co.
22. Luisada, A.A. and F. Portaluppi, *The main heart sounds as vibrations of the cardiohemic system: old controversy and new facts.* Am J Cardiol, 1983. **52**(8): p. 1133-6.
23. Shaver, J.A., R. Salerni, and P.S. Reddy, *Normal and abnormal heart sounds in cardiac diagnosis. Part I: Systolic sounds.* Curr Probl Cardiol, 1985. **10**(3): p. 1-68.
24. Sakamoto, T., et al., *Hemodynamic Determinants of the Amplitude of the First Heart Sound.* Circ Res, 1965. **16**: p. 45-57.
25. Kusukawa, R., et al., *Hemodynamic determinants of the amplitude of the second heart sound.* J Appl Physiol, 1966. **21**(3): p. 938-46.
26. Heintzen, P., *The genesis of the normally split first heart sound.* Am Heart J, 1961. **62**: p. 332-43.
27. Rosner, S.W. and S. Rodbard, *Beat-to-Beat Variation in the Split Second Heart Sound.* Am J Cardiol, 1964. **13**: p. 333-9.
28. Tavel, M.E., *Clinical Phonocardiography & External Pulse Recording*. 3rd ed. 1978, Chicago: Year Book Medical Publishers Inc.
29. Newton, A., K.B. Desser, and A. Benchimol, *Apexcardiogram in the assessment of left ventricular function.* Adv Cardiol, 1985. **32**: p. 81-95.
30. Durand, L.G. and P. Pibarot, *Digital signal processing of the phonocardiogram: review of the most recent advancements.* Crit Rev Biomed Eng, 1995. **23**(3-4): p. 163-219.
31. Rangayyan, R.M. and R.J. Lehner, *Phonocardiogram signal analysis: a review.* Crit Rev Biomed Eng, 1988. **15**(3): p. 211-36.
32. Jacobs, J.E., K. Horikoshi, and M.L. Petrovick, *Feasibility of Automated Analysis of Phonocardiogram.* Journal of the Audio Engineering Society, 1969. **17**(1): p. 49-54.
33. Yoganathan, A.P., et al., *Use of the fast Fourier transform in the frequency analysis of the second heart sound in normal man.* Med Biol Eng, 1976. **14**(4): p. 455-60.
34. Yoganathan, A.P., et al., *Use of the fast Fourier transform for frequency analysis of the first heart sound in normal man.* Med Biol Eng, 1976. **14**(1): p. 69-73.
35. Karpman, L., et al., *Sound envelope averaging and the differential diagnosis of systolic murmurs.* Am Heart J, 1975. **90**(5): p. 600-6.
36. Beyar, R., et al., *Heart-sound processing by average and variance calculation--physiologic basic and clinical implications.* IEEE Trans Biomed Eng, 1984. **31**(9): p. 591-6.
37. Gerbarg, D.S., et al., *Analysis of phonocardiogram by a digital computer.* Circ Res, 1962. **11**: p. 569-76.
38. Sarkady, A.A., R.R. Clark, and R. Williams, *Computer analysis techniques for phonocardiogram diagnosis.* Comput Biomed Res, 1976. **9**(4): p. 349-63.
39. Liang, H., S. Lukkarinen, and I. Hartimo, *Heart Sound Segmentation Algorithm Based on Heart Sound Envelogram*, in *Computers in Cardiology*. 1997. p. 105-108.

40. Nigam, V. and R. Priemer, *Accessing heart dynamics to estimate durations of heart sounds.* Physiological Measurement, 2005. **26**: p. 1005-1018.
41. Iwata, A., et al., *Algorithm for detecting the first and the second heart sounds by spectral tracking.* Med Biol Eng Comput, 1980. **18**(1): p. 19-26.
42. Joo, T.H., et al., *Pole-zero modeling and classification of phonocardiograms.* IEEE Trans Biomed Eng, 1983. **30**(2): p. 110-8.
43. Wang, W., et al., *Analysis of the first heart sound using the matching pursuit method.* Med Biol Eng Comput, 2001. **39**(6): p. 644-8.
44. Sava, H., P. Pibarot, and L.G. Durand, *Application of the matching pursuit method for structural decomposition and averaging of phonocardiographic signals.* Med Biol Eng Comput, 1998. **36**(3): p. 302-8.
45. Durand, L.G., et al., *Evaluation of FFT-based and modern parametric methods for the spectral analysis of bioprosthetic valve sounds.* IEEE Trans Biomed Eng, 1986. **33**(6): p. 572-8.
46. Akay, Y.M., et al., *Noninvasive acoustical detection of coronary artery disease: a comparative study of signal processing methods.* IEEE Trans Biomed Eng, 1993. **40**(6): p. 571-8.
47. Gamero, L.G. and R. Watrous, *Detection of the First and Second Heart Sound Using Probabilistic Models*, in *IEEE Engineering in Medicine and Biology Society*. 2003. p. 2877-80.
48. Gill, D., N. Gavriely, and N. Intrator, *Detection and Identification of Heart Sounds Using Homomorphic Envelogram and Self-Organizing Probabilistic Model*, in *Computers in Cardiology*. 2005. p. 957-960.
49. Lin, Z. and J.D. Chen, *Advances in time-frequency analysis of biomedical signals.* Crit Rev Biomed Eng, 1996. **24**(1): p. 1-72.
50. Wood, J.C. and D.T. Barry, *Quantification of first heart sound frequency dynamics across the human chest wall.* Med Biol Eng Comput, 1994. **32**(4 Suppl): p. S71-8.
51. Wood, J.C., A.J. Buda, and D.T. Barry, *Time-frequency transforms: a new approach to first heart sound frequency dynamics.* IEEE Trans Biomed Eng, 1992. **39**(7): p. 730-40.
52. Chen, D., et al., *Time-frequency analysis of the first heart sound. Part 2: An appropriate time-frequency representation technique.* Med Biol Eng Comput, 1997. **35**(4): p. 311-7.
53. Obaidat, M.S., *Phonocardiogram signal analysis: techniques and performance comparison.* J Med Eng Technol, 1993. **17**(6): p. 221-7.
54. Livanos, G., Ranganathan, N., Jiang, J. , *Heart sound analysis using the S transform*, in *Computers in Cardiology* 2000. p. 587-590.
55. Stein, P.D., et al., *Frequency spectra of the first heart sound and of the aortic component of the second heart sound in patients with degenerated porcine bioprosthetic valves.* Am J Cardiol, 1984. **53**(4): p. 557-61.
56. Durand, L.G., et al., *Comparison of pattern recognition methods for computer-assisted classification of spectra of heart sounds in patients with a porcine bioprosthetic valve implanted in the mitral position.* IEEE Trans Biomed Eng, 1990. **37**(12): p. 1121-9.

57. Guo, Z., et al., *Artificial neural networks in computer-assisted classification of heart sounds in patients with porcine bioprosthetic valves.* Med Biol Eng Comput, 1994. **32**(3): p. 311-6.
58. Bentley, P.M., P.M. Grant, and J.T. McDonnell, *Time-frequency and time-scale techniques for the classification of native and bioprosthetic heart valve sounds.* IEEE Trans Biomed Eng, 1998. **45**(1): p. 125-8.
59. Wismüller, A., Lange, O., Dersch, D. R., Leinsinger, G. L., Hahn, K., Pütz, B., and Auer, D, *Analysis of Biomedical Image Time-Series.* Int. J. Comput. Vision, 2002. **46**(2): p. 103-128.
60. Cuesta-Frau, D., J.C. Pérez-Cortés, and G. Andreu-García, *Clustering of electrocardiograph signals in computer-aided Holter analysis.* Computer Methods and Programs in Biomedicine, 2003. **72**(3): p. 179-196.
61. Adolph, R.J., J.F. Stephens, and K. Tanaka, *The clinical value of frequency analysis of the first heart sound in myocardial infarction.* Circulation, 1970. **41**(6): p. 1003-14.
62. Clarke, W.B., Austin, S.M., Pravib, M.S., Griffen, P.M., Dove, J.T., McCullough, J., Schreiner, B.F., *Spectral Energy of the First Heart Sound in Acute Myocardial Ischemia.* Circulation, 1978. **57**: p. 593-598.
63. Chen, D., et al., *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, 1997. **35**(5): p. 455-61.
64. Heckman, J.L., et al., *Frequency analysis approach to the origin of the first and second heart sounds.* Am Heart J, 1982. **104**(6): p. 1309-18.
65. Chen, D., et al., *Estimation of pulmonary artery pressure by spectral analysis of the second heart sound.* Am J Cardiol, 1996. **78**(7): p. 785-9.
66. Xu, J., 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.* IEEE Trans Biomed Eng, 2001. **48**(3): p. 277-83.
67. Tranulis, C., et al., *Estimation of pulmonary arterial pressure by a neural network analysis using features based on time-frequency representations of the second heart sound.* Med Biol Eng Comput, 2002. **40**(2): p. 205-12.
68. Xiao, Y., et al., *The Phonocardiogram Exercise Test.* IEEE Engineering in Medicine and Biology, 1999. **18**(4): p. 111-115.
69. Cohn, P.F., et al., *Diastolic Heart Sounds during Static (Handgrip) Exercise in Patients with Chest Pain.* Circulation, 1973. **47**: p. 1217-1221.
70. DeGroff, C.G., et al., *Artificial neural network-based method of screening heart murmurs in children.* Circulation, 2001. **103**(22): p. 2711-6.
71. Thompson, W.R., et al., *Automated cardiac auscultation for detection of pathologic heart murmurs.* Pediatr Cardiol, 2001. **22**(5): p. 373-9.
72. Debiais, F., et al., *Time-frequency analysis of heart murmurs. Part II: Optimisation of time-frequency representations and performance evaluation.* Med Biol Eng Comput, 1997. **35**(5): p. 480-5.
73. Syed, Z., et al., *A Framework for the Analysis of Acoustical Cardiac Signals.* IEEE Trans Biomed Eng, 2007. **54**(4): p. 651-662.

74. Tanaka, S., Y. Matsumoto, and K. Wakimoto, *Unconstrained and non-invasive measurement of heart-beat and respiration periods using a phonocardiographic sensor.* Med Biol Eng Comput, 2002. **40**(2): p. 246-52.
75. Benchimol, A. and E.G. Dimond, *The Normal and Abnormal Apexcardiogram. Its Physiologic Variation and Its Relation to Intracardiac Events.* Am J Cardiol, 1963. **12**: p. 368-81.
76. Manolas, J., P. Wirz, and W. Rutishauser, *Relationship between duration of systolic upstroke of apexcardiogram and internal indexes of myocardial function in man.* Am Heart J, 1976. **91**(6): p. 726-34.
77. Antani, J.A., H.H. Wayne, and W.J. Kuzman, *Ejection phase indexes by invasive and noninvasive methods: an apexcardiographic, echocardiographic and ventriculographic correlative study.* Am J Cardiol, 1979. **43**(2): p. 239-47.
78. Manolas, J. and W. Rutishauser, *Diastolic amplitude time index: a new apexcardiographic index of left ventricular diastolic function in human beings.* Am J Cardiol, 1981. **48**(4): p. 736-45.
79. Manolas, J., et al., *Usefulness of noninvasive detection of left ventricular diastolic abnormalities during isometric stress in hypertrophic cardiomyopathy and in athletes.* Am J Cardiol, 1998. **81**(3): p. 306-13.
80. Garrard, C.L., 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, 1970. **42**(3): p. 455-62.
81. Swatzell, R.H., Jr., et al., *The use of the systolic time interval for predicting left ventricular ejection fraction ischemic heart disease.* Am Heart J, 1977. **93**(4): p. 450-4.
82. Asmar, R., et al., *Assessment of arterial distensibility by automatic pulse wave velocity measurement. Validation and clinical application studies.* Hypertension, 1995. **26**(3): p. 485-90.
83. Gavriely, N., *Breath Sounds Methodology*. 1995: CRC.
84. Nikitin, N.P. and K.K. Witte, *Application of tissue Doppler imaging in cardiology.* Cardiology, 2004. **101**(4): p. 170-84.
85. D'Hooge, J., et al., *Regional strain and strain rate measurements by cardiac ultrasound: principles, implementation and limitations.* Eur J Echocardiogr, 2000. **1**(3): p. 154-70.
86. Weidemann, F., et al., *Can strain rate and strain quantify changes in regional systolic function during dobutamine infusion, B-blockade, and atrial pacing--implications for quantitative stress echocardiography.* J Am Soc Echocardiogr, 2002. **15**(5): p. 416-24.
87. Greenberg, N.L., et al., *Doppler-derived myocardial systolic strain rate is a strong index of left ventricular contractility.* Circulation, 2002. **105**(1): p. 99-105.
88. Kohler, B.U., C. Hennig, and R. Orglmeister, *The principles of software QRS detection.* IEEE Eng Med Biol Mag, 2002. **21**(1): p. 42-57.
89. Allen, J.B., *Short Term Spectral Analysis, Modification and Synthesisby Discrete Fourier Transform.* IEEE Transactions on Acoustics, Speech, and Signal Processing, 1977. **25**(3): p. 235-238.

90. Mallat, S.G., *A Theory for Multiresolution Signal Decomposition: The Wavelet Representation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989. **11**(7): p. 674-693.
91. Stockwell, R.G., Mansinha, L., Lowe R.P., *Localization of the complex spectrum: The S-Transform.* IEEE Trans. Signal Process., 1996. **44**(4): p. 998-1001.
92. Cohen, L., *Time-frequency distributions- A review.* Proc. of the IEEE, 1989. **77**: p. 941-981.
93. Duda, R.O. and P.E. Hart, *Pattern Classification and Scene Analysis*. 1973, New York: Wiley.
94. Rabiner, L.R. and B.W. Juang, *Fundamentals of Speech Recognition*. Prentice Hall Signal Processing Series, ed. A.V. Oppenheim. 1993, Englewood Cliffs, NJ: Prentice Hall
95. Johnson, S.C., *Hierarchical Clustering Schemes.* Psychometrika, 1967. **2**: p. 241-254.
96. Ward, J.H., *Hierarchical Grouping to Optimize an Objective Function.* Journal of the American Statistical Association, 1963. **58**(301): p. 236-244.
97. Fisher, R., *The Statistical Utilization of Multiple Measurements.* Annals of Eugenics, 1938. **8**: p. 376-386.
98. Amit, G., Gavriely, N., Intrator, N., *Cluster Analysis and Classification of Heart Sounds.* Biomedical Signal Processing and Control, 2009. **4**: p. 26-36.
99. Amit, G., Gavriely, N., Intrator, N., *Respiratory Modulation of Heart Sound Morphology.* American Journal of Physiology – Heart and Circulatory Physiology, 2009. **296**: p. 796-805.
100. Woody, C.D., *Characterization of an adaptive filter for the analysis of variable latenc.* Medical and Biological Engineering and Computing, 1967. **5**(6): p. 539-554.
101. Berndt, D. and J. Clifford, *Using dynamic time warping to find patterns in time series*, in *AAAI Workshop on Knowledge Discovery in Databases*. 1994. p. 229-248.
102. Keogh, E. and M. Pazzani, *Scaling Up Dynamic Time Warping for Data Mining Applications*, in *The 6th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2000. p. 285-289.
103. Syeda-Mahmood, T., D. Beymer, and F. Wang, *Shape-based Matching of ECG Recordings*, in *The 29th Annual International Conference of the IEEE EMBS*. 2007. p. 2012-2018.
104. Caiani, E.G., et al., *Warped-average template technique to track on a cycle-by-cyclebasis the cardiac filling phases on left ventricular volume*, in *Computers in Cardiology*. 1998. p. 73-76.
105. Witkin, A.P., *Scale-Space Filtering: A New Approach to Multi-Scale Description*, in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. 1984. p. 150-153.
106. Amit, G., N. Gavriely, and N. Intrator, *Automatic Segmentation of Heart Signals*, in *The 18th EURASIP Conference BIOSIGNAL* 2006. p. 6-8.
107. Millioz, F., J. Huillery, and N. Martin. *Short Time Fourier Transform Probability Distribution for Time-Frequency Segmentation*. in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. 2006. Toulouse,France.

108. *Heart Sounds and Murmurs (Online resource)*. 2006, Texas Heart Institute.
109. Amit, G., Gavriely, N., Lessick, J., Intrator, N., *Acoustic Indices of Cardiac Functionality*, in *International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS)*. 2008. p. 77-83.
110. Amundsen, B.H., et al., *Noninvasive myocardial strain measurement by speckle tracking echocardiography: validation against sonomicrometry and tagged magnetic resonance imaging.* J Am Coll Cardiol, 2006. **47**(4): p. 789-93.
111. Amit, G., Gavriely, N., Intrator, N., *Cluster Analysis and Classification of Heart Sounds.* Biomedical Signal Processing and Control, to appear.
112. Amit, G., Gavriely, N., Intrator, N., *Respiratory Modulation of Heart Sound Morphology.* American Journal of Physiology – Heart and Circulatory Physiology, submitted.
113. Ahdesmaki, M., et al., *Robust regression for periodicity detection in non-uniformly sampled time-course gene expression data.* BMC Bioinformatics, 2007. **8**: p. 233.
114. Fisher, R., *Tests of significance in harmonic analysis.* Proceedings of the Royal Society of London, 1929. **125**: p. 54-59.
115. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the Royal Statistical Society. Series B, 1995. **57**(1): p. 289-300.
116. Durand, L.G., et al., *Spectral analysis and acoustic transmission of mitral and aortic valve closure sounds in dogs. Part 1. Modelling the heart/thorax acoustic system.* Med Biol Eng Comput, 1990. **28**(4): p. 269-77.
117. Gunopulos, D. and G. Das. *Time Series Similarity Measures and Time Series Indexing (Tutorial)*. in *6th ACM SIGKDD international conference on knowledge discovery and data mining*. 2001. Boston, MA: ACM.
118. Keogh, E.J. and M.J. Pazzani, *An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback*, in *4th International Conference on Knowledge Discovery and Data Mining*. 1998. p. 239-243.
119. Salvador, S. and P. Chan, *Toward accurate dynamic time warping in linear time and space.* Intelligent Data Analysis, 2007. **11**(5): p. 561-580
120. Zhou, M. and M.H. Wong, *A segment-wise time warping method for time scaling searching.* Information Sciences, 2005. **173**: p. 227-254.
121. Perng, C.S., et al., *Landmarks: A New Model for Similarity-Based Pattern Querying in Time Series Databases*, in *16th International Conference on Data Engineering*. 2000. p. 33-42.
122. Burges, C.J.C., *A tutorial on Support Vector Machines for pattern recognition.* Data Mining and Knowledge Discovery, 1998. **2**(2): p. 121-167.
123. Ahonen, J., et al., *Pharmacokinetic-pharmacodynamic relationship of dobutamine and heart rate, stroke volume and cardiac output in healthy volunteers.* Clin Drug Investig, 2008. **28**(2): p. 121-7.
124. Scharf, S.M., et al., *Effects of normal and loaded spontaneous inspiration on cardiovascular function.* J Appl Physiol, 1979. **47**(3): p. 582-90.
125. Ishikawa, K. and T. Tamura, *Study of respiratory influence on the intensity of heart sound in normal subjects.* Angiology, 1979. **30**(11): p. 750-5.

126. Scharf, S.M., et al., *Respiratory phasic effects of inspiratory loading on left ventricular hemodynamics in vagotomized dogs.* J Appl Physiol, 1992. **73**(3): p. 995-1003.

APPENDIX A – INFORMED CONSENT FORMS

Dobutamine stress echo study

טופס הסכמה מדעת להשתתפות בניסוי רפואי בבני אדם

מספר הבקשה בוועדת הלסינקי (למילוי על-ידי מזכירות הוועדה): _____

<u>אני החתום מטה</u>:

	שם פרטי ומשפחה:
	מספר תעודת זהות:
מיקוד:	כתובת:

א) מצהיר/ה בזה כי אני מסכים/ה להשתתף בניסוי רפואי, כמפורט במסמך זה.

ב) מצהיר/ה בזה כי איני משתתף בזמן חתימת מסמך זה, בניסוי רפואי אחר הכרוך בשימוש במוצר מחקר כלשהו, וכי אני מתחייב/ת לא להשתתף בכל ניסוי רפואי אחר הכרוך בשימוש במוצר מחקר במשך כל תקופת ניסוי זה.

ג) מצהיר/ה בזה כי הוסבר לי על-ידי:

שם החוקר/חוקר המשנה המסביר:

1. כי החוקר הראשי (שם הרופא): <u>ד"ר יהונתן לסיק</u> קיבל ממנהל המוסד הרפואי, בו ייערך הניסוי, אישור לביצוע הניסוי הרפואי בבני-אדם, כמשמעותו בתקנות בריאות העם (ניסויים רפואיים בבני-אדם) תשמ"א-1980 (להלן הניסוי הרפואי).

2. כי לחוקר הראשי ולחוקרי המשנה יש/~~אין~~[1] זיקה[2] ליוזם הניסוי[3].
אם יש – פרט: החוקר הראשי הוא גם יוזם הניסוי

3. כי הניסוי הרפואי נערך בנושא:
<u>הקלטת אותות לב מבית החזה</u>

4. כי אני חופשי/ה לבחור שלא להשתתף בניסוי הרפואי, וכי אני חופשי/ה להפסיק בכל עת את

---

[1] מחק את המיותר

[2] קשר של העסקה בשכר, או קשר מסחרי או עסקי, או קשר משפחתי או אישי, וכל קשר אחר, לרבות קשר של כפיפות בעבודה, שיש בו כדי לעורר חשש לקיום ניגוד עניינים או תלות, ולמעט החזר הוצאות או תשלום עבור השתתפות בוועדות לפי נוהל זה.

[3] אם החוקר הראשי הוא גם יוזם הניסוי, יש לציין זאת במפורש.

השתתפותי בניסוי, כל זאת מבלי לפגוע בזכותי לקבל את הטיפול המקובל.

5. כי במקרה של מילוי שאלון – אני רשאי/ת שלא לענות על כל השאלות שבשאלון או על חלק מהן.

6. כי מובטח לי שזהותי האישית תשמר סודית על-ידי כל העוסקים והמעורבים במחקר ולא תפורסם בכל פרסום, כולל בפרסומים מדעיים.

7. כי המוסד הרפואי פעל להסדרת כיסוי ביטוחי הולם של החוקרים, הרופאים והצוות הרפואי העוסקים בניסויהקליני מפני תביעות שיוגשו ע"י משתתפים בניסוי הקליני ו/או תביעות צד ג' הקשורות עם הניסוי הקליני בין בתקופת ביצוע הניסוי ובין לאחריו. אין באמור כדי לפגוע בזכויותיי על פי כל דין.

8. כי במקרה הצורך, לפי המלצת החוקר הראשי, קיימת האפשרות שאמשיך לקבל את מוצר המחקר ללא תשלום גם לאחר סיום הניסוי הרפואי לתקופה של שלוש שנים, כאשר לא נמצא לי טיפול רפואי חליפי מתאים. זאת, בין היתר, בתנאי שהמוצר עדיין לא אושר לשימוש בהתוויה המבוקשת בניסוי במדינת ישראל, ולא ניתן לקבלו מקופת החולים/שירותי הבריאות בו/בהם אני מבוטח/ת.
ההחלטה לגבי המשך מתן מוצר המחקר נתונה בידי ועדת הלסינקי המוסדית, והכל כפוף לקיומו של תכנית טיפול ומעקב.

9. כי מובטחת לי נכונות לענות לשאלות שיועלו על-ידי וכן האפשרות להיוועץ בגורם נוסף (לדוגמא רופא-משפחה, בני משפחה וכו'), באשר לקבלת החלטה להשתתף בניסוי הרפואי ו/או להמשיך בו.

10. כי בניסויים רפואיים בהם משתתפות נשים בגיל הפוריות, במקרה של הריון במהלך הניסוי הרפואי, האישה תקבל ייעוץ (על-ידי החוקר) לגבי השפעות שייתכנו על העובר ולגבי גורל ההריון, כולל האפשרות של הפסקת ההריון.

11. כי בכל בעיה הקשורה לניסוי הרפואי אוכל לפנות לפרופ'/ד"ר יהונתן לסיק
מספר טלפון/משיבון: 04-8542342, בכל שעות היממה.

ד) הנני מצהיר/ה כי נמסר/ה לי מידע מפורט על הניסוי הרפואי, על פי הנושאים המפורטים להלן:

1. מטרות הניסוי;

האזנה לקולות לב וקולות נשימה מקובלת מזה שנים רבות ומתבצעת בדרך כלל על ידי הרופא באמצעות הסטטוסקופ. ההאזנה לקולות מספקת לרופא מידע חשוב על תפקוד הלב והנשימה. גם הקלטת קולות מהגוף מוכרת וידועה ומשמשת לצרכים קליניים ומחקריים. הקלטה של קולות מכמה מתמרים (מיקרופונים) בעת ובעונה אחת משמשת כדי לזהות קשרי גומלין בין הקולות באזורים שונים של בית החזה ומאפשרת אבחון יותר טוב של מצב החולה והבנה משופרת של יצירת הקולות.

2. המספר בקירוב של המשתתפים בניסוי הרפואי;

בניסוי הרפואי ישתתפו כ-40 נבדקים.

3. התקופה הצפויה למשך ההשתתפות בניסוי;

ההשתתפות בניסוי היא חד פעמית, למשך זמן הבדיקה בלבד

ההשתתפות הניסוי תאריך את זמן הבדיקה בכ-15 דקות. זמן הבדיקה הכולל הוא כ-40 דקות.

4. שיטות: תיאור מוצר המחקר, תיאור ההליכים השונים במשך תקופת הניסוי (טיפול ומעקב), תוך הבחנה ברורה בין ההליכים המחקריים לבין ההליכים המקובלים ברפואה; ציון הסיכויים של המשתתף לקבל כל אחד מהטיפולים המוצעים בניסוי (כולל פלצבו, במידה שקיים);

מערכת ההקלטה בה נשתמש בניסוי כוללת שלושה מיקרופונים בגודל של כ 3 ס"מ כל אחד שיוצמדו לדופן החזה באמצעות סרט אלסטי או סרט דביק (פלסתר). גלי הקול מועברים דרך מגבר למחשב עליו הם מוקלטים ובו מתבצע עיבוד האותות. במקביל נקליט גם את אות ה EKG הסטנדרטי ואותות ממכשיר האקוקרדיוגרף.

מהלך הניסוי לנבדק במהלך בדיקת מאמץ אקוקרדיוגרפית:
במהלך כל הקלטה שתימשך 60 שניות תתבקש/י לשכב על צד שמאל, להימנע מלדבר ולנשום נשימה קלה ככל האפשר. מלבד זאת אין כל דרישה מגבילה או צורך בהשתתפותך הפעילה. במהלך בדיקת המאמץ האקוקרדיוגרפית יבוצעו כ-10 הקלטות. בסה"כ תאריך השתתפותך בניסוי את הבדיקה בכ-15 דקות.

מהלך הניסוי לנבדק במהלך בדיקה אקוקרדיוגרפית שגרתית או מתנדב:
במהלך כל הקלטה תתבקש/י לשכב על צד שמאל, להימנע מלדבר ולנשום נשימה קלה ככל האפשר. במהלך שתי הקלטות תתבקש/י לבצע מאמץ של כיווץ כף היד כנגד מכשיר המפעיל התנגדות, ולהתמיד בהפעלת מאמץ קבוע במשך 3 דקות. שתי הקלטות נוספות יבוצע במצב מנוחה ללא צורך בהשתתפותך הפעילה. בסה"כ תאריך השתתפותך בניסוי את הבדיקה בכ-15 דקות.

5. היתרונות הצפויים למשתתף או לאחרים, כתוצאה מהניסוי;

השתתפותך בניסוי היא על בסיס התנדבותי ונועדה לאפשר לנו להבין טוב יותר את קשרי הגומלין בין קולות הלב, הפעילות הנשימתית ותרשים ה EKG במצבי חולי שונים. התועלת הצפויה מהשתתפותך בניסוי היא מדעית ואין אנו צופים שתהינה לתוצאות ההקלטה השפעות כלשהן (חיוביות או שליליות) על מצבך או על הטיפול בך.

6. הסיכונים הידועים ו/או אי-הנוחות שניתן לחזותם למשתתף במחקר;
במידה שיש בניסוי הרפואי סיכון למשתתף - הסבר על הטיפול הרפואי שיקבל במקרה של פגיעה בבריאותו והאחריות לנתינתו;

ההקלטה היא לחלוטין בלתי פולשנית. אין אנו צופים תופעות לוואי או סיבוכים כלשהם.

7. נסיבות בהן עלולה השתתפותו בניסוי הרפואי להיפסק בהחלטת החוקר או היוזם;

השתתפותך בניסוי עשויה להיפסק בהחלטת החוקר במידה והאיכות הטכנית של ההקלטה תהיה בלתי מספקת.

8. לפי העניין, החוקר ימסור למשתתף מידע על תוצאות רפואיות אפשריות של החלטת <u>המשתתף</u> על הפסקת השתתפותו בניסוי הרפואי לפני סיומו;

אין אתה חייב להשתתף בניסוי והשתתפותך בו לא תשפיע על הטיפול בך. כמו"כ אתה רשאי להפסיק את השתתפותך בכל עת ואין אתה חייב הסבר על הסבות לכך.

9. הסבר על טיפולים חלופיים, ועל יתרונותיהם וחסרונותיהם, באם ישנם כאלה, למשתתף;

10. מידע רלוונטי אחר (כפי שנמסר על-ידי יוזם הניסוי):

ה) הנני מצהיר/ה בזה כי את הסכמתי הנ"ל נתתי מרצוני החופשי וכי הבינותי את כל האמור לעיל. כמו-כן, קיבלתי עותק של טופס הסכמה מדעת זה, נושא תאריך וחתום כדין.

עם חתימתי על טופס הסכמה זה, הנני מתיר ליוזם הניסוי הרפואי, לוועדת הלסינקי המוסדית, לגוף המבקר במוסד הרפואי ולמשרד הבריאות גישה ישירה לתיקי הרפואי, לשם אימות שיטות הניסוי הרפואי והנתונים הקליניים. גישה זו למידע הרפואי שלי תבוצע תוך שמירת סודיות, בהתאם לחוקים ולנהלים של שמירת סודיות.

ו) במקרים שבהם הניסוי הרפואי כרוך במתן שירותים: ביצוע בדיקות רפואיות או באספקת אביזרים, תכשירים או משתלים, הנני מצהיר/ה בזה כי אני יודע/ת ומסכים/ה שהמידע על השתתפותי בניסוי הרפואי יועבר לרופא המטפל שלי בקופת החולים/שירותי הבריאות[1] בה/בהם אני מבוטח/ת.
ידוע לי כי בקופת החולים/שירותי הבריאות לא ייעשה במידע זה כל שימוש, אלא לצורך טיפול ומעקב רפואיים בלבד.

תאריך	חתימת המשתתף/ת בניסוי	שם המשתתף/ת בניסוי הרפואי

במקרה הצורך[2]

תאריך	חתימת העד	מספר תעודת זהות	שם העד הבלתי תלוי

<u>הצהרת החוקר/חוקר המשנה</u>:

ההסכמה הנ"ל נתקבלה על-ידי, וזאת לאחר שהסברתי למשתתף/ת בניסוי הרפואי כל האמור לעיל וכן וידאתי שכל הסבריי הובנו על-ידו/ידה.

תאריך	חתימתו	שם החוקר/חוקר המשנה המסביר

[1] מחק את המיותר

[2] במקרה שהמשתתף בניסוי, או נציגו החוקי, אינו מסוגל לקרוא את טופס ההסכמה מדעת, עד בלתי תלוי חייב להיות נוכח במשך ההסבר על מהות הניסוי הרפואי. לאחר שהמשתתף או נציגו החוקי הביע את הסכמתו בעל-פה להשתתפות בניסוי, העד יחתום על טופס ההסכמה, תוך ציון תאריך החתימה.

Respiratory pressure study

טופס לנבדק

שם פרטי:_____ שם משפחה:_____
ת.ז: _____ גיל: _____

- נא לסמן X אם הינך מקיימ/ת את אחד הסעיפים:

1. אני ידוע כסובל/ת ממחלה לבבית איסכמית או כל מחלה לבבית אחרת _____
2. אני סובל/ת מהפרעות קצב לבביות _____
3. אני סובל/ת מיתר לחץ דם _____
4. אני סובל/ת ממחלה ראתית כרונית _____
5. אני סובל/ת ממחלה ראתית אקוטית _____
6. אני סובל/ת ממחלה לבבית- ראתית כרונית / אקוטית _____
7. אני מעשן/ת או עישנתי עד לפני 12 חודשים _____ ואם כן : מס' קופסאות ליום :___
8. אני סובל/ת ממחלת עור מדבקת _____
9. אני סובל/ת ממחלת עור המתבטאת באלרגיה לחומר מסויים _____
10. יש לי צלקות רחבות ע"פ עור צוואר ו/או ע"פ עור בית החזה _____

- נא לקרוא את מהלך הניסוי, ולודא כי הבנת אותו. אם הינך מרגיש/ה כי אין ביכלותך לבצע את הניסוי אזי את/ה מתבקש/ת לא לחתום בסוף הטופס.

מטרת הניסוי היא לבדוק ולהוכיח את הקשר הקיים בין הלחצים המשתנים בתוך בית החזה במחזור הנשימה (שאיפה ונשיפה) לשינויים בקולות הלב.

במשך הניסוי את/ה מתבקש/ת לנשום באופן נורמלי דרך הפה בלבד (ישנה אפשרות לסגור את האף בקליפס ספיציפי) ולעצור את הנשימה לפרק זמן קצר, הכל נעשה מול התנגדויות מסויימות; דבר שיכול לדמות מצבים בהם אדם מפעיל לחצים גדולים בעת נשימה. תוך כדי הנשימה יוקלטו קולות לב, ECG וקולות נשימה, וימדדו הלחצים שאת/ה מפעיל/ה בעת הנשימה.

- להקלטת הלחצים שאת/ה מפעיל/ה תוך כדי הנשימה, את/ה מתחבר/ת לפיה מקרטון אשר דרכה תנשום/ תנשמי.
  ע"מ לוודא שהנשימה שלך מתבצעת אך ורק דרך הפה, ישנה אפשרות לסגור את האף ע"י קליפס ספיציפי.
- ההתנגדות לנשימה מיושמת דרך צינורות (פלסטיק) צרים (בקוטר של 0.5 cm) באורכים שונים. ככל השצינור ארוך יותר, ההתנגדות גדולה יותר. ההתנגדויות מולן תנשום/תנשמי נעות בין 0 התנגדות (ללא חיבור שום צינור) ועד התנגדות מס' 4, הצינור האורך ביותר.
- הנשימה תתבצע לפי שני פרוטוקולים:

פרוטוקול 1. : נשימה רגילה – 10 שניות
הפסקת נשימה – 15 שניות
נשימה רגילה – 15 שניות

פרוטוקול 2. : נשימה רגילה דרך הפה למשך 40 שניות.
כל פרוטוקול יבוצע פעמיים עם כל התנגדות. ובכך סה"כ אורך הניסוי הינו מקסימום 40-30 דק'.

- החיבורים: בשביל שנוכל לחבר את החיבורים המתאימים תתצרכ/י להוריד את החלק העליון של הלבוש .

ע"מ להקליט את קולולת הלב, נצמיד לבית החזה, בגובה מרווח בין צלעי רביעי, 2 סמוגרפים במרחק שווה ימנית ושמאלית לקו האמצע.
חיבורי ה- ECG יהיו: כתף ימין, כתף שמאל, ואזור בטן שמאלית עליונה.
להקלטת קולות הנשימה: נחבר סמוגרף לצוואר, ימנית לקו האמצע.

- המיזוג בחדר יותאם לנוחיותך.
- בחדר ימצאו הבודק והנבדק בלבד.
- תוך כדי הניסוי צריך לשמור על השקט בכדי לוודא הקלטה נכונה.
- קצב הניסוי יהיה לפי נוחיותך, במקרה ואת/ה מתעייפ/ת אפשר להפסיק לכמה דקות ולהמשיך בעת שתהיה/תהיי מוכנ/ה.
- את תוצאות הניסוי יראו הבודק והמנחים שיפענחו את התוצאות.

אני (שם פרטי ושם משפחה) _____ נושא ת.ז _____
מצהיר כי אני קראתי את הטופס במלואו והבנתי אותו בשלימותו ,ויודע/ת כי אני מקיים את התנאים המוכללים ושאני מסוגל לבצע את התרגילים המוסברים הנ"ל.

תאריך: _____ חתימה: _____

תוצאות פרק זה פורסמו במאמר:

Amit, G., Gavriely, N., Intrator, N.: Respiratory Modulation of Heart Sound Morphology. *American Journal of Physiology–Heart and Circulatory Physiology* 296:796-805, 2009.

דיון ומסקנות (פרק 6)

הקשר בין התהליכים הפיזיולוגיים היוצרים את האותות הויברו-אקוסטיים לבין המורפולוגיה של האותות הנקלטים על פני בית החזה הוא קשר מורכב, המושפע מהדינמיקה של השריר המתכווץ, מזרימת הדם ומפעולת המסתמים, כמו גם מבקרה עצבית והורמונלית ומתכונות התווך דרכו עוברות התנודות המכניות. האותות הויברו-אקוסטיים, בהיותם תוצאה ישירה של מחזור הלב המכני, נושאים מידע אודות תפקוד מערכת הלב, כלי הדם והנשימה. השיטות והתוצאות שתוארו בעבודה המוצגת כאן מספקות תובנות חדשות הן אודות הכלים החישוביים המתאימים לעיבוד אותות אלה והן אודות הקשרים הפיזיולוגיים בין המורפולוגיה של האותות לבין תפקוד הלב והפעילות הנשימתית. בפרק הדיון מפורטות ונדונות תובנות אלה ומתוארות המסקנות העולות מהעבודה לגבי ישימות השיטות לניטור רציף ולא פולשני של התפקוד המכני של הלב והנשימה. כמו-כן, מתוארות מגבלות העבודה ומוצעים כיוונים למחקר עתידי.

וברוחב הפס של האות. ניתוח באמצעות התמרת זמן-תדר (S-transform) והפעלת Principal Component Analysis שימש לאפיון השינויים ב-S1 ולהגדרת מדדים אקוסטיים חד-מימדיים המתארים את הדינמיקה המורפולוגית של S1 תוך כדי מבחן המאמץ. המדדים האקוסטיים הושוו לנתוני ייחוס אקוקרדיוגרפיים המייצגים את התפקוד הסיסטולי הגלובלי של החדר, ואשר חושבו בכל אחד משלבי מבחן המאמץ באמצעות טכניקה הקרוייה 2D Strain Echocardiography. בין המדדים האקוסטיים והאקוקרדיוגרפיים נמצא מתאם גבוה בכל הנבדקים, אשר חיזק את הישימות של ניטור התפקוד הסיסטולי של החדר באמצעות ניתוח הקול הראשון של הלב. תוצאות אלה פורסמו במאמר:

Amit, G., Gavriely, N., Lessick, J., Intrator, N.: Acoustic Indices of Cardiac Functionality. *International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS) 2008*, Vol 2, pp 77-83.

ניתוח נוסף של השינויים המורפולוגיים של S1 במהלך מבחן מאמץ פרמקולוגי נעשה באמצעות שיטת האַשְׁכּוּל (clustering) והסיווג שתוארה בפרק 3. החלוקה האוטומטית של אותות S1 על-פי המורפולוגיה שלהם בייצוג בזמן או בזמן-תדר זיהתה את תבנית השינויים הספקטרליים ב-S1, הנגרמים על-ידי תגובת המאמץ. הקשר בין המורפולוגיה של S1 ורמת המאמץ הודגם ביכולת של מסווג (classifier) לחזות את השלב במבחן המאמץ מתוך האות בדיוק ממוצע גבוה של 7%±86 עבור כל הנבדקים. התוצאות המוצגות כוללות גם ניתוח של השפעת ייצוג האות, פונקציית המרחק ושיטת הסיווג על ביצועי השיטה.

הערכת התפקוד הנשימתי על-ידי ניתוח אותות לב (פרק 5)

הפרק החמישי עוסק בניתוח ההשפעה של הפעילות הנשימתית על קולות הלב ובחינת האפשרות לחזות פרמטרים של תפקוד נשימתי כגון פאזת הנשימה (שאיפה, נשיפה, עצירת נשימה) ומידת ההתנגדות לנשימה על-ידי ניתוח השינויים המורפולוגיים באותות האקוסטיים. בניתוח האותות שהוקלטו תוך כדי שינויים בדפוס הנשימה הודגמה שונות מורפולוגית ניכרת של S1 ו-S2 בזמן נשימה רגילה, שאינה קיימת בעת עצירת נשימה. שונות זו היתה בעלת התנהגות מחזורית, וזמן המחזור ששוערך מתוך מדידת תבנית ההשתנות המורפולוגית של הקולות הראה התאמה טובה למשך מחזור הנשימה. ניצפו הבדלים משמעותיים בתכונות של S1 ו-S2 בזמן שאיפה ובזמן נשיפה. הקול הראשון נחלש ומושהה בזמן שאיפה, בעוד הקול השני מתחזק ומקדים. מורפולוגיה של פיצול (split) בין מרכיבי האות זוהתה בזמן שאיפה הן ב-S1 והן ב-S2, בהתאם לידוע מהספרות. כאשר רמת ההתנגדות לנשימה עולה, שינויים אלה הופכים מודגשים יותר וכן מתרחשים בשלב מוקדם יותר במחזור הנשימה. המורפולוגיות האופייניות של קולות הלב, הקשורות לפאזת הנשימה ולמידת המאמץ הנשימתי, זוהו על-ידי clustering אוטומטי, ואלגוריתם הסיווג הצליח לחזות את פאזת הנשימה ואת מידת ההתנגדות לנשימה מתוך המורפולוגיה של S1 בדיוק ממוצע גבוה של 7%±87 ו-7%±82, בהתאמה. הניתוח הכמותי של קולות הלב עשוי, אם כן, לשמש גם לניטור של התפקוד הנשימתי והקשר בין מערכות הלב והריאות.

העיקריים. בפרק זה מוצגות מספר שיטות לשְׁיווּר (alignment) בסיסי של אותות בתחום הזמן, ובהן phase-shift averaging ו-dynamic time warping. בהמשך, מתוארת שיטה לזיהוי ארועים משמעותיים באותות נמוכי תדר כגון אותות דופק, המבוססת על ייצוג הסיגנל במרחב בעל סקלה משתנה (scale space) וזיהוי נקודות קיצון במרחב זה. בין הנקודות שזוהו בכל פעימה של האות מבוצע יישור על-ידי הפעלת dynamic time-warping, ונקודות שרמת החזרתיות שלהן בכל הפעימות גבוהה ומידת השונות שלהן בזמן ובמשרעת נמוכה, נבחרות כנקודות סגמנטציה. שיטה אחרת עושה שימוש בפרמטרים סטטיסטיים של ייצוג הזמן-תדר לזיהוי איטרטיבי (iterative) של המרכיבים העיקריים בייצוג זה של הסיגנל. על המרכיבים שזוהו מבוצע אִשכּוּל (clustering) שמסווג אותם לקבוצות, כך שנוצרת סגמנטציה אוטומטית שאינה עושה שימוש במידע חיצוני לחלוקת האות למחזורי לב, וכן אינה זקוקה לידע כלשהו אודות מבנה הסיגנל הצפוי. מודגמת הפעלת שיטות אלה בהצלחה על אותות לב ויברו-אקוסטיים. תוצאות פרק זה תוארו במאמרים:

Amit, G., Gavriely, N., Intrator, N.: Automatic Segmentation of Heart Signals, *Proceedings of the 18th EURASIP Conference BIOSIGNAL 2006*:6-8.

Amit, G., Gavriely, N., Intrator, N.: Segmentation and Alignment of Heart Signals. In preparation.

הערכת התפקוד המכני של הלב על-ידי ניתוח אותות לב (פרק 4)

הפרק הרביעי של העבודה מתאר את יישומן של שיטות ניתוח האות לבחינת הקשר בין התפקוד המכני של הלב לבין האותות הויברו-אקוסטיים, ולהסקת מידע פיזיולגי אודות תפקוד הלב מתוך תכונות האות. יישום ראשון, ששימש כהוכחת היתכנות לגישה המחקרית של העבודה, היה זיהוי ארועים במחזור הפעילות המכנית של הלב מתוך אותות הדופק והקולות ואמידה אוטומטית של מרווחי זמן במחזור הלב, תוך השוואה לנתוני ייחוס אקוקרדיוגרפיים. עיבוד האותות הויברו-אקוסטיים כלל סגמנטציה יוריסטית של S1 ו-S2 ושימוש במיקום הקולות לזיהוי המיקום של ארועים סיסטוליים ודיאסטוליים באות הדופק בעורק הקרוטידי (carotid pulse) ובהולם חוד הלב (apex pulse). במקביל, הוקלטו אותות אקו-דופלר שאפשרו שחזור של דפוס זרימת הדם מהעליה לחדר ומהחדר לאבי העורקים. המתאם בין האירועים שזוהו באות האקוקרדיוגרפי ובאותות הויברו-אקוסטיים הודגם הן באופן ויזואלי והן בהשוואה כמותית בין הערכים הנמדדים של זמני הזרקת הדם (ejection time) ומילוי החדר (filling time). תוצאות אלה פורסמו במאמר:

Amit, G., Gavriely, N., Lessick, J., Intrator, N.: Automatic Extraction of Physiological Features from Vibro-Acoustic Heart Signals: Correlation with Echo-Doppler. *Computers in Cardiology 2005*: 299-302.

ביישום שני של ניתוח כמותי של קולות לב נעשה שימוש בתכונות ספקטרליות של הקול הראשון, S1, לשם ניטור רציף של התפקוד הסיסטולי של החדר השמאלי. באותות S1 שהוקלטו תוך כדי שינוי כושר ההתכווצות (contractility) של הלב בזמן בדיקת מאמץ פרמקולוגית זוהתה תבנית של עלייה באנרגיה הספקטרלית

במסגרת עבודה זו בוצעו שני סוגים של הקלטות של אותות לב ויברו-אקוסטים: הניסוי הראשון כלל הקלטות מ-11 נבדקים תוך כדי ביצוע מבחן מאמץ פרמקולוגי שגרתי, ואיסוף מקביל של מידע אקוקרדיוגרפי על התפקוד הסיסטולי של חדר שמאל. הניסוי השני כלל הקלטות מ-12 נבדקים תוך שינויים בדפוס הנשימה בזמן מעבר מנשימה רגילה לעצירת נשימה ובזמן שינוי רמת ההתנגדות לנשימה, תוך הקלטת אות הלחץ הרגעי בזמן מחזור הנשימה. הפרק השני של העבודה מתאר את המערכות ששימשו לאיסוף המידע, את הפרוטוקולים הניסויים ואת העיבוד הראשוני שבוצע לאותות.

שיטות עיבוד וניתוח האות (פרק 3)

הפרק השלישי של העבודה עוסק בשיטות החישוביות ששימשו לניתוח אותות הלב, והוא כולל בתחילתו סקירה אודות הבסיס התיאורטי של כמה מהשיטות. עיקר הפרק מוקדש לטיפול בשתי בעיות: אִשְׁכּוּל (clustering) וסיווג מורפולוגיים של קולות לב, וסגמנטציה של אותות לב.

אִשְׁכּוּל וסיווג של קולות לב

אותות ביו-רפואיים מחזוריים, ובכללם אותות לב ויברו-אקוסטיים, מתאפיינים בשונות מורפולוגית משמעותית מפעימה לפעימה, שהינה תוצאה של השונות הטבעית של התהליכים הפיזיולוגיים היוצרים אותם. מטרת פרק זה במחקר היתה לזהות ולאפיין באופן אוטומטי את המורפולוגיות החוזרניות של הסיגנל ואת הקשר הסיבתי בין המצב הפיזיולוגי לשינוי המורפולוגי באות, על-מנת לחזות פרמטרים פיזיולוגיים מתוך מאפייני קולות הלב. לשם כך, אנו מתארים שיטת ניתוח אות הכוללת קדם-עיבוד (preprocessing), ייצוג במרחב זמן-תדר ואִשכּוּל היררכי של האותות לקבלת אשכולות (clusters) המייצגים את המורפולוגיות החוזרניות העיקריות. ייצוג של כל פעימה במרחב המוגדר עלידי המרחקים למרכזי האשכולות העיקריים מאפשר לבנות מסווג (classifier) או רגרסור (regressor) אשר משמשים לבחינת הקשר שבין מורפולוגית הסיגנל לפרמטרים פיזיולוגיים. נבדקה התלות של ביצועי השיטה המוצעת בתכונות שונות של ניתוח האות, בכללן הבחירה באופן ייצוג האות הגולמי, אופן מדידת המרחק בין האותות וסוג אלגוריתם הסיווג. אותות סימולטיביים שימשו להשוואת הרגישות של שיטות ייצוג זמן-תדר שונות לשינויים בתכונות האות. בהמשך, בפרקים 4 ו-5, מופעלת ונבחנת שיטה זו לניתוח קולות לב אמיתיים ולחיזוי פרמטרים של תפקוד הלב והנשימה מתוך האותות האקוסטיים.

העבודה המתוארת בפרק זה פורסמה במאמר:

Amit, G., Gavriely, N., Intrator, N.: Clustering and Classification of Heart Sounds. *Biomedical Signal Processing and Control* 4 (2009) 26-36.

סגמנטציה של קולות לב

סגמנטציה של האות נועדה לחלק אותו למקטעים המייצגים מרכיבים נפרדים או ארועים שונים במחזור הלב. השונות של האותות הויברו-אקוסטיים מפעימה לפעימה מקשה על זיהוי אוטומטי של מרכיבי הסיגנל ומחייבת שימוש בשיטות מתקדמות על מנת לבצע החפפה (registration) של מחזורי האות וזיהוי של המרכיבים

תקציר

רקע – אותות לב ויברו-אקוסטיים (פרק 1)

הלב הוא מערכת מכנית שתפקידה להזרים דם לאברי הגוף. הלב בנוי מארבעה מדורים: שתי עליות אשר אוספות את הדם הנכנס ממערכת הווריד‍ים, ושני חדרים הדוחפים את הדם למערכת העורקים. זרימת הדם מווסתת על-ידי זוג מסתמים בין העליות לחדרים וזוג מסתמים בין החדרים לעורקים. הפעילות המכנית המחזורית של הלב כוללת התכווצות הקרוייה סיסטולה והתרפות הקרוייה דיאסטולה. מערכת הנשימה משפיעה על הפעילות המכנית הלבבית על-ידי שינויי לחצים הגורמים להבדלים בתפקוד החדר השמאלי והימני. בזמן מחזור הלב נוצרות תנודות מכניות של דפנות הלב, המסתמים והדם, אשר נקלטות על-פני בית החזה כאותות בתחום התדר של הקול והתת-קול (infrasound). הקול הראשון של הלב (S1) נשמע בזמן תחילת ההתכווצות הסיסטולית של החדרים, והקול השני (S2) נשמע בסוף הסיסטולה ותחילת הדיאסטולה. התנודות נמוכות התדר של הלב וכלי הדם נקלטות כאותות דופק (pulse) באיזור חוד הלב (apex) ובאיזור עורק הקרוטיד. אותות אלה, המכונים אותות ויברו-אקוסטיים, הינם תוצר ישיר של התהליכים המכניים של מחזור הלב, ועל-כן נושאים מידע פיזיולוגי רב-ערך אודות תהליכים אלה. למרות השימוש הרפואי רב השנים בהאזנה לקולות הלב, ניתוח כמותי של אותות אלה נדחק מהמחקר בשנים האחרונות, שהתמקד בשיטות הדמיה מתקדמות כגון אקוקרדיוגרפיה וטומוגרפיה ממוחשבת (CT). שיטות רבות ערך אלה דורשות ציוד מורכב, וכן מפעילים ומפרשים מיומנים. יתרה מכך, שיטות אלו אינן מתאימות לניטור רציף מחוץ לסביבת בית החולים.

מטרת עבודת מחקר זו היתה פיתוח שיטות עיבוד וניתוח מתקדמות להסקה של מידע פיזיולוגי מתוך אותות ויברו-אקוסטיים, והערכת הישימות של שיטות אלה לניטור בלתי-פולשני ורציף של תפקוד הלב ויחסי הגומלין בין מערכות הלב והנשימה. הפרק הראשון של העבודה כולל סקירה מקיפה של הרקע הפיזיולוגי והרפואי, המקור והתכונות של האותות, הספרות המחקרית ותיאור מטרות העבודה.

שיטות מחקריות (פרק 2)

המתודולוגיה הניסויית בעבודה זו התבססה על ביצוע הקלטות של אותות ויברו-אקוסטיים תוך כדי שינוי מבוקר של התנאים הפיזיולוגיים המשפיעים על התהליכים המכניים של מחזור הלב. במקביל, נאספו מידע ואותות נוספים (ECG, לחץ נשימה, אקו-דופלר) ששימשו כנקודות יחוס למצב הפיזיולוגי בזמן ההקלטה. דפוסי השינוי באותות הויברו-אקוסטיים, שזוהו באמצעות ניתוח ממוחשב, הושוו למידע היחוס על מנת ללמוד על היחסים בין המצב הפיזיולוגי של הלב לתבניות המורפלוגיות של האותות הויברו-אקוסטיים.

תמצית

לב האדם הוא משאבה אלקטרו-מכנית המזרימה דם לרקמות הגוף בתהליך מחזורי של כיווץ והרפייה. התהליכים המכניים במחזור הפעילות של הלב יוצרים תנודות נמוכות-תדר אשר נקלטות על פני בית החזה כאותות קוליים או תת-קוליים, הקרויים אותות ויברו-אקוסטיים. אותות אלה נושאים מידע חשוב אודות הפעולה והתפקוד הפיזיולוגים של מערכת הלב וכלי הדם, ואודות יחסי הגומלין בין מערכת הלב ומערכות גוף אחרות כגון מערכת הנשימה. בעבודת המחקר המוצגת כאן פותחו שיטות ניתוח אותות וזיהוי תבניות לצורך הסקה אוטומטית של מידע פיזיולוגי מתוך אותות לב ויברו-אקוסטיים. באמצעות שיטות אלה נבחן הקשר בין האותות הויברו-אקוסטיים לבין התפקוד המכני של הלב, והוערכה הישימות של שימוש באותות אלה לניטור בלתי-פולשני רצוף של תפקוד הלב. אנו מתארים אוסף רחב של שיטות לניתוח אותות ויברו-אקוסטיים, בהן שיטות לקדם-עיבוד (preprocessing), זיהוי אירועים ותתי-מרכיבים (סגמנטציה), יצוג על-ידי התמרות זמן-תדר, שִׁיווּר (alignment) של מחזורי האות, מיצוי מאפיינים (feature extraction) עמיד ברעש, אִשׁפּוּל (clustering) מורפולוגי וסיווג (classification). בפרט, אנו מציגים שיטת ניתוח אות לזיהוי המורפולוגיות האופייניות השונות של קולות הלב וסיווגן למצבים פיזיולוגיים. שיטות אלה שימשו לניתוח אותות לב שהוקלטו תוך כדי שינוי מבוקר של התנאים הפיזיולוגיים.

בניתוח של קולות לב שהוקלטו בזמן מבחן מאמץ פרמקולוגי אופיינה תבנית של עלייה באנרגיה הספקטרלית וברוחב הפס של הקול הראשון של הלב (S1) בזמן מאמץ. ניתוח זמן-תדר של תבנית זו שימש להגדרת מדדים אקוסטיים המתארים את הדינמיקה המורפולוגית של S1. נמצא מִתאָם גבוה בין המדדים האקוסטיים לבין מדדים אקוקרדיוגרפיים של התפקוד הסיסטולי של החדר השמאלי. בנוסף, שיטת הניתוח האוטומטי על-ידי אִשׁפּוּל מורפולוגי זיהתה את התבנית הזו והצליחה לחזות בדיוק גבוה את רמת המאמץ הלבבי מתוך המורפולוגיה של S1.

כמו כן נעשה שימוש בכלים שפותחו בעבודה זו לצורך ניתוח של קולות לב שהוקלטו תוך כדי שינויים בפעילות הנשימתית. בניתוח זה אופיינו שינויים תלויי זמן (temporal changes) ושינויי תצורה (morphological changes) הנגרמים לקולות הלב S1 ו-S2 על-ידי מערכת הנשימה. שיטה זו מאפשרת לחזות בדיוק גבוה פרמטרים אופייניים של דפוס הנשימה, דוגמת הפאזה של מחזור הנשימה, רמת ההתנגדות לנשימה ולחץ הנשימה הרגעי, מתוך המורפולוגיה של קולות הלב.

תוצאות מחקר מספקות תובנות חדשות אודות הכלים החישוביים המתאימים לעיבוד כמותי אוטומטי של אותות לב ויברו-אקוסטיים ואודות הקשרים הפיזיולוגיים בין מורפולוגית האות לבין התפקוד המכני של מערכות הלב והנשימה. שיטות חישוביות כאלו עשוית בעתיד להוות טכנולוגיה לא פולשנית חדשה לניטור קרדיו-פולמונרי רציף, ואף לשפר את יכולת הזיהוי של פגיעה בתפקוד המכני של הלב עקב מחלות של מערכות הלב, כלי הדם ודרכי הנשימה.

TEL AVIV UNIVERSITY אוניברסיטת תל-אביב
הפקולטה למדעים מדוייקים ע"ש ריימונד ובברלי סאקלר
בית הספר למדעי המחשב

ניתוח אוטומטי של אותות לב ויברו-אקוסטיים: שילוב שיטות עיבוד אות ולמידה חישובית לשם הערכה וניטור לא פולשניים של תפקוד הלב

חיבור לשם קבלת תואר
'דוקטור לפילוסופיה'
מאת

גיא עמית

עבודה זו בוצעה בהדרכת:
פרופ' נתן אינטרטור ופרופ' נעם גבריאלי

הוגש לסנאט של אוניברסיטת תל-אביב
אפריל 2009