MIT Automated Auscultation System

by

Zeeshan Hassan Syed

S.B., Computer Science and Engineering
Massachusetts Institute of Technology (2003)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2003 [illegible]

Author . . . . . .
Department of Electrical Engineering and Computer Science
May 9, 2003

Certified by . . . . . .
John V. Guttag
Professor, Computer Science and Engineering
Thesis Supervisor

Accepted by . . . . . .
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

MIT Automated Auscultation System

by

Zeeshan Hassan Syed

S.B., Computer Science and Engineering

Massachusetts Institute of Technology (2003)

Submitted to the Department of Electrical Engineering and Computer Science
on May 9, 2003, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

At every annual exam, the primary care physician uses a stethoscope to listen for cardiac abnormalities. This approach is non-invasive, inexpensive, and fast. It is also highly unreliable. Over 80% of the people referred to cardiologists as suffering from the most commonly diagnosed condition, mitral valve prolapse (MVP), do not have this condition.

Working in conjunction with cardiologists at MGH, we developed a robust, low cost, easy to use tool that can be employed to diagnose MVP in the office of primary care physicians. The system fuses signals from an electronic stethoscope and a two-lead EKG, and uses software running on a desktop or laptop computer to make a diagnosis. We also provide a number of novel audiovisual diagnostic aids. These allow physicians to visualize both individual heart beats and a visual-prototypical heart beat constructed from a sequence of beats. They also permit doctors to listen to an audio-prototypical heart-beat, audio enhanced heart-beats that amplify clinically significant sounds, and slowed down heart-beats that make it easier to separate clinically relevant cardiac events.

We tested our system on 51 patients. The number of false positives was reduced to approximately 10%. While there is no generally accepted statistic on false negatives, anecdotal experience indicates that our system also outperforms physicians in this respect.

Thesis Supervisor: John V. Guttag
Title: Professor, Computer Science and Engineering

Acknowledgments

I would like to express my sincerest gratitude to Professor John Guttag, who has been a continued source of guidance, support and inspiration throughout the course of my education at MIT. He has contributed significantly to my development, both personally and professionally, and has assisted me in recognizing and addressing my weaknesses. His technical savvy, keen insight into subtle issues and vision spanning several disciplines has allowed me to evolve into a more effective researcher. I have also learnt a great deal from numerous discussions with him on topics beyond academics and, in particular, from the various anecdotes recounting his personal experiences. I am a much improved, more well-rounded person today and I credit this all to his dedication towards his students.

I am also indebted to my colleagues at the MIT Laboratory for Computer Science, who made it a pleasure to work on this project. Dorothy Curtis was invaluable with her persistent stream of quality ideas and cannot be thanked enough for being available at all hours of the day to bounce ideas off. Michel Goraczko expedited work on this project tremendously by assisting with a vast array of technical issues. Gregory Harfst, with his cheerful disposition, made my years at LCS a thoroughly enjoyable experience and I have learnt a great degree from his all-encompassing knowledge and his unflappale personality, which were often the only guiding lights on the days before demos. I am also very grateful to him for sharing the LaTeX skills necessary to write up this thesis. Eugene Shih was of considerable help in allowing me to fulfil my responsibilities to 6.170 and I learnt a great deal from him during discussions on a number of different issues related to this project. Jason Gift and Gina Yi also assisted me tremendously with a pool of signal processing knowledge I could always draw upon.

A very special word of thanks to Ali Shoeb and Asfandyar Qureshi. I have thoroughly enjoyed working with the two of them and the memories of our many animated discussions in the hallways of Macgregor and in Room 529 at LCS are amongst my most cherished. I am very proud to have developed a deep personal friendship with

two of the most wonderful individuals I have ever met.

This project could not have been possible without the infinite knowledge of Dr. Robert Levine and the committment and enthusiasm of Dr. Francesca Nesta. They provided us with the clinical background needed for our study as well as the dataset used for refining hypothesis and evaluating our system. Both Dr. Levine and Dr. Nesta were quick to explain the physiological concepts that guided our work and to encourage progress. Their contribution was a critical ingredient of the success achieved by our work. I would also like to thank Dr. Nathaniel Sims and Dr. Rueben Mezrich , who never failed to surprise with their extensive knowledge of contemporary technologies and have played important roles in promoting the use of computers in medicine by suggesting a wide range of very interesting practical applications.

Finally, I would like to dedicate my efforts over the last four years to my parents, who have made many sacrifices to allow me to be where I am today. Without a doubt, they are the two most important people to me in the world. They have taught me to value integrity, work hard and demonstrate a concern for others. I am eternally indebted for everything that they have empowered me to achieve.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Heart auscultation, the process of interpreting the sounds produced by the heart, is a fundamental tool in the diagnosis of cardiac disease. It serves as the most commonly employed technique in primary health care and in circumstances where sophisticated medical equipment is not available (such as remote areas or developing countries). However, detecting relevant symptoms and forming a diagnosis based on sounds heard through a stethoscope is a skill that can take years to acquire and refine. Part of this difficulty stems from the fact that heart sounds are often separated from one another by less than a period of 30 ms [1]. In addition, the signals characterizing cardiac disorders typically have far less energy than normal heart sounds. This makes the task of acoustically detecting abnormal activity a challenge.

Even once the ability to perform auscultation is acquired, there is no organized way to impart it to others. The percentage of programs that incorporate structured teaching of auscultation is only 27.1 % for internal medicine and 37.1 % for cardiology [2]. This constitutes a further challenge to learning how to listen to heart sounds.

It would be advantageous if the benefits of auscultation could be obtained with a reduced learning curve, using equipment that is low-cost, robust and easy to use. The impact of the success of such an effort can be understood in light of the fact that contemporary "gold standard" tests are expensive and often unnecessary. In fact,

as many as 80 percent of patients referred to cardiologists have only benign heart murmurs or normal hearts [1, 3, 4]. These cases represent a severe inefficiency as far as medical care is concerned, since the cost of a visit to a cardiologist (including associated echocardiography) runs anywhere from $300 to $1000 in the United States [1]. Such false positives also constitute a significant waste of time for both patients and cardiologists, and also the source of much unnecessary emotional anxiety for patients and their families. In addition to this, there are also many forms of heart disease that remain asymptomatic, and thereby undetected, for several years until they eventually deteriorate into serious medical disorders.

1.2 Goal

The main goal of this project can be understood as designing, implementing and evaluating a software application system capable of performing automated auscultation. Since mitral valve prolapse (MVP) is the most common heart disorder in the industrialized world at present, and is also the most common cause of referral to cardiologists, we focus exclusively on detecting the presence of this disease in patients[1]. However, a future goal of ours is to extend use of this system to other pathologies. Consequently, we seek to develop a framework that is sufficiently non-specific to MVP to allow our software to deal with other heart diseases with minimal modification.

We intend for the application developed to be employed in the office of a primary case physician to assist him or her in determining whether patients require consultations with cardiac specialists and further examination by means of an echocardiogram. The system aims to make a recommendation and support its decision with audio-visual diagnostic aids. In this way, the software mediates between patient and doctor.

The need for insightful audio-visual diagnostic aids is driven by the requirement that the system work in a transparent manner that makes it easy to validate its

[1]This decision is motivated both by the greater impact of a system capable of diagnosing MVP correctly and by the ease of collecting data to produce such a detector

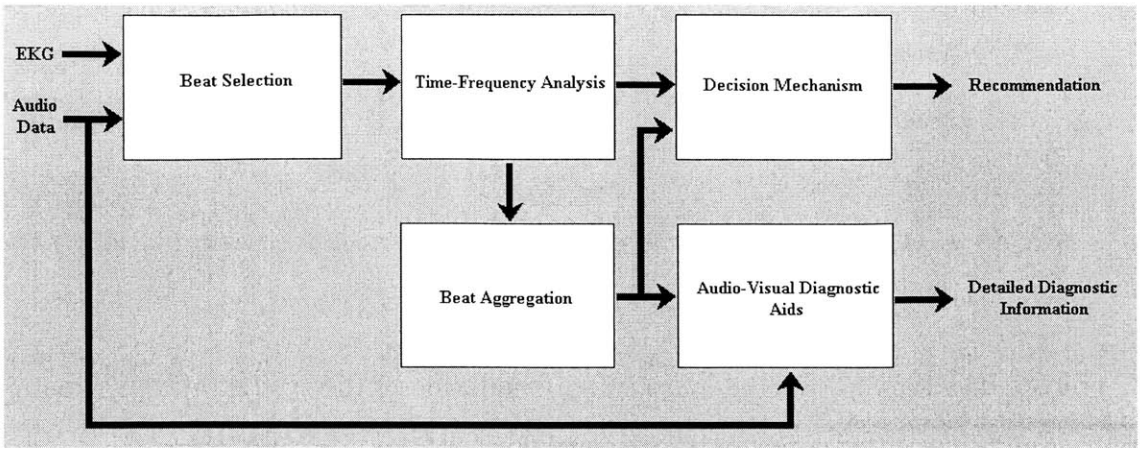


Figure 1-1: Sub-Problem Decomposition of Automated Auscultation

performance in light of physiological knowledge. By adopting a totally glass-box approach and isolating the features necessary to diagnose cardiac disorders, we hope to demystify auscultation, thus making it easier to impart this skill to others. During the course of the project we strive to find out specifically what distinguishes between MVP and non-MVP cases by extracting characteristic acoustic signals associated with the underlying anatomy of the heart. Doing so makes it easier to communicate the ability to perform auscultation to others. Moreover, this knowledge also permits the possibility of refining existing physiological models.

1.3 Problem Decomposition

The task of performing automated auscultation can be structured into five distinct sub-problems:

- *beat selection* to restrict analysis to a subset of the total beats recorded for the patient, with attention being focussed on those beats that are determined to contain most diagnostic information based on a set of medically relevant criteria (e.g., beat length) and are free of noise

- *time-frequency analysis* of heart sounds to determine physiologically significant features

- *beat aggregation* to further synthesize features by combining information across multiple heart-beats

- a *decision mechanism* that maps feature values to a recommendation that can be used to assist primary case physicians

- *audio-visual diagnostic aids* to abet validation of the system and provide information beyond a mere indication of whether the disease is present in a patient or not

Each of these sub-problems can be solved independently, with the overall solution being an integration of these parts as shown in Figure 1-1. This approach allows for modularity whereby changes can be made separately to any of the components shown based on clinical trials. This is particularly important from the perspective of extending the system to deal with other heart disorders. The only component that needs to be changed to allow for this is the decision mechanism. The bulk of the functionality developed does not need to be modified.

1.4 Organization of Thesis

In what follows, Chapter 2 presents the context necessary to understand this project better. A detailed description of mitral valve prolapse is provided and this is followed by a discussion of heart murmurs, with emphasis on murmurs due to MVP. Chapter 3 then presents an overview of the system and describes the approach adopted to address the problems of beat selection, time frequency analysis, beat aggregation and the development of a decision mechanism to generate a recommendation to output. Test results evaluating the performance of our system and physiologically relevant findings are also discussed here. This is followed by an in-depth explanation of the suite of diagnostic audio-visual aids in Chapter 4. Previous work in the area of

automated auscultation is presented next in Chapter 5. Finally Chapter 6 concludes this thesis and discusses future work.

Chapter 2

Background

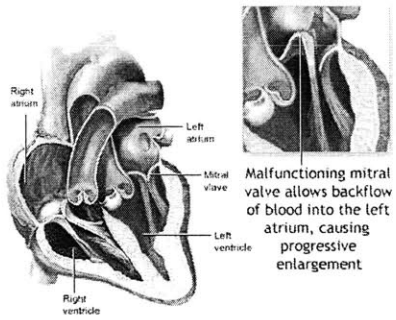
In this chapter we present the context necessary for this project. We start with a detailed discussion of mitral valve prolapse in Section 2.1. This is followed by a description of heart murmurs in Section 2.2, with emphasis on murmurs resulting from MVP.

The material presented here assumes familiarity with the physiology of the human heart. Readers without a medical background may find the information provided in Appendix A helpful in understanding the subsequent sections.

2.1 Mitral Valve Prolapse

Mitral valve prolapse or MVP [5, 6, 7, 8] is a heart disorder frequently diagnosed in healthy people and is for the most part harmless. The reduction to heart murmurs from rheumatic fever has left MVP as the most common valvular disorder in industrialized nations. Estimates indicate that anywhere between 2-5 % of the population in the United States suffers from this condition.

As mentioned in Appendix A the function of the mitral valve is to keep blood flowing in one direction through the left side of the heart by preventing the backflow of blood when the heart contracts. The mitral valve derives its name from its shape, which resembles a tall bishop's hat, called a miter. The valve is made up of two triangular shaped fibrous membrane leaflets, thin but tough, that are attached to



©A.D.A.M.

Figure 2-1: Mitral Valve Prolapse

strong parachute-like chords known as the chordae tendineae. These in turn are connected to cone-shaped muscles, the papillary muscles, that attach to the lower portion of the interior wall of the ventricles.

When the heart contracts, the two leaflets billow up to close off the opening between the upper atrium and the lower ventricle on the left side of the heart. The condition known as prolapse arises when the shape or dimensions of the leaflets of the valve are not ideal, preventing them from closing properly and leading them to balloon out. The flapping of the leaflets may result in a clicking sound. In acute cases, the prolapsing of the mitral valve may also allow a slight flow of blood back into the left atrium and this condition is called mitral regurgitation (Figure 2-1. It is this that gives rise to abnormal heart sounds called murmurs. These shall be focus of Section 2.2.

The vast majority of people with MVP have no discomfort at all. Some individuals may report mild symptoms such as shortness of breath, dizziness and either skipping or racing of the heart. More rarely, chest pain is reported. However, these symptoms may not necessarily be related to MVP and as a result it is difficult to make a diagnosis based solely on whether or not a patient exhibits such behavior. Instead, diagnosis proceeds either by means of auscultation (whereby a doctor uses a stethoscope to listen to the sounds produced by the heart) or by means of an echocardiogram as shown in Figure 2-2. An echocardiogram functions by emitting very high frequency sounds waves. These travel through the layers of skin and muscle to produce an image

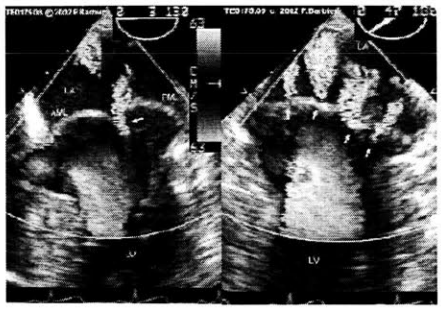


Figure 2-2: Echocardiogram for an MVP Patient

of the heart that can be seen on a screen in a manner analogous to using a radar or sonar imaging. Although an echocardiogram is the gold standard for evaluating the presence of MVP, it is relatively expensive and this makes a strong case for promoting the use of auscultation to detect MVP.

In most cases, patients diagnosed as suffering from MVP require no special treatment. However, in the case of mitral regurgitation the flow of blood back into the left atrium causes an increased risk of acquiring bacterial endocarditis (an infection in the lining of the heart). To prevent this, many physicians and dentists prescribe antibiotics before certain surgical or dental procedures. Also, patients with significant mitral regurgitation generally need to be followed more closely by their physicians. In certain cases, surgical repair or valve replacement may be necessary if the condition worsens. In addition, anti-arrhythmics (drugs which regulate the heart rhythm) may be needed to control irregular heart rhythms. Vasodilators (drugs that dilate blood vessels) also help reduce the workload of the heart and digitalis may be used to strengthen the heart-beat.

Table 2.1 details the evaluation and management of MVP disorders of increasing severity.

2.2 Heart Murmurs

When a valve is stenotic or damaged, the abnormal turbulent flow of blood produces an audible "swooshing" sound known as a *murmur* [54, 9, 10].

Risk Category	Echo Evaluation	Other Tests	Treatment
Low MVP without valvular deformity or regurgitation	Echocardiogram every 5 yr	• Initial ECG • 24-hr Holter monitor • Graded exercise stress test	• Education and reassurance • For palpitations: beta-blockers, dietary changes, and regular exercise
Mild MVP with valvular deformity and no regurgitation	Echocardiogram every 2-3 yr	• Initial ECG • 24-hr Holter monitor • Graded exercise stress test • Stress echocardiogram	• Oral antibiotic prophylaxis • Treat even mild hypertension • Encourage weight loss if needed • Treat palpitations as above
Moderate MVP with valvular deformity and mild regurgitation	Echocardiogram every 2-3 yr	• Initial ECG • 24-hr Holter monitor • Graded exercise stress test • Stress echocardiogram	• Oral antibiotic prophylaxis • Treat even mild hypertension • Encourage weight loss if needed • Treat palpitations as above
High MVP with moderate-to-severe regurgitation	Doppler echocardiogram every yr	• Initial ECG • 24-hr Holter monitor • Graded exercise stress test • Stress echocardiogram • Others based on signs and symptoms	• As above and closely monitor cardiac function and replace mitral valve when necessary

Table 2.1: Evaluation and Management of Mitral Valve Prolapse

2.2.1 General Characteristics

Timing

Murmurs are generally distinguished as systolic and diastolic by timing them against S1 and S2. Furthermore, in the case of systolic murmurs, it is possible to listen carefully to tell whether the murmur completely fills that phase of the cycle (in which case it is said to be holosystolic), or has discrete start and end points. Murmurs with discrete start and end points are classified as early, mid, or late systolic, depending upon the timing. Regurgitant murmurs, like mitral valve insufficiency, tend to fill the entire phase, while ejection murmurs, like aortic stenosis, usually have notable start and end points within that phase.

Severity

The intensity of the murmur is next, graded according to the Levine scale [11, 12]:

- I - Lowest intensity, difficult to hear even by expert listeners
- II- Low intensity, but usually audible by all listeners
- III - Medium intensity, easy to hear even by inexperienced listeners, but without an audible trill
- IV - Medium intensity with an trill
- V - Loud intensity with a trill that is audible even with the stethoscope placed on the chest with the edge of the diaphragm
- VI - Loudest intensity with a trill that is audible even with the stethoscope raised above the chest

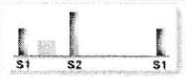
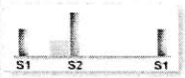
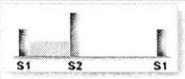
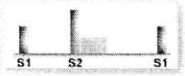
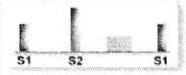
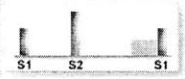
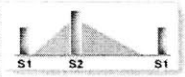
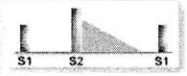
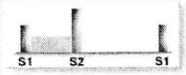
Location

A murmur may not be audible over all areas of the chest. The exact locations at which a murmur may be heard vary according to the underlying pathology. For clinical purposes, it is important to note where the murmur is heard best and what other sites it radiates to (if it radiates).

Murmur Shape and Sound Quality

Murmurs may possess many different morphologies. The shape of the murmur may be continuous (uniform/constant), a plateau (constant through systole), a crescendo (increasing), a decrescendo/diminuendo (decreasing) or a crescendo-decrescendo (diamond-shaped murmur). Common descriptive terms for sound quality include rumbling, blowing, machinery, scratchy, harsh, or musical.

Figure 2-3 illustrates different murmur patterns based on timing and morphology.

	**Early Systolic Murmur** - begins with S1 and ends before or about the middle of systole.
	**Mid Systolic Murmur** - begins after S1 and ends before S2.
	**Late Systolic Murmur** - begins at about the middle of systole and ends at the time of S2.
	**Holosystolic Murmur** - begins with S1 and ends with, or continues somewhat beyond, S2.
	**Early Diastolic Murmur** - begins with S2.
	**Mid Diastolic Murmur** - begins after S2.
	**Late Diastolic Murmur** - occurs immediately prior to S1.
	**Continuous Murmur** - has both systolic and diastolic components.
	**Crescendo** - the loudness of the murmur increases progressively. The systolic component of a patent ductus arteriosus murmur is of this type.
	**Decrescendo** - the loudness of the murmur decreases progressively. The murmurs of aortic and pulmonic regurgitation are examples of this type.
	**Crescendo-Decrescendo** - the loudness of the murmur increases and then decreases. This configuration is typical of systolic ejection murmurs.
	**Plateau** - the loudness of the murmur remains relatively constant. Holosystolic murmurs are representative of this type.

Reproduced from http://www.nurspeak.com/tools/docs/murmur_patterns.doc

Figure 2-3: Murmur Patterns

Dynamic Manoeuvres

It may be possible to hear murmurs better in certain positions than in others. Once again, this depends on the specific pathology afflicting the patient. Moreover, the audibility of murmurs may also be accentuated by inspiration and expiration depending on the nature of the murmur involved.

Clinical Significance

Murmurs may or may not be clinically significant. This follows from the fact that whereas a murmur may be caused by normal blood flow through an impaired valve, it may also be created by high flow through a normal valve. Pregnancy is a common high-volume state where these physiologic flow murmurs are often heard. Anemia and thyrotoxicosis can also cause high-flow situations where the murmur is not pathologic itself, but indicates an underlying disease process. Children also frequently have innocent murmurs which are not due to underlying structural abnormalities.

2.2.2 Murmurs due to Mitral Valve Prolapse

The murmurs associated with mitral valve prolapse are somewhat complex. Following a normal S1 and an initial briefly quiet systole, the valve suddenly prolapses, resulting in a mid-systolic click. The click is characteristic of MVP and even without a subsequent murmur, its presence alone is enough for the diagnosis. Immediately after the click, a brief crescendo-decrescendo murmur occurs, which can be seen to peak during mid to late systole. This is usually heard best at the apex. The murmur is a result of the turbulent backflow of blood towards the end of systole. As the right ventricle contracts during systole, the pressure in this chamber continues to increase. Eventually, it becomes sufficiently high to force open the damaged mitral valve, pushed blood back into the right atrium. This flow of blood through the small orifice between the flaps of the valve gives rise to a high frequency murmur just before S2.

In contrast to most other murmurs, MVP is enhanced by Valsalva manoeuvres [1] and decreased by squatting. This is because manoeuvres such as Valsalva decrease the volume of the left ventricle causing the prolapse to occur sooner. Other manoeuvres such as squatting increase venous return and diastolic filling, thereby enhancing the ventricular volume. This helps maintain tension along the chordae and keeps the valve shut. In addition to these manoeuvres, MVP murmurs are also heard better with patients lying down for reasons identical to the Valsalva manoeuvre [14].

The presence of significant mitral regurgitation often leads to a holosystolic murmur. The mitral valve in such cases is compromised to an extent that it permits backflow of blood for the entire systolic period rather than simply towards the end of systole when the pressure in the ventricles is sufficiently high to force blood back into the atria. The quality of the murmur is usually described as blowing, and is often associated with an S3 because of the left atrial volume overload. Although S1 is due to a combination of mitral and tricuspid valve closure, the mitral valve is the louder aspect. Because the valve closure in mitral regurgitation is incomplete, S1 may be noticeably quieter. Finally, in severe regurgitation, the pressure in the left ventricle quickly equalizes with venous pressure in the left atrium during the start of diastole. The result is that the aortic valve may close prematurely and may occasionally result in a widely split S2.

Heart murmurs arising from MVP are best heard at the apex as shown in Figure A-20 and radiate into the axilla. Another useful site for listening to heart murmurs is the para-sternum or the fourth intercostal space.

This information related to the timing of heart murmurs and the sites where they are most audible forms the foundation of our approach to diagnosing cardiac disorders.

[1]The Valsalva manoeuvre [13] is used with patients who have suspected heart abnormalities and is performed by attempting to forcibly exhale while keeping the mouth and nose closed.

2.3 Summary

This chapter presented the background necessary to understand our system better. The disorder of mitral valve prolapse was described in significant detail. We also discussed the different characteristics of heart murmurs, focussing specifically on murmurs due to MVP. By developing an understanding of the disorder we are trying to detect and the acoustic information it gives rise to, we hope to make it easier to follow the approach adopted by our software. This will be the subject of Chapter 3.

Chapter 3

Detection

This chapter details the approach adopted in developing algorithms to perform automated auscultation. More specifically, in what follows, we discuss our solutions to each of the subproblems proposed earlier in Section 1.3. We start out with a brief reminder of what these subproblems are in Section 3.1 and provide an overview of our system. The issue of beat selection is then addressed in Section 3.2. This is followed by an explanation of the time-frequency approach employed to analyze selected heart-beats in Section 3.3. Section 3.4 next describes the mechanism to aggregate information from multiple beats. The process of translating the features resulting from this process into a classification is subsequently detailed in Section 3.5. Finally, we conclude with an evaluation of the performance of our system.

3.1 System Overview

The goal of developing a system to perform automated auscultation can be achieved by tackling each of three problems independently. The recorded audio data and the EKG signal are first used to extract beats considered to contain the most diagnostic information. These beats are then examined using a time-frequency approach, and finally the results of this analysis are employed to produce a diagnosis.

The system was developed in MATLAB [15]. This decision was based on the exhaustive library of signal processing routines available under this platform.

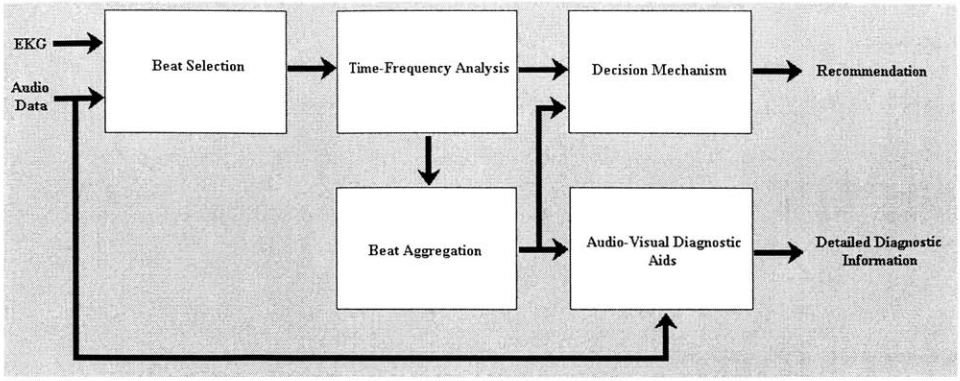


Figure 3-1: Sub-Problem Decomposition of Automated Auscultation Revisited

3.2 Beat Selection

As discussed in Section 2.2.2 mitral valve prolapse can be diagnosed by focussing on murmurs and clicks during the second half of systole. In order to make use of this information, it is first necessary to segment the recorded audio signals into individual heart-beats.

Some beats might be corrupted by noise or may fail to meet certain medically relevant criteria that can be used to separate beats with increased diagnostic information from others. We choose not to analyze these beats.

This section details the algorithms employed for both signal segmentation and beat selection.

3.2.1 Segmentation

To develop a detector for MVP, we focus almost exclusively on the presence of murmurs during systole. This results in the need to segment the recorded audio signal into systolic regions, which can be achieved by locating the QRS complex and S2 (with the interval between the two corresponding to systole).

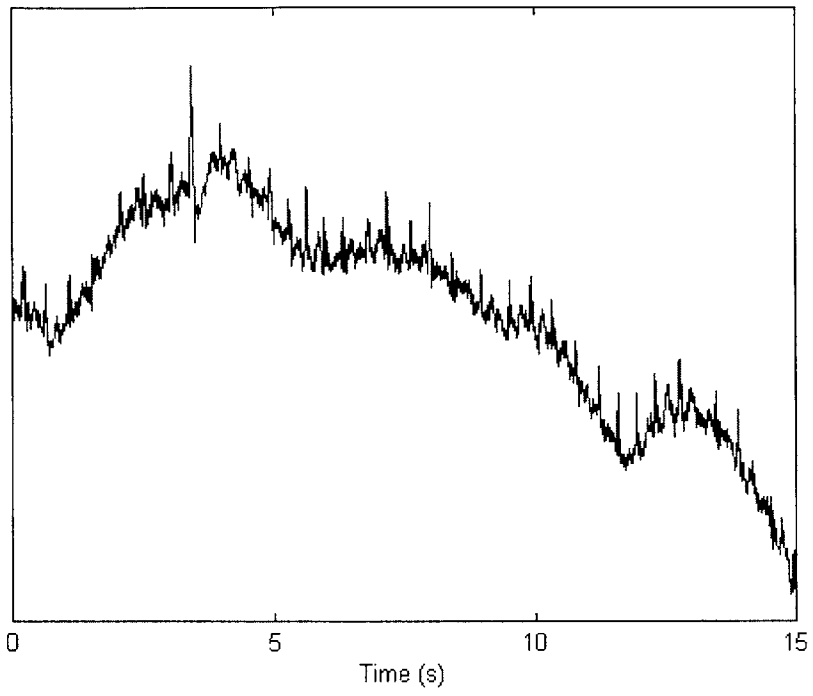


Figure 3-2: EKG Baseline Fluctuation

EKG Filtering

The algorithms that will be presented in Sections 3.2.1 and 3.2.1 to locate the QRS complex and S2 depend critically on the EKG signal recorded. This may suffer from baseline fluctuation (Figure 3-2) and high-frequency noise (Figure 3-3). In order to deal with this, we bandpass filter all EKG data prior to using it in the algorithms developed.

The baseline fluctuation is low frequency and is removed by excluding all frequencies below 1.5 Hz. The high frequency noise can be similarly dealt with by only considering frequencies in the EKG signal upto 100 Hz. There might still be noise in the 1.5 - 100 Hz range, but the signal-to-noise ratio in this region is generally quite good since the physiological activity overshadows the noise. Hence, no additional attempts are made to deal with noise in the passband.

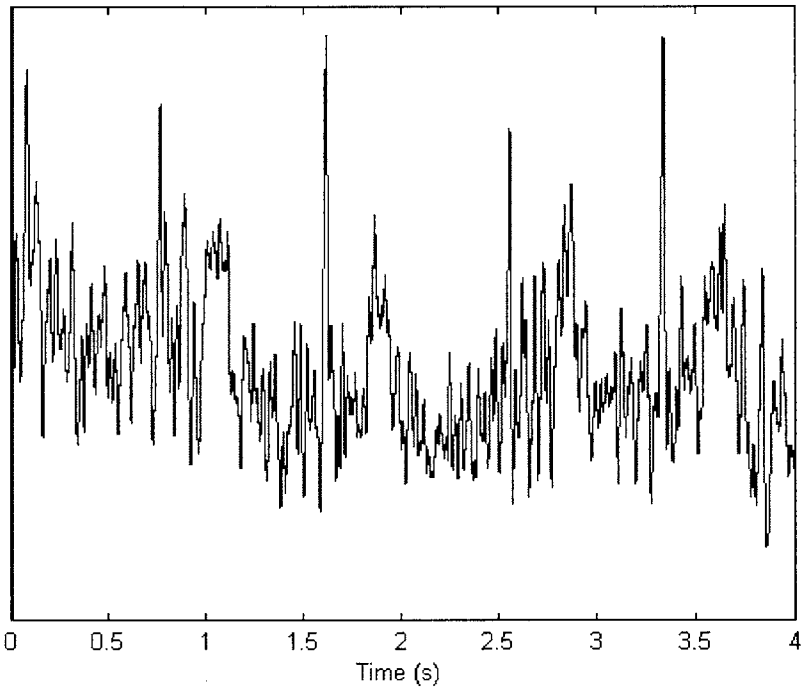


Figure 3-3: High-frequency Noise in EKG

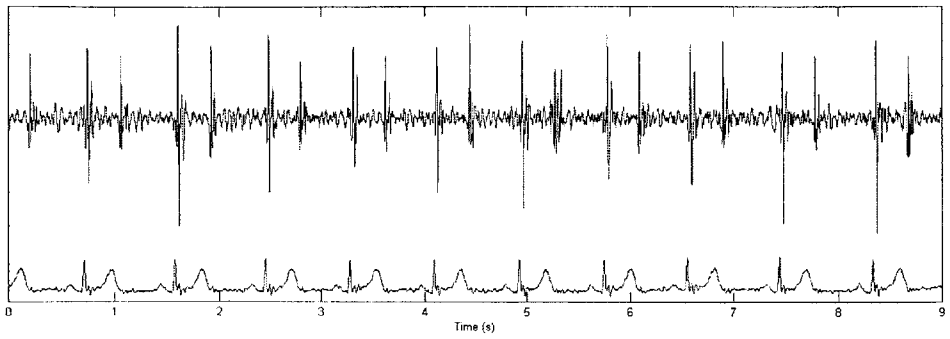


Figure 3-4: Simultaneously Recorded Audio and EKG Signals

An FIR approximation to the infinite impulse response of an ideal filter [20] is used to filter the EKG signal. This employs a Hamming window of length 50001.

QRS Detection

The beginning of systole is marked by the first heart sound (S1). This is preceded by the QRS complex as shown in Figure 3-4, which plots both the EKG signal and the simultaneously recorded audio signal.

Hence, the problem of detecting the onset of systole can be mapped to the task of detecting the QRS complexes in the audio signal[1]. There are many well-known approaches to doing so [16]. We make use of a modified version of the algorithm proposed originally by Fraden et al [17], which has empirically proven to be robust in the face of electromyographic noise and powerline interference [16].

Let W be a one-dimensional array of sample points of the synthesized, digitized

[1]It is important to note that the QRS complex and S1 are two different physiological events. More specifically, the QRS complex occurs slightly before S1. Hence, in order to precisely locate S1, one would need to search for a peak in the audio data immediately after the QRS complex. However, since the information relevant to diagnosing MVP is to be found exclusively in the last half of systole, we do not make a distinction between the QRS complex and S1. The separation between these two events is not considered to be significant and instead, for the remaining discussion, we choose to use the QRS complex as an approximate indication of the beginning of systole

EKG. An amplitude threshold is calculated as a fraction of the peak value of the EKG signal:

$$\text{amplitude threshold} = 0.4 \max [W]$$

The scaling factor of 0.4 corresponds to the optimal value for this parameter determined experimentally in [17].

The raw data is then rectified:

$$W0(n) = W(n) \text{ if } W(n) \geq 0$$
$$W0(n) = -W(n) \text{ if } W(n) < 0$$

Following this, the rectified ECG is passed through a low-level clipper. If W0(n) is greater than or equal to the amplitude threshold:

$$W1(n) = W0(n)$$

Otherwise, if W0(n) is less than the amplitude threshold:

$$W1(n) = \text{amplitude threshold}$$

The first derivative is then calculated at each point of the clipped, rectified array as follows:

$$W2(n) = W1(n + 1) - W1(n - 1)$$

Finally, a QRS candidate is declared at every point where W2(n) exceeds the fixed constant threshold:

$$W2(n) \geq 0.33$$

The value of 0.33 was chosen empirically for our dataset.

The presence of noise in the EKG signal may give rise to multiple candidates in proximity to actual QRS complexes. As a first attempt, these can be removed by adding an extra step at the end of the algorithm where all QRS candidates are

discarded except those corresponding to a local peak in the underlying EKG signal. More specifically, if a QRS candidate corresponds to the peak value of the EKG signal over a window of 100 ms centered at the position of the QRS candidate, it is retained. All other QRS candidates are ignored. An alternative approach developed by us is to search over a window of 100 ms centered at each QRS candidate for the peak value in the EKG signal and keep a track of all the peaks found. Close by QRS candidates would map to the same peak and the positions of the final set of peaks could be returned as the locations of the QRS complexes. In our system we make use of the second approach because of its increased tolerance for synchronization errors. Whereas the first approach would remove all QRS candidates unless they exactly corresponded to the QRS complexes, the second approach tolerates minor deviations and only uses the QRS candidates as indicators of a close by QRS complex, which can then be pinned down by examining the audio signal.

S2 Detection

The problem of locating S2 is considerably more challenging and far less studied than the task of finding the beginning of systole. The major challenge is that the EKG does not provide us with any clear indication of the onset of S2.

As mentioned in Section A.4.2 the T wave precedes S2, but the lag between the T wave and S2 is more variable and longer than the delay between the QRS complex and S1. Another way of looking at this is to realize S1 occurs at the end of the QRS complex and S2 takes place at the end of the T wave. Whereas the QRS complex is a narrow spike and has an end that is close to its peak value in amplitude, there is no clear indication of the end of the T wave since it can be arbitrarily wide. Therefore, locating the T wave does not help significantly to focus the search for S2.

One approach for addressing this issue is to make use of the physiological information that the length of systole is normally 300 ms long [50]. However, as shown in Figure 3-5 the duration of systole varies drastically between patients. Since we are concerned with the timing of events during the last half of systole this is critical. For example, if the duration of systole is less than 300 ms, we may accidentally treat the

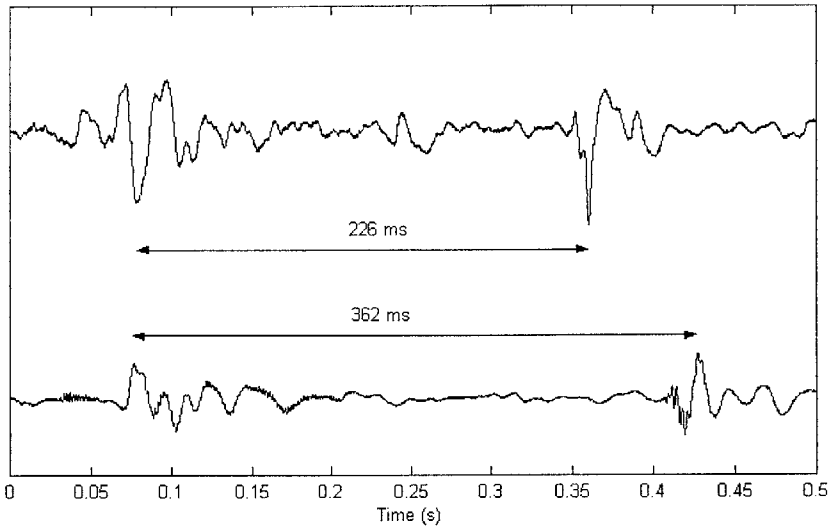


Figure 3-5: Variation in Systolic Duration for Two Patients in Our Dataset

increased energy due to S2 as a heart murmur. Similarly, if the duration of systole is greater than 300 ms, we may miss late systolic murmurs by incorrectly thinking of them as being diastolic.

Hence, we need to develop some mechanism to isolate the position of S2 on a per patient basis. Our approach is based on using both the EKG data and a recorded audio stream to yield a robust means of locating S2.

The algorithm that we developed for this purpose starts off with knowledge of the positions of the QRS complexes. Denoting the location of the i-th QRS complex in time by q_i and the one-dimensional array of points of the digitized EKG by W, we define:

$$\text{beginpt}_i = q_i + 60 \text{ ms}$$
$$\text{endpt}_i = q_i + \tfrac{2}{3}(q_{i+1} - q_i)$$

The candidate T wave corresponding to the i-th QRS complex, t_i is then declared to be:

$$t_i = \text{maxpos}\ [W(\text{beginpt}_i : \text{endpt}_i)]$$

In other words, the T wave corresponding to the i-th QRS complex is declared to be at the position of the peak in the EKG signal between the times beginpt_i and endpt_i. The decision to search over this interval follows from the physiological knowledge that the peak in the T wave occurs at least 60 ms after the QRS complex and is normally within the two thirds of the cardiac cycle immediately following QRS.

Using this information, and denoting the corresponding simultaneously recorded audio signal by X, we declare the i-th S2 candidate, s_i, to be:

$$s_i = \text{maxpos}\ [X(t_i : t_i + 150\ \text{ms})]$$

Hence, S2 is declared to be at the position of the peak in the audio signal between the candidate T wave and a period of 150 ms after it. Once again, the choice of 150 ms is motivated by the clinical information that S2 lies within this period following the peak of the T wave.

It is important to point out here that although the location of S2 can be determined independently for each systole in this manner, such an approach would lead to the algorithm being susceptible to acoustic noise. Artifacts such as the movement of the stethoscope during systole while recording and background noise may lead to peaks in the audio signal in addition to S2 as shown in Figure 3-6. These peaks may have amplitudes greater than that of S2, resulting in an incorrect determination of when the second heart sound occurs.

To remedy this problem we make use of the fact that although the length of systole may vary drastically between patients, it is relatively uniform for heart-beats belonging to any particular patient. Hence, we focus on calculating the median systolic length for each patient and then use this value to approximately predict the position of S2 following each QRS complex. More specifically, using the values of q_i and s_i obtained earlier, we calculate for all possible values of i:

$$\text{qslength}_i = s_i - q_i$$

Then:

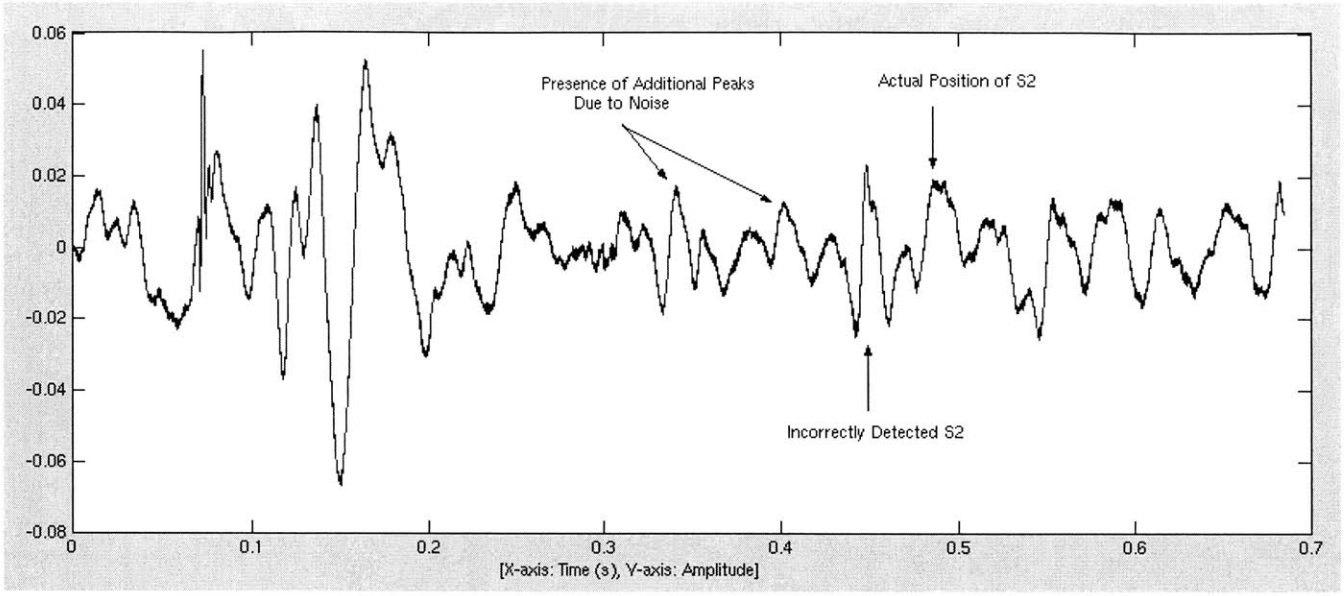


Figure 3-6: Beat with Additional Peaks Before S2 due to Artifacts

$$\text{medlength} = \text{median value of qslength}$$

And for all i, the position of the i-th S2 using the median systolic length, meds_i, is declared to be:

$$\text{meds}_i = \text{q}_i + \text{medlength}$$

Although the actual i-th S2 may vary slightly in position from med_i, the imprecision is insignificant given the various difficulties discussed earlier associated with the task of accurately locating the second heart sound.

3.2.2 Length-biased Admission of Noise-Free Beats

The algorithms detailed in Section 3.2.1 allow us to isolate the systolic portions of the audio signal for further analysis as illustrated on the left of Figure 3-7. Some of these segments may, however, be corrupted by noise. In addition, others may fail to meet

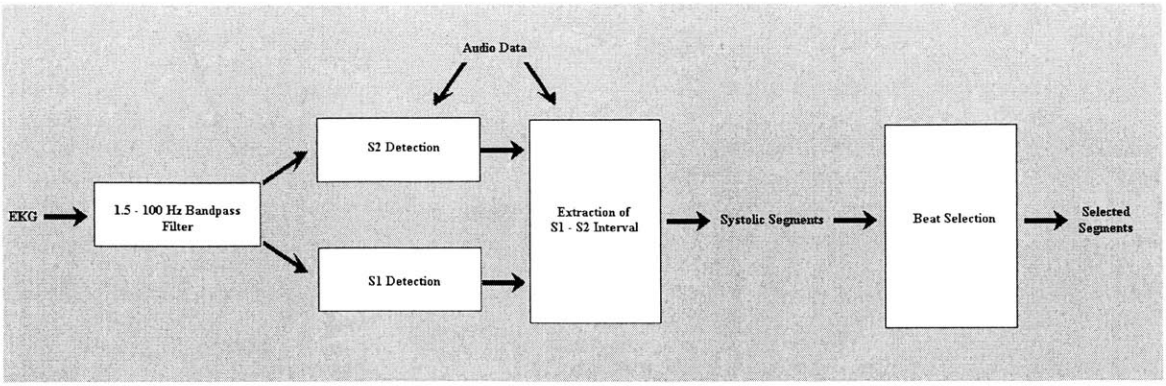


Figure 3-7: Block Diagram of System to Extract Systolic Segments with Beat Selection

the biological criteria indicative of the presence of diagnostic information. For this purpose, it is necessary to add a screening step prior to the analysis stage (as shown on the right in Figure 3-7) that determines whether or not a particular heart-beat should be examined further.

To detect the presence of noise in a beat we make use of the information that cardiac events during the middle of systole typically do not have significantly greater energy than the first and second heart sounds. Beats with peak amplitudes during the middle half of systole that are more than the amplitude of both heart sounds are consequently declared to be noisy. The increase in energy is attributed to the presence of artifacts and the beats are discarded. Figure 3-8 gives an example of a beat labelled as being noisy.

The algorithm to determine whether the i-th recorded beat is noisy starts off with the positions of the QRS complex and S2 corresponding to that beat. Denoting these by q_i and s_i respectively, the length of the systolic portion for that beat can be calculated as follows:

$$\text{systoliclength} = s_i - q_i$$

Using this value and letting X be a one-dimensional array corresponding to the

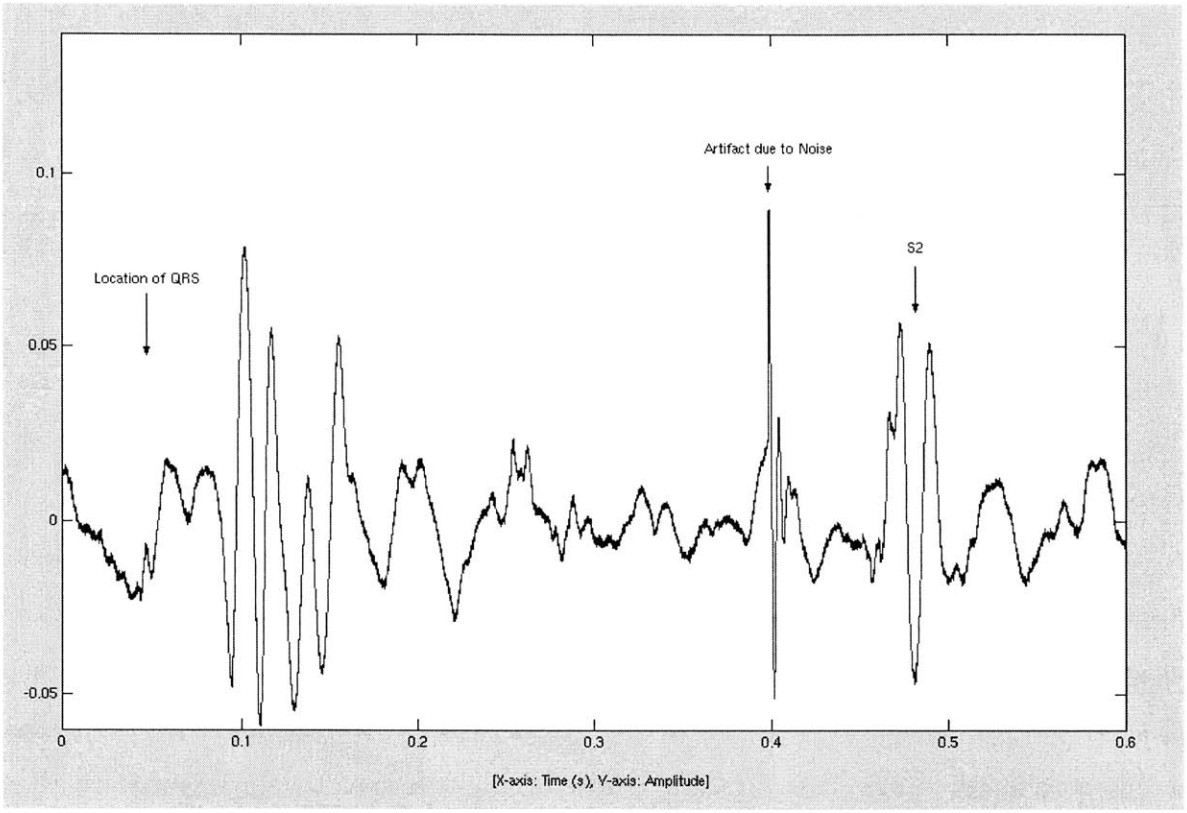


Figure 3-8: Noisy Beat Discarded by our System

recorded audio signal, the peak amplitude of the beat during the middle half of systole is defined as:

$$\text{maxpeakamp}_i = \max[X(q_i + \tfrac{\text{systoliclength}}{4} : s_i - \tfrac{\text{systoliclegth}}{4})]$$

The amplitude of the first heart sound (which lies within a period of 100 ms after the QRS complex) is then found by:

$$\text{s1amp}_i = \max[X(q_i : q_i + 100 \text{ ms})]$$

Since s_i is roughly chosen to correspond to the peak of S2, the amplitude of the second heart sound can be determined by conducting a localized search over a window of 50 ms centered at s_i[2]:

$$\text{s2amp}_i = \max[X(s_i - 25 \text{ ms} : s_i + 25 \text{ ms})]$$

If maxpeakamp_i is more than the value of both s1amp_i and s2amp_i, the beat is discarded.

In addition to removing noisy beats, we also focus on selecting beats according to the length of their R-R interval[3]. Systolic segments that are preceded by a long diastole are preferred to those following shorter diastolic periods. The reason for this is that a longer diastole allows the ventricles more time to fill with blood. This leads to an increase in the volume of blood passing through the heart during systole, which in turn produces a more audible murmur in the presence of MVP [18, 19].

The algorithm to achieve length biased beat selection starts by calculating the median length for the R-R intervals. Letting W(i) be the length of the R-R interval associated with the i-th beat we define:

[2]This value was chosen based on the observation that the maximum shift of S2, for any beat, from the predicted position of S2 (based on the median systolic length for the patient as described in Section 3.2.1) is less than 50 ms.

[3]The R-R interval for any beat is defined as the distance from the R wave of the previous beat to the R wave of the current one. Since the duration of systole does not vary significantly within a patient, the approach of picking beats with longer R-R intervals corresponds to selecting beats that have longer diastoles before them

$$W_{median} = \text{median } [W(i)]$$
$$W_{deviation} = \text{standard dev } [W(i)]$$

Following this, we set upper and lower thresholds that control which beats are admitted:

$$W_{upper} = W_{median} + 0.3\ W_{deviation}$$
$$W_{lower} = W_{median} - 0.3\ W_{deviation}$$

All beats within the range W_{lower} - W_{upper} are further examined as will be discussed in Section 3.3.1. It is important to point out here that in order to increase confidence in the output recommendation it is necessary to analyze several beats. For our purposes, we chose to select at least 20 beats for every 30 seconds of recording time. Since the initial range W_{lower} - W_{upper} usually contains fewer beats, we define the following procedure for widening the range of admission:

$$W_{upper}' = W_{upper} + 0.05\ W_{deviation}$$
$$W_{lower}' = W_{lower} - 0.025\ W_{deviation}$$

Where W_{upper}' is the updated value of the upper threshold and W_{lower}' is the update value of the lower threshold given original threshold values of W_{upper} and W_{lower}. It is in this step of the algorithm that the bias in favor of longer beats is introduced. In particular, we widen the range of admission by 0.05 times the standard deviation upwards but by only half that factor in the reverse direction.

Although the use of an upper threshold is not necessary given our attempt to bias our algorithm in favor of selecting longer beats, we nevertheless choose to ignore the longest beats in the signal if they differ significantly from the median value of the R-R interval length. In such cases these beats are considered to be outliers or segmentation errors[4], and are therefore not examined further.

[4]Such as missing an entire beat in the middle due to a weak QRS or the presence of noise in EKG

3.3 Time-Frequency Analysis

As mentioned in Section 2.2.2, the characteristic signature of mitral valve prolapse is the presence of energy at higher frequencies during the last half of systole. Hence, both the frequency content and the timing of events in the cardiac cycle are important, and the problem is well-suited for time-frequency analysis.

3.3.1 Filter Bank

In order to achieve a time-frequency decomposition of heart sounds we make use of a filter bank that separates the audio signal into its constituent frequency band components.

We considered using a joint Fourier transform to derive the frequency content of the heart-beats as a function of defined time intervals within the cardiac cycle. However, the joint Fourier transform is essentially a smoothing function in time. It requires an important portion of the systolic waveform (up to one-third) to analyze frequency content and therefore blurs the rapidly-changing events recorded from the heart. Thus, a prominent first heart sound can potentially produce the appearance of an early-systolic murmur. Similarly, late-systolic MVP murmurs may also appear to be less clearly late-systolic. An alternative analysis was therefore necessary to maintain high temporal fidelity in order to detect changes in energy at higher frequencies over time.

To accomplish this, we use a filter bank that analyzes each heart-beat using a series of sharp frequency filters to eliminate overlap between the bands. In all, 16 bands are used, each spanning a 50 Hz interval from 50 to 850 Hz. Since no windowing in time is necessary, this method provides a sharp delineation of the timing of cardiac events.

In more detail, the filter bank consists of 16 bands, each corresponding to an FIR approximation to the infinite impulse response of an ideal filter [20]. The length of the Hamming windows used is 50001. As in Section 3.2.1, this leads to filters with peak approximation error of -53 dB and transition bands with a width of around 3.5

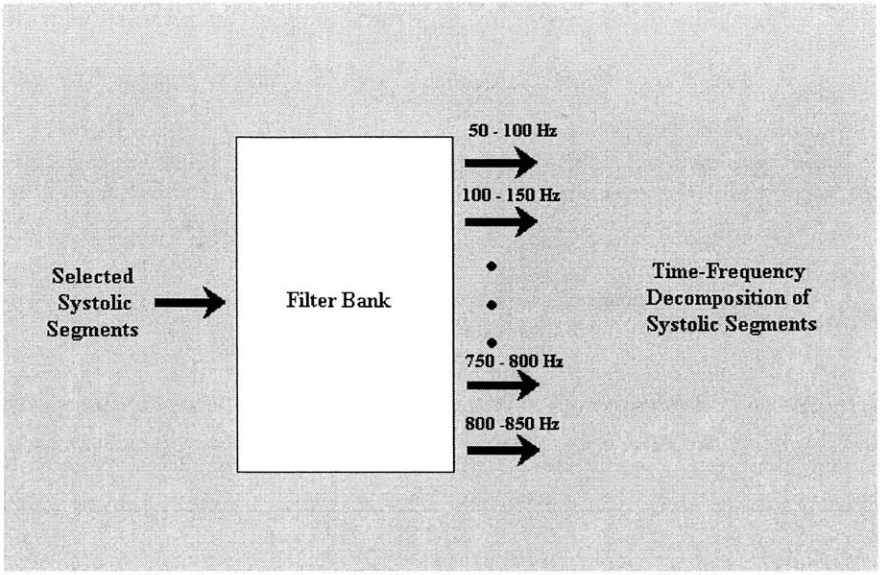


Figure 3-9: Block Diagram Representation of Filter Bank

Hz. Figure 3-9 shows the block diagram for the filter bank.

It is important to point out here that the width of the transition band for each of the filters in the filter bank should be significantly lower than the total width of each frequency band. In other words, the transition from passband to stopband should be sufficiently sharp to prevent the energy content of any frequency band from being colored by that of adjacent bands. This is crucial since the amplitude of low frequency energy in the recorded signal is generally several orders of magnitude greater than the amplitude of high frequency energy (Figure 3-10). Hence, a too wide transition band (i.e., one that does not significantly attenuate the low frequency content of the signal received from neighboring bands) would lead to low frequency energy swamping the energy at higher frequencies. This would conceal trends at higher frequencies and reduce the visibility of the high frequency murmur signature.

Figures 3-11 and 3-12 illustrate this effect for a patient suffering from moderate mitral valve prolapse with late systolic regurgitation. Figure 3-11 displays the time-frequency decomposition achieved for a filter bank using Hamming windows of length

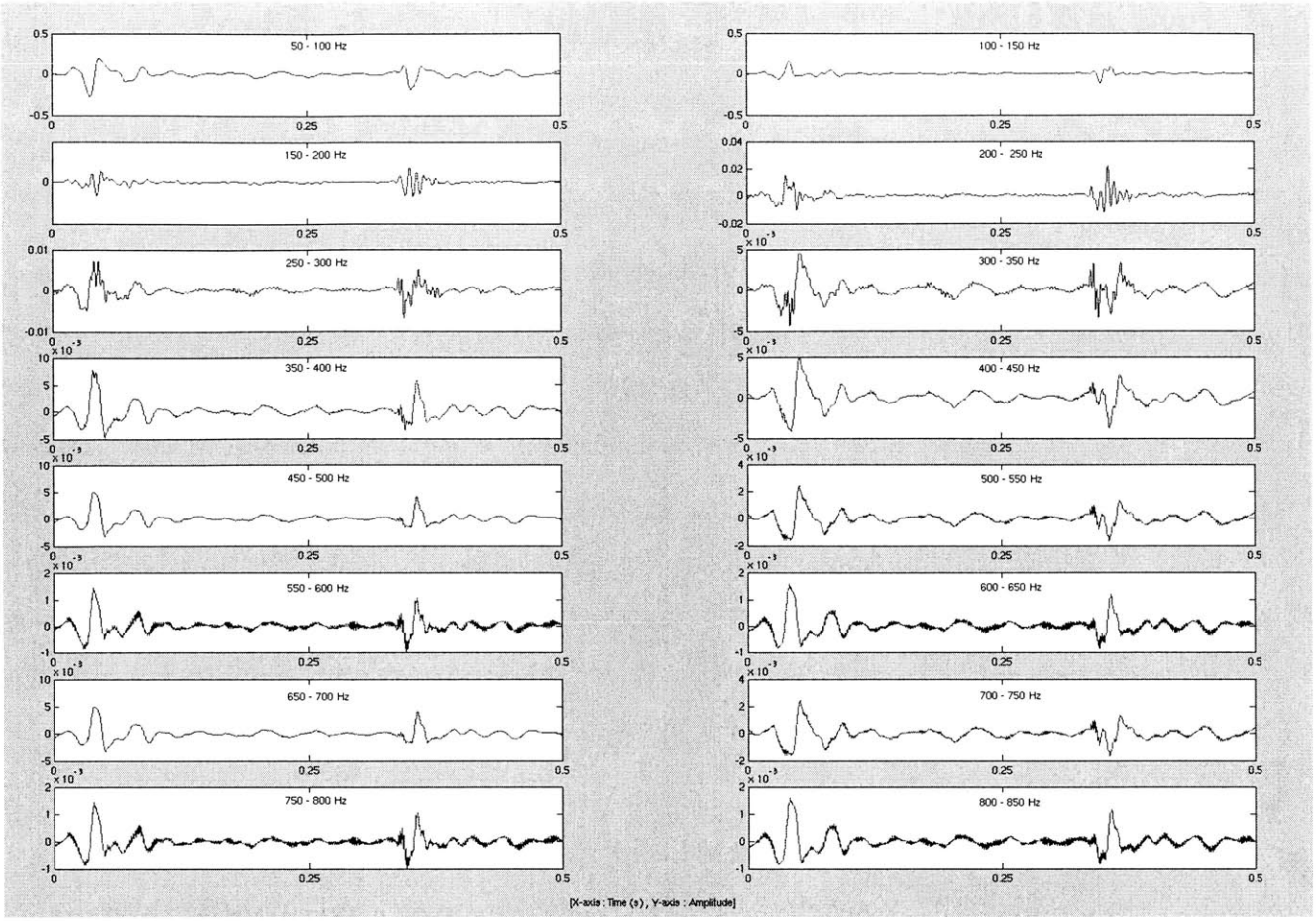


Figure 3-10: Variation in Amplitude of Different Frequency Bands for Filter Bank

50001 (corresponding to a transition band of width 3.5 Hz for each filter). The presence of high frequency energy just prior to S2 can be readily discerned from the plots, and it is easy to recognize that the patient suffers from MVP. This is not the case in Figure 3-12, which shows the time-frequency decomposition for a filter bank employing Hamming windows of length 1001 (transition band width of 176.2 Hz). At lower frequencies there is not much difference between the two figures. However, in Figure 3-12 energy at lower frequencies suppresses information in the higher frequency energy bands.

3.3.2 Band Aggregation

The filter bank shown in Figure 3-9 divides the signal into 16 bands, each spanning 50 Hz. We could instead have decided upon a smaller number of wider bands, or

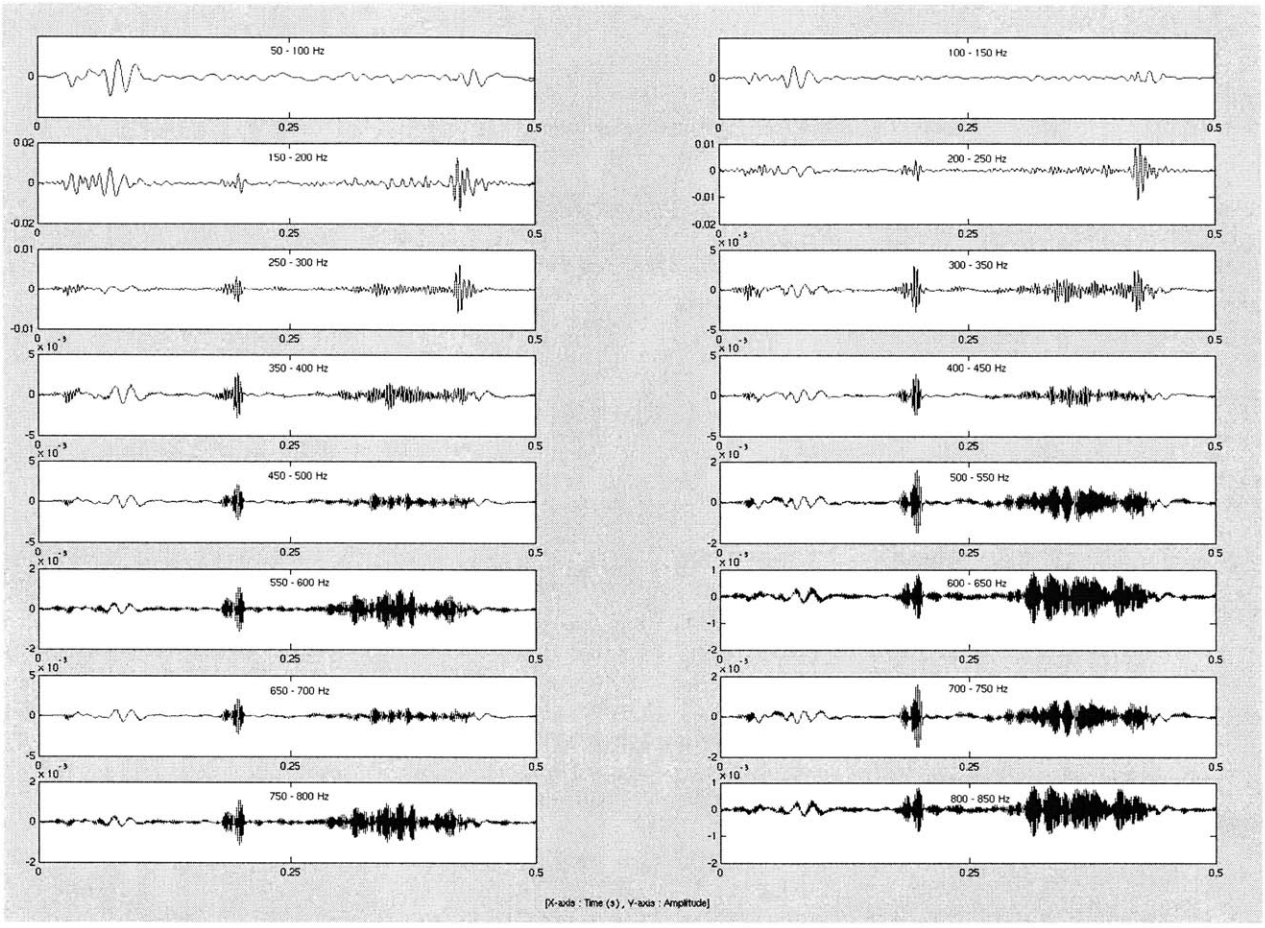


Figure 3-11: Time-Frequency Decomposition Corresponding to a Filter Bank Using Hamming Windows of Length 50001 (Transition Band Width of 3.5 Hz)

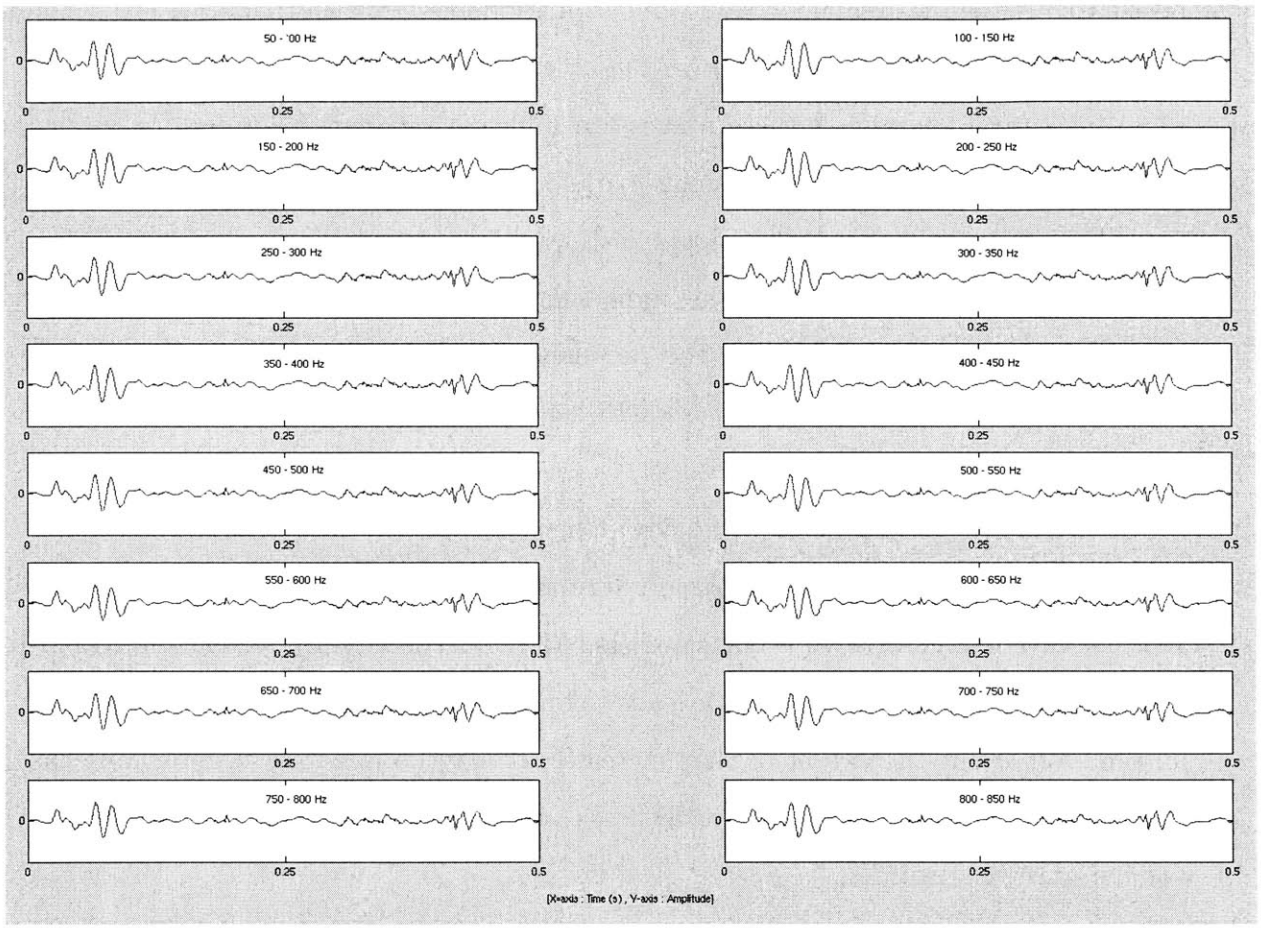


Figure 3-12: Time-Frequency Decomposition Corresponding to a Filter Bank Using Hamming Windows of Length 1001 (Transition Band Width of 176.2 Hz)

more bands, each spanning a decreased interval. The subsequent discussion justifies the choice made. It first presents the case for dividing the signal into many bands. This is followed by counter-arguments in favor of increasing the range of frequencies spanned by each band.

The decision to divide the audio signal into 16 different frequency bands (as opposed to just a low frequency band and a high frequency band) follows from the fact that heart murmurs may differ in the range of frequencies over which they lead to an increase in energy. If these ranges are sufficiently narrow, aggregating all the frequency bands together into a single band washes out useful diagnostic information. This effect is illustrated in Figures 3-13 and 3-14, which once again display the time-frequency decomposition of the audio signal corresponding to patient shown in Figures 3-11 and 3-12. Although the presence of high frequency energy in systole prior to S2 is clearly visible in Figure 3-13, aggregating the higher frequency bands obscures it in Figure 3-14.

On a closely related note, by maintaining a sufficiently fine granularity in frequency we hope to be able to provide information beyond a mere determination of whether or not mitral valve prolapse exists. More specifically, by being able to pinpoint the frequencies at which MVP leads to an increase in energy (and by observing information about the extent of that increase), it may be possible to determine the exact size of the opening between the leaflets of the mitral valve and the corresponding volume of regurgitation.

The motivation for band aggregation stems largely from the practical consideration of robust performance in the presence of high frequency noise. Empirically, this noise appears to be localized in a narrow range of frequencies and seems to corrupt only one or (in the extreme) two of the bands output by the filter bank[5]. Under such circumstances, a strong case exists for merging multiple bands together.

[5]It is important to point out that for a small number of files in the dataset there was noise at all frequencies, possibly due to an extremely noisy environment or the audio sensor being incorrectly positioned on the chest. Band aggregation is of limited use in such a scenario. Instead, the technique is intended as a means of dealing with files that are only moderately corrupted by noise in some bands and still possess sufficient information at other frequencies for a correct diagnosis to be made

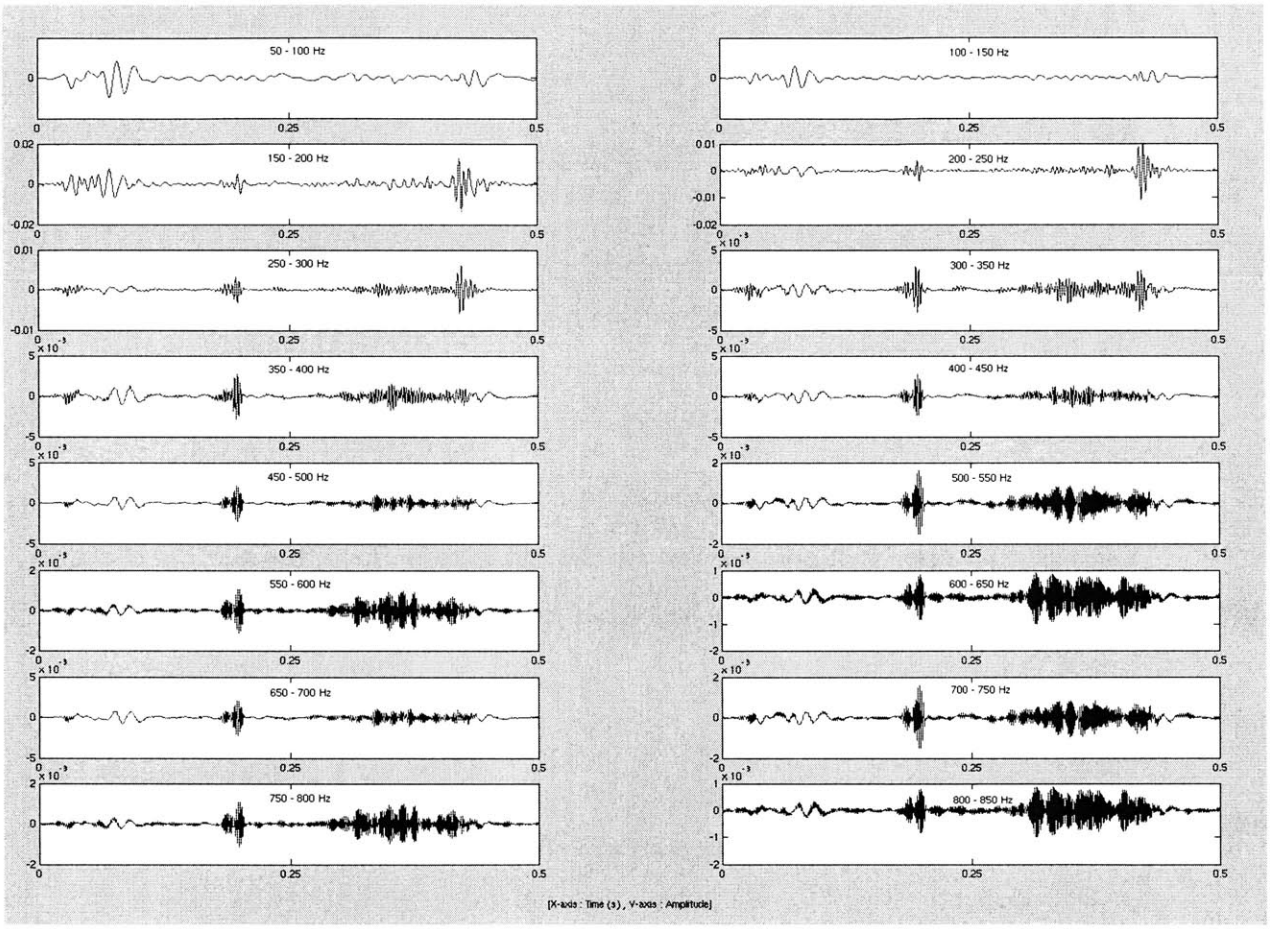


Figure 3-13: Unaggregated Output of Filter Bank (16 Output Bands)

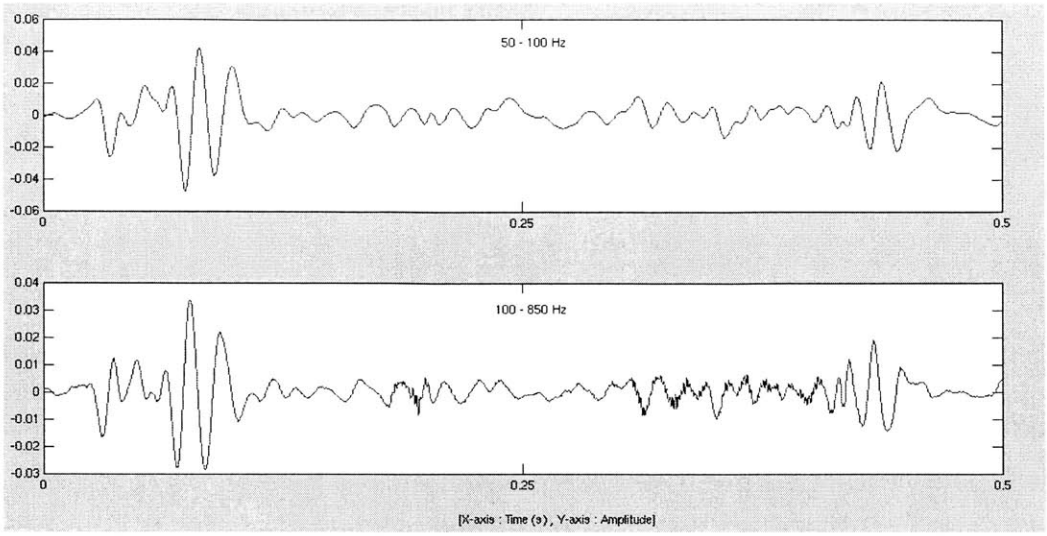


Figure 3-14: Output of Filter Bank with Aggregation of all Frequencies Above 100 Hz into One Band (2 Output Bands)

The eventual approach adopted by our system is to perform limited band aggregation. In particular, we create the following four composite bands (Figure 3-15):

- 50 - 150 Hz
- 150 - 350 Hz
- 350 - 550 Hz
- 550 - 850 Hz[6].

Also, since the energy in the low frequency bands is generally several orders of magnitude greater than the energy in bands at higher frequencies (Section 3.3.1) the process of aggregation must perform some normalization prior to combining different

[6]Noise was generally more prevalent at frequencies above 600 Hz (possibly due to high frequency pulmanory sounds). As a result, a greater degree of aggregation was performed for high frequency bands

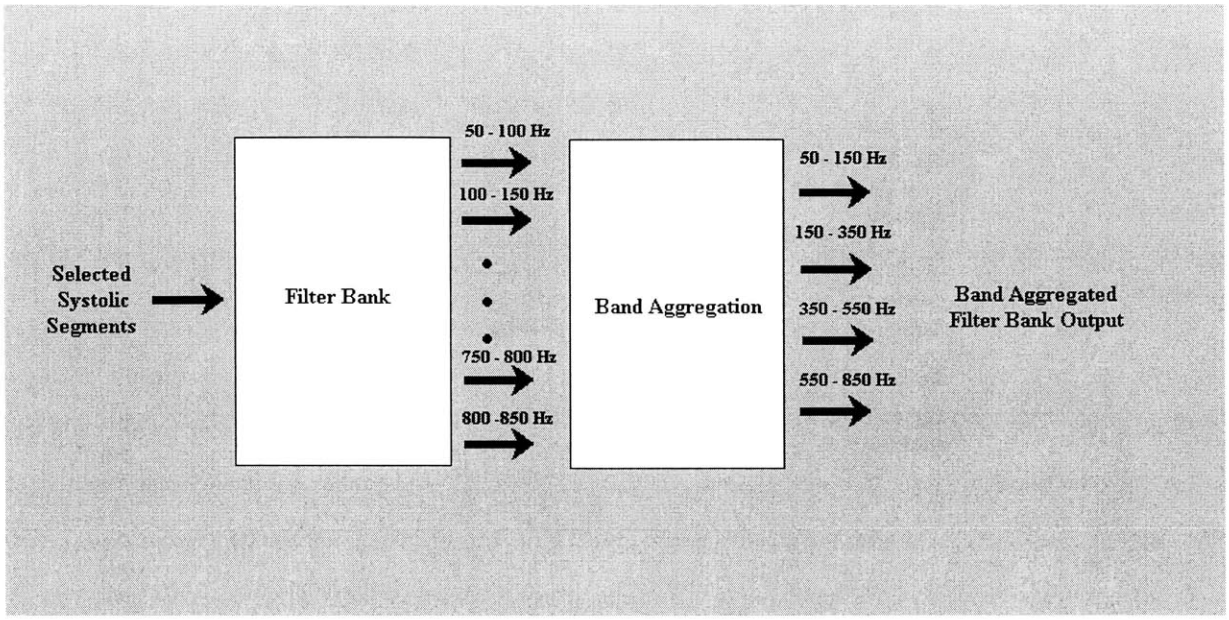


Figure 3-15: Block Diagram Representation of Filter Bank with Band Aggregation

bands together. Failure to do so will result in the lower frequencies heavily dominating the higher ones when the composite bands are formed[7].

In order to address this issue, we use an aggregation algorithm that scales bands by the reciprocal of their maximum value prior to combination. Figure 3-16 shows the effect of this approach when used to aggregate all frequency bands above 100 Hz into a single band for the patient shown in Figure 3-14. As expected, this method works much better than the non-normalized aggregation illustrated in Figure 3-14, i.e., there is considerably less loss of diagnostic information here.

The workings of the fair aggregation algorithm can be illustrated by means of an example of how the process of creating a composite band 150 - 350 Hz is performed.

[7]It is for this reason that we prefer to first divide the signal into 16 bands and then aggregate these following normalization rather than using a filter bank that directly outputs the four bands required. Also, since none of the murmurs in our dataset displayed a localized increase in energy that necessitated finer granularity than 50 Hz, we did not see the need to perform the additional computation of dividing the signal up into more than 16 bands and then aggregating these together.

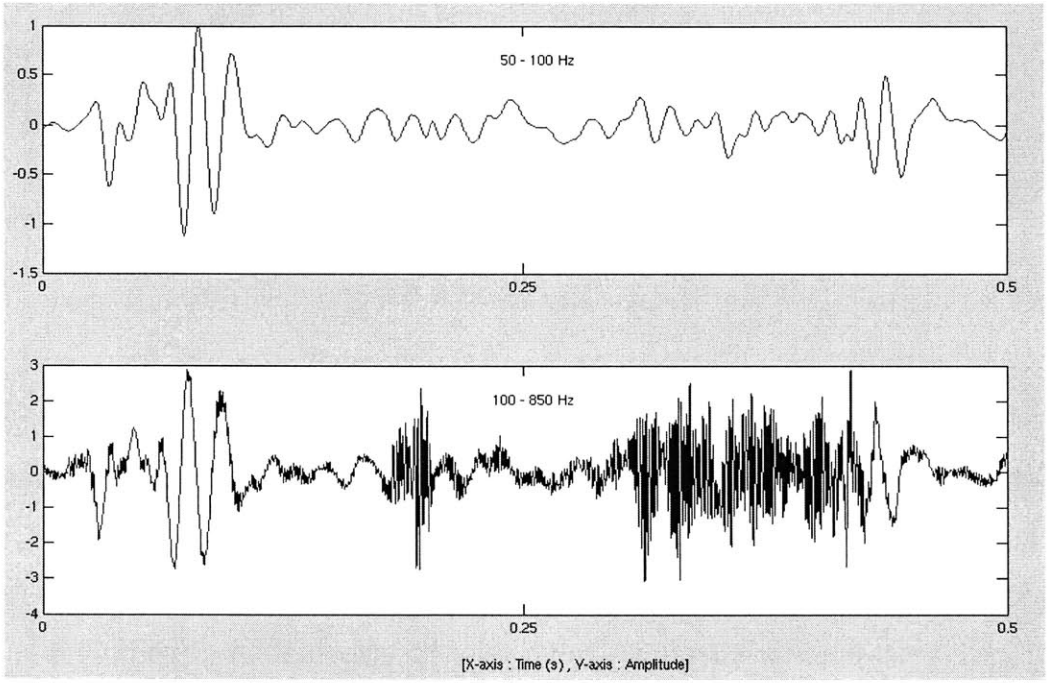


Figure 3-16: Output of Filter Bank with Post-Normalization Aggregation of Frequencies Above 100 Hz into One Band (2 Output Bands)

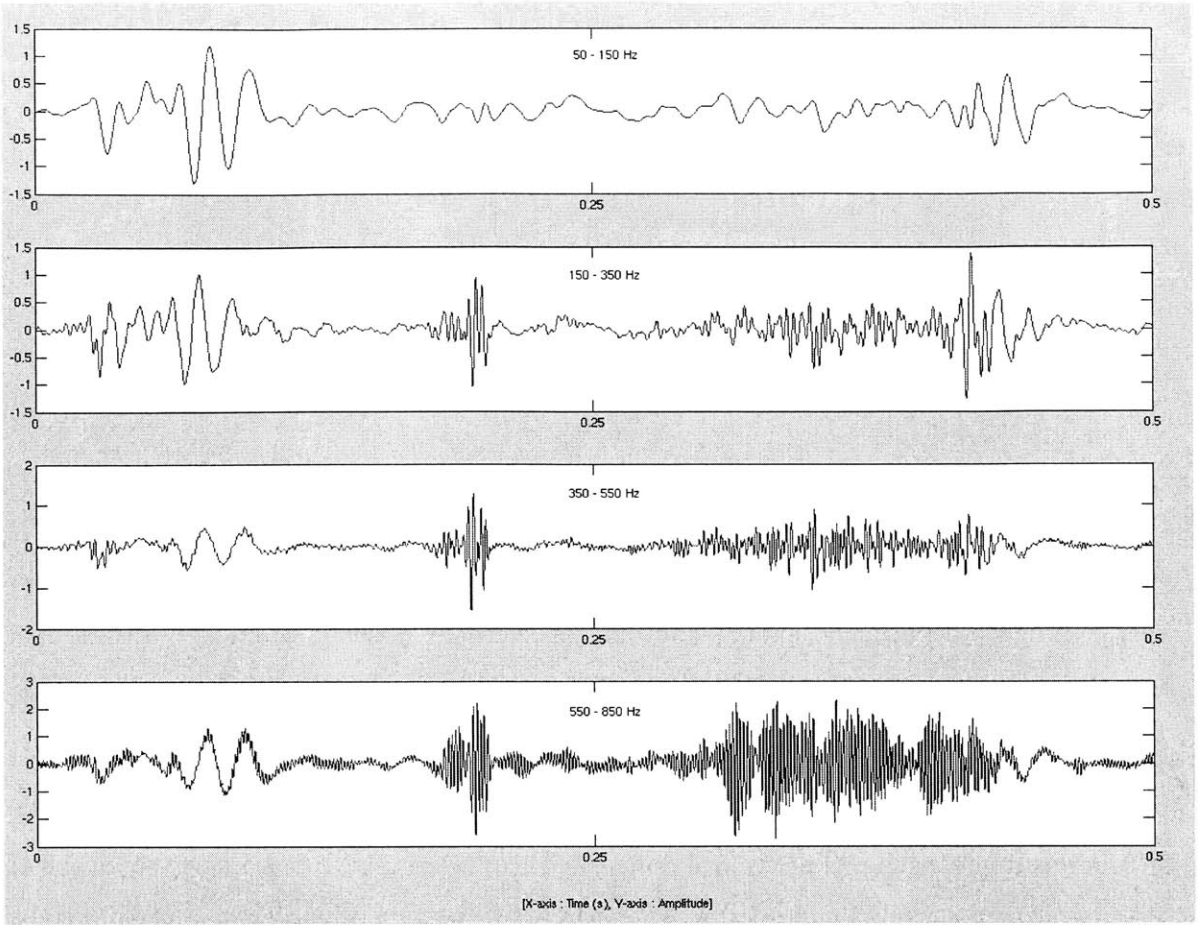


Figure 3-17: Output of Filter Bank with Post-Normalization Aggregation of Frequencies into Four Different Bands (4 Output Bands)

Defining $X_{150}(n)$, $X_{200}(n)$, $X_{250}(n)$ and $X_{300}(n)$ as the outputs of the filter bank corresponding to the 150 - 200 Hz, 200 - 250 Hz, 250 - 300 Hz and 300 - 350 Hz bands, the composite band $X_{150-350}$ is given by:

$$X_{150-350} = \frac{X_{150}}{max[X_{150}]} + \frac{X_{200}}{max[X_{200}]} + \frac{X_{250}}{max[X_{250}]} + \frac{X_{300}}{max[X_{300}]}$$

Figure 3-17 displays the output of this simple band aggregation approach for the patient shown earlier suffering from moderate mitral valve prolapse with late systolic regurgitation.

While dramatically better than unscaled aggregation this procedure suffers from the problem of destructive interference. More specifically, in a manner similar the

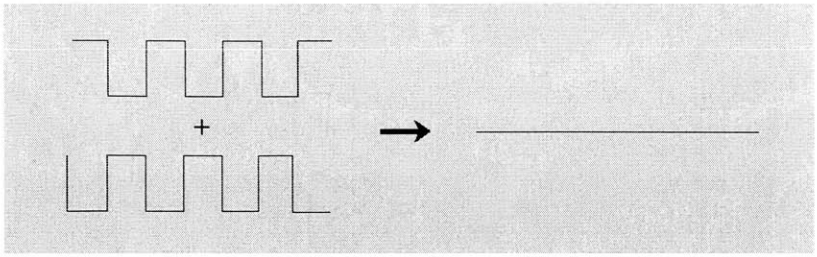


Figure 3-18: Extreme Case of Destructive Interference

extreme case shown in Figure 3-18, positive and negative values from different bands output by the filter bank may cancel each other, leading the composite band to incorrectly indicate diminished energy content (or even the absence of energy altogether) at any time instant. This problem can be countered simply by calculating the absolute value at every time instant for all the 16 bands output by the filter bank (i.e., the time-envelope characterization for these bands) and passing along these signals to the ensuing stages. Since we are concerned exclusively with the energy content of the signal as it is, it does not matter whether the signal is positive or negative since we are concerned exclusively with the amplitude of the signal at any point of time.

A complete block diagram representation of the time-frequency analysis component of the system is presented in Figure 3-19.

Figure 3-20 shows the output of our limited band aggregation approach with time-envelope characterization for the same audio recording displayed in Figures 3-17.

3.4 Beat Aggregation

The time-frequency decomposition discussed in Section 3.3 provides us with the frequency components for each selected beat. In order to observe the characteristic trends persisting amongst the majority of the beats, we also develop a mechanism to merge information from multiple systolic segments to create a single representative heart-beat for the patient. In other words, we assimilate information from the selected beats to generate the time-frequency decomposition of a hypothetical "typical" beat

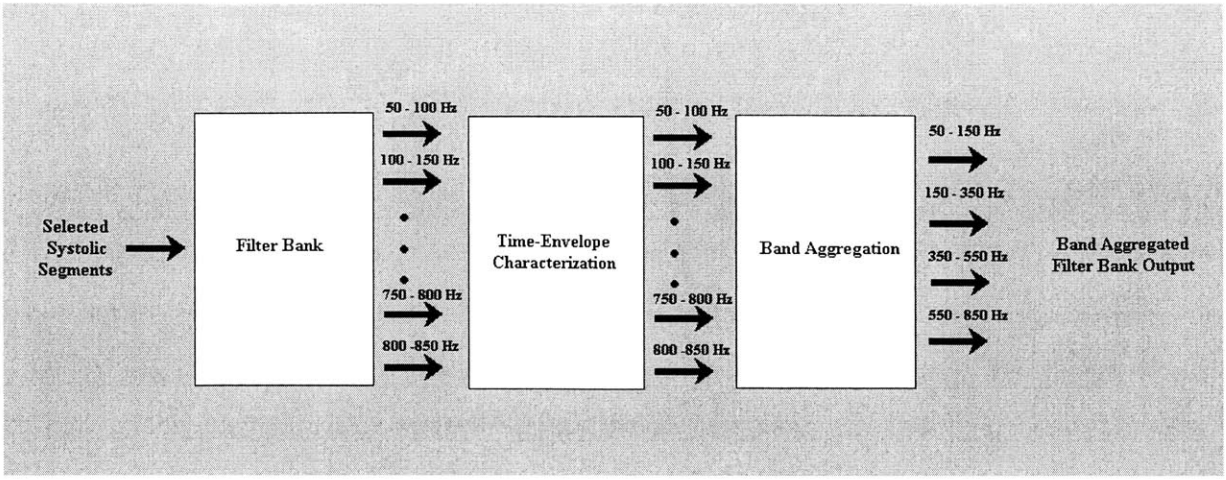


Figure 3-19: Block Diagram Representation of Filter Bank with Band Aggregation and Time-Envelope Characterization

for the patient.

The process of beat aggregation for the purpose of detection starts with the time-frequency decomposition of all beats as described in Section 3.3. This divides each beat into its band-aggregated components at different frequencies (Section 3.3.2). Again, these components are time-envelope characterized, i.e., we calculate the absolute value at every time instant for the component signals (Section 3.3.2)[8].

With this information in hand, we proceed by lining up beats in time and find the median four amplitudes at any time instant for each of the four bands as shown in the third box in Figure 3-21.

Letting $X_{i,j}$ be a one-dimensional array corresponding to the j-th frequency band of the i-th beat, this step of the prototypical beat calculation can be represented as finding the median four elements along every column of the array:

[8]As was the case with band aggregation, this requirement is particularly important while performing beat aggregation to avoid destructive interference whereby the positive and negative values from different beats may cancel each other. This would lead to the aggregate incorrectly indicating diminished energy content (or even the absence of energy altogether) at any time instant

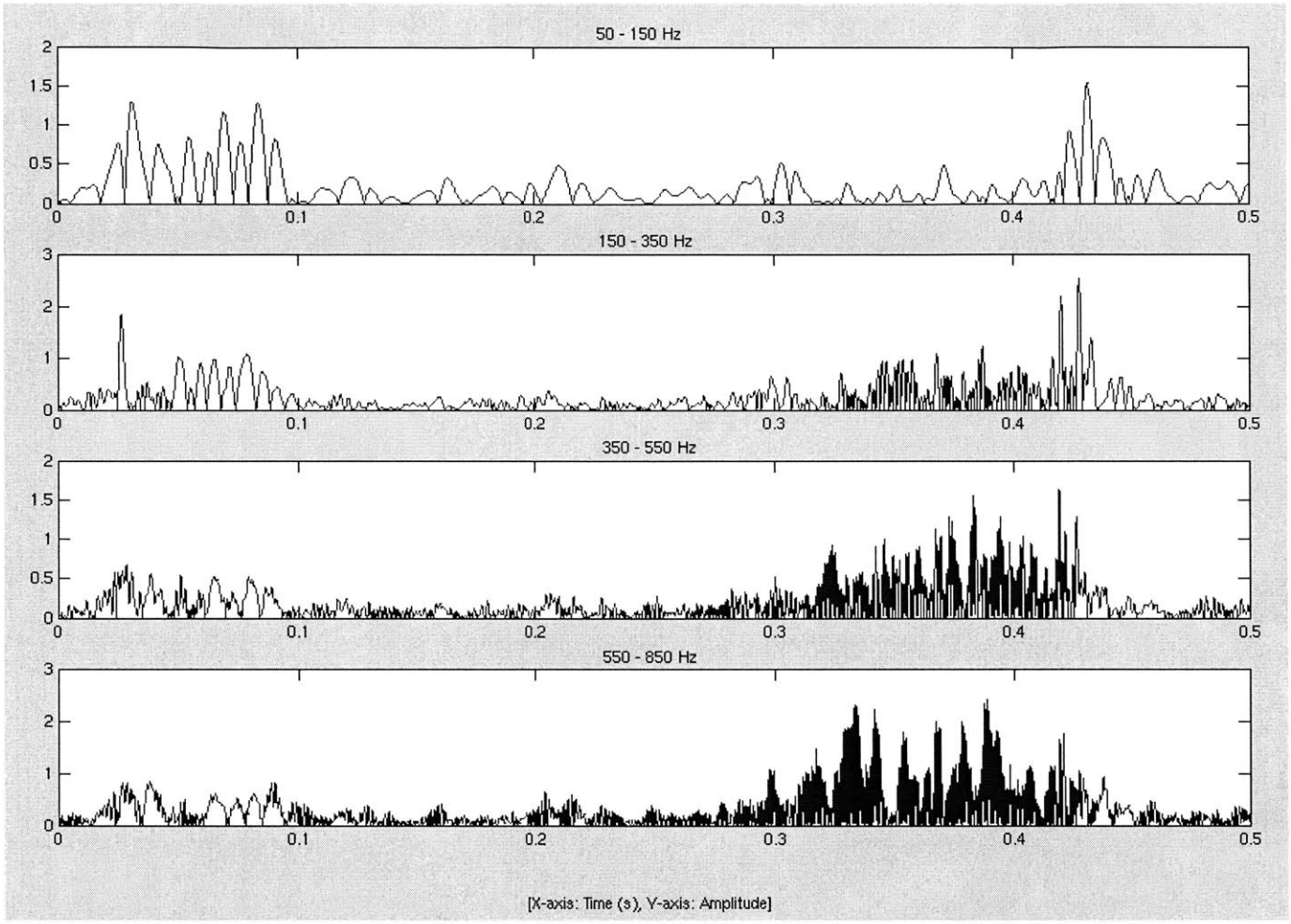


Figure 3-20: Output of Filter Bank with Time-Envelope Characterization and Post-Normalization Aggregation of Frequencies into Four Different Bands (4 Output Bands)

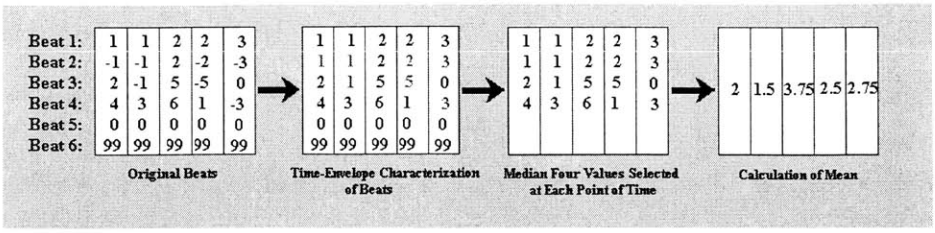


Figure 3-21: Example Prototypical Beat Construction

$$\begin{pmatrix} X_{1,j} \\ X_{2,j} \\ \cdots \\ X_{n-1,j} \\ X_{n,j} \end{pmatrix}$$

for all possible values of j[9].

The mean of these median amplitudes is then calculated for each range of frequencies output by the filter bank. This is illustrated on the right in Figure 3-21. Representing $Y_{1,j}$, $Y_{2,j}$, $Y_{3,j}$ and $Y_{4,j}$ as one-dimensional arrays that each contain one of the median four elements at every time instant for the j-th band, this step corresponds to calculating the mean of the array:

$$\begin{pmatrix} Y_{1,j} \\ Y_{2,j} \\ Y_{3,j} \\ Y_{4,j} \end{pmatrix}$$

for all possible values of j.

The end result is a time-frequency decomposition of the prototypical beat. If Z_j is the one-dimensional array containing the means of the median four values calculated at every time instant for the j-th band, the prototypical heart-beat has a time-frequency decomposition given by:

- 50 - 150 Hz: Z_1

[9]I.e., 50 - 150 Hz, 150 - 350 Hz, 350 - 550Hz and 550 - 850 Hz.

- 150 - 350 Hz: Z_2
- 350 - 550 Hz: Z_3
- 550 - 850 Hz: Z_4

One of the key advantages of pooling multiple beats is that it allows us to derive a representation of the sound actually generated by the heart while discarding random or systematic noise. Since we only examine the median four amplitudes at each time-band pair, artifacts leading to increased energy in the signal are treated as outliers and removed except in the circumstance where these artifacts occur at precisely the same instance in the cardiac cycle for at least 50% of the beats. The probability that the timing of the artifacts associated with noise will be strongly correlated to the cardiac cycle and that they will be produced several times during recording is low.

The strategy of taking the mean of the median four beats adds further robustness to noise. In particular, if there is considerable variation between the median four amplitudes at any point of time, this approach reconciles these differences by calculating the average amplitude rather than breaking the tie in favor of any of the beats. Another approach would have been simply to use the mean of the overall signal rather than the mean of the medians, but we rejected this idea since it drastically reduced immunity to artifacts. This follows from the observation that the noise in our dataset falls into the category of impulsive, salt-and-pepper noise [21, 22]. Salt-and-pepper noise is characterized by large spikes at isolated times in the signal. Only a small fraction of the samples are affected, but the error in these samples is often great. Median filters are well suited to deal with this kind of noise [23, 24]. By taking the mean of the signal we would factor in data corresponding to noisy outliers, which could potentially be significantly larger than the non-noisy signal. This would lead to the appearance of energy when in fact there was none.

Although the construction of the prototypical heart-beat ignores variations between beats, it does not result in a significant loss of relevant information. Patients suffering from mitral valve prolapse show evidence of the disorder on the majority of the recorded beats. In such a case, the calculation of the median beats leads to

four beats, all possessing the MVP signature. Conversely, although noise might cause normal patients to have one or more beats that appear to contain energy at higher frequencies just before systole, it is extremely improbable that the median four beats will all suffer this effect.

3.5 Decision Mechanism

The process of beat selection (Section 3.2) and the time-frequency analysis (Section 3.3) together provide information regarding how energy is distributed over different frequency ranges for heart-beats belonging to any patient. The beat aggregation mechanism supplements this by revealing persisting trends in the signals recorded. In order to make a recommendation to physicians, all this information must be mapped to a classification.

3.5.1 Band-Specific Thresholding

To obtain a classification we employ a decision mechanism using band-specific thresholding.

60% Energy Scaled Lag Feature Construction

Since MVP is characterized by increased energy content at higher frequencies during the last half of systole (Section 2.2.2), we focus on locating the peaks in energy at higher frequencies for every selected beat. The position of maximum signal amplitude is determined separately for each range of frequencies output by the filter bank (Section 3.3.1). This search is limited to a region from mid-systole to slightly after S2 since we expect all diagnostic information to be present there. In particular, if X_i, j is a one-dimensional array of audio data corresponding to the j-th frequency band of the i-th selected beat [10], we define $\text{peakpos}_{i,j}$, the position of maximum amplitude for the j-th frequency band of the i-th beat to be:

[10] I.e., an array containing the amplitude of the recorded audio signal for the systolic segment between the i-th S1 and S2 with only the components in the j-th band included

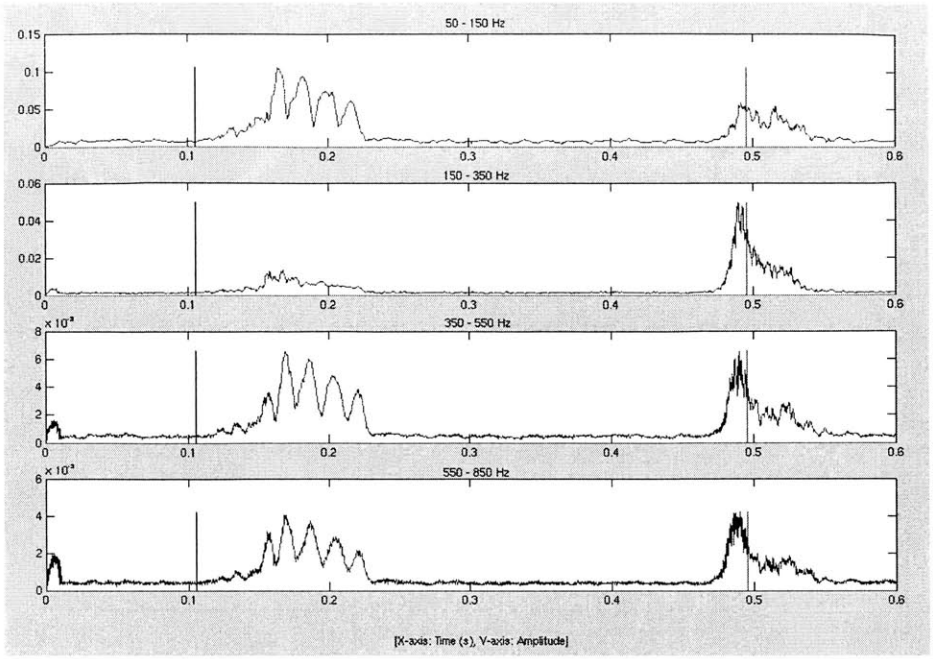


Figure 3-22: Time-Frequency Decomposition for Non-MVP Patient (1)

$$\text{peakpos}_{i,j} = \text{maxpos}\ [\text{X}_{i,j}(\frac{length(X_{i,j})}{2} : \text{length}(\text{X}_{i,j}))]$$

In the absence of MVP, the peaks in energy at higher frequencies are solely the result of the harmonics associated with S2. As a result, for normal patients and those with benign murmurs, the maximum signal amplitude occurs at or very close to S2. Figures 3-22, 3-23 and 3-24 illustrate this effect[11].

On the other hand, for patients suffering from MVP there is substantial energy content at higher frequencies during the last half of systole and the position of maximum signal amplitude tends to shift significantly prior to S2 (Figures 3-25, 3-26 and 3-27). However, this is not always the case. Occasionally, the peaks at higher frequencies become flatter and significantly wider, extending well into systole (Figures 3-28, 3-29 and 3-30). In some cases, additional peaks may also appear well before S2(Figures 3-31, 3-32 and 3-33).

[11]The vertical black line on the left for all the figures that follow indicates the detected location of the QRS complex. The vertical red line on the right corresponds to the predicted position of S2.

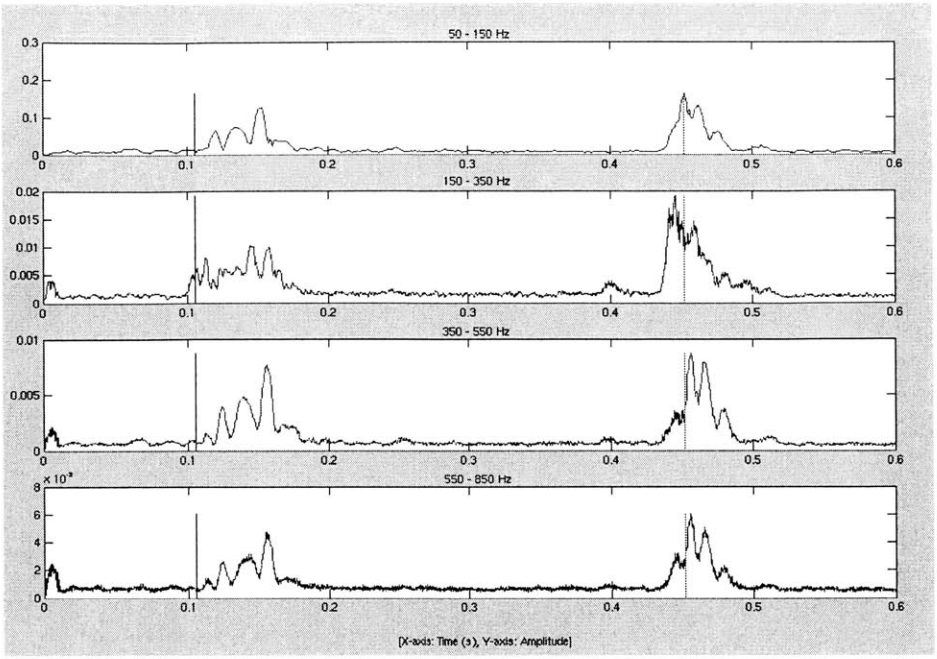


Figure 3-23: Time-Frequency Decomposition for Non-MVP Patient (2)

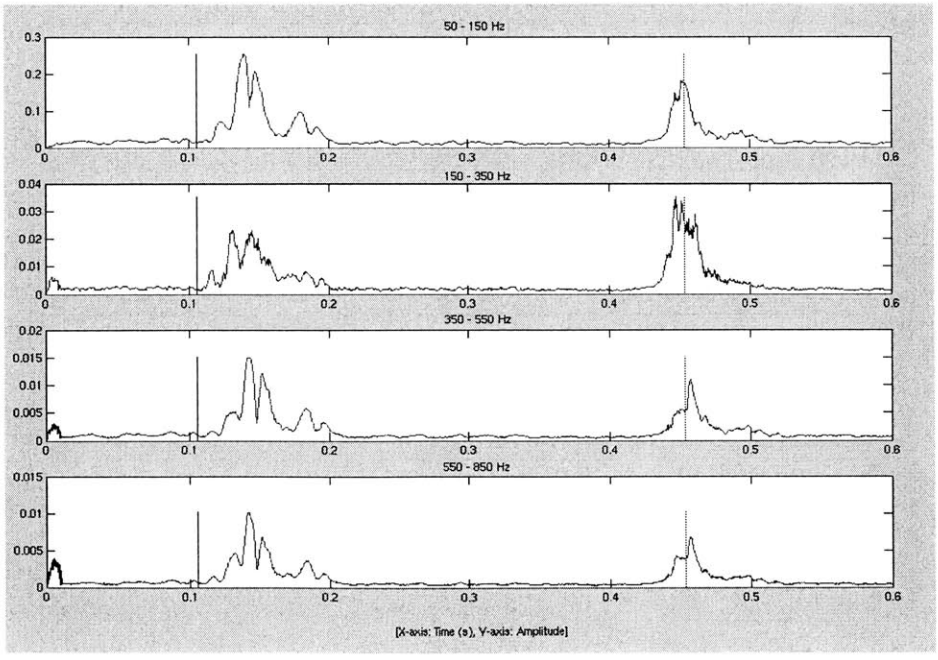


Figure 3-24: Time-Frequency Decomposition for Non-MVP Patient (3)

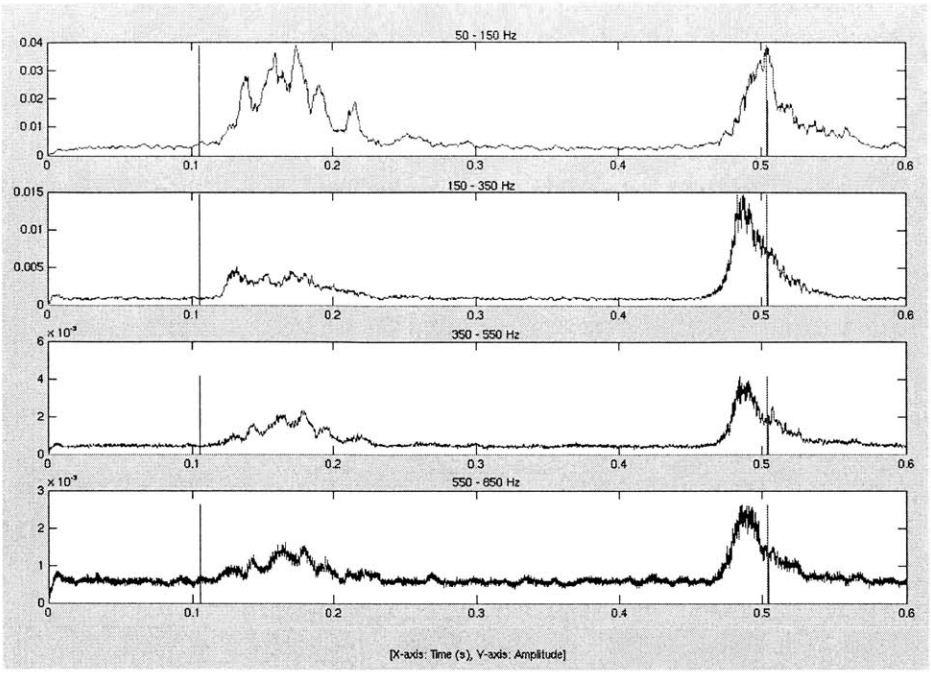


Figure 3-25: Time-Frequency Decomposition for MVP Patient with High Frequency Peaks Shifted Significantly Prior to S2 (1)

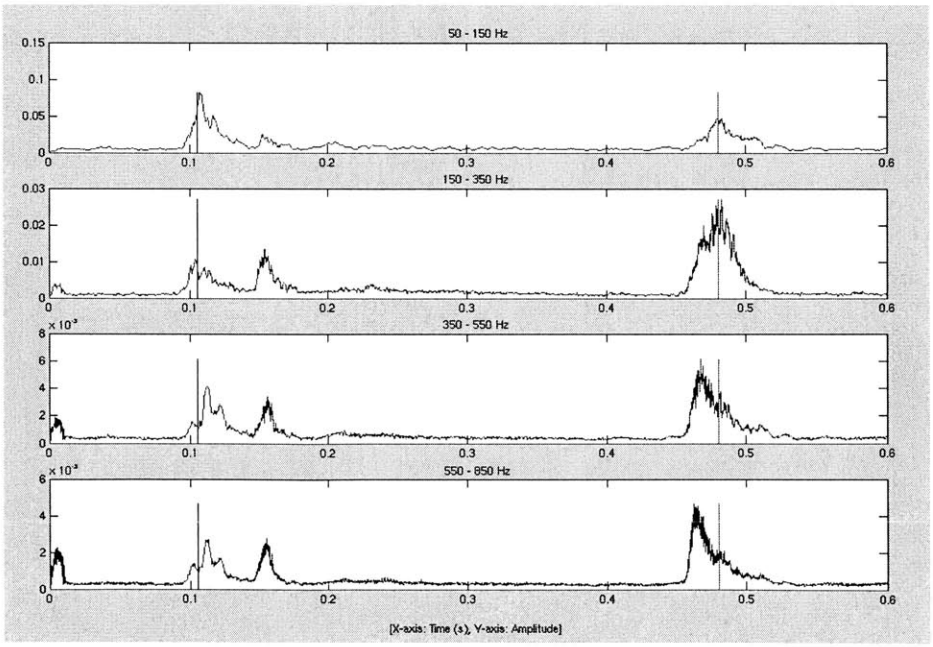


Figure 3-26: Time-Frequency Decomposition for MVP Patient with High Frequency Peaks Shifted Significantly Prior to S2 (2)

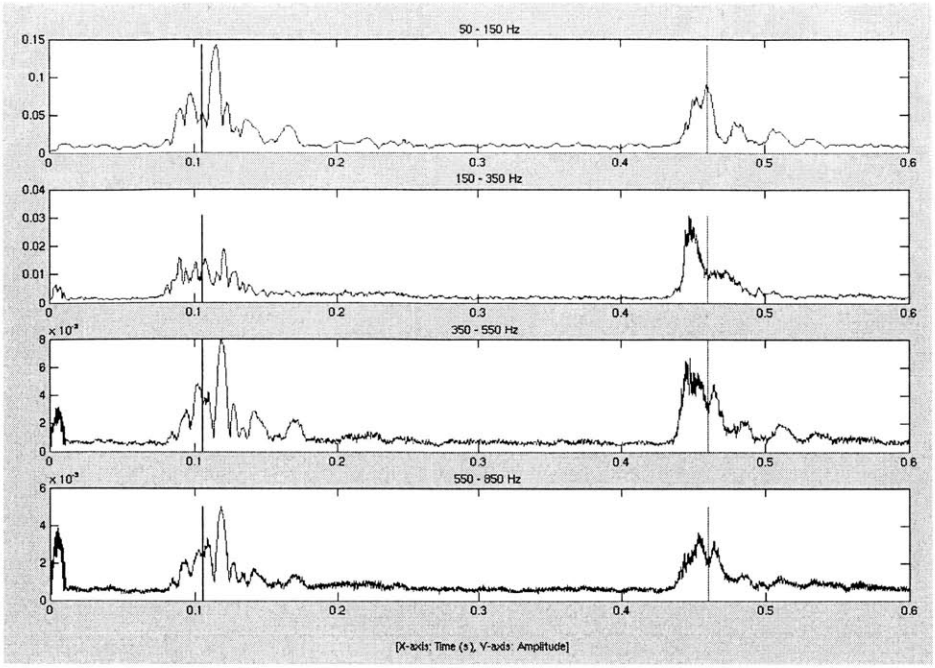


Figure 3-27: Time-Frequency Decomposition for MVP Patient with High Frequency Peaks Shifted Significantly Prior to S2 (3)

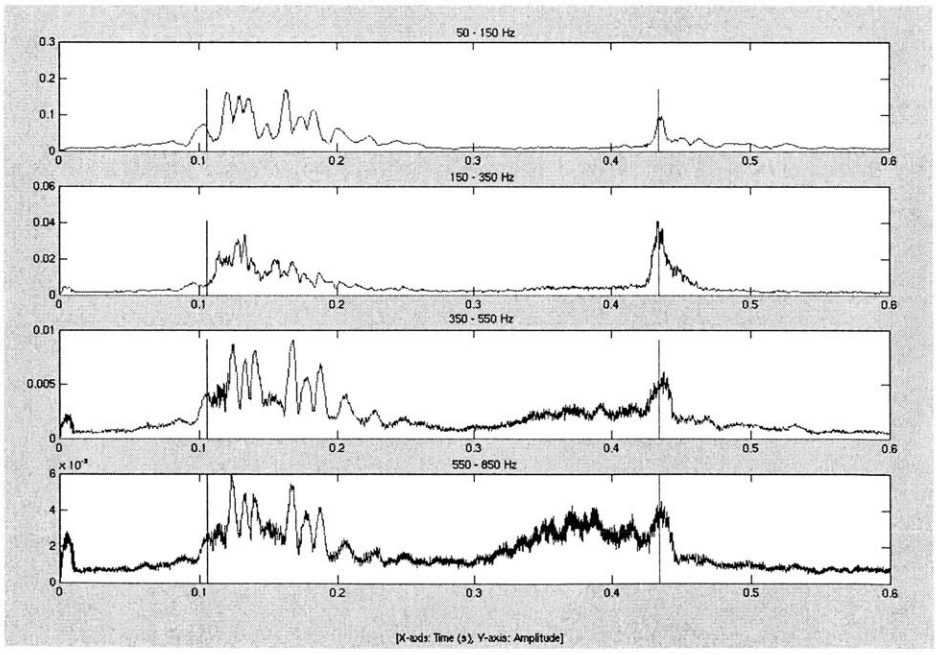


Figure 3-28: Time-Frequency Decomposition for MVP Patient with Wider High Frequency Peaks Extending into Systole (1)

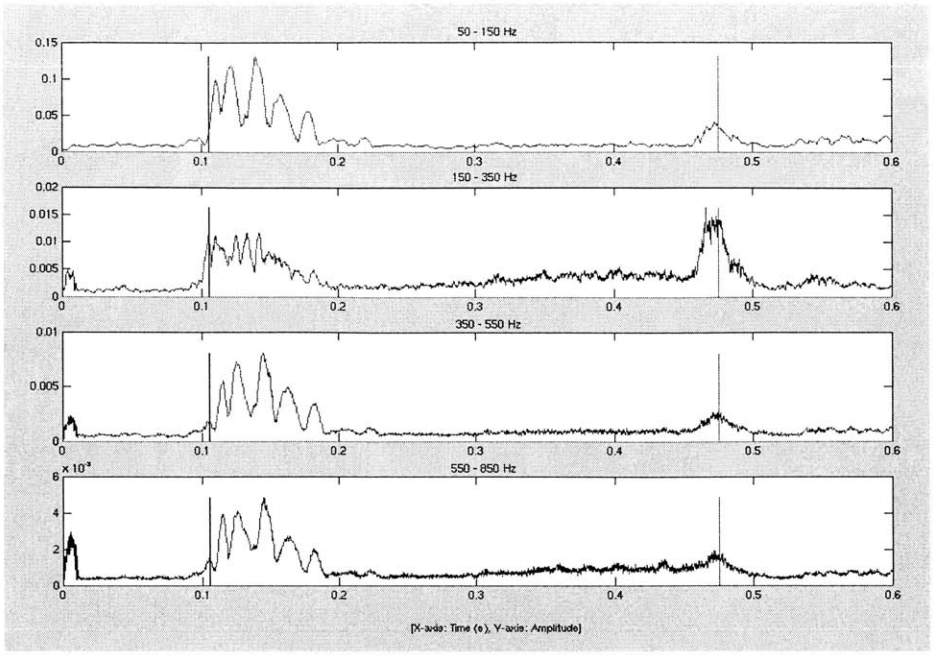


Figure 3-29: Time-Frequency Decomposition for MVP Patient with Wider High Frequency Peaks Extending into Systole (2)

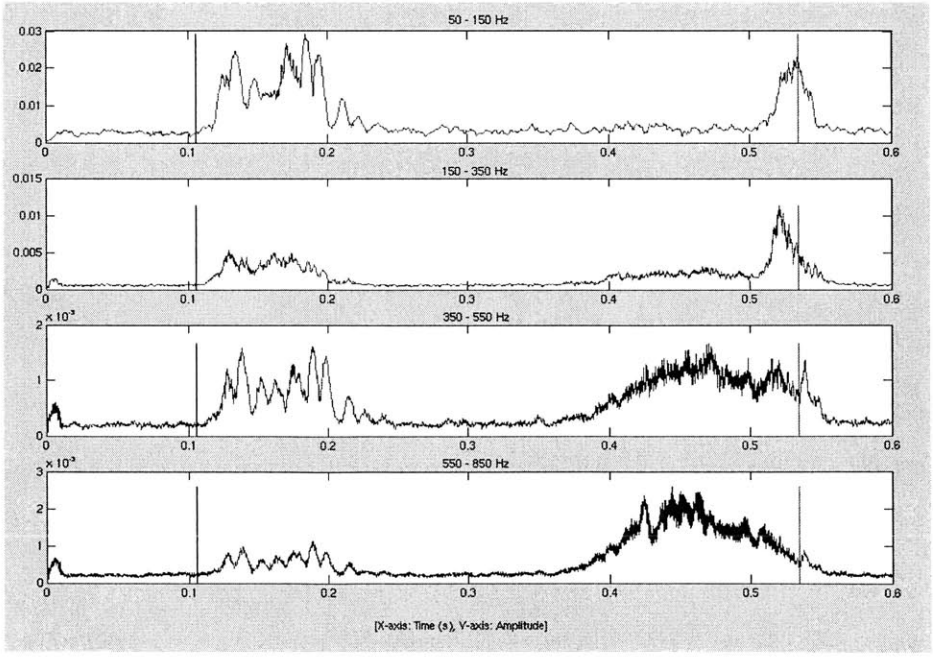


Figure 3-30: Time-Frequency Decomposition for MVP Patient with Wider High Frequency Peaks Extending into Systole (3)

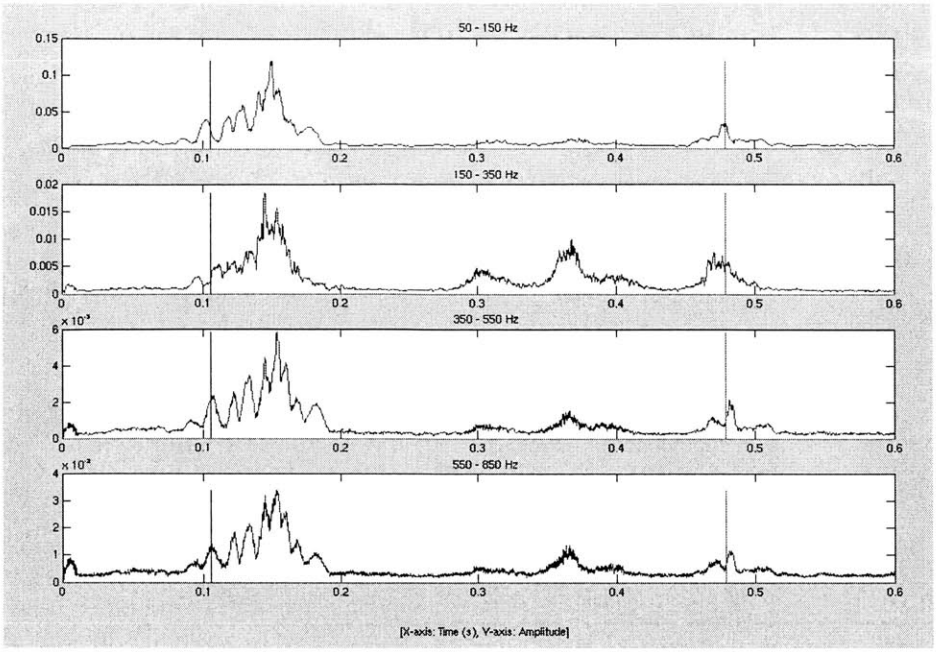


Figure 3-31: Time-Frequency Decomposition for MVP Patient with Additional High Frequency Peaks (1)

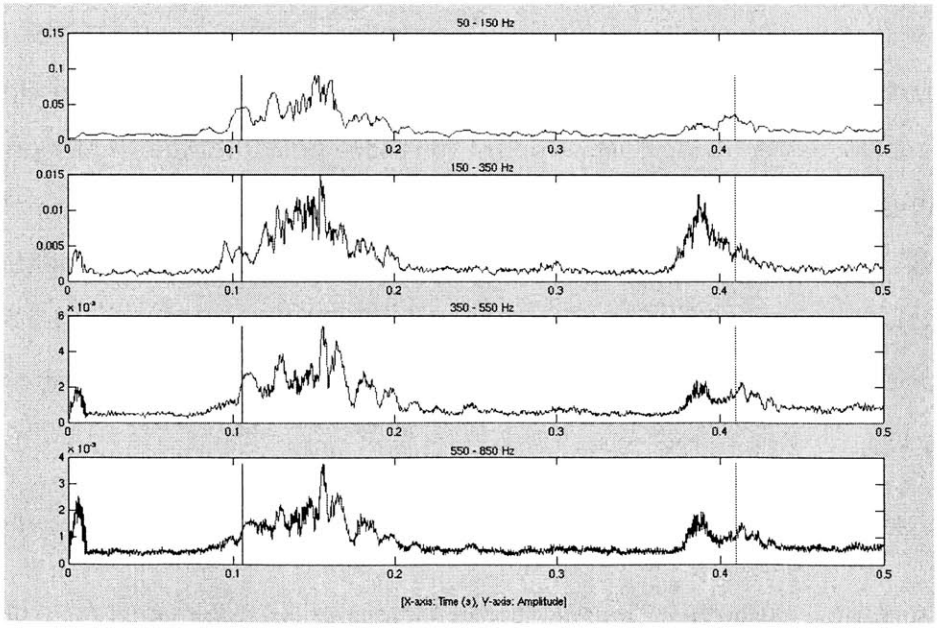


Figure 3-32: Time-Frequency Decomposition for MVP Patient with Additional High Frequency Peaks (2)

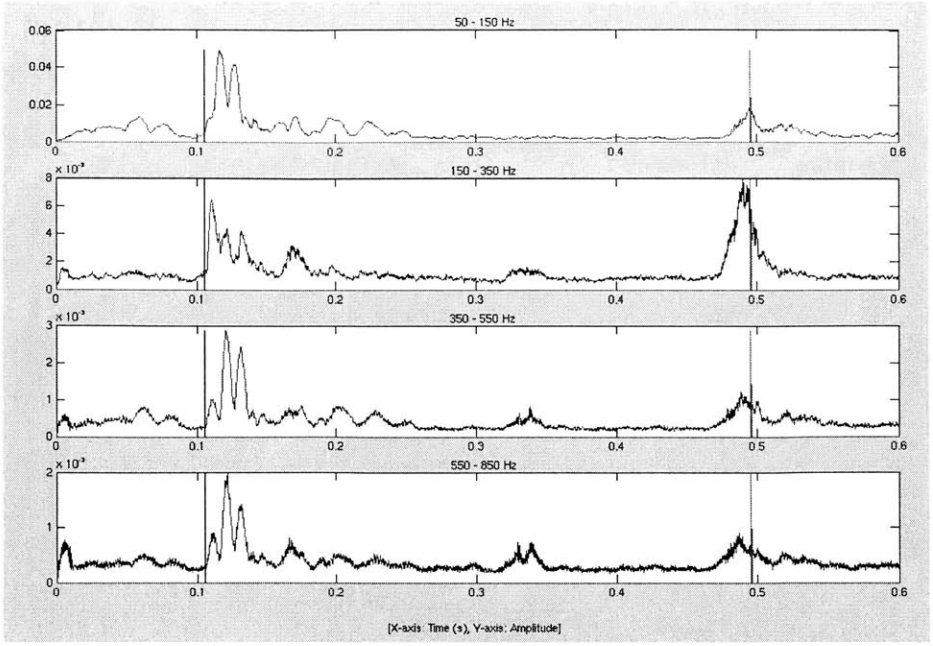


Figure 3-33: Time-Frequency Decomposition for MVP Patient with Additional High Frequency Peaks (3)

As a result of these phenomena, simply examining the position of maximum signal amplitude is not sufficient. Specifically, it prevents us from successfully distinguishing normal patients and those with benign murmurs from patients in whom the onset of MVP leads to flatter, wider peaks or the presence of additional peaks prior to S2. This is illustrated in Figures 3-31 and 3-32.

Based on this behavior, a more sophisticated approach calculates the earliest point in the last half of systole where the signal amplitude first exceeds 60% of the peak value[12]. In particular, for the j-th frequency band of the i-th beat, we define $60\text{peakpos}_{i,j}$ to be:

$$60\text{peakpos}_{i,j} = \text{first-index-true } [\mathrm{X}_{i,j}(\frac{length(X_{i,j})}{2} : \text{length}(\mathrm{X}_{i,j})) \geq 0.6 \text{ peakpos}_{i,j}]$$

The intuition behind calculating this value is that it allows us to measure exactly how early on during the last half of systole the presence of considerable energy can be detected. We can use this result to compute the lag between the earliest occurrence of

[12]This value was experimentally determined to be optimal by examining all values from 25% to 95% in increments of 5.

energy and S2 (i.e., the time interval by which energy at higher frequencies precedes S2) by:

$$60\text{prec}_{i,j} = \text{length}(X_{i,j}) \text{ - } 60\text{peakpos}$$

Where $60\text{prec}_{i,j}$ is the lag between the earliest occurrence of energy and S2 for the j-th frequency band of the i-th beat.

Finally, since the S1-S2 intervals can vary dramatically between patients as discussed in Section 3.2.1 we scale $60\text{prec}_{i,j}$ by the duration of systole. In other words, the final feature used, $60\text{precscaled}_{i,j}$, is given by:

$$60\text{precscaled}_{i,j} = \frac{60prec_{i,j}}{systoliclength}$$

This is then employed by the classifier to make a recommendation.

Threshold-based Classification on a Per Beat Basis

Using the values of $60\text{precscaled}_{i,j}$, we first classify each individual beat. To do this, we define a threshold value, t_j, for each frequency band j such that if $60\text{precscaled}_{i,j}$ is at least t_j, the beat will be declared as being indicative of MVP. We declare the i-th beat as belonging to a patient suffering from MVP if for any value of j:

$$60\text{precscaled}_{i,j} \geq t_j$$

The thresholds for each band were determined by means of examination to be [13]:

- 150 - 350 Hz : -0.045
- 350 - 550 Hz : -0.045
- 550 - 850 Hz : -0.02

Since the 50 - 150 Hz band corresponds to low frequencies, we did not consider the value of the feature for it and there was no associated threshold value.

If a beat is not classified as being characteristic of MVP it is by default assumed to be belonging to a patient who is either normal or has a benign murmur.

[13]In Section we argue that the values selected are relatively stable in that small variations do not lead to a dramatic change in the results achieved.

Classification on a Per File Basis

Given the classification for each beat, we can assign an overall label of MVP to a file bearing recordings for a patient if 40% of the selected beats are indicative of MVP[14]. Again, by default, the absence of MVP is assumed to correspond to a patient who is either normal or has a benign murmur.

Classification on a Per Patient Basis

As will be explained in Section 3.6.1, the dataset for this project contains multiple files with audio recordings for each patient (corresponding to different patient positions and recording sites). The overall patient recommendation is MVP if any of the files are labelled to belong to a patient suffering from the disorder. This is to gear the decision making process towards minimizing false negatives. Patients who are not classified as MVP are assumed to be normal or with benign heart murmurs.

Prototypical Beat Detection

Although the construction of the prototypical heart-beat (Section 3.4) results in the loss of information related to the variation between beats, it has desirable properties such as increased immunity to noise. To benefit from these effects we can replace the decision mechanism that performs beat to beat detection with one that simply examines the prototypical beat in a manner identical to that discussed in Section 3.5.1. Rather than considering every beat, we only examine the high-frequency bands of the prototypical beat and declare the patient to be suffering from MVP if any of the threshold criteria are met for it.

We have designed our system to allow for both modes of decision making (i.e., beat to beat detection and prototypical beat detection). The specific mode of operation can be determined by the user depending on the acceptable trade-off between noise reduction and preservation of the variation between beats.

[14]The value 40% was chosen to minimize the error rate obtained on a per file basis while varying this parameter between 0% to 100% in increments of 5.

3.6 Evaluation

This section presents the results obtained by our system. We start with a discussion of our dataset and then quantify the performance of the detector on a per-patient.

3.6.1 Data

The dataset for this project consists of audio signals and simultaneous EKG recorded for fifty-one patients. Thirty of these have normal hearts or benign murmurs, whereas the remaining twenty-one suffer from MVP.

For each of the patients, two recordings were made (leading to a total of 102 files). Signals were collected from the apex and left lower sternal border (Section A.4.3) with the patient lying down[15].

Most of the data was collected at Massachusetts General Hospital by Dr. Francesca Nesta from patients referred for an echocardiogram. In addition to this, Dr. Nesta also collected data during field studies from families with one or more members diagnosed as suffering from MVP. The data collected from families was recorded at the homes of the subjects being examined.

The dataset for this project was challenging for two reasons. Firstly, the files belonging to patients referred for an echocardiogram corresponded to individuals who had already been diagnosed by primary physicians as possessing MVP. Consequently, the majority of the non-MVP files in our dataset actually represented mistakes made earlier by doctors. A false positive rate of less than 100% on these files would already be an improvement upon the existing performance of primary care practitioners.

In addition to this, the files collected during field studies were recorded in an uncontrolled environment. As a result, there is considerable noise in these signals and some of the artifacts observed were the sound of children playing and that of people talking. For a small number of files, beats are also corrupted by dogs barking.

[15] As discussed in Section 2.2.2, murmurs due to MVP are best heard from the apex and left lower sternal border for patients lying down. As a result, only the files corresponding to these sites and positions were considered.

Another important point to mention is that all files were labelled using the results of an echocardiogram examination. This is widely used by physicians as the "gold standard" for diagnosing MVP. We therefore employ these label as the ground truth in evaluating our system.

All data was sampled at 44 kHz with 16-bit quantization.

3.6.2 Results

Sections 3.1 - 3.5 detail the internal workings of our system. What follows here is a discussion of how well our software performs. We focus on examining the extent to which we are able to realize our goal of improving contemporary health care. As our results show, our system performs significantly better than primary care physicians, promising a drastic reduction in the number of incorrect referrals to cardiologists. Though the evidence is less conclusive, it also suggests that our system can reduce the number of patients classified as being non-MVP when they do, in fact, suffer from the disorder. Since such patients may require treatment, it is critical to catch them at an early stage of the disease.

Classification Rates

Our detector achieves similar results for both beat to beat detection (Section 3.5.1) and prototypical beat detection for our dataset (Section 3.5.1). The prototypical beat detection approach performs marginally better than the beat to beat detection in terms of false positives and both methods work identically for false negatives. This is supported by our intuition and can be explained from the fact that a number of our files are corrupted by the presence of noise (especially those collected during field studies). The beat by beat approach incorrectly classifies files not discarded by the beat selection step (Section 3.2.2) but still possessing significant high frequency energy as MVP. The prototypical detection approach conversely pools noisy beats and is more robust in the face of this effect.

As a result of the near identical performance of the two detection strategies, we

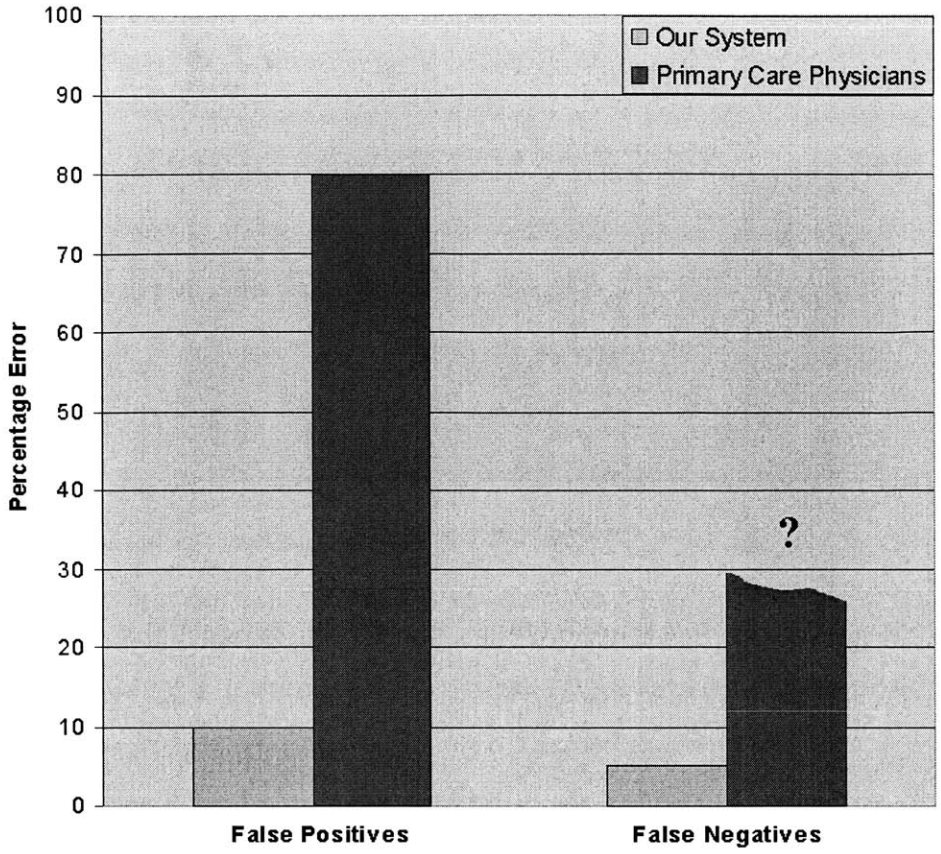


Figure 3-34: Comparison of Our System to Primary Care Physicians

focus only on prototypical beat detection, contrasting the results achieved by this approach to the classification rates obtained for primary care physicians.

The performance of our detector is presented in Table 3.1 and Figure 3-34.

	MVP	Non-MVP
Correctly Diagnosed	20 (95%)	27 (90%)
Incorrectly Diagnosed	1 (5%)	3 (10%)

Table 3.1: Performance of Detector (in terms of number of patients)

Given these results, our system can be seen to convincingly outperform primary care physicians in terms of false positives. Specifically, our software reduces the

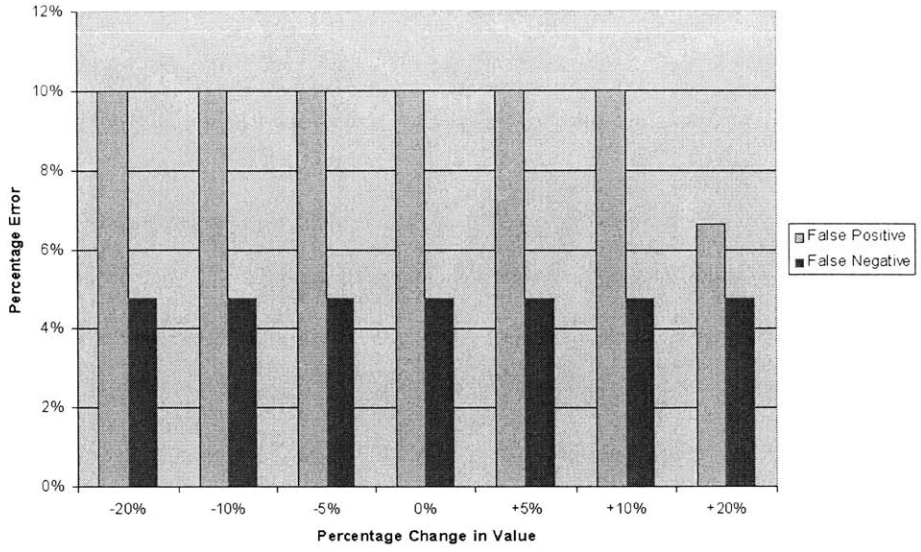


Figure 3-35: Sensitivity of 150 - 350 Hz Threshold Value

widely reported false positive rate of 80% to 10%. It is difficult to compare the performance in terms of false negatives since the prevalence of undiagnosed cases is poorly documented in the medical literature. Nevertheless, anecdotal evidence suggests that our system fares significantly better than primary care physicians on this front as well. Dr. Nesta, a trained cardiologist, rated each patient on the degree of MVP heard during auscultation[16]. 21% of the patients suffering from MVP were assigned the same scores as non-MVP patients. With high probability these patients would have been missed by primary care physicians. Conversely, our system obtains only a solitary false negative, corresponding to a patient suffering from minimal MVP with no regurgitation. The patient did not require any treatment.

Sensitivity Analysis

Figures 3-35 - 3-37 show how the false positive and false negative rates are affected by changes to the thresholds chosen for the high frequency bands.

As can be seen, the values chosen are fairly stable. Small changes do not lead to significantly different diagnoses. This gives us some confidence that our system will

[16]This was done after seeing the results of the echocardiogram classification

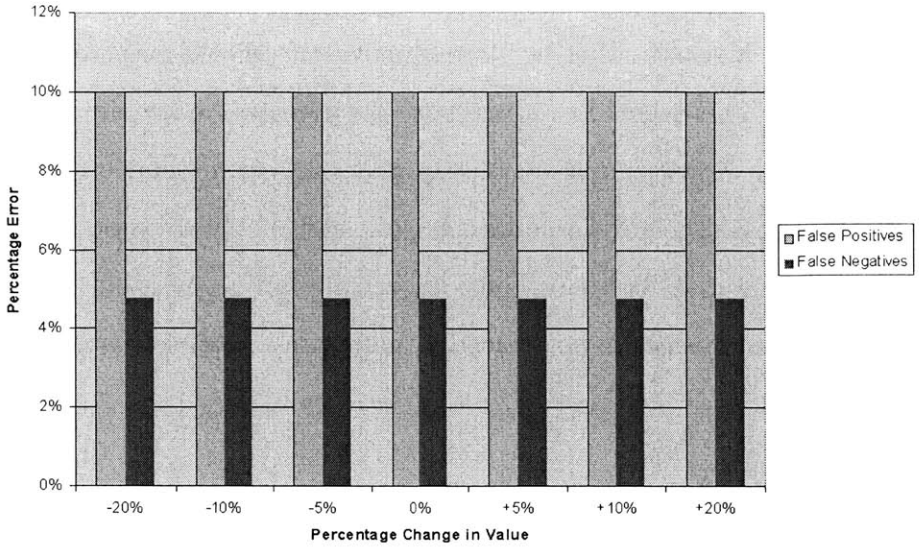


Figure 3-36: Sensitivity of 350 - 550 Hz Threshold Value

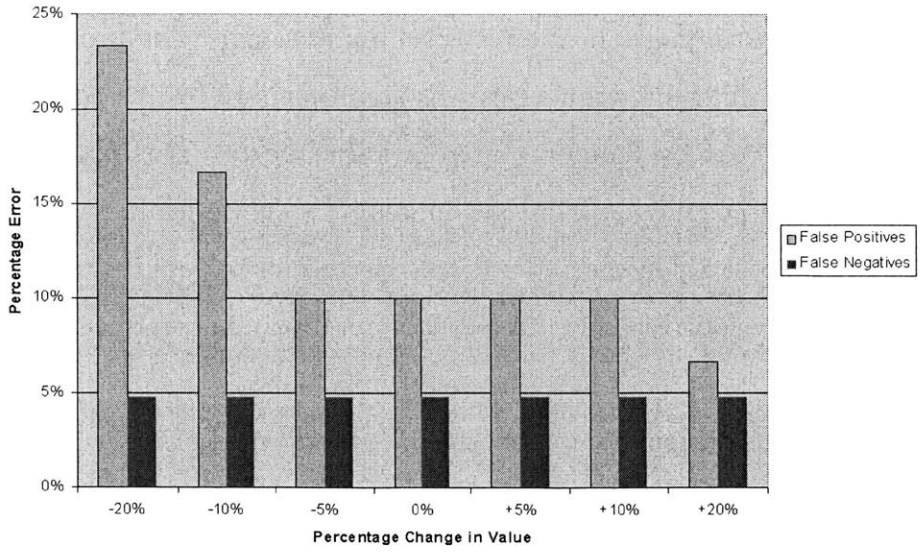


Figure 3-37: Sensitivity of 550 - 850 Hz Threshold Value

maintain similar performance beyond the scope of our dataset.

Looking at the data in Figures 3-35 and 3-37, it appears that we can improve the performance of our system further by increasing the thresholds for the 150-350 Hz and 550-850 Hz bands. This seems to reduce false positives while keeping false negatives constant. We do not however pursue this approach since we believe that for a larger dataset an increase in the threshold values will lead to a corresponding increase in the false negative rate. This is not considered to be an acceptable trade-off for the decrease in false positives, and we prefer to operate in a conservative mode.

3.7 Summary

This chapter detailed the system developed to detect MVP. We examined the process of segmenting heart sounds based on a novel approach to locate S2. We also presented a mechanism for selecting noise-free beats that are physiologically relevant. To study the heart sounds obtained in this manner, we proposed a time-frequency decomposition that aggregated various frequency components in a number of different ways to reveal diagnostic information and make the system more robust to the presence of random and systematic noise in the recorded audio signal. The process of merging information across different frequencies was supplemented by the construction of a prototypical heart-beat by temporal aggregation of beats. This process of pooling multiple beats together further increased immunity to noise while emphasizing characteristic trends observed in the audio data collected for patients. Finally, a decision mechanism was developed to use the features obtained and produce a classification to be output.

When tested, the system comprehensively outperformed primary care physicians. False positives were reduced from over 80% to almost 10%. It was further argued that the system reduces false negatives considerably (from a lower bound of around 21% to 5%). We also presented the results of a sensitivity analysis to increase confidence that our system generalizes well and has scope beyond the dataset. These results are highly encouraging and promise to lead to extensive deployment of our system.

Chapter 4

Diagnostic Aids

This chapter presents the suite of audio-visual diagnostic aids developed for the purpose of increasing confidence in the recommendations made by the system. We start with a discussion of the difficulties associated with traditional auscultation and motivate the need for a toolbox of diagnostic aids. Section 4.2 then details the process of visually displaying the prototypical heart-beat for each patient. This merges information from across multiple beats and presents it to users in a compact form. The functionality of reduced rate playback of heart sounds (with preservation of frequency content) is explained next in Section 4.3. This is intended to tease apart events in the cardiac cycle by increasing the separation between them in time. Doing so makes it easier to distinguish soft murmurs from other heart sounds only a few milliseconds away. Finally, Section 4.4 describes another auditory aid that calculates an enhanced audio-prototypical beat for each patient and plays this back.

4.1 Difficulties Associated with Traditional Auscultation

The ability to perform auscultation has traditionally been restricted to skilled cardiologists. One reason for this is that the energy in the high frequency audio signature indicative of MVP is several orders of magnitude smaller than the energy at lower

frequencies. Hence, to successfully diagnose the presence of mitral valve prolapse it is necessary to filter out the lower frequency component of the signal. It is only with several years of practice that the human ear can be trained to achieve this successfully.

Another difficulty associated with cardiac auscultation stems from the fact that in the case of certain patients suffering from mild MVP with trace regurgitation, the only indication of a heart disorder is the presence of a peak at higher frequencies right at the end of systole (i.e., just prior to S2). Although a computer is able to recognize this effect given a signal that has been sampled at a sufficiently high rate, it is difficult for the human ear to recognize distinct events such as these that are only a few milliseconds apart.

4.2 Prototypical Heart-Beat Visualization

The first of our diagnostic aids addresses the issues presented in Section 4.1 by providing an effective means of visualizing the diagnostic information contained in the audio signal.

4.2.1 Scaled Display of Different Frequency Bands

The problem of the MVP signature having a significantly lower energy content relative to the remainder of the signal can be solved by scaling the energy in higher frequency bands and displaying it separately from the rest of the signal. More specifically, since information at the frequencies of interest gets washed out by heart sounds at lower frequencies, we display the high and low frequency components separately to prevent any adverse interaction between the two. This division of the audio signal into different frequency ranges is identical to the time-frequency decomposition discussed in Section 3.3.2. The visualization aid developed makes use of this fact by plotting the band-aggregated output of the filter bank directly to a screen. This achieves the desired separation while avoiding the need for redundant computation.

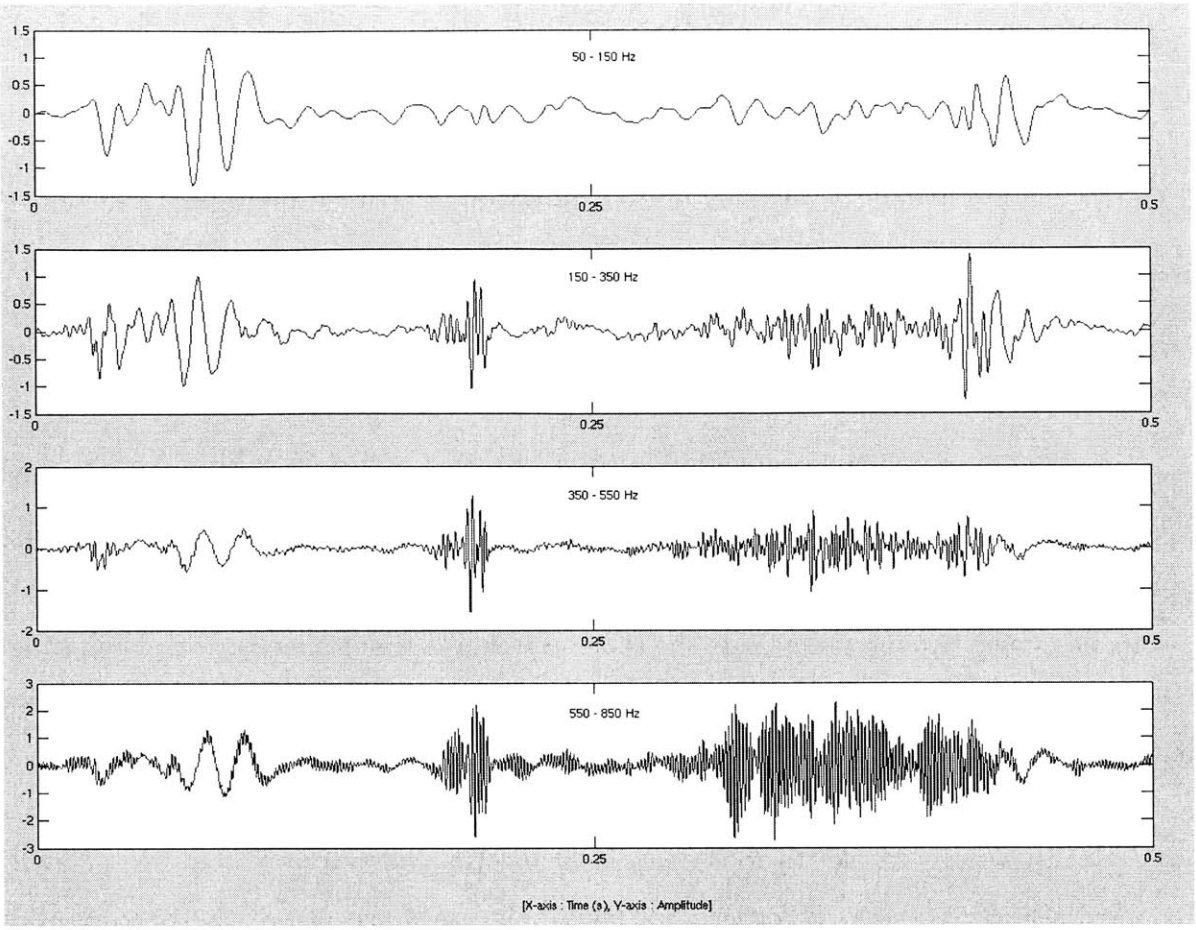


Figure 4-1: Time-Frequency Visualization of a Sample Beat Corresponding to an MVP Patient

4.2.2 Display of Entire Beats

The issue of distinguishing between events that are close together is addressed by providing a fixed snapshot of an entire beat at a time. This allows the position of two or more events, and the interval between them, to be examined in detail for any given beat. More generally, viewing entire beats allows comparisons in morphology, amplitude, location etc. to be conducted between different parts of the signal. Whereas listening only presents a fraction of the total information in the beat at any time, a visual display (such as the one presented in Figure 4-1) can be used to output the content of the signal in its entirety, allowing the separation of various events in both time and frequency to be noted.

4.2.3 Beat Aggregated Display

One of the key difficulties associated with visualizing the audio information recorded for any patient is the sheer amount of content that must be absorbed by the user. More specifically, information from multiple beats must be presented in a manner that facilitates decision making by revealing relevant information.

The sequential display of individual beats clearly prevents us from achieving this goal. In particular, such an approach greatly increases the difficulty of comparing non-adjacent beats and as such renders inter-beat comparisons impossible to any reasonable extent. At the same time, the other extreme of simultaneously displaying the time-frequency decomposition of all selected beats for a patient leads to information overload, making it difficult to focus on important features.

Ideally, we would like to focus on displaying the characteristic features of each patient's heart sounds. In order to achieve this, we make use of the prototypical heart-beat discussed in Section 3.4. This merges information from different beats and creates a single beat for each patient.

The visual aid developed plots the four frequency bands of this beat to a screen as shown in Figures 4-2 and 4-3. In addition to plotting the four bands, the positions of the QRS complex (vertical black line) and S2 (vertical red line) are also displayed and the region of interest in the signal for detecting MVP is highlighted. As the plots demonstrate, it is relatively straightforward to discriminate MVP visually, even though this might not be the case by listening to the recorded heart sounds[1]

[1]Sound files for both patients are provided on the web at `http://maas.lcs.mit.edu/sounds/`. The patient shown in Figure 4-2 corresponds to the file labelled "Normal" on the website, whereas the "Tricky MVP" file presents the signal recorded for the patient in Figure 4-3.

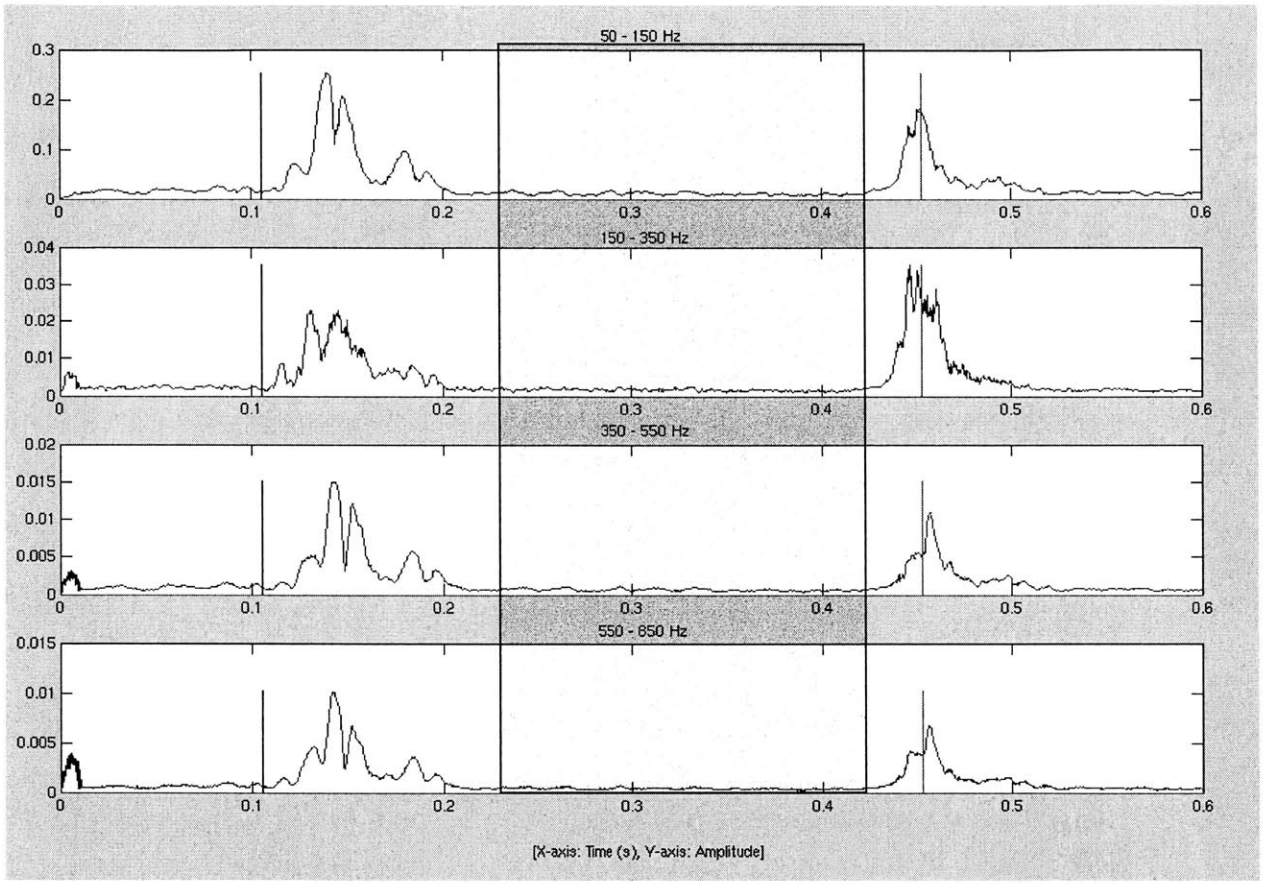


Figure 4-2: Time-Frequency Visualization of Prototypical Beat for a Non-MVP Patient

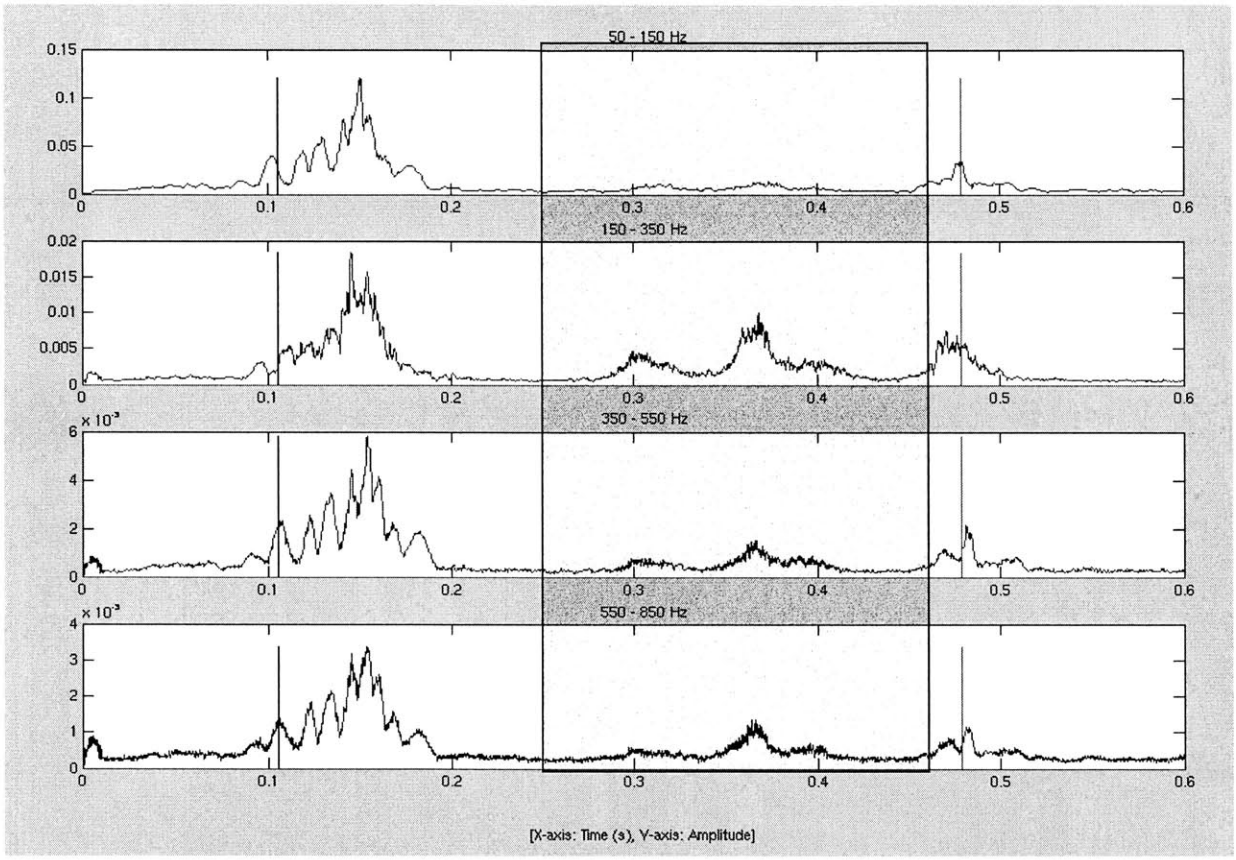


Figure 4-3: Time-Frequency Visualization of Prototypical Beat for an MVP Patient

4.3 Reduced Rate Playback with Preservation of Frequency Content

The second of the two diagnostic aids developed to assist primary care physicians focusses on providing the ability to play back heart sounds at slowed down rates. This facilitates differentiating between various acoustical events by making the separation between them more noticeable.

A possible approach to achieve a reduced rate playback would be to upsample the recorded signal and play it back at the same sampling rate at which it was initially recorded. This technique does not work well however, as it modifies the frequency content of the signal [25]. Since mitral valve prolapse is characterized by the presence of energy at higher frequencies, this effect is unacceptable.

As an alternative to upsampling, we make use of existing research on the subject of variable rate playback of sound without loss of fidelity [26, 27, 28, 29, 30, 31, 32]. We employ the phase vocoder implementation for timescale modification of audio developed at Columbia University [32]. This choice was based on the open-source, publicware nature of the system as well as the fact that it was coded up in MATLAB. This allowed it to be readily incorporated into our system.

A sample of the slowed down heart sounds produced by the approach discussed here can be found on the web at `http://maas.lcs.mit.edu/sounds/`.

4.4 Enhanced Audio-Prototypical Heart-Beat Playback

In addition to providing the functionality to play back heart sounds at a reduced rate, we also allow for the playback of a slightly modified version of the prototypical beat. Since the time-envelope characterization step leads to a signal that has a positive value at every time instant and greatly changes the auditory characteristics of the signal, we allow for the construction of a slightly imprecise version of the prototypical

beat that does not include this step. The rest of the procedure for constructing the representative beat is unchanged and the resulting signal obtained is called the audio-prototypical heart-beat.

The audio-prototypical heart beat can be reconstituted from its different frequency components merely by adding these components together. Since information at higher frequencies is more relevant for diagnosing MVP, we normalize all the frequency bands prior to combination. Specifically, in a manner identical to that presented in Section 3.3.2 for aggregating bands, we scale each frequency band by its maximum amplitude. This leads to an increase in the high frequency content of the signal.

Although removing the time-characterization step during the construction of the audio-prototypical beat leads to the possibility of destructive interference, this effect is unavoidable in order to obtain an interpretable beat. In practice however, the loss of information is normally insignificant and subjectively, the audio-prototypical heart-beat makes it considerably easier to hear murmurs due to MVP.

A sample of the enhanced audio-prototypical beats produced by the approach discussed here can be found on the web at `http://maas.lcs.mit.edu/sounds/`.

4.5 Summary

This chapter discussed the suite of audio-visual aids developed during the course of this project to assist primary care physicians. These tools address the challenges associated with traditional auscultation. We proposed the concept of displaying the prototypical heart-beat on a screen to provide a compact visualization of the relevant information in the recorded signal. The use of visual aids was supplemented by auditory presentation tools. To differentiate between cardiac events in temporal proximity of each other, we introduced the facility to play back heart sounds at a reduced rate with preservation of the frequency content. We also constructed a version of the prototypical beat suited for listening and provided the functionality to play this back.

Chapter 5

Related Work

The problem of coming up with an acoustic-based device that is able to assist in the differentiation between pathological heart murmurs, benign murmurs and normal heart sounds is a relatively old one. Recent advances in digital stethoscopy, which allow audio signals to be captured with a degree of sensitivity sufficient to prevent loss of relevant information, have led to a flurry of activity [33, 34, 35, 36, 37, 38, 39, 40, 41]. The bulk of these efforts address goals that differ significantly from ours. In particular, there is a general trend amongst these studies to focus merely on determining whether a heart murmur exists or not, with a lack of emphasis on differentiating between benign and pathological cases. Moreover, [38] and [39] are simply observations of the time-frequency analysis of pathological heart murmurs, with no visible push towards developing automated diagnostic algorithms. Finally, [34] focusses exclusively on the examination of sounds produced by prosthetic heart valves. Since actual heart sounds are considerably different, this work is of limited applicability.

In what follows, we discuss some of the other, more relevant attempts to develop a software capable of detecting cardiac disorders [42, 45, 46, 47, 48, 49], contrasting these approaches to our system.

One of the key differences between our work and the studies that will be discussed shortly is our goal of achieving a system capable of performing automated auscultation in a completely transparent manner. Not only does this make it easier to validate the system in light of physiological knowledge, but moreover, it demystifies auscultation

and facilitates imparting this skill to others.

With the exception of the effort by Myint et al, the systems presented all employ neural nets at one or more points along the decision path. These combine the features in some unobservable manner and greatly reduce the transparency of the system. In addition, neural nets also suffer from the problem that they focus on isolating statistically relevant trends in the dataset chosen for analysis. There is no way to incorporate clinical knowledge and this makes it difficult to ensure that the combination of features chosen has scope beyond the training set.

Associated with the need for validation is the requirement that sufficient data be available to formulate and test hypothesis. Whereas in our case we were able to benefit from a collaboration with cardiologists at Massachusetts General Hospital to collect a large number of recordings, some of the studies that will be discussed shortly were not so fortunate. The extreme case is represented by Myint et al, who were only able to evaluate the performance of their classifier on a single heart murmur. A number of other studies presented also suffer from the same problem and had to be content with making the most of a limited set of data by using recordings as both training and test data. This prevented them from effectively measuring how well the error in their system had generalized[1]. Moreover, given the small number of training samples, it is reasonable to expect that the machine learning techniques used may have overfit the data. As a result, the success rates claimed in the remainder of this chapter should be viewed with a certain degree of caution.

On a closely related note, whereas the diagnostic aids we provide attempt to increase clinical confidence in the recommendations made by our system, there is no similar effort in any of the other projects. Instead, decisions must be accepted or rejected by themselves. It is expected that aids such as the visualization of the prototypical beat and playback of the slowed down heart sound will yield additional information regarding the condition of the patient's heart, beyond simply providing a classification of the underlying pathology.

[1] i.e., the extent to which the training error was representative of the actual error that would be obtained for any random input provided to the system

5.1 DeGroff et al

The work conducted by DeGroff et al [42] uses a neural net to distinguish between innocent and pathological murmurs. For each patient, a small sample of three characteristic beats is isolated by one of the investigators, and a normalized energy spectrum of the audio data is obtained by applying a Fast Fourier Transform to the heart sounds. Various spectral resolutions (1, 3, and 5 Hz) and frequency ranges (0 to 90, 0 to 150, 0 to 210, 0 to 255, and 0 to 300 Hz) are then employed as an input into the neural net, with the values leading to the most favorable results being chosen as the optimal setting for these parameters. Using this approach, false positive and false negative rates of 0% are claimed for the system. The dataset used for this project consists of 69 recordings collected from patients with innocent murmurs and those suffering from aortic insufficiency, aortic stenosis, atrial septal defect, mitral regurgitation, patent ductus arteriosus, pulmonary stenosis, pulmonary insufficiency, peripheral pulmonary stenosis and ventricular septal defects. The error rate was obtained using a leave-one-out cross-validation scheme [43].

An important difference between our approach and the work of DeGroff et al is that our system is completely automated. More specifically, no user input is required at any stage beyond the recording the original signal. This is not the case here, with characteristic beats being isolated manually. This human provided information greatly simplifies the problem at hand. Recognizing which beats are characteristic requires a highly trained individual with the ability to detect the disorder the system is trying to diagnose.

In addition to this, DeGroff et al also focussed exclusively on detecting moderate murmurs. These cases typically have higher energy and are easier to identify. For our purposes, we attempted to develop a system that was also able to address the more interesting problem of diagnosing mild cardiac disorders.

Another dissimilarity between our work and that of DeGroff et al is that we focus on diagnosing heart disorders in adults. The study presented here addresses the problem of detecting and classifying pediatric murmurs alone. This tends to be an

easier problem since murmurs can generally be heard more clearly in children. This is due to the fact that children's hearts beat faster than adult hearts, leading to increased turbulence. This produces a more distinguishable murmur. In addition, the chest walls in children typically have less fat and muscle than in adults. This makes it easier to hear pediatric murmurs than those in adults [44].

DeGroff et al also differ from our work in that they do not adopt a clear time-frequency approach. Given the fact that heart disorders such as MVP lead to changes in frequency content of the signal localized at different points in the cardiac cycle, a time-frequency approach would appear to be physiologically motivated. Instead, the work discussed here simply calculates the normalized energy spectrum of the entire signal without any consideration of when changes in frequency take place.

5.2 Shino et al

Shino et al [45] also employ neural nets in their approach to detecting cardiac disorders. After segmenting the recorded signal into systolic periods, a neural net is employed as a first step to recognize the presence of murmurs. If a murmur is found, a second neural net is used to classify the detected murmur as being either pathological or benign, utilizing spectral analysis.

In more detail, the systolic period is divided into 10 consecutive bins and the amplitude is normalized so that the variance of each bin is equal to unity. Then, the average variances of the systolic heart sound in each bin over all available beats and variances of the first and second heart sound are fed into a neural net to detect the presence of a systolic murmur. If a murmur is found, then to further classify it into a pathological or benign murmur, a 128 point Fast Fourier Transform is applied to the systolic portion and averaged over several available beats. The mean frequency and the peak value are then used for murmur separation. A false positive rate of 6.7% and a false negative rate of 10.6% is claimed for a dataset consisting of 44 pathological murmurs, 61 innocent murmurs and 36 normal cases. A four-fold cross-validation approach [43] is used to determine the error rates provided.

One of the key differences between our approach and the one presented here is that Shino et al choose to assign labels to their dataset based on doctors looking at the signals rather than by means of an echocardiogram, widely used by physicians as the "gold standard" for diagnosing MVP. Although a case might be made for such an approach on the basis of the fact that a system should not attempt to classify files without conceivable audio diagnostic information as belonging to patients with heart murmurs, we choose to avoid this due to the possibility that even in recordings which seem normal to humans there may very well be features that could be recognized by a computer and used to make a correct decision. Focussing solely on cases where a murmur can be heard clearly would have restricted our system to performing at best as well as its human counterparts. It is well known that this is not good enough[2].

Shino et al also focus on detecting and classifying murmurs in children. Once again, this is a deviation from our goal of developing a system capable of diagnosing heart disorders in adults.

5.3 Barschdorff et al

Barschdorff et al [46] make use of a two-step diagnostic process as well. As a first step, a neural net is used to detect the presence of a murmur. A second neural net then classifies the particular disease leading to that murmur.

In order to achieve this, the signal is first segmented into systolic and diastolic intervals according to a synchronously recorded EKG. This is achieved by finding the QRS complex of the EKG wave and then feeding the heart rate, patient age and gender into a neural net to calculate the duration of systole. Artifact free chunks of the signal are then obtained by removing the segments that have a variance greater than some threshold. Following this, features are calculated by applying the wavelet transform. These are then used to detect whether a murmur exists or not by feeding them into

[2]DeGroff et al also use human experts to label files. However, echocardiographic confirmation are performed on 100% of the patients with pathological heart murmurs and 70% of the patients with innocent heart murmurs. This greatly reducing the effect described here.

a neural net. If a murmur is found to be present, a second neural net classifies it as being pathological or benign. On a sample set of 61 patients, a false positive rate of 0% and a false negative rate of 12.5% is claimed for the first neural net used to detect the presence of a heart murmur. 28 of these had no heart disease or lacked a considerable murmur. The remaining 33 patients had recognizable noise phenomena, due to diagnosed heart diseases. The error rates were obtained by training on half the patients (14 without heart murmurs and 17 with pathological murmurs) and testing on the remaining ones. Due to the lack of sufficient data to divide the signals into a training and test set, no figures were provided for the success of the second neural net attempting to discriminate between innocent and pathological murmurs.

Barschdorff et al differ significantly from our system in that they treat the absence of a considerable murmur as a normal case, irrespective of the echocardiogram findings. For reasons similar to those presented earlier while discussing Shino et al, we avoid adopting such an approach since it would restrict our system to perform only as well as its human counterparts. Moreover, the task of detecting significant murmurs (due to moderate and severe cardiac disease) was considered to be a drastic oversimplification of the problem at hand. Borderline cases are the ones most often misdiagnosed and were therefore thought to be of most interest. As a result, we label recordings according to their echocardiogram classifications rather than any human perception of cardiac disorder. Our system then attempts to detect the presence of heart disease accurately, irrespective of the severity of the condition.

As was the case with DeGroff et al and Shino et al, Barschdorff et al also diverge from our approach in that they exclusively focus on pediatric murmurs. Once again, this constitutes an important deviation from our approach geared towards performing auscultation on adults.

5.4 Reed et al

Reed et al [47] also adopt an approach to detecting cardiac disorders where neural nets are a key component of the system. Heart sounds are first segmented by hand

into sample segments, each consisting of a single heart-beat cycle. Each segment is transformed using a 7 level wavelet decomposition, based on a Coifman 4th order wavelet kernel (chosen due to its relative symmetry and fast execution). The resulting transform vectors, 4096 values in length, are reduced to 256 element feature vectors by discarding the 4 levels with shortest scale. In addition to substantially simplifying the neural network in the classifier which follows, this step also reduces noise. The magnitudes of the remaining coefficients in each vector are calculated, then normalized by the vectors energy. Finally, each feature vector is classified using a three layer neural network (256 input nodes, 50 hidden nodes, and 5 output nodes). A 100% accuracy is claimed. The dataset for this project consisted of one normal patient and one patient for each of the heart disorders considered (mitral valve prolapse, coarctation of the aorta, ventricular septal defect and pulmonary stenosis). Due to the lack of sufficient data, the training set and the test set were the same.

One of the main dissimilarities between our system and the work of Reed et al is that Reed et al require heart sounds to be segmented manually. This is a significant deviation from our system, which functions in a completely automated manner without requiring any human input beyond the recording of the original signal.

5.5 Watrous et al

Unlike the majority of the efforts to develop software capable of detecting heart disorders using audio data, the approach taken by Watrous et al [48] incorporates physiological information.

The system makes use of a wavelet based bank of correlators to match-filter recorded signals to known templates for S1, S2, S3, S4, clicks and heart murmurs. More specifically, the signal is passed through a set of correlators where the wavelets are chosen to specifically match the different heart sounds. Some of these sounds may match with an arbitrary subset of the chosen wavelets and in particular the heart sounds may vary from person to person and show variations from heart-beat to

heart-beat in an individual. As a result, the outputs from the correlators are treated as features and are passed to decision blocks consisting of neural nets, which then segment the signal into the various heart sounds, clicks and murmurs. There is also a block to detect a noisy input signal. The outputs from these are then passed into a second layer of decision blocks which attempt to classify the murmur.

This work is in progress at the moment and no results are provided.

The study conducted by Watrous et al seems to focus primarily on segmenting the recorded signal rather than classifying heart murmurs. Mathematical models are developed for S1, S2, S3 and S4, but there is no such described attempt to do so for heart murmurs. This ties in to the broader issue that given the vast differences in murmur signatures (even for murmurs corresponding to the same pathology) it is difficult to develop templates for each of these. The idea for adopting a match-filter approach was inspired by the communications problem of detecting a symbol in noise, but this does not readily apply to the problem at hand since the symbol in this case may vary drastically and does not possess a well-defined form.

5.6 Myint et al

Finally, another interesting attempt at developing a system such as the one under consideration is the study conducted by Myint et al [49]. This first segments the signal into diastolic and systolic portions making use of the fact that the peaks in the signal due to diastole will generally be more spaced out than in systole (and when heart rate increases, the duration of diastole will decrease whereas that of systole will remain relatively fixed). A single murmur is then isolated and is subjected to a time-frequency analysis by calculating its spectrogram. The peak magnitude, frequency at peak, average magnitude and the standard deviation of magnitude are then computed for each localized spectrum. From these, the averages and standard deviations of the peak magnitudes and their frequencies across the murmur duration are found. Small standard deviations and "reasonable" average values indicate mitral regurgitation. This achieves a detection rate of 100% when tested on a single murmur isolated for a

patient with mitral regurgitation.

In a manner identical to DeGroff et al, the work of Myint et al also differs from our system in that characteristic beats must be isolated manually. Once again, doing so drastically simplifies the problem since recognizing which beats are characteristic requires the ability to detect the disorder to start out with. Assuming this functionality is available results in an incomplete solution.

5.7 Summary

Tables 5.1 and 5.2 summarize the key differences between our work and the various approaches discussed in the previous sections[3].

Criterion	Our	DeGroff	Shino	Barschdorff	Reed	Watrous	Myint
Supported by Audio-Visual Aids	Yes	No	No	No	No	No	No
Fully Automated	Yes	No	No	Yes	Yes	Yes	No
Time-frequency Approach	Yes	No	Yes	Yes	Yes	Yes	Yes
Adult Auscultation	Yes	No	No	No	Yes	Yes	Yes

Table 5.1: Comparison of Our System with Previous Approaches to Automated Auscultation

Criterion	Our	DeGroff	Shino	Barschdorff	Reed	Watrous	Myint
Patients in Dataset	51	69	141	61	5	N/A	1
Echocardiogram Used as Gold Standard	Yes	Yes	No	No	Yes	N/A	Yes
Training and Test Sets Disjoint	Yes	Yes	Yes	Yes	No	N/A	Yes

Table 5.2: Comparison of Validation Strategies for Our System and Previous Approaches to Automated Auscultation

[3]As described in Section 5.5, Watrous et al do not provide any test results for their study.

Chapter 6

Summary, Future Work and Conclusions

6.1 Summary

Cardiac auscultation is a quick, inexpensive means of detecting mitral valve prolapse (MVP). However, it is also highly error prone. The skill requires years of training and imposes exacting demands on the human ear. It would be advantageous if the benefits of auscultation could be obtained with a reduced learning curve, using equipment that is low-cost, robust and easy to use. To do this, we propose a new detection paradigm that interposes software running on a personal computer between the doctor and the patient.

The main goal of this project was designing, implementing and evaluating a software application system capable of performing automated detection of MVP. In addition to making a diagnosis, the software supports its decision with effective presentation tools.

Our system divides the problem into five major components; *beat selection* to segment the heart signal and obtain characteristic heart-beats that should be analyzed further, *time-frequency analysis* to produce features needed for classification, *beat aggregation* to augment this set by additional features that combine information across multiple heart-beats, a *decision mechanism* to produce a diagnosis, and a suite of

audio-visual diagnostic aids geared towards providing information beyond a yes-no classification to doctors.

The process of beat selection starts off by fusing audio information with the EKG to segment the heart signal. The position of the QRS complex is located using existing algorithms and a new algorithm is proposed to perform S2 detection. Once the systolic segments have been isolated in this manner, they are screened for noise. Noise-free heart-beats that meet a set of physiological criteria are then analyzed further.

The systolic sections obtained from the beat selection subsystem are subjected to a time-frequency decomposition using a filter bank. The frequency components are then aggregated in a number of different ways to reveal diagnostic information and make the system more robust to the presence of random and systematic noise in the recorded audio signal.

The combination of frequency bands is supplemented by the aggregation of the selected beats. In order to observe the characteristic trends persisting amongst the majority of the beats, we develop a mechanism to pool multiple beats together. This allows for the creation of a single representative heart-beat for the patient called the "prototypical heart-beat". The time-frequency decomposition of the prototypical beat produces additional features that can be used for detecting cardiac disease.

The decision mechanism uses the output of the time-frequency analysis and beat aggregation subsystems to make a diagnosis. We support two different detection schemes. The first one considers each beat separately and combines the individual classifications for each beat to make a final decision. The second examines the information encapsulated within the prototypical beat and produces a classification for this single beat.

Since the prototypical heart-beat provides a compact representation of the acoustic signal recorded, it is also a critical component of the toolbox of audio-visual diagnostic aids. We provide the functionality to both view and listen to a "typical" heart-beat for the patient. In addition, we allow users to play back slowed down heart-sounds without distorting the high-frequency content content of the signal.

One of the most important challenges associated with developing our system was

the decision to gear this project towards outperforming human experts. The job performed currently by primary care practitioners is far from satisfactory. Our goal was to match the "gold standard" diagnosis, which is based on an echocardiogram. This approach greatly increased the difficulty of our work; it is far easier to communicate with clinicians and understand specifically what they are listening to than to develop a detector based on the vaguely correlated factor of there being "energy before S2"[1]. As a result, we had to first dedicate considerable time to analyzing the signal and determining what the relevant diagnostic information was, before actually developing a classifier that made use of it. This led to the need to design a suite of audio-visual tools before work on the actual detector could proceed. Since these tools were powerful enough to allow us to discriminate the presence of heart disorders (despite our non-clinical backgrounds) they were considered to be increasingly helpful to primary care physicians.

Closely related to this issue was the need to develop a detector that incorporated the observed diagnostic information. For example, we were required to translate the visual trends associated with prototypical heart-beat plots for MVP patients into a logical decision mechanism that could be used to perform classification. Moreover, since false negatives for our system correspond to patients diagnosed incorrectly to be non-MVP (some of whom might require additional treatment), we had to ensure that this error rate remained as low as possible. To do so, the decision mechanism had to place a greater emphasis on avoiding false negatives than reducing false positives. Determining how to achieve this and figuring out the extent to which trading false negatives for false positives is acceptable were difficult questions we had to answer during the course of our work.

In terms of the time-frequency analysis of systolic segments, a key challenge was to determine precisely how to decompose the recorded signal into its different frequency components (i.e., using a Short Time Fourier Transform, filter banks or wavelets). Once this was complete, we needed to aggregate information from across different frequency bands in ways that would reveal characteristic trends associated with MVP.

[1]Which is the only information we had available to us when we started this project.

This required figuring out which bands to aggregate and deriving a strategy on how this aggregation should deal with bands that have widely differing amplitudes. In addition to this, we needed to determine the number of frequency components to divide the signal into. The requirement of increasingly fine granularity in frequency argued in favor of many narrow frequency bands. Conversely, the need for minimizing the effect of noise localized in frequency suggested the use of a small number of wider bands.

Another problem that we needed to resolve was that of segmenting heart signals. Although QRS detection is well studied, the task of locating S2 is considerably less frequently addressed. One option available was to use the average length of systole to predict the location of S2. This approach did not work well since the length of systole deviates sharply from its mean value. Consequently, an algorithm that focussed on actually detecting the presence of S2 was needed. This algorithm had to be sufficiently robust to avoid incorrectly declaring artifacts due to noise as S2.

Finally, associated with this issue was the challenge of characterizing noise in the recorded signals and determining when a beat should be discarded as being noisy. This was particularly difficult since deviations from normal heart activity could be the result of MVP and not noise. There was also a need in general to figure out which beats were "better" than others. Since a key goal of our system is its explainability, we had to derive a set of physiological criteria that could be used a metric to weigh heart-beats.

When tested on 51 patients, 21 of whom suffered from MVP, the system displayed a false positive rate of 10%. This is a considerable improvement over the false positive rate of over 80% associated with primary care physicians. Associated with this was a false negative rate of 5%. Although the false negative rate for doctors is less well documented, our guess is that our system outperforms humans on this front as well. This is based on the observation that our system detected MVP's that skilled cardiologists did not hear.

It is difficult to quantify the performance of the audio-visual aids developed during the course of this project. However, we have obtained extremely positive feedback

on their use from doctors and are encouraged by their ability to allow people who cannot detect the presence of MVP acoustically to do so visually.

6.2 Future Work

Our project provides a foundation that can be expanded upon in a number of interesting ways. Some possibilities include:

- Expanding the use of our system to heart pathologies other than mitral valve prolapse
- Developing an alternative beat selection scheme that weights beats rather than simply admitting them or not
- Improving upon the physiological criteria used to admit beats
- Finding better ways to characterize noise in the recorded signals
- Making use of adaptive signal processing techniques to remove the noise from signals
- Employing an array of stethoscope sensors to capture data simultaneously for multiple sites
- Augmenting the decision criterion with additional features

We are also eager to clinically deploy our system, and are in the process of developing an industrial-quality version of our software. This will be deployed at Emory University in the summer of 2003 to assist medical students in learning how to perform auscultation. We hope to extend this in the near future with the use of our software in the offices of primary care physicians. This is expected to yield considerably more data, allowing us to test our system more exhaustively and refine our hypotheses.

6.3 Conclusions

Our system provides a promising approach to diagnosing at least one important class of cardiac abnormalities. It drastically reduces the error rate associated with primary care physicians, and should minimize the number of incorrect referrals to cardiologists. This would lead to considerable savings of time and money, and spare patients misdiagnosed as suffering from MVP substantial emotional anxiety. In addition, we also expect the system to significantly decrease the number of false negatives. Since these cases often require treatment, our software allows cardiac disease to be detected early on, before the the condition worsens.

The process of combining automated auscultation with the use of diagnostic aids is of great importance. Not only does it allow the system to be validated, but, more significantly, it provides users with information beyond a single-bit classification. This permits doctors to focus on those aspects of the abnormality that vary between patients. These include determining the extent of the disease and the morphology of the heart murmur. The information obtained can be used to develop a better understanding of events taking place at the anatomical level for each patient.

The new detection paradigm proposed by this project offers significant improvement to the way physical exams are currently conducted and promises to be of immense value as a teaching tool in medical schools.

We expect the suite of audio-visual diagnostic aids to assume a central role in assisting primary care physicians in performing routine physical examinations. By making use of the functionality provided, we expect doctors to perform auscultation in a more reliable manner. This can be done without any significant changes to the way physical examinations are conducted at present. No new equipment is required[2] and there are no additional maneuvers to perform. This limits the training needed to make use of our tools to simply understanding how to interact with the user interface of the software. Moreover, since the relevant signals can be collected during the course

[2]Our system only makes use of an electronic stethoscope with EKG leads and a personal computer, both of which exist in the offices of primary care physicians today.

of a normal checkup, there is no increase in the time taken to conduct the physical exam itself. The additional time required to analyze the signals using the techniques proposed is incremental and the physician may choose to employ only a subset of the aids he or she finds most helpful personally. Also, since the signals are available electronically, they can be saved to disk. This has three major advantages:

- The signals can be analyzed later
- The signals can be compared to previous recordings for the patient or recordings for other similar patients
- The signal can be replayed several times to listen to events of interest

In addition to this, another advantage of this approach is that the signals recorded can be duplicated and referred to experts for more help (much the same way as X-ray images are presently viewed). This leads to a broader idea; the new paradigm for physical examinations does not necessitate the presence of a primary care physician. The signals can be analyzed automatically using the system developed, with ambiguous cases referred to cardiologists. This could be done remotely and would be particularly useful in distant areas and those parts of the world where skilled medical help is not available.

Because of the transparent nature of the tools developed, we anticipate that the audio-visual diagnostic aids will be used extensively in teaching environments. The ability to play back slowed down heart sounds allows instructors to more precisely indicate the events of interest in the audio data. In addition to this, the visual display of the prototypical beat permits audio information to be correlated to what can be seen. Since it is generally easier to recognize heart disorders visually than acoustically, this promotes a deeper understanding of the audio signals resulting from cardiac disease. The playback of the enhanced audio-prototypical beat also indicates to students what the characteristic high-frequency signal associated with disorders such as MVP sounds like. Finally, the ability to replay the recordings provides students with an opportunity to listen to heart sounds multiple times and learn to recognize them more accurately.

Our study holds great potential for improving contemporary health care. We are excited by our achievements and envision significant changes to the way auscultation is conducted at present. The use of computers to assist in the interpretation of heart sounds promises to lead to an inexpensive, quick and increasingly reliable means of diagnosing cardiac disorders in the years to come.

Appendix A

Heart Physiology

The following serves as a brief overview of the physiology of the heart. The anatomy of the heart is first presented in Section A.1. This is followed by a detailed discussion of the cardiac cycle in Section A.2. Finally, we conclude with a description of the electrical (Section A.3) and acoustical (Section A.4) activity associated with the heart.

A.1 Anatomy

The human heart [50, 51, 52, 53] is a muscular organ lying in the thoracic cavity behind the lungs as shown Figure A-1. It is composed of cardiac muscle tissue that is able to contract and relax rhythmically, and by doing so, pumps blood throughout the body.

The heart is divided into two sides separated by a middle wall known as the septum. The right side of the heart receives oxygen-deficient blood from the various regions of the body and delivers it to the lungs. In the lungs, oxygen is absorbed by the blood and the oxygenated blood is then received by the left side of the heart, which delivers it to the rest of the body. This is illustrated in Figure A-2.

The heart has four separate compartments or chambers. The upper chambers on either side of the heart, which are called atria (singular atrium), receive and collect the blood coming to the heart. The right atrium receives blood from the inferior and superior vena cava and the left atrium is supplied by the pulmonary veins. The atria

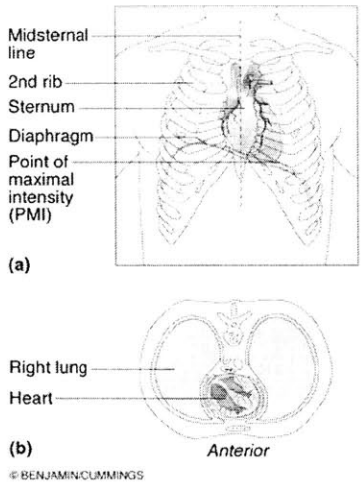


Figure A-1: Position of Heart in the Thoracic Cavity

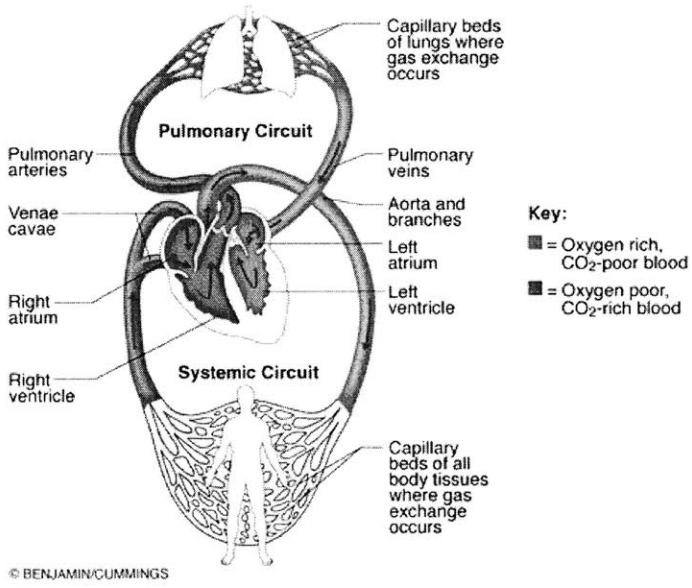


Figure A-2: The Human Circulatory System

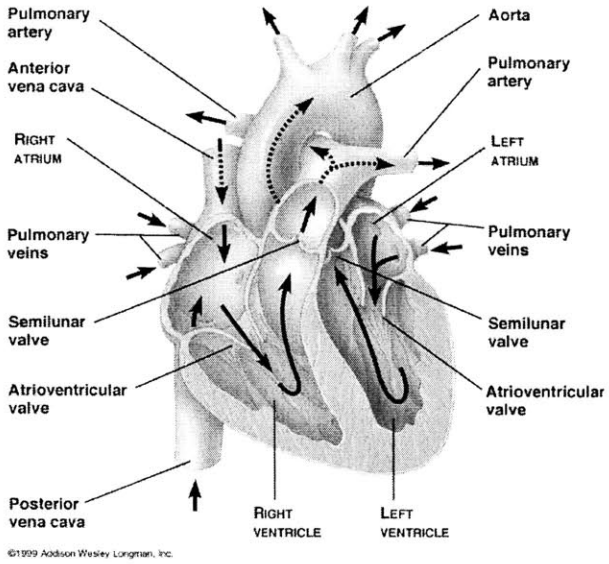
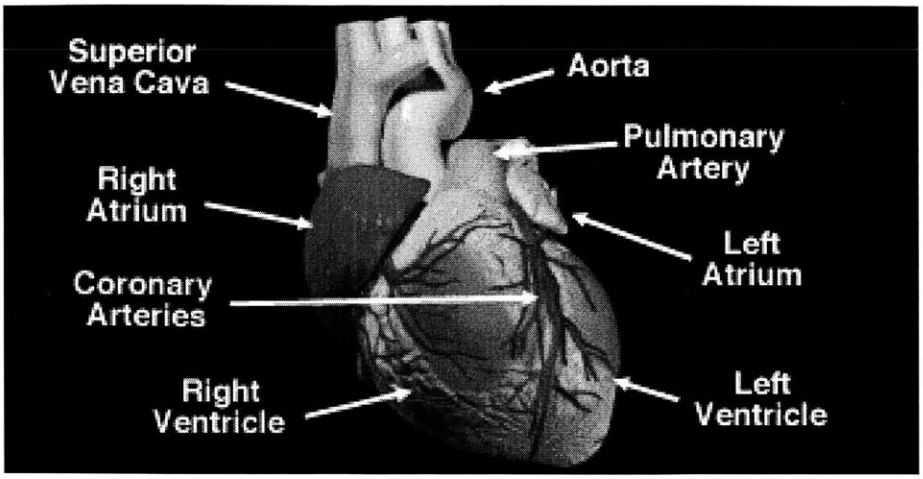


Figure A-3: Longitudinal Section Through the Human Heart

then deliver blood to the powerful lower chambers, called ventricles, which pump blood away from the heart through powerful, rhythmic contractions. Blood leaves the right ventricle through the pulmonary artery and similarly, the left ventricle is connected to the aorta. The first branches of the aorta are small arteries known as the coronary arteries. These supply blood to the heart itself.

Figures A-3 and A-4 provide visual representations of the human heart.

The atria are separated from their respective ventricles by the atrioventricular (AV) valves. These prevent the back flow of blood from the ventricles to the atria when the ventricles contract. The valve that separates the right atrium from the right ventricle is known as the tricuspid valve, whereas the bicuspid or mitral valve separates the left atrium from the left ventricle. Two more valves, the semilunar (SL) valves, help separate the two ventricles from the arteries to which they lead. The valve separating the right ventricle from the pulmonary artery is called the pulmonary valve, whereas the valve separating the left ventricle from the aortic is known as the aortic valve. These valves prevent the back flow of blood from the arteries into the ventricles. Collectively, there are a total of four valves in the heart with no valves



Reproduced from http://home.earthlink.net/ avdoc/infocntr/htrhythm/hrphysio.htm

Figure A-4: Three Dimensional View of the Human Heart

being present to prevent back flow from the atria into the veins [1].

Figure A.1 provides a transverse section through the heart showing the position and normal opening and closing of the atrial and semilunar valves.

A.2 Cardiac Cycle

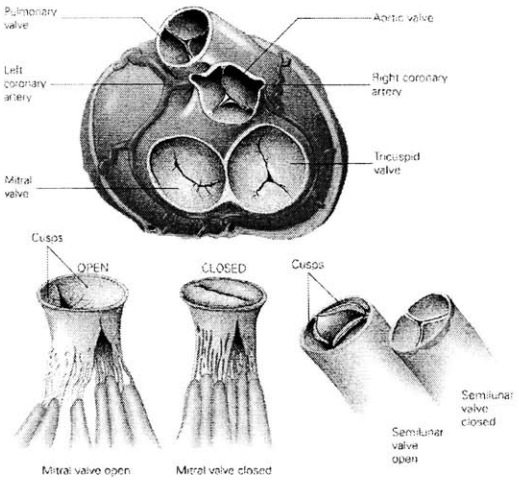
The periodic pumping action of the heart that results in the unidirectional flow of blood through the human body is known as the cardiac cycle. At rest, the heart beats about 70 times per minute, with each beat having a corresponding duration of approximately 800 ms. The heart rate and the duration of each beat vary significantly between people and may also have different values for the same person (depending on the activity being performed).

At the highest level, there are two main stages in a heart-beat: *systole* and *diastole*. Systole may be further divided into atrial systole and ventricular systole[2]. This is illustrated in Figure A-6.

At a more detailed level, diastole and ventricular systole may be subdivided further

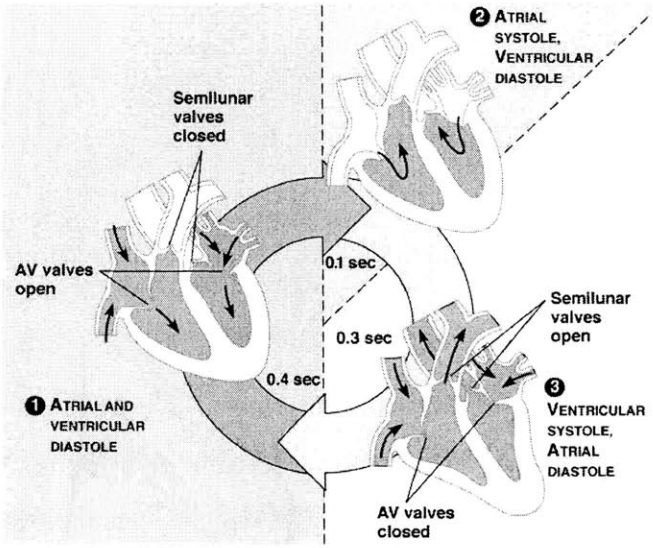
[1]Instead, the veins themselves have a series of valves along their length to prevent the backflow of blood

[2]The term systole is often used loosely to refer to ventricular systole alone



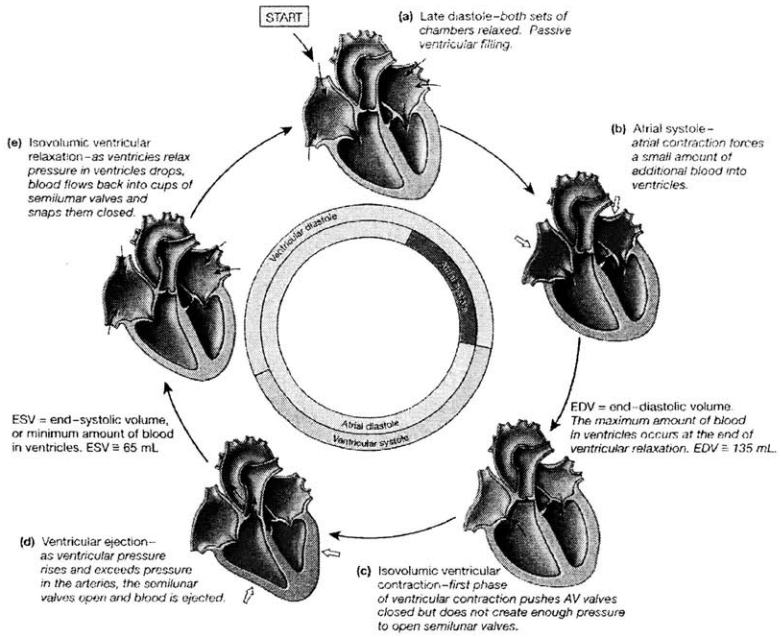
Reproduced from http://home.earthlink.net/ avdoc/infocntr/htrhythm/hrphysio.htm

Figure A-5: Transverse Section through Heart Showing Normal Functioning of Heart Valves



Reproduced from http://meds.queensu.ca/medicine/physiol/undergrad/phase2/Lecture5f.ppt

Figure A-6: Diastole, Atrial Systole and Ventricular Systole



Reproduced from http://207.233.44.253/wms/reynolmj/lifesciences/lecturenote/bio3/Chap23.ppt

Figure A-7: Detailed Cardiac Cycle

as shown in Figure A-7.

A.2.1 Atrial Systole

During *atrial systole* the atria contract, pushing a small fraction of their volume of blood into the ventricles to fill them to maximum capacity. Blood arriving at the heart can no longer enter the atria during this period, and this results in blood flowing back up the jugular vein, causing the first discernible wave in the jugular venous pulse.

A.2.2 Ventricular Systole

The first stage of ventricular systole is known as the *isovolumic ventricular contraction*. During this part of the cardiac cycle, the thick muscular walls of the ventricles contract raising the pressure in the ventricles and causing the mitral and tricuspid (AV) valves to be closed once this pressure exceeds that of the atria. As the ventricles continue to contract isovolumetrically (i.e., without a change in the volume of blood), the pressure inside them increases, approaching the pressure in the aorta and

pulmonary artery.

Once the pressure in the ventricles exceeds that in the aorta and pulmonary artery, *ventricular ejection* begins. The semilunar valves open, allowing blood to exit the ventricles. This causes the volume (and correspondingly the pressure) in the ventricles to decrease rapidly. As more blood enters the arteries, pressure there builds until the flow of blood reaches a peak. Right ventricular contraction pushes the tricuspid valve into the atrium and increases atrial pressure, creating a small wave into the jugular vein that is normally simultaneous with the carotid pulse.

A.2.3 Diastole

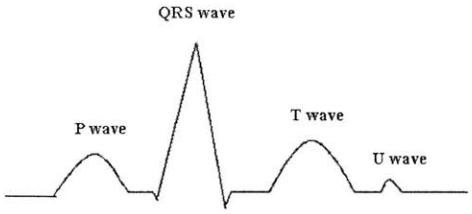
As the pressure continues to fall in the ventricles it eventually falls below the pressure in the arteries causing blood in the arteries to start flowing back towards the ventricles. This causes the semi-lunar valves to close and is known as *isovolumic ventricular relaxation.* Throughout this period and ventricular systole, the atria, which are in diastole, continue to fill with blood, causing atrial pressure to rise gradually and leading to the back flow of blood after it hits the closed AV valve. This causes the second discernible wave of the jugular venous pulse.

Isovolumic ventricular relaxation is followed by *late diastole.* During this stage the atrial pressure exceeds that in the ventricles, causing the mitral and tricuspid (AV) valves to be pushed open. Ventricular volume increases rapidly as blood flows passively from the atria into the ventricles.

A.3 Electrical Activity

A.3.1 Electrocardiogram

The electrocardiogram is a record of the electrical activity occurring in the heart during one cardiac cycle. The EKG tracing is composed of three distinct deflections or waves and well defined intervals as shown in Figure A-8. The physiological relevance of these will be explained in the next subsection.



Reproduced from http://faculty.ucc.edu/biologypotter/Heart_Physiology/sld011.htm

Figure A-8: Sample EKG Tracing

A.3.2 Conduction System of the Heart

The heart is innervated by the autonomic nervous system, which increases or decreases heart rate. It does not however initiate contraction as is the case with the other muscles in the human body. Instead, the heart is able to beat without any direct stimulus from the nervous system. This is due to the fact that the heart has an intrinsic regulating system called the *conduction system.* This is composed of specialized muscle tissue and is responsible for generating and distributing the action potentials that stimulate the cardiac muscle fibres to contract. The conduction system tissue is found in the sinoatrial (SA) node, atrioventricular (AV) node, atrioventricular (AV) bundle (bundle of His [3]), bundle branches, and conduction myofibres (Purkinje fibres). All cardiac muscle is termed as being capable of *self-excitation.* This means that cardiac muscle is able to spontaneously and rhythmically generate the action potentials that result in contraction of the muscle. The normal resting rate of self-excitation of the sinoatrial node is about 70 times per minute in adults and is much greater than the rest of the conduction system. Hence, as shall be explained shortly, the SA node drives the rest of the cardiac muscle and is called the pacemaker.

As shown in Figure A-9, the onset of the depolarization in the heart that leads to contraction begins within a focus of pacemaker cells found at the upper right border of the heart, just below the opening to the superior vena cava. These cells are responsible for the initial activation and are collectively known as the sinoatrial node. Since the SA node spontaneously generates action potentials faster than other

[3]Named after theSwiss physician Wilhelm His, Jr.

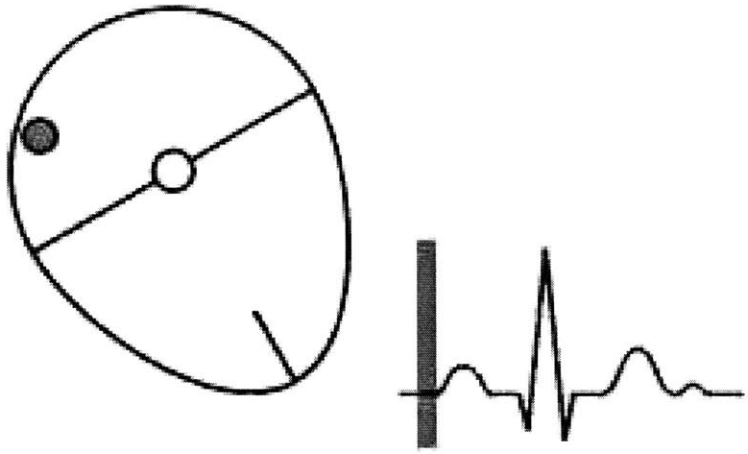


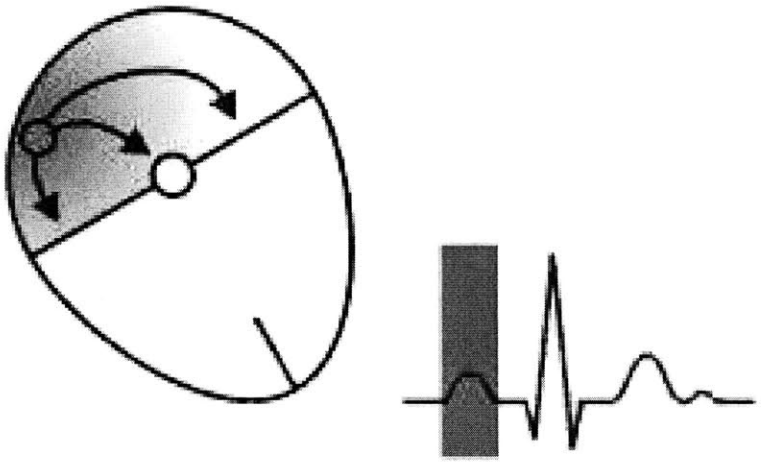
Figure A-9: Sinoatrial Nodal Impulse Origin

components of the conduction system, nerve impulses from the SA node spread to the other areas and stimulate them so frequently they are not able to generate action potentials at their own inherent rates. Thus the faster SA node sets the rhythm for the rest of the heart. The rate set by the SA node may be altered by nerve impulses from the autonomic nervous system but these impulses do not initiate contraction.

The action potential generated by the SA node travels downwards, leftwards and posteriorly as a wave through both atria, depolarizing each cell in turn. This propagation of charge is responsible for the P wave that appears on the electrocardiograph and is illustrated in Figure A-10.

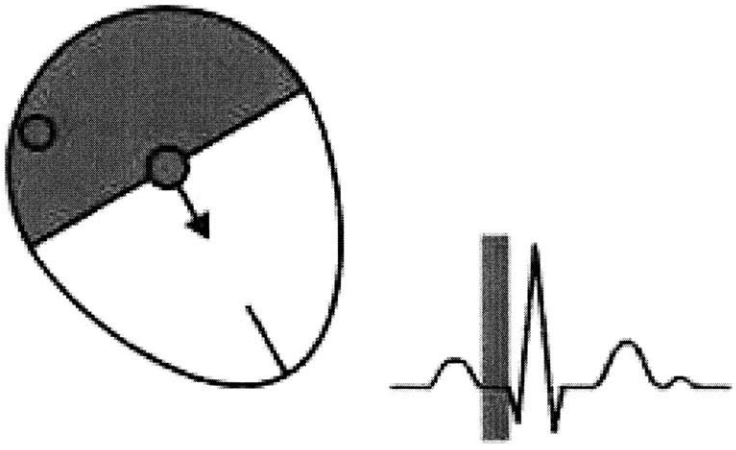
Eventually, the depolarization reaches the AV node (Figure A-11) near the center of the heart at the top of the interventricular septum. This is responsible for conducting the impulse from the atria to the ventricles and introduces a delay in the propagation of the action potential while doing so. By virtue of the fact that the atrioventricular node is small no depolarization voltage is recorded and an isoelectric PR segment is seen on the electrocardiograph.

The depolarization proceeds down the septum, along the bundle of His as shown in Figure A-12, before splitting to follow the left and right bundle branches. From there it proceeds onwards to the conduction myofibres (Purkinje fibres), which are



©Cardionetics

Figure A-10: Atrial Depolarization



©Cardionetics

Figure A-11: Atrioventricular Nodal Depolarisation and Activation of Bundle of His

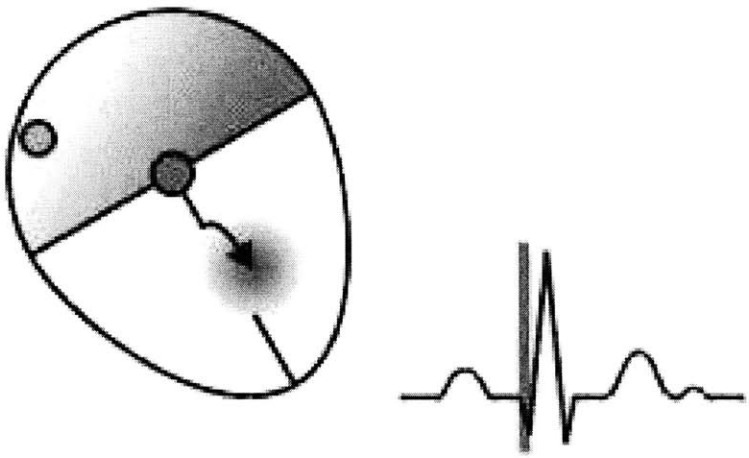


Figure A-12: Septal Depolarization

responsible for distributing the action potential and thus depolarize the myocardial cells of the ventricles. The left bundle branch is activated first, with the septal depolarization proceeding from left to right. This gives rise to a small negative deflection within the electrocardiograph called the Q wave. As the impulse travels down the septum the atria and sinoatrial node start to polarize.

The wave of depolarization continues down the septum into the ventricular free wall. The left ventricular wall mass is significantly greater than the right and as a result of this, the mean vector of depolarization of the ventricular free wall is to the left. Figure A-13 illustrates this.

The depolarization takes place from the endocardium to the epicardium (Figure A-14) and is responsible for the R wave. Atrial, atrioventricular nodal and bundle of His polarization continue but are obscured by the higher action potential in the ventricles. The rim of the ventricular muscle below the atrioventricular groove is the last to be depolarized. The direction of depolarization leads to the S wave.

The depolarization of all the ventricular muscle cells is followed by an isoelectric ST segment that corresponds to the plateau of the action potential of all fibres. Figure A-15 provides a visual representation of this.

Polarization occurs when the membrane potential returns to the baseline. This

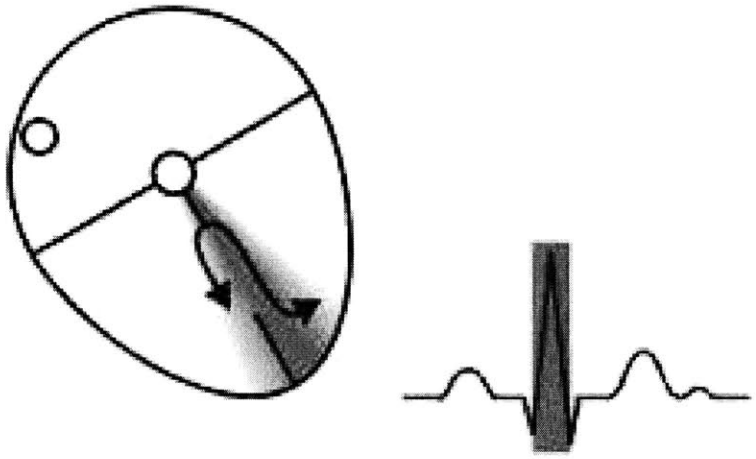


Figure A-13: Early Ventricular Free Wall Depolarisation

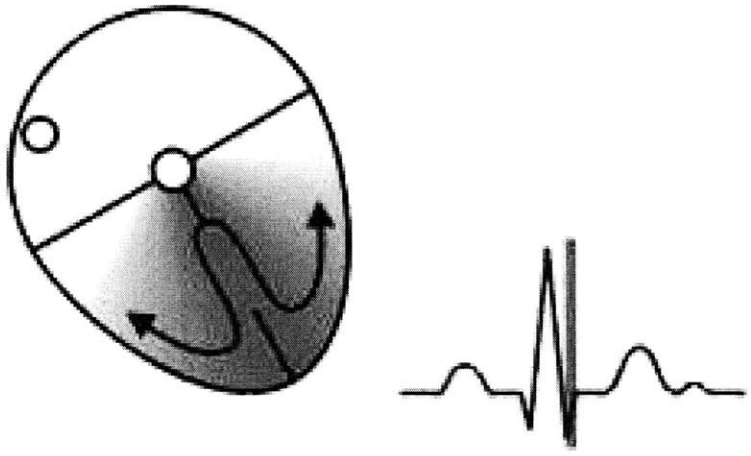


Figure A-14: Late Ventricular Free Wall Depolarisation

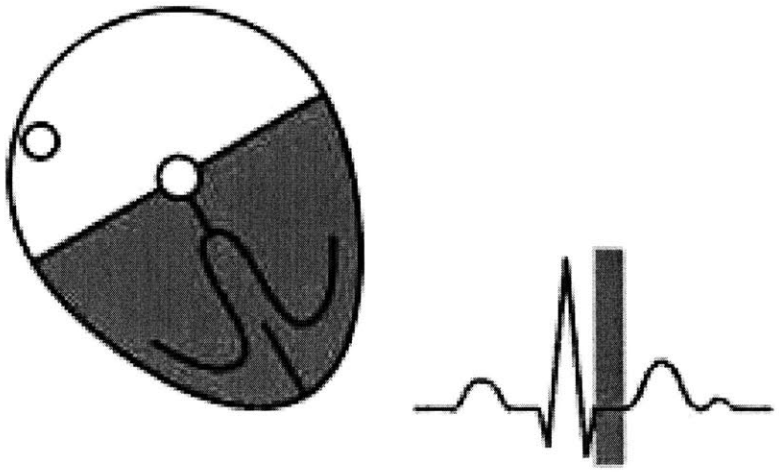


Figure A-15: Ventricular Systole

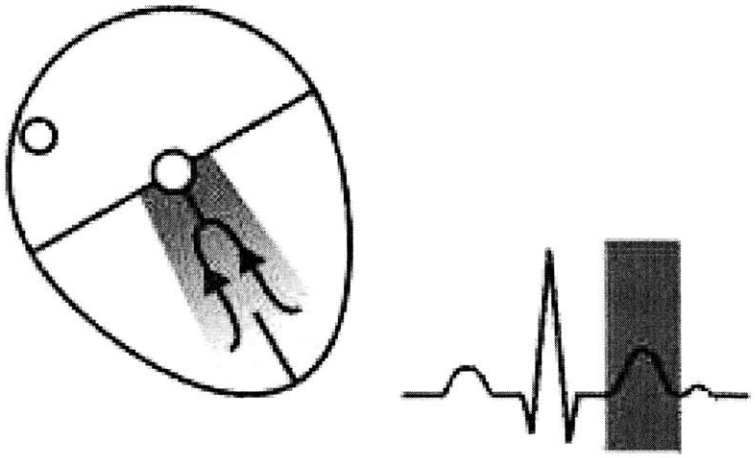


Figure A-16: Ventricular Systole Wall Polarization

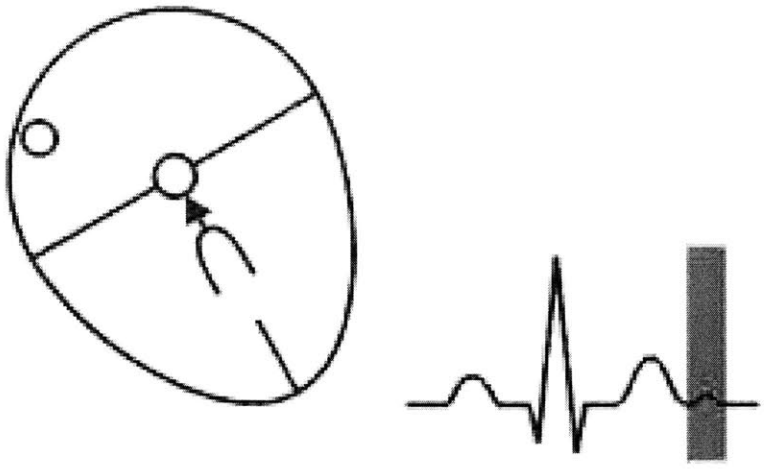


Figure A-17: Bundle of His Polarization

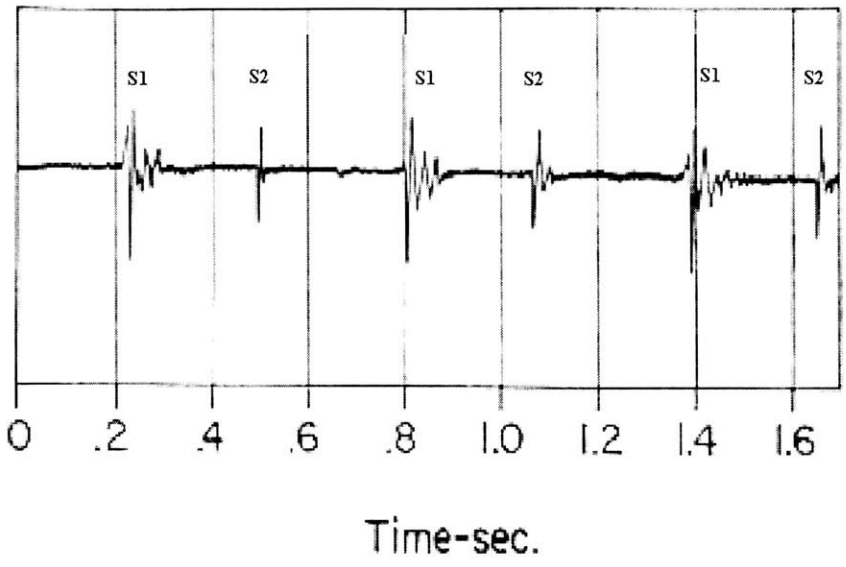
leads to a deflection in the same direction as depolarization because both polarity and direction are negated, causing the T wave as shown in Figure A-16.

The bundle of His has the greatest capacitance within the conduction cycle and acts as the final part of the heart to polarize. The positive deflection caused by this is so small that it is often unclear. When detected it is referred to as the U wave (Figure A-17).

A.4 Acoustical Activity

A.4.1 Phonocardiogram

The phonocardiogram is a record of the acoustical activity occurring in the heart during one cardiac cycle. The phonocardiogram tracing is composed of two major heart sounds as shown in Figure A-18. The physiological relevance of these will be the focus of the next subsection.



Reproduced from http://meds.queensu.ca/medicine/physiol/undergrad/phase2/Lecture5f.ppt

Figure A-18: Sample Phonocardiogram Tracing

A.4.2 Heart Sounds

The most obvious of the sounds [54] associated with the normal function of the heart are the first and second sounds, or S1 and S2. These demarcate systole from diastole.

S1 is the "lub" sound, which marks the approximate beginning of systole and is generated when the increase in intraventricular pressure during contraction exceeds the pressure within the atria, causing a sudden closing of the AV (tricuspid and mitral) valves. This occurs slightly after the QRS complex in the EKG.

S2 is the "dub" sound, which occurs at the end of systole, when the ventricles begin to relax, causing the pressure within the heart to become less than that in the aorta and pulmonary artery. This leads to a brief back flow of blood that result in the semilunar (aortic and pulmanory) valves snapping shut. S2 takes place at the end of the T-wave.

Although S1 and S2 are each considered to be discrete sounds, each is generated by the the near-instantaneous closing of two separate valves. For most practical pur-

poses it is sufficient to consider each of these sounds as being single and instantaneous. However, it is worth remembering the exact order of the closures since certain conditions can split each heart sound into its separate valve components and knowing the order of valve closure makes understanding the different reasons for the splitting of heart sounds easier. During S1, the closing of the mitral valve slightly precedes the closing of the tricuspid valve, while in S2, the aortic valve closes just before the pulmonary valve. This follows from the fact that the pressure during systole in the left ventricle is much greater than in the right, resulting in the mitral valve closing before the tricuspid in S1. Similarly, because the pressure at the start of diastole in the aorta is much higher than in the pulmonary artery, the aortic valve closes first in S2.

When listening to a patient's heart, the cadence of the beat usually distinguishes S1 from S2. Because diastole takes about twice as long as systole, there is a longer pause between S2 and S1 than there is between S1 and S2. However, rapid heart rates can shorten diastole to the point where it is difficult to discern which is S1 and which is S2. This issue can be dealt with by making use of information in the EKG signal.

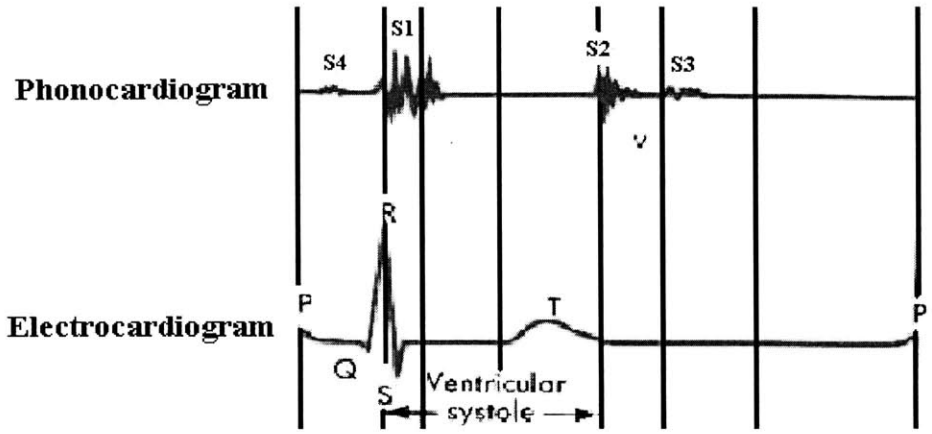
In addition to these two sounds, abnormal function of the heart may give rise to two additional acoustical signals that occur independently of each other. S3 is a third heart sound and is due to rapid passive ventricular filling. It occurs in dilated congestive heart failure, severe hypertension, myocardial infarction, or mitral incompetence. Similarly, S4 is a fourth heart sound that is associated with atrial contraction against a stiffened ventricle. This is often associated with aortic stenosis or hypertensive heart disease and may also occur in heart failure.

Figure A-19 illustrates the position of these heart sounds in the cardiac cycle.

A.4.3 Auscultation Sites

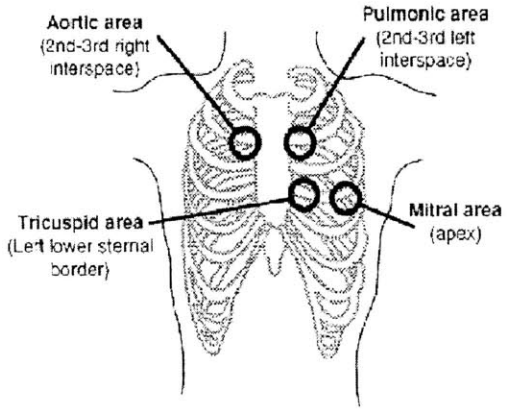
Generally, the sounds produced by each valves is best heard over a particular region of the chest as shown in Figure A-20.

Normally, only S1 and S2 can be heard using the stethoscope. The S3 sound is of-



Reproduced from http://meds.queensu.ca/medicine/physiol/undergrad/phase2/Lecture5f.ppt

Figure A-19: Occurrence of Heart Sounds in the Cardiac Cycle



Reproduced from http://meds.queensu.ca/medicine/physiol/undergrad/phase2/Lecture5f.ppt

Figure A-20: Auscultation Sites in the Human Body

ten prominent in children however, and in certain cases, S4 may also be distinguished for normal patients.

Bibliography

[1] Pease, A. "If the Heart Could Speak", Pictures of the Future, Fall 2001, Seimens Webzine

(http://w4.siemens.de/FuI/en/archiv/pof/heft2_01/artikel19/)

[2] Craige, E. "Should Auscultation be Rehabilitated?", New England Journal of Medicine 1988, 318:1611-3.

[3] Mangione et al., "Cardiac Auscultatory of Internal Medicine and Family Practice Trainees: a Comparison of Diagnostic Proficiency", JAMA 1997;278:717-722.

[4] Mangione et al, "The Teaching and Practice of Cardiac Auscultation During Internal Medicine and Cardiology Training: a Nationwide Survey", Ann Intern Med 1993;119;47-54.

[5] "Facts about Mitral Valve Prolpase", National Heart, Lung and Blood Institute Information Booklet

(http://www.nhlbi.nih.gov/health/public/heart/other/mvp/mvp_fs.htm)

[6] Hoffman, R. "Mitral Valve Prolapse", Conscious Choice

(www.consciouschoice.com/holisticmd/hmd093.html)

[7] MEDLINEplus Medical Encyclopedia : Mitral Valve Prolapse

(http://www.nlm.nih.gov/medlineplus/ency/article/000180.htm)

[8] The Physician and Sportsmedicine : Mitral Valve Prolapse

(http://www.physsportsmed.com/issues/1996/07_96/joy.htm)

[9] Heart Sounds and Murmurs

(http://www.mcevoy.demon.co.uk/Medicine/Medicine/ClinExamn/Murmurs.html)

[10] Heart Murmurs and Other Abnormal Heart Sounds, NurSpeak

(http://www.nurspeak.com/tools/murmurs.htm)

[11] Bates, B. "A Guide to Physical Examination", 4th edition. J.B. Lippincott Co., Philadelphia 1987.

[12] Yancone-Morton, L. "Cardiac assessment", RN 54:(12)28-35, Dec. 1991.

[13] Valsalva Maneuver, HealthAtoZ

http://www.healthatoz.com/healthatoz/Atoz/ency/valsalva_maneuver.html

[14] Dr. Robert Levine, Private Communication.

[15] MATLAB is a trademark of The Mathworks, Inc., 3 Apple Hill Drive, Natick, MA 01760-2098, USA.

(http://www.mathworks.com/)

[16] Friesen, G. et al, "A comparison of the noise sensitivity of nine QRS detection algorithms", IEEE Transactions on Biomedical Engineering, Vol. 37, No. 1, January 1990

[17] Fraden, J. et al, "QRS Wave Detection", Med. Biol. Eng. Comput., vol 18, pp.125-132, 1980.

[18] Dr. Nathaniel Sims, Private Communication.

[19] Dr. Francesca Nesta, Private Communication.

[20] Oppenheim, A. et al, "Discrete-Time Signal Processing", Prentice Hall Signal Processing Series

[21] Lim, J., "Two-Dimensional Signal and Image Processing", Prentice Hall, 1990

[22] Gonzalez, R. et al, "Digital Image Processing" Addison-Wesley, 1993.

[23] Borda, R. et al, "Error reduction in small sample averaging through the use of the median rather than the mean", Electroenceph Clin Neurophysiol, v. 25 (1968), pp. 391-392.

[24] Tukey, J. "Exploratory Data Analysis" Reading, MA: Addison- Wesley, 1971.

[25] Oppenheim, A. et al, "Signals and Systems", Prentice Hall, 1996

[26] Fairbanks, G. et al, "Method for time or frequency compressionexpansion of speech", IEEE Trans. Audio and Electroacoustics, AU-2, pp. 712, 1954.

[27] Flanagan, J. et al "Phase Vocoder", Bell System Technical Journal, November 1966, 1493-1509.

(http://www.ee.columbia.edu/ dpwe/e6820/papers/FlanG66.pdf)

[28] Lee, F., "Time compression and expansion of speech by the sampling method", J. Audio Eng. Soc., Vol. 20:3, pp. 738742, 1972.

[29] Dolson, M., "The phase vocoder: A tutorial", Computer Music Journal, vol. 10, no. 4, pp. 14 - 27, 1986.

[30] Suzuki, R. et al, "Time-scale modification of speech signals using crosscorrelation functions", IEEE Transactions on Consumer Electronics, Vol. 38:3, pp. 357363, 1992.

[31] Laroche, J., "New Phase Vocoder Technique for Pitch-Shifting, Harmonizing and Other Exotic Effects", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. Mohonk, New Paltz, NY. 1999.

(http://www.ee.columbia.edu/ dpwe/papers/LaroD99-pvoc.pdf)

[32] Ellis, D., "A Phase Vocoder in MATLAB", Columbia University

(http://www.ee.columbia.edu/ dpwe/resources/matlab/pvoc/)

[33] Cathers, I. "Neural network assisted cardiac auscultation", Artif Intell Med. 1995;7:5366.

[34] Guo, Z. et al. "Artificial neural networks in computer assisted classification of heart sounds in patients with porcine bioprosthetic valves", Med Biol Eng Comp. 1994;32:311316.

[35] Coehn, M. et al. "Comparative Approaches to Medical Reasoning", River Edge, NJ: World Scientific; 1995:271288.

[36] Okuni, M. et al, "Trial of a new cardiac mass screening system in school children", Jpn Circ J. 1978;42:4952.

[37] Reynolds, J. "Heart disease screening of preschool children", Am J Dis Child. 1970;119:488493.

[38] White, P. et al, "Time frequency analysis of heart murmurs in children" Proc IEEE. 1997;3:14.

[39] El-Asir, B. et al, "Time frequency analysis of heart sounds". Proc IEEE. 1996;12:553558

[40] Leung, T. et al, "Analysing pediatric heart murmurs with discriminant analysis", Proc IEEE. 1998;18:16281631.

[41] Sarkady, A. et al, "Computer analysis techniques for phonocardiogram diagnosis", Comput Biomed Res. 1976;9:349363.

[42] DeGroff, C. et al, "Artificial Neural NetworkBased Method of Screening Heart Murmurs in Children", Circulation 2001;103:2711-2716.

[43] Winston, P. "Artificial Intelligence", Addison-Wesley Pub Co, 3rd edition, 1992

[44] Dewan, S. et al, "Heart Murmurs", Good Health Magazine, Feb/Mar 1998.

(http://www.seton.net/Wellness/GoodHealthMagaine/GoodHealthMagaineAr0A67/-FebMarBJJILoveYourH09829/HeartMurmurs.asp)

[45] Shino, H. et al, "Detection and classification of systolic murmur for phonocardiogram screening", Proc. of the 18th Intl Conf. of the IEEE Eng. in Med. and Biol. Soc., Volume 1, pp. 123124. Amsterdam, The Netherlands.

[46] Barschdorff, D. et al, "Automatic phonocardiogram signal analysis in infants based on wavelet transforms and artificial neural networks", Computers in Cardiology 1995, pp. 753756. IEEE, Vienna, Austria.

[47] Reed, T. et al, "The Analysis of Heart Sounds for Symptom Detection and Machine-Aided Diagnosis", 2nd Conference on Modelling and Simulation in Biology, Medicine and Biomedical Engineering, Delft, The Netherlands

[48] Watrous, R. et al, "Wavelet based bank of correlators approach for phonocardiogram signal classification", Proc. of the IEEE-SP Intl Symp. on Time-Frequency and Time-Scale Analysis, pp. 7780. Pittsburgh, PA.

[49] Myint, W. et al, "An Electronic stethoscope with diagnosis capability"

[50] Banerji, A. et al. "Circulation : Cardiovascular System", Life Science Discipline, LA Mission College

(http://207.233.44.253/wms/reynolmj/lifesciences/lecturenote/bio3/Chap23.ppt)

[51] Gulaff, E. "Cardiac Cycle", Perfusion Education, KFSH&RC

(http://www.kfshrc.edu.sa/perfusion/education/CardiacCycle.ppt)

[52] Fisher, J. "Cardio-Vascular Physiologiy", Queen's University Faculty of Health Sciences

(http://meds.queensu.ca/medicine/physiol/undergrad/phase2/Lecture5f.ppt)

[53] "The ECG", Cardionetics

(http://www.cardionetics.com/docs/healthcr/ecg.htm)

[54] The Auscultation Assistant, Physiology

(http://www.wilkes.med.ucla.edu/Physiology.htm