

Audio-Visual Tools for Computer-Assisted Diagnosis of Cardiac Disorders

Zeeshan Syed, Daniel Leeds, Dorothy Curtis, John Guttag
Massachusetts Institute of Technology
{zhs, ddleeds, dcurtis, guttag}@csail.mit.edu

Francesca Nesta, Robert A. Levine
Harvard Medical School and Massachusetts General Hospital
{fnesta, rlevine}@partners.org

Abstract

The process of interpreting heart sounds is restricted by human auditory limitations. Shortcomings such as insensitivity to frequency changes, slow responses to rapidly occurring changes in acoustic signals and an inability to discriminate the presence of soft pathological sounds are the source of inaccuracies and persist even with experience. This restricts both the practice and teaching of auscultation.

In this paper we propose and evaluate a suite of presentation tools for computer-assisted auscultation. We explore the use of digital signal processing techniques to slow down heart sounds while preserving frequency content, differential enhancement across frequency scales to amplify pathological disease signatures, visualization of the signal to measure changes in signal energy across time and presentation of a representative prototypical signal for the patient.

1. Introduction

When a heart valve is stenotic or damaged, the abnormal blood flow patterns produce a series of audible vibratory sounds known as *murmurs* [1]. Murmurs may also be produced as a result of septal defects, or compromised communication between blood vessels and the heart [2]. Physicians detect these disorders by listening to heart sounds at different locations across the torso. In many cases, murmurs heard during routine physical examinations offer important clues to the presence of previously undetected and asymptomatic cardiac disease. Consultations with cardiologists and various tests such as echocardiography may then be used to establish a firm diagnosis and form the basis for treatment of an underlying disorder.

The process of interpreting heart sounds is known as cardiac auscultation. It serves as a first line of defense against cardiac disease, but is associated with numerous difficulties. Different cardiac ailments produce a potentially overwhelming set of acoustic pathological events, and correctly identifying a disorder requires discrimination of subtle variations in the timing characteristics and spectral properties of heart sounds [3]. This analysis is further complicated by natural variations in heart sounds introduced by factors such as the sex, age, habitus and dynamic state of the patient. Inaccuracies also arise due to the presence of innocent or benign murmurs, which may have no pathological significance. Up to 50% of school aged children have physiological murmurs, with a ratio of innocent murmurs to pathological murmurs of 10:1 [4,5]. Even in adults, anxiety, stress, fever, anemia, overactive thyroid glands and pregnancy may cause benign murmurs. Typically, these cases are distinguished by examining the intensity of sounds, in addition to their timing and frequency content.

Sounds that are interesting from the perspective of auscultation are often short lived and separated from one another by less than 30 milliseconds [6]. Pathological signals indicative of cardiac diseases are also often much quieter than other heart sounds and their audibility varies across successive heart beats. It is not surprising then, that the ability to perform auscultation accurately has traditionally been restricted to skilled cardiologists. Even with extensive experience, physicians may often disagree about heart sounds, especially in the case of phenomena associated with brief sounds [7]. In [8] and [9] these inaccuracies are attributed to human auditory limitations, which include insensitivity to frequencies, slow responses to rapidly occurring changes in acoustic signals and an inability to unmask soft sounds in the proximity of loud ones. These shortcomings affect both the practice and teaching of auscultation.

In this paper we propose a suite of presentation tools for computer-assisted diagnosis of cardiac disorders. Our efforts are focused on reducing the subjective component of auscultation, by:

- Slowing down heart sounds with preservation of frequency content to allow better time-resolution of events,
- Amplification of frequencies to enhance soft pathological disease signatures
- Graphical visualization of the raw audio signal to precisely determine energy changes in the signal and indicate how events are organized over an interval of time, and
- Presentation of a prototypical beat that sums up activity across an examination to create a representative biological signal and reduces the effect of noise, including the presence of lung sounds

The structure of this paper is as follows. Sections 2 and 3 describe the audio and visual presentation tools of the MIT Automated Auscultation System (MAAS). The techniques at the heart of these tools, and the

results of applying these to cardiac data are presented. Section 4 then discusses related work. A summary and conclusions are presented in Section 5.

2. Audio Aids

Audio files for this section can be accessed online at <http://maas.csail.mit.edu/sounds>.

2.1. Reduced Rate Playback with Preservation of Frequency Content

The first of our audio aids allows physicians to play back heart sounds at slowed down rates. This facilitates differentiating between various acoustical events by making the separation between them more pronounced. In doing so, subjective demands on the ear to tease apart temporally adjacent artifacts is reduced.

To prevent distortion of the frequencies in the raw acoustical signal, we draw upon existing techniques for time-scale modification of audio [10]. We make use of

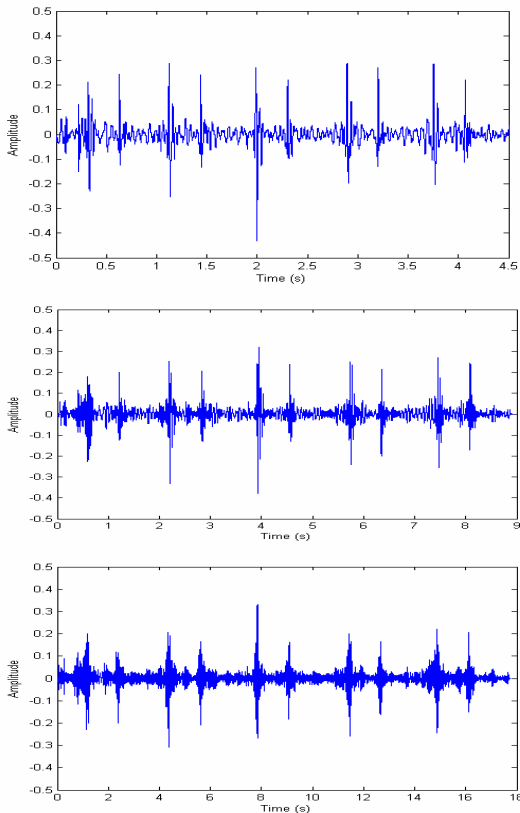


Figure 1: Original audio signal from a normal heart (top) compared to versions of this signal slowed down by factors of two (middle) and four (bottom)

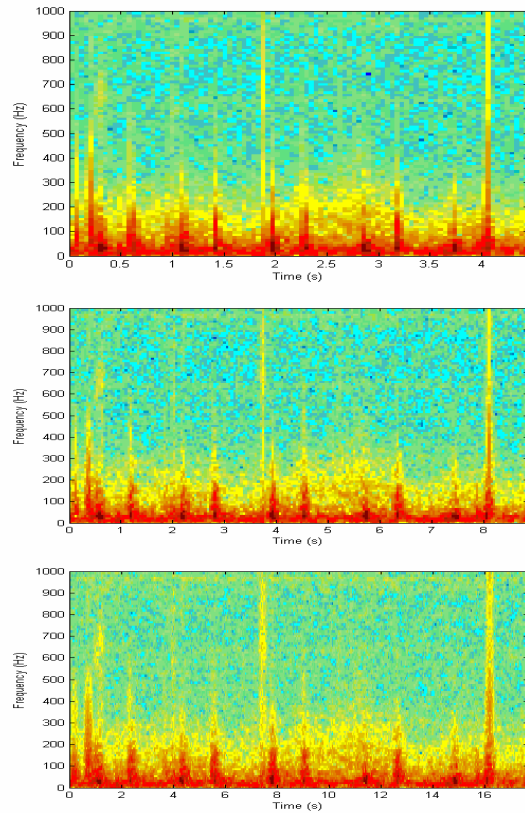


Figure 2: Spectrograms for the signals shown in Figure 1, comparing the raw audio (top) with versions of the signal slowed down by a factor of two (middle) and four (bottom)

phase vocoding, which effectively stretches the time-base of a spectrogram to slow down a sound while retaining its short-time spectral characteristics. If the spectrogram is narrowband (i.e., the window of analysis is sufficiently long to provide sufficient frequency resolution to identify individual harmonics) the pitch of the signal is preserved. The phase associated with each bin in the modified spectrogram then needs to be modified to provide a consistent transition of phase over time. We use the MATLAB phase vocoder implementation [11] to achieve this.

Figure 1 plots both the original audio signal for a normal heart and the corresponding signals obtained by slowing down the recording by factors of two and four. Figure 2 presents the spectrograms for the signals shown in Figure 1. Slowing down heart sounds does not significantly alter the time-frequency content of the original signal.

2.2. Selective Amplification of Frequencies

One of the key challenges associated with auscultation, described in [12], is the fact that the amplitude of pathological activity is often two to three orders of magnitude less than the baseline signal corresponding to the first and second heart sounds (i.e., S1 and S2). This requires physicians to have ears sensitive to soft sounds. To address this issue we allow for the selective amplification of the different frequency ranges identified in [12]. Specifically, we use time-frequency decomposition and then reconstitute the original signal by differentially amplifying the signal for each band. The degree of amplification for each band is a configurable parameter, and in most cases was adjusted to normalize each band relative to the baseline S1-S2 signal.

Figure 3 shows the original time-domain signal for a normal heart and illustrates the effects of selective amplification. Figure 4 does the same for a patient suffering from mitral regurgitation (MR). In the case of the normal heart, energy at higher frequencies is focused at the locations of S1 and S2. Amplifying high frequency bands preserves this effect. For the patient with MR, there is increased energy during systole. The amplification process reveals this phenomenon more clearly. The spectrogram shows increased energy at high frequency relative to the baseline S1-S2 signal, and the time-domain signal shows increased activity during systole, particularly close to S2. Amplification also affects the amplitude of the time-domain signal in Figure 4 more than the signal in Figure 3. This is due to the presence of greater high frequency content in the patient with MR, and consequently more energy is

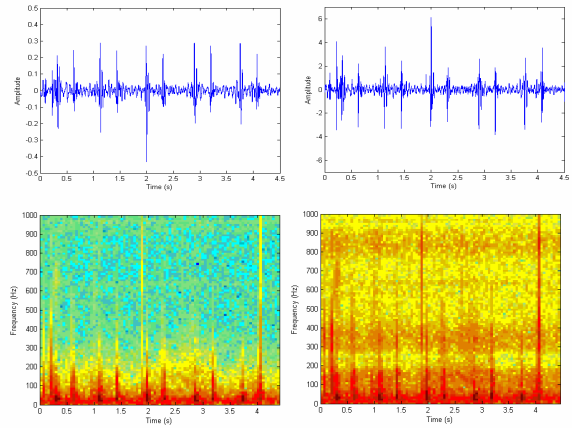


Figure 4: Original audio and spectrogram from a normal heart (left) compared to the amplified version (right)

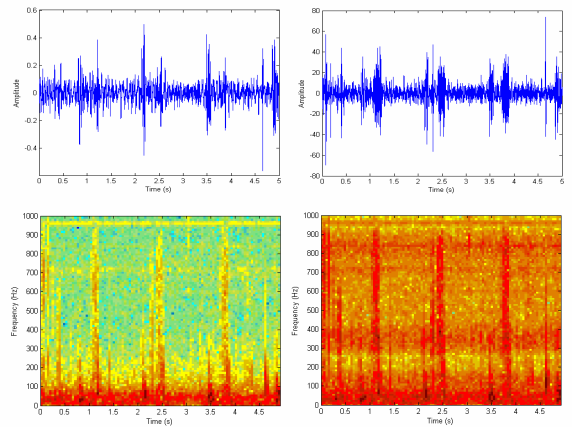


Figure 4: Original audio and spectrogram from an MR patient (left) compared to the amplified version (right)

amplified in this case.

2.3. Prototypical Beat Playback

One of the key difficulties associated with auscultation is the sheer amount of information that must be absorbed by the physician. In [12] and [13], a mechanism is proposed to create a composite prototypical heart beat for patients corresponding to diseased activity. This provides the benefits of compacting data, reducing noise and enhancing the recurring nature of pathological activity while reducing physiologically irrelevant variations across individual beats. Disease findings can be robustly analyzed without focusing attention on events that do not persist or arrive randomly. We adopt this prototypical representation of audio as an aid and allow for the playback of such signals to reduce the demand on

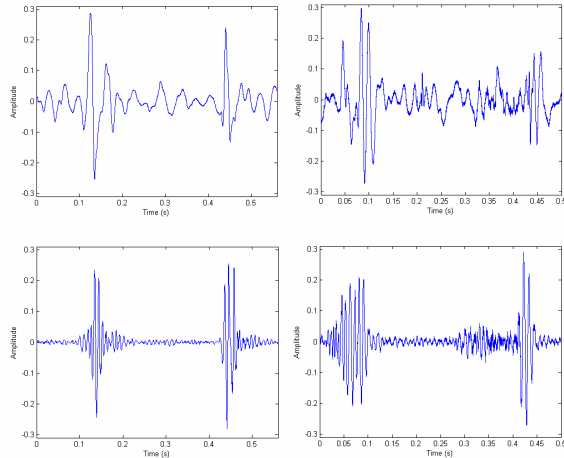


Figure 5: Individual (top) and prototypical beats (bottom) for normal hearts (left) and MR (right)

physicians to keep track of information across examinations.

Frequency content of the prototypical beat can be selectively amplified as described earlier to enhance repeating sounds that are soft.

Figure 5 compares individual beats from a normal heart and a patient with MR to the respective prototypical beats. In both cases, the prototypical beats remove variations due to noise. For the patient with MR, recurrent activity corresponding to the late systolic murmur seen during echocardiography is preserved.

3. Visual Aids

3.1. Annotated Visualization of Raw Data

The issue of distinguishing between events that are close together in time can be addressed by allowing the slowed down playback of heart sounds. An alternative approach is to graphically display fixed snapshots of beats. If the audio signal is sampled at a high enough frequency, visualizing the data makes it easier to discriminate distinct events. In addition, the magnitude of the signal at different locations, and the spacing between events can be measured accurately.

We supplement the on-screen display of heart beats with annotations of clinically relevant events such as the R and T waves (derived from simultaneous EKG) and S2. This allows fusion of information across multiple modalities and reduces the complexity of auscultation by decreasing the number of tasks a physician is required to carry out in parallel. Instead of having to segment the signal, establish the location of events within the cardiac cycle, and process all this

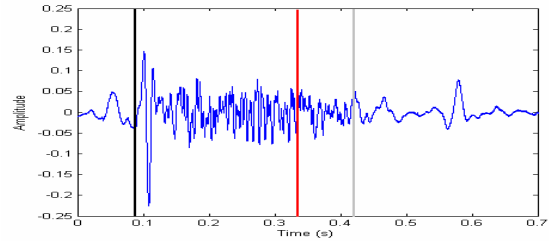


Figure 6: Graphical display of an annotated heart beat for a patient with MR. The black vertical line (left) marks the R-wave, the red line (middle) marks the T-wave and the grey line (right) marks S2.

information in real-time to make a decision, physicians can scroll through plots of each heart beat at their own pace, focusing more closely on the analysis of a signal where cardiac activity relative to events is pre-labeled.

The graphical display of audio data provides the additional advantage of reducing the bias introduced by the human ear, which is not equally sensitive to all frequencies of sound within the audible range of 20-20,000 Hz. Signals at maximum human sensitivity (close to 3,500 Hz) are often perceived as louder than those with similar amplitude at lower or higher frequencies [14], and this effect is diminished by visualizing the data rather than listening to it.

Figure 6 provides an annotated plot of a heart beat for a patient with MR.

3.2. Scaled Frequency Bands

In the presence of the high amplitude S1-S2 baseline, soft pathological sounds may often be suppressed visually. We propose a mechanism to scale the energy at higher frequencies and display signals corresponding to different frequency ranges separately to resolve this issue.

We use the time-frequency decomposition and normalization described in Section 2.2 to scale the frequency bands of the audio signal. The output of this process is then displayed.

Figures 7 and 8 illustrate this mechanism for a normal heart and a patient with MR respectively. The scaled frequency band display makes it easier to recognize the pathological condition. For the normal heart, even after selective amplification, there is no significant high frequency energy. For the patient with MR, the murmur is difficult to discern in the raw audio recording but is revealed by frequency analysis, which also characterizes the activity as a high-frequency murmur. Based on physical acoustics, this can be interpreted as a flow domain with high driving

pressures, corresponding to a pathological murmur and not a benign one. The clear late-systolic timing also indicates that the murmur is not aortic stenosis or a simple MR, but mitral valve prolapse (MVP).

3.3. Prototypical Beat Display

The prototypical beats created using the techniques described in [12] and [13] can be graphically displayed in addition to being played back. Figure 9 presents an example of this approach.

The use of the prototypical beat for visualization purposes provides two key advantages over display of the raw or frequency-scaled signal (Figures 6–8). Firstly, the prototypical beat extracts relevant activity while removing noise. Secondly, the prototypical beat is able to collapse information from across the entire duration of recording. This allows for a single beat to be analyzed and removes the need to iterate over displays of multiple beats, as is the case for raw or frequency-scaled signals. An improvement is provided in terms of both the quality of information displayed, and the time needed to assess it.

4. Related Work

Our suite of audio-visual tools builds upon the auscultatory framework described in [12]. We make use of processing blocks provided to segment the signal, choose beats for further analysis, decompose the signal into frequency bands and create prototypical signals most likely to contain pathological activity.

Most commercially available electronic stethoscopes are currently packaged with visualization tools for filtering out heart sounds and displaying the resulting data. We augment this approach by providing annotations of data based on the segmentation algorithms described in [12] and also develop a collection of different audio-visual aids that can supplement existing software products.

The ideas of selective amplifying heart sounds, and allowing for variable rate playback of the raw audio signal can also be found in [8] as part of a more general discussion on the future of auscultation and conceivable tools that could improve current practice. We formalize these concepts and develop a toolbox that provides an implementation of both ideas. In [15], the notion of a frequency shifting stethoscope is proposed, which converts low input frequencies below the 20 Hz cutoff of the audible range into a range of change of frequency in the mid-auditory range. A salient feature of our work is that we supplement the emphasis in [8] and [15] on amplifying low frequency

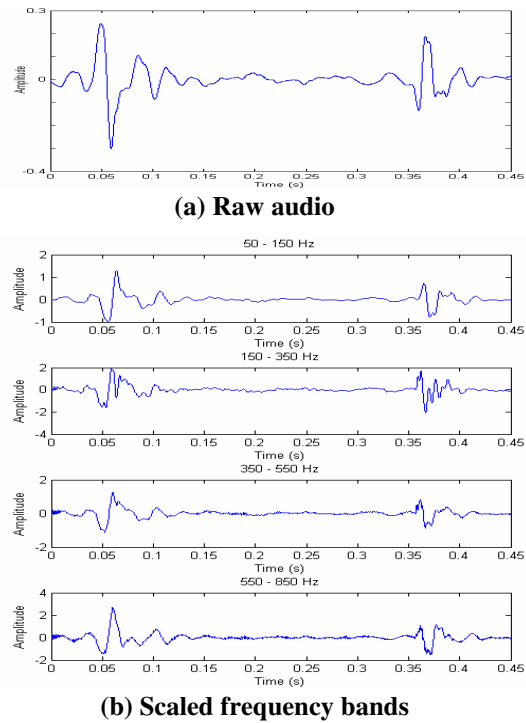


Figure 7: Display of raw audio and scaled frequency bands for a normal heart

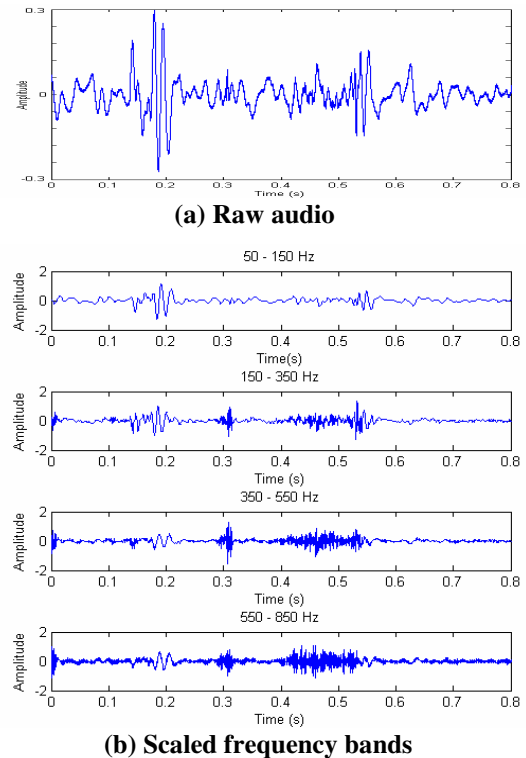


Figure 8: Display of raw audio and scaled frequency bands for a patient with MR

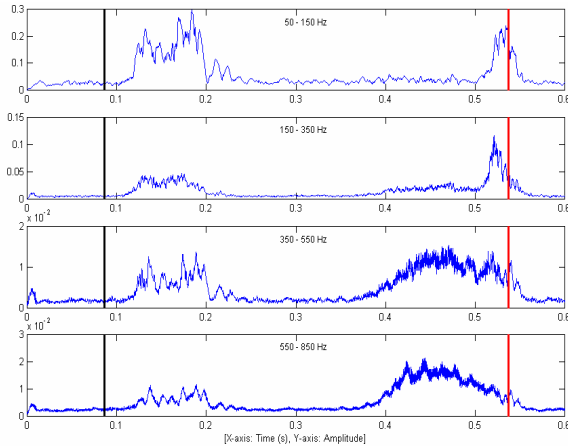


Figure 9: Prototypical beat for a patient with MR. The black line (left) gives the position of the prototypical R wave in EKG while the red line (right) corresponds to the prototypical S2

ranges corresponding to diastolic gallop sounds and diastolic murmurs due to mitral stenosis (MS), with a mechanism to amplify soft high frequency sounds that are associated with regurgitation. In [7], a graphic system to display heart sounds with low frequency amplifications is described as failing to record high frequency murmurs such as those associated with aortic or mitral regurgitation. Our system is better suited to detect these kinds of events.

We also introduce the use of the prototypical beat as an essential component of computer-assisted diagnosis. We did not find any prior work in this area.

Many computer systems exist geared towards automated auscultation (a discussion of these works is provided in [16]). Our approach to computer-assisted diagnosis differs in that we do not strive for a single-bit classification. We propose tools that could yield additional information regarding a patient's cardiac health. This includes determining the extent of the disease and the patient-specific morphology of the heart murmur. We believe this information allows for a better understanding of events taking place at the anatomical level for each patient.

5. Summary and Conclusions

In this paper, we presented a suite of audio-visual presentation tools that can be used to assess cardiac health during physical exams. We addressed the problem of auscultation from the perspective of the subjective demands it places on human faculties, and proposed software aids that make it easier to recognize soft, short-lived heart sounds recurring within pathological frequency ranges. Conversations with

cardiologists suggest that our toolbox is able to facilitate diagnosis of cardiac disorders. By making it easier to hear relevant activity, our work also allows for learning modules that can be used in medical schools to teach auscultation in a structured manner.

6. Acknowledgements

We thank the Center for Integration of Medicine and Innovative Technology (CIMIT), the Deshpande Center for Technological Innovation, the AETNA Foundation, and the MIT Project Oxygen Partnership for funding, and Meditron for providing stethoscopes.

7. References

- [1] P Stein, "A Physical and Physiological Basis for the Interpretation of Cardiac Auscultation: Evaluations Based Primarily on the Second Sound and Ejection Murmurs", Futura Publishing Company; 1981.
- [2] T Biancaniello, "Innocent Murmurs: A Parent's Guide", *Circulation*, 109, pp.162-163; 2004.
- [3] B Shipton, H Wahba, "Valvular Heart Disease: Review and Update", *American Family Physician*, 63(11), pp.2201-2208; 2001.
- [4] S Gibson, "Clinical Significance of Heart Murmurs in Children", *Medical Clinics of North America*, 30, 35-36; 1946.
- [5] C Caceres, L Perry, *The Innocent Murmur: A Problem in Clinical Practice*, Little Brown and Company; 1967.
- [6] A Pease, "If the Heart Could Speak", *Pictures of the Future*; Fall 2001.
- [7] M Tavel, D Brown, D Shander et al, "Enhanced Auscultation with a New Graphic Display System", *Archives of Internal Medicine*, 154, 893-898; 1994.
- [8] M Tavel, "Cardiac Auscultation: A Glorious Past But Does It Have a Future?", *Circulation*, 93, 1250-1253; 1996.
- [9] M Selig, "Stethoscopic and Phonoaudio Devices: Historical and Future Perspectives", *Am Heart J*, 126, 262-268; 1993.
- [10] J Flanagan, R Golden, "Phase Vocoder", *Bell System Technical Journal*, 1493-1509; 1966.
- [11] D Ellis, "A Phase Vocoder in MATLAB", Columbia University.
- [12] Z Syed, D Leeds, D Curtis, F Nesta, RA Levine, J Guttag, "A Framework for the Analysis of Acoustical Cardiac Signals", *submitted to IEEE TBME*.
- [13] Z Syed, J Guttag, "Prototypical Representation of Biological Signals", *in preparation*.
- [14] L Beranek, "Acoustical Measurements", *Acoustical Society of America*; 1988.
- [15] G Holloway, W Martin, "An Electronic Frequency Shifting Stethoscope for Heart Sounds", *Journal of Bioengineering*, 2, 59-64; 1978.
- [16] Z Syed, "MIT Automated Auscultation System", Master's thesis, Massachusetts Institute of Technology, 2003.