Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Software Enhanced Learning of Cardiac Auscultation

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

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

Abstract

Listening to heart sounds during physical exams can offer useful clues to the presence of cardiac disease. Cardiac auscultation is non-invasive, inexpensive and fast. It is also highly unreliable, and requires good hearing and considerable expertise.

In this paper, we describe an audio-visual tool designed to help people learn to be better at cardiac auscultation. The use of digital signal processing techniques makes pathological findings more recognizable in both audio and visual representations. This reduces dependence on the ability of people to hear relevant information.

1. Introduction

Cardiac auscultation, the process of interpreting heart sounds, is part of the first line of defense against heart disease. Unfortunately, detecting relevant symptoms and performing screening based on the sounds heard through a stethoscope is difficult. Evidence suggests that as many as 80% of the individuals referred to cardiologists as the result of auscultatory exams have benign murmurs or normal hearts [1-3]. And there is no way of knowing how many problems are missed.

A number of factors contribute to the difficulty of cardiac auscultation. Relevant pathological activity is often soft, short-lived and occurs in proximity to loud, normal activity: a typical murmur is 1000 times softer than normal heart sounds and can last for as little as thirty milliseconds. The acoustic information is also inconsistent across the course of an examination, owing to natural variation and noise. These factors make it difficult to precisely pinpoint a sound in the audio signal, and to identify markers indicative of heart disease.

Unsurprisingly, cardiac auscultation is difficult to teach. Only 27.1% of the educational programs for internal medicine, and 37.1 % for cardiology, incorporate the structured teaching of auscultation [4].

In this paper, we describe a system that addresses some of the difficulties associated with learning how to perform cardiac auscultation. We describe a software based system that uses a variety of audio-visual to highlight clinically relevant information in heart sounds. The focus of our work is to decompose information into time-correlated audio and visual streams, where abnormal activity can be recognized independently in each signal. This process of providing multiple perspectives on acoustic cardiac data helps develop a more complete understanding of auscultation.

The structure of this paper is as follows. Sections 2 and 3 briefly describe the digital signal processing techniques at the heart of our work and present an overview of the system. Section 4 then discusses related work. Summary and conclusions appear in Section 5.

2. Audio-Visual Aids

In this section, we describe the audio-visual tools at the heart of our work. These techniques form the basis of computer aids to assist in diagnostic aspects of cardiac auscultation [5]. Section 3 presents a tool that incorporates the result of applying many of these techniques into a system for computer assisted cardiac auscultation.

2.1. Scrolling Display of Heart Sounds

As a first step, we allow for the playback of heart sounds with the simultaneous scrolling display of a visual representation of the sounds. Visualizing the data in this form facilitates the recognition of distinct events, especially in the proximity of loud artifacts. This is particularly important in dealing with cases such as trace late-systolic murmurs, where the murmurs are several orders of magnitude softer than the baseline S1-S2 signal corresponding to the two heart sounds. The visual display indicates the precise moment when the acoustical activity of interest is taking place.

To maximally correlate the activity being seen with that being heard, we filter the audio signal before display to approximate the effect of the human ear, which perceives sounds at certain frequencies as being louder than those with similar amplitudes at higher or lower frequencies [6].

The display of the raw signal is supplemented by information that relates sounds to the location of points such as the R and T waves in time-correlated ECG. The display of the R wave in the ECG signal helps show the student when the first heart sound (S1) begins and the T wave helps identify the second heart sound (S2). This allows for auscultation to be carried out in a more modular manner, with students able to focus on listening to the sounds without having to mentally perform segmentation in parallel.

2.2. Amplification of Pathological Frequencies

In order to address the low amplitude nature of some pathological sounds, we allow for the amplification of frequency bands of the original signal corresponding to pathological ranges [7]. This reduces the dependence on having hearing that is sensitive enough to isolate relevant activity.

The resulting signal can be played back as described in Section 2.1 with an associated scrolling display that allows visual and acoustic activity to be related.

2.3. Slowed Down Heart Sounds

We allow for heart sounds to be slowed down and played back at reduced rates to prolong short events and make them easier to hear. This process increases the time separation between sounds, facilitating resolution of significant events during the cardiac cycle.

To restrict the distortion of frequencies, we employ phase vocoding [8], which preserves the pitch characteristics of the raw signal. The sounds produced can then be listened to directly, or in conjunction with any of the techniques discussed in Sections 2.1 and 2.2.

2.4. Prototypical Beat Construction

In [7] and [9] we described a mechanism to create a representative aggregate signal from multiple heart beats. The mechanism is considerably more complex than simple averaging, which often obscures rather than highlights clinically relevant information.

Properly done, the pooling of beats improves the quality of the data by reducing noise and enhancing events that are recurrent in nature. At the same time, the construction of a prototypical signal reduces physiologically irrelevant variations permitting the analysis of disease findings in a robust manner without emphasis on events that do not persist or are random in nature. This allows greater focus on important diagnostic information. The prototypical beat is constructed to preserve the high energy content corresponding to murmurs that repeat across beats with varying intensity. The construction also removes lung sounds and background noise that occupy frequency ranges corresponding to murmurs, but are not consistent with each beat.

The prototypical beat can be combined with any of the techniques described in Sections 2.1–2.3.

3. The System

The learning system incorporates implementations of the techniques presented in Section 2, a repository of heart sounds and a user interface. It runs on both Windows and Linux.

The learning system uses a set of fifty heart sound recordings. These were selected from a larger collection that contains recordings from several sources:

- Patients referred to the Massachusetts General Hospital (MGH) for an echocardiogram based on the suspicion of valvular disorders.
- Members of families with a history of mitral valve problems.
- Members of the Computer Science and Artificial Intelligence Laboratory at MIT (for additional normal heart sounds).

Each recording is about thirty seconds long, and contains the raw audio and time-correlated ECG. Both signals were sampled at 44.1 kHz, with 16 bit quantization. The data was collected using a Meditron electronic stethoscope with 2-lead ECG.

We processed the recordings from these cases using the techniques described in Section 2 and created a collection of aids for each case. These include a slowed down version of the original audio signal, a graphical representation of the prototypical beat (the visual prototype) and a recording synthesized from the prototypical beat (the audio prototype). Heart sounds were slowed down using the techniques described in Section 2.3 to achieve a reduction in playback rate by a factor of 2. The visual prototypical beat was constructed such that the information can be displayed separately for each of four different frequency bands

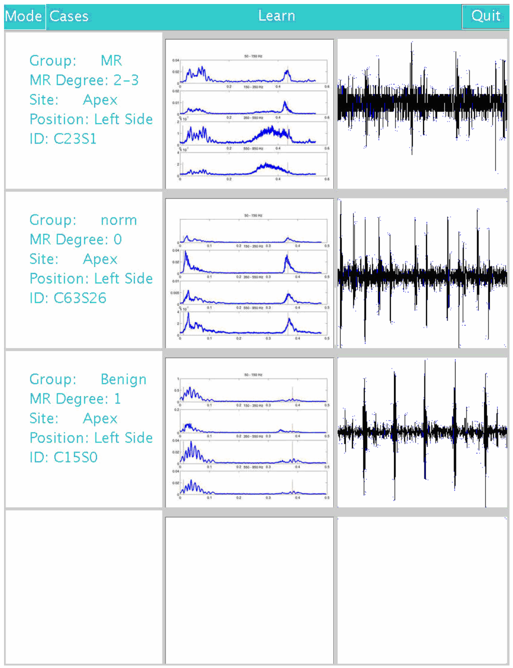

Fig. 1. Learning mode with information for individuals with MR (top), normal heart sounds (middle) and benign murmurs (bottom) is displayed in parallel. Students are provided the functionality to play back any of these sounds. The plots on the extreme right correspond to raw audio tracings of each record. The visual prototype is displayed in the middle and shows how energy evolves in the signal for each frequency band.

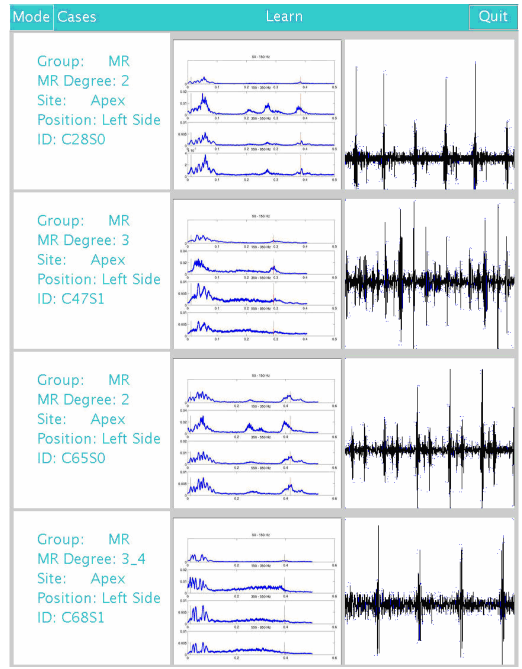

Fig. 2. Learning mode with recordings for multiple MR patients: In this case, the simultaneous display of the visual prototypes for the heart sounds allows students to develop an appreciation of varying disease signatures.

(50Hz-150Hz, 150Hz-350Hz, 350Hz-550Hz, and 550Hz-800Hz). For the audio prototypical beat, selective amplification is further applied to these frequency ranges to make low amplitude information in the higher bands more easily recognizable.

The system has two modes: learning mode and quiz mode. In the learning mode the student can select from several categories of heart sounds. There are mitral regurgitation (MR) heart sounds, normal heart sounds, and heart sounds with benign murmurs. The MR heart sounds are further subdivided into sections depending upon the degree of mitral regurgitation. Once the student has selected a particular heart sound to study, the system displays information about the case, the visual prototypical beat, and the original audio recording.

The case information includes the classification of the case (MR, normal or benign), the degree of mitral regurgitation as seen on the echocardiogram, the site where the heart sound was recorded, the position of the patient and the case id. The visual prototypical beat shows four frequency bands. The amplitudes are all positive so that the cumulative power can be seen more easily. Then the student can listen to the original recording, the slowed down recording, or the audio prototypical beat. We allow for up to four heart sounds to be displayed in parallel, facilitating comparisons between recordings that share the same underlying pathology or belong to different classes. The user interface allows students to play any of the sounds while viewing the graphical representation of all signals.

Figure 1 shows the learning mode with a selection of heart sounds. Notice that the visual prototype highlights the presence of increased energy during late systole for the MR case (on top of the figure). In contrast, the energy in the recordings corresponding to the normal heart and the benign murmur is concentrated at S1 and S2 instead.

Figure 2 provides another view of the learning mode. In this case, different recordings, each corresponding to an MR case are compared. The murmur appears in the higher frequency bands as before, but can occupy different ranges. The disorder also exhibits varying morphologies, corresponding to cases such as late systolic, mid-systolic and holo-systolic murmurs.

In the quiz mode, the learning tool begins by selecting a random heart sound. The student can then choose whether to look at the visual prototypical beat, or listen to the original recording, the slowed down

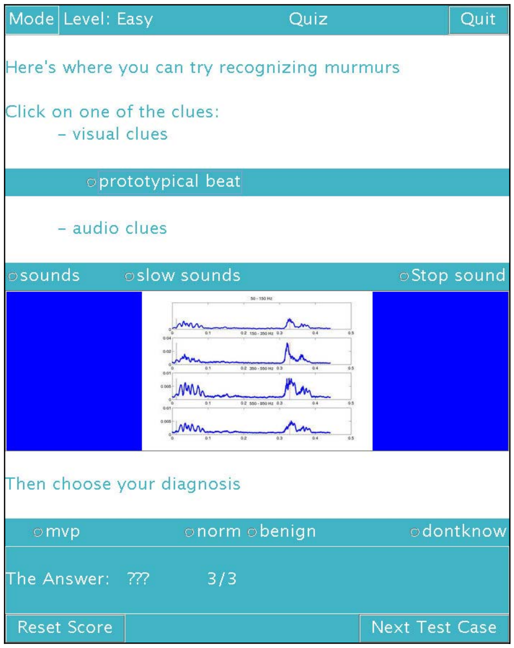

Fig. 3. Quiz mode: In this case, the system has selected a random heart sound from the 'easy' level and the student has chosen to display the visual prototypical beat. Towards the bottom of the screen, the "3/3" indicates that the student has correctly identified 3 recordings in 3 attempts.

version, or the audio prototype. The system keeps track of the number of correct answers compared to the total attempts. Figure 3 illustrates the functionality provided by the quiz mode.

4. Related Work

A number of commercially available systems (e.g. [10]) exist as teaching bundles for auscultation. These essentially function as sound distributors, allowing recordings from a single sensor to be shared across multiple headsets. Our work differs from this approach is that we approach the teaching of auscultation from the perspective of making it easier to extract relevant information from raw audio signals, as opposed to sharing the recording in real-time across individuals. We do not address the need for hardware that can be used to transduce recordings and deliver them simultaneously to multiple destinations.

The idea of using a software-based heart sound teaching system also appears in [11], which allows for the graphical display of raw audio along with vocal explanations and prompts. Our work differs in the use of digital signal processing techniques that transform the raw audio into various representations geared towards making pathological activity more visible or audible. This allows students to understand diagnostic markers better, before attempting to learn how to detect cardiac disorders.

5. Summary and Conclusions

We have described a tool to aid in the learning of auscultation. Through the use of digital signal processing techniques we enhance the information available: the slowed down heart sounds allow the student to hear more clearly sounds that occur very close in time and the visual prototypical beat highlights the timing and shape of the murmurs.

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] A Pease, "If the Heart Could Speak", *Pictures of the Future*, Fall 2001.

[2] S Mangione and L Nieman, "Cardiac Auscultatory Skills of Internal Medicine and Family Practice Trainees: A Comparison of Diagnostic Proficiency", *JAMA*, 278; pp. 717-22, 1997.

[3] S Mangione, L Nieman, E Gracely and D Kaye, "The Teaching and Practice of Cardiac Auscultation During Internal Medicine and Cardiology Training: A Nationwide Survey", *Ann. Intern. Med.*, 119; pp. 47-54, 1993.

[4] E Craige, "Should Auscultation be Rehabilitated", *New England Journal of Medicine*, 318; pp. 1611-13, 1988.

[5] Z Syed, D Leeds, D Curtis, F Nesta, RA Levine and J Guttag, "Audio-Visual Tools for Computer-Assisted Diagnosis of Cardiac Disorders", *19^{th} IEEE Int. Symp. on Computer-Based Medical Systems*, 2006.

[6] L Beranek, "Acoustical Measurements", *Acoustical Society of America*, 1988.

[7] 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.*

[8] J Flanagan, R Golden, "Phase Vocoder", *Bell System Technical Journal*, 1493-1509; 1966.

[9] Z Syed, J Guttag, "Prototypical Representation of Biological Signals", *in preparation*.

[10] Welch Allyn Meditron Teaching Bundle.

[11] X Shouzhong, Z Shiyong, C Zehan, G Fei and P Yuli, "An Audiovisual and Quantitative Heart Sound Teaching System", *proceedings of the 20^{th} IEEE Int. Conf. on Eng. in Med. and Bio.*, 20; pp. 3361-63, 1998.