A Probabilistic Model for Phonocardiograms Segmentation Based on Homomorphic Filtering

Gill D [1], Intrator N [2], Gavriely N[3]

[1] Department of Statistics, The Hebrew University of Jerusalem, Jerusalem 91905, Israel, [2]School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel, [3]Rappaport Medicine Faculty, Technion IIT, Haifa 31096, Israel

gill@mta.ac.il

Abstract. This work presents a novel method for automatic detection and identification of heart sounds. Homomorphic filtering is used to obtain a smooth envelogram of the phonocardiogram, which enables a robust detection of events of interest in heart sound signal. Sequences of features extracted from the detected events are used as observations of a hidden Markov model. It is demonstrated that the task of detection and identification of the major heart sounds can be learned from unlabelled phonocardiograms by an unsupervised training process and without the assistance of any additional synchronizing channels.

1 Introduction

Phonocardiography is a continuous noninvasive, low-cost but accurate monitoring method for valves functioning; it is easily repeatable with no risk to the patient. However, heart diagnosis by auscultation requires high skills and experience of the listener. Automatic detection and identification of heart sounds plays an essential role in automatic diagnosis of phonocardiograms, especially in cases when additionally synchronizing channels like Electrocardiograms (ECG) and carotid pulse (CP) are not present.

Segmentation refers to partitioning the PCG signal into cardiac cycles, and detection and identification of the main events (e.g.: S1, S2, S3, S4, murmurs, and other sounds) and intervals (systole, diastole) in each cycle.

2 Methods

A number of techniques have been proposed for segmentation. Many of them are based on the information conveyed by the envelope of the signal and attempt to detect certain events where they cross a predetermined threshold. These methods include envelope extraction using discrete wavelet decomposition and reconstruction [1] or using the magnitude of the analytic signal formed using the PCG and its Hilbert transform [2]. In the present approach, homomorphic filtering [3] is used to extract a smooth envelogram, which enables the detection of events that are suspected to be S1, S2 or others.

A monocomponent AM-FM signal can be expressed as a product of its amplitude modulation (AM) and frequency modulation (FM) components:

$$x(t) = a(t) \cdot f(t) \qquad a(t) > 0 \tag{1}$$

We denote:

$$\hat{x}(t) = \ln|x(t)| = \ln a(t) + \ln|f(t)| . \tag{2}$$

In cases where $x(t) = 0$ we add a small positive value, and then we have

$$\hat{x}(t) = \ln a(t) + \ln|f(t)| . \tag{3}$$

By using a low-pass linear filter L whose pass-band covers the typical frequencies of the AM component and attenuates the typical high frequencies of the FM component, we obtain:

$$\hat{x}_l(t) = L[\ln a(t)] + L[\ln|f(t)|] \approx \ln a(t)\,. \tag{4}$$

The reversal procedure of recovering $a(t)$ is done by an exponential operation. The recovered AM component is called *homomorphic envelogram* and has the advantage of scalable smoothness, which handles the problems of splits and serrated peaks (see figure 1).

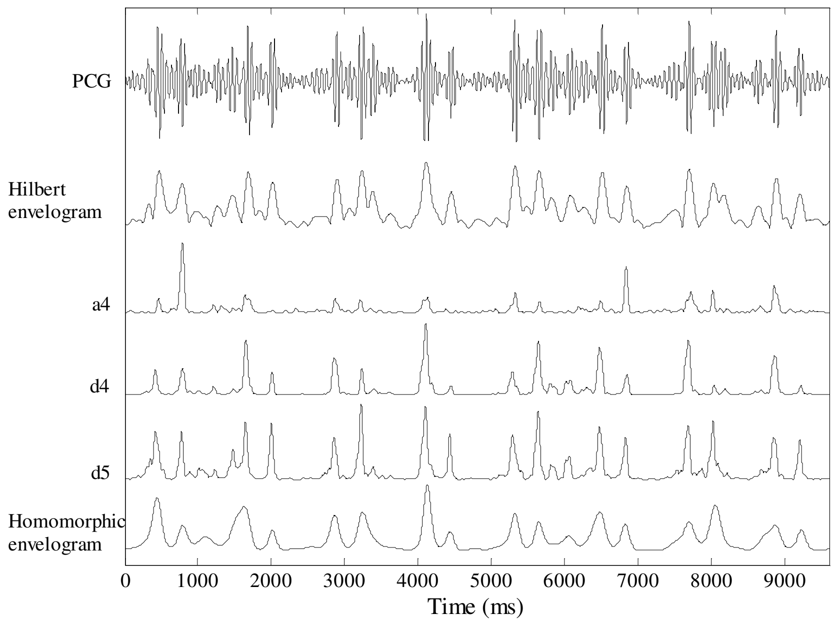

Figure 1. A phonocardiogram of a subject recorded in a CT room (top) and several envelopes: Hilbert envelogram, normalized average Shannon energy of 4th-level approximation (a4), normalized average Shannon energy of 4th-level detail (d4), normalized average Shannon energy of 5th-level detail (d5), and homomorphic envelogram of the signal's Shannon energy.

We use hidden Markov model (HMM) [4] whose states simulate the true events, and the observations within each state include features extracted from the envelogram's peaks. In order to prevent events misses, all envelogram's peaks were collected. Using the Baum-Welch algorithm, HMM with three states and a multivariate normal distribution of observations was trained. The training set consisted of sequences of features vectors that were extracted from all the event locations. The features extracted from each event, selected by the detection process, were: time duration form current peak to the following one, the amplitude of the current envelogram's peak, and the second derivative of the envelogram at the peak. Since empirically the transition from detected S1 to detected S2 shows much higher probability than other transitions, we find the maximal entry in the transition probability matrix, M, denoted by $m_{i^*j^*}$ and identify the state i^* as S1 and state j^* as S2. The Viterbi algorithm computes the most probable sequence of states ($\hat{I}$) given the model parameters and a sequence of observations:

$$\hat{I} = \arg\max_I P\left(I \,\middle|\, \left\{ \Delta t_i, |x(t_i)|, \frac{\partial^2 a(t_i)}{\partial t^2} \right\}_{i=1}^{n-1}, \lambda \right) \tag{5}$$

Where: t_i is the time of the i-th peak, Δt_i is the time difference between the (i+1)-th and the i-th peaks, a is the homomorphic envelogram, x is the original signal, and n is the number of peaks.

3 Results

44 phonocardiograms were taken from 17 subjects, aged from 13 to 72 years, and were recorded for 30-60 seconds from several auscultation locations. The environmental conditions included: quiet office, hospital computed tomography (CT) room (noise from the CT scanner, patients moving, and talking), and hospital echo room (noise from patients moving, and people talking). The recording was done using a passive piezoelectric sensor at sampling rate of 4kHz. The recordings were filtered with a pass band from 20Hz to 250Hz.

For each of S1 and S2 we define a true positive (TP) when the corresponding heart sound is detected within a symmetric appropriate window and is the nearest detected event to a human expert annotation. A false negative (FN) is considered when there is a failure to detect the event correctly within a window. A false positive (FP) is considered when there is an event detection outside the appropriate window or if the detection is within the appropriate window but is not the nearest detected event to the expert annotation. The sensitivity (Se) is defined as: Se=TP/(TP+FN) and Positive predictivity (PP) is defined as: PP=TP/(TP+FP). A four-fold cross-validation was performed for model evaluation. The evaluation results are given in Table 1.

	Se	PP
S1	98.6%	96.9%
S2	98.3%	96.5%

Table 1.The algorithm results compared to annotations by an expert.

4 Discussion and Conclusions

As shown in the results, the homomorphic filtering is sensitive enough to reveal low SNR events and still overcome the problem of detecting too many events caused by heart sounds splits, serrated peaks and noisy environment. The results show that this method competes successfully with other state of the art methods [1,5] and yet it is simple and unsupervised. Adding supervised learning and time-frequency representation features can strengthen the method and enlarge the variety of problems it can handle such as automatic detection of other heart sounds as S3, S4, murmurs, and valves defects. In addition, the method can be generalized to other biomedical signals.

References

[1] Liang H, Lukkarinen S, Hartimo I. A heart sound segmentation using wavelet decomposition and reconstruction. Proceedings of the 19th Annual International Conference of the IEEE-EMBS, Chicago 1997;1630-1633.

[2] Sarkady AA, Clarck RR, Williams R. Computer analysis techniques for phonocardiogram diagnosis. Comput Biomed Res. 1976 Aug;9(4):349-63.

[3] Oppenheim AV. Generalized linear filtering. In: Gold B, Rader CM, editors. Digital Processing of Signals, New-York, McGraw-Hill Book Company, 1969; ch. 8.

[4] Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 1989; 77(2):257-286.

[5] Gamero LG, Watrous R. Detection of the first and second heart sound using probabilistic Models. Proceedings of the 25th Annual International Conference of the IEEE-EMBS, Cancun, Mexico, 2003; 2877-2880.