Robust Heart Rate Estimation from Noisy Phonocardiograms

David B Springer[1], Thomas Brennan[2], Jens Hitzeroth[3], Bongani M Mayosi[3], Lionel Tarassenko[1], Gari D Clifford[1,4]

[1] University of Oxford, Oxford, UK
[2] Massachusetts Institute of Technology, Cambridge, USA
[3] University of Cape Town and Groote Schuur Hospital, Cape Town, South Africa
[4] Emory University, Atlanta, USA

Abstract

Accurate heart rate estimation is a fundamental process when analysing phonocardiograms (PCGs). While this is trivial in noise-free recordings, it becomes a difficult task in PCGs corrupted by various noise sources. While numerous PCG-based heart rate estimation techniques have been explored in the literature, no comparison between these techniques has been performed to identify the best-performing method in noisy recordings.

This paper evaluates various denoising, normalisation, envelope extraction and heart rate estimation techniques on 585 noisy recordings made using four different devices. The best-performing algorithm correctly estimated the heart rate in 471 (80.5%) of these PCGs, while correctly estimating the heart rate in 86% of the PCGS from a separate (publicly available) test dataset.

1. Introduction

An essential process in the automatic analysis of phonocardiograms (PCGs) is heart rate (HR) estimation. This is difficult in noisy PCGs due to the fact that noise sources, such as motion artifacts and speech and lung sounds, interfere with the heart sounds of interest in both the time and frequency domains [1]. Furthermore, PCGs recorded by non-experts, such as in the case of mobile-phone based applications, are likely to be of lower quality, necessitating robust HR estimation.

Many different HR estimation techniques have been explored in the literature. However, no comparison has been made between these techniques to find the best-performing method. This paper compares a number of pre-processing, signal envelope detection, normalisation and HR estimation algorithms on noisy heart sound recordings from multiple data sources to find the best-performing algorithm for accurate estimation of the HR from PCGs with a large amount of noise contamination.

The most common HR estimation technique when using the PCG is peak detection in the autocorrelation after various transformations [2–4].

A critical step for the derivation of the autocorrelation is signal envelope extraction. The three envelope extraction techniques that have been used extensively are the homomorphic envelogram [2, 5], the Hilbert envelope [3, 6] and the average normalised Shannon energy [7]. A fourth, classic method for envelope extraction, based on full-wave rectification and low-pass filtering (FWR-LPF) was investigated in this paper.

Methods for HR estimation in the autocorrelation of PCG signals include single peak detection [2] and computing the periodicity using singular value decomposition [3]. In addition, multiple peak detection was investigated.

Wavelet denoising has been found to be a suitable method for removing noise contamination from PCG signals [6]. However, there is limited agreement on the optimal wavelet to use when analysing PCGs. The Morlet wavelet [8], Daubechies wavelet family [6, 9] and Symlet wavelet family [10] have all been motivated for PCG analysis.

Normalisation of PCGs is performed to limit the variation between recordings from different patients and different recording devices. Methods include division by the maximum amplitude in the signal [4, 10], subtracting the mean and dividing by the standard deviation of the signal [7] or dividing by a percentile value of the signal [2].

This paper investigates the combination of these parameters and methods that yield the best-performing HR estimation algorithm.

2. Methods

2.1. Datasets

Three different data sets, using a range of six different recording devices, were used in this study.

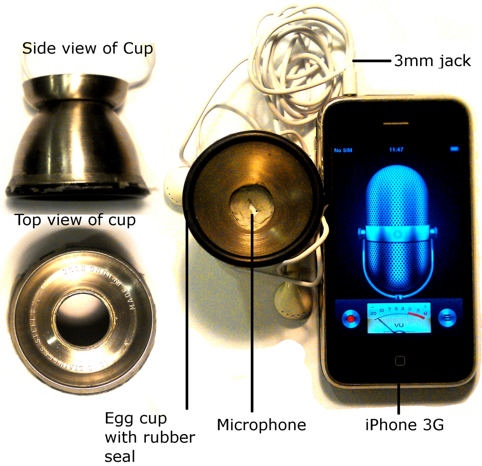

Figure 1. Hand-made stethoscope with hands-free kit attached and plugged into an iPhone 3G.

The first data set contains recordings from 150 patients from the cardiology clinic at the Groote Schuur hospital in Cape Town, South Africa[1]. These patients had various heart conditions, including biological and artificial valve replacements, congestive heart failure, pacemakers and congenital disorders. All recordings were made by an untrained research assistant in order to replicate the role of an untrained health care worker.

PCG recordings were made using three different devices: the first was a 3M Littmann 3200 electronic stethoscope; the second PCG recording device used was an iPhone 3G mobile phone; and the third was a a Nokia 3110 Classic mobile phone. Both mobile phones were equipped with a hand-made stethoscope attachment, shown in Figure 1. This stethoscope attachment was made by placing the microphone of a standard iPhone hands-free kit into the neck of a metal egg-cup, based on the work of Kuan [11].

The reference HR for each recording was derived by finding the median time between peaks in synchronous photoplethysmography (PPG) data recorded with each PCG recording using a Nonin Onyx II 9560 finger pulse oximeter.

The second data set consisted of 405 synchronous PCG and ECG recordings from 123 de-identified adult patients attending the Massachusetts General Hospital for cardiac screening or in-home recordings of people suffering from mitral valve prolapse (MVP), 83 of whom were found to have murmurs [12]. The recordings were made using a Welch-Allyn Meditron Elite electronic stethoscope. The reference HR for each PCG signal was found using a Pan-Tomkins detector in the synchronous ECG signals [13].

[1]This study was approved by the Human Research Ethics committee, Health Science Faculty, University of Cape Town (HREC REF: 568/2010)

Finally, the test dataset in this study is a publicly available dataset [14], consisting of 111 annotated recordings from two sources: twenty-one normal recordings using the iStethoscope Pro iPhone application and 90 normal recordings using the DigiScope device. The reference HR from these recordings is found by computing the median time between the manually annotated first heart sound positions from each recording.

2.2. Data Exclusion

To ensure accurate reference HR estimates, poor-quality PPG (from the Cape Town dataset) and ECG signals (from the Massachusetts dataset) were excluded.

The quality of the reference PPG signals used in the mobile stethoscope dataset was found by using the qSQI algorithm [15]. In order to ensure absolute precision of the reference HR estimates in this study, any PPG signal with a signal quality value of below 0.9 was excluded. This led to the exclusion of six Littmann, four iPhone and five Nokia recordings.

The quality of the ECG signals used in the MIT database was computed using the method derived by [16]. This classifier was trained using the original data used by the researchers, being similar to the data used in this study. Using this classifier, 108 ECGs were found to be of low quality. From the remaining recordings, 150 were randomly selected in order to match the number of recordings from the mobile stethoscope dataset.

2.3. Analysis Overview

This analysis performed an exhaustive search over 41 different wavelet denoising parameters, four envelope detection methods, five normalisation thresholds and three HR estimation techniques to derive a HR estimate for each recording. This lead to a total of 14,760 different combinations. These steps are described in the following sections.

2.4. Pre-processing

All PCG recordings were downsampled to 1000 Hz using a polyphase anti-aliasing filter. The frequency content of the fundamental heart sounds is below 500 Hz [1] and hence the Nyquist-Shannon sampling criterion was satisfied.

As discussed in Section 1, wavelet denoising has been shown to be advantageous for PCG analysis. However, there is little agreement on the optimal wavelet and decomposition level to use. For that reason, a number of wavelets (Morlet, Daubechies 4-10, Symlet 18 and Biorthogonal 2.8) and decomposition levels (3-6) were used to decompose the signal using the discrete wavelet transform

Table 1. Results on training sets, showing the best-performing algorithm parameters on each set, the total number of recordings in each set and the number of recordings within the defined tolerance (± 5 bpm)

Dataset	Total PCGs	Wavelet	Decomposition Level	Normalisation Level (%)	Envelope Detection	HR Estimation	PCGs within tolerance
Nokia	145	Daubechies 4	4	95	Hilbert	Single Peak	120
iPhone	146	No Wavelet	NA	95-97	FWR-LPF	Single Peak	107
Littmann	144	Biorthogonal 2.8	3	100	Hilbert	Single Peak	122
MIT	150	Daubechies 8	4	100	Hilbert	Single Peak	144
		Daubechies 10	4	99	Hilbert	Single Peak	144
Combined	585	Biorthogonal 2.8	3	95	Hilbert	Single Peak	471

(DWT)[2], and then reconstructing the signal with the exclusion of wavelet approximations and details outside of the chosen decomposition level. The effect of this was the filtering out of frequencies in the signal outside the desired frequency range.

This led to a total of 41 different pre-processing methods; 40 using wavelet decomposition and one with no wavelet denoising.

2.5. Envelope Detection

As discussed in Section 1, four different envelope extraction methods were investigated. These were: the FWR-LFP method[3], the extraction of the Hilbert Envelope [5], homomorphic filtering [2] and finding the average Shannon energy envelope [7].

2.6. Normalisation and Autocorrelation

It was decided to follow the normalisation process used by [2] where the normalised envelope is found by first subtracting the median and then dividing by a percentile value of the absolute value of the signal. In order to optimize the normalisation percentile value, a range from 95 to 100 % was tested.

Thereafter, the autocorrelation waveform was computed. The autocorrelation is the cross-correlation of a signal with itself, which accentuates repeating patterns in noisy signals. This is useful when analysing a semi-periodic signal such as the PCG.

2.7. Heart Rate Estimation Methods

Three HR estimation algorithms were tested. The first method was the detection of the first salient peak in the autocorrelation [2]. This method of HR estimation computes the maximum peak within in a permissible range (dictated by realistic values of HR) within the autocorrelation function. The HR limits used in this study were 30 - 140 beats per minute (bpm), as resting HRs outside of this range were not expected. The HR, hr, was calculated using:

$$hr = 60/lag_{peak}\, bpm \tag{1}$$

where lag_{peak} is the lag time to the detected peak in the autocorrelation.

The second HR estimation method was to find multiple peaks in the autocorrelation. This HR estimation algorithm is an extension of the previous method. As before, the location of the peak with the greatest amplitude within the permissible range in the autocorrelation is found. Thereafter, the largest peak within a permissible lag from the previously found peak location is found. In order to compute the optimal number of peaks to find, a range of three to nine peaks was used. The HR is found by computing the median lag between the identified peaks and using Equation 1.

Finally, a HR estimation method, introduced by [3], finds windows of the autocorrelation function that are most similar using singular value decomposition (SVD) and bases the HR on the width of these windows.

3. Results

HR monitors should be able to compute the HR to within 10 % of the reference HR, or within five beats per minute, whichever is larger [17]. Therefore, the most successful algorithm was that which estimated the HR within these bounds for the highest number of recordings.

The combinations of parameters which led to the highest number of correctly derived HRs for each data set are shown in Table 1. In order to find the best-performing algorithm across all datasets, the best-performing algorithm on all 585 training recordings was found. This is shown in the last line of Table 1.

The results from each algorithm from Table 1 on the test set are shown in Table 2.

[2] In the case of the Morlet wavelet, the continuous wavelet transform was used as the discrete wavelet transform is not possible.

[3] This was a 20 Hz cut-off, second-order, zero-phase, Butterworth low-pass filter

Table 2. Results on test sets, showing number of PCGs within $\pm$ 5 bpm tolerance using best-performing algorithm from each training set (e.g., "Nokia" heading refers to the best-performing algorithm found using the Nokia dataset in Table 1).

		PCGs within $\pm$ 5 bpm tolerance using best-performing algorithm from each training set:					
Dataset	Total PCGs	Nokia	iPhone	Littmann	MIT		Combined
iStethoscope	21	20	18	20	20	20	19
Digiscope	90	74	70	76	71	70	76

4. Discussion

Single peak detection along with the Hilbert envelope extraction led to the best HR estimates in all but one of the training datasets, illustrating the superiority of these methods.

Lower percentage values for normalisation for the Nokia and iPhone datasets are expected, due to their higher susceptibility to noise spikes when using a low-cost stethoscope.

The best-performing algorithm on the "combined" training set would be most generalisable to new datasets, and as expected, performed as well as other algorithms on the test sets. Therefore, it can be concluded that this method, using single peak detection in the autocorrelation waveform after preprocessing and normalisation yields the most accurate HR estimates across a variety of PCGs.

Acknowledgments

David Springer is funded by the Rhodes trust. Many thanks to Dr. Francesa Nesta, Prof. John Guttag, Dr. Zeeshan Syed and Dorothy Curtis for the use of the MIT auscultation data.

References

[1] Abbas AK, Bassam R. Phonocardiography Signal Processing. Morgan & Claypool Publishers, 2009.

[2] Schmidt SE, Holst-Hansen C, Graff C, Toft E, Struijk JJ. Segmentation of heart sound recordings by a duration-dependent hidden Markov model. Physiological Measurement April 2010;31(4):513–29.

[3] Kumar D, Carvalho P, Antunes M, Paiva RP, Henriques J. Noise detection during heart sound recording using periodicity signatures. Physiological Measurement May 2011; 32(5):599–618.

[4] Kao WC, Wei CC. Automatic phonocardiograph signal analysis for detecting heart valve disorders. Expert Systems with Applications June 2011;38(6):6458–6468.

[5] Rezek I, Roberts S. Envelope Extraction via Complex Homomorphic Filtering. Technical Report TR-98-9. Technical report, Imperial College, London, 1998.

[6] Messer SR, Agzarian J, Abbott D. Optimal wavelet denoising for phonocardiograms. Microelectronics Journal December 2001;32(12):931–941.

[7] Liang H, Lukkarinen S, Hartimo I. Heart sound segmentation algorithm based on heart sound envelogram. In Computers in Cardiology, volume 24. Lund, Sweden, 1997; 105–108.

[8] Oskiper T, Watrous R. Detection of the first heart sound using a time-delay neural network. In Computers in Cardiology. Memphis, Tennesee, August 2002; 537–540.

[9] Ahlstrom C, Länne T, Ask P, Johansson A. A method for accurate localization of the first heart sound and possible applications. Physiological Measurement March 2008; 29(3):417–28.

[10] Dokur Z, Ölmez T. Heart sound classification using wavelet transform and incremental self-organizing map. Digital Signal Processing November 2008;18(6):951–959.

[11] Kuan KL. A framework for automated heart and lung sound analysis using a mobile telemedicine platform. Masters thesis, Massachusetts Institute of Technology, 2010.

[12] Syed Z, Leeds D, Curtis D, Nesta F, Levine RA, Guttag J. A framework for the analysis of acoustical cardiac signals. IEEE Transactions on Biomedical Engineering 2007; 54(4):651–662.

[13] Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering March 1985;32(3):230–6. ISSN 0018-9294.

[14] Bentley P, Nordehn G, Coimbra M, Mannor S. The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results. http://www.peterjbentley.com/heartchallenge/index.html, 2011.

[15] Li Q, Clifford GD. Dynamic time warping and machine learning for signal quality assessment of pulsatile signals. Physiological Measurement September 2012;33(9):1491–501.

[16] Behar J, Oster J, Li Q, Clifford GD. ECG signal quality during arrhythmia and its application to false alarm reduction. IEEE Transactions on Biomedical Engineering June 2013;60(6):1660–6.

[17] Association for the Advancement of Medical Instrumentation. Cardiac monitors, heart rate meters, and alarms, 2002.

Address for correspondence:

David Springer
University of Oxford, Wellington Square, Oxford OX1 2JD
david.springer@eng.ox.ac.uk