# PCG EVENT DETECTION AND CLASSIFICATION USING RIGHT SINGULAR VECTORS *

Erin Budd and Maryhelen Stevenson
*University of New Brunswick*
*Fredericton, New Brunswick Canada*

## ABSTRACT

An unsupervised methodology for detecting and classifying recurring sounds in the phonocardiogram (PCG) without reference to any other biological signal is proposed. Peaks in the orthonormal time domain singular vectors of a time-frequency representation of the PCG are used as both detection statistics and classification features. Applying the proposed strategy to the first three singular vectors detects 97.5% of the 1224 fundamental heart sounds in 42 PCG's.

## INTRODUCTION

A phonocardiogram (PCG) is a graphic record of the heart sound signal. It is a nonstationary signal that consists of a recurring cycle of two types of acoustic events, which are referred to as the *fundamental heart sounds* $S_1$ and $S_2$. The PCG may also include other cardiac events, such as $S_3$, $S_4$, murmurs, clicks, and snaps, which may or may not occur in every cardiac cycle. Non-cardiac events, such as noise and artifacts due to breathing and motion, may also occur. The fundamental heart sounds and the murmurs, if present, potentially contain information regarding valve functionality and heart disease [1]. Detecting and labelling these transients is a first step towards extracting diagnostic information from the heart sounds and the intervals between them. A low cost, automated system that requires no specialized equipment or secondary signals for labelling events in the PCG could be useful as part of either a learning tool for cardiac auscultation or a screening aid for valvular dysfunction. Such a system would be particularly beneficial in environments where more expensive diagnostic tests, such as the echocardiogram, may be unavailable or subject to long wait times.

The automated identification of the events in the PCG is a challenging task. All events in the PCG may vary in duration, temporal spacing, amplitude, and frequency characteristics from one heart beat to the next in a recording from one subject, as well as between subjects. Furthermore, the PCG events may overlap in either time or frequency or both.

Using time cues from a synchronously-recorded peripheral biological signal such as the electrocardiogram (ECG), the fundamental heart sounds can be accurately segmented from the PCG either manually or automatically [2]. Approaches to PCG segmentation without reference to an ECG typically segment the PCG by isolating peaks in an energy profile, *e.g.* the Shannon Energy Envelogram (SEE) [3] or a wavelet-based profile [4]. Timing guidelines regarding the expected durations of $S_1$, $S_2$, and their separating intervals are then used to reduce the number of invalid claimed events, recover missed events, and assign labels $S_1$ and $S_2$ to the remaining segments. The accuracy of strategies that rely on timing guidelines tends to degrade when presented with rapidly varying heart rates or overlapping events.

The objective of the proposed work is to detect recurring events in the PCG and group them into three classes, *i.e.* $S_1$, $S_2$, and *Other*, by identifying similar agglomerations of coefficients in a time-frequency representation of the PCG.

## METHODOLOGY

### Time-Frequency Matrix

Let $X$ denote a $1 \times N$ array of PCG samples. It is assumed that $X$ contains multiple occurrences of $S_1$ and $S_2$. An example of two PCG cycles is shown in Figure 1 (a). The fundamental heart sounds were labelled manually with reference to its ECG.

A time-frequency representation of $X$ is obtained by correlating it with a bank of $M$ complex Morlet wavelets,

*This work was supported by the Natural Sciences and Engineering Research Council of Canada.

$\psi_m(k), m = 1,2,\ldots,M$, each given by

$$\psi_m(k) \quad = \quad \frac{1}{\sqrt{a_m}\pi^{\frac{1}{4}}} e^{-\frac{k^2}{2a_m^2}} e^{-j2\pi f_m k}, \qquad (1)$$

where $a_m$ denotes the scale of $\psi_m(k)$, $f_m = \frac{f_0}{a_m}$ is the center frequency of $\psi_m(k)$, and $f_0 = \frac{6}{2\pi}$. The scales $\{a_m\}$ are chosen such that the corresponding values of $\{f_m\}$ span the frequency range of the PCG events of interest.

The $1 \times N$ array of the $m^{th}$ correlator's output is given by the continuous wavelet transform (CWT):

$$C_m(n) = \sum_{k=1}^{N} \psi_m(k-n)X(k). \qquad (2)$$

The $m^{th}$ correlator acts as a matched filter whose magnitude response is a Gaussian function with a center frequency at $f = f_m$ and a bandwidth proportional to $\frac{1}{2\pi a_m}$.

Let $W$ denote a matrix of the squared magnitude of the outputs of the bank of $M$ correlators, thus

$$W(m,n) = |C_m(n)|^2. \qquad (3)$$

Matrix $W$ is a $M$ frequency bin $\times$ $N$ time bin representation of $X$. Figure 1 (b) is a contour plot of the portion of $W$ that corresponds to the PCG of Figure 1 (a).

Event Detection

Given a $M \times N$ matrix $W$ of real values with $M < N$, there always exists orthonormal matrices, $U = [u_1,\ldots,u_M]$ of size $M \times M$ and $V = [v_1,\ldots,v_M]$ of size $N \times M$ such that

$$W = USV^T, \qquad (4)$$

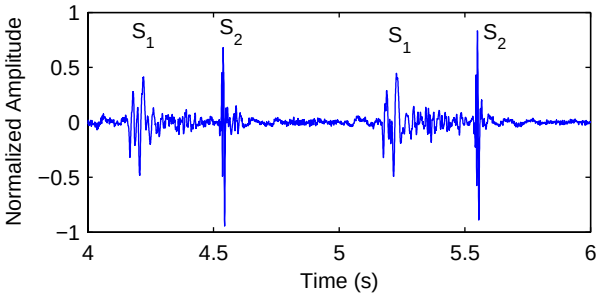
where $S = \mathrm{diag}(s_1 \geq s_2 \geq \ldots \geq s_M)$. The values $s_i$ are the *singular values* of $W$, $u_i$ is the $i^{th}$ *left singular vector* of $W$, and $v_i$ is the $i^{th}$ *right singular vector* of $W$. The pair $(u_i, v_i)$ are chosen to satisfy $\sum_{n=1}^{N} v_i(n) > 0$, where $v_i(n)$ is the $n^{th}$ element of $v_i$.

According to (4), the $n^{th}$ column of $W$, $W_n$, can be reconstructed as a weighted sum of the vectors $u_i$:

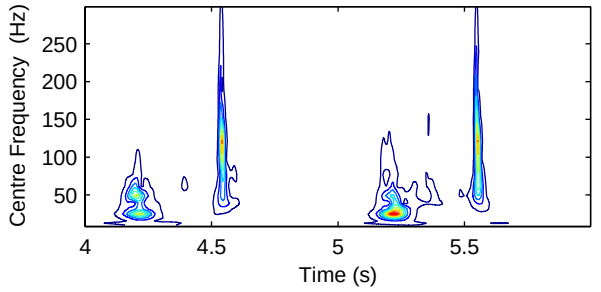
$$W_n = \sum_{i=1}^{M} s_i v_i(n) u_i = \sum_{i=1}^{M} \alpha_{i,n} u_i, \qquad (5)$$

where the weight associated with $u_i$ in the reconstruction of $W_n$ is:

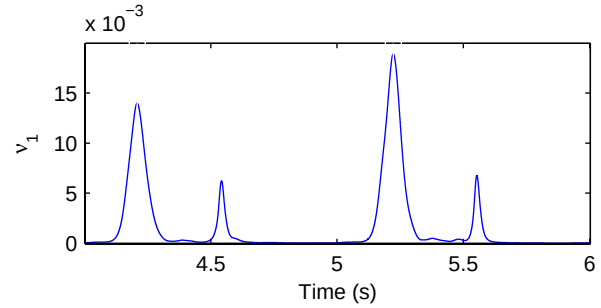
$$\alpha_{i,n} = s_i v_i(n) = u_i^T W_n. \qquad (6)$$




(a) Two seconds of the PCG.




(b) Contour plot of time-frequency coefficients.




(c) First right singular vector, $v_1$.

Figure 1: Example of the proposed detection strategy applied to PCG # 1006113.

The second equality in (6) follows from the orthonormality of the $\{u_i\}$.

For $W$ as given in (3), each of the first $D$ vectors $u_i, i = 1,2,\ldots,D$, can be thought of as a template of frequency bin values that resembles energetic and/or commonly occurring $W_n$, *i.e.* those that are associated with $S_1$, $S_2$, and murmurs. As $\alpha_{i,n}$ is proportional to $v_i(n)$, the time indices $n$ associated with peaks of $v_i$ can be used to identify the time at which recurring energetic events occur.

Figure 1 (c) shows the portion of the first right singular vector $v_1$ corresponding to the two second segment of the PCG in Figure 1 (a). Note that a peak in $v_1$ occurs when any $S_1$ or $S_2$ is present. For the PCG of Figure 1,

the $v_1$ peaks associated with $S_1$ events tend to be higher than those associated with $S_2$ events.

The event detection strategy consists of identifying recurring energetic $W_n$ by finding the values of $n$ associated with the peaks in $v_i(n)$, $i = 1,2,\ldots,D$. Let $p_i^{(k)}$ denote the local peak in the $k^{th}$ region, $k = 1,2,\ldots,K_i$ of $v_i(n)$ that exceeds a threshold $\rho_i$. Let $n = n_i^{(k)}$ denote the value of $n$ associated with the peak $p_i^{(k)}$. For each $k$, specify a segment $N_i^{(k)}$ of fixed size $N_{seg}$, centered at $n = n_i^{(k)}$. The resulting set of $K_1 + K_2 + \ldots + K_D$ segments constitute the *Claimed* events. As some events produce peaks in multiple singular vectors, the set of *Claimed* events is pruned with the goal of retaining only one segment per event.

Feature Extraction

The objective of feature extraction is to obtain a $D$-element vector $[p_1^{(k)}, p_2^{(k)}, \ldots, p_D^{(k)}]$ for each of the *Claimed* events. The $i^{th}$ feature is defined as $p_i^{(k)} = \max v_i(n) > \rho_i, n \in N^{(k)}$.

## EXPERIMENTATION

Data Description

The database of PCG's that was used for this experiment was obtained from Laboratory of Biomedical Engineering, Institut de recherches cliniques de Montréal (IRCM), University of Montreal. The PCG's were recorded using an Irex Medical System model 120-131 microphone placed externally on the tricuspid area of the chest wall. A standard lead II ECG was also recorded. The various events in the PCG's were not individually labelled by an expert.

All subjects had bioprosthetic valves implanted in their hearts. The recordings include examples of "normal" systolic murmurs, and "abnormal" stenotic or regurgitant murmurs.

Implementation

Each PCG was prepared for processing as follows [4]:(1) subtracting the sample mean, (2) low-pass filtering using a Butterworth filter of order 3 with a cutoff frequency of 300 Hz, (3) normalizing the data to lie between $\pm 1$, and (4) extracting 13 s of the PCG.

Matrix $W$ was computed for each PCG using $M = 16$ wavelet scales, which were chosen such that the values of $f_m$ were logarithmically spaced in the range $[\frac{10}{f_s}, \frac{300}{f_s}]$, where $f_s = 5$ kHz denotes the sampling frequency. The detection algorithm described above was applied to each PCG with $D = 3$ and $\rho_i = \frac{1}{N}\sum_{n=1}^{N} v_i(n)$. A fixed segment width of $N_{seg} = 625$ sa (125 ms) corresponding to one quarter of the shortest expected heart cycle of 500 ms, was used. With the exception of $N_{seg}$, all of the parameters listed here were chosen either experimentally or with reference to [4].

For each *Claimed* event, a feature vector of length $D = 3$ was extracted. We expect that the feature vectors are sufficiently similar for different events from the same class so as to form class-specific clusters in $D$-dimensional space. Classification was performed using unsupervised $K$-means clustering with $K = 3$. Cluster labels were assigned manually as follows. The cluster with the largest number of $S_1$'s was labelled as $C_{S_1}$. Similarly, labels $C_{S_2}$ and $C_{Other}$ were assigned to the remaining two clusters.

Detection Results

The accuracy of the detection stage was evaluated by comparing each of the *Claimed* events to the *True* $S_1$ and *True* $S_2$ segments in each PCG, which were first obtained via a semi-automated procedure that used the ECG as reference. A *Claimed* event was tallied as a $S_1$ *Hit* or a $S_2$ *Hit* if it overlapped with either a *True* $S_1$ or a *True* $S_2$, respectively. If not a *Hit*, a *Claimed* event was recorded as an *Other*. Percentages of *Hit* events were tabulated as $P_H = \frac{\#S_1 Hits + \#S_2 Hits}{\#True}$. For the purpose of comparison, the event detection algorithm described herein was also used to segment the SEE of the PCG [3]. The SEE of the PCG was calculated using overlapping windows of duration 0.2 s with an intra-window shift of 0.2 ms (1 sa).

The event detection results are listed in Table 1. The number of *Claimed* events increases with $D$. No more than two events were missed in 40 of the 42 PCG's. At $D = 3$ the detection results are comparable to those using the SEE; in addition, the proposed algorithm inherently provides information for event classification.

Classification Results

Clustering was performed on the feature vectors extracted from the *Claimed* events. As shown in the ex-

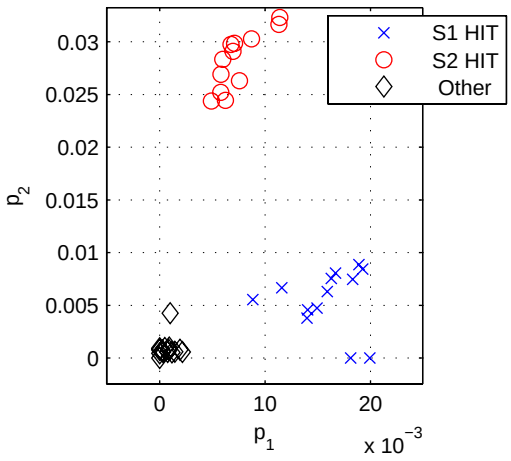



Figure 2: Projection of the first two proposed features for the *Claimed* events of PCG # 1006113.

Table 1: Event Detection Results for 42 PCG's

| Method | # *Hits* ($P_H$) 1224 (%) | # Others |
|---|---|---|
| D=1 | 1072 (87.6) | 45 |
| D=2 | 1181 (96.4) | 229 |
| D=3 | 1193 (97.5) | 327 |
| SEE | 1211 (98.9) | 405 |

ample of Figure 2, the feature vectors extracted from the *Others* are typically clustered close to the origin. The *Other* events (*i.e.* murmurs, other cardiac event, noise) tend to have smaller coefficients in $W_n$ because these events are less similar to the Morlet wavelet than are the $S_1$'s and $S_2$'s. The distributions of the remaining features in this example are also typical of most of the PCG that were analyzed in this work: the features associated with the $S_1$ *Hits* and $S_2$ *Hits* have maximum values along different dimensions (*e.g.* $p_1$ and $p_2$, respectively).

In general, the $S_1$ *Hits* formed a cluster separate from the $S_2$ *Hits*. Indeed, only 5.4% of $S_1$ *Hits* were labelled as $C_{S_2}$ and 15.0% of $S_2$ *Hits* were labelled as $C_{S_1}$. This result indicates that the class of $S_1$ events is sufficiently different from the class of $S_2$ events so as to be distinguished using unsupervised approaches.

The most common sources of classification error were: (1) the grouping of one or more *Others* in $C_{S_2}$; and (2) the misclassification of $S_1$ *Hits* in $C_{Other}$. The PCG's with low $P_H$ also tended to have poor classification results because there was typically an underrepresentation of one class of event in the set of *Claimed* events. That class was then clustered with the *Others*, whereas the well-detected class was divided between two clusters.

## CONCLUSIONS

The results demonstrate that PCG events can be detected and classified using a novel strategy that identifies peak values of the right singular vectors of a time-frequency representation of the PCG. The significance of the work is that the PCG events are segmented without using either subject-specific diagnostic information or a training set of labelled PCG events. Moreover, the events are classified without using *a priori* information regarding the frequency characteristics of different PCG event classes or timing guidelines regarding the expected durations of events and intra-event intervals. Future work includes developing a strategy to automatically assign cluster labels and to incorporate timing-based features to enhance the proposed classification strategy.

## ACKNOWLEDGEMENTS

The authors would like to thank Dr. L.-G. Durand and Dr. P. Pibarot of IRCM, Canada, for providing the data for this work.

## REFERENCES

[1] D. Labus, *Heart Sounds Made Incredibly Easy*. Lippincott Williams and Wilkins, 2005.

[2] R. Lehner and R. Rangayyan, "Three-channel microcomputer system for segmentation and characterization of the phonocardiogram," *IEEE Transactions on Biomedical Engineering*, vol. 34, no. 6, pp. 485–489, 1987.

[3] H. Liang, S. Lukkarinen, and I. Hartimo, "A heart sound segmentation algorithm based on heart sound envelogram," in *Computers in Cardiology*, 1997, pp. 105–108.

[4] S. Rajan, E. Budd, M. Stevenson, and R. Doraiswami, "Unsupervised and uncued segmentation of the fundamental heart sounds in phonocardiograms using a time-scale representation," in *Proc. of the 28th IEEE EMBS Annual International Conference*, 2006, pp. 3732–3735.