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

Unsupervised and Uncued Segmentation of the Fundamental Heart Sounds in Phonocardiograms Using a Time-Scale Representation

S. Rajan, E. Budd, M. Stevenson, R. Doraiswami

Abstract— **A methodology is proposed to segment and label the fundamental activities, namely the first and second heart sounds, S_1 and S_2, of the phonocardiogram (PCG). Information supplementary to the PCG, such as a cue from a synchronously acquired electrocardiogram (ECG), subject-specific prior information, or training examples regarding the activities, is not required by the proposed methodology. A bank of Morlet wavelet correlators is used to obtain a time-scale representation of the PCG. An energy profile of the time-scale representation and a singular value decomposition (SVD) technique are used to identify segments of the PCG that contain the fundamental activities. The robustness of the methodology is demonstrated by the correct segmentation of over 90% of 1068 fundamental activities in a challenging set of PCGs which were recorded from patients with normally functioning and abnormally functioning bioprosthetic valves. The PCGs included highly varying fundamental activities that overlapped in time and frequency with other aberrant non-fundamental activities such as murmurs and noise-like artifacts.**

I. INTRODUCTION

Phonocardiograms (PCGs) are recordings of acoustic waves produced by the pumping action of the heart. The PCG, or the heart sound signal, is a nonstationary signal with a recurring sequence of two types of acoustic events, S_1 (the first heart sound) and S_2 (the second heart sound), termed herein as the *fundamental activities*. The PCG may also include other cardiac events, such as murmurs, clicks, and snaps. These other events, along with noise and artifacts due to breathing and motion, are termed herein as the *non-fundamental activities*. The non-fundamental activities may overlap the fundamental activities in either time or frequency or both. 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. This makes the detection of fundamental activities a non-trivial task.

The activities in the PCG that relate to a given heart disease are contained in the specific intervals of the cardiac cycle [1], which are identified relative to S_1 and S_2. Thus, the identification of the fundamental activities plays an important role in the interpretation of PCGs. Due to the limitations of the human auditory system and the lack of physician training in cardiac auscultation [2], the PCG is often discarded as a diagnostic test for valvular dysfunction in favor of expensive tests, such as the echocardiogram, which are sometimes unavailable and subject to long wait times. A system that automatically identifies and labels each of the fundamental activities in the PCG without a synchronized electrocardiogram (ECG) could potentially be the first step toward either a diagnostic aid to physicians or a learning tool for cardiac auscultation. Such a tool could be readily available in a physician's office at low cost.

An example containing two cardiac cycles of a PCG signal with various activities is shown in Fig. 1 (a). It is clear from its synchronous ECG signal in Fig. 1 (d) that the QRS complex in the ECG provides a time cue for the segmentation of PCG activities. Indeed, most of the segmentation approaches reported in the literature require information from an extra biological signal. Readers are directed to [3] for a detailed survey of such methods. Literature that describes automated *uncued* segmentation of PCGs, i.e. segmentation without the use of a secondary signal, is sparse.

Unlike most approaches to PCG segmentation described in the literature, the proposed methodology is uncued and unsupervised: no training data or peripheral biological signals are required. Because the data set used in this work is not expertly labelled, no subject-specific or data-set-specific prior information is used to "fine-tune" the segmentation procedure. The prior information that is employed in the proposed methodology is limited to the following general characterizations of the PCG: (1) each subject's heart rate is less than 160 beats per minute; (2) the frequency range of the fundamental activities is limited to $[10, 300]$ Hz; and (3) the fundamental activities resemble the Morlet wavelet [3],[4]. It is assumed that the fundamental activities have higher energy than the non-fundamental activities.

The proposed methodology is different from the recently proposed time-domain-based uncued and unsupervised segmentation approaches that use discrete wavelet transforms [5], [6] and a matching pursuit method using Gabor atoms [1]. Unlike [5] and [6], the proposed segmentation methodology is done in the joint time-scale domain, and spectral energy information is utilized to distinguish the fundamental activities from other activities. Segmentation is based on the characteristics of the activities and there is no *ad hoc* choice of window lengths. Although both the proposed methodology and that described in [1] are time-scale-based, uncued, and unsupervised, the proposed methodology is the simpler as it avoids the use of recursive decomposition and localized clustering to collect activities that form a cardiac cycle.

The segmentation methodology proposed in this paper uses a time-scale representation of the PCG that is calculated

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

S. Rajan is with DRDC Ottawa,3701 Carling Avenue, Ottawa, Ontario, Canada K1A 0Z4 sreeraman.rajan@drdc-rddc.gc.ca

E. Budd, M. Stevenson, and R. Doraiswami are with the Department of Electrical and Computer Engineering, University of New Brunswick, Fredericton, New Brunswick, Canada E3B 5A3

using a bank of Morlet wavelet correlators. Segmentation is accomplished in three steps. First, peaks in the energy profile derived from the time-scale representation are identified and used to obtain segments containing activities. These segments are essentially two-dimensional joint time-scale representations of the activities. Second, the singular values of each segment are used to identify which segments contain fundamental activities. Third, information regarding the sequential repetition of the fundamental activities and the durations of the time intervals separating them are exploited to label the segments containing fundamental activities as either S_1 or S_2.

II. METHODOLOGY

A. Time-Scale Representation of the PCG

Let X denote a $N \times 1$ vector of samples of the PCG that is recorded from one subject. It is assumed that X is sufficiently long to contain multiple occurrences of both types of fundamental activities, S_1 and S_2. Let Y denote the 2-dimensional time-scale matrix representation of X which is obtained by passing the PCG through a bank of M Morlet wavelet correlators,

$$Y(m,n) = \sum_{k=1}^{N} \psi_m(k-n)X(k), \tag{1}$$

where $\psi_m, m = 1, \ldots, M$, denotes the Morlet wavelets. The Morlet wavelets were chosen because of their resemblance to the fundamental activities, namely S_1 and S_2 [3],[4] . The Morlet wavelet ψ_m is given by

$$\psi_m(k) = \pi^{-\frac{1}{4}} e^{-\frac{k^2}{2a_m^2}} e^{j\omega_m k}, \tag{2}$$

where $\omega_m = \frac{\omega_0}{a_m}$ is the distinct center frequency associated with ψ_m and $\omega_0 = 5$ rad/s is the center frequency of the mother Morlet wavelet. The scaling parameter a_m is related to ω_m and the window duration of ψ_m such that the same number of oscillations is present in each ψ_m. It is important to distribute ω_m appropriately over the typical frequency range of fundamental activities with the intent that one or more ψ_m will closely match any occurrence of a fundamental activity. This permits the use of the same set of correlating wavelets to obtain the time-scale representation of each PCG without any prior information regarding the specific frequency content of the S_1 and S_2 in a given PCG.

B. Segmentation of the PCG

The objective of segmentation is to divide Y along its temporal dimension such that each fundamental activity is approximately contained in a segment. Segmentation is based on the energy profile $E(n)$ of the PCG defined as follows:

$$E(n) = \sum_{m=1}^{M} |Y(m,n)|^2, n = 1, 2, \ldots, N. \tag{3}$$

Provided that the $\{\psi_m\}$ match the fundamental activity sufficiently well, a local peak in $E(n)$ will occur when the fundamental activity is present [3]. As shown in Fig. 1 (b), the dominant peaks in the energy profile are well localized with the fundamental activities. An automated process excises a segment of Y, denoted $A_i, i = 1, \ldots, I$, for each peak in $E(n)$. The $M \times N_i$ matrix A_i contains either a fundamental activity or a non-fundamental activity, and N_i is the number of contiguous time bins associated with A_i. To minimize any loss of information, the segmentation process is applied until every time bin of Y is assigned to a segment.

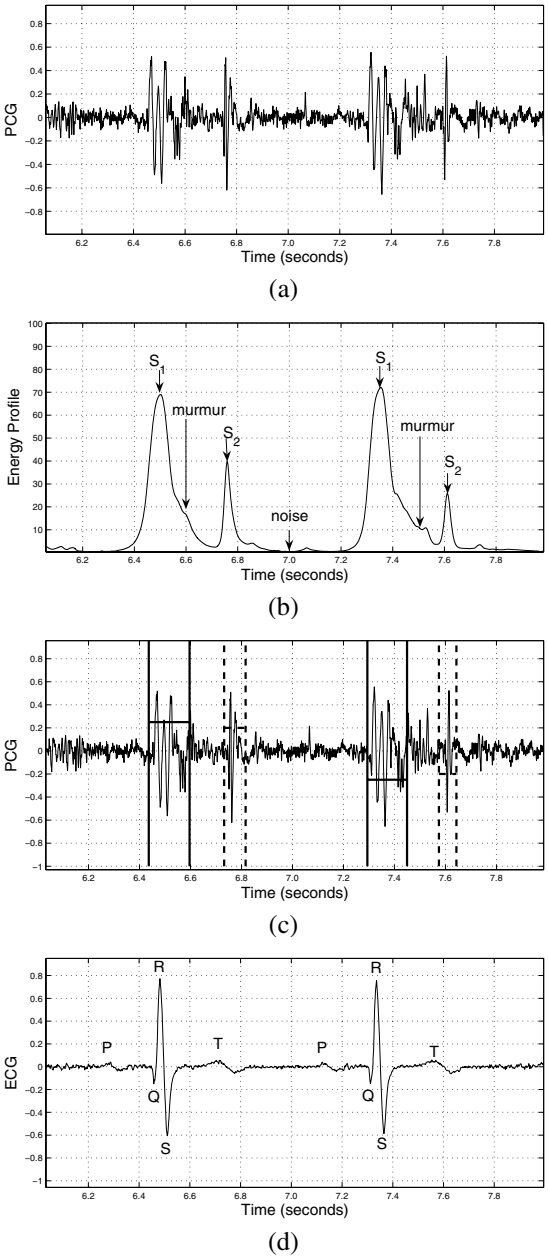

Fig. 1. An example of two cardiac cycles of a PCG and its corresponding segmentation results. (a) PCG. (b) Energy profile. (c) Segmentation output (solid line indicates segmented and labelled S_1; dashed line indicates segmented and labelled S_2). (d) ECG.

C. Segment Classification Using Singular Values

Because the fundamental activities are highly variable and training data are not available, an unsupervised approach is

formulated to classify each segment A_i as either a fundamental activity or a non-fundamental activity.

Using SVD, A_i can be written as [7]

$$A_i = U_i S_i V_i^T, \tag{4}$$

where U_i is a $M \times r$ matrix whose columns span the column-space of A_i, S_i is a $r \times r$ diagonal matrix of the singular values, $\{\sigma_{i,j}\}, j = 1, 2, \ldots, r$, of A_i, V_i is a $N_i \times r$ matrix whose columns span the row-space of A_i, and r is the rank of A_i.

Although the fundamental activities are nonstationary, each occurrence of a fundamental activity in a PCG is a structured narrowband signal with a limited number of spectral components [8] that is expected to match well with one or more of the Morlet wavelets $\{\psi_m\}$ [3],[4]. As a result, one can expect that a small number of singular vectors will typify each fundamental activity well. The singular values associated with these singular vectors will be large relative to all other singular values. Thus when A_i includes a fundamental activity, the largest d singular values, together with their associated singular vectors, will tend to characterize the fundamental activity; the remaining $r - d$ singular values/vectors will characterize the noise.

In the case that A_i does not include a fundamental activity, e.g. A_i contains either a murmur or artifact (both of which are typically broadband [1]), we might expect that a good characterization of A_i may require a larger number of singular values/vectors. Thus the dominant singular values are expected to be larger when a fundamental activity is present relative to when one is not present. A test statistic may be derived using the singular values to determine the presence of a fundamental activity. A decision strategy for classifying each A_i based solely on its d dominant singular values can be loosely formulated as: *assign A_i to the class of fundamental activities if its d dominant singular values are relatively large; assign A_i to the class of non-fundamental activities if its d dominant singular values are relatively small*. An automated classification procedure based on this strategy is presented in Section III-B.

D. Fundamental Activity Labelling

Label S_1 or S_2 is assigned to each of the segmented fundamental activities by exploiting the following characteristics of the PCG: (1) fundamental activities occur as a recurring sequence of S_1 and S_2, (2) the duration of the time interval between S_1 and S_2 (systole) is longer than the duration of the time interval between S_2 and S_1 (diastole), and (3) the heart rate is less than or equal to 160 beats per minute.

III. IMPLEMENTATION AND RESULTS

The database of PCGs that was used for this experiment was obtained from Laboratory of Biomedical Engineering, Institut de recherches cliniques de Montréal (IRCM), Université de Montréal. The PCG were recorded using an Irex Medical System model 120-131 microphone on the tricuspid area of the chest wall. The resolution of the analog-to-digital converter was 12 bits. A standard lead II ECG was also recorded. All subjects had bioprosthetic valves implanted in their hearts. The functionality of each subject's bioprosthetic heart valve was classified as either normal or abnormal by an experienced cardiologist. The various events in the PCGs were not labelled by the cardiologist.

PCGs with a high degree of variability and irregularity were selected in order to demonstrate the robustness of the proposed methodology. The proposed methodology was applied to 13 seconds of PCGs recorded from each of 42 different subjects: 21 subjects had normally functioning artificial heart valves while 21 subjects had abnormally functioning artificial heart valves. In total, there were 534 occurrences of each of S_1 and S_2. In these PCGs, the heart rate varied from 72 to 160 beats per minute. Signal characteristics such as heart rate, signal-to-noise ratio, energy, frequency spectrum content, event duration, and the presence of non-fundamental activities changed rapidly within each recording and also between recordings.

Three stages of pre-processing prepared each PCG for segmentation: (1) subtracting the sample mean, (2) low-pass filtering using a Butterworth filter of order 3 with a cutoff of 300 Hz, and (3) normalizing the data to lie between ± 1.

A. Time-Scale Representation of the PCG

The success of the methodology rests on a strong resemblance between the fundamental activities and the Morlet wavelets used in the correlator bank. A bank of $M = 31$ Morlet wavelet correlators is used to obtain the time-scale representation of the PCG. The center frequencies $\{w_m\}$ of the correlators were chosen to be logarithmically spaced between 10 Hz and 300 Hz. The high concentration of correlators helps to ensure a strong resemblance between the fundamental activities and one or more of the Morlet wavelets. The choice of a redundant representation also allows the same bank of correlators to accommodate frequency variations in S_1 and S_2 across all subjects and heart rates.

B. Segmentation of the PCG

Following the segmentation and the SVD of each A_i in Y, an unsupervised threshold-based approach was used to classify the segments as either containing a fundamental activity or not. The most dominant singular value of each segment, i.e. $d = 1$, was compared to a data-adaptive threshold. As explained in Subsection II-C, the most dominant singular vectors of the time-scale representation of the PCG are expected to typify the fundamental activities (if they are present), whereas the less dominant singular vectors are expected to typify the non-fundamental activities. Since X, the vector of samples of the PCG, includes both fundamental and non-fundamental activities, the median singular value of its time-scale representation Y was chosen heuristically to be the classification threshold. If the dominant singular value of A_i was greater than this typically low threshold, then A_i was assigned to the class of fundamental activities. Otherwise, A_i was assigned to the class of non-fundamental activities. This decision rule ensured that all fundamental activities were assigned correctly. Finally, each A_i that was

TABLE I

PERCENTAGE OF ACCURATELY SEGMENTED S_1 AND S_2 OF NORMAL AND ABNORMAL PCGS

Fundamental Activity	Normal † (269)	Abnormal (265)
S_1	90.3 (243)	90.2 (239)
S_2	91.1 (245)	90.6 (240)

† In this study, a "normal" PCG refers to a PCG that was recorded from a subject with a normally functioning bioprosthetic heart valve.

classified as containing a fundamental activity was labelled as either S_1 or S_2.

C. Results

The performance of the proposed segmentation methodology was measured as a percentage of the "true" S_1 and S_2 activities that were "accurately" segmented. A fundamental activity S_1 or S_2 was counted as accurately segmented if the labelled S_1 or S_2 region of the PCG aligned with the appropriate events in its corresponding, synchronously recorded, ECG signal. The proposed methodology correctly segmented 90.5 % of the fundamental activities. For example, since the region of the PCG that is labelled S_1 (i.e. solid line) in Fig 1 (c) coincides with the peak in the QRS complex of the ECG in Fig 1 (d), it was counted as a correctly segmented S_1. As the proposed methodology is uncued, the ECG signal was used for performance evaluation only. A summary of the results is presented in Table I.

A particularly promising aspect of the results is the small amount of degradation (i.e. less than 1 %) in performance for abnormal PCGs as compared to normal PCGs. Abnormal PCGs present a greater challenge because their fundamental activities are subject to greater variations in energy and have non-fundamental activities interspersed. Although the energy of the artifacts may be high relative to that of the fundamental activities the fundamental activities in the regions of the artifacts were, in general, correctly segmented.

D. Discussion

For most PCGs in the data set, the proposed methodology was robust to the irregularities and complexities of the fundamental activities as well as to the artifacts and noise that are inherent to PCGs.

The performance of the proposed methodology is adversely affected when any of the following occurs within a single PCG: the energies of the fundamental activities vary significantly, the heart rate varies considerably, there is low signal-to-noise ratio of the fundamental activities relative to the non-fundamental activities, or heart murmurs have higher energy and are structured. A murmur that has structure resembles a narrowband signal. If the murmur also has relatively high energy, the output of the correlators due to the murmur may have energy greater than that of either one or both of the fundamental activities. This may result in segmentation errors.

IV. CONCLUSIONS AND FUTURE WORKS

The results of the proposed methodology for PCG segmentation demonstrate that fundamental activities can be identified without the use of a synchronous ECG, labelled training data, or subject-specific characteristics of the fundamental activities. The results also show that a time-scale representation is an appropriate choice for the segmentation of PCG signals. The success of the proposed methodology rests on the choice of correlating signals, in this case the Morlet wavelets, that resemble the fundamental activities. This paper has also illustrated that the dominant singular values of the time-scale representation of PCG segments can be used to distinguish between those that contain fundamental activities and those that do not.

Although the choice of the dominant singular values of the time-scale representations of PCG segments for segmenting the fundamental activities has been demonstrated in this paper, the study of other time-scale features for the segmentation and labelling of both fundamental and non-fundamental activities of the PCG is ongoing. The implementation of the proposed methodology as part of an automated and uncued procedure to enhance the interpretation of PCGs by physicians, as either a diagnostic aid or a learning tool, is also being considered.

V. 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] H. Sava, P. Pibarot, and L. Durand, "Application of the matching pursuit method for structural decomposition and averaging of phonocardiographic signals," *Medical and Biological Engineering and Computing*, vol. 36, no. 3, pp. 302–308, 1998.

[2] S. Mangione and L. Z. Nieman, "Cardiac auscultatory skills of internal medicine and family practice trainees: A comparison of diagnostic proficiency," *Journal of American Medical Association*, vol. 278, no. 9, pp. 717–722, 1997.

[3] S. Rajan, "A wavelet-based correlator approach for segmentation and classification of the first two sounds in a PCG signal," Ph.D. dissertation, The University of New Brunswick, Canada, May 2004.

[4] S. Rajan, R. Doraiswami, M. Stevenson, and R. Watrous, "Wavelet based bank of correlators approach for phonocardiogram signal classification," in *Proc. of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, 1998, pp. 77–80.

[5] H. Liang, S. Lukkarinen, and I. Haritmo, "A heart sound segmentation algorithm using wavelet decomposition and reconstruction," in *Proc. Int. Conf. in IEEE Engineering in Medicine and Biology*, vol. 4, 1997, pp. 1630–1633.

[6] H. Liang and I. Hartimo, "A feature extraction algorithm based on wavelet packet decomposition for heart sound signals," in *Proc. of the IEEE-SP Int. Symposium on Time-Frequency and Time-Scale Analysis*, 1998, pp. 93–96.

[7] I. T. Joliffe, *Principal Component Analysis*. Springer-Verlag New York, Inc., 2002.

[8] H. Sava and J. T. McDonnell, "Spectral composition of heart sounds before and after mechanical heart valve implantation using a modified forward-backward Prony's method," *IEEE Trans. on Biomedical Engineering*, vol. 43, no. 7, pp. 734–742, 1996.