Efficient method for events detection in phonocardiographic signals

Juan Martínez-Alajarín and Ramón Ruiz-Merino

Departamento de Electrónica, Tecnología de Computadoras y Proyectos,
Universidad Politécnica de Cartagena,
C/ Dr. Fleming s/n, 30202 Cartagena, SPAIN

ABSTRACT

The auscultation of the heart is still the first basic analysis tool used to evaluate the functional state of the heart, as well as the first indicator used to submit the patient to a cardiologist. In order to improve the diagnosis capabilities of auscultation, signal processing algorithms are currently being developed to assist the physician at primary care centers for adult and pediatric population. A basic task for the diagnosis from the phonocardiogram is to detect the events (main and additional sounds, murmurs and clicks) present in the cardiac cycle. This is usually made by applying a threshold and detecting the events that are bigger than the threshold. However, this method usually does not allow the detection of the main sounds when additional sounds and murmurs exist, or it may join several events into a unique one. In this paper we present a reliable method to detect the events present in the phonocardiogram, even in the presence of heart murmurs or additional sounds. The method detects relative maxima peaks in the amplitude envelope of the phonocardiogram, and computes a set of parameters associated with each event. Finally, a set of characteristics is extracted from each event to aid in the identification of the events. Besides, the morphology of the murmurs is also detected, which aids in the differentiation of different diseases that can occur in the same temporal localization. The algorithms have been applied to real normal heart sounds and murmurs, achieving satisfactory results.

Keywords: Phonocardiogram, heart murmurs, event detection, segmentation, signal processing

1. INTRODUCTION

Cardiac auscultation is, together with the electrocardiogram (ECG), the first basic analysis tool used to evaluate the functional state of the heart, as well as the first indicator used to submit the patient to a cardiologist. However, modern medical image techniques (mainly echocardiography) have made auscultation lost its importance as a diagnosis tool. Because of the development of these new techniques, the teaching of auscultation has moved backwards, so a generalized lost of the basic auscultatory skills of the physicians has become very common nowadays.[1] Besides, the frequent use of the image diagnosis techniques and their high cost make them available in few hospitals only, which increases delays in waiting lists.

However, the low cost and the potential capabilities of auscultation, together with the recent developments in microelectronics and signal processing techniques, have increased in the last years the interest in auscultation. This is a cheap, quick, simple and non-invasive analysis tool, but lacks objectivity. For that reason, there is currently great interest in the development of automatic systems capable of aiding the physician in the diagnosis of the valvular state of the heart by analyzing the sounds and murmurs present in the cardiac cycle.[2] The ideal system should be able to acquire the cardiac sounds from the patient and provide an objective diagnosis, using processing algorithms to analyze and interpret the phonocardiographic signal in an automatic way.

There exists many papers in which custom algorithms have been developed for the distinction between normal sounds and two or more valvular diseases.[2–5] In many cases,[2] the basic events of the cardiac cycle are not identified, nor processed individually; however the energy fingerprint of the whole cardiac cycle is associated

Further author information: (Corresponding author: J. M.-A.)
J. M.-A.: E-mail: juanc.martinez@upct.es
R. R.-M.: E-mail: ramon.ruiz@upct.es

with the corresponding valvular disease. This method is simple, but does not take into consideration many information that could be useful to improve the diagnosis.

During auscultation, the physician detects the cardiac events, which are later evaluated according to some properties (intensity, temporal localization, duration, ...). Finally, the physician establishes an association between the properties of that events (already identified as sounds or murmurs), and the valvular condition of the heart.

The phonocardiogram (PCG) is the graphical representation of the heart sounds. Phonocardiography overcomes some of the drawbacks of traditional auscultation, providing important information in the detection of heart valve disorders. However, the interpretation of the phonocardiograms still remains a very difficult task, due to the many variables that have influence on the generation and transmission of heart sounds.

In this paper we present a hierarchical structure that provides a methodology for the analysis and interpretation of the phonocardiographic signal, following a procedure similar to that of the physician during auscultation. This structure is based on signal abstraction, and is organized in four levels, each one with several processing blocks aimed at very specific tasks, that cover all the processing needed from the acquisition of the phonocardiogram to the final diagnosis. A briefly description of the processing blocks of the hierarchy will be presented, and then we will focus on some of the most important tasks, especially in the event detection algorithm. The method used improves other event detection methods and provides an efficient technique to detect all the important cardiac events, no matter the intensity of the sounds or murmurs, or the proximity between events.

The application of this system is aimed at the medical prediagnosis, especially in primary attention. It can be used as a low cost screening technique to make objective the auscultatory perception of the physician, and to give priority to waiting lists for the echocardiographic analysis of the Cardiology Services.

2. DATA ACQUISITION

To develop the processing algorithms, sounds acquired in the Cardiology Service of the Hospital General Universitario of Murcia and sounds of a commercial database[6] have been used. The former have been recorded with an electronic stethoscope (model Androscope i-stethos, from Andromed) using the audio input of a laptop, with a sampling frequency of 22050 Hz and 16 bits accuracy.

To record the sounds, a graphical user interface has been developed specifically for this task (Fig. 1). This environment includes the usual characteristics in this kind of interfaces, like the following: the duration of the records can be selected by the user, the data recorded are saved as WAV and ASCII files, acquisition of a second data channel (peripheral pulse), visualization of PCG and pulse signals, reproduction of PCG signal, and zoom capabilities. The peripheral pulse signal is recorded not to act as a reference signal, but to validate the segmentation computed from the PCG signal solely.

Besides, some other not so frequent capabilities have been also implemented, like the recording of observations of the patient auscultation, selection of the sensor used (bell or diaphragm) and the area of auscultation, as well as the recording of auscultation maneuvers, like Valsalva, sitting, handgripping, ..., which allow the discrimination among several types of pathological conditions. In this way, all the necessary information is recorded to be further processed.

3. HIERARCHICAL STRUCTURE

An excellent methodology of the problem of defining hierarchical structures for signal processing can be realized using several levels of abstraction,[7] each one associated with some local information that will be shared between them. In our system, besides the multiple levels and their associated data, several *specialists* per level have been defined, which are processing blocks aimed to specific tasks of analysis or extraction of information of the related signals. Each specialist is defined with some input data, some grouping and processing operations for the computation of new signals and characteristics, and some output data. These output data will be available to all other specialists from the same and higher levels.

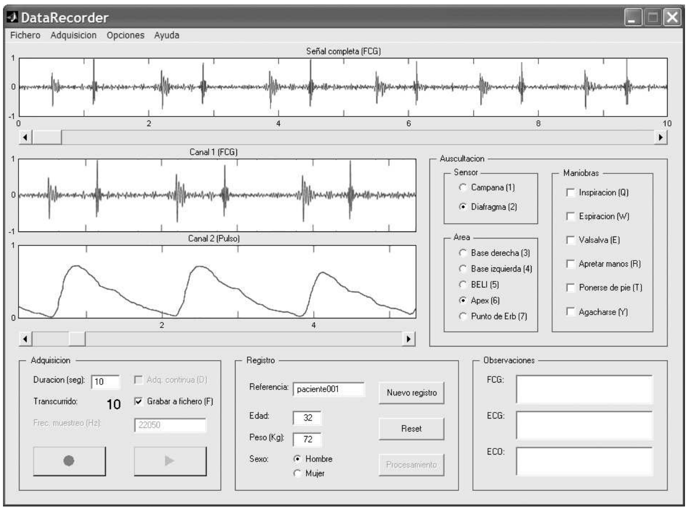

Figure 1. Graphical user interface developed to acquire phonocardiographic signals.

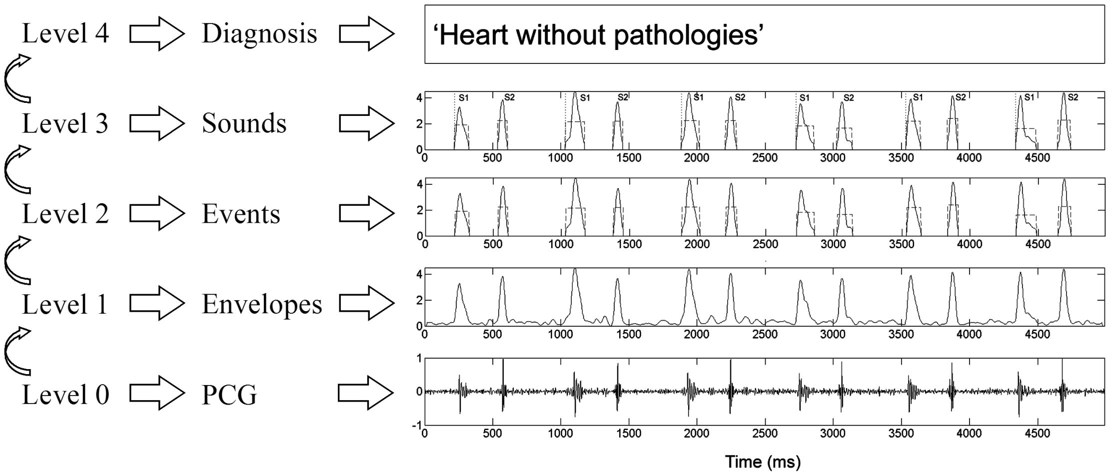

Figure 2. Main signals associated with the different levels of the PCG processing hierarchy. Level 0 (PCG) does not really belong to the hierarchy; it is only the input signal to the hierarchy.

The lowest level of the hierarchy performs intensive numeric computations, whereas as we climb the hierarchy, there is a progressive information abstraction process, passing from quantitative information (lower levels) to qualitative information (higher levels).

Other methodologies similar to that proposed in this paper (tasks specialists in hierarchical systems for information processing) have been also applied successfully to other biomedical environments, like the detection of cardiac ischaemia,[8] or the identification of the main heart sounds in fetal phonocardiographic signals,[9] using meta-knowledge areas hierarchical organized.

The proposed hierarchy is organized in four levels, each one associated with a main signal (Fig. 2):

- Level 1: envelopes. The main signals in this level are the envelopes of the PCG, which give an approximate representation of several instantaneous magnitudes extracted from the PCG, like the energy or frequency.

- Level 2: events. In this level, the main signals are the events detected in PCG. These events correspond to portions of the envelopes that should be analyzed in higher levels to obtain important information needed for the final diagnosis.
- Level 3: sounds: The signals associated with this level are the events already delimited and identified as sounds, clicks, or murmurs, and the PCG segmented in cardiac cycles.
- Level 4: diagnosis. In the highest level, the associated signal is the diagnosis achieved after processing the PCG in the lower levels. The diagnosis expresses the pathological condition of the heart in a form to be easily understood by the physician and the patient.

The signal that acts as input to the hierarchy (Level 0) corresponds to the sound waves that the heart generates (that is, the PCG, when they are graphically represented), although this signal does not correspond to any level in the hierarchy. Next the specialists in each level of the hierarchy are described (Fig. 3).

3.1. Level 1: Envelopes

- Signal conditioning: the PCG is decimated from the original sampling frequency of the recording to the working sampling frequency (4410 Hz) and then its range is normalized (mean = 0 and standard deviation = 1) to avoid differences in intensity caused by physiological characteristics of the patient (obesity, . . .).
- Filtering: the decimated and normalized PCG is then filtered with a Chebyshev type I low-pass filter (4th order, 0.5 dB ripple, 800 Hz cutoff frequency) to remove the high frequencies due to ambient noise and interferences, and then with a Chebyshev type I high-pass filter (4th order, 0.5 dB ripple, 20 Hz cutoff frequency) to remove the low frequencies recorded due mainly to muscle movements.
- Envelopes: the envelopes of instantaneous amplitude, energy, and frequency of PCG are computed as a moving average of these instantaneous magnitudes. This is done because the envelopes are easier to process than PCG, while they retain enough information. The computation of the envelopes is presented in Section 4.1, and further details about the instantaneous frequency can be found in Section 4.2.

3.2. Level 2: Events

- Cardiac rhythm: the average heart rate of the PCG is computed from the autocorrelation signal of the combined envelopes. In this way, information in temporal domain as well as in frequency domain is considered, so the average heart rate can be computed more accurately. No auxiliary signals are used to compute the average heart rate, only the PCG. Details about this task can be found in Section 4.3
- Signal abstraction: this task determines the portions of the PCG that are interesting to be further processed.
- Event detection: the events that form the PCG are accurately detected. This is necessary because in the previous specialist, two or more events could be grouped in only one macroevent, especially in those cases where a murmur follows or precedes a sound. The algorithm for an accurate event detection is detailed in Section 4.4.

3.3. Level 3: Sounds

- Sound identification: once the events have been correctly detected, they are identified and labeled with the name of the cardiac event (S1, S3, diastolic murmur, . . .).
- Cycle segmentation: the PCG signal is segmented in cardiac cycles using the average heart rate and the labeled events.
- Computation of prototypes: a prototype of each pathology is built from those cardiac cycles that present a certain degree of similarity among them. Other cycles will be then compared to the prototype cycle to determine how similar they are.

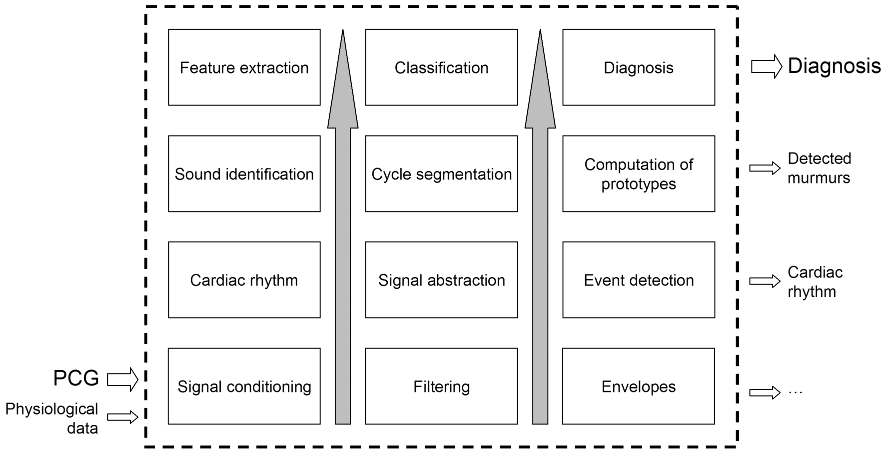

Figure 3. Specialists of each one of the four levels of the PCG processing hierarchy.

3.4. Level 4: Diagnosis

- Feature extraction: a set of discriminant features is extracted from the identified events and the segmented cycles. These features will be used as inputs of the classifier.
- Classification: a supervised classifier associates the set of extracted features with the pathologies that can be detected by the system.
- Diagnosis: using the previous classification and the physiological data of the patient, the final diagnosis is presented to the physician or patient in a format that is easily understood.

4. DESCRIPTION OF THE ALGORITHMS

This section describes the algorithms of some of the tasks indicated in Section 3: computation of the envelopes, computation of the instantaneous frequency, determination of the average heart rate, and accurate event detection. We will focus on the event detection algorithm, since it is a crucial task for the processing of the higher levels of the hierarchy.

Since the final objective is to mimic a traditional stethoscope (improving it including diagnosis capabilities), the processing algorithms do not use any auxiliary signals like ECG, the peripheral pulse or the carotid pulse, and rely solely on the processing of the PCG signal.

4.1. Envelopes of Instantaneous Magnitudes

After the original PCG signal has been decimated, normalized and filtered, three instantaneous magnitudes (amplitude, energy and frequency) are computed from the resulting signal, $x(n)$, defined as follows:

- instantaneous amplitude (IA):
$$y_{IA}(n) = |x(n)|\,, \tag{1}$$
- instantaneous energy (IE):
$$y_{IE}(n) = x(n)^2\,, \tag{2}$$
- instantaneous frequency[10] (IF):
$$y_{IF}(n) = \frac{F_\mathrm{s}}{2\pi}\left[\frac{\arg[z(n+1)] - \arg[z(n-1)]}{2}\right] \bmod 2\pi\,, \tag{3}$$

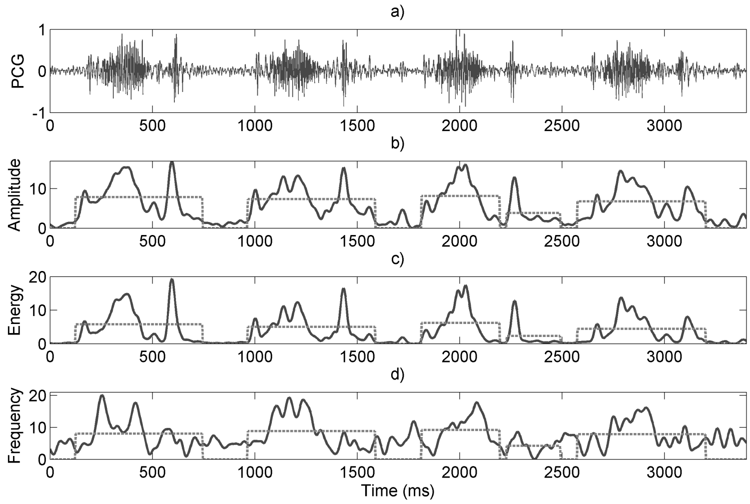

Figure 4. Representation of envelopes (solid lines) and abstracted signals (dash lines) of a real abnormal sound (mid-systolic murmur of aortic stenosis): a) PCG, b) amplitude, c) energy, and d) frequency.

where F_s is the sampling frequency, and $z(n)$ is the analytic signal (see Section 4.2).

After that, the envelopes are computed for each one of these signals by applying a moving average filter,[11] with a window length of 60 ms, and 57 ms of overlapping. Each sample in the envelope signal was computed as

$$G_j(m) = \frac{1}{P} \cdot \sum_{i=1}^{P} y_j(i) \cdot w(i)\,, \tag{4}$$

where $m = 1, \ldots, M$, $j = \{IA, IE, IF\}$, M is the length of the envelope signals, P is the number of samples in the window (corresponding to 60 ms), and w is the triangular window (used to smooth the envelopes).

Then, the envelopes were normalized (mean = 0 and standard deviation = 1) to avoid different range scales, with

$$H_j = \frac{G_j - \mu_{G_j}}{\sqrt{\sigma_{G_j}}}\,, \tag{5}$$

where μ_{G_j} and σ_{G_j} are, respectively, the mean and variance of G_j. The last step involved the subtraction of the minimum value of each normalized envelope from all the signal, in order to achieve a minimum envelope value of 0. Finally, all the values smaller than -20 dB than the maximum peak in the amplitude envelope were zeroed in the three envelopes.

These three envelopes allow to obtain different types of information. The amplitude envelope gives information about the amplitude of the PCG, and in some cases (where the noise level is low), it is enough to detect the cardiac events. Secondly, the energy envelope provides useful information about the energy of the signal, remarking those sounds with high intensity (these sounds are often the first and second heart sounds). On the contrary, weak sounds with low amplitude (like the fourth heart sound, or noise) will probably be buried, but this is not a problem since this envelope will be used only to detect the sounds with higher intensity (for the rest of the sounds, the amplitude envelope will be used). Finally, the IF envelope gives important information about the frequency content of the signal (the meaning of this information is discussed in Section 4.2), necessary to complement the temporal information obtained from the amplitude and energy envelopes. Fig. 4 represents the PCG signal acquired for an abnormal sound (mid-systolic murmur of aortic stenosis) and the three envelopes calculated (solid lines), together with the abstracted signals (dash lines).

4.2. Instantaneous Frequency

The *instantaneous frequency*, f_i, is commonly defined as the derivative of the phase of the *analytic signal*,

$$f_i(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt}\,, \tag{6}$$

where $\phi(t)$ is the phase of the analytic signal $z(t) = a(t)\exp[j\phi(t)]$, generated from the real signal, $x(t)$, by

$$z(t) = x(t) + j\mathcal{H}[x(t)]\,, \tag{7}$$

where $\mathcal{H}$ is the Hilbert Transform.

For discrete signals, $x(n)$, the discrete instantaneous frequency (DIF) is defined[12] as

$$f_i(n) = \frac{F_s}{2\pi}\left[\frac{\arg[z(n+1)] - \arg[z(n-1)]}{2}\right] \bmod 2\pi\,, \tag{8}$$

where F_s is the sampling frequency and $z(n)$ is the analytical signal, defined by

$$z(n) = x(n) + j\mathcal{H}[x(n)]\,. \tag{9}$$

Initially, the IF was considered as the average frequency of the signal, but, finally, it has been demonstrated[13] that for multicomponent signals (like the PCG), these two concepts represent different quantities. Nevertheless, the IF has been proposed[14] as a useful technique to characterize the frequency information of the PCG. Specific methods for estimating the IF for PCG signals have been also developed.[15] The IF can be also calculated as the first moment with respect to frequency from the Wigner-Ville distribution, and some parameters can be extracted from the IF signal in order to classify the sounds recorded.[10] The IF is computationally less intensive than other time-frequency techniques such as wavelets or the spectrogram, and it can characterize the frequency information of the PCG with enough accuracy to allow the identification and classification of the sounds.

4.3. Average Heart Rate

Since no auxiliary signals are used in the processing algorithms to achieve the final diagnosis, the heart rate detection must rely solely on the PCG. When the sounds analyzed are normal and contain only the first and second sounds (S1 and S2), the heart rate detection can be performed easily by detecting the peaks higher than a certain threshold. However, when the PCG corresponds to an abnormal sound, these considerations can not be taken into account, since the presence of additional sounds, clicks, or murmurs does not allow to decide easily the type of each sound, and so, the labeling of the sounds becomes a difficult task and the heart rate detection can not be performed from the peaks detected above the threshold.

The method proposed in this paper to compute the average heart rate of the PCG is based on the autocorrelation function (ACF) of the product (Q) of the three modified envelopes H_j, previously normalized so their maximum value was 1:

$$Q = \frac{H_{IA}}{\max(H_{IA})}\cdot\frac{H_{IE}}{\max(H_{IE})}\cdot\frac{H_{IF}}{\max(H_{IF})}\,, \tag{10}$$

$$ACF = \frac{Q * Q}{\max(Q * Q)}\,. \tag{11}$$

In this way, the resulting signal ACF is in the range $[0, 1]$, has $2M - 1$ samples, its maximum value is 1 and is located in the middle point, and it is symmetric with respect to this point, so only half the signal can be analyzed (for example, its right half, which corresponds to the last M samples). This leads to a typical signal like that depicted in Fig. 5a, where several relative maxima can be detected (Fig. 5b). These peaks form a set of candidates, but only some of them, those who really define the heart rate (Fig. 5c), will be selected to compute the average heart rate, and the rest will be discarded.

The criterion used is that a candidate peak will be selected to compute the heart rate if it is larger than all the candidates at its right (in the case of the right half of the ACF signal). If not, the candidate will be discarded. For example, in Fig. 5b, the candidates B and C have been discarded. Finally, the average heart rate can be computed as the mean value of the intervals defined by the selected relative maxima A, D, E, F, and G (Fig 5c).

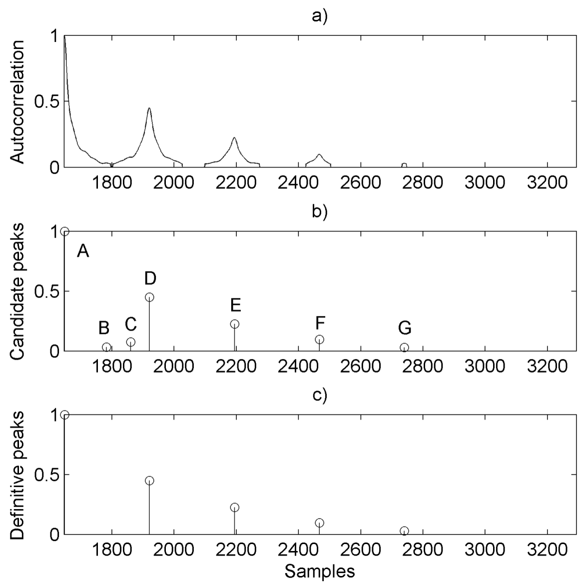

Figure 5. Detection of the average heart rate for a normal sound: a) right half of the autocorrelation signal, b) set of candidate peaks, c) set of relative maxima that really defines the heart rate. Horizontal scale represents the sample number of the autocorrelation signal.

4.4. Event Detection

The signal abstraction specialist uses a very basic method to obtain the portions of the PCG (or envelopes) where there is useful information to process (events), and the portions where only ambient noise exists (silences). For that reason, in some cases (especially where the PCG record contains murmurs), an event can be formed really by two or more events (then it is called a *macroevent*). An accurate separation between the individual events of a macroevent is then mandatory, in order to perform correctly the further processing.

The delimitation of the individual events inside a macroevent is a difficult task, because the events are usually very near, and their separation may be not clear using threshold techniques. Time-frequency techniques like the spectrogram could then aid in the delimitation of events, but these methods are time consuming when they are applied repetitively.

The analysis of sets of local parameters extracted from each relative maximum peak[7] is a method that can be applied to the amplitude or energy envelopes, in order to determine if the peak corresponds to an event or it is a small variation of the envelope. For that reason it is desirable to have a smooth envelope, so that very small peaks do not increase the computational complexity of the algorithm. Each event consists of a main point (A), and 5 additional envelope points (B, C, D, E, and F). Each point i is defined by a value (x_i^j) and a temporal localization (t_i^j), where $j = \{A, B, C, D, E, F\}$, $i = 1, \ldots, N$, and N is the number of peaks. The meaning of each point is the following:

- A: relative maximum peak,
- B: nearest relative minimum at the left of A,
- C: nearest relative minimum at the right of A,
- D: nearest point at the left of A that satisfies that $x^A/x^D \geq K_1$, where $K_1 > 1$,
- E: nearest point at the right of A that satisfies that $x^A/x^E \geq K_1$,
- F: nearest point to A that satisfies that $x^F \geq x^A$.

Also it is possible that an event consists of two or more peaks (e.g., a murmur or a split second sound). The peaks that form the event can be joined later, by analyzing the characteristics of amplitude, energy and frequency extracted from the different envelopes.

Next the pseudo-code of the event detection algorithm is presented:

1. all the relative maxima and minima of the amplitude envelope are computed (vectors Mx and Mn, respectively),
2. for each relative maximum Mx_i, the pairs (t_i^j, x_i^j) are computed, with $j = \{A, B, C, D, E, F\}$,
3. those relative maxima that satisfy at least one of the following conditions are removed:
   (a) if $\frac{x_i^A}{\max_i(x_i^A)} < K_2$, or
   (b) if $|t_i^B - t_i^C| < K_3$, or
   (c) if $|t_i^F - t_i^A| < K_4$,
4. the pairs (t_i^j, x_i^j) are recomputed for the remaining Mx_i,
5. those adjacent events that satisfy all of the following conditions are joined:
   (a) $\frac{\max(\mu_{i+1}^{AI}, \mu_i^{AI})}{\min(\mu_{i+1}^{AI}, \mu_i^{AI})} < K_5$, where μ_i^{AI} is the mean value of the event i in the AI envelope, defined from t_i^B to t_i^C,
   (b) $\frac{\max(\mu_{i+1}^{EI}, \mu_i^{EI})}{\min(\mu_{i+1}^{EI}, \mu_i^{EI})} < K_6$, where μ_i^{EI} is the mean value of the event i in the EI envelope, defined from t_i^B to t_i^C,
   (c) $\frac{\max(\mu_{i+1}^{FI}, \mu_i^{FI})}{\min(\mu_{i+1}^{FI}, \mu_i^{FI})} < K_7$, where μ_i^{FI} is the mean value of the event i in the FI envelope, defined from t_i^B to t_i^C,
6. a vector of parameters R_i is computed for each remaining event i defined by a pair (t_i^A, x_i^A),

Constants K_1 to K_7 must be experimentally determined. Once the events have been delimited by the pairs (t_i^j, x_i^j) and the vectors of parameters R_i have been computed (forming the parameter matrix $\mathbf{R}$), each event is first analyzed to determine if it is a murmur. The criterium used is that the area under the event in the IF envelope must be high, and the duration of the event must be above a threshold.

When an event is detected as a murmur, the temporal morphology of the murmur is also analyzed. This is carried out by dividing the duration of the murmur in three parts (beginning, middle and ending), and computing the normalized energy of each part with respect to the energy of the whole murmur. The relations between the energies of the three parts allow the determination of the morphology of the murmur: decrescendo, crescendo, crescendo-decrescendo, decrescendo-crescendo, and constant.

5. RESULTS

The two lower levels of the hierarchy have been developed and validated, whereas the third and fourth levels are currently under development. Concerning the results obtained for the two lower levels, two real signals will be used as test: a) a crescendo-decrescendo mid-systolic murmur of aortic stenosis (AS), and b) a constant mid-systolic murmur of mitral regurgitation (MR). These signals are illustrated in Figs. 4a and 6a (AS) and 7a (MR).

After the heart sounds have been acquired, the PCG is the input to the hierarchical system, which carries out the processing tasks defined for each specialist and level. Initially, the envelopes of instantaneous amplitude, energy and frequency are computed (Figs. 4b, 4c and 4d).

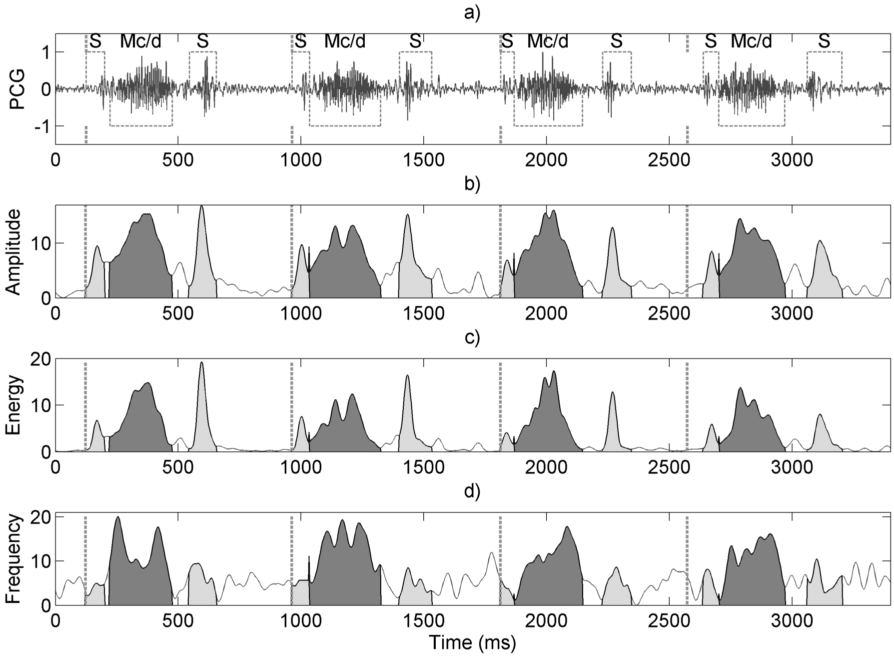

Figure 6. Representation of events detected in an abnormal record (mid-systolic murmur of aortic stenosis): a) PCG, b) amplitude, c) energy, and d) frequency. In a), events detected as sounds have been represented with dash lines over the 0 baseline and with the label "S", and events detected as murmurs have been represented under the 0 baseline and with the label "M". The label "c/d" stands for a crescendo-decrescendo murmur. In b), c), and d), sounds and murmurs have been represented, respectively, in light and dark gray, and the beginning of the cardiac cycles has been indicated as dash lines.

Signal abstraction is performed after the computation of the envelopes, obtaining those portions of the PCG which may need further processing to obtain the final diagnosis. For some PCG signals, especially those which contain murmurs, like Fig. 4a, usually two or more cardiac events appear joined in a unique abstracted signal, as it can be seen in Figs. 4b, 4c and 4d in dash lines. In these figures, S1, the mid-systolic murmur, and S2 have been joined in a unique macroevent in three of the four cycles of the signal. An accurate detection of events is thus needed to correctly detect all the individual events in the PCG.

Concerning the computation of the heart rate from the autocorrelation signal of the envelopes, this method has proven very robust. The average heart rate (vertical dash lines in Figs. 6b, 6c, and 6d) was computed accurately not only for normal and additional sounds, but also for murmurs, including continuous murmurs.

It is necessary to remark the importance of a correct detection and delimitation of all the events for the final diagnosis, since this task is really that of mimics the behavior of the physician during auscultation. These events are the basic units of analysis for the upper half of the hierarchy, so an incorrect delimitation would lead to important errors in the upper levels.

The method used to improve the event detection and delimitation has successfully separated the macroevents shown in Figs. 4b, 4c and 4d in the atomic events that form these signals, shown in Fig. 6a. The events detected as sounds (S1 and S2) have been labeled as "S", whereas murmurs have been labeled as "M".

In Figs. 4b, 4c and 4d, it is shown how the intensity and energy of the murmur (dark gray) are often larger than those of the main sounds, S1 and S2 (light gray). For this reason, an approach based on thresholds to detect the main sounds would not work. Besides, it is clearly shown that the instantaneous frequency for the mid-systolic murmur is bigger than for the normal sounds, which corresponds with higher frequencies in the murmur than in the normal sounds (Fig. 6d).

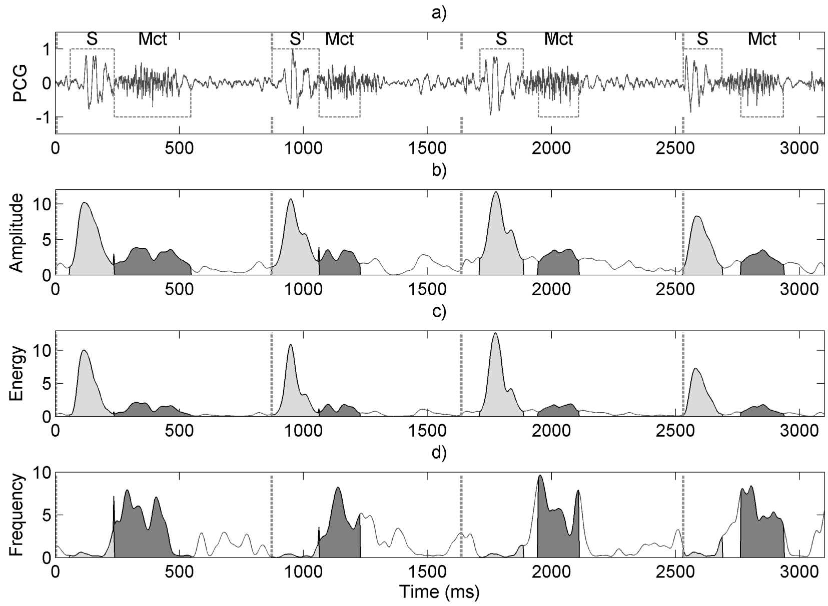

Figure 7. Representation of events detected in an abnormal record (mid-systolic murmur of mitral regurgitation): a) PCG, b) amplitude, c) energy, and d) frequency. In a), events detected as sounds have been represented with dash lines over the 0 baseline and with the label "S", and events detected as murmurs have been represented under the 0 baseline and with the label "M". The label "ct" stands for a constant murmur. In b), c), and d), sounds and murmurs have been represented, respectively, in light and dark gray, and the beginning of the cardiac cycles has been indicated as dash lines.

In Fig. 7 it is shown how this method is also capable of separating S1 and the mid-systolic murmur for the MR record, whereas this was not possible applying methods based on thresholds. The basic events of the MR record was thus successfully detected. It can be seen in Figs. 7b, 7c, and 7d how the intensity and energy of S1 is larger than that of the murmur, whereas the frequencies of the murmur are bigger than for S1.

Once the events have been delimited, they have been labeled as sounds ("S") or murmurs ("M") in Fig. 7a. In order to discriminate between AS and MR records, further analysis is needed for the murmurs of these two signals. The morphology of the murmurs, together with the analysis of some other properties extracted from the envelopes (mean and maximum values, and area enclosed under the event), allows the discrimination between these two pathologies, which occur in the same temporal localization (systole). The murmurs of these two pathologies were successfully discriminated, as shown in the labels in Figs. 6a and 7a, detecting the AS murmur as a crescendo-decrescendo one (Mc/d), and the MR murmur as a constant intensity one (Mct).

6. CONCLUSIONS

In this paper we have presented a hierarchical structure based on signal abstraction for the analysis and interpretation of phonocardiograms. This hierarchy provides a methodology to process the PCG that follows a similar procedure to that of the physician during cardiac auscultation. Besides, the modular structure allows a high degree of independency for the different processing blocks.

The signals extracted from the PCG (amplitude and energy in the temporal domain, and instantaneous frequency in the frequency domain) contains enough information to properly characterize the cardiac sounds and murmurs. Besides, the autocorrelation signal of the envelopes has proven to be very robust to obtain the average heart rate even in cases with additional sounds and murmurs, using no auxiliary signals.

One of the most important tasks of the hierarchy is the correct detection of the cardiac events. We have applied a method based on the extraction of a set of parameters from relative maximum peaks for the accurate detection and delimitation of the events that form the cardiac cycle, achieving very successful results. This algorithm greatly improves the threshold method often used to detect events, and works correctly in records with additional sounds and murmurs. The algorithm can even separate between different events that may appear joined in the temporal domain (usually a murmur and a main sound).

The discrimination between sounds and murmurs, in order to further analyzed the murmurs, also achieved good results, and the three-parts murmur division also recognized successfully the temporal morphology of the murmurs. Anyway, these two techniques will be refined as more records are available, in order to improve the generalization capabilities of these algorithms.

Concerning the hierarchical structure, results obtained so far have been very good in detecting and characterizing the events, so future results obtained with the complete hierarchy seem very promising to aid in the diagnosis of the valvular state of the heart.

ACKNOWLEDGMENTS

This work has been supported by Fundación Séneca of Región de Murcia and Ministerio de Ciencia y Tecnología of Spain, under grants PB/63/FS/02 and TIC2003-09400-C04-02, respectively.

REFERENCES

1. S. Mangione and L. Z. Nieman, "Cardiac auscultatory skills of internal medicine and family practice trainees. a comparison of diagnostic proficiency," *JAMA* **278**(9), pp. 717–722, 1997.
2. T. Ölmez and Z. Dokur, "Classification of heart sounds using an artificial neural network," *Pattern Recognition Letters* **24**, pp. 617–629, 2003.
3. T. S. Leung, P. R. White, W. B. Collis, E. Brown, and A. P. Salmon, "Classification of heart sounds using time-frequency method and artificial neural networks," in *Proceedings of the 22nd Anual International Conference, Engineering in Medicine and Biology Society*, **2**, pp. 988–991, (Chicago, IL., USA), 23–28 July 2000.
4. J. E. Hebden and J. N. Torry, "Identification of aortic stenosis and mitral regurgitation by heart sound analysis," in *Computers in Cardiology*, **24**, pp. 109–112, 1997.
5. W. W. Myint and B. Dillard, "An electronic stethoscope with diagnosis capability," in *Proceedings of the 33rd Southeastern Symposium on System Theory*, pp. 133–137, 2001.
6. D. Mason, *Listening to the heart*, F. A. Davis Co., 2000.
7. E. E. Milios and S. H. Nawab, "Signal abstractions in signal processing software," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **37**, pp. 913–928, June 1989.
8. R. Ruiz, S. Barro, J. Presedo, F. Palacios, and J. Vila, "Abstraction of information in electrocardiographic monitoring for the detection of myocardial ischaemia," in *V International Symposium on Biomedical Engineering*, pp. 291–294, 1994.
9. J. T. E. McDonnell, "Knowledge-based interpretation of foetal phonocardiographic signals," *IEE Proceedings F, Radar and Signal Processing* **137**, pp. 311–318, October 1990.
10. Z. Sharif, M. S. Zainal, A. Z. Sha'ameri, and S. H. S. Salleh, "Analysis and classification of heart sounds and murmurs based on the instantaneous energy and frequency estimations," in *Proceedings of the IEEE TENCON*, **2**, pp. 130–134, 2000.
11. H. Liang, S. Lukkarinen, and I. Hartimo, "Heart sound segmentation algorithm based on heart sound envelogram," in *Computers in Cardiology*, pp. 105–108, 1997.
12. L. Griffith, "Rapid measurement of digital instantaneous frequency," *IEEE Transactions ASSP* **23**, pp. 207–222, 1975.
13. P. J. Loughlin and B. Tacer, "Comments on the interpretation of instantaneous frequency," *IEEE Signal Processing Letters* **4**, pp. 123–125, May 1997.
14. H. Yoshida, H. Shino, and K. Yana, "Instantaneous frequency analysis of systolic murmur for phonocardiogram," in *Proceedings of the 19th International Conference, Engineering in Medicine and Biology Society*, **4**, pp. 1645–1647, (Chicago, IL., USA), 30 October – 2 November 1997.
15. L. Sun, M. Shen, and F. H. Y. Chan, "A method for estimating the instantaneous frequency of non-stationary heart sound signals," in *Proceedings of the 2003 International Conference on Neural Networks and Signal Processing*, **1**, pp. 798–801, 2003.