Phonocardiography Signal Segmentation for Telemedicine Environments

Santiago Murillo Rendón[1], Cristian Castro Hoyos[1],
Carlos M. Travieso-Gonzales[2], and Germán Castellanos-Domínguez[1]

[1] Universidad Nacional de Colombia, Manizales, Colombia
[2] Universidad de Las Palmas de Gran Canaria, España

Abstract. In this paper, phonocardiography (PCG) segmentation methodology based on envelope detection is developed by using a time-scale representation and a synthetic electrocardiogram signal (EKG). The heart cycle duration is calculated by autocorrelation of S1-S2 sounds that are synchronized with the synthetic EKG. Two algorithms for noisy signal removal are implemented to ensure the detection of signals with low signal to noise ratio. Approach is tested in a PCG database holding 232 recordings. Results show an achieved accuracy up of 90%, thus, overperforming three state-of-the-art PCG segmentation techniques used to compare the proposed approach. Additionally, the synthetic EKG is built by estimation of heart rate length, thus it does not use a patient recording EKG, reducing the computational cost and the amount of required devices.

Keywords: Heart Sound Segmentation, Phonocardiogram, Telemedicine, Autocorrelation.

1 Introduction

Auscultation is the basic diagnostic tool for every medical physician, because of its simplicity, associated low-cost, and its non-invasive nature. The heart sound recorded by auscultation or phonocardiography (PCG) is the primary mechanical analysis tool for the heart function. Other recent techniques, as echocardiography, suggest more sensible methods to estimate the cardiac mechanics. Nevertheless, their associated high cost and the elaborated medical knowledge needed for their evaluation make them unsuitable for uncontrolled environment applications, such as *telemedicine*.

In cardiac auscultation, the physician finds natural heart sounds originated by the opening and closing valves, named S1 and S2 sounds, respectively. The systolic interval starts at the beginning of S1 and extends until the beginning of S2, while the diastolic interval starts at the beginning of S2 and extends until the next S1. Occurrence of additional sounds in the periods, called little and big silences (intervals between S1 and S2), are clear evidence of myocardium mechanical failures. Each period composed by these systolic and diastolic periods is known as cardiac cycle and it is the basic unit for the sound heart analysis. Identification of those additional sounds requires a higher medical ability and their recognition can vary because of the inherent limitations in the listener and the limited sensibility of the human ear in the low-frequency range [1].

I. Rojas, G. Joya, and J. Cabestany (Eds.): IWANN 2013, Part II, LNCS 7903, pp. 124–134, 2013.

Hence, there is a need for development of automatic aided diagnosis tools, providing a more objective mechanism for identification of heart sounds.

To this end, several segmentation techniques are found in the literature that, however, commonly need for additional information extracted from signals as the electrocardiogram (EKG) or the carotid pulse [2]. That is, techniques must have more sources and thus increase the complexity and cost of the acquisition system, making its usage prohibitive in telemedicine applications. Particularly, a comparison of envelope-based segmentation methods is given in [3], where most of them make use of feature extraction techniques such as Shannon's energy envelope, Hilbert transform and other cardiac features. Other approaches use the energy estimated from the Short Time Fourier Transform spectrogram or the scalogram computed with the continuous wavelet transform [4, 5]. Afterwards, peak detection and time/amplitud thresholding are used in [6] for locating the S1 and S2 sounds. Nonetheless, energy is not always constant in the components of the PCG, since the auscultation focuses, pathologies, breathing and, in general, some other artifacts produce changes in the time-energy distribution, which in turn induce changes in the amplitude peaks of the signal. Moreover, the time threshold or window for locating sounds assumes a priori their length duration, leading to bartered segmentation, i. e., diastolic-systolic segments instead of the normal systolic-diastolic heart cycle. In [7], noisy signals are evaluated for a noise-robust detection method by repeating the signal recording when heart sound characteristics are masked by noise. However, the heart cycle estimation method implies a high computational cost that can not be accomplished in near real-time application.

In this paper, a PCG segmentation methodology based on envelope detection is discussed. The approach takes into consideration the changes in the Heart Rate as a mitigating factor, aiming to adapt the cardiac cycle length to each segment length. In the approach, the PCG envelope is computed from the signal time-scale representation. Then, the cardiac cycle length is estimated from the autocorrelation of 3-second windowed envelope signal. Lastly, the beginning of the QRS segment and the T-wave termination, on EKG, are synchronized to the beginning of S1 and S2 sounds, on PCG, respectively. Since the approach does not required any time or amplitude thresholding, issues related to the parameter tuning are avoided. Hence, the location and synchronization of the synthetic EKG signal ensures a segment-by-segment detection without bartering of the detected locations. Also two noise detection stages are implemented: one based on predetection of noisy signals, and other by evaluating the time-frequency periodicity similar to [7]. All the used methods are especially chosen to work efficiently in telemedicine applications. The performance is measured in terms of segmentation accuracy, that is, expert observation on whether the heart sound segment is properly located against the EKG. The paper is then organized as follows, the next section presents a theoretical background information. Afterwards, the proposed scheme for segmentation of PCG signals and state all the parametrization in the methodology is presented. In the fourth section, results of accuracy segmentation are shown. Finally some conclusions and future work are given as a summary of the research.

2 Method

2.1 Noise Predetection

The basic idea behind this stage is to estimate the energy variation along the PSG signal, $x(t), t \in T_a$. To this end, the Shannon energy is calculated, as follows [8]:

$$\eta(t) = -x(t)^2 \log x(t)^2 \tag{1}$$

So, if the signal is corrupted by noise within a given estimation segment, T_a, its energy should increase, being noticeably higher that the average energy value, and thus allowing to infer the presence of noise. Signals with strong noise indicators should be either rejected or recorded again. With this in mind, the average Shanon energy is calculated as:

$$\hat{\eta} = \boldsymbol{\mathcal{E}} \{\eta(t) : \forall t \in T_a\} \tag{2}$$

where $\boldsymbol{\mathcal{E}} \{\cdot\}$ stands for expectation operator.

2.2 Time-Frequency Enhancement of PCG Signal

Continuous Wavelet Transform: Wavelet analysis that comes from the basic definition of Fourier theory is based on the representation of a time-varying signal via a space domain transformation, that is, to the time-frequency domain, as follows:

$$X(\tau, w) = \int_{T_a} [x(t)\, w^*(t - \tau)]\, e^{-j\omega t} dt \tag{3}$$

where w is a window function segmenting the signal in portions that are assumed to be stationary and τ is a time offset. The window function gives the main feature of the STFT, i.e., the width of the window (called support) that provides the capability of managing the resolution of the representation. Nonetheless, the use of basis functions as well as the window components in transformation enables to analyze the signal at different frequency bands with different resolutions, in other words, every spectral component is not analyzed equally as for the STFT; this approach is then called Wavelet Transform (WT) that maps a one-dimensional signal into a two-dimensional representation, time and scale, allowing to give local (at a given resolution) information about the frequencies occurrences. The CWT is then defined as:

$$W(a, b) = \int_{T_a} x(t)\, \Psi^*_{a,b}(t) dt \tag{4}$$

where $*$ denotes complex conjugation; variables a and b are the scale and translation, respectively, and compose the new dimension of the WT. The set of Wavelets is then generated by translations and dilations of a single Wavelet function called the mother Wavelet, i.e., $\Psi_{a,b}(t) = a^{-1/2}\Psi((t - b)/a)$.

Time-Scale Representation of PCG Signal: After obtaining the time-scale signal representation, its energy must be computed. To get a smoother energy estimation, instead of Shannon operator in Eq. (2), the scalogram is used that is defined as follows:

$$W_{SCAL}(t,a) = \frac{1}{ca^2}\left|W(t,a)\right|^2 \tag{5}$$

where c is a energy normalization parameter such as the averaged energy of the signal and the averaged energy of the scalogram becomes equal.

Scalogram then provides with the temporal localization of the energy foci, which are expected to belong to the fundamental components of the signal. Thus, an envelope of the PCG signal can be obtained with more accuracy, if the PCG is portioned into short duration frames, when adding the energy concentration over time instants. Therefore, the following marginal estimation is accomplished:

$$\varepsilon(t) = \sum_{\forall a}\left|W_{SCAL}(t,a)\right|^2 \tag{6}$$

Since all short time duration segments of a long-term PCG signal can be assumed to be quasi-stationary, the autocorrelation of the envelope of such segments should provide with information about the periodicity of the PCG signal. Particularly, for a given PSG signal segment that lasts T_a holding several S1 and S2 events, the autocorrelation that should show the averaged duration of the heart sound cycle is estimated as follows:

$$\boldsymbol{r}_{\varepsilon,\varepsilon}(t) = \begin{cases} \sum\limits_{m=1}^{T_a-m-1} \varepsilon(n+m)\,\varepsilon(m), & m \geq 0 \\ \varepsilon(-m), & m < 0 \end{cases} \tag{7}$$

Taking into account that for the autocorrelation function both properties hold: periodicity, i.e, $\boldsymbol{r}(\tau) = \boldsymbol{r}(\tau + T)$, $\forall t \in T$, and maximum value, $\max_{\forall \tau \in T}\{\boldsymbol{r}(\tau)\} = \boldsymbol{r}(0)$, then one can infer that $\boldsymbol{r}(0) = \boldsymbol{r}(T)$, being T the period of the PCG signal that includes both S1 and S2 events. Since analysis frame in Eq. (7) lasts more that T, that is, $T_a > kT$, with $k \approx 3\ldots,4$, then the estimated correlation function will hold $2k$ equally located peaks. The peak sequence relates the following matching events:(S1 coincides with S1)$\rightarrow$ (S1-S2)$\rightarrow$ (S1-S1)$\rightarrow$ (S1-S2)$\rightarrow$ ... Therefore, the distance between odd peaks is the estimated value of T.

It must be quoted that the obtained in Eq. (7) estimation of the segment duration T directly from the PCG signal avoids the usage of additional information, like the EKG or the carotid pulse recording.

2.3 Refined Time-Frequency Periodicity Analysis

Estimation of duration T, so far, assumes a high value of signal-to-noise ratio. However, regularity of peaks is strong affected by the presence of noise. To cope with this issue, a refined periodicity estimation is provided based on the frequency band analysis of time-frequency plane.

In general, the spectrogram, given in (4), is linearly divided in F sub-bands, from each one of which the corresponding energy envelopes is accomplished by Eq. (6). Assuming a moderate signal-to-noise ratio, the same peak sequence relating the before considered matching events takes place, but for each one of the split bands. This paper considers $F = 15$. Afterwards, a matrix is constructed with sets of the autocorrelation envelops gathering 5 neighboring sub-bands at the same time, where for each matrix the following periodicity value is introduced: $\rho_i = (\lambda_{i2}/\lambda_{i1})^2$, (being λ_{i1} the first and λ_{i2} - the second singular value). In case of strong periodicity, either relationship should hold: $\rho_1 > \rho_2 > \rho_3$ or $\rho_1 < \rho_2 < \rho_3$, otherwise, one can infer the persistent noise in signal [7].

2.4 Synchronism for Segmentation

The intrinsic relationship among heart sounds and EKG provides us a tool for detecting or better determining whether an event in the envelope of the PCG signal is systolic or not, however, the simultaneous EKG is not always available. Having define the length of the heart cycle, we generate a synthetic EKG signal of the same length using the nonlinear model proposed in [9]. The EKG signal starts on the onset of the QRS (that is the associated beginning of a systolic sound) and has a duration of T. Synchronization refers then to identification of the first systolic sound over each analyzed PCG part. This is done when comparing the distance between the R-peaks and the T-wave end in the synthetic EKG signal with the first and second local maximums of the envelope, that is, the distance between the R-peak and the first local maxima of ε is minimum when compared with T if the signal starts some close at systolic sound, if not, we compare the second local maxima of ε with the end of the T-wave with a tolerance parameter and determine if it is a diastolic sound, in that case the segmentation is done from the next local maxima which is supposed to be a systolic sound. The T value is recalculated for each new heart cycle (the synthetic EKG adopts the new T value), that allows to adapt the segmentation to the heart rate variability, and therefore, a better segmentation performance can be achieved.

3 Experimental Setup

The methodology presented in this work is displayed in Fig. 1.

3.1 Database

In this work, the heart sound (Phonocardiogram, PCG) database for murmur detection of Control and Signal Processing Group at the Universidad Nacional de Colombia - Manizales is used. This database is composed of 29 patients, each one with 8 recordings, namely 4 recordings in the traditional auscultation focuses(aortic, pulmonary, mitral and tricuspid) in diaphragm and bell modes. All recordings have a duration of 20 seconds and present between 10 and 35 heart cycles according to heart rate of each patient. An electronic stethoscope and the meditron software were used. Signals were recorded in .wav at sampling frequency $44.1kHz$ and resolution 16 bit. A labeling

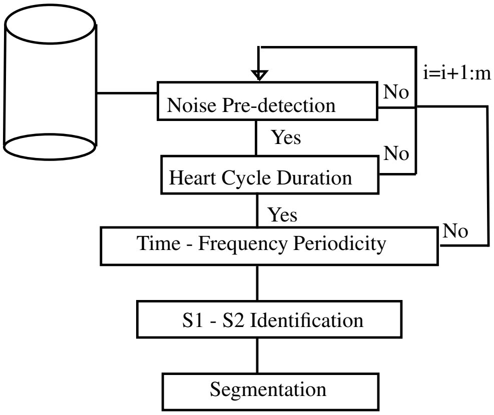

Fig. 1. Algorithm of the proposed methodology

process was made by a medical team, normal or pathological are the possible classes. The pathological state corresponds to the existence of abnormal sounds, especially the appearance of murmurs. Murmur intensity depends on auscultation focus, that is the principal reason to make the recording over all the 4 foci.

3.2 Measurements

For performance estimation, we use a traditional pan and Tompkins EKG segmentation algorithm, this is to know the start and end of each heart cycle, given the EKG and PCG correlation. To validate the method segmentation, the PCG segmentation should coincide with pan-tompkins results in at least 10%, in other case the part of signal is considered incorrectly segmented. We use this information to measure the method true and false detection, as a made in [10]: $\hat{P}_F = N_F/(N_D + N_F)$ and $\hat{P}_D = N_D/(N_D + N_M)$, where N_D is the number true positives, N_F false negatives, and N_m false negatives. Thus, $\hat{P}_F$ corresponds with the probability of false detection and $\hat{P}_D$ the probability of detection. Fig. 2 explains the procedure.

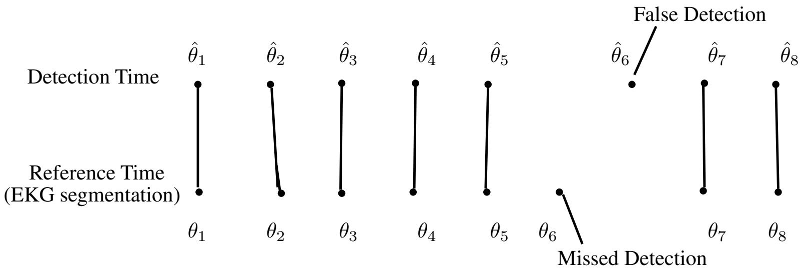

Fig. 2. Explanation of measurements

3.3 Experiments

Signals are evaluated to identify a strong presence of noise as in Sec.2.1. It is important because the presence of noise in signals affects the time-frequency behavior. Then, T is estimated for each signal. The T in some signals might not be calculated, this occurs in signals with noise even after pre-detection stage. After that, we probe the signals with other noise detection method, here the periodicity in both frequency and time as in Sec. 2.3 is measured. Signals without periodicity are identified as noisy instances. The last stage consists in segmenting the clean signals, the measurements described in Sec. 3.2 are calculated with the idea of evaluating the method performance.

4 Results and Discussion

4.1 Predetection Stage

All 232 recorded signals were evaluated, in total 18 signals was detected as noisy signals for the bell mode and for the Diaphragm mode a total of 20 signals result noisy, the focus procedence source of noisy signals for each mode are showed in table 1. Note that several signals from pulmonary focus has been detected as noisy in relation with the other focuses, this is because of the presence of respiratory sounds. In this focus the proximity with the lung structures causes masking in heart sounds. In other words, this experiment shows the high presence of internal noise sources. In noise terms, this focus is followed by the mitral focus, here exist perturbations associated to lung structure too, due to the focus localization in the intercostal space. The other focuses are affected by different noise sources i.e., digestive sounds or human voice. It is important to emphasize that owing to the data acquisition conditions all the recordings are affected by external noise, this can justify the detection of some signals in this stage.

Note that noisy segments show a high energy with respect of the average energy, Fig. 3 exhibit a signal with noisy segments, here is evident the strong noise presence and its influence in energy values.

4.2 Heart Cycle Duration

For the rest of signals, after noise pre-detection, the heart cycle duration is estimated as in Sec. 2. For comparison purposes the EKG signal is used to identify the real heart frequency. In these terms, the estimated heart cycle duration is considered good if it matches at least in $\pm 10\%$ the real value. As a result of this comparison we find 10

Table 1. Noise Pre-Detection

Focuses of Auscultation							
Bell Mode				Diaphragm Mode			
A	M	P	T	A	M	P	T
1	6	8	3	5	5	5	5

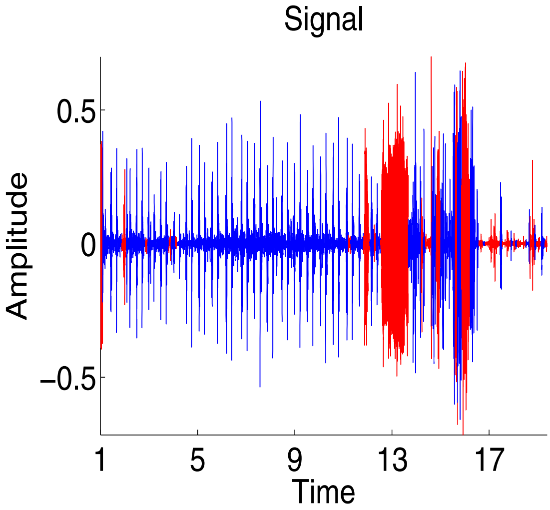

Fig. 3. Noisy Signal

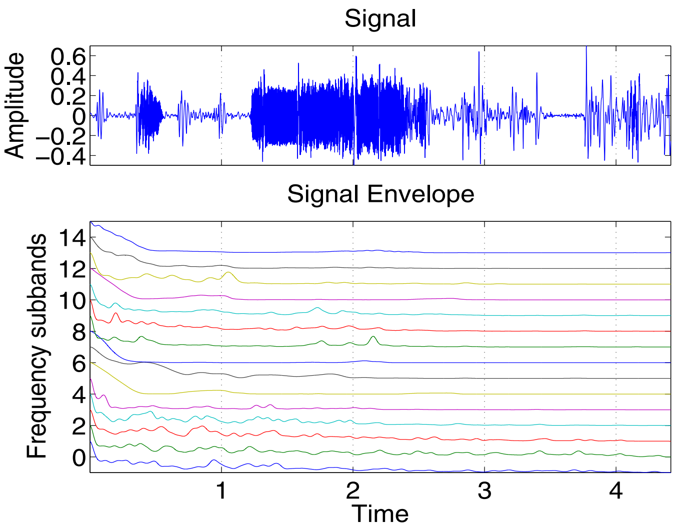

Fig. 4. Noisy signal

signals where T can not be calculated, this is only a 4.3% of the signals in the original dataset. The estimation problems occur because of the noise present in these signals that can not be detected in the predetection stage; the periodicity analysis is an additional method to identify this kind of noisy signals.

4.3 Refined Procedure of Computing Time-Frequency Periodicity

After this analysis, 15 signals were identified with periodicity problems. Fig.4 is an example of noisy signal and its frequency problems, the low homogeneity between the autocorrelations obtained of the 15 frequency subbands shows it. Nonetheless, in Fig. 5, it is showed a signal with periodic morphology, note the uniform dynamic along the signal through all the subbands, namely the peaks appear at the same time.

4.4 Achieved Segmentation

Lastly, after segmenting a total of 1741 heart cycles, according to the EKG reference, the present method identified 1516 of these heart cycles. It represents a sensibility of 90.88% and false detection rate of 9.11%, these values are calculated as explained in Sec.2. These results show the high efficiency of the exposed method and present that a good detection of noisy signals is a necessary stage for PCG signal segmentation. The present work shows the goodness of noise detection as a stage in automatic phonocardiography segmentation, and secondly, the use of segmentation methodologies based on envelope detection together with the synthetic EKG can improve the good segmentation rates.

Compared with other works, the present method has superior performance, for instance in [3], segmentation results for some kinds of envelope extraction methods are shown i.e., Shannon envelope shows segmentation efficiency for abnormal cases between 75.5% and 89.4%, and for normal cases from 65.9% to 78.2%), while Hilbert envelope shows low segmentation rates. Also this work say that CSCW shows segmentation efficiency between 96.2% and 100% for normal cases and between 72.7% and 88.2% for abnormal cases. In [11], discussed approach gives an accuracy from 88.29%

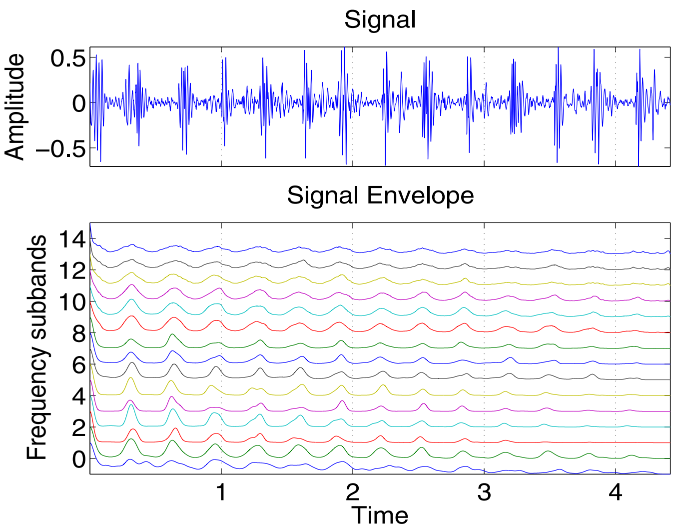

Fig. 5. Noisy signal

to 66.77% for Morlet envelope and Hilbert envelope, respectively. In [8] it is showed an accuracy of about 95.51%, but taking into account that they use a-priory information as the average heart rate and it is supposed a continuous hear rate. The presented method is robust to heart rate variability and even talking about patients with some arrhythmia it is possible to carry out a proper segmentation.

5 Conclusions and Future Work

An efficient segmentation methodology for PCG signals is presented that allows segmentation accuracy of up to 90% in most of the database registers. Method performance show that cycle heart length estimation by autocorrelation and synchronization with the synthetic EKG provides a good segmentation strategy with the joint analysis of the envelope of the original signal, it is important to note that the proposed method is chosen to work in telemedicine environments. This can enhance the traditional methods of PCG segmentation ensuring a correct length calculation and S1-S2 detection.

As regards the telemedicine application, it is important to note that the noise conditions for heart sound acquisition can not be adequate, for that reason it is necessary to implement strategies that contribute to identify the quality of recorded signals, for instance, the periodicity and energy methodologies used in the present work are a good response to this problem. The use of an adaptive synthetic EKG from a T estimation in each heart cycle segment, allows the segmentation even in patients with a large heart rate variability.

The low accuracy rates presented along the state of the art are understood, because the noisy signals diminish the methodologies performance, that as a result of the envelope distortions. Nevertheless, it is important to work in segmentation methods based on envelope, because its simplicity enables it to work in telemedicine applications. Despite the noise analysis, it is important to use a refined acquisition protocol, that allows to identify if it exists acceptable conditions for recording signals, that is, an environment free of noise and to get the lowest quantity of excluded signals.

Finally as a future work, we propose to find better methods to estimate the heart cycle duration, because a better evaluation of time-frequency signal periodicity is highly dependant on this, furthermore, it can improve the segmentation performance by decreasing the false detection rate. Also it is important to improve the synchronization methodology, that because there exist signals with attenuated sounds, that produce losses of peaks, which also decreases the accuracy rates.

Acknowledgments. This work is supported by the "*Aprendizaje de máquina a partir de múltiples expertos en clasificación multiclase de señales de voz" project associated with "Jóvenes Investigadores*" program by COLCIENCIAS and Universidad Nacional de Colombia - Manizales, the project "*Servicio de Monitoreo Remoto de Actividad Cardíaca para el Tamizaje Clínico en la red de Telemedicina del Departamento de Caldas*" and the "*Grupo de Control y Procesamiento Digital de Señales Código 20501007205*"

References

1. Ari, S., Kumar, P., Saha, G.: A robust heart sound segmentation algorithm for commonly occurring heart valve diseases. Journal of Medical Engineering and Technology 32(6), 456–465 (2008)
2. Wang, P., Kim, Y., Ling, L.H., Soh, C.B.: First heart sound detection for phonocardiogram segmentation. In: 27th Annual International Conference of the Engineering in Medicine and Biology Society, IEEE-EMBS 2005, pp. 5519–5522 (January 2005)
3. Choi, S., Jiang, Z.: Comparison of envelope extraction algorithms for cardiac sound signal segmentation. Expert Systems with Applications 34(2), 1056–1069 (2008)
4. Malarvili, M.B., Kamarulafizam, I., Hussain, S., Helmi, D.: Heart sound segmentation algorithm based on instantaneous energy of electrocardiogram. In: Computers in Cardiology, pp. 327–330 (September 2003)
5. Yuenyong, S., Nishihara, A., Kongprawechnon, W., Tungpimolrut, K.: A framework for automatic heart sound analysis without segmentation. BioMedical Engineering OnLine 10, 13+ (2011)
6. Tseng, Y.-L., Ko, P.-Y., Jaw, F.-S.: JawDetection of the third and fourth heart sounds using Hilbert-Huang transform. BioMedical Engineering OnLine 8 (February 2012)
7. Kumar, D., Carvalho, P., Antunes, M., Paiva, R.P., Henriques, J.: Noise detection during heart sound recording using periodicity signatures. Physiological Measurement 32(5), 599 (2011)
8. Ari, S., Saha, G.: On a robust algorithm for heart sound segmentation. Journal of Mechanics in Medicine and Biology 7(2), 129–150 (2007)
9. Clifford, G.D., McSharry, P.E.: Generating 24-hour ECG, BP and respiratory signals with realistic linear and nonlinear clinical characteristics using a nonlinear model. In: Computers in Cardiology, pp. 709–712. IEEE (2004)
10. Sörnmo, L., Laguna, P.: Bioelectrical Signal Processing in Cardiac and Neurological Applications. Elsevier Academic Press (June 2005)
11. Zhong, L., Guo, X., Ji, A., Ding, X.: A robust envelope extraction algorithm for cardiac sound signal segmentation. In: 2011 5th International Conference on Bioinformatics and Biomedical Engineering, iCBBE, pp. 1–5 (2011)