computer methods and programs in biomedicine xxx (2009) xxx–xxx

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

A novel method for pediatric heart sound segmentation without using the ECG☆

Amir A. Sepehri[a,*], *Arash Gharehbaghi*[a], *Thierry Dutoit*[b], *Armen Kocharian*[c], *A. Kiani*[c]

[a] *ICT Research Center, Amir Kabir University, Tehran, Iran*
[b] *TCTS Laboratory, Faculte Polytecnique de Mons, Belgium*
[c] *Children Heart Center, Medical University of Tehran, Iran*

ARTICLE INFO

Article history:
Received 22 October 2008
Received in revised form
23 October 2009
Accepted 27 October 2009

Keywords:
Congenital heart diseases
Segmentation
Heart sounds
Arash_Band
End-pointing
Cardiac cycles
Murmurs
Pediatric heart sound
Children's heart sound

ABSTRACT

In this paper, we propose a novel method for pediatric heart sounds segmentation by paying special attention to the physiological effects of respiration on pediatric heart sounds. The segmentation is accomplished in three steps. First, the envelope of a heart sounds signal is obtained with emphasis on the first heart sound (S_1) and the second heart sound (S_2) by using short time spectral energy and autoregressive (AR) parameters of the signal. Then, the basic heart sounds are extracted taking into account the repetitive and spectral characteristics of S_1 and S_2 sounds by using a Multi-Layer Perceptron (MLP) neural network classifier. In the final step, by considering the diastolic and systolic intervals variations due to the effect of a child's respiration, a complete and precise heart sounds end-pointing and segmentation is achieved.

1. Introduction

In the past 10 years, some researchers have tried to develop a noninvasive test system for congenital heart diseases detection based on heart sounds analysis techniques [1–4]. There has been a great improvement in developing such a system using manual segmentation along with the electrocardiogram (ECG) signal [5,6]. However, one important obstacle in developing an all-automatic congenital heart diseases detection system has been automatic segmentation of systolic and diastolic periods as well as extraction of S_1 and S_2 sounds. If we could automatically segment the pediatric heart sounds signal and extract S_1 and S_2, then it would be possible to make an all-automatic system to detect and determine congenital heart diseases by noninvasive acoustical method as a stand alone system. Such a system could be used by hospital technicians rather than by a trained physician for disease's detection. In general, we can use a synchronous 12-lead ECG signal for complete segmentation but, in cases of infants or newborn children, such an ECG signal acquisition has its own inconveniencies.

☆ We would like to thank Professor Joel Honcq for his assistance and his ideas in this work.
* *Corresponding author at*: TCTS (EE), Rue de l'Epargne 37, 7000 Mons, Belgium. Tel.: +32 0 65 84 48 86; fax: +32 0 65 84 19 46.
E-mail address: a.sepehri@capis.be (A.A. Sepehri).
0169-2607/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2009.10.006

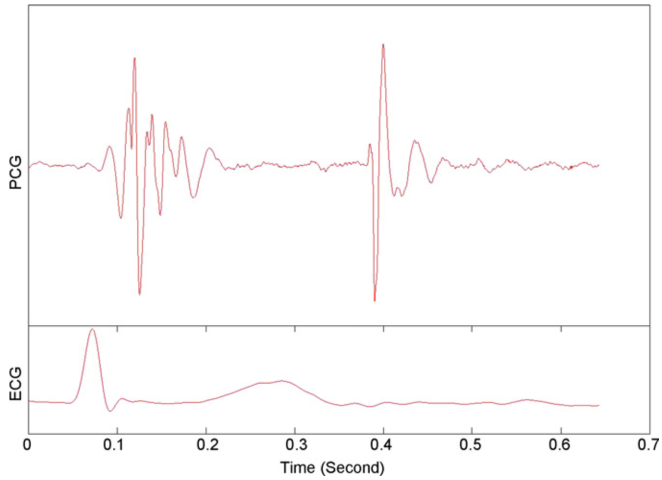

Fig. 1 – A complete heart sound cycle with ECG.

A number of recent works have been done on adult heart sound signal segmentation [7–9]. However, in some cases the end-pointing is done by using the ECG signal.

Based on spectral and timing properties of pediatric heart sound signals, we present an algorithm which extracts cardiac cycles along with S_1 and S_2 sounds from a heart sounds signal. In children, respiration causes a variation of cardiac cycle duration. But, the effect of respiration on diastolic duration is by far more tangible than systolic duration. Based on this effect, we present a novel method for a complete segmentation of children's heart sounds signals. We extract the basic heart sounds by taking into account the repetitive and spectral characteristics of S_1 and S_2 sounds and using a Multi-Layer Perceptron (MLP) neural network classifier.

To evaluate the performance of our work, we trained the method using 60 normal and pathological samples of our data-bank. We then tested the method on 60 other samples separately. We have correctly identified S_1 and S_2 sounds in more than 93% of the cases. The results show that the method opens a way to an all-automatic congenital heart diseases detection through heart sounds analysis in children.

2. Heart sound segmentation difficulties

An electrocardiogram is recording of electrical activities of the heart cells gathered from the surface of the body [10]. Phonocardiogram (PCG) signals are recordings of acoustical waves produced by the mechanical action of heart which are semi-periodic [11]. A normal heart by itself initiates two sounds; the first heart sound (S_1) and the second heart sound (S_2). Mechanical activities of a heart are always originated from its electrical activities [10]. S_1 signal is always after the QRS-complex in ECG signal and S_2 after T-Wave. Fig. 1 shows a synchronous recording of a PCG and its corresponding ECG signal.

Cardiac cycles can be determined by using ECG. However, for congenital heart disease detection, we need to extract S_1, S_2, systolic and diastolic duration in each cardiac cycle. Such segmentation is done manually using the ECG signal for end-pointing of the cardiac cycles. In children with left or right hypertrophic ventricle, axis deviation of heart causes an abnormality on ECG signal which complicates the manual segmentation [10]. Besides, the ECG signal acquisition has its own inconveniencies in children.

Heart sounds analysis is performed on special segments of a cardiac cycle. Depending on circumstances, a PCG signal may be mixed with different environmental noises, artifacts and pathological sounds. Artifacts are random sounds could be initiated from different sources such as movement of the stethoscope. When there is a murmur, the beginning of a heart sound is sometimes completely covered by murmurs. Therefore, automatic segmentation of a child heart cycle without using the ECG is a complicated task.

3. Systolic and diastolic intervals variations

A systolic interval starts at the beginning of S_1 and lasts until the beginning of S_2. A diastolic interval starts at the beginning of S_2 and lasts until the end of the cycle [12]. S_1 sound is a result of closure of mitral and tricuspid valves. A short time after the onset of S_1, aortic and pulmonary valves start opening and the preloaded blood ejects into the arteries [12]. The blood ejection lasts until the closure of the aortic and pulmonary valves [12]. This period of time is called the ejection time (ET) and the part of the systolic duration before the ET is named Pre-Ejection period (PEP). Respiration has significant influences on right ventricular size, diastolic filling time and also heart rate in children [13,14]. The influences of respiration on the left ventricular dimension and filling time are the same as the right ventricular in children [15]. The effect of respiration on diastolic filling time and heart rate is age-dependent [14]. On the other hand, the influence of respiration on the ET is in the opposite direction with respect to the PEP [16]. Therefore, the whole systolic time which is the sum of PEP and ET remains almost constant in contrast with the diastolic period which has tangible variations due to respiration [16]. To verify this assumption experimentally, we compute the variance of systolic period over the variance of diastolic period for each of the 120 subjects of our data-bank, denoted by SOD as follow:

$$\mathrm{SOD} = \frac{\text{variance of systolic periods}}{\text{variance of diastolic periods}}$$

Fig. 2 shows the SOD for 120 normal and abnormal cases of the data-bank.

As it is seen in the graph, the SOD is less than one for all the subjects. It is worth noting that the diastolic function is age-dependent [14,17] and the aforementioned results could be valid only for children. The effect of respiration on systolic and diastolic intervals for a child is depicted in Fig. 3.

4. The segmentation method

The method is based on spectral properties of pediatric heart sounds and the influences of respiration on systolic and diastolic timing of cardiac cycles. The segmentation is accomplished in three steps. In the first step (4.1), we obtain an envelope of the heart sounds signal by computing the spectral energy. The spectral energy of the basic heart sounds is limited to a specific narrow frequency band. The envelope shows

Please cite this article in press as: A.A. Sepehri, et al., A novel method for pediatric heart sound segmentation without using the ECG, Comput. Methods Programs Biomed. (2009), doi:10.1016/j.cmpb.2009.10.006

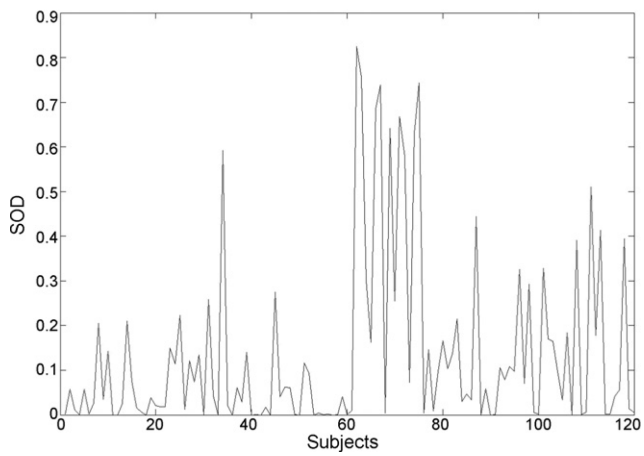

Fig. 2 – Variance of systolic period over the variance of diastolic period (SOD) for 120 normal and abnormal children.

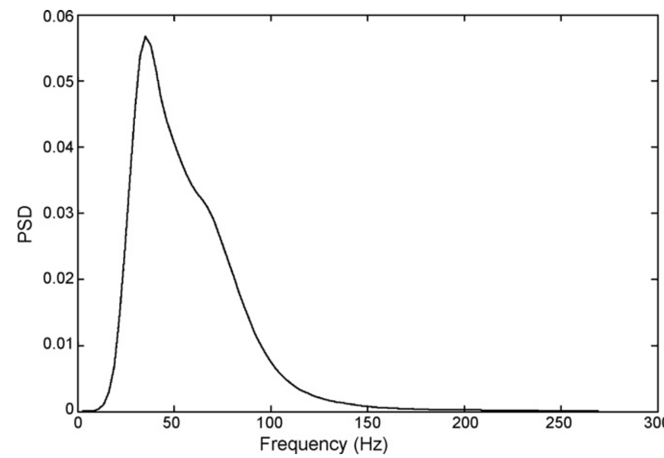

Fig. 4 – The average PSD of the basic heart sounds for 60 subjects.

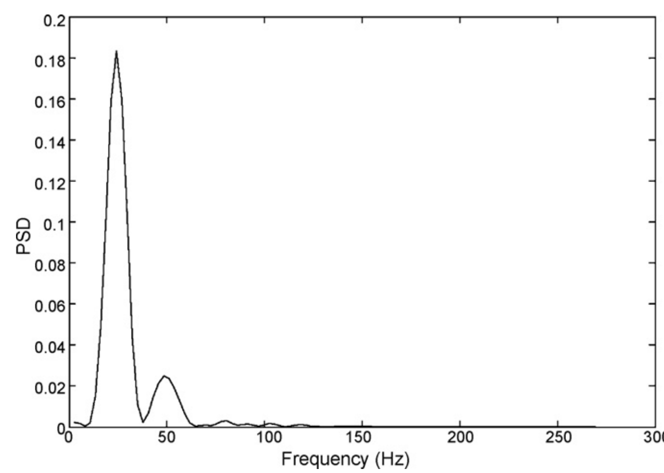

Fig. 5 – The average PSD of S_3 and S_4.

large peaks where the first or the second heart sounds exist. However, artifacts and other sounds can also cause large peaks on the envelope. If we find large peaks of the envelope with their related forward and backward minimums we can detect all possible sounds in the signal. In the second step (4.2), by using a neural network classifier, it is possible to discriminate between the basic heart sounds and the other sounds. In the final step (4.3), endpoints of the cardiac cycles are indicated by taking into account the respiration effects on systolic and diastolic durations.

4.1. *Sounds extraction*

We extract sounds by finding an envelope for the signal which emphasizes the basic heart sounds. The envelope is obtained by taking into account the spectral energy of the basic heart sounds. In order to evaluate the frequency contents of the basic heart sounds, we select 600 cardiac cycles from our data-bank and obtain related power spectral density function (PSD)

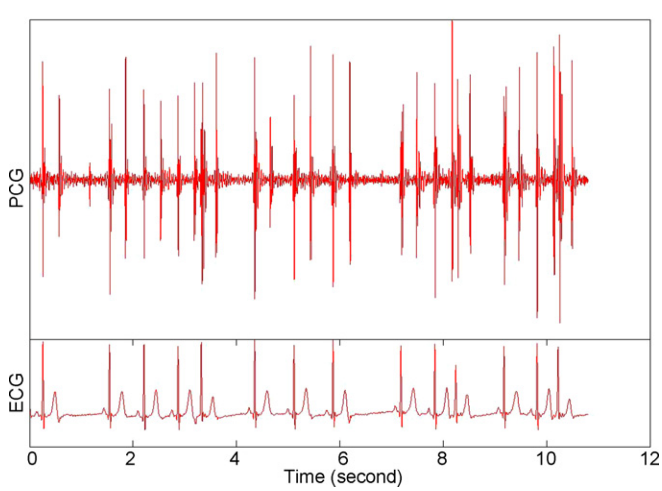

Fig. 3 – Cardiac cycle variations with respiration. The systole periods has much less variation than the diastole periods.

using the FFT technique. Fig. 4 shows the average of PSD for S_1 and S_2 over the 60 instances.

It is seen that a large amount (more than 90%) of the spectral energy are inside frequency band of 30–120 Hz. Previous researchers reached the same conclusions [18,19]. The frequency bands of murmurs extend up to 600 Hz. Third heart sound (S_3) and forth heart sound (S_4) are diastolic low frequency sounds. In order to investigate the PSD of S_3 and S_4, we compute PSD average for 120 S_3 sounds and 20 S_4 sounds existing in our data-bank. The results are depicted in Fig. 5.

As seen in Fig. 5, S_3 and S_4 sounds have frequency bands of about 15–50 Hz. We have experienced that AR modeling gives satisfactory results for power spectral density estimation for pediatric heart sound segmentation. If we model the signal as the output of an IIR filter which is stimulated by a white noise process with zero mean and variance δ, the PSD is calculated by [6]:

$$PSD(f) = \frac{T\delta^2}{\left|1 + \sum_{k=1}^{p} a_k e^{-j2k\pi f}\right|^2} \qquad (1)$$

Please cite this article in press as: A.A. Sepehri, et al., A novel method for pediatric heart sound segmentation without using the ECG, Comput. Methods Programs Biomed. (2009), doi:10.1016/j.cmpb.2009.10.006

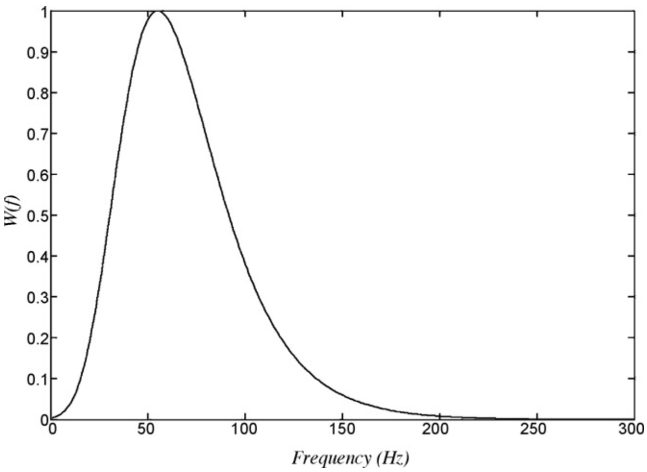

Fig. 6 – The frequency window $W(f)$ for $\mu = 55$ and $\delta = 25$.

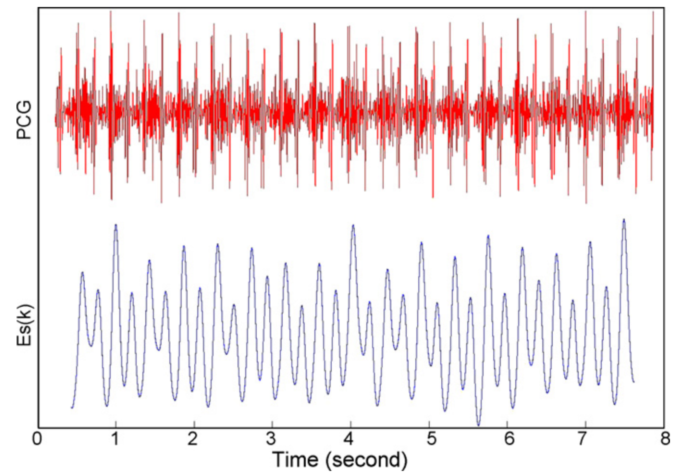

Fig. 7 – The PCG (top) and its envelope $Es(k)$ (bottom) for a children with VSD. On the envelope $Es(k)$, the basic heart sounds are completely recognizable among a harsh systolic murmur.

where a_k is the filter coefficients, p is the filter order, f is frequency and T is sampling interval. Now, we obtain a special envelope of the input signal using above formula and paying special attention to the spectral properties of the basic heart sounds. In order to obtain the envelope, we segment the signal in samples of 100 ms with overlap factor of 75%. Then, we calculate the short time spectral energy in each interval (the sample rate of 2 kHz and the resolution of 16 bit/s). The energy of an interval k, over the frequencies f_1 and f_2 is calculated by:

$$En(k) = \sum_{f=f_1}^{f_2} PSD(f)\ W(f) \tag{2}$$

where f_1 and f_2 represent frequencies of 3 dB below the maximum frequency amplitude. $W(f)$ is a modified Gaussian window defined as:

$$W(f) = \exp\left(-\exp\left(-\frac{f-\mu}{\delta}\right) - \frac{f-\mu}{\delta} + 1\right) \tag{3}$$

The frequency window is depicted in Fig. 6 for $\mu = 55$ and $\delta = 25$.

Choosing $\mu = 55$ and $\delta = 25$ the window covers the frequency contents of the basic heart sounds. Considering the PSD of the basic heart sounds, it is seen that important frequency contents of these sounds are inside the window. Therefore, the modified Gaussian window retains the energy of S_1 and S_2 segments, and attenuates the energy of murmurs and other sounds. Passing the $En(k)$ through a low pass filter results in a smoother curve (Es) for further analysis. We use an 8th order FIR filter with cutoff frequency of 10 Hz. Fig. 6 shows an up-sampled $Es(k)$ for a PCG of a child suffering from VSD (Ventricular Septal Defect).

As seen in Fig. 7, each sharp peak on the envelope, incorporates a sound. This property of the envelope provides the possibility to extract all sounds. In order to extract sharp peaks of the envelope, we use a special filter whose output $y(n)$, for an auxiliary input $x(n)$, is defined as:

$$Y(n) = x(n) - \frac{1}{2}(x(n-L) + x(n+L)) \tag{4}$$

where L is a constant number indicating the filter sensitivity. Our experimental results indicate that for $L = 50$ ms, the filter provides an efficient discrimination of the sharp peaks over the $Es(k)$.

In order to extract sounds, we compute the absolute value of derivative of $Es(k)$ and denote it as $Ed(k)$. Then, we determine sharp peaks of both $Es(k)$ and $Ed(k)$ separately by using Eq. (4). A sharp peak of $Ed(k)$ that occurs immediately after a sharp peak of the $Es(k)$, indicates end point of the corresponding sound. The last peak of the $Ed(k)$ prior to a sharp peak of the $Es(k)$ indicates the beginning of the corresponding sound.

4.2. S_1 and S_2 recognition

We separate S_1 and S_2 from other sounds using a neural network classifier. Periodicities of S_1 and S_2 in conjunction with their spectral properties are distinctive features that discriminate them from other sounds. In the first step, we have some extracted sounds constituted of K number of S_1 sounds, L number of S_2 sounds and M number of other sounds. The total number of the sounds (N) is:

$$N = K + L + M$$

Considering the periodicity of the basic heart sounds, using cross-correlation technique results in a useful feature for separation of the sounds. For an extracted sound k, the following feature indicates how well the sound is repetitive:

$$C_k = \sum_{j=1}^{N} \gamma_{kj} \tag{5}$$

$$\gamma_{kj} = Max(\Phi_{kj}(t)) \tag{6}$$

where $\Phi_{kj}(t)$ is normalized cross-correlation function of the two extracted sounds, k and j. C_k is relatively high for repetitive sounds, e.g. S_1 and S_2. In contrast, it is small for non-repetitive sounds, e.g. artifacts. Although, C_k is a suitable feature to

Table 1 – Results of the pediatric heart sound segmentation for the test data-bank.

Normal cases	Abnormal cases	Total number of cycles	Missed cycles	Misclassification	Efficiency (%)
20	40	823	19	34	93.6

reject random sounds like artifacts, it is unable to distinguish between the basic heart sounds and other repetitive sounds, e.g. S_3 and S_4. In order to recognize the basic heart sounds from S_3 and S_4, we use spectral energy of the extracted sounds over special frequency bands. We use the Arash-Band technique to obtain distinctive frequency bands whose energies provide maximum discrimination between the two classes [1]. We categorize S_3 and S_4 in one class and apply the Arash-Band technique for S_1 and S_2 separately. The Arash-Band for S_1 and S_2 are, 40–100 Hz and 0–50 Hz, respectively. As it is noted in Section 4.1, the murmurs have spectral energy above the frequency of 100 Hz. Therefore, we can use this feature for distinction of the basic heart sounds from murmurs. A feature vector with the following components can finally be defined for each extracted sound:

- C_k: Sum of cross-correlation of a sound with all other sounds.
- E_{40-100}: The energy of a sound in frequency band of 40–100 Hz.
- E_{0-50}: The energy of a sound in frequency band of 0–50 Hz.
- E_{100}: The energy of a sound in frequency band of more than 100 Hz.

Then, we use a three layers MLP artificial neural network (ANN) for recognizing the basic heart sounds. The ANN has 4 neurons and 12 neurons in input and mid-layers, respectively. The last layer of the classifier has 1 neuron which is assigned to one of the two states:

- The basic heart sounds.
- Artifact or other sounds.

The ANN is trained with the back propagation error method. We use second momentum for better training phase.

4.3. *Distinguishing S_1 from S_2*

The neural network classifier provides a series of labels. S_1 labels and S_2 labels are one after the other. Here the only problem is to determine whether the first label indicates S_1 or S_2 sound.

Separating S_1 and S_2 sounds plays an important role in pediatric heart sound segmentation. We have mentioned above that in children, diastolic period has a greater variation than systolic period due to the respiration effects. Taking into account this physiological property and using mathematical tools we can separate S_1 from S_2 and consequently systolic and diastolic periods as shown below.

Considering systolic and diastolic duration variances, we define the reduced variance function, as:

$$V(d) = \frac{Var(d)}{Mean(d)} \qquad (7)$$

where $Var(d)$ is the variance of variable d and $Mean(d)$ is the average of variable d. $V(d)$ represents relative variance of d. Now we label all the N detected sounds by:

$$L_k, \quad k = 1, \ldots, N$$

In fact, the labels are the time occurrences of the detected sounds which are measured to the start points of the sounds. We categorize the labels into two groups as even labels (LE_i), and odd labels (LO_i). We also define the following parameters:

$$T_{1i} = LE_{(i+1)} - LE_i \qquad (8)$$

$$T_{2i} = LO_{(i+1)} - LO_i \qquad (9)$$

$$d_{1i} = LE_i - LO_i \qquad (10)$$

$$d_{2i} = LO_{(i+1)} - LE_i \qquad (11)$$

$$i = 1, \ldots, N/2$$

d_1 or d_2 which results in a greater $V(d)$ belongs to diastolic interval and the ones with smaller $V(d)$ value belongs to systolic interval. The group (even or odd), which comes right after diastolic interval, belongs to S_1 sounds and the other group represents S_2 sounds.

5. Results

We have collected 1200 s of synchronous PCG and ECG signals taken from 120 children. We have used 60 samples as the training data-bank and the remaining 60 samples as the test data-bank.

The test data-bank contains 823 cardiac cycles. Each data-bank contains 20 recordings, taken from normal children and 40 recordings taken from children with congenital heart diseases. We used the test data-bank to evaluate the performance of the method. Table 1 shows results of the method.

As it is shown in Table 1, the efficiency of our method is higher than 93% which is a significant result.

6. Conclusions

Based on spectral and timing properties of a child heart sounds signal, we have developed a three step algorithm for a complete segmentation of pediatric heart sounds signal. In the first step, we have detected all different sounds existed in a heart sound signal. In the second step, we have recognized the basic heart sounds among all extracted sounds by using a neural network. In the final step, we have distinguished between S_1 and S_2 based on the children respiration effect on cardiac timings. Results show that the algorithm is efficient and could be used as an effective tool for a complete segmentation in a computerized screening of congenital heart diseases system.

Conflicts of interest

There is no conflict of interest.

REFERENCES

[1] A.A. Sepehri, J. Hancq, T. Dutoit, A. Gharehbaghi, A. Kocharian, A. Kiani, Computerized screening of children congenital heart diseases, Computer Methods and Programs in Biomedicine 92 (November) (2008) 186–192.
[2] J.P. de Vos, M.M. Blanckenberg, Automated pediatric auscultation, IEEE Transactions on Biomedical Engineering 54 (February) (2007) 244–252.
[3] C. Degroff, S. Bhatikar, J. Hertbearg, Artificial neural network-based method for screening heart murmurs in children, Circulation (June) (2001) 2711–2716.
[4] T. Sekiya, A. Watanabe, M. Saito, The use of modified constellation graph method for computer-aided classification of congenital heart diseases, IEEE Transactions on Biomedical Engineering 38 (August) (1991) 814–820.
[5] A.A. Sepehri, H. Leich, A. Gharehbaghi, End pointing of heart sounds using neural network, in: Global Signal Processing Conference (GSPx), Santa Clara, CA, USA, 2004.
[6] A. Haghighi-Mood, J.N. Torry, A sub-band energy tracking algorithm for heart sound segmentation, IEEE Computers in Cardiology (1995).
[7] H. Liang, S. Lukkavineu, I. Hartimo, Heart sound segmentation algorithm based on heart sound envelogram, IEEE Computers in Cardiology 24 (1997).
[8] H. Liang, S. Lukkavineu, I. Hartimo, A heart sound segmentation algorithm using wavelet decomposition and reconstruction, in: IEEE/EMBS Proceedings 19th International Conference, Chicago, IL, USA, 1997.
[9] T. Oskiper, R. Watrous, Detection of the first heart sound using a time-delay neural network, IEEE Computers in Cardiology 29 (2002) 537–540.
[10] R. Janson, Review of complexes, in: A Simplified Approach to Electrocardiography, W. B. Saunders Company, Philadelphia, USA, 1986.
[11] A.G. Tilkian, Phonocardiography and external pulse recording: phonocardiogram, in: Understanding Heart Sounds and Murmurs, Chapter 4, W. B. Sanders Company, Philadelphia, USA, 1984.
[12] T.E. Andreoli, C.C.J. Carpenter, Structural and function of the normal heart and blood vessels, in: Cecil Essentials of Medicine, Chapter 3, W. B. Saunders Company, Philadelphia, USA, 2007.
[13] G. Norgard, H. Vik-Mo, Effects of respiration on right ventricular size and function: an echocardiography study, Pediatric Cardiology 13 (1992) 136–140.
[14] Y. Zhendong, Effects of age and respiration on right ventricular diastolic filling pattern in normal children, Pediatric Cardiology 19 (1998) 218–220.
[15] B.L. Lendrum, A.M. Mondkar, J. Brian Harris, B. Smulevitz, I. Carr, Respiratory variation in echocardiographic dimensions of left and right ventricles in normal children, Pediatric Cardiology 1 (1979) 39–45.
[16] P.S. Nandi, V.M. Pigott, D.H. Spodick, Sequential cardiac responses during the respiratory cycle: patterns of change in systolic intervals, Chest 63 (March) (1998) 380–385.
[17] F.A. Bu'Lock, M.G. Mott, R.P. Martin, Left ventricular diastolic function in children measured by Doppler echocardiography: normal values and relation with growth, British Heart 73 (February) (1995) 334–339.
[18] A.P. Yoganathan, R. Gupta, F.E. Udwadia, J. Wayen Miller, W.H. Corcoran, R. Sarma, J.L. Johnson, R.J. Bing, Use of the fast Fourier transform for frequency analysis of the first heart sound in normal man, Medical and Biological Engineering (January) (1976).
[19] D. Chen, L.G. Durand, H.C. Lee, Time-frequency analysis of the first heart sound. Part 1. Simulation and analysis, Medical & Biological Engineering & Computing (July) (1997).

Please cite this article in press as: A.A. Sepehri, et al., A novel method for pediatric heart sound segmentation without using the ECG, Comput. Methods Programs Biomed. (2009), doi:10.1016/j.cmpb.2009.10.006