A Cardiac Sound Localization and Identification Method for Electronic Stethoscope

Ning Zhou, Nan Hu*
School of Electronic and Information Engineering,
Soochow University
Suzhou, China
Email: 2267226559@qq.com, * hunan@suda.edu.cn

Shengsheng Cai, Xingguo Xu
Mintti Medical Technology Co. , Ltd.
Hefei, China
Email: caishengsheng@mintti.cn, xuxingguo@mintti.cn

***Abstract*—In the study of electronic stethoscope, analysis of phonocardiogram depends on the localization and identification of the first cardiac sound (CS) S1 and the second CS S2, which are basic components of the CS signal. However, in realistic environments the recorded CS signal is often mixed with the respiratory sound (RS) and other ambient interferences, which may cause failure in CS component localization and identification. In this paper, a CS localization and identification method is proposed, which is based on two steps: the rough CS localization step and the identification and mending step. The rough CS localization step involves calculating the approximate Shannon entropy, to roughly localize the CS components in low computational complexity. On account of the influence of RS and other ambient interferences, the aforementioned first step may fail to localize some CS components. By calculating and comparing the ratio of low-frequency power to high-frequency power for each CS component, the second step of the method can amend the rough CS localization results and identify which type of CS component they belong to. At last, the estimation of heart rate (HR) is also easily derived from the CS localization and identification results. Experiments using data recorded in various conditions shows the efficiency of the proposed method.**

Keywords—phonocardiogram; cardiac sound localization; heart rate estimation.

I. INTRODUCTION

The development of digital signal processing technology provides the possibility of automatic noninvasive low-cost diagnosis of heart and lung diseases by using electronic stethoscopes, in place of the traditional auscultation art mastered by only a few skilled physicians. The study of phonocardiogram (PCG) is based on analyzing the cardiac sound (CS) signal recorded by electronic stethoscope, which is a kind of complex sound signal that can vary among various subjects in various health statuses. The problem of cardiopulmonary sound overlap is always met when we processing the data recorded by electronic stethoscope in a real environment, as the CS and the respiratory sound (RS) overlap not only in the time domain but also in the frequency band 60-320 Hz [1]. Besides, there are many environment noises interfere the recorded data, which also affect the efficiency of auscultation. The key to address this problem is correctly segmenting the data where CS does exist and then separating CS and RS out for following analysis. In this context, correct localization and identification of CS components plays an important role in assessing the heart and lung status using electronic stethoscopes. Generally, CS includes four components in one period: the first component S1, the second component S2, the third component S3 and the fourth component S4 [2], where S1 and S2 are audible parts. The S1 represents the first CS of the CS signal and the S2 represents the second CS of the CS signal. The sound signal of healthy children and young people includes the third heart sound, which is a normal phenomenon, and the fourth heart sound is a pathological heart sound. In this paper, we only concern the localization and identification of S1 and S2.

In the literature, localization of CS embedded in RS signals is usually realized by thresholding the extracted feature sequences according to some characteristics of RS and CS signals [3-6]. Adaptive thresholding is the most practical method because it does not require any prior information. The accuracy of these CS localization methods depends, of course, on the corresponding thresholds and the effectiveness of extracted features for distinguishing CS from RS. A multiresolution products based method using wavelet coefficients was proposed in [4] for localizing CS. By the method of variance fractal dimension trajectory [5], localization of CS was achieved in lung sounds recordings. Third-order cumulant was used in [7] to identify the nonlinear data parts and a time-frequency analysis method was proposed. A multi-scale mean shift CS localization method was given in [8], where the data was divided into linear and nonlinear parts firstly by higher-order statistics [9] and wavelet coefficients. The Shannon entropy was used to localize the CS components [10]. This Shannon entropy based method was claimed to achieve excellent localization results, while it needs a large computational complexity. In fact, most of the existing CS localization methods only consider the scenarios when the subjects are taking breath at low or medium flow rates in quite environment, while they may deteriorate when the subjects are taking deep breath or the ambient is very noisy.

In this paper, the problem of CS localization and identification is addressed, and a method is proposed which has two steps: the rough CS localization step and the identification and mending step. By calculating the approximate Shannon entropy with low computational complexity, the first step can roughly localize the CS components S1 and S2. By calculating and comparing the

This work was supported by the National Natural Science Foundation of China [Grant 61601316].

978-1-5386-7946-3/18/$31.00 ©2018 IEEE

ratio of low-frequency power to high-frequency power for each CS component, the second step of the method can amend the rough CS localization results, which are easily affected by RS at high flow rates and other ambient interference, and identify which type of CS component they belong to. Ultimately, the estimation of heart rate (HR) can be easily derived from the CS localization and identification results. Data recorded from experiments in various realistic scenarios, that including the subjects are taking deep breath or the ambient is very noisy, are used to verify the efficiency of the proposed method.

II. METHOD

The proposed CS localization and identification method includes two steps: the first step is to roughly localize the CS components by estimating the approximate Shannon entropy of the recorded data, and the second step is to amend the rough localization results that are easy to deteriorate caused by RS and ambient noise and identify which type of CS components they belong to.

A. Rough localization

As probability can be used to measure the uncertainty of information, Shannon entropy was used to design a CS localization method in [10]. Assume that a segment of recorded data with N samples is $\{X_1, X_2, \cdots, X_N\}$, and then the probability density function of the data distribution estimated using this data segment is given as follows:

$$p(x) = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{h}K\left(\frac{x - X_n}{h}\right) \tag{1}$$

where $K(x) = \exp(-x^2/2)/\sqrt{2\pi}$ is the Gaussian kernel, and h is Gaussian kernel bandwidth. The method in [10] calculated the Shannon entropy of $p(x)$

$$H(p) = -\int_{-\infty}^{\infty} p(x)\log p(x)\,dx \tag{2}$$

and compared $H(p)$ with an adaptive threshold to determine whether this data segment is located in a certain CS component.

However, an accurate estimation of $H(p)$ will cause a large computational complexity, and hence instead in this paper we calculate the approximate Shannon entropy

$$\hat{H}(p) = -\sum_{k=1}^{K} wp(x_k)\log p(x_k) \tag{3}$$

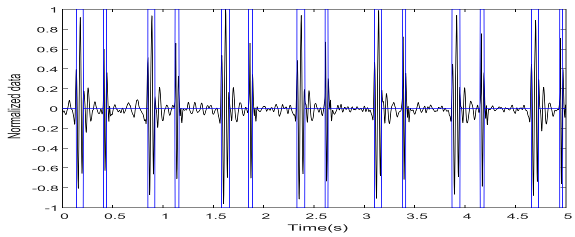
where $x_k, k = 1, 2, \cdots, K$ are samples from the uniform sampling in $[\mu - 10\sigma, \mu + 10\sigma]$ with sampling interval $w = 20\sigma/(K-1)$, and μ and σ are mean and standard deviation of this data segment, respectively. The Gaussian kernel bandwidth for $p(x)$ is given by $h = 1.06\sigma/N^{0.2}$ as suggested in [9]. The detection criterion of this rough localization procedure is given as follows:

$$\hat{H}(p) \underset{H_0}{\overset{H_1}{\gtrless}} \alpha(\mu + \sigma) \tag{4}$$

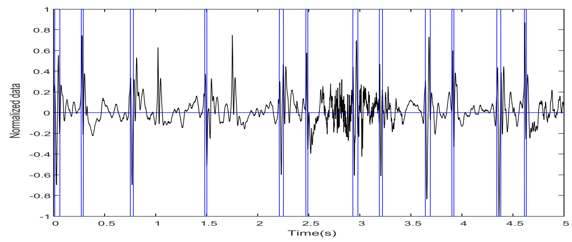
where H_1 denotes that CS component is detected in this data segment, H_0 denotes the opposite of H_1, and α is a preset constant parameter. The estimate of $\hat{H}(p)$ greatly reduces the computational complexity of Shannon entropy based CS localization method.

Commonly, the length of each data segment should be set to be much shorter than that of S1 or S2. Hence, in each CS component there will be several continuous or discontinuous data segments that are decided to be located in CS. In our proposed method, we connect the detected data segments that are close to each other together, whether they are continuous or discontinuous, and consider them to belong to the same CS component.

In many papers, it was claimed that though troubled by the high computational complexity, the Shannon entropy based CS localization method in [10] can achieve excellent performance. This may be true when the subjects are breathing at low or medium flow rates in quite environment. When the subjects are taking a deep breath or the CS is interfered, this method is easy to deteriorate. Fig.1 (a) and Fig.1 (b) show the performances of the method in [10] when one subject is taking a normal breath and a deep breath, respectively, where the black line represents the signal data and the blue line represents the localization result of CS components. It can be noticed that compared with the satisfactory performance when normal breath, in the deep breath scenario, the detection of some CS components is possibly to be missed. In this context, we should amend this rough localization result.



(a)



(b)

Fig.1. Examples of localization results of the method in [10]: (a) the result for a batch of normal breath data; (b) the result for a batch of deep breath data.

B. *Identification and Mending*

As the example depicted in Fig. 1 (b), though there may be some CS components not detected by the aforementioned first step, for a long enough time, there will be one or more couples of S1 and S2 detected in the rough localization step. In our method, if the distance between two adjacent detected CS components is less than a certain short time interval, we consider that the front one is S1 and the next one is S2, and they belong to the same CS cycle. This preliminary identification of CS components is reasonable, as the distance between the adjacent S1 and S2 is always much shorter than that between the adjacent S2 and S1. The left unclassified CS components will be further identified by using the information derived from the already classified ones.

The identification of the unclassified CS components is based on the fact that S1 has more low-frequency components than S2, which is shown in Fig. 2. If the vector of fast Fourier transform (FFT) result of some detected CS component data is $\mathbf{X}_f = \left[\mathbf{X}_{fl}^{\mathrm{T}}\ \mathbf{X}_{fh}^{\mathrm{T}}\right]^{\mathrm{T}}$, where $(\cdot)^{\mathrm{T}}$ denotes transpose and $\mathbf{X}_{fl}$ and $\mathbf{X}_{fh}$ are low-frequency component vector and high-frequency component vector, respectively, we calculate the ratio of low-frequency power to high-frequency power

$$R = p_{fl} / p_{fh} \tag{5}$$

for each detected CS component, where $p_{fl} = \left\|\mathbf{X}_{fl}\right\|_2^2$ and $p_{fh} = \left\|\mathbf{X}_{fh}\right\|_2^2$ are low-frequency power and high-frequency power, respectively. Of course, R for S1 is much larger than that for S2.

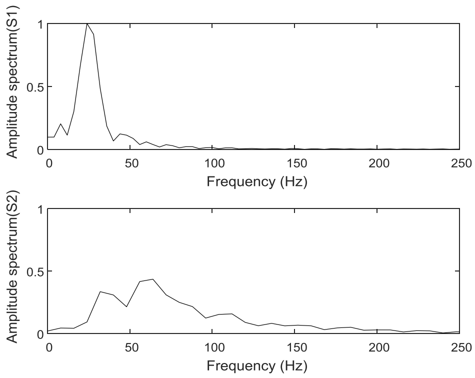


Fig. 2. An example of amplitude spectrum of CS components, where the upper figure shows S1 and the lower figure shows S2.

With the aid of R, the unclassified CS components now can be identified by using the following criterion:

$$\left|R - \bar{R}_2\right| \underset{\tilde{H}_2}{\overset{\tilde{H}_1}{\gtrless}} \left|R - \bar{R}_1\right| \tag{6}$$

where $\bar{R}_1$ and $\bar{R}_2$ are the average of Rs corresponding to the preliminary identified S1 and S2, respectively, and $\tilde{H}_1$ and $\tilde{H}_2$ denote that the unclassified CS component are identified as S1 and S2, respectively.

The localization result can be amended after all CS components have been identified. The main task is to localize the missed partner of the CS component identified using (6) in the same CS cycle. The key is using the average distance between adjacent S1 and S2 and the average distance between adjacent S2 and S1. The last procedure is to amend the range of each CS component. As mentioned in [11], S1 generally lasts for 0.14s and S2 lasts for 0.11s, and hence we amend the range of S1 to be 0.14s outward from the center and 0.11s for S2 in the same way. This range mending is useful for further study of separation of cardiac and respiratory signals.

At last, HR can be easily estimated by

$$hr = 60 / \bar{T}_1 \tag{7}$$

where $\bar{T}_1$ is the average distance between adjacent S1 and S1.

III. Result and Outcome

The efficiency of the proposed method is directly tested by data recorded using a electronic stethoscope developed by Mintti Medical Technology Co., Ltd., and three healthy subjects aged from 24 to 27 years was recruited. The data acquisition sampling frequency was 8 kHz and the auscultation position was on the left chest. The length of each data segment for approximate Shannon entropy calculation was 0.02s, and α was fixed to be 0.1.

A. *Localization Performance*

The first subject was asked to take deep breath in a quiet environment, and we use a short fragment (3.7s) of the recorded data to show the performance of the proposed method in this scenario, depicted in Fig. 3, in comparison with the method in [10]. It can be noticed that only the proposed method can correctly localize all the CS components and designate their lasting ranges.

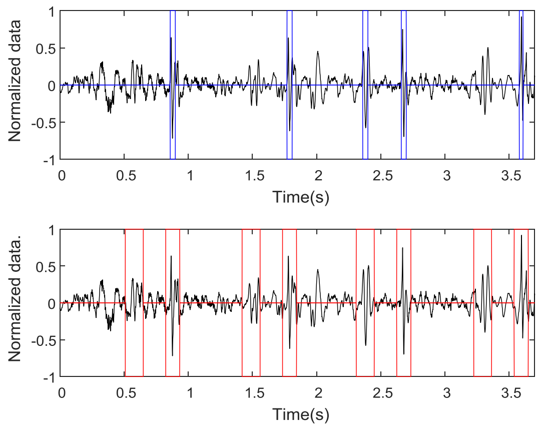


Fig. 3. Localization results for data recorded in the deep breath scenario: the upper figure shows the result of the method in [10], which is marked by blue blocks; the lower figure shows the result using the proposed method, which is marked by red blocks.

The second subject was asked to take deep breath in a noisy environment, when someone aside was speaking loudly. A short fragment (3.7s) of the recorded data was used to show the performance of the proposed method in this scenario, depicted in Fig. 4, in comparison with the method in [10]. It can be noticed that even in this harsh condition, the proposed method can still localize all the CS components and designate their lasting ranges, while the compared one missed some CS component caused by the ambient interference.

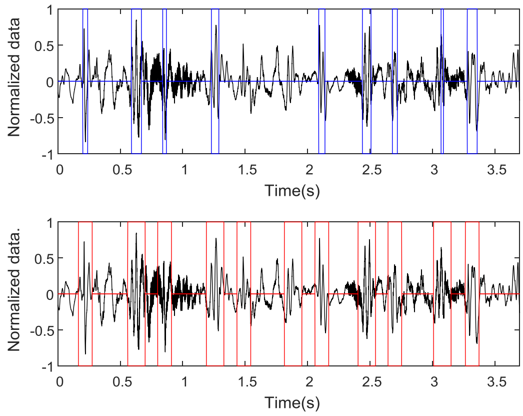


Fig. 4. Localization results for data recorded in the deep breath and ambient interference scenario: the upper figure shows the result of the method in [10], which is marked by blue blocks; the lower figure shows the result using the proposed method, which is marked by red blocks.

Table I shows the accuracy of CS localization in three scenarios for normal breathing, deep breathing in quiet environment and deep breathing in noisy environment, by using longer recorded data. The proposed method has much better localization performance compared with the method in [10], especially when the CS signal is interfered by RS signal or ambient noise. The proposed method achieves 100% localization accuracy in normal breathing. Though its performance will degrade by interference, it still has 97.2% localization accuracy even when interfered by both RS signal and ambient noise.

TABLE I. LOCALIZATION ACCURACY IN VARIOUS SCENARIOS

Experimental scenario	**Total number of CS components**	**Rate of correct localization**	
		Method in [10]	***Proposed method***
normal breathing	435	98.4 %	100 %
deep breathing in a quiet environment	552	88.2 %	99.8 %
deep breathing in a noisy environment	392	72.2 %	97.2 %

B. HR Estimation Performance

The data recorded from the third subject lasts 14s, and is used to verify the efficiency of HR estimation from the localization and identification results.

Firstly, the subject was asked to take deep breath in a quiet environment without a break. The following figures show the localization result and HR estimation result for this data. HR was estimated to be 0 at first because that the estimation of HR needs at least two whole CS cycles. Fig.5 shows the location results of the method in [10], where some of the CS components were not located, which affected the estimation of HR in the lower figure of Fig.5. Fig.6 shows the result of the proposed method, where it can locate the CS components completely and improve the accuracy of HR estimation. It is interesting to notice that the fluctuating HR was well consistent with respiration.

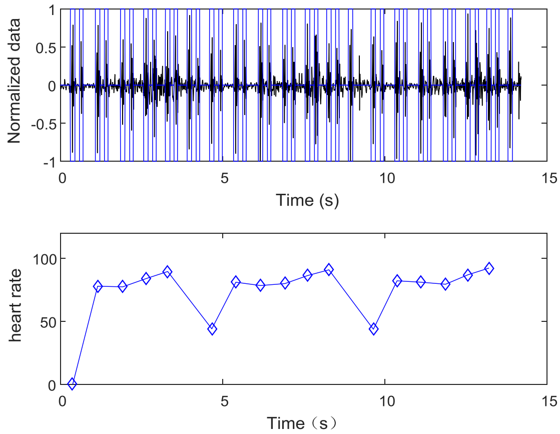


Fig. 5. Results of CS localization and HR estimation by the method in [10] in a deep breath scenario: the upper figure shows the localization result, the lower figure shows the result of HR estimation.

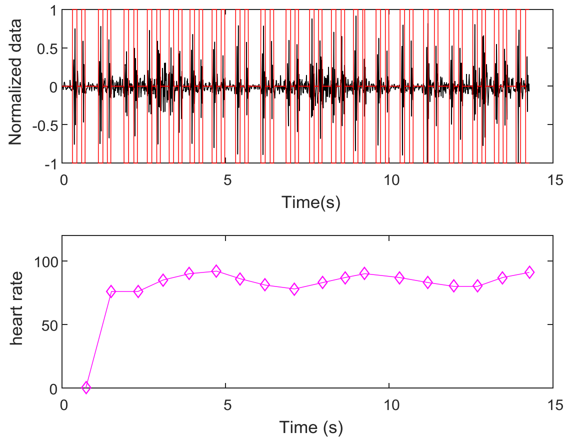


Fig. 6. Results of CS localization and HR estimation by the proposed method in a deep breath scenario: the upper figure shows the localization result; the lower figure shows the result of HR estimation.

Secondly, the subject was asked to take normal breath in a quiet environment, while after a certain time duration the electronic stethoscope was removed from the auscultation position for about 3s and then put back to the previous place.

Fig. 7 shows the localization result and HR estimation result for this data. It was shown that the CS localization and HR estimation still worked for this particular auscultation interruption and resuming scenario. When the electronic stethoscope was removed the HR estimation quickly fell to zero, while when resuming auscultation the localization and HR estimation quickly recovered. It can be noted that it does not affect the calculation of the heart rate in the proposed method when there is a pause in the procedure of auscultation.

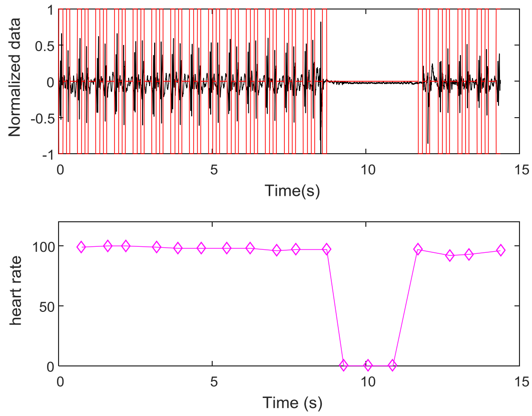


Fig. 7. Results of CS localization and HR estimation by the proposed method in a normal breath scenario with a 3s interruption: the upper figure shows the localization result; the lower figure shows the result of HR estimation.

IV. Conclusion

A CS component localization and identification method, together with HR estimation, has been proposed. The approximate Shannon entropy was used to roughly localize the CS components in real time. By calculating and comparing the ratio of low-frequency power to high-frequency power for each CS component, the rough CS localization results was identified and amended. HR estimation was derived from the amended localization results. Experiments performed in various real scenarios verified the efficiency of the proposed method.

References

[1] S. Debbal, F. R. Bereski. "Spectral analysis of the PCG signals," Internet. J. Med. Technol., vol.1, 2006.

[2] I. Hossain and Z. Moussavi, "An overview of heart-noise reduction of lung sound using wavelet transform based filter," in Proc. IEEE Eng. Med. Biol. Soc, 2003, pp. 458-461.

[3] V. Iyer, P. Ramamorthy, H. Fan, and Y. Ploysongsang, "Reduction of heart sounds from respiratory sounds by adaptive filtering," IEEE Trans. Biomed. Eng. vol.33, no.12, pp.1141-1148, December. 1986.

[4] Z. Moussavi, D. Flores, and G. Thomas, "Heart sound cancellation based on multi-scale products and linear prediction," in IEEE Eng. Med. Biol. Mag. Vol. 54, no. 2 , pp. 234-243. September. 2005.

[5] J. Gnitecki and Z. Moussavi, "Variance fractal dimension trajectory as a tool for heart sound localization in lung sounds recordings,"in Proc. Eng. Med. Biol. Soc., 2003, pp. 2420-2423.

[6] M. Kompis and E. Russi, "Adaptive heart-noise reduction of respiratory sounds recorded by a single microphone," in Proc. 14th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc, 1992, pp. 691-692.

[7] F. Jin and F. Sattar, "Tracking and time-frequency analysis an nonlinearity of tracheal sounds," Med. Bio Eng. Comp. vol. 47, no. 4, 2009, pp. 457-461.

[8] F. Jin, Farook Sattar, "A multiscale mean shift localization approach for robust extraction of heart sounds in respiratory signals," IEEE Int Conf. Acou. vol. 32, no. 3, pp. 1291-1295. 2013.

[9] L. Hadjileontiadis and S. Panas, "Adaptive reduction of heart sounds from lung sounds using fourth-order statistics," IEEE Trans. Biomed. Eng. vol. 44, no. 7, pp. 642–648, July. 1997.

[10] A. Yadollahi and Z. M. K. Moussavi, "A Robust Method for Heart Sounds Localization Using Lung Sounds Entropy," IEEE Trans. Biomed. Eng. vol. 53, no. 3, March 2006.

[11] E. Kaniusas, Biomedical signals and sensors II: linking acoustic and optic biosignals and biomedical sensors. Berlin, Heidelberg: Springer-Verlag, 2015.