The Automatic Repairing Method Addressing Clipping Distortions and Frictional Noises in Electronic Stethoscope

Ning Zhou, Jiajun Wang, Bing Sun, Renyu Liu, Nan Hu*
School of Electronic and Information Engineering, Soochow University
Suzhou, China
e-mail: 20165228036@stu.suda.edu.cn, jjwang@suda.edu.cn, sunbing@suda.edu.cn, 1002280496@qq.com, *hunan@suda.edu.cn

***Abstract*—The auscultation signal collected by the electronic stethoscope may be sometimes accompanied by various interferences, including external speech/acoustic interferences, clipping distortions, frictional noises, etc. The external speech/acoustic interferences can be eliminated by adaptive filtering, with the aid of an extra recording sensor. However, clipping distortions and frictional noises cannot be addressed by this methodology, and how to automatically repair them has not been fully discussed in the literatures, which affects the signal quality and further the cardiopulmonary sound automatic diagnosis. In this paper, the repairing method that automatically addresses clipping distortions and frictional noises for electronic stethoscope is developed. A simple signal difference method is introduced to automatically detect the clipping distortion regions, and these regions are repaired by the Hermite interpolation. The regions that frictional noises exist are detected by employing Mel-frequency cepstral coefficients (MFCCs) and support vector machine (SVM), and they are repaired by involving the empirical mode decomposition (EMD) as well as correlation coefficients. The proposed method can automatically detect, locate and ultimately repair multiple regions of clipping distortions and frictional noises, and applying it in recorded real auscultation data proves its efficiency.**

Keywords-auscultation; electronic stethoscope; clipping distortion; frictional noise

I. INTRODUCTION

Auscultation is one of the most important techniques for physicians to know the condition of patients in first time. With the rapid development of Internet of things (IoT), various types of electronic stethoscopes keep emerging, rendering real-time monitoring of patients, telemedicine and automatic diagnosis using cardiopulmonary sound data through artificial intelligence (AI).

The electronic stethoscope converts weak physiological sound signals, such as cardiac sound and respiratory sound, into electrical signals through a transducer on the head of the stethoscope. In practical applications, the recorded data may be sometimes polluted by noises or interferences. For example, in clinical usage the ambient speech and acoustic sounds are usually recorded simultaneously, and they can be eliminated by adopting adaptive filtering with the aid of an extra reference sensor recording the ambient interferences. However, there are still other noises or interferences that cannot be addressed by this scheme, which includes noise of various organs of the body, clipping distortions caused by excessive pressure on the head of stethoscope, frictional noises [1] caused by the friction between the head of stethoscope and clothes, etc. The typical way of denoising in auscultation is adopting wavelet decomposition [2-5]. According to the characteristics of wavelet coefficients at different scales, multi-level wavelet decomposition threshold was used to remove noise in [5]. A group of nonlinear chirp signals were used as atoms for matching pursuit to eliminate noise [6]. An improved algorithm combining empirical mode decomposition (EMD) and wavelet was proposed for phonocardiogram signal denoising in [7]. Singular value decomposition (SVD) and wavelet packet transform was used to remove the noise in phonocardiogram signal in [8].

For the auscultation data where clipping distortions and frictional noises exists, direct application of aforementioned denoising methods may be not appropriate, as the characteristics of these two interferences are very different from the conventional noises which are usually modeled as Gaussian distributed. A clipping distortion restoration method using cubic spline interpolation was proposed in [9], while how to automatically detect and locate the clipping distortion regions was not mentioned.

In this paper, the problem of automatic detecting and repairing regions of clipping distortions and frictional noises is addressed. For clipping distortions, by thresholding the differenced signal, the regions where distortion happened are detected and located, and they are further repaired in turn by using Hermite interpolation method [10]. For frictional noise interference, the Mel-frequency cepstral coefficients (MFCCs) [11] are used to extract its characteristics and then the support vector machine (SVM) [12] is employed to detect the frictional noise regions. In this regions, EMD is used to decompose the data into intrinsic mode function (IMF) components, and correlation coefficients are calculated to judge which parts in IMF components should be reserved. The repaired signal is obtained by summing selected IMFs or parts of IMFs, and the truncation areas are also repaired by Hermite interpolation. The real data recorded by a electronic stethoscope is used to verify the efficiency of the proposed method.

II. REPAIRING CLIPPING DISTORTIONS

When collecting cardiopulmonary sound signals, if an excessive pressure is given on the head of the electronic stethoscope, the recorded signal will be distorted in a short

978-1-7281-0641-0/19/$31.00 ©2019 IEEE

time slot. This phenomenon may happen when the patient is uncooperative, for example in application in pediatric clinic. A typical example where clipping distortion happened is shown in Figure 1.

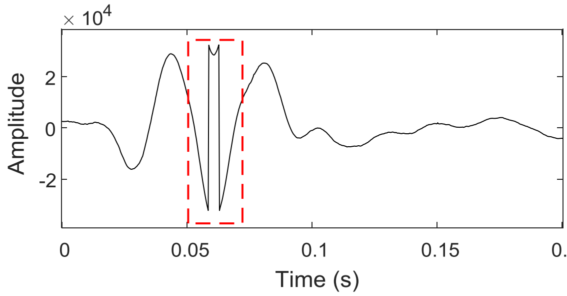

Figure 1. A typical example of clipping distortion, where the red dashed box covers the region where clipping distortion happened.

There may be some AI methods that can be applied to automatically detect clipping distortion, while in fact it is not essential to utilize these complicated methods. It can be observed from Figure 1 that, at the moments when the recorded data converts from normal state to saturation state or in reverse, the amplitude of recorded signal changes drastically. Hence, we propose to use the following simple routine to automatically detect clipping distortion regions.

Firstly, we calculate $\mathbf{x}_d$, which is the difference of the original N-sample signal vector $\mathbf{x}$:

$$x_d(1) = 0,$$
$$x_\mathrm{d}(n) = x(n) - x(n-1), \quad n = 2,3,...,N. \tag{1}$$

Secondly, we set a threshold Th, which is typically given as

$$Th = \alpha \cdot \max(|x|), \tag{2}$$

where $0 < \alpha < 1$ is a preset constant. We find out the time points whose absolute value in $\mathbf{x}_d$ is greater than Th, and then use these time points to form a set $\Delta = \{n_1, n_2, ..., n_L\}$. The adjacent points in the set Δ that have opposite signs are paired, and they are determined as the starting and ending points of a clipping distortion region. If a single unpaired time point is found near the start or end of the whole data, then the previous or later data should be combined for additional pairing. According to the pairing results, K nonoverlapping regions of clipping distortion $\lfloor n_{1,\mathrm{begin}}, n_{1,\mathrm{end}} \rfloor, \lfloor n_{2,\mathrm{begin}}, n_{2,\mathrm{end}} \rfloor, ..., \lfloor n_{K,\mathrm{begin}}, n_{K,\mathrm{end}} \rfloor$ can be obtained.

Lastly, if $K \geq 1$, we need to repair these regions in turn from left to right. In order to guarantee the continuity and smoothness of the repaired data, the Hermite interpolation [10] method is employed. For the kth region, we use the following scheme to obtain its interpolation points: on the original signal $\mathbf{x}$, go backward in time to find the third zero cross point (ZCP) $n_{K,\mathrm{left3}}$ from point $n_{k,\mathrm{begin}}$, and go forward in time to find the first ZCP $n_{k,\mathrm{right1}}$ from point $n_{k,\mathrm{end}}$. The collection of interpolation points for the kth clipping distortion region is given by

$$\Psi_k = \lfloor n_{k,\mathrm{left3}}, n_{k,\mathrm{begin}} - P \rfloor \cup \lfloor n_{k,\mathrm{end}} + P, n_{k,\mathrm{right1}} \rfloor, \tag{3}$$

where P =1~10 is the number of reserved data points considering transition time of saturation in electronic stethoscope. For the $T_k + 1$ time points in $\Psi_k : n_0, n_1, ..., n_{T_k}$, we estimate the derivative $x'(\mathrm{n_j})$ at each time point. The fitting point $\tilde{x}(\mathrm{m})$ with $m \in \lfloor n_{K,begin} - P + 1, n_{K,end} + P - 1 \rfloor$ is ultimately given by

$$\tilde{x}(m) = \tilde{x}_1(m) + \tilde{x}_2(m), \tag{4}$$

where

$$\tilde{x}_1(\mathrm{m}) = \sum\nolimits_{i=0}^{T_k} \left\{ \left[1 - 2(m - n_i) h_i \right] f_i^2(m) x(n_i) \right\},$$
$$\tilde{x}_2(m) = \sum\nolimits_{i=0}^{T_k} \left\{ (m - n_i) f_i^2(m) x'(n_i) \right\},$$
$$f_i(m) = \sum\nolimits_{\substack{j=0 \\ j \neq i}}^{T_k} (m - n_j) / (n_i - n_j) \text{ and } h_i = \prod\nolimits_{\substack{j=0 \\ j \neq i}}^{T_k} 1 / (n_i - n_j).$$

III. Addressing Friction Noises

The generation mechanism and characteristics of frictional noises is far more complicated than that of clipping distortions. In this paper, we combine MFCCs [11] with SVM [12] for detection and localization of frictional noise interfered regions, where the signal characteristics are extracted by MFCCs and then classified by SVM. We divide data $\mathbf{x}$ into segments with 0.2s (50% overlapping) and calculate the power spectrum for each data segment

$$|Y(k)|^2 = \left| \sum_{m=1}^{M} x(m) h(m) \mathrm{e}^{\frac{-j2\pi mk}{M}} \right|^2, 1 \leq k \leq N_{\mathrm{FFT}}/2 + 1, \tag{5}$$

where $\mathbf{h}$ denotes the Hanning window and N_{FFT} is the point number of fast Fourier transform (FFT). Then we convert the linear frequency into the Mel frequency by

$$f_{mel}(f) = 2959 \times \log_{10}(1 + f/700), \tag{6}$$

and construct a filter bank with Q linearly spaced triangular filters (50% overlapping) in Mel frequency domain to obtain

$$y(q) = \sum_{k=1}^{N_{FFT}/2+1} |Y(k)|^2 \Psi_q(k), q = 1,2,...,Q, \tag{7}$$

where Ψ_q is the qth filter in the designed filter bank. The qth element of MFCCs is ultimately given by

$$c(q) = \sqrt{\frac{2}{Q}} \sum_{l=0}^{Q-1} \log[y(l+1)] \cos\left[q \cdot \frac{2l-1}{2} \cdot \frac{\pi}{Q} \right], q = 1,2,...,Q \tag{8}$$

The vector of MFCCs $\mathbf{c}$ for each data segment is input into a linear SVM, which is given by

$$f(\mathbf{c}) = \mathrm{sign}(\mathbf{w}^\mathrm{T} \mathbf{c} + b), \tag{9}$$

where $\mathbf{w}$ and b are pre-trained slope vector and intercept variable, respectively. Only when $f(\mathbf{c}) > 0$, frictional noises are determined to appear in this data segment. If several adjacent data segments are determined to be interfered by frictional noises, they are merged in to the same region of frictional noise.

If we detect one or more regions of frictional noise, we still need to repair these regions in turn from left to right. For the kth region, we use EMD [7] to decompose the corresponding data $\mathbf{x}_k$ into M_k IMFs, which can be represented as

$$\mathbf{x}_k = IMF_1 + IMF_2 + ... + IMF_{M_k} . \quad (10)$$

Among the derived IMFs, IMF_1 mainly contains very high frequency components, so it can be deemed that there are no components of cardiac or respiratory sounds. $IMF_4, IMF_5, ..., IMF_{M_k}$ only contains components of cardiopulmonary sound signals. It can also be noticed that in IMF_2 and IMF_3 the components of frictional noises and cardiopulmonary sound signals exist simultaneously.

Now the main task is deriving the cardiopulmonary sound components embedded in IMF_2 and IMF_3. It can be noticed that they right correspond to the common components in IMF_2 and IMF_3. Therefore, we just need to find these areas by the following procedure: firstly, IMF_2 and IMF_3 are filtered by the same bandpass Butterworth filter; secondly, they are segmented into 0.02s frames (50% overlapping); thirdly, correlation coefficients between each corresponding aligned data frame of IMF_2 and IMF_3 are calculated; lastly, reserve the data frames whose correlation coefficients are larger than a preset threshold Th_{corr}. When the distance between two adjacent reserved data frames is less than 0.02s, they are merged into the same data frame. Thus ultimately P nonoverlapping data frames are obtained, and the corresponding data point intervals are $\left[\tilde{n}_{1,\text{begin}}, \tilde{n}_{1,\text{end}}\right]$, $\left[\tilde{n}_{2,\text{begin}}, \tilde{n}_{2,\text{end}}\right]$, ..., $\left[\tilde{n}_{P,\text{begin}}, \tilde{n}_{P,\text{end}}\right]$. The data points of IMF_2 and IMF_3 in the selected P data point intervals are reserved, and the rest ones are set to be zeros, which leads to the new components $\overline{IMF}_2$ and $\overline{IMF}_3$. The repaired version of $\mathbf{x}_k$ is given by

$$\tilde{\mathbf{x}}_k = \overline{IMF}_2 + \overline{IMF}_3 + IMF_4 + ... + IMF_{M_k} . \quad (11)$$

We may meet data discontinuity at the edge data points in the selected P data point intervals, so here the Hermite interpolation is also employed to guarantee the continuity and smoothness of the repaired data. To obtain the fitting data points in the edge of the pth interval, we use $\left[2\tilde{n}_{p,\text{begin}} - \tilde{n}_{p,\text{end}}, \tilde{n}_{p,\text{begin}} - L\right] \cup \left[\tilde{n}_{p,\text{begin}} + L, \tilde{n}_{p,\text{end}} - L\right] \cup \left[\tilde{n}_{p,\text{end}} + L, 2\tilde{n}_{p,\text{end}} - \tilde{n}_{p,\text{begin}}\right]$ as the corresponding interpolation points.

IV. Experimental Results

We use the data recorded by the electronic stethoscope Smartho-D2 (Figure 2) produced by *minttihealth*, to verify the efficiency of the proposed repairing method. Two healthy subjects were recruited and the data was recorded in a quiet environment.

Figure 2. Smartho-D2 produced by *minttihealth*.

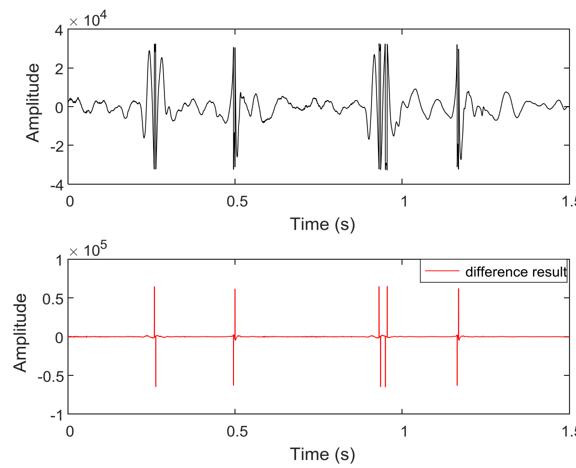

Figure 3. Original signal and its difference result in the first experiment. Upper figure: original signal; Lower figure: difference result.

A. Repairing Clipping Distortions

The first data was recorded from the left anterior chest of the first subject. To generate clipping distortions, the head of electronic stethoscope was strongly and rapidly pressed several times during auscultation. The sampling rate was 4kHz and the time duration was 1.5s.

The upper figure in Figure 3 shows the original data, and the lower one shows the signal difference results obtained by (1). It can be observed that when clipping distortion occurs, the corresponding difference result will be abrupt. Here we set $Th = 0.5 \cdot \max(|x|)$, and when the difference result is larger than Th the starting or end point of a clipping distortion can be determined. Accordingly, 5 regions of clipping distortions have been automatically detected and located, and the results are shown in Figure 4. It can be noticed that all regions of clipping distortions have been correctly localized, including two closely spaced ones (0.0136s).

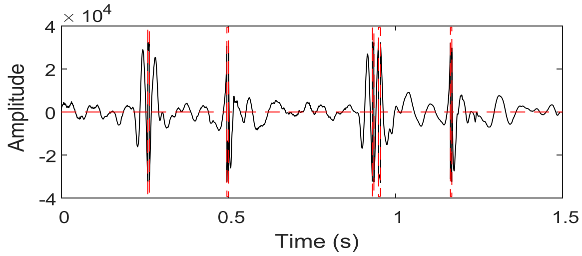

Figure 4. Clipping distortion region localization results, where the dashed boxes show the locations of these regions.

After localizing the regions of clipping distortions, the proposed method repaired these regions in turn, and the repaired result is shown in Figure 5. From Figure 5, it can be observed that the repaired result reserved the characteristics of cardiac sound well, and it also sounds natural through hearing assessment. Another advantage that can be noticed is that the proposed repairing method can easily address closely spaced regions of clipping distortions.

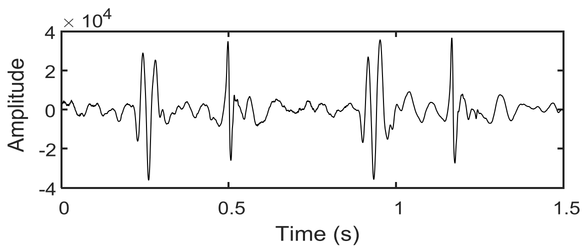

Figure 5. The data after repairing the clipping distortion regions.

B. *Repairing Frictional Noises*

The second experiment was used to examine the performance in addressing frictional noises. The data was recorded from the right anterior chest of the second subject. During data recording, we deliberately rubbed the head of electronic stethoscope on the subject's clothes in a short time, to generate frictional noises. The acquisition time was 3s and the sampling rate was 4kHz.

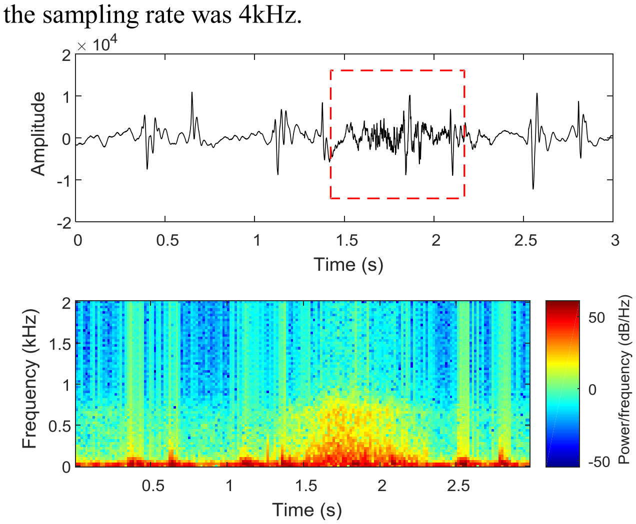

Figure 6. The waveform and spectrum of the data recorded in the second experiment. Upper figure: signal waveform; Lower figure: power spectra obtained by short time Fourier transform (STFT).

The waveform of the recorded data and its corresponding STFT results are shown in Figure 6. It is shown that a region interfered by frictional noises exists in this data, and it introduces a lot of high frequency components. By employing the MFCCs and SVM based frictional noise detection method, where the parameters of SVM was pre-trained by using a database including 214 segments of frictional noise interfered data and 325 segments of normal data. The localized time interval is also shown in Figure 6, where $n_{begin} = 1.4s$ and $n_{end} = 2.2s$ marked by the red box.

We decompose the data in the time interval $\lfloor n_{begin}, n_{end} \rfloor$ by EMD, and the IMFs are shown in Figure 7.

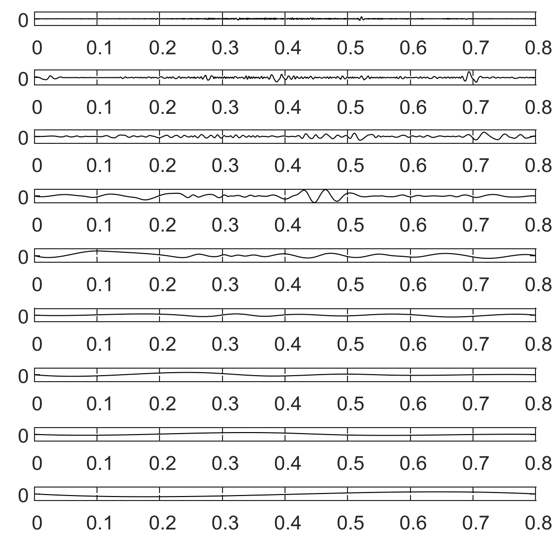

Figure 7. The IMFs obtained by EMD in $\lfloor n_{begin}, n_{end} \rfloor$, and from the top to bottom they are arranged as IMF_1, IMF_2, …. IMF_9.

The correlation coefficients between corresponding data frames of IMF_2 and IMF_3 were calculated and they are shown in the bottom figure of Figure 8.

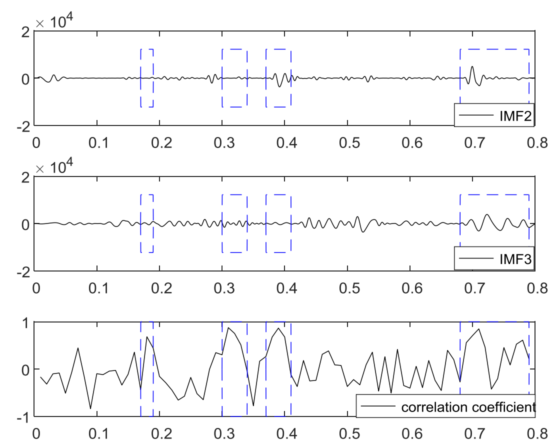

Figure 8. Bandpass filtered IMF_2, IMF_3 and their correlation coefficients.

In Figure 8, the high correlation coefficients between IMF_2 and IMF_3 indicate the locations of cardiopulmonary sound signal components in IMF_2 and IMF_3. We reserved the data frames whose correlation coefficients are greater than the preset threshold value $Th_{corr} = 0.45$. When the distance between two adjacent reserved data frames is less than 0.02s, they were merged into the same data frame. Thus, $P = 4$ nonoverlapping data intervals were obtained, as shown in the blue dashed boxes.

By using the selected 4 data intervals in IMF_2 and IMF_3, as well as IMF_4 ~IMF_9 , and further applying the Hermite interpolation, we obtained the repaired results as shown in Figure 9. From the STFT results in Figure 9, the repaired results are very similar to the data that not influenced by frictional noises. By hearing assessment, the repaired results also sound very natural.

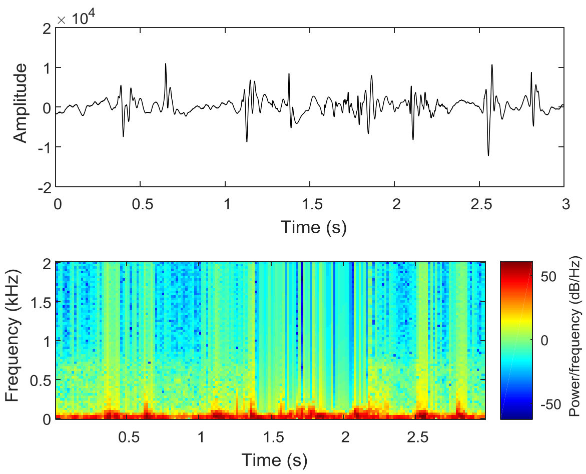

Figure 9. The repaired results addressing frictional noises.

V. CONCLUSION

The problem of automatic detecting and repairing clipping distortions and frictional noise in electronic stethoscope was addressed. For clipping distortions, they were detected and located by thresholding the differenced signal and the Hermite interpolation was used to repair the corresponding regions. For frictional noises, the interference regions were detected by employing MFCCs and SVM. EMD was used to decompose the original data into IMF components and the cardiopulmonary sound components embedded in IMFs were further extracted by using correlation coefficients, and then the auscultation data can be repaired by summing and interpolation. The data recorded by electronic stethoscope was used to verify the efficiency of the proposed method.

REFERENCES

[1] A. Akay, "Acoustics of friction," Journal of the Acoustical Society of America, 2002, 111(4):1525-48.

[2] P. Varady, "Wavelet-Based Adaptive Denoising of Phonocardiographic Records," International Conference of the IEEE Engineering in Medicine & Biology Society, IEEE, 2001.

[3] Z. Lu, Y. Xing-Hai, "The application of an improved wavelet threshold denoising method in heart sound signal," Cross Strait Quad-regional Radio Science & Wireless Technology Conference, IEEE, 2011.

[4] T. Omari, F. Bereksi-Reguig, "An automatic wavelet denoising scheme for heart sounds," International Journal of Wavelets, Multiresolution and Information Processing, 2015, 13(03):1550016.

[5] F. Liu, Y. Wang, "Research and Implementation of Heart Sound Denoising," Second International Asia Symposium on Intelligent Interaction & Affective Computing & Second International Conference on Innovation Management, 2010.

[6] F. Hedayioglu, M. G. Jafari, S. S. Mattos, M. D. Plumbley, M. T. Coimbra, "Denoising and segmentation of the second heart sound using matching pursuit," Proc. 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'12), IEEE, 2012.

[7] S. Ha, C. We, G. Jing, "An improved empirical mode decomposition-wavelet algorithm for phonocardiogram signal denoising and its application in the first and second heart sound extraction," International Conference on Biomedical Engineering & Informatics, 2014.

[8] A. Mondal, I. Saxena, H. Tang, P. Banerjee, "A noise reduction technique based on nonlinear kernel function for heart sound analysis," IEEE Journal of Biomedical & Health Informatics, 2018, 22(3):775-784.

[9] D. Emmanouilidou, E. D. Mccollum, D. E. Park, M. Elhilali, "Computerized Lung Sound Screening for Pediatric Auscultation in Noisy Field Environments," IEEE Transactions on Biomedical Engineering, 2018, 65(7): 1564-1574.

[10] Grundy, R. E, "The application of Hermite interpolation to the analysis of non-linear diffusive initial-boundary value problems," Ima Journal of Applied Mathematics, 2018, 70(6):814-838.

[11] M. A. Hossan, S. Memon, M. A. Gregory, "A novel approach for MFCC feature extraction," Signal Processing and Communication Systems (ICSPCS), 2010 4th International Conference on, IEEE, 2011.

[12] P. P. Dahake, K. Shaw, P. Malathi, "Speaker dependent speech emotion recognition using MFCC and Support Vector Machine," 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), IEEE, 2016.