*Journal of Biomedical Engineering and Technology, 2020, Vol. 8, No. 1, 14-24*
Available online at http://pubs.sciepub.com/jbet/8/1/3
Published by Science and Education Publishing
DOI:10.12691/jbet-8-1-3

# Spectrum Characteristics Analysis and Recognition of CHD Heart Sound in Five Auscultation Locations

**Sheng MIAO[1], Jian'e Dong[1,*], Jingyu Hou[2], Zhong Lihui[1]**

[1]Bigdata and Intelligence Engineering College of Southwest Forestry University, Kunming650224, China
[2]School of Information Technology, Deakin University, Victoria, Australia
*Corresponding author: wumingzhicao@126.com

*Received October 20, 2020; Revised November 21, 2020; Accepted November 30, 2020*

**Abstract** Cardiac Auscultation is widely used in the diagnosis of congenital heart disease (CHD) due to its non-invasive and cost-effective procedure. Heart sound analysis can provide effective auxiliary diagnosis information and aid in automatically screening patients. However, there are different signal spectrum characteristics in different auscultation locations, and there are no unified standards in the selection of auscultation locations during the heart sound analysis. This paper addresses the problem of auscultation locations and the selection of spectrum characteristics in the heart sound identification of CHD patients. 385 cases of normal and CHD heart sound signals were used to extract three groups of representative spectral features: Power Spectral Density (PSD), Mel-Frequency Cepstrum Coefficients (MFCCs) and Instantaneous Frequency Cumulative(IFC) from five different auscultation locations which are Aortic area(A), Pulmonic area(P), Tricuspid area(T), Mitral area(M), and Second aortic valve area(E). Significance detections based on p-value and Gaussian kernel support vector machine (SVM) were used to test these spectrum characteristics in five auscultation locations. The results show that the spectrum energy and the difference of auscultation areas concentrated within the frequency range of 20-150Hz. There is no statistical significance in terms of the spectrum characteristics of CHD in area A ($p>0.05$) compared with the statistical significance in other auscultation areas ($p<0.01$). The classification performance of the SVM method that uses spectrum characteristics of area E was the best overall. The experiment and analysis results show that, in CHD heart sound recognition, E-area signals have the best recognition effect, signals from TMP areas could be used as a reference, and A-area signals should be used with cautions. This discovery provides a practical guide to the clinical auscultation and signal processing of Congenital Heart Disease.

***Keywords:*** *Heart Sound (HS), Congenital Heart Disease (CHD), auscultation locations, spectrum analysis, signal processing*

**Cite This Article:** Sheng MIAO, Jian'e Dong, Jingyu Hou, and Zhong Lihui, "Spectrum Characteristics Analysis and Recognition of CHD Heart Sound in Five Auscultation Locations." *Journal of Biomedical Engineering and Technology*, vol. 8, no. 1 (2020): 14-24. doi: 10.12691/jbet-8-1-3.

## 1. Introduction

Heart sound is an important biomedical signal, which contains rich information about heart movement [1]. It also plays important roles in the diagnosis of coronary artery disease (CAD) [2]. Heart sound has a potential of being used for the diagnosis of CAD, and can provide effective auxiliary diagnosis information in the early diagnosis [3]. Compared with other diagnostic methods, Auscultation of heart sounds is a non-invasive, low-cost and effective method [4].

Among cardiovascular diseases, congenital heart disease causes more serious risks to adolescents due to its high incidence and mortality rates, especially in recent decades [5]. Congenital heart disease is very harmful because it is more likely to occur in young child population, especially in newborn babies [5,6]. The incidence rate of congenital heart disease is increasing in developing countries such as China [7]. Most developed countries in Europe and America have monitored and counted the patients with CHD [8], and a lot of studies have been done on the recognition of CHD patients by using heart sounds [4].

Heart sound is a quasi-periodic but non-stationary random signal. During the process of heart sound collection, it is easy to be interfered by noise. Therefore, it is a challenge to recognize and analyze it [9].

In the research of heart sound processing, researchers mainly focus on heart sound segmentation and classification. Segmentation is mainly to segment a heart sound into fundamental heart sounds: S1, S2, systolic and diastolic in a time domain. Segmentation is usually a step of heart sound preprocessing. Hidden Semi-Markov Model (HSMM) is the most popular segmentation method at present [10] with some improved algorithms based on it [11]. For statistical data, this method achieved higher accuracies in the identification of fundamental heart sounds S1 and S2 (94.54±5.15% and 93.96 ±5.01%

respectively) compared to the method that uses pathological sounds for identification in the segmentation process [9]. Nevertheless, it is still impossible to achieve a 100% accuracy in heart sound segmentation, especially for pathological sounds or noisy signals. However, heart sound segmentation is not a necessary step for heart sound recognition. Therefore many researchers focus on heart sound analysis without segmentation [12,13,14].

The main purpose of heart sound recognition is to distinguish the Pathological (abnormal) heart sounds from the collected sounds in order to diagnose patients with cardiovascular disease [3]. According to the literature review [9], statistical features, wavelet transform, Mel-frequency cepstral coefficients (MFCC), instantaneous frequency and phase based on Hilbert transform, spectrum characteristics (such as STFT, FFT, Power spectrum etc.) were the most extracted features for the detection of pathological sounds. For the methods that use classifiers, most studies classified pathological heart sounds using learning based approaches such as Artificial Neural Network (ANN) and Support Vector Machine (SVM). This is because other methods such as decision trees were not suitable for complex feature classification. Compared with the back-propagation artificial neural network (BP-ANN) and hidden Markov models (HMM), least square SVM can provide similar results but requires less training time, and it was found that, compared with other kernels, Gaussian Radial Basis Kernel Function (GRKF) produced the best results for SVM in classifying normal and abnormal signals.

Deep learning model (such as CNN) is also a research hotspot in recent years, but due to the limitation of heart sound data volume, this model also needs further development [12,13,15,16,17].

The development of heart sound detection technologies requires high-quality, rigorously validated, and standardized databases of heart sound signals obtained under various healthy and pathological conditions. However, the incompleteness and disunity of data records make it difficult to compare techniques and data sets [2,4].

An online survey via Twitter, Facebook and Reddit received 328 responses between 24 July and 22 August 2016, and the results showed that only about half of the respondents were able to correctly identify proper locations for heart sound collection. Therefore, studying the effectiveness of auscultation area selection is crucial to the identification of heart diseases [18].

When listening to and transducing heart sounds, physicians usually use five locations. These five locations are named according to the positions from which the valves can be best heard [2]:

• Aortic area(A)—centred at the second right intercostal space.

• Pulmonic area(P)—in the second intercostal space along the left sternal border.

• Tricuspid area(T)—in the fourth intercostal space along the left sternal edge.

• Mitral area(M)—at the cardiac apex, in the fifth intercostal space on the midclavicular line.

• The second auscultation location of aortic valve (E) — in the third and fourth intercostal space on the left border of sternum.

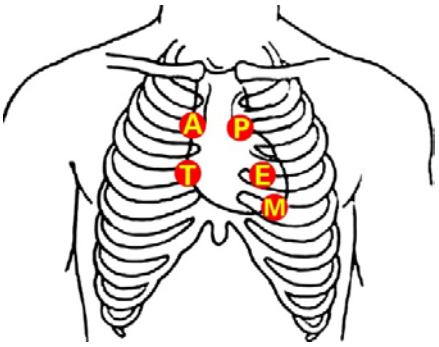


**Figure 1.** Five auscultation locations of heart sounds

Figure 1 shows these five auscultation locations. The heart sounds collected from different auscultation locations are different in time domain waves. Different auscultation locations have different effects on disease detection. However, many heart sound databases and literatures do not clearly describe the auscultation locations from which heart sounds were collected.

For example, the most commonly used heart sound database, 2016 PhysioNet/CinC Challenge Database, contains eight independent heart sound databases which store a total of 3153 heart sound recordings from 764 subjects/patients, lasting from 5s to over 120s [2]. These records, however, do not clearly indicate the location of heart sound collection in each case, which makes it impossible to analyze and compare the recognition effect of each location. At present, there are few existing literatures that clearly explain the effect of acquisition location of heart sound signals. Furthermore, there is no research on the comparison and analysis of heart sound acquisition locations [1,2,3,19,20,21].

This study focuses on the recognition of CHD heart sounds from different auscultation signals. We collected heart sounds of normal people and patients with CHD from five auscultation locations, extracted the spectrum characteristics of each auscultation location, and classified the characteristics of different auscultation locations using statistical methods and SVM. The results enable us to identify the auscultation location from which the collected heart sound signals are the best for distinguishing the characteristics of patients with CHD, as well as to provide a guide to the location selection for heart sound auscultation.

# 2. Materials and methods

## 2.1. Study Database

The database of this study consists of 1925 heart sound recordings from 385 de-identified child patients. These heart sound (HS) data were collected from 2017 to 2020 in Fuwai Yunnan Cardiovascular Hospital in Kunming, China. The frequency response of the stethoscope was 20-1000 Hz. The sample rate was 5000 Hz with 16 bit quantization. The recording time of each signal was 20s, and signals were saved in ‘blvm’ format.

Of the 385 patients, 174 were normal controls, 211 had murmurs that were related to CHD and were marked as abnormal or pathological heart sounds, 85 had aortic

disease and 126 had other miscellaneous conditions (tricuspid regurgitation, endocarditis, asymmetric septal hypertrophy). For each patient, five signals were collected from five auscultation locations. Each location is marked with an alphabetic character. The location order by which the heart sounds were recorded was the mitral valve area (M), pulmonary valve area (P), aortic valve area (A), second auscultation location (E), and tricuspid valve area (T), i.e., the order M-P-A-E-T, with patients in a supine position. All diagnoses were verified by at least one clinical expert through echocardiographic examination.

## 2.2. Spectrum Characteristics extraction

Heart sound is a kind of quasi-periodic random signal. The duration of a cardiac cycle in adolescents is about 0.5 to 0.8ms. Signals in cardiac cycles are similar but with some differences.

We divided the cardiac cycles of each heart sound to ensure the reliability and accuracy of the analysis. For each HS signal case (about 6s), we took ten cycles. We analyzed the spectral characteristics of each cycle in turn, and then averaged them. Three typical spectral features of heart sound were selected for analysis: Power spectrum, Mel-frequency cepstrum coefficient and instantaneous frequency. They are effective features for heart sound recognition, and have been proved effective in CVD heart sound recognition. In this study, they were selected to calculate the spectrum distribution of different auscultation locations. Details of these three features are as follows.

### 2.2.1. Power Spectrum

Heart sound has been proved a non-stationary random signal, so the fast Fourier transform (FFT) cannot be used to calculate the spectrum distribution directly. For random signals, power spectral density (PSD) can better reflect the frequency energy distribution. Power spectral has proved its effectiveness in heart sound recognition [22]. Actually each heart sound cycle can be mapped into one frame using R-R interval based ECG [23], and the signal length of each frame is consistent. Then hamming windows are used to segment heart sound signals. The hamming windows are expressed as:

$$w(n) = 0.5[1 - \cos(\frac{2\pi n}{M-1})] \quad 0 \le n \le M-1 \qquad (1)$$

where M is the length of window. In this study, the length of window is 1024 which ensures the signal frequency resolution.

Power spectrum of each intercepted data can be calculated by the autocorrelation function. Details are as follows.

Let $X(n) = S(n) * w(n)$ be signals intercepted through Hamming windows, $S(n)$ is the heart sound signal of a cardiac cycle, and the autocorrelation function of $X(n)$ can be expressed as:

$$R_x(m) = E[X(n)X(n+m)] \qquad (2)$$

The Power spectrum then can be calculated as:

$$S_x(k) = \sum_{m=0}^{N} R_x(m) e^{-\frac{2\pi}{N}km} \qquad (3)$$

where k is the frequency sampling point, $N$ is the length of the signal.

The power spectrums of all signals divided by Hamming windows are added as the signal of one cardiac cycle, and the power spectrums of ten cardiac cycles are averaged as the power spectrum density of one case.

### 2.2.2. Mel-Frequency Cepstrum Coefficient

For acoustic signals, human hearing has been proved not equally sensitive to all frequency bands showing a linear behavior in low frequencies and a logarithmical behavior in high frequencies.

Mel-Frequency Cepstrum Coefficients(MFCC) has a nonlinear spectrum which is widely used in PCG recognition [9,24].

For a given signal, we use window functions to frame it, and filter it by using the mel triangle filter. Then we convert the logarithmic spectrum into the Mel scale by using equation (4) (Wang et al 2007).

$$F_{mel}(\mathrm{f}) = 2595\log_{10}(1 + \mathrm{f}/700) \qquad (4)$$

where f is the frequency of signal and $F_{mel}$ is the MFCC. In this study, Hamming windows are used to frame the signal, and the number of triangle filters is set as 24, therefore, 24 groups of a heart sound period signal are generated through a triangle filter. 96 cepstrum coefficients are taken for each group of signals, so 24 * 96 features can be extracted for each heart sound.

### 2.2.3. Instantaneous frequency

Instantaneous frequency can represent frequency characteristics at each time point, therefore, it is used as an important feature in Cardiovascular heart sound recognition, and its validity has been proved [9]. Nevertheless, instantaneous frequency is only meaningful if the signal has a single spectral component, which is not the case for stethoscope signals. Therefore, the empirical mode decomposition (EMD) is applied to decompose the signal into simple components named intrinsic mode functions (IMFs)

The EMD algorithm decomposes the given signal as:

$$X(n) = \sum_{j=1}^{n} c_j + r_n \qquad (5)$$

The procedure depicted above was illustrated in the work of Huang et al [25]. For the IMF marked as $c_n$, it is transformed into a complex signal by Hilbert transform:

$$z(n) = c(n) + H[c(n)] = a(n)e^{j\varphi(n)} \qquad (6)$$

where $a(n)$ and $\varphi(n)$ are the amplitude and phase of the signal respectively, $j$ is the imaginary sign. $H(\bullet)$ is the Hilbert transform. Instantaneous frequency can be defined as the difference of instantaneous phase [26]:

$$\omega(n) = diff(\varphi(n)) \qquad (7)$$

where $\omega(n)$ is the instantaneous frequency. The above calculation steps show that the empirical mode decomposition (EMD) is applied to decompose the signal into intrinsic mode functions (IMFs), and each IMF can be calculated as one $\omega(n)$.

$\omega(n)$ is a time domain signal and can be represented in a two-dimensional coordinate system in which the horizontal axis is the time, and the vertical axis is the corresponding instantaneous frequency value. In order to convert the instantaneous frequency $\omega(n)$ into a time independent feature, we extract the frequency of the instantaneous frequency and call it Instantaneous Frequency Cumulant (IFC). By observing the distribution of instantaneous frequency cumulant values in the heart sound cycle, we can compare the difference in each auscultation location.

## 2.3. Analysis Methods and Tools

To show the difference of each spectrum feature from different auscultation locations, we use significance test and SVM classifier in this study to detect differences.

### 2.3.1. Significance Test

Statistical p-value test is used for significance test. P-value detection can reflect the difference between two groups of features [27]. We use it to detect the significant difference of different features between normal and Pathological heart sounds.

In this paper, we expect to test whether there are significant differences between the two groups of features extracted from normal and abnormal heart sounds. To this end, we set two groups of features as group A and group B, and test whether there is a significant difference between the two groups by A/B hypothesis significance test.

Take the sample mean $\mu$ as an example, $\mu_A$ and $\mu_B$ are the means of group A and group B respectively. We make the following two assumptions: $H_0$ assumes that there is no significant difference between the two groups, while $H_1$ assumes that there is a significant difference between the two groups. The hypotheses can be expressed as:

$$H_0 : \mu_A - \mu_B = 0$$

$$H_1 : \mu_A - \mu_B \neq 0$$

In order to ensure the validity of statistical results, it is necessary to analyze the relationships among the confidence interval, statistical power and minimum sample [28]. There are two types of errors in hypothesis testing.

Type I error: the original hypothesis is correct, but it is rejected, denoted by alpha (α).

Type II error: the original hypothesis is wrong, but it is not rejected, denoted by beta (β).

The value of statistic power is 1-β, and the value of confidence interval (CI) is 1-α.

The ratio between the sample sizes $n_A$ and of two groups A and B is defined as:

$$\kappa = \frac{n_A}{n_B} \tag{8}$$

The minimum number of samples size can be expressed as:

$$n_A = \kappa n_B \text{ and } n_B = \left(1+\frac{1}{\kappa}\right)\left(\sigma\frac{z_{1-\alpha/2}+z_{1-\beta}}{\mu_A-\mu_B}\right)^2 \tag{9}$$

where $z = \frac{\mu_A - \mu_B}{\sigma\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}}$, κ is the matching ratio, σ is the standard deviation, The Z value can be obtained by looking up the standard normal distribution table, For example, if $\alpha$=0.05, then $z_{1-\alpha/2}=z_{0.975}$, whose value is $z_{0.975}=1.96$ according to the normal distribution table.

Each spectrum feature in five auscultation locations is used to do the significance test. In order to ensure the validity of the statistical results, statistical power and confidence interval need to reach a high degree of credibility, then the minimum sample size can be estimated.

### 2.3.2. SVD Features Compression

The dimension of extracted spectrum features of each signal is high, which is not conducive to the recognition of subsequent classifiers. In order to compress the high-dimensional features and keep the effectiveness of the features, Singular Value Decomposition (SVD) is used in our study. Compared with Principal Component Analysis (PCA), SVD has a higher numerical stability and does not need the zero mean value. So it is an effective compression method [29]. SVD can be described as follow:

$$A = U\Sigma V^T \tag{10}$$

Where A is an $m \times n$ source feature matrix, $U$ is an $m \times m$ square matrix, and Σ is an $m \times n$ matrix, all elements of Σ except diagonal ones are 0, and the diagonal elements are called singular values, and $V$ is an $n \times n$ matrix named singular value vector.

The singular values of Σ are extracted as features, then the $m \times n$ dimensional feature matrix can be compressed as a $1 \times n$ dimensional matrix.

### 2.3.3. SVM Feature Classification

Most of the recent studies classified pathological heart sounds using learning based Support Vector Machine (SVM)), which is one of the widely used classifiers. In the early study, Gaussian Radial Basis Kernel Function (GRKF) produced the best results in classifying normal, aortic stenosis, pulmonary stenosis, tricuspid insufficiency and mitral insufficiency heart sounds compared to other kernel functions such as Linear Kernel Function (LKF), Polynomial Kernel Function (PKF) and so on. In this study, we select Gaussian Radial Basis Kernel for SVM, calculate AUC and draw ROC curves to compare the effect of spectrum feature classification in different auscultation locations.

At present, the discrimination parameters widely used for measuring the performance are as follows:

TP represents the true positive number, TN represents the true negative number, FN represents the false negative number, and FP represents the false positive number. Using these calculation parameters, some performance measures are defined as follows [2]:

$$recall = \frac{tp}{tp + fn} \tag{11}$$

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn} \tag{12}$$

$$Precision = \frac{tp}{tp + fp} \tag{13}$$

#### 2.3.4. Analysis Tool

Matlab (2017b) was used as an analysis tool to extract spectrum characteristics from heart sounds. In addition to some common functions and tools, the main functions used were: hamming() windows used to smooth signal, periodogram() used to calculate power spectrum, Melbankm() used to produce melfilter, and normfit() used to do Significance analysis. The EMD algorithm and instantaneous frequency calculation were implemented using the Matlab EMD tools package. SVM toolbox was used to generate classification results and ROC (Receiver Operating Characteristic) curves.

## 3. Results

Three groups of spectrum features were analyzed using 1925 heart sound recordings collected from 385 patients. These features were extracted from heart sounds obtained from five auscultation locations of patients. The heart sounds contained normal and pathological ones.

The results of each feature group were obtained as follows.

### 3.1. Spectrum Characteristics Distribution

Figure 2 shows the power spectrum distribution from 0 to 500Hz of the normal and pathological heart sounds in five auscultation locations. There are ten curves in total, and the black curves show the normal signals while the red curves show the abnormal signals. Figure 3 shows the spectrum difference distributions of normal and abnormal signals in each auscultation area. Five lines were used to represent the power spectrum signal difference distribution of normal heart sounds and Pathological heart sounds in five auscultation locations.

It can be seen from Figure 2 that the power spectrum of HS was mainly located within the range of 20-150Hz. In general, the power spectrum of Pathological heart sounds was higher than that of normal ones. Regarding different power spectrums of normal and abnormal signals shown in Figure 3, the difference of signals in each location was relatively large within the band range of 20-75hz, P area had the biggest gap within 100-200Hz band, the gap of area M and E was even bigger compared with the small gap of area A and T. However the power spectrum of abnormal signal in area T had the maximum value at 89Hz, 177Hz, 352Hz and 440Hz, which led to big gaps in these frequency points.

Figure 4 shows ten group MFCC distributions of the normal and pathological heart sounds in five auscultation locations. As introduced in the aforementioned methods, a HS signal passed 24 sets of triangle filters with 24 banks of MFCC being calculated. Since the power spectrum of HS was mainly located within the low frequency band and the pass-band frequency of triangle filter was from low to high, only the first ten banks of mel-coefficients were drawn for comparison and analysis. There were ten curves in total, and the black curves show the normal signals while the red curves show the abnormal signals. Figure 3b shows the mean MFCCs calculated from nine banks, and the normal and abnormal signals from five acquisition locations were compared.

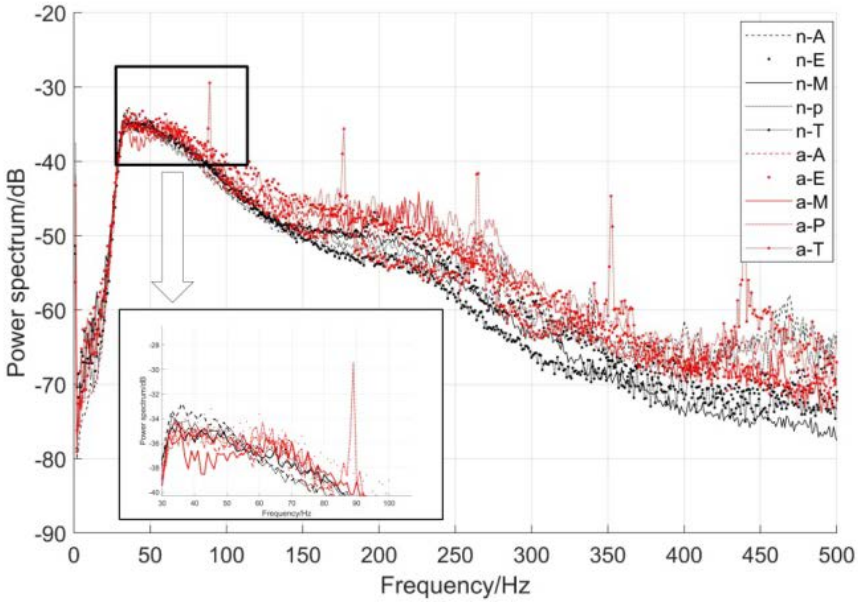


**Figure 2.** Power spectrum distribution of normal and Pathological heart sound signals in five auscultation locations (Curve naming used the format '*-#', where '*' represents either normal (n) or abnormal (a), and '#' represents the signal acquisition area (A to T))

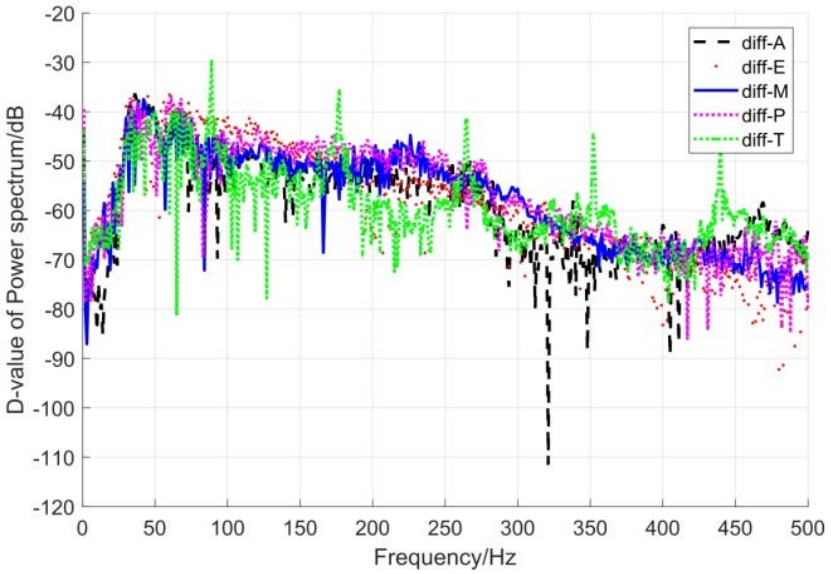


**Figure 3.** Spectrum difference distribution of normal and abnormal signals in five auscultation areas

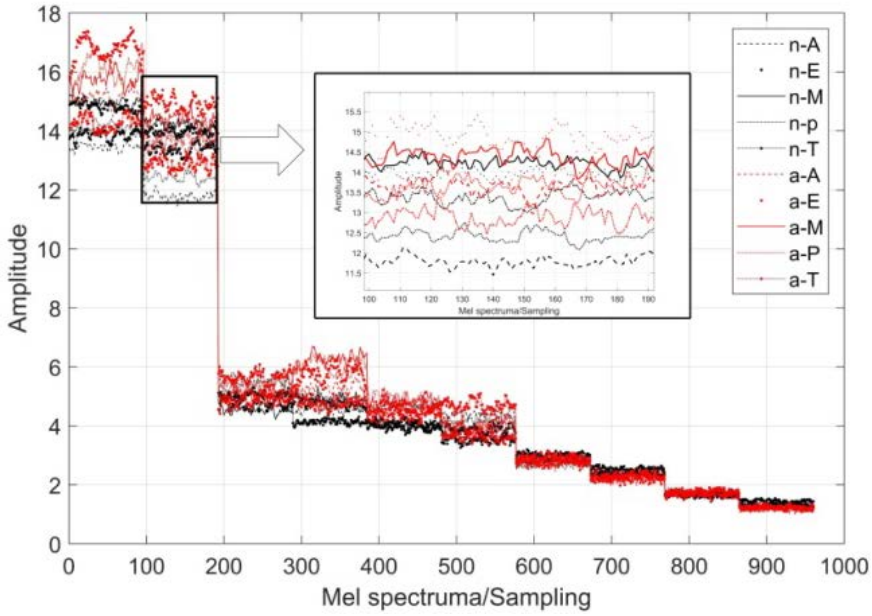


**Figure 4.** Ten banks of MFCC of normal and pathological heart sounds in each auscultation location (Curve naming used format 'diff-#', where '#' represents the signal acquisition area (A to T))

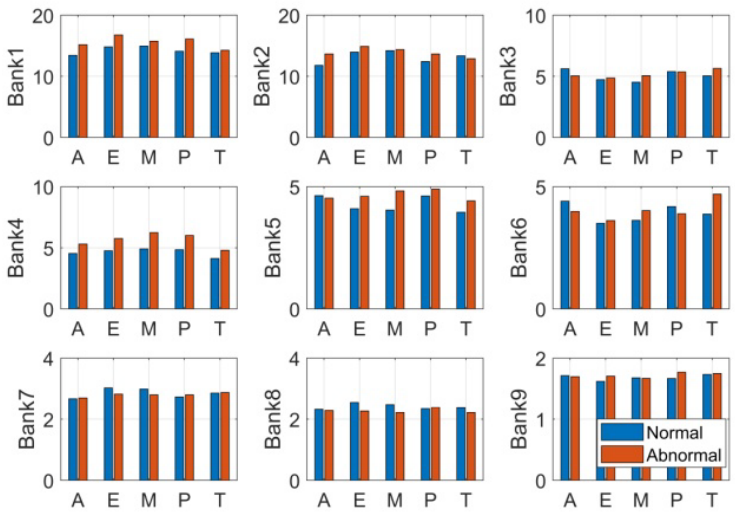


**Figure 5.** Comparison of mean MFCCs for the first nine banks ('A to T' represent signal acquisition location)

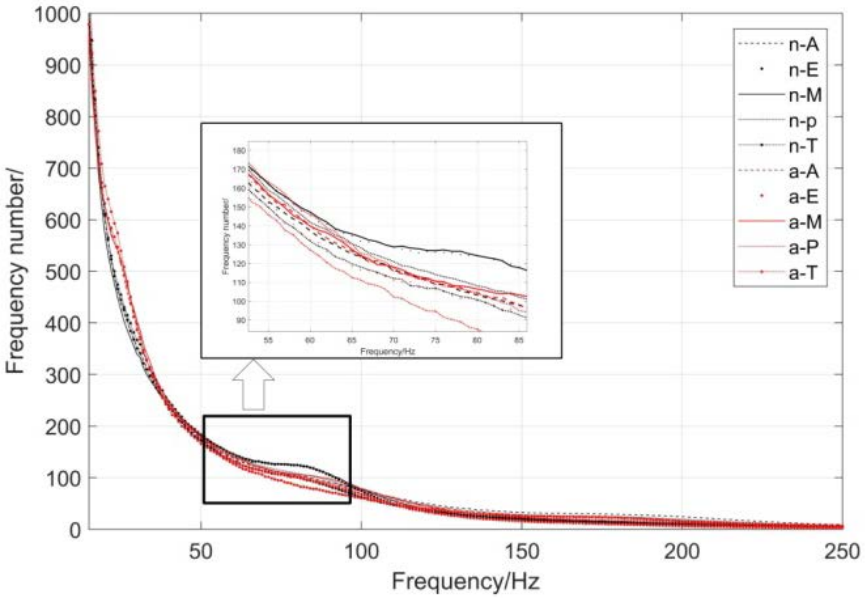


**Figure 6.** Power spectrum distribution of normal and Pathological heart sounds in each auscultation location (Curve naming used the format 'diff-#', where '#' represents the signal acquisition area (A to T))

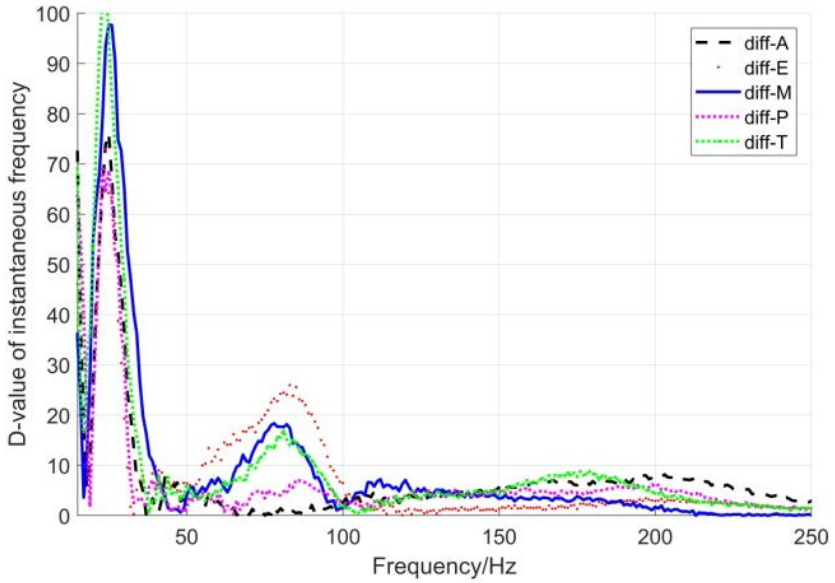


**Figure 7.** Power spectrum distribution of normal and pathological heart sounds in each auscultation location (Curve naming used the format 'diff-#', where '#' represents the signal acquisition area (A to T))

As depicted in Figure 4 and Figure 5, the amplitude of MFCC in a low frequency bank was big and decreased with the increase of frequency band. The difference of the normal and abnormal signals also had a big gap in a low frequency band. In the bank1and bank2, areas AEP had significant differences compared with areas MT. Areas EMP of bank4 and bank5 had bigger gaps than other areas. Area T had an obvious gap in bank6, while the differences of these areas were not obvious in bank3, bank7, 8 and 9.

Figure 6 shows the distribution of instantaneous frequency cumulant (IFC) on 25Hz-250Hz frequency band. Each point on the curve represents the count value of the instantaneous frequency in a heart sound period at that frequency point. There are ten curves in total, and the black curves show the normal signals while the red curves show the abnormal signals. Figure 7 shows the distributions of instantaneous frequency differences between normal and abnormal HSs.

Figure 6 and Figure 7 show that the Instantaneous frequency cumulative (IFC) values were mainly located within the 100Hz band, the IFC value and difference of frequency band over 100Hz were very small. There were two frequency bands with great differences for the IFC value distribution. They were 25-40Hz and 60-90Hz. For the first band, areas MT had a bigger gap, and area P had the biggest gap in the second band. In the high frequency part, the differences were very small (Less than ten counts).

## 3.2. P-value Test Results

In order to detect the significant difference between normal and abnormal signals in different auscultation locations, we used the statistical method to detect spectrum characteristics.

First of all, the demand of minimum sample size under different statistic power and confidence interval was

analyzed. The statistic power was set from 0.5 to 0.9, while the Type I error α was set as 0.05 with the confidence interval (CI) 0.95, $\kappa$=1, the variance and mean of each group were calculated by samples, and the group A minimum sample size for each feature was calculated according to expression (10), and results are shown in Table 1. According to the calculation results in Table 1, our data volume met the analysis requirement for power=0.8.

**Table 1. Calculation of the minimum size of each group's features under different power (CI=0.95)**

| Sample size | Power 0.5 | Power 0.6 | Power 0.7 | Power 0.8 | Power 0.9 |
|---|---|---|---|---|---|
| power spectrum | 48 | 62 | 77 | 98 | 132 |
| MFCC | 69 | 88 | 111 | 141 | 198 |
| IFC | 39 | 50 | 63 | 80 | 107 |

And then, we calculated the mean and variance (Sigma) of each spectrum feature and used P-value to test the mean value. Table 2, Table 3 and Table 4 respectively show the test results for power spectrum, Mel-Frequency Cepstrum Coefficient and instantaneous frequency cumulant for normal and abnormal HSs. It can be seen from the detection results of three types of spectrum features in five auscultation locations that there was no significant difference in three types of spectrum characteristics from auscultation area A with statistical significance ($p>0.05$) compared with other auscultation areas ($p<0.05$).

**Table 2. Variance (Sigma) of power spectrum in each auscultation location (CI=0.95), p-value in normal and pathological heart sounds**

| Auscultation location | Normal HS sigma(10e-4) | Abnormal HS sigma(10e-4) | p-value |
|---|---|---|---|
| A | [0.3898 0.4833] | [0.4282 0.6007] | 0.2653 |
| E | [0.4097 0.5077] | [0.0809 0.1134] | 3.69E-04 |
| M | [0.4371 0.5416] | [0.4743 0.6671] | 1.96E-07 |
| P | [0.4290 0.5320] | [0.6788 0.9523] | 4.27E-07 |
| T | [0.4309 0.5344] | [0.0774 0.1086] | 3.96E-10 |

**Table 3. Variance (Sigma) of MFCC in each auscultation location (CI=0.95), p-value in normal and pathological heart sounds**

| Auscultation location | Normal HS sigma(10e-4) | Abnormal HS sigma(10e-4) | p-value |
|---|---|---|---|
| A | [1.8521 2.2967] | [1.8166 2.5486] | 0.5588 |
| E | [1.5687 1.9440] | [1.6885 2.3689] | 0.0135 |
| M | [1.4811 1.8354] | [1.7420 2.4502] | 0.0039 |
| P | [1.8306 2.2700] | [1.9673 2.7600] | 0.0088 |
| T | [1.5643 1.9398] | [1.8896 2.6510] | 0.0013 |

**Table 4. Variance(Sigma) of instantaneous frequency cumulant in each auscultation location (CI=0.95), p-value in normal and pathological heart sounds**

| Auscultation location | Normal HS sigma | Abnormal HS sigma | p-value |
|---|---|---|---|
| A | [28.8823 35.7011] | [30.0264 42.1253] | 0.2908 |
| E | [26.1018 32.2843] | [28.4766 39.9510] | 0.005 |
| M | [24.8606 30.7491] | [33.6204 47.2880] | 0.0024 |
| P | [27.4918 33.9824] | [32.6749 45.8410] | 0.036 |
| T | [26.5955 32.8744] | [32.3978 45.4522] | <0.001 |

## 3.3. SVM Classification Results

To show the recognition effect of heart sound spectrum features in each auscultation location of Pathological heart sounds, SVM classifier was used to classify these features. In order to reduce the calculation time of SVM, we used SVD to compress the features. The compression results of each group of features were shown in Table 5.

**Table 5. features compression results using SVD**

| Feature | Source dimension | Compression dimension |
|---|---|---|
| Power spectrum | 1024*10 | 1*10 |
| Mel coefficient | 96*24 | 1*24 |
| Instantaneous frequency | 1024*10 | 1*10 |

Compressed features from each auscultation location were classified by SVM. The receiver operating characteristic (ROC) curve was drawn to compare the classification accuracy of each auscultation location. Figure 8 shows the power spectrum recognition results, Figure 9 shows the MFCC recognition results and Figure 10 shows the IFC recognition results.

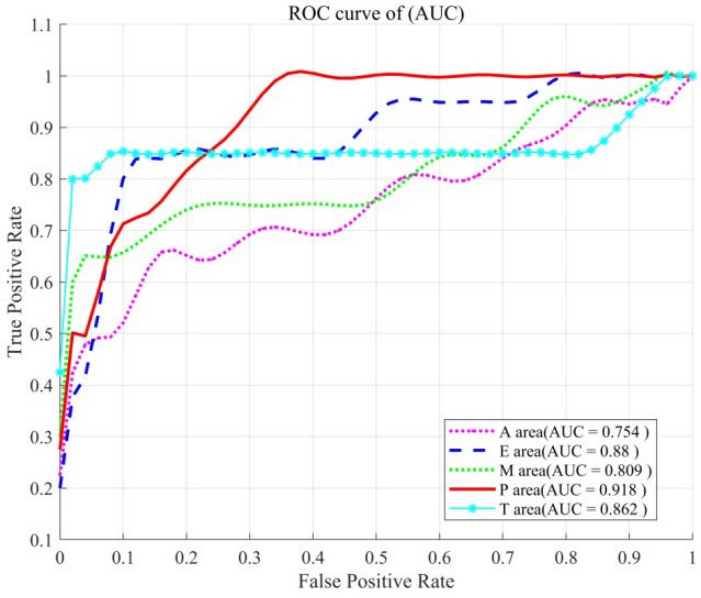


**Figure 8.** ROC for power spectrum recognition score in each auscultation location (AUC=Area Under Curve)

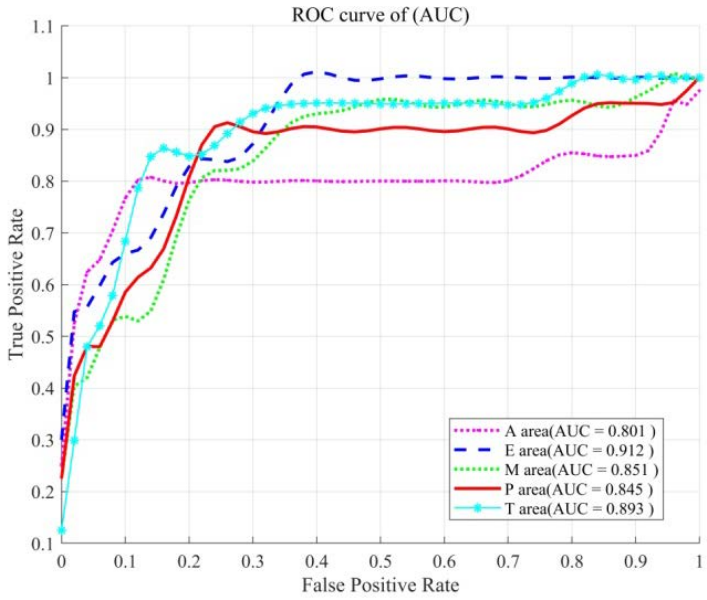


**Figure 9.** ROC for MFCC recognition score in each auscultation location

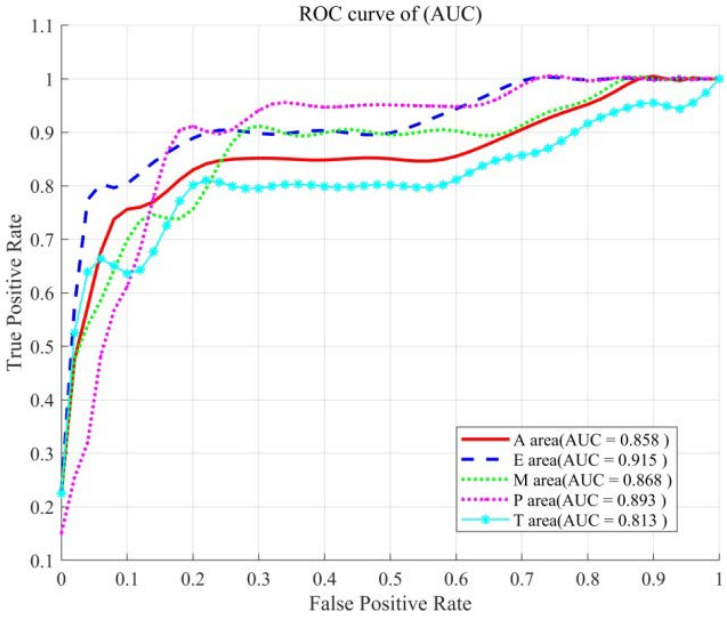


**Figure 10.** ROC for Instantaneous frequency cumulant (IFC) recognition score in each auscultation location

**Table 5. Results of power spectrum comparison within each auscultation location (Se= sensitivity, Sp= specificity, Acc=accuracy)**

| Auscultation location | Se | Sp | Acc |
|---|---|---|---|
| A | 0.843 | 0.665 | 0.754 |
| E | 0.877 | 0.883 | 0.88 |
| M | 0.836 | 0.782 | 0.809 |
| P | 0.967 | 0.869 | 0.918 |
| T | 0.897 | 0.827 | 0.862 |

Se=Sensitivity, Sp=Specificity, Acc= Accuracy.

**Table 6. Results of mel-coefficient comparison within each auscultation location (Se= sensitivity, Sp= specificity, Acc=accuracy)**

| Auscultation location | Se | Sp | Acc |
|---|---|---|---|
| A | 0.843 | 0.759 | 0.801 |
| E | 0.914 | 0.883 | 0.912 |
| M | 0.787 | 0.915 | 0.851 |
| P | 0.871 | 0.819 | 0.845 |
| T | 0.884 | 0.903 | 0.893 |

**Table 7. Results of Instantaneous frequency cumulative within each auscultation location (Se= sensitivity, Sp= specificity, Acc=accuracy)**

| Auscultation location | Se | Sp | Acc |
|---|---|---|---|
| A | 0.844 | 0.872 | 0.858 |
| E | 0.947 | 0.883 | 0.915 |
| M | 0.943 | 0.793 | 0.868 |
| P | 0.971 | 0.814 | 0.893 |
| T | 0.784 | 0.827 | 0.813 |

Table 5 - Table 7 show the classification results of the three features in each auscultation location. It can be seen from Table 5 that for the power spectrum, areas E and P got the highest score, for the MFCC, areas E and T got the highest score, and for the Instantaneous Frequency Cumulant, areas E and P got the highest score. Area A had the lowest score and the overall score of area M is lower from Table 6 and Table 7.

## 4. Discussion

Heart sound auscultation is a non-invasive detection method, but the auscultation position of heart sound is different, as well as the corresponding auscultation effect. Choosing the auscultation position correctly is helpful to detecting congenital heart disease quickly, which is also an embodiment of patient-oriented [30]. There are two advantages for clinical diagnosis to study the signal characteristics of different auscultation positions of heart sounds. One is to provide reference information to doctors helping them quickly select and locate auscultation points. Another is to provide a guide to heart sound analysis for different auscultation locations, so as to improve the success rate of abnormal heart sound recognition.

Our experimental results and analyses showed that the signal spectrum characteristics of different auscultation locations were different. The recognition success rate could be effectively improved by selecting signals from different locations.

The acquisition of heart sound data for patients with congenital heart disease is challenging, because of the sporadic occurrence of cases and a long time of data collection. In order to ensure the consistency of the data, patients need to keep the same body position and the position of acquisition point should be as correct as possible.

Due to the above reasons, it is difficult to collect HS data in a large-scale. At present, hundreds of signals we used were only a small number of sample signals. Through traditional analysis, we found that the signals in each location followed some rules when they were used for identifying congenital heart disease. Nevertheless, in order to improve the validity and correctness of the analysis, more data and classification analyses are needed.

We analyzed the minimum number of samples required for each group of features under different power settings. The results showed that our sample size ensured the statistical power was not less than 0.8 with the IC=0.95.

In terms of power spectrum distribution, signal energy is mainly concentrated within the 20-150HZ frequency band, which is consistent with the results of literature analysis [3]. For the difference of power spectrum, the difference of signals in each region was not very big, and the power spectrum feature recognition rate in region P was the highest, reaching 91%.

Regarding the MFCC characteristics, the low-frequency part was quite different, and area E had the highest classification correctness rate, which was up to 90%.

From the point of view of IFC, the instantaneous frequency was mainly concentrated within the frequency band 100Hz, and the difference of each region was mainly concentrated within this frequency band. The difference ofthe regions M and T was similar, but the difference of region P was the biggest within 50-100hz. Area E had the highest classification correctness rate.

Based on the analysis of the three spectrum characteristics of signals in each area, we observed that signals in area A had no significant differences between normal and abnormal signals in P-value detection, but had significant differences in other four areas. The classification effect of signals in area A was poor with the three spectrum features, so signals in area A are not recommended for the recognition of heart sound signals of CHD. The classification effect of signals in the areas of M and P was generally good, which was about 85%, so they could be used as auxiliary reference signals. Generally, the classification success rate of signals in the areas of E and T was relatively high, especially for the signals in area E. In conclusion, the recognition of CHD could firstly consider the signals collected from area E, followed by area T and areas M and P.

Because the number of signals and cases was limited in this research, the experimental results could be used as a reference in clinical applications, and more cases are needed to verify the effectiveness of our method.

In the follow-up work, we will consider the cooperative recognition that uses signals from multi locations, for example, selecting signals from more than two auscultation locations for joint recognition. Another work is to improve the recognition success rate. The main problems are how to select and combine the signals of each location, and how to make the feature fusion more effective.

## 5. Conclusion

The analysis results show that in CHD heart sound recognition, E-area signals have the best recognition effect, signals from the areas T, M and P can be used as references, and A-area signals should be used with cautions.

## Acknowledgements

This study is supported by National Natural Science Foundation of China (61941204), Major science and technology projects of Yunnan Province (2018ZF017), Joint agricultural project of Yunnan Provincial Department of Science and Technology (2017FG001(-074), 2018FG001(-108), 2018FG001-(101), 2017FG001(-034), 2018FG001(-059)). We also would like to thank Yunnan Fuwai Hospital and Yunnan University for providing data collection support.

## References

[1] Azra, Saeidi, Farshad, Almasganj. 3D heart sound source localization via combinational subspace methods for long-term heart monitoring. Biomedical Signal Processing & Control. 2017; 31: 434-43.

[2] Liu C, Springer D, Li Q, Moody B, Juan RA, Chorro FJ, et al. An open access database for the evaluation of heart sound algorithms. Physiological Measurement. 2016; 37(12): 2181-213.

[3] Schmidt SE, Holst-Hansen C, Hansen J, Toft E, Struijk JJ. Acoustic Features for the Identification of Coronary Artery Disease. IEEE Transactions on Biomedical Engineering. 2015; 62(11): 2611-9.

[4] Clifford GD, Liu C, Moody BE, Roig JM, Mark RG. Recent advances in heart sound analysis. Physiological Measurement. 2017; 38(8):E10.

[5] E C, FH B, SA C, MWM dL, L S, TL G. Newborn pulse oximetry screening in the context of a high antenatal detection rate for critical congenital heart disease. Acta Paediatrica. 2019.

[6] Vlachos A, Osorio DS, Atsidaftos E, Kang J, Lababidi ML, Seiden HS, et al. Increased Prevalence of Congenital Heart Disease in Children With Diamond Blackfan Anemia Suggests Unrecognized Diamond Blackfan Anemia as a Cause of Congenital Heart

Disease in the General Population: A Report of the Diamond Blackfan Anemia Registry. circulation genomic & precision medicine. 2018; 11(5): e002044.

[7] Li Q, Zhu W, Zhang B, Wu Y, Yuan Y, Zhang H, et al. The MALAT1 gene polymorphism and its relationship with the onset of congenital heart disease in Chinese. Bioscience Reports. 2018; 38(3).

[8] Fredriksen P, Ingjer E, Thaulow E. Physical activity in children and adolescents with congenital heart disease. Aspects of measurements with an activity monitor. Cardiology in the Young. 2000; 10(2): 98-106.

[9] Lau I, Sun Z. Three-dimensional printing in congenital heart disease: A systematic review. Journal of Medical Radiation Sciences. 2018.

[10] Springer DB, Tarassenko L, Clifford GD. Logistic Regression-HSMM-Based Heart Sound Segmentation. IEEE Transactions on Biomedical Engineering. 2016; 63(4): 822-32.

[11] Tsao Y, Lin T-H, Chen F, Chang Y-F, Cheng C-H, Tsai K-H. Robust S1 and S2 heart sound recognition based on spectral restoration and multi-style training. Biomedical Signal Processing and Control. 2019; 49: 173-80.

[12] Thiyagaraja SR, Ram D, Shrestha PL, Anurag C, Thompson MA, Anumandla PT, et al. A novel heart-mobile interface for detection and classification of heart sounds. Biomedical Signal Processing & Control. 2018; 45: 313-24.

[13] Fahad HM, Ghani Khan MU, Saba T, Rehman A, Iqbal S, Diaspro A. Microscopic abnormality classification of cardiac murmurs using ANFIS and HMM. Microscopy Research & Technique. 2018.

[14] Langley P, Murray A. Heart sound classification from unsegmented phonocardiograms. Physiological Measurement. 2017; 38(8).

[15] Liu Y, Guo X, Zheng Y. An Automatic Approach Using ELM Classifier for HFpEF Identification Based on Heart Sound Characteristics. Journal of Medical Systems. 2019.

[16] Zhang W, Han J, Deng S. Heart sound classification based on scaled spectrogram and tensor decomposition. Expert Systems with Applications. 2017; 84: 220-31.

[17] Whitaker BM, Suresha PB, Liu C, Clifford G, Anderson D. Combining sparse coding and time-domain features for heart sound classification. Physiological Measurement. 2017.

[18] Mallinson. A survey into paramedic accuracy in identifying the correct anatomic locations for cardiac auscultation. British Paramedic Journal. 2017; 2(2): 5.

[19] Kang S, Doroshow R, Mcconnaughey J, Shekhar R. Automated Identification of Innocent Still's Murmur in Children. IEEE Transactions on Biomedical Engineering. 2017; 64(6): 1326-34.

[20] Dominguez-Morales JP, Jimenez-Fernandez AF, Dominguez-Morales MJ, Jimenez-Moreno G. Deep Neural Networks for the Recognition and Classification of Heart Murmurs Using Neuromorphic Auditory Sensors. IEEE Transactions on Biomedical Circuits & Systems. 2017: 1-11.

[21] Kouras N, Benidir M, Boutana D, editors. The Empirical Mode Decomposition and the SVD for Abnormal Heart Sound Signals detection and Time-Frequency Analysis. Iet International Conference on Intelligent Signal Processing 2016: IET.

[22] Teo SK, Bo Y, Ling F, Yi S, editors. Power Spectrum Analysis for Classification of Heart Sound Recording. Computing in Cardiology Conference; 2016.

[23] Jurak. Heart sounds analysis using probability assessment. Physiological Measurement. 2017; vol. 38, no. 8, :pp. 1685-700.

[24] Chen W, Zhao Y, Lei S, Zhao Z, Pan M. [Study of biometric identification of heart sound base on Mel-Frequency cepstrum coefficient]. Journal of Biomedical Engineering. 2012; 29(6): 1015-20.

[25] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings Mathematical Physical & Engineering Sciences. 1998; 454(1971): 903-95.

[26] Hu H, Jing K, Guan L, editors. Instantaneous frequency estimation based on empirical mode decomposition2008.

[27] Rice WR. A Consensus Combined P-Value Test and the Family-Wide Significance of Component Tests. Biometrics. 1990; 46(2): 303-8.

[28] Ian, White. Sample Size Calculations in Clinical Research. Journal of the Royal Statistical Society. 2008.

[29] McGivney DF, Pierre E, Ma D, Jiang Y, Saybasili H, Gulani V, et al. SVD Compression for Magnetic Resonance Fingerprinting in the Time Domain. IEEE Transactions on Medical Imaging. 2014; 33(12): 2311-22.

[30] Pretorius E, Cronje ML, Strydom O, editors. Development of a pediatric cardiac computer aided auscultation decision support system. International Conference of the IEEE Engineering in Medicine & Biology; 2010.