

# Abnormal heart sounds detection based on the scaled time-frequency representation and feature selection

Wenjie Zhang<sup>1</sup>, Jiqing Han<sup>1</sup>, Shiwen Deng<sup>2</sup>

<sup>1</sup>School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

<sup>2</sup>School of Mathematical Sciences, Harbin Normal University, Harbin, China

## Abstract

*Abnormal heart sounds detection is of great value for the pre-diagnosis of heart disease. Since the pathological information is usually contained in each heart cycle, it is essential to segment the heart sounds before further analysis. In order to extract more detailed information, the time and frequency domain analysis of the heart cycle is necessary. So the time-frequency representation (TFR) of each heart cycle is extracted in this study. However, the durations of the heart cycles between different samples are usually not the same. As a result, the sizes of the corresponding TFR are different which prohibits the direct comparison between them. To solve this problem, the TFRs are scaled to a fixed size through the bilinear interpolation method. Nevertheless, the scaled TFR contains some noises which are useless for classification. For the purpose of removing noises, a feature selection method based on similarity is applied. Then, the selected features are input to a support vector machine (SVM) for classification. At last, the proposed method is evaluated on the dataset offered by the PhysioNet/Computing in Cardiology Challenge 2016. The overall score 85.40% is achieved by a 5-fold cross-validation. It is the average of the sensitivity(75.03%) and specificity(95.76%). The best overall performance of our method on the challenge is 84%.*

## 1. Introduction

Cardiovascular diseases have long been one of the main causes of human death worldwide [1]. However, the patients can take effective measures to cure the diseases before it is too late. So it is of great value to detect the cardiovascular diseases in the early stage. The cardiovascular diseases usually causes the abnormal blood flow between the atria and ventricles. As a result, the pathological information is reflected on the heart sounds. Thus it is possible to detect the heart diseases through detecting the abnormal heart sounds. Auscultation is one of the commonly used method to detect the abnormal heart sounds in clin-

ical practice [2]. Nevertheless, auscultation requires the subjective experience of the professional physicians. Besides, the patients also must be present. To solve these problems, the automated analysis of the heart sounds has been popular in recent years.

The automated analysis of heart sounds usually contains three steps: segmentation, feature extraction and classification. The aim of segmentation is to detect the fundamental heart sounds (FHS) which are the important physical characteristics of the hearts, including S1 and S2. The FHSs provide references for feature alignment in the feature extraction step by dividing the heart sounds signals into heart cycles. In order to focus on the feature extraction and classification, the state-of-the-art segmentation method, i.e., the hidden semi-Markov model (HSMM) based segmentation method [3], is used in this study.

In the feature extraction stage, it is intended to extract the features which have a positive correlation with the pathological changes of the non-stationary heart sounds signals. Since the TFR of heart sounds signals has been successfully proposed before to investigate the dynamic properties reflecting the physiological or pathological episodes [4,5], the TFR of each heart cycle is extracted in this study. In the literature, the main methods of the TFR extraction include parametric and non-parametric approaches. More specifically, the parametric approaches are mainly based on the time-dependent autoregressive (AR) modeling and their extensions [6]. And the parametric approaches are commonly implemented by means of short time Fourier transform (STFT) [7], S-transform (ST) [8], discrete wavelet transform (DWT) [9] and Choi-Williams distribution (CWD) [10]. Among the methods, the STFT has a fixed resolution in the frequency bands whereas ST and WT have multi-resolution. CWD and AR have better resolution in the frequency domain. However, better resolution does not necessarily results in a better representation. For comparison, the five TFR methods are implemented in this study. After the extraction of TFR, the features are difficult to compare directly since the sizes of the TFR are not the same which are caused by the differ-

ent lengths of heart cycles. In order to align the TFR and maintain its structure at the same time, the bilinear interpolation which has the good ability to keep the structure is employed. Thus the extracted TFRs are scaled to the same size. Nevertheless, the scaled TFR contains some noises which are useless for classification. For the purpose of removing noises, a feature selection method based on similarity is applied [11]. At last, the classification task is performed by the SVM using a Gaussian kernel.

In this study, we investigated the scaled time-frequency representation and feature selection for abnormal heart sounds detection, as shown in Figure 1. As the process of detecting the dynamic properties of the heart sounds, five kinds of TFRs are employed. In order to align the TFRs, the bilinear interpolation is applied. The noises of the scaled TFRs are eliminated by a feature selection method. The results are compared using the sensitivity, specificity and the overall score.

## 2. Methods

### 2.1. Data collection

The dataset is offered by the PhysioNet/Computing in Cardiology Challenge 2016 [12]. It is contributed by seven different research teams. In the dataset, there are 3420 recordings, including 2755 normal heart sounds recordings and 665 abnormal heart sounds recordings. The heart sounds signals are all resampled to 2000 Hz.

### 2.2. Preprocessing and segmentation

Since the information of the heart sound signals are mainly concentrated above 50 Hz, the signals are first filtered by high pass Order 6 Butterworth filter. Then the signals are segmented using the state-of-the-art HSMM based method. After segmentation, some researchers analysis the signals according to the S1, systolic, S2 and diastolic periods in each heart cycle. However, some abnormal heart sounds appears in the FHSs and continue to present in the systolic and diastolic, it is difficult to analysis separately. Thus the heart cycles are the basic analysis units in this study.

### 2.3. Feature extraction

In order to extract the TFRs of the heart cycles, both the parametric and non-parametric approaches are employed. For the parametric AR model of  $p$ th order, it is described as follows,

$$x[t] = \sum_{i=1}^p a[i]x[t-i] + e[t]; e[t] \sim N(0, \sigma_e^2[t]) \quad (1)$$

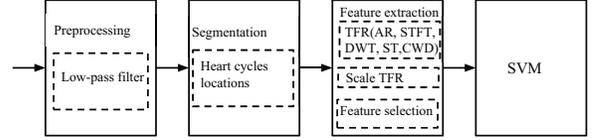


Figure 1. The framework of this study.

where  $a[i]$  is the linear prediction coefficients and  $e[t]$  is the prediction error. Equation (1) can be effectively solved by the Burg's method. The spectral of  $x[t]$  is then estimated by

$$S_{AR}[\tau, f] = \frac{\sigma_e^2[\tau]}{|1 - \sum_{i=1}^p a[i]e^{-j2\pi fi}|^2} \quad (2)$$

For the non-parametric model STFT, it has a fixed resolution and the spectral is obtained by

$$S_{STFT}[\tau, f] = \sum_{t=1}^N x[t]w[t-\tau]e^{-j2\pi ft} \quad (3)$$

where  $w[t-\tau]$  is the window function.

To solve the fixed resolution problem, the ST and DWT is proposed. The ST is defined as,

$$S_{ST}(\tau, f) = \int_{-\infty}^{+\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-j2\pi ft} dt \quad (4)$$

where  $h(t)$  is the heart sounds signal. The window function  $e^{-\frac{(\tau-t)^2 f^2}{2}}$  is a function of the frequency  $f$  which results in the high time resolution in high frequency and high frequency resolution in low frequency. The DWT is mainly achieved by the mother wavelet,

$$S_{DWT}(\tau, f) = \int_{-\infty}^{+\infty} h(t) \frac{1}{\sqrt{2f}} \psi\left(\frac{t-\tau 2^f}{2f}\right) dt \quad (5)$$

where  $\frac{1}{\sqrt{2f}} \psi\left(\frac{t-\tau 2^f}{2f}\right)$  is the mother wavelet. However, the DWT only decompose the signal in low frequency, some important information in high frequency is omitted. To solve this problem, the wavelet packet transformation (WPT) is developed.

In order to have a better resolution, the quadratic time-frequency distributions (QTFD) is introduced. CWD is one of the QTFDs which aims to reduce undesirable cross terms without degrading the resolution and it is defined as,

$$S_{CWD}(\tau, f) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{\nu^2/\sigma}} h\left(t - \frac{1}{2}\nu\right) h\left(t + \frac{1}{2}\nu\right) \cdot e^{-(t-\tau)^2/(4\nu^2/\sigma)} e^{-j2\pi f\nu} dt d\nu \quad (6)$$

Thus the TFRs  $\mathcal{S}(\tau, f) \in \mathbb{R}^{T \times F}$  of each heart cycle is obtained. However, the sizes of different TFRs between different subjects are usually not the same. To overcome this problem, the bilinear interpolation is employed. Suppose that the TFRs after interpolation are  $\mathcal{S}' \in \mathbb{R}^{T' \times F'}$ . The main steps are as follows:

1. Set  $s_T = T/T'$  and  $s_F = F/F'$ .
2. Suppose that  $\mathcal{S}'(\tau', f')$  is one element in  $\mathcal{S}'$ . Set  $\mu = \tau' \cdot s_T$  and  $\xi = f' \cdot s_F$ . Set  $\Delta\mu = \mu - \lfloor \mu \rfloor$  and  $\Delta\xi = \xi - \lfloor \xi \rfloor$ . Round down the value of  $\mu$  and  $\xi$  as  $\mu = \lfloor \mu \rfloor$  and  $\xi = \lfloor \xi \rfloor$ .
3. Then, the element  $\mathcal{S}'(\tau', f')$  is represented as

$$\begin{aligned} \mathcal{S}'(\tau', f') &= \mathcal{S}(\mu, \xi) \cdot (1 - \Delta\mu) \cdot (1 - \Delta\xi) \\ &\quad + \mathcal{S}(\mu + 1, \xi) \cdot \Delta\mu \cdot (1 - \Delta\xi) \\ &\quad + \mathcal{S}(\mu, \xi + 1) \cdot (1 - \Delta\mu) \cdot \Delta\xi \\ &\quad + \mathcal{S}(\mu + 1, \xi + 1) \cdot \Delta\mu \cdot \Delta\xi \end{aligned}$$

4. Iterate step 2 and step 3 until all the elements in  $\mathcal{S}'$  are obtained.

Take the sizes of the  $\mathcal{S}'$  between different heart cycles to be equal. Thus, the sizes of the TFRs are scaled to be equal through the bilinear interpolation method.

## 2.4. Feature selection and classification

The scaled TFRs are contaminated with noises and some features are useless for classification. To exclude these features, a feature selection method is proposed according to the ReliefF criterion. The features of the scaled TFRs are scored by,

$$\begin{aligned} R_{score}(\mathbf{x}_r) &= \sum_{m=1}^M \left( \sum_{s=1}^k \frac{1}{k} (\mathbf{X}(m, r) - \mathbf{X}(NM(m)_s))^2 \right. \\ &\quad \left. - (\mathbf{X}(m, r) - \mathbf{X}(NH(m)_s))^2 \right) \end{aligned} \quad (7)$$

where  $\mathbf{x}_r$  is the  $r$ th feature of the scaled TFRs,  $\mathbf{X} \in \mathbb{R}^{M \times R}$  is the data matrix,  $NM(m)_s$  and  $NH(m)_s$  indicate the  $s$ th nearest data instances to  $\mathbf{X}(m, r)$  with the same class label and different class, respectively. After scoring, the features with higher score are considered more discriminative. Thus the features with lower scores are excluded by a threshold value.

SVM is a supervised machine learning model which aims to find the maximum-margin hyperplanes between different data in order to discriminate them. It has been proven a powerful classifier and is used in heart sounds classification successfully in the literature. So the selected features are trained and evaluated by SVM in this study.

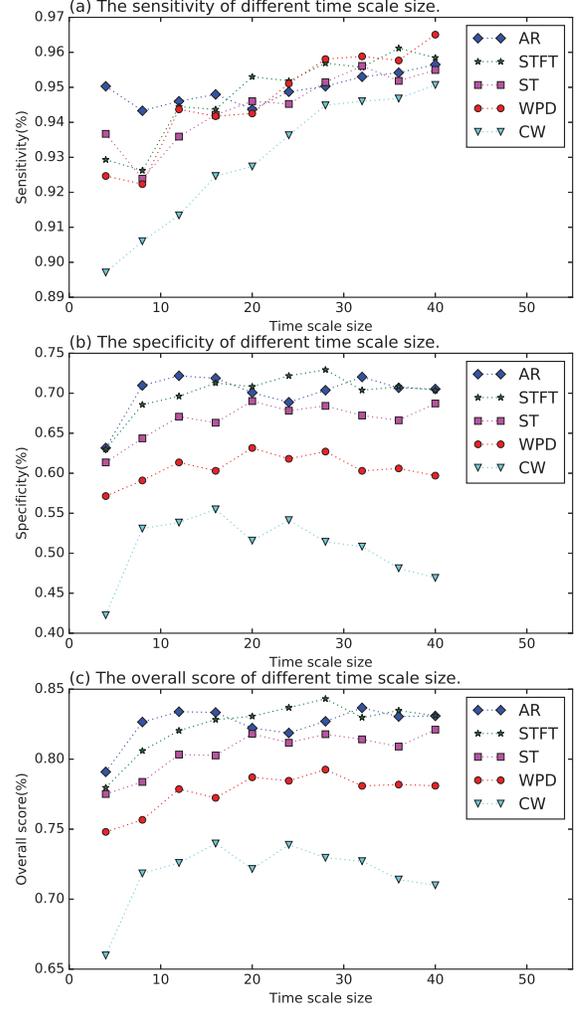


Figure 2. The performance of different time scale sizes.

## 3. Results and discussion

The dataset is evaluated by a 5 fold cross-validation method using the SVM classifier. The SVM is implemented with a radial basis function (RBF) kernel and the parameter of the RBF kernel is 0.1. Different lengths of the scale size are evaluated in the experiment. The different time scale sizes are experimented with a fixed frequency scale size 28, as shown in Figure 2. It can be seen that the AR and STFT method has the better overall performance for detecting the abnormal heart sounds. The best overall score 84.31% is achieved with a time scale size 28 by STFT. Also, the frequency scale sizes are evaluated, as shown in Figure 3. Still the AR and STFT method has the better overall performance. The best overall score 84.33% is achieved with a time frequency size 32 by STFT.

At last, the feature selection is performed. And the performance is improved, as shown in table 1.

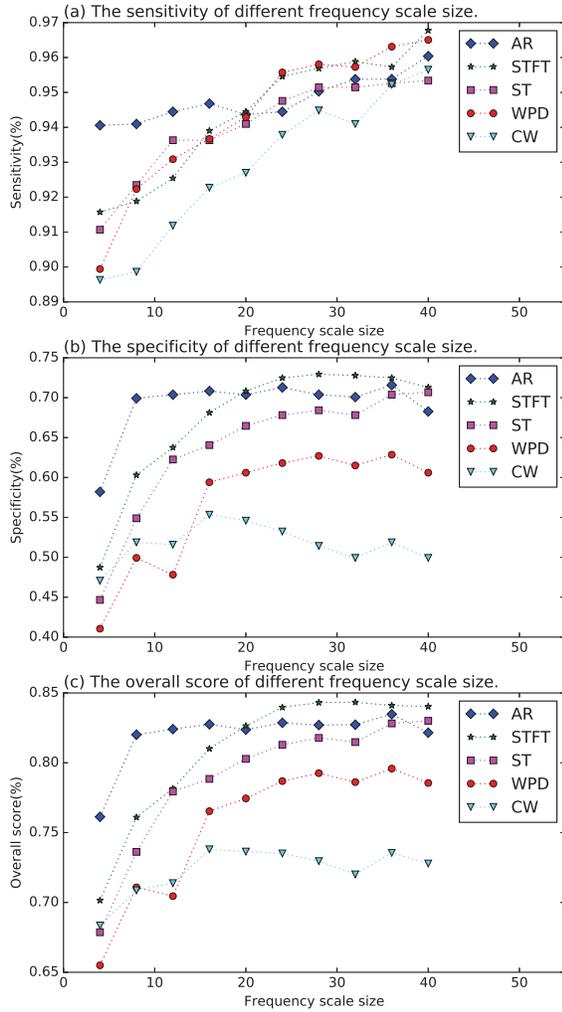


Figure 3. The performance of different frequency scale sizes.

	Sens	Spec	Overall	Scale	Select
AR	69.92	94.64	82.28	$36 \times 32$	461
STFT	<b>75.03</b>	<b>95.76</b>	<b>85.40</b>	$28 \times 28$	314
ST	72.78	95.02	83.90	$36 \times 28$	403
WPD	65.26	93.98	79.62	$28 \times 28$	314
CW	63.15	87.22	75.19	$16 \times 16$	102

Table 1. The best of performance different methods.

#### 4. Conclusion

In this study, a abnormal heart sounds detection method based on scaled TFR and feature selection is proposed. The sizes of TFRs are scaled to the same size through the interpolation algorithm which enables the direct comparison. Then the performance of different TFRs are compared. STFT achieves the best overall score 85.40% after feature selection.

#### References

- [1] Dawber TR, Moore FE, Mann GV. Ii. coronary heart disease in the framingham study. *International journal of epidemiology* 2015;44(6):1767–1780.
- [2] Hanna IR, Silverman ME. A history of cardiac auscultation and some of its contributors. *The American journal of cardiology* 2002;90(3):259–267.
- [3] Springer DB, Tarassenko L, Clifford GD. Logistic regression-hsmm-based heart sound segmentation. *IEEE Transactions on Biomedical Engineering* 2016;63(4):822–832.
- [4] Haghghi-Mood A, Torry N. Time-frequency analysis of systolic murmurs. In *Computers in Cardiology*. IEEE, 1997; 113–116.
- [5] Chen T, Xiang L, Zhang M. Recognition of heart sound based on distribution of choi-williams. *Research on Biomedical Engineering* 2015;31(3):189–195.
- [6] Avendano-Valencia L, Godino-Llorente J, Blanco-Velasco M, Castellanos-Dominguez G. Feature extraction from parametric time–frequency representations for heart murmur detection. *Annals of Biomedical Engineering* 2010; 38(8):2716–2732.
- [7] Barma S, Chen BW, Man KL, Wang JF. Quantitative measurement of split of the second heart sound (s2). *IEEEACM Transactions on Computational Biology and Bioinformatics* 2015;12(4):851–860.
- [8] Moukadem A, Dieterlen A, Hueber N, Brandt C. A robust heart sounds segmentation module based on s-transform. *Biomedical Signal Processing and Control* 2013;8(3):273–281.
- [9] Patidar S, Pachori RB. Classification of heart disorders based on tunable-q wavelet transform of cardiac sound signals. In *Chaos Modeling and Control Systems Design*. Springer, 2015; 239–264.
- [10] Quiceno-Manrique A, Godino-Llorente J, Blanco-Velasco M, Castellanos-Dominguez G. Selection of dynamic features based on time–frequency representations for heart murmur detection from phonocardiographic signals. *Annals of biomedical engineering* 2010;38(1):118–137.
- [11] Cai Y, Yang M, Yin H. Relieff-based multi-label feature selection. *International Journal of Database Theory and Application* 2015;8(4):307–318.
- [12] et al L. An open access database for the evaluation of heart sound algorithms. *Physiol Meas* (In Press) 2016;.

Address for correspondence:

Jiqing Han  
 School of Computer Science and Technology / Harbin Institute of Technology  
 92 Xidazhi St /Nangang, Harbin / China  
 jqhan@hit.edu.cn

Shiwen Deng  
 School of Mathematical Sciences / Harbin Normal University  
 1 Shida Rd /Hulan, Harbin / China  
 dengswen@gmail.com