



Heart Sound Recognition of Congenital Heart Disease Based on MFCC and BFCC Feature Extraction

Zhu Lili¹, Pan Jiahua², Shi Jihong¹, Yang Hongbo², Tan Chaowen¹, Wang Weilian^{1,*}

¹School of Information, Yunnan University, Kunming, China

²Yunnan Fuwai Cardiovascular Disease Hospital, Kunming, China

Email address:

857407372@qq.com (Zhu Lili), wlwang_47@126.com (Wang Weilian)

*Corresponding author

To cite this article:

Zhu Lili, Pan Jiahua, Shi Jihong, Yang Hongbo, Tan Chaowen, Wang Weilian. Heart Sound Recognition of Congenital Heart Disease Based on MFCC and BFCC Feature Extraction. *Asia-Pacific Journal of Information Science and Technology*. Vol. 1, No. 1, 2019, pp. 1-6.

Received: December 18, 2018; **Accepted:** January 17, 2019; **Published:** February 1, 2019

Abstract: Cardiac auscultation is a basic method for the initial diagnosis and screening of congenital heart disease (CHD). The analysis of heart sound signals is conducive to the realization of machine-assisted auscultation. In this paper, the original heart sound signal was denoised and reconstructed by wavelet transform. The Shannon-envelope of the heart sound signal was extracted and the cardiac cycles were segmented. The feature parameters of MFCC and BFCC with a frame length of 2048 were extracted from S1 as the starting point, and a set of 32-dimensional feature was obtained. It was used as input to the BP neural network for learning, training, and testing of the network. 60 heart sounds of CHD and 60 normal heart sounds were selected randomly for training. Also the heart sounds of 40 CHD and 40 normal heart sounds were selected randomly to test after training. The experimental results show that the sensitivity is 85.11%, and the specificity is 88.89% of BP network by using BFCC features. The network sensitivity of using the MFCC is 86.96%, and the specificity is 90.91%. The BP network by using MFCC feature parameter has a more significant recognition ratio than the one using BFCC feature parameter as the input.

Keywords: Congenital Heart Disease (CHD), Wavelet Denoising, MFCC, BFCC, BP Neural Network

基于MFCC与BFCC特征提取的先心病心音识别

朱莉莉¹, 潘家华², 施继红¹, 杨宏波², 谭朝文¹, 王威廉^{1*}

¹云南大学信息学院, 昆明, 中国

²云南省阜外心血管病医院, 昆明, 中国

邮箱

857407372@qq.com (朱莉莉), wlwang_47@126.com (王威廉)

摘要: 心脏听诊是目前先心病初诊和筛查的基本手段, 对心音信号的分析研究有助于实现机器辅助听诊。本文采用小波变换对原始心音信号去噪、重构, 提取信号的香浓包络并进行心动周期分段定位。以S1为起点分别进行帧长为2048点的MFCC和BFCC的32维特征参数提取, 并作为BP神经网络的输入用于网络学习训练和测试。随机选取先心病病例心音和正常心音各60例进行训练, 异常、正常心音各40例进行测试。实验结果表明, BFCC特征参数作为输入训练的网络敏感度为85.11%, 特异度为88.89%; 而MFCC特征参数作为输入训练的网络敏感度为86.96%, 特异度达到90.91%。MFCC特征参数较BFCC特征参数作为输入的网络具有更好的识别效果。

关键词: 先心病, 小波变换, MFCC, BFCC, BP神经网络

1. 引言

先天性心脏病(Congenital heart disease, CHD)几乎占据了婴儿主要出生缺陷疾病的1/3,是最常见的心脏病[1]。先心病的诊断分为初诊和确诊两个阶段。确诊主要依靠超声心动图,而初诊和筛查主要靠心脏听诊。心音能反映心脏工作是否正常。心音通常包括第一、第二、第三、第四心音(S1-S4)以及心脏杂音,研究心音和心音杂音对于心血管疾病,尤其是先心病的诊断具有重要意义。心音在采集过程中往往有噪声干扰,为消除噪声有采用基于dbN小波进行心音去噪[2]的,有采用平移不变小波阈值去噪方法[3]的。在心音信号包络提取上,有归一化香浓化包络,希尔伯特黄变换包络,以及维奥拉积分提取包络[4]。对心音信号的分段定位,有采用多尺度矩分析来定位心音的每个心动周期,以及根据近似循环特性,寻找极值点来定位心动周期的[5]。对心音信号的特征提取研究,有使用小波包倒谱系数作为特征参数[6]的,有使用将双声道能量熵比作为特征参数[7]的,还有提取主IMF分量的关联维数,以及多重分形谱的宽度作为特征向量[8]的。但这些研究所使用的数据量偏少,所提取的特征欠缺完备性。心音信号分类识别算法,有通过BaumWelch参数重新估计和隐马尔科夫模型(HMM)的组合[9]进行识别的,有使用基于遗传算法(GA)的训练周期模糊神经网络(RFNN)进行分类识别[10]的,有基于最小二乘支持向量机(LS-SVM)[11]、高斯混合模型(GMM)[12]等方法进行心音样本识别。这些文献中的识别算法都较好的进行识别,但是需计算量庞大,导致识别过长。近年来语音信号处理和识别获得了较好的发展,能够有效地对语音信号进行分析。心音同属于音频信号,可借助语音信号处理技术对先心病心音进行分析研究。本文采用小波变换对心音信号进行去噪和重构处理、用归一化香浓公式提取包络,并精确地进行了心音分段定位。用语音中常用的Mel频率倒谱系数(Mel-Frequency Cepstrum Coefficient,

MFCC)[13]和Bark频率倒谱系数(Bark frequency Cepstral Coefficient, BFCC)[14]对心音信号进行特征提取,并做了对比性研究。

本文所用CHD心音数据来源于课题组在云南省阜外心血管病医院临床采集样本,正常心音数据来源于课题组在云南省德宏州中小学进行先心病筛查的健康儿童采集样本。实验中随机选取了心音数据库中异常数据100例,正常数据100例,采样频率为5000HZ。本文选用db6小波对信号进行去噪重构,提取归一化香浓包络进行分段定位,采用Mel频率倒谱作为特征参数,应用BP神经网络对正常心音和异常心音分类识别。

2. 预处理

2.1. 小波变换

小波变换的主要过程是根据选定的小波基和分解层数,选择适当的阈值和阈值函数进行阈值处理,对处理过后的信号进行重构得到去噪后的信号[15]。小波变换的小波函数的选择并不是唯一的,不同的小波函数对同一信号会产生不同的效果。

信号 $f(t)$, 它的小波变换公式如公式(1)所示:

$$WT(\alpha, \beta) = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{+\infty} f(t) * \varphi\left(\frac{t-\beta}{\alpha}\right) dt \quad (1)$$

α 是尺度,控制小波函数的缩放, β 是平移量,控制小波函数的平移。小波函数的平移和收缩允许时频窗在高频变窄而在低频变宽[16]。

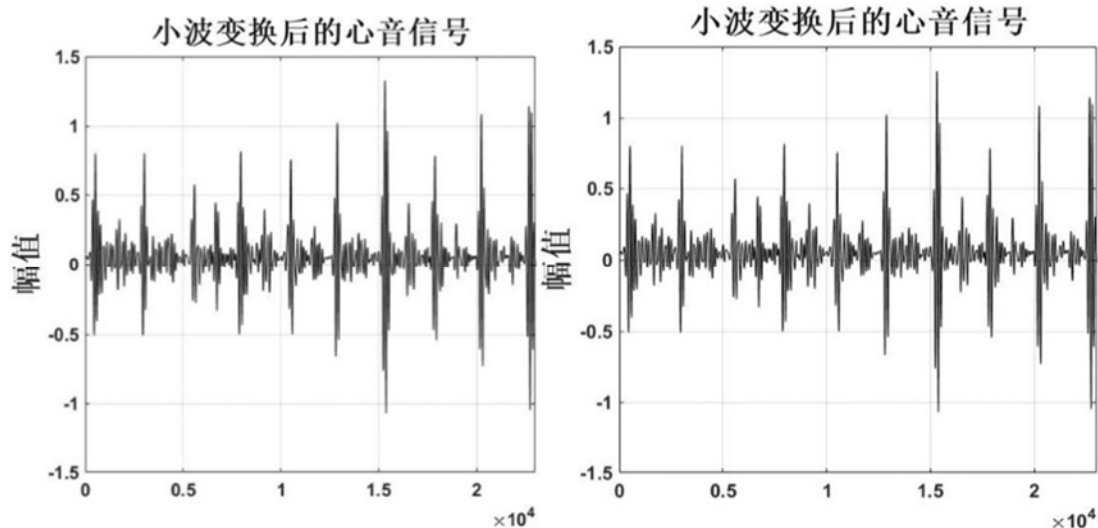


图1 原始信号与小波重构后的信号。

本文选择的小波基是db6小波,分解层数为8层,选择的是固定阈值和软阈值函数对信号进行阈值处理,重构后得到去噪信号。对某一例信号进行小波去噪的结果如图1

所示。左侧是原始心音信号图,右侧为去噪重构后的信号图。由图中可以观测到,小波变换后的信号去除了许多毛刺和干扰。

2.2. 分段定位

对信号进行分段定位,能够统一所有的心音信号的起始点,便于信号的特征提取。一个心动周期包含S1, S1S2间期, S2和S2S1间期,如图2所示。

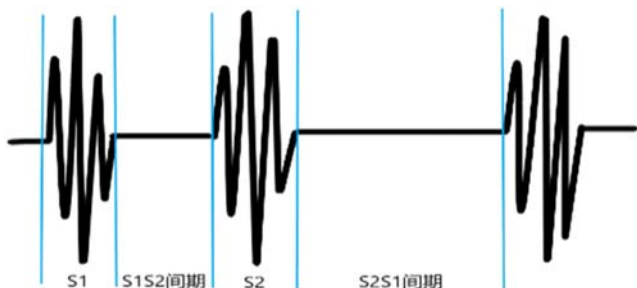


图2 一个心动周期。

本文所有信号都是以S1作为信号的起始点。为了能够清晰地分辨出心音信号的S1起始点,首先要提取信号



图3 MFCC特征参数提取过程。

- (1) 预加重: 由于低频段能量大,而高频段能量不足。对信号进行预加重,提高高频段的信噪比,使信号变得平坦。
- (2) 分帧加窗: 其主要目的是为了将短时非平稳随机的心音信号在极短的时间内看作是平稳信号,便于分

析研究。帧长一般取信号中认为相对平稳的长度,帧移取帧长的1/2或1/3。为了克服泄露现象,选择Hamming窗。一帧原始信号和一帧加窗信号如图4所示。

3. 特征提取

3.1. MFCC

Mel频率倒谱系数的特征提取在语音信号中是使用比较广泛的一种方法。MFCC的特征提取首先对信号进行预处理,经过分帧加窗后的信号再进行快速傅里叶变换,通过Mel三角滤波器组滤波及尺度变换,取对数能量后进行离散余弦变换即可。MFCC特征参数提取过程如图3所示。

析研究。帧长一般取信号中认为相对平稳的长度,帧移取帧长的1/2或1/3。为了克服泄露现象,选择Hamming窗。一帧原始信号和一帧加窗信号如图4所示。

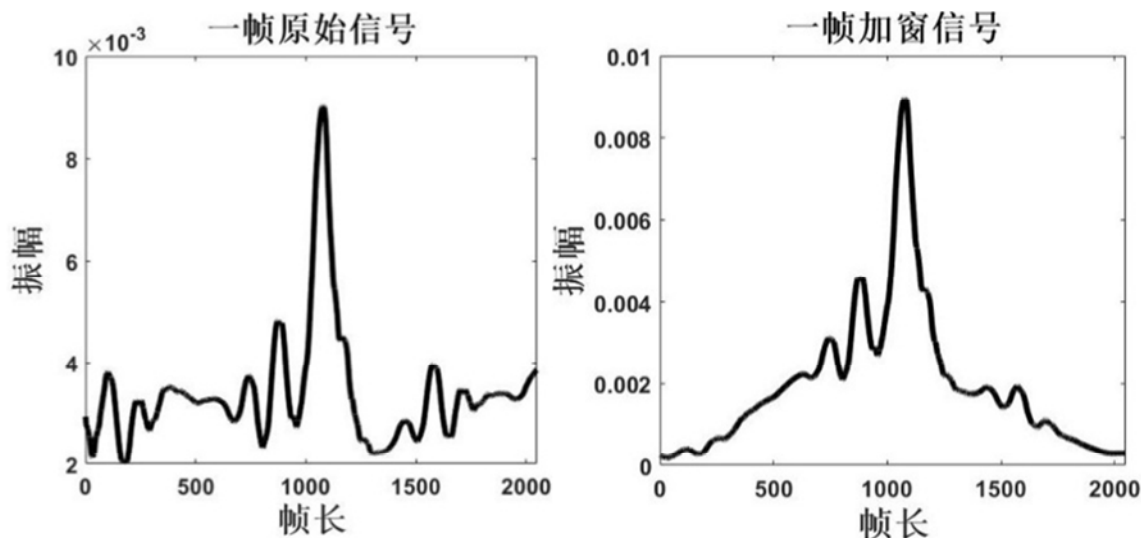


图4 一帧原始信号和一帧加窗信号。

- (3) FFT: 快速傅里叶变换,通过公式2,将心音信号 $x(i)$ 从时域变换到频域。

$$X(k) = \sum_{i=0}^{N-1} x(i)e^{-j2\pi ik/N}; (0 \leq i, k \leq N-1) \quad (2)$$

- (4) Mel滤波器组: 根据Mel频率与线性频率之间的公式3,将FFT后的信号进行滤波和刻度转换。Mel三角滤波器结构如图5所示。

$$Mel(f) = 2595 \log_{10}(1 + f/700) \quad (3)$$

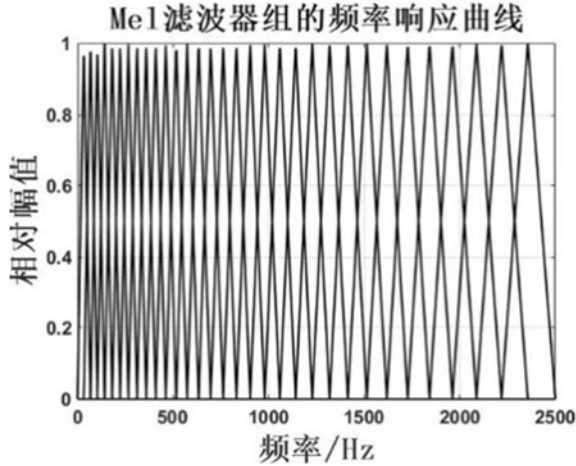


图5 Mel三角滤波器组。

- (5) 取对数能量: 根据公式4计算每个滤波器输出的对数能量, 实现对心音频谱的压缩, 得到对数频谱 $S(m)$ 。



图6 BFCC特征参数提取过程。

BFCC特征提取的滤波器组选用的是Bark滤波器组, Bark滤波器结构如图7所示。

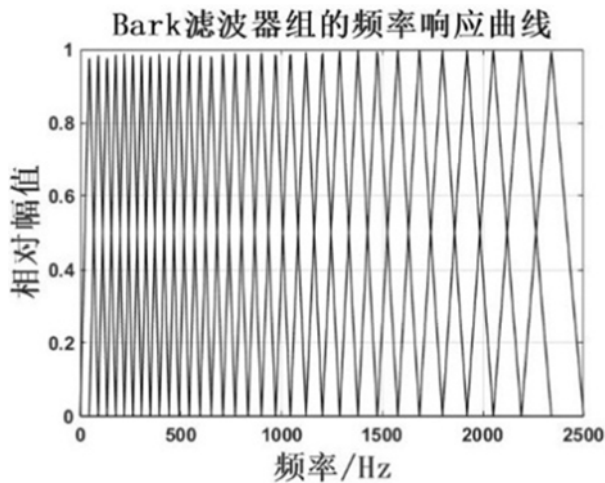


图7 Bark滤波器组。

实验中通过Bark滤波器实现信号从频域到Bark域的转变, 得到的Bark域系数通过离散余弦变换, 转换成时域的BFCC特征参数。Bark频率与线性频率之间的关系如公式6所示。

$$Bark(f) = 6 \ln \left[\frac{f}{600} + \left[\left(\frac{f}{600} \right)^2 + 1 \right]^{0.5} \right] \quad (6)$$

3.3. 特征提取

在语音中的MFCC和BFCC特征参数的提取过程中, 采样频率是8000HZ, 一般选用的帧长是256或128。虽然

$$S(m) = \ln \left(\sum_{k=0}^{N-1} |X(k)|^2 H_m(k) \right); (0 \leq m \leq M) \quad (4)$$

M 是滤波器个数, $H_m(k)$ 是每个滤波器的传递函数。

离散余弦变换: 根据DCT公式5, 将处理过的信号转换到时域的MFCC, 即MFCC特征参数 $C(n)$ 。

$$C(n) = \sum_{m=1}^M S(m) \cos \left(\frac{\pi n(m-0.5)}{M} \right) \quad (5)$$

3.2. BFCC

BFCC的特征提取是语音中另一种特征提取的方法, 它和MFCC特征提取很相似[17]。预加重信号, 使信号平坦, 分帧加窗后进行快速傅里叶变换, 通过Bark滤波器组滤波, 取对数能量, 进行离散余弦变换得到BFCC特征参数。BFCC特征参数提取过程如图6所示。

心音和语音都是音频信号, 但心音信号与语音信号的产生原理和信息内容等都是不同的。心音信号的采样频率是5000HZ, 经过试验研究, 心音信号的MFCC特征参数提取和BFCC特征参数提取中的帧长选用2048, 帧移选用1024, 实验效果最佳。一个完整心动周期包含收缩期和舒张期。一般收缩期的时长是0.1秒到0.3秒左右。当帧长取2048时, 相当于包含了半个心动周期的信息。一个完整收缩期信息量总是能被包含在特征中。经过试验总结, 本文实验中的加窗选择的都是Hamming窗, 选择的滤波器个数是32个, 最终形成了32维的特征信号。

4. 实验结果

4.1. 分类识别

机器识别的过程是一个样本学习匹配的过程。近年来, 神经网络成为了热点研究的识别器。本文使用的是能够前向传递信息, 后向反馈误差的BP(Back Propagation)神经网络。实验中使用的是有着一层输入层, 一层隐含层和一层输出层的三层BP神经网络。根据特征提取的32维MFCC特征信号和32维BFCC特征信号, 输入层节点采用32个。实验的目的是判别信号属于异常还是正常, 输出节点以为2个。隐含层的节点根据多次实验结果分析, 最终采用10个节点。虽然MFCC特征参数和BFCC特征参数是不同的特征, 但是BP神经网络学习过程中的网络结构是一样的。根据梯度下降法对BP神经网络进行训练, 直至网络达到所要求的收敛误差。

4.2. 测试结果

实验分为两个阶段: 学习过程和测试过程。

学习过程：即网络的训练过程。本文随机选取120例信号(60例正常信号和60例异常信号)作为网络的训练集进行训练。直至达到收敛要求。

测试过程：检验网络训练的效果。本文随机选取80例信号(40例正常信号和40例异常信号)作为网络的测试集。如表1和表2，分别是MFCC特征参数作为输入的网络测试结果和BFCC特征参数作为输入的网络测试结果。

表1 MFCC特征参数网络测试结果。

信号情况	测试数	正确识别数	正确识别率(%)	敏感度(Sensitivity)(%)	特异度(Specificity)(%)
正常	40	36	90.00	86.96	
异常	40	34	85.00		90.91

表2 BFCC特征参数网络测试结果。

信号情况	测试数	正确识别数	正确识别率(%)	敏感度(Sensitivity)(%)	特异度(Specificity)(%)
正常	40	35	87.50	85.11	
异常	40	33	82.50		88.89

由表1可以看出，由120例MFCC特征参数作为训练集用于BP神经网络的学习过程，使用80例信号作为测试集。在40例正常信号中能够正确识别出36例，正确识别率达到90.00%；在40例异常信号中能够正确识别出34例，正确识别率是85.00%。MFCC特征参数作为训练集的BP网络识别过程的敏感度达到86.96%，特异度达到90.91%。

由表2可以发现，由120例BFCC特征参数作为训练集用于BP神经网络的训练过程，80例信号对网络效果进行检测。在40例正常信号中能够正确识别35例信号，正确识别率是87.50%；在40例异常信号中能够正确识别33例信号，正确识别率是82.50%。BFCC特征参数作为训练集的BP网络识别过程的敏感度是85.11%，特异度是88.89%。

从上述两个表中可以观察到，两种不同的特征参数进行BP网络的训练得到的结果是不同的，如图8所示。

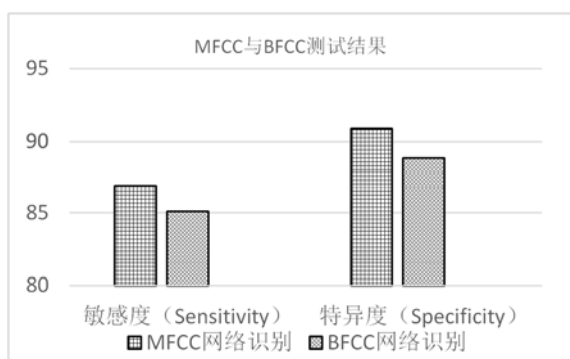


图8 MFCC和BFCC网络测试结果。

观察图8可以看出，MFCC特征参数进行训练识别的效果要比BFCC特征参数进行训练识别的效果好。MFCC特征参数训练的网络的敏感度为86.96%，特异度为90.91%，比BFCC特征参数训练的网络的敏感度和特异度高。说明MFCC特征参数更能反映心音信号特征，更能正确识别。

5. 结论

心音信号是对心脏情况的反馈，对心音进行研究是极具意义的。通过对心音信号的研究分析，可以实现机器辅

助听诊。本文通过小波变换对原始心音信号进行去噪处理，提取去噪后信号包络，便于分段定位，截取始终以S1作为起始点的心音信号，有利于下一步特征提取的进行。经过试验研究分析，选用帧长2048，帧移1024分别对预处理过的信号进行MFCC特征参数和BFCC特征参数提取。将特征提取到的32维信号作为BP神经网络的输入进行网络训练。实验结果表明，BFCC特征参数训练的网络敏感度为85.11%，而特异度是88.89%；MFCC特征参数训练的网络敏感度是86.96%，特异度达到90.91%。两者相比较，MFCC特征参数训练的网络比BFCC特征参数训练的网络有着更好的识别效果。

致谢

项目基金：基于人工智能与互联网+的先天性心脏病初诊辅助诊断技术与平台研发(No.2018F017);先天性心脏病心音库构建与心音特征提取方法研究(No.61261008)。

参考文献

- [1] 高燕,黄国英.先天性心脏病病因及流行病学研究进展[J].中国循证儿科杂志,2008,3(3):213-222。
- [2] 田文雪.基于香农熵的心音信号检测方法研究[J].无线互联科技,2017(6):3-4。
- [3] 郭兴明,何彦青,卢德林,等.平移不变小波在心音信号去噪中的应用[J].计算机工程与应用,2014,50(24):209-212。
- [4] 郭兴明,蒋鸿,郑伊能.基于改进的维奥拉积分方法提取心音信号包络[J].仪器仪表学报,2016,37(10):2352-2358。
- [5] Yan Z, Jiang Z, Miyamoto A, et al. The moment segmentation analysis of heart sound pattern.[J]. Computer Methods & Programs in Biomedicine, 2010, 98(2):140-150.
- [6] Abo-Zahhad M, Ahmed S M, Abbas S N. Biometrics from heart sounds: Evaluation of a new approach based on wavelet packet cepstral features using HSCT-11 database ☆ [J]. Computers & Electrical Engineering, 2016, 53:346-358.

- [7] 孙科学,魏敏,宋江珉,等.双声道心音能量熵比的提取与识别研究[J].计算机技术与发展,2017,27(12):166-169。
- [8] 郭兴明,张文英,袁志会,等.基于EMD关联维数和多重分形谱的心音识别[J].仪器仪表学报,2014,35(4):827-833。
- [9] Arslan A, Yildiz O. Automated Auscultative Diagnosis System for Evaluation of Phonocardiogram Signals Associated with Heart Murmur Diseases[J]. Gazi University Journal of Science, 2018, 31(1).
- [10] Uyar K, İlhan A. Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks[J]. Procedia Computer Science, 2017, 120:588-593.
- [11] 许莉莉,师炜,郭学谦,等.基于最小二乘支持向量机的心音分类识别研究[J].中国医疗设备,2017,32(4):38-41。
- [12] 钟丽莎,万江中,黄志伟,等.基于GMM的心音信号生物识别方法研究[J].中国医疗器械杂志,2013(2):92-95。
- [13] Chen T, Zhao S, Shao S, et al. Non-invasive diagnosis methods of coronary disease based on wavelet denoising and sound analyzing[J]. Saudi Journal of Biological Sciences, 2017, 24(3):526.
- [14] 段瑞琪.基于MFCC和短时能量混合的心音身份识别[D].燕山大学,2015。
- [15] Gulzar T, Singh A, Sharma S. Comparative Analysis of LPCC, MFCC and BFCC for the Recognition of Hindi Words using Artificial Neural Networks[J]. Acta Materialia, 2014, 56(13):3132-3145.
- [16] 张磊邦,唐荣斌,蒋建波,等.基于小波变换的心音特征值提取算法研究[J].电子测量技术,2014,37(6):42-46。
- [17] Rehman F U, Kumar C, Kumar S, et al. VQ based comparative analysis of MFCC and BFCC speaker recognition system[C]// International Conference on Information and Communication Technologies. 2017:28-32.