

DOI: 10.16289/j.cnki.1002-0837.2016.05.011

## 基于小波变换和 K-means 聚类算法的心电信号特征提取

王瑞荣<sup>1</sup>, 余小庆<sup>1\*</sup>, 朱广明<sup>1</sup>, 王 敏<sup>2</sup>

(1. 杭州电子科技大学生命信息与仪器工程学院, 浙江杭州 310018; 2. 杭州红十字会医院, 浙江杭州 310003)

**摘要:** 目的 研究一种基于小波变换和 K-means 聚类算法的心电信号特征提取方法, 根据特征点信息判断心电是否正常。方法 利用小波变换和形态学滤波方法去除工频干扰、肌电干扰和基线漂移等主要的噪声之后, 利用 K-Means 聚类算法提取出心电信号的 QRS 波群、P 波和 T 波这 3 个主要的特征点, 实现心电智能诊断。结果 实验数据取自 MIT-BIH 数据库, 多次实验结果显示 QRS 波群的阳性检测度( $P^+$ ) 达到 99.68% 和灵敏度( $S_e$ ) 达到 99.21%, P 波和 T 波的检测准确度分别达 91.43% 和 97.01%。结论 相对于其它方法, 本文心电特征提取方法准确度较高, 具有一定参考价值; 在移动医疗和临床医疗方面具有一定实用价值。

**关键词:** 心电信号; 小波变换; K-means; QRS 波群; P 波; T 波

**中图分类号:**      **文献标识码:** A      **文章编号:** 1002-0837(2016)05-0368-04

**Feature Extraction of Electrocardiogram Signal Based on Wavelet Transform and K-means Clustering Algorithm.** Wang RuiRong, Yu Xiaoqing, Zhu Guangming, Wang Min. Space Medicine & Medical Engineering, 2016, 29(5): 368-371

**Abstract:** **Objective** To study a method to extract the features of ECG signal based on wavelet transform and K-means clustering algorithm so as to provide useful information for detection of cardiac disease or abnormality. **Methods** After removing the power frequency interference, electromyography (EMG) interference and baseline drift with wavelet transform and morphological filtering methods, the K-Means clustering algorithm was adopted to extract the QRS complex, P wave and T wave, and the characteristics of the ECG signal, thus the automatic delineation of ECG was realized. **Results** The experimental data was taken from the MIT-BIH database. The results showed that the detection rate of the positive ( $P^+$ ) was 99.68% and sensitivity ( $S_e$ ) of the QRS complex was 99.21%. While the extraction precision of P and T waves were 91.43% and 97.01% respectively. **Conclusion** The method in this paper has higher accuracy as compared with some other methods in ECG feature extraction, and it can provide reference for other researchers, and is practical in mobile medical and clinical.

**Key words:** electrocardiogram signal; wavelet transform; K-means; QRS complex; P wave; T wave

**Address correspondence and reprint requests to:** Yu Xiaoqing. College of Life Information Science & Instrument Engineering, Hangzhou Dianzi University, Hangzhou Zhejiang 310018, China

正常心电信号主要有 QRS 波群、P 波和 T 波等特征。根据这几个特征点来判断心电信号的异常与否。心电信号分析包括三大部分: 信号预处理、特征提取以及诊断分类。本文重点介绍前两者的处理算法。心电信号预处理常用的方法有数字滤波<sup>[1]</sup>、小波阈值滤波、自适应滤波<sup>[2]</sup>和形态滤波<sup>[3]</sup>。QRS 复合波的检测算法主要有小波变换法<sup>[4]</sup>、差分阈值法<sup>[5]</sup>和神经网络法<sup>[6]</sup>。P 波、T 波的检测是心电信号特征检测的一大难题, 正常心电信号有较好的检测效果, 但是异常心电信号常常伴随 P 波和 T 波的畸变, 导致检测效果不佳。

本文采用小波变换结合形态滤波法对心电信号进行预处理, 采用 K-means 聚类算法(差分阈值法)进行 QRS 波群的特征提取, 采用波峰斜率特征法进行 P、T 波的检测。传统数字滤波器容易造成信号的失真, 完全小波变换滤波计算复杂度较高, 自适应滤波需要期望信号参与才能达到较好的滤波效果。小波变换结合形态学滤波法综合了小波变换和形态学滤波的优势, 使算法的复杂度减小, 同时提高信号信噪比。K-Means 聚类算法相对于小波变换和神经网络算法更简单、快速, 同样能达到很好的检测效果。

### 1 方法

#### 1.1 基于小波变换结合形态滤波的信号预处理

修回日期: 2016-07-08

\* 通讯作者: 余小庆 185743125@qq.com

基金项目: 国家自然科学基金(61374005)

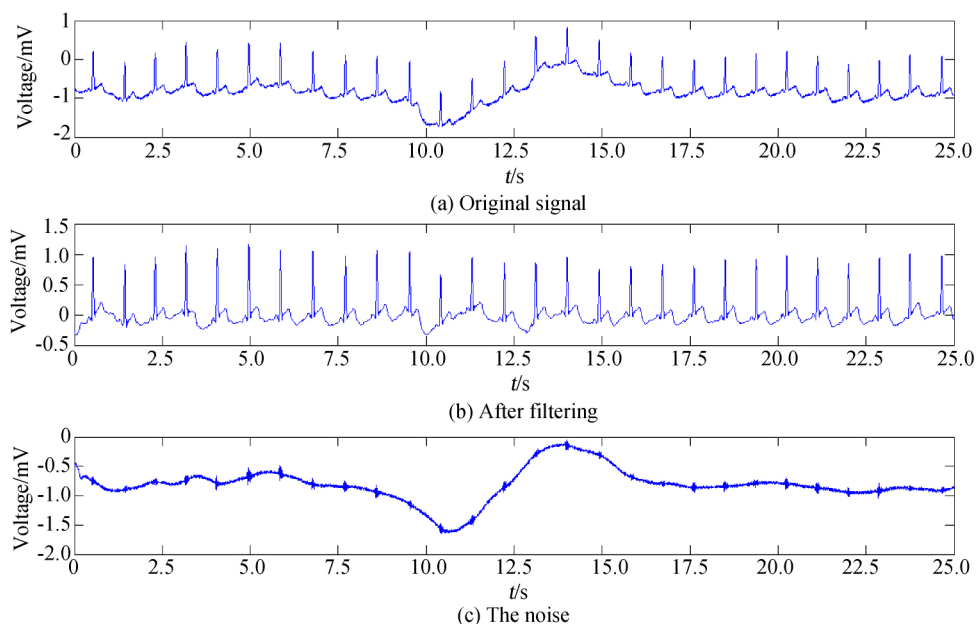


图1 信号预处理结果

Fig. 1 Results after morphological filtering

预处理过程需要去除工频干扰、肌电干扰和基线漂移这3种主要的噪声。肌电干扰和工频干扰相对心电信号属于高频信号,而基线漂移属于低频信号,文中采用小波变换去除高频信号,而采用形态学滤波去除低频信号,主要分为以下步骤:

1) 将信号进行4尺度分解,根据小波分解的特点得到肌电干扰基本在1和2尺度上,工频干扰在2、3尺度上,心电信号在3、4尺度上;

2) 将1、2、3尺度上的细节系数进行阈值处理,设置阈值来确定该细节系数是否置0;

3) 将处理后的尺度系数重构信号,再进行形态学滤波,去除基线漂移;

4) 得到滤波后的信号,如图1所示,第一幅是带有噪声的原始信号,第二幅是预处理以后的信号,第三幅是被滤除的噪声。

## 1.2 基于 K-means 聚类算法的 QRS 波群提取

文中将心电信号的斜率绝对值作为聚类数据集  $X = \{x_1, x_2, x_3, \dots, x_n\}$ , 以及将要生成的数据子集数目  $K=2$ , 得到子集  $C = \{c_1, c_2\}$ ,  $c_1$  和  $c_2$  分别有各自的聚类中心  $u_1$  和  $u_2$ , 计算该类内各个点到聚类中心  $u_i$  的距离, 若该点距  $u_1$  的距离小, 则该点属于  $c_1$ , 反之属于  $c_2$ , 归类结束后, 分别计算各类的距离平方和, 见式(1)。

$$J(c_k) = \sum_{x \in c_k} \|x_i - u_k\|^2 \quad k=1, 2 \quad (1)$$

聚类结束的标志是各类总的平方和, 见公式(2), 达到最小值<sup>[7]</sup>。

$$J(C) = \sum_k J(c_k) \quad (2)$$

此时可根据聚类中心值, 得到最后的分类子集  $C$  以及分类结果  $V = \{v_1, v_2, v_3, \dots, v_n\}$ 。QRS 波群集中在斜率较大的区域, 此时便定位出 QRS 波群的位置。

定位到 QRS 波群的区域后, 根据 R 波是 QRS 波群中最大波峰值这个特点, 可以很快定位到 R 波。检索到最大值的点是判定 R 波波峰的前提条件, 还必须满足波峰的判别条件, 若不满足波峰条件, 判断该点左右哪个方向比该点值大, 若右边值大, 则继续向右搜索直到找到满足波峰条件的位置, 才确定 R 波的位置。检索到 R 波以后, 根据 R 波的位置检索 Q 和 S。Q 和 S 是 R 波的两边波谷位置。

根据以上条件得到检索结果如图2所示。QRS 起终点检索方法与 QS 检索类似, 起终点的检索判断标志是波峰条件, 检索结果见图2。

## 1.3 P 波和 T 波提取

P 波和 T 波集中在 4、5 尺度上, 但 5 尺度的 R 波失真严重会影响 P 波和 T 波的检索, 文中选择在 4 尺度的近似信号上检索 P、T 波。P、T 波是 R-R 期间较明显的波峰, 为此检索出 R-R 期间所

有波峰点,从 R-R 期间的前半段中检索最大的波峰位置就是 T 波,后半段检索最大波峰位置就是 P 波。P、T 波起终点检索方法是往两边分别检索出波谷的位置,检索结果见图 3。

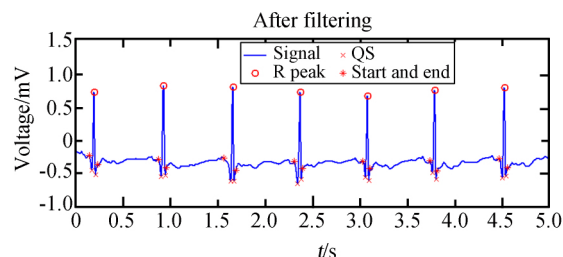


图 2 QRS 复合波检测结果  
Fig. 2 Detection results of the QRS complex

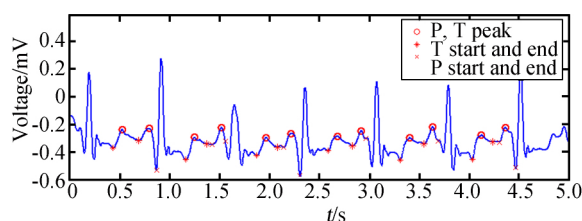


图 3 P 波和 T 波检测结果  
Fig. 3 Detection results of P and T waves

## 2 实验

实验所用数据来自 MIT-BIH,该数据库是由美国麻省理工学院提供的研究心律失常的数据库,共 48 组数据,存在 41 种心律失常,数据来自真实的病患。为了验证小波结合形态学滤波方法的有效性,选取 MIT-BIH 数据库中 30 s 的 119 号信号作为实验信号,信号中加入 50 Hz 的工频干扰,200 Hz 的高频干扰和 0.1 Hz 挤基线漂移干扰,分别进行小波阈值去噪、FIR 数字滤波器去噪、文中算法和自适应去噪,对比这几种算法的信噪比 SNR 和计算时间,结果如表 1 所示。

MIT-BIH 数据库中数据存在许多异常,其中 100,114,121,209 等是正常心电,102,217 等存在起搏心跳异常心电,109,111 等存在左束支传导

表 1 滤波方法的比较

Table 1 Comparison between several filtering methods

Method	SNR (db)	Cost time (s)
Wavelet transform	8.843	0.373
FIR digital filter	1.954	0.056
This work	9.532	0.204
Self-adaptive filter	6.905	0.172

阻滞异常,124,212,231 等存在右束支传导阻滞异常,200 等存在室性早博异常,109,121 等存在严重基线漂移,文中方法依然能准确检测出特征点,实验结果见表 2,FP 代表错检,FN 代表漏检,TP 代表正确检测。QRS 波群幅度较大,特征明显,检测难度较低,而 P 波和 T 波幅度小,易受干扰,并且在 QRS 波群检测出来的基础上进行检测,导致 P、T 波的检测准确度略低于 QRS 波群检测,但仍能分别达到 91.43% 和 97.01% 的检测准确度。

表 2 QRS 波群检测实验结果

Table 2 Detection results of QRS complex

ECG	Total	FP	FN	TP
QRS	5160	16	41	5119
P	5159	134	154	5005
T	5159	398	442	4717

文中所提算法与其他方法,对比结果见表 3,以敏感度( $S_e$ )和阳性检测率( $P^+$ )来评估检测算法的可靠性。从表 3 都可以得出,文中提出的方法在 QRS 波群的检测上达到 99.21% 的灵敏度和 99.68% 的阳性检测率。从表中可以看出,文中算法相比其他算法具有较高的准确率。

## 3 讨论与结论

心电信号的 QRS 波群、P 波和 T 波等特征的检测是心电自动分析诊断的关键步,获得了准确

表 3 QRS 检测准确率对比  
Table 3 Comparison of QRS complex detection accuracy

Method	$P^+$ (%)	$S_e$ (%)
This work	99.68	99.21
Adaptive threshold and principal component analysis <sup>[5]</sup>	99.71	96.28
Wavelet based derivative approach <sup>[8]</sup>	99.65	98.91
Neural network algorithm <sup>[6]</sup>	98.75	98.75
Curve length transform (CLT) <sup>[9]</sup>	97.86	98.59
Fast Fourier transform <sup>[10]</sup>	99.93	96.59

的特征,才能对心电信号进行诊断。本文提出的采用小波变换结合形态学对信号进行滤波,采用 K-means 聚类算法提取 QRS 波群,在这两者的基础上提取 P 波和 T 波。小波变换结合形态学滤波方法抑制了噪声,提高了信噪比,为 QRS 波群等特征检测提供了良好的基础,且 K-means 聚类算法充分利用心电信号的特点,准确定位 QRS 波群区域,在 QRS 波群区域内定位幅度较大的波峰 R 波也变得较为容易。整个信号特征提取的过程,小波变换、形态学滤波和 K-means 聚类算法都充分发挥了各自的算法优势,取得了较高的特征提取准确率。

本文最后将提出的方法与其他算法进行对比,证明了文中算法的优势,具有一定的理论意义和使用价值。心电信号的特征提取是心电分析的关键步,只有准确提取特征,才能准确分类,减少诊断误差,达到人们所需要的智能诊断效果。

#### [参考文献]

- [1] Mahesh SC, Aggarwala RA, Uplane MD. Interference reduction in ECG using digital FIR filters based on Rectangular window[J]. WSEAS. Trans. Sig. Proc., May. 2008, 5(4): 340-349.
- [2] 张翔,周严. 基于自适应滤波的心电图中呼吸干扰的抑制[J]. 生物医学工程杂志, 2011, 28(1): 58-62.
- [3] 李延军,严洪,王增丽. 心电基线漂移去除方法的比较研究[J]. 航天医学与医学工程, 2009, 22(5): 381-386.
- [4] Li YJ, Yan H, Wang ZL. A comparative study on removal methods of ECG baseline wandering[J]. Space Medicine & Medical Engineering, 2009, 22(5): 381-386.
- [5] Rai HM, Trivedi A, Chatterjee K, et al. R-peak detection using Daubechies wavelet and ECG signal classification using radial basis function neural network[J]. Journal of the Institution of Engineers (India): Series B, 2014, 95(1): 63-71.
- [6] Rodriguez R, Mexicano A, Ponce-Medellin R, et al. Adaptive threshold and principal component analysis for features extraction of electrocardiogram signals[C]//Computer, Consumer and Control (IS3C), 2014 International Symposium on. IEEE, 2014: 1253-1258.
- [7] Hasan MA, Ibrahimy MI, Reaz MBI. NN-Based R-peak detection in QRS complex of ECG signal[C]//4th Kuala Lumpur International Conference on Biomedical Engineering, 2008, 1&2: 217-220.
- [8] Anil KJ. Data clustering: 50 years beyond K-means[J]. Pattern Recognition Letters, 2010, 31(8): 651-666.
- [9] Anita P, Talele KT. ECG feature extraction using wavelet based derivative approach[M]. Technology Systems and Management. Springer Berlin Heidelberg, 2011: 239-247.
- [10] Tekeste T, Bayasi N, Saleh H, et al. Adaptive ECG interval extraction[C]//Circuits and Systems (ISCAS), 2015 IEEE International Symposium on. IEEE, 2015: 998-1001.
- [11] Mironovova M, Bifla J. Fast Fourier Transform for feature extraction and neural network for classification of electrocardiogram signals[C]//Future Generation Communication Technology (FGCT), 2015 Fourth International Conference, IEEE, 2015: 1-6.

作者简介: 王瑞荣,男,博士,教授,研究方向为仪器科学与技术、生物信号处理。E-mail: wangrr@hdu.edu.cn