

# 基于小波变换和支持向量机的心电信号 ST 段分类

杨宇<sup>1</sup>, 司玉娟<sup>1,2</sup>, 宋晓洋<sup>1</sup>

(1. 吉林大学通信工程学院, 长春 130012; 2. 吉林大学珠海学院电子信息系, 广东 珠海 519041)

**摘要:** 为完成 ECG( Electrocardiogram) 信号特征点提取, 并对 ST 段分类, 提出了一种基于离散小波变换和支持向量机的 ST 分类算法。首先对信号进行预处理, 完成噪声消除, QRS 波群检测和提取特征值; 然后计算 ST 段平均值、曲线面积和标准差, 并结合使用 SVM( Support Vector Machine) 对 ST 段进行分类。Matlab 仿真结果表明, 小波去噪效果明显, ST 段未出现失真现象, 特征点提取完整。经 MIT-BIT 数据库验证, 分类结果显示交叉验证准确率平均值为 80.70%, 训练准确率平均值为 91.83%, 测试准确率平均值为 74.28%。

**关键词:** 特征点提取; 分类; 小波变换; 支持向量机

中图分类号: TN911

文献标识码: A

DOI:10.19292/j.cnki.jdxp.2016.03.003

## ST Segment Classification of ECG Signals Based on Wavelet Transform and Support Vector Machine

YANG Yu<sup>1</sup>, SI Yujuan<sup>1,2</sup>, SONG Xiaoyang<sup>1</sup>

(1. College of Communication Engineering, Jilin University, Changchun 130012, China;

2. Department of Electronic Information Science & Technology, College of Zhuhai, Jilin University, Zhuhai 519041, China)

**Abstract:** To complete the ECG signal feature points extraction and the classification of ST segment, we put forward an algorithm based on the discrete wavelet transform, combined with the first derivative and the SVM (Support Vector Machine). The algorithm can accomplish the signal preprocessing, noise elimination, QRS complex detection and extraction of characteristic value, calculating the average ST segment, curve area and the standard deviation, and the simple classification of ST segment by using the SVM combined with the three sets of data. The matlab simulation results show that the wavelet denoising is effective and has no distortion, and completely extract ST segment feature points. The data are downloaded from the MIT-BIT database, the classification results show that cross-validation average accuracy is 80.70%, the average accuracy of training is 91.83%, the average testing accuracy was 74.28%.

**Key words:** feature point extraction; classification; wavelet transform; support vector machine(SVM)

## 0 引言

心电信号主要由 QRS 波群、P 波、T 波等主要波形构成, 不同的波段表征不同的心脏活动, 是分析 ECG( Electrocardiogram) 信号的重要依据<sup>[1]</sup>。由于噪声频率和 ST 段频率有重叠部分, 为提取分析 ECG 特征点, 有效去除各种噪声, 准确定位并提取 ST 段成为分析心电信号的关键和研究 ST 段的重点。准确地识别 ST 段并分类有助于心肌缺血、冠心病等症状的诊断<sup>[2]</sup>。

收稿日期: 2015-08-29

基金项目: 吉林省重点科技攻关基金资助项目(20150204039GX); 吉林省长春市重大科技攻关专项基金资助项目(14KG064); 广东省省级科技计划基金资助项目(2013B010101020)

作者简介: 杨宇(1991—), 男, 湖北鄂州人, 吉林大学硕士研究生, 主要从事通信与信息系统研究, (Tel) 86-15948384148( E-mail) yangxiaojian@yahoo.com; 通讯作者: 司玉娟(1963—), 女, 长春人, 吉林大学教授, 博士生导师, 主要从事通信与信息系统研究, (Tel) 86-43604424766( E-mail) siyj@jlu.edu.cn。

目前为止,很多学者做了大量的工作,主要可分为两类:一些学者主要研究心电信号的形态,这类方法准确率高,不易出错,但计算复杂,无法实时检测分析;另一些学者主要研究ST段的动态波形,利用二维线性关系确定ST段形态,这类算法计算量小,实时性好,但由于ST段振幅较小,波形容易受到噪声影响,误差较大。T波检测算法主要有窗口搜索法,对消法和圆弧逼近法<sup>[3]</sup>。目前比较常用的ST段检测算法有多项式拟合ST段算法和神经网络算法,但多项式拟合算法易受噪声干扰,而神经网络算法则存在输入量过大而导致计算过程复杂的问题。很多学者借用小波算法思想运用小波分解和重构实现去噪和特征点提取<sup>[3-6]</sup>,Llamedo等<sup>[4]</sup>通过小波算法实现ECG信号的ST特征点提取,但未进行分类。Bustamante等<sup>[5]</sup>使用小波变换实时提取心电ST段特征点,但分类过于复杂,实际应用困难。具体涉及到ST段的分类识别的工作较少。

笔者使用SVM(Support Vector Machine)机器学习方法分类,选择ST段平均值,曲线面积和标准差作为参考特征向量,将ST段分为正常、抬高和压低3类。使用离散小波变换实现信号的预处理和滤波,检测出QRS波群和T波、P波及相应的起始点和终点,最后使用支持向量机对ST段进行分类。笔者使用的算法简单,适用性强,实时应用方便。

## 1 算法简介

### 1.1 小波变换

小波算法具有良好的时频局部化特性,能实现信号从时域到时间-尺度平面的转换,通过多尺度分析可以在不同分辨率下观察信号不同的局部化特征,因此,小波分析是进行时频局部特性分析的理想数学工具<sup>[7]</sup>。近年来,小波算法被广泛应用于信号和图像的处理中,具有良好的时频分析特性,一经提出便得到迅速发展。

小波变换常用于处理复杂的非平稳信号,是强大的时间-频率分析和信号编码工具<sup>[8]</sup>。利用小波可以同时捕获高频短时域和低频长时域信号。通过调制转换,突出显示信号形态的细微变化尺度,因此该方法对瞬态、非周期性和其他非平稳信号特性分析都非常适用。小波技术的另一个关键优势是包含多种小波函数,从而能为信号的处理提供多种适当的选择方式。

给定一个基本函数 $\psi(t)$ ,令

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (1)$$

其中 $a, b$ 均为常数,且 $a > 0$ 。设 $x(t) \in L^2(R)$ ,即平方可积,则 $x(t)$ 的小波变换定义为

$$W_x(a,b) = \frac{1}{\sqrt{a}} \int x(t) \psi^*\left(\frac{t-b}{a}\right) dt = \int x(t) \psi_{a,b}^*(t) dt = \langle x(t), \psi_{a,b}(t) \rangle \quad (2)$$

$W_x(a,b)$ 是信号 $x(t)$ 的小波变换, $a$ 是尺度因子, $b$ 是时移。 $\psi(t)$ 称为基本小波, $\psi_{a,b}(t)$ 是基本小波经伸缩和移位后产生的小波基函数。

在式(2)中, $b$ 用于对基本小波进行时域平移的中心位置, $a$ 用于对基本小波进行伸缩。由 $\psi(t)$ 变成 $\psi(t/a)$ ,当 $a > 1$ 时,若 $a$ 越大,则 $\psi(t/a)$ 的时域支撑范围(即时域宽度)越大;当 $a < 1$ 时, $a$ 越小,则 $\psi(t/a)$ 的宽度越窄。

### 1.2 支持向量机

在机器学习中,支持向量机也称为支持向量网络,通过分析数据和模式识别,完成监督学习的算法,常用于分类和回归分析以及相关的学习算法。给定一组训练样本,使用支持向量机训练算法建立一个模型,将新的实例分类到最接近的类别,使其成为非概率的线性分类。图1为小波变换的时间频域特性。支持向量机模型非常具有代表性,通过测试数据的映射关系,使不同的分类类别间间隔明显加宽。模型建立后,输入新的测试数据,然后映射到相同的空间中,

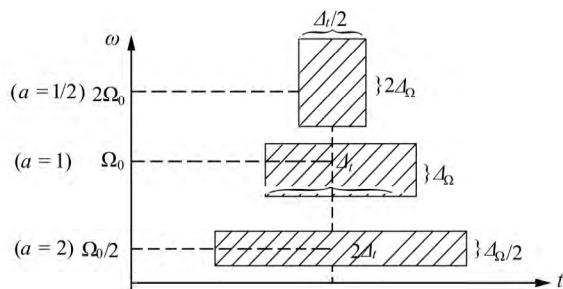


图1 小波变换的时间频域特性

Fig. 1 Time and frequency domain characteristic of wavelet transform

根据间隔大小预测数据所属类别。

除了进行线性分类外,支持向量机还能有效执行非线性分类,通过内核函数,映射到高维特征空间。更准确的说,支持向量机是由一个平面或一组高维空间构成,可用于分类,回归和其他任务。幅度越大,泛化误差越小,一般通过计算超平面到最近的训练数据点的最大距离,确定分类标准。

## 2 算法设计

### 2.1 噪声处理

在提取信号特征点之前,首先要去除噪声的干扰。ECG信号中的主要噪声为肌电干扰和基线漂移。肌电干扰的频率约为50 Hz,基线漂移频率很低,约在0.5 Hz以下,与ST段的频率非常接近。笔者采用离散小波变换去除噪声。

使用离散小波变换将心电信号分解为8层不同的频域尺度,根据心电信号和噪声的频率关系,将与噪声频率重叠的频域设定阈值,之后使用逆小波变换重构信号去除噪声。信号采样率为250 Hz,近似系数cA8的对应频率介于0~0.49 Hz,为去除基线漂移,cA8直接设置为0。通过估计细节系数cD1~cD4的噪声幅度,根据方程设定阈值

$$h = \sigma \sqrt{2 \log N} \quad (3)$$

其中 $N$ 为噪声估计的长度, $h$ 为阈值, $\sigma$ 定义如下

$$\sigma = \text{median}(|cD_1|) \quad (4)$$

该阈值为自适应阈值,能在保留重要特征信息的前提下,获得低噪声信号。

### 2.2 ECG 特征点检测

要从ECG信号中提取分析QRS波,通常采用求导法寻找波形中最大的斜率。在小波分解的信号中,cD6的能量系数明显小于细节系数的平均值。故为检测QRS波群,只需结合细节系数cD2、cD3、cD4、cD5的重构<sup>[9]</sup>。求出信号的一阶导数,定位正负极值对过零点,即R波峰值,并删除多检点,补偿漏检点,若两个R波间隔小于设定的最小R-R间距的一半,删除二者中幅度小的R波;当R波间隔大于1.6倍R-R间隔时,逐步减小阈值,重新进行检测。在检测的R波左右两端,利用所求得的一阶导数值变化,确定QRS波起点和终点。

在QRS波左右两边同时使用加窗法,找寻窗口内极值点,并计算对应距离,极值过零点分别确定为P波和T波最高点,再根据斜率关系确定P波和T波的起始点和终点。

### 2.3 ST段分类

大量实验研究表明,ST段抬高常伴随有心肌缺血,ST段压低常见于冠心病。ST段的正确分类能为医疗诊断提供很大的帮助,缩短诊断时间,为病人提供及时治疗。

ST段是QRS波群的终点到T波的起点部分,笔者使用MIT-BIT中数据,每路数据集包含V4和V5两路信号。选择3个特征向量:ST段均值、ST段面积和ST段均方根(RMS)作为判断标准。利用支持向量机将ST段分类为3种:正常、压低和抬高,并根据交叉验证法验证其准确率。

## 3 实验结果及对比

为验证算法的实用性,使用该算法对MIT-BIT库中不同类型的心电数据进行处理,记录实验结果。标记出QRS波群,T波、P波和对应的起始点,并提取出其中的ST段,将提取结果分类为正常,抬高和压低3种。

图2为去噪后的ECG波形及特征点定位,从图2中可以看出,该算法能正确检测到ECG信号的特征点和对应波形的起始和终止位置。图3为 $c$ 和 $g$ 对准确

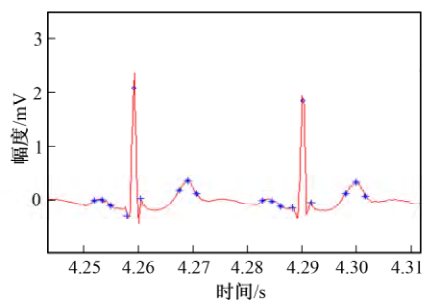


图2 去噪后的ECG波形以及特征点定位

Fig. 2 ECG waveform and feature point location after denoising

率的影响,适当地选择其值能提高分类的准确率。

图 4 为分类结果。从图 4 中可看到预测测试集分为 3 类,使用其预测的 SVM 模型对实际测试集分类。经过计算,采用 MIT-BIT 数据库分类结果准确率平均值为 74% 左右。

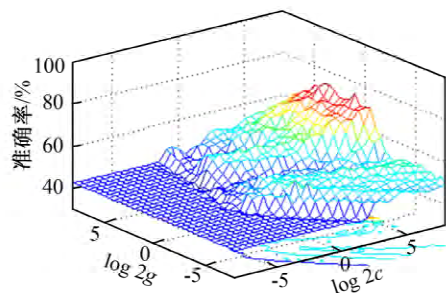


图 3 参数  $c$  和  $g$  对准确率的影响  
Fig. 3 The influence of parameter  $c$  and parameter  $g$  on the accuracy rate

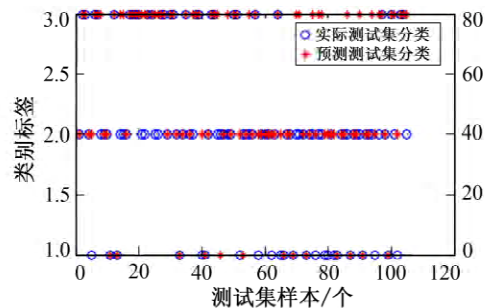


图 4 分类结果  
Fig. 4 Classification result

MIT-BIT 库中数据导联方式有 V1V2V3V4V5 和 MLIH 方式,表 1 为测得的 QRS 波群个数和文献的比较和文件分类的准确率。

表 1 实验测得的 R 波和分类准确率

Tab. 1 Experimental measured R wave and classification accuracy

文件	实验测得 R	文献 R <sup>[10]</sup>	CA 准确率/%	训练准确率/%	检测准确率/%
100	2 273	2 273	79.04	91.42	66.67
101	1 868	1 865	80.00	100.00	51.42
102	2 189	2 187	79.04	79.04	50.42
103	2 083	2 084	74.28	90.47	65.71
104	2 239	2 230	92.38	98.09	89.52
105	2 624	2 572	66.67	91.42	63.81
106	2 063	2 027	92.38	98.11	37.14
107	2 175	2 137	89.52	97.14	90.47
108	1 710	1 763	85.71	95.33	74.29
109	2 532	2 532	84.76	85.71	64.76
111	2 142	2 124	72.38	95.24	58.10
112	2 539	2 539	78.10	82.86	84.76
113	1 794	1 795	66.67	92.38	68.57
114	1 169	1 879	66.67	89.52	59.05
115	1 953	1 953	80.95	92.38	79.05
116	2 399	2 412	80.00	91.43	46.67
117	1 535	1 535	91.43	95.33	80.95
118	2 278	2 275	95.24	96.10	90.47
119	1 987	1 987	78.10	82.86	75.24

表 1 为实验检测到的 R 波和文献 [10] 中检测到的 R 波比较以及分类的准确率分析。由表 1 中可知,该算法能很好地完成 R 波检测。表 1 中 CA 准确率为交叉验证,其目的是为了得到可靠稳定的模型,比例越高,效果越好<sup>[11]</sup>。采用心电数据库中信号文件作为数据来源,将每段信号前 10% 作为训练,信号 90% 作为检测信号,选取 3 个特征参数作训练分类标准,包括 ST 段平均值,ST 段斜率和 ST 段面积。实验结果显示,CA 准确率平均值为 80.70%,训练准确率平均值为 91.83%,检测准确率平均值为 74.28%。

该算法的优点是利用离散小波变换、一阶求导和支持向量机方法,完成信号的预处理,噪声消除,特征点检测和 ST 段分类。结果显示,该算法适用于大部分心电信号。

## 4 结 语

笔者提出了一种基于离散小波变换的算法,结合一阶求导和支持向量机,完成信号的预处理,噪声消除,特征点检测和 ST 段分类。与传统方法相比,该算法简单,通用性强。Matlab 仿真表明,该算法耗时小,在 MIT-BIT 库中无较大误差,在 ST 段的分类满足其基本要求。为提高分类的准确率,需要更合适的特征向量,寻找参数作为使用的方法。

### 参考文献:

- [1] RASHKOVSKA A, AVBELJ V. Signal Processing Methods for ST Variability Assessment in ECG [C] // International Convention on Information & Communication Technology Electronics & Microelectronics. Opatija: IEEE, 2013: 331-334.
- [2] FAN C H, HSU Y, YU S N, et al. Detection of Myocardial Ischemia Episode Using Morphological Features [C] // Conf Proc IEEE Eng Med Biol Soc. Osaka [s. n.], 2013: 7334-7337.
- [3] AMAL A, RESHMI G. Cardiac Ischemia Diagnosis Using Stress ECG Analysis [C] // Third International Conference on Advances in Computing and Communications. Cochín: IEEE Computer Society, 2013: 204-207.
- [4] LIAMEDO M, MARTINEZ J P, ALBERTAL M, et al. Morphologic Features of the ECG for Detection of Stress-Induced Ischemia [C] // Computing in Cardiology. Zaragoza [s. n.], 2013: 591-594.
- [5] BUSTAMANTE C A, DUQUE S I, OROZCO-DUQUE A, et al. ECG Delineation and Ischemic ST-Segment Detection Based in Wavelet Transform and Support Vector Machines [J]. Health Care Exchanges, 2013, 64(3): 1-7.
- [6] ANTONIO FASANO, VALERIA VILLANI. Baseline Wander Removal in ECG and AHA Recommendations [C] // Computing in Cardiology Conference. Zaragoza: IEEE, 2013: 1171-1174.
- [7] SAHOO G K, ARI S, PATRA S K. ECG Signal Analysis for Detection of Cardiovascular Abnormalities and Ischemic Episodes [C] // Information & Communication Technologies. JeJu Island [s. n.], 2013: 1055-1059.
- [8] KRIMI S, OUNI K, ELLOUZE N. T-Wave Detection Based on an Adjusted Wavelet Transform Modulus Maxima [J]. International Journal of Biological Sciences, 2006, 24(1): 1-4.
- [9] POORNACHANDRA S. Wavelet-Based Denoising using Subband Dependent threshold for ECG Signals Processing [J]. Digital Signal Processing, 2008, 18(1): 49-55.
- [10] PAN J, TOMPKINS W J. A Real-Time QRS Detection Algorithm [J]. IEEE Transactions on Biomedical Engineering, 1985, 32(3): 230-236.
- [11] TANG Y, GUO W, GAO J. Efficient Model Selection for Support Vector Machine with Gaussian Kernel Function [C] // Computational Intelligence and Data Mining, 2009. CIDM09. IEEE Symposium on. Nashville, TN: IEEE, 2009: 40-45.

(责任编辑: 刘东亮)