

基于数学形态学及支持向量机的心率失常识别

刘雄飞, 晏晨伟*, 胡志坤

(中南大学 物理与电子学院, 长沙 410083)

(*通信作者电子邮箱 benaive@qq.com)

摘要:为实现对不同类型的心电图自动分析,研究并提出了一种顺序筛选极大值的R波定位算法,并采用支持向量机(SVM)进行最后的心律失常心拍识别。定位算法以数学形态学为基础,结合心电图自身特点,定义R波筛选区间,避免了传统算法中的阈值选择;定位R波峰后以R波峰为中心提取不同类型的心率失常的心拍,选择径向基(RBF)支持向量机进行识别分类。使用MIT-BIH心率失常数据库文件进行实验仿真,结果表明,算法对含不同类型心拍的心电图R波峰正确检测率较高(99.36%),学习后的SVM能有效识别早搏、房颤、束支传导阻滞、正常等不用类型心拍,总体识别率达到99.75%。

关键词:心电图;数学形态学;R波检测;心律失常分类;支持向量机

中图分类号: TP391.9 **文献标志码:** A

Arrhythmia classification based on mathematical morphology and support vector machine

LIU Xiongfei, YAN Chenwei*, HU Zhikun

(School of Physics and Electronics, Central South University, Changsha Hunan, 410083, China)

Abstract: To achieve automatic analysis for different types of ElectroCardioGraph (ECG), a sequential screening method for maximum value was brought to detect R wave, while Support Vector Machine (SVM) was used to identify arrhythmia heart beats finally. The localization algorithm based on mathematical morphology combined with characteristics of ECG defined R-wave screening interval to avoid threshold selection in traditional algorithm. After R-peaks being positioned, various types of arrhythmia heart beats were extracted with R wave crest as its center and classified by selecting Radial Basis Function (RBF) or SVM. The results of the simulation experiment on the MIT-BIH database files indicate that this algorithm acquired high relevance ratio at 99.36% for ECG with different types of heart beats. After learning, the SVM can effectively identify as many as 4 types, such as atrial premature beat, premature ventricular beat, bundle branch block and normal heart beat, the overall recognition rate is 99.75%.

Key words: ElectroCardioGram (ECG); mathematical morphology; R wave detection; arrhythmia classification; Support Vector Machine (SVM)

0 引言

心电图自动分析作为远程心电监护的重要组成部分,主要是对心电图进行预处理、QRS波检测以及心率失常识别分类。QRS波检测中首要的工作是对R波进行准确定位,R波具有幅值大的特点,目前对R波检测常用方法有差分法、数学形态学法^[1-3]及小波变换法^[4-5]。差分法、数学形态学算法简单但对高频噪声敏感;小波变换法能有效地降低噪声干扰,定位准确度高,但计算量大,不适合在一些嵌入式平台使用。心率失常是指心脏冲动的起源部位、心搏频率和节律以及冲动传导的任一异常,常见有房(室)性早搏、房颤、左(右)束支传导阻滞等。文献[6]采用的是人工神经网络对室性早搏进行识别、对训练集要求高。文献[7]结合医生识读心电图判别依据,运用逻辑判别树方法对室性异常心拍进行识别,识别率较高。

本文使用MIT-BIH数据库^[8-9]的心电文件,基于数学形态学基础,利用开运算凸显R波峰;并根据心电图周期性特点,提出一种用于定位R波峰的顺序筛选极大值法;根据注释文件,以R波峰为中心提取不同类型的心率失常的心拍,

使用支持向量机(Support Vector Machine, SVM)实现对室性早搏、房性早搏、束支传导阻滞及正常心拍的识别分类。

1 基本原理

1.1 数学形态学用于R波定位基本原理

数学形态学以集合论为理论基础,其基本思想是用具有简单形态的结构元素去度量一维或多维信号中的对应形态,达到对信号结构分析与识别的目标,其基本运算有腐蚀、膨胀、开运算及闭运算。设 f 为长度为 N 的一维离散信号, g 为所选择的一维结构元素,长度为 M ,且有 $N > M$,则腐蚀(\ominus)及膨胀(\oplus)算法定义如下:

$$\begin{cases} (f \ominus g)(n) = \min_{m=0,1,\dots,M-1} \{f(n+m) - g(m)\}; \\ n = 0, 1, \dots, N-M \\ (f \oplus g)(n) = \max_{m=0,1,\dots,M-1} \{f(n-m) + g(m)\}; \\ n = M-1, M, \dots, N-1 \end{cases} \quad (1)$$

腐蚀、膨胀算法通过不同组合可实现形态开(\oslash)和闭(\oslash)运算,即:

收稿日期:2012-11-01;修回日期:2012-12-24。

作者简介:刘雄飞(1960-),男,湖南长沙人,教授,主要研究方向:数字信号处理、心电图检测算法、车牌定位识别算法;晏晨伟(1989-),男,江西南昌人,硕士研究生,主要研究方向:心电图检测、嵌入式软件;胡志坤(1976-),男,湖北鄂州人,副教授,主要研究方向:复杂系统的在线监测与故障诊断。

$$\begin{cases} f \circ g = [(f \ominus g) \oplus g](n) \\ f \cdot g = [(f \oplus g) \ominus g](n) \end{cases} \quad (2)$$

开闭运算用于一维形态滤波时,开运算具有削平信号的波峰作用,闭运算具有能填平信号波谷的作用^[10-11];开(闭)运算后与原始信号做差值运算可得到只含波峰(谷)的信号 $PE(VE)$,如图1所示,即:

$$\begin{cases} PE = f - f \circ g \\ VE = f - f \cdot g \end{cases} \quad (3)$$

数学形态定位法能有效地消除基线漂移干扰,但对高频噪声敏感;使用数学形态法定位 R 波的经典过程^[10]是:对心电信号分别作开闭运算得到突显的 R 波峰的 PE ,及突显 Q 波和 S 波波谷的 VE ;通过调整阈值搜索对应的正负极大值对来定位 R 波。

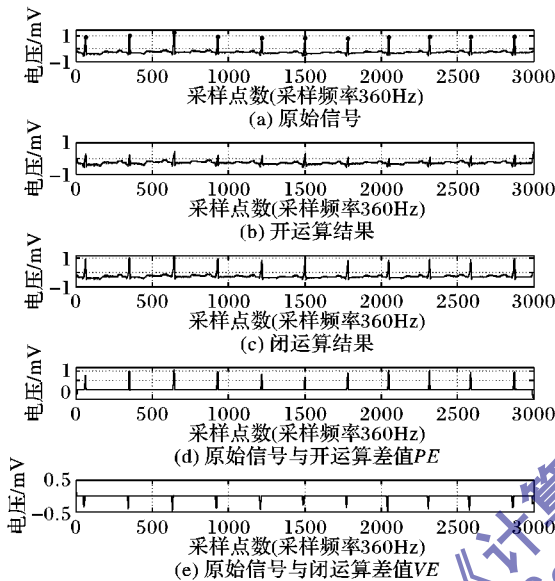


图1 T100 心电片段进行形态学运算的结果

1.2 顺序极大值筛选法

波形态运算得到的信号 PE 中除了有突显的 R 波峰外,还存在其他如 P 波、T 波及较强高频噪声干扰引起的波峰,因此需要对这些波峰或极大值进行筛选。本文根据心电节拍的周期性特点,提出及采用一种顺序极大值筛选法,即直接在信号 PE 中,在一定范围内比较筛选出极大值来定位 R 波值。

根据医学知识,正常人心拍范围为 $60 \sim 100 \text{ min}^{-1}$ 。病理性心率过快一般不超过 140 min^{-1} ;设心拍率为 $rr \text{ min}^{-1}$, f_s (Hz) 为心电信号采样频率,则定义比较范围 $area = f_s * 60 / rr$,即 RR 区间长度;其中令 $rr = 150 \text{ min}^{-1}$,算法流程如下:

步骤1 预处理:对 PE 进行差分运算得到所有极大值点的位置 P_0 ,设定阈值 pth 为 PE 前 $2 * f_s$ 点中最大极值,消除 PE 中幅值小于 $pth/9$ 的极大值点,得到预处理后的极大值点位置索引向量 P ,维数为 N 。

步骤2 极大值筛选:设定标记 $flag$ 对 P 中极大值点进行筛选,在 Matlab 中代码如下:

```
len = length(P);           得到极大值个数;
flag = zeros(1, len);      设定标记向量;
for i = 2 to len - 1;
    IF flag(i) == 0          判断是否更改标记;
        IF P(i) - P(i - 1) < area
            IF PE(P(i)) < PE(P(i - 1))
                flag(i) = 1;
            ELSE
                flag(i - 1) = 1;
```

```
END
END                          判断与前一个极值点的距离及大小;
IF P(i + 1) - P(i) < area
    IF PE(P(i)) < PE(P(i + 1))
        flag(i) = 1;
    ELSE
        flag(i + 1) = 1;
    END
END                          判断与后一个极值点的距离及大小;
END
END
P(flag == 1) = [];          删除标记为1的点;
```

步骤3 重复步骤2直至 P 维数不变,即极大值点个数不再改变;算法结束, P 即为 R 波峰所在位置。

信号 PE 中 R 波峰值较为突显,算法能剔除以 R 波峰为中心的 $area$ 范围内的极大值,从而达到检测 R 波峰的作用;对于 R 波幅值或心率变化的心电图,算法不需要进行传统的阈值调整,通过定义的比较范围 $area$ 来避免两个 R 波峰值作比较进而避免漏检。

1.3 支持向量机分类原理

支持向量机 (SVM) 是建立在统计学习理论及结构风险最小原理基础的机器学习方法,能够解决小样本、非线性及高维的模式识别问题,可用于心率失常识别^[12-13]。SVM 采用核函数 φ 将低维空间不可分的高维输入数据 x 非线性映射到高维空间中,在高维空间中构造分类超平面,使样本线性可分类。分类超平面在原低维空间可表示为分类决策面:

$$w \cdot \varphi(x) + b = 0 \quad (4)$$

决策面应满足以下约束条件:

$$y_i w_i \cdot \varphi(x_i) + b_i \geq 1 - \zeta_i; i = 1, 2, \dots, n \quad (5)$$

n 为样本数; $\zeta_i \geq 0$ 为实际应用需要引入的松弛变量。寻找最优分类超平面问题转化成二次规划问题:

$$\min \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \zeta_i \right)$$

$$\text{s. t. } y_i w_i \cdot \varphi(x_i) + b_i \geq 1 - \zeta_i; i = 1, 2, \dots, n \quad (6)$$

其中引入惩罚因子 C , 综合考虑分类器的泛化能力和分类准确率,引入拉格朗日函数^[14-15]得到优化问题(6)的对偶形式:

$$\begin{cases} \max \left(\sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \right) \\ \sum_{i=1}^n \alpha_i y_i = 0, 0 \leq \alpha_i \leq C; i = 1, \dots, n \end{cases} \quad (7)$$

$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j)$; 求解问题(7)得到 α_i , 继而得到分类函数:

$$f(x) = \text{sign} \left(\sum_{x_i \in SVs} \alpha_i y_i K(x_i, x) + b \right) \quad (8)$$

使用 SVM 进行分类的关键问题是选择核函数^[14], 使输入的非线性可分数据经变换后线性可分。

2 实验过程及结果

MIT-BIH 心率失常数据库有 52 条记录不同心率失常的心电数据,并每条心电数据有相应注释文件,注明了心率失常的类型及位置^[8]。实验首先对数据进行分析,选择了其中 20 条包含不同类别、形态明显的心电数据^[9],采用上述算法进行 R 波定位,检测结果如表1所示。表1中: N 表示正常搏动, V 表示室性早搏, A 表示房性早搏, L 表示左束支传导阻滞, R 表示右束支传导阻滞;括弧的意义为该类型节拍数 < 10。

对表1中数据分析可知,算法对正常心拍的检出率较高,对于包含V、A、L、R都能得到较好的检测率,其中信号T232部分QRS波形过于畸形成, R波峰倒置或幅值过低, T116、T205是由于心电图部分片段R波幅值过低导致漏检。

检测出R波后,根据注释文件,选择6条包含不同类别心率失常的心电数据;为提取包含P波及QRS波等信息的不同心率失常心拍,以顺序极大值筛选算法正确检出的R波峰位置为中心,向左右分别取150点,作为单个心拍样本;随机取各条数据中不同心率失常心拍的10%作为SVM训练集,其余90%作为检测集(表2),为简化分类结果,将束支传导阻滞归为一类。

表1 本文算法对MIT-BIH心电数据的检测结果

信号文件名	总心拍数	正确检出数	含心拍类型	检出率/%
100	2273	2273	N, (V), A	100.00
103	2084	2084	N, (A)	100.00
105	2572	2563	N, V	99.65
109	2532	2523	V, L	99.65
111	2124	2123	R	99.95
112	2539	2539	N, (A)	100.00
115	1953	1953	N	100.00
116	2412	2364	N, V, (A)	98.01
117	1535	1535	N, (A)	100.00
119	1987	1986	N, V	99.95
121	1863	1851	N, (V), (A)	99.36
122	2476	2476	N	100.00
123	1518	1518	N, (V)	100.00
205	2656	2628	N, V, (A)	98.95
212	2748	2746	N, L	99.93
215	3363	3301	N, V, (A)	98.17
220	2048	2043	N, V	99.76
221	2427	2376	N, V	97.90
230	2256	2255	N, (V)	99.96
232	1780	1718	A, R	96.52
合计	5146	44855		99.36

表2 用于SVM训练及检测的各类心率失常心拍

信号文件名	顺序极大值筛选算法检出的心拍个数			
	N	V	A	L&R
111	0	0	0	2072
116	2278	78	0	0
119	1543	443	0	0
212	922	0	0	1824
221	2031	343	0	0
232	0	0	1331	383
合计	6774	864	1331	4279

在Matlab下安装的libsvm工具箱中,设定好参数,用函数svmtrain进行训练,svmpredict进行检测。选择径向基(RBF)核函数 $K(x_i, x_j) = \exp(-g \|x_i - x_j\|)^2$,调整惩罚系数 C 及参数 g ,得到对检测集最佳分类结果如表3所示($C = 10000, g = 0.0001$),训练集回验率为100%。训练好的SVM分类器能对房性早搏、室性早搏、左(右)束支传导阻滞进行有效的识别,对检测集总体识别正确率为99.75%。

3 结语

本文在数学形态学基础上,提出了一种顺序极大值筛选法来提取R波峰,适用于心率变化、及包含不同类型心拍的心电图检测,实验总体检测率为99.36%。

利用支持向量机解决小样本、高维、非线性的模式识别问题的优势,有效避免了其他分类方法需要准确检测心电信号多个特征点的问题,提高了分类识别的准确率。

采用径向基SVM支持向量机对不同心率失常心电数据进行识别分类,总体识别率为99.75%,能准确识别室性早搏、房性早搏等不同心率失常心拍。

表3 SVM对不同类型心拍识别结果

心拍类型	检测心拍数	正确识别心拍数	识别率/%
N	6097	6090	99.89
V	708	707	99.86
A	1198	1181	98.58
L&R	3852	3847	99.87
合计	11855	11825	99.75

参考文献:

- [1] ZHANG FEI, LIAN YONG. QRS detection based on multiscale mathematical morphology for wearable ECG devices in body area networks [J]. IEEE Transactions on Biomedical Circuits and Systems, 2009, 3(4): 220-228.
- [2] 田学隆, 闫春红, 俞亚青, 等. 基于小波变换的R波检测算法[J]. 生物医学工程学杂志, 2006, 23(2): 257-261.
- [3] 陈永利, 段会龙. 基于数学形态学和信号包络提取的QRS波检测[J]. 中国生物医学工程学报, 2007, 26(3): 332-336.
- [4] BANERJEE S, MITRA M. ECG feature extraction and classification of anteroapical myocardial infarction and normal subjects using discrete wavelet transform [C]// 2010 International Conference on Systems in Medicine and Biology. [S.L.]: IEEE, 2010: 55-60.
- [5] 杨思军, 郝继飞, 闫文杰, 等. 小波滤波与QRS波检测[J]. 计算机工程与应用, 2011, 47(12): 239-241.
- [6] 熊平, 徐敏, 雷英俊. 基于小波变换和BP神经网络的室性早搏(PVC)识别[J]. 中国医学物理学杂志, 2010, 27(2): 1762-1765.
- [7] 李坤阳, 胡广书. 基于心电图分析的心律失常分类[J]. 清华大学学报: 自然科学版, 2009, 49(3): 416-418.
- [8] MOODY G, MARK R. The MIT-BIH arrhythmia database [DB/OL]. (2005-07-22) [2012-10-01]. <http://ecg.mit.edu/>.
- [9] 宋喜国, 邓亲恺. MIT-BIH心率失常数据库的识读及应用[J]. 中国医学物理学杂志, 2004, 21(4): 230-232.
- [10] 刘少颖, 卢继来, 胡广书, 等. 基于数学形态学和小波分解的QRS波群检测算法[J]. 清华大学学报: 自然科学版, 2004, 44(6): 852-855.
- [11] TADEJKO P, RAKOWSKI W. Mathematical morphology based ECG feature extraction for the purpose of heartbeat classification [C]// CISIM '07: Proceedings of the 6th International Conference on Computer Information Systems and Industrial Management Applications. [S.L.]: IEEE, 2007: 322-327.
- [12] GHOSH D, MIDYA B L, KOLEY C, et al. Wavelet aided SVM analysis of ECG signals for cardiac abnormality detection [C]// 2005 Annual IEEE INDICON. Piscataway, NJ: IEEE Press, 2005: 9-13.
- [13] BAKHSIPOUR A, POOYAN M, MOHAMMADNEJAD H, et al. Detection with ECG analysis using wavelet transform and support vector machines [C]// Proceedings of the 17th Iranian Conference of Biomedical Engineering, Computing in Cardiology. Piscataway, NJ: IEEE Press, 2010: 405-408.
- [14] 曹玉珍, 李广, 范增飞. 基于小波变换特征提取的支持向量机心搏分类研究[J]. 天津大学学报, 2007, 40(7): 811-815.
- [15] 杨煦, 张跃. 基于支持向量机的心律失常识别方法[J]. 计算机工程与设计, 2007, 28(18): 4442-4445.