

文章编号: 2095-2163(2020)06-0051-05

中图分类号: TP305

文献标志码: A

基于12导联的心电特征可视化分析

王佳琪¹, 袁永峰²

(1 东北林业大学 信息与计算机工程学院, 哈尔滨 150040; 2 哈尔滨工业大学 计算机科学与技术学院, 哈尔滨 150001)

摘要: 近年来,心脑血管患病数超过2.9亿,死亡率仍处在上升阶段,防治心血管病刻不容缓,心电技术尤为重要。临床诊断通常采用12导联系统,同时配合可视化技术可以较为直观地反映心脏的信号传播活动。利用12导联ECG可视化技术分析心电特征对于不同疾病的差异程度,首先利用GQRS-极值法对心电信号进行波峰识别,其次将提取到的心电特征进行统计分析,最后通过信息增益的方法分析各特征之间的差异。对于疾病异常情况,可以通过本文的心电特征可视化分析技术快速反映心脏不同部位的病变程度。

关键词: 12导联; 疾病; 心电特征; 差异程度; 可视化

Visual analysis of ECG characteristics based on 12 leads

WANG Jiaqi¹, YUAN Yongfeng²

(1 College of Information and Computer Engineering, Northeast Forestry University, Harbin 150040, China;

2 School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

[Abstract] In recent years, the number of cardiovascular and cerebrovascular diseases has exceeded 290 million, and the mortality rate is still in the rising stage. The prevention and treatment of cardiovascular diseases cannot be delayed. Therefore, electrocardiography is particularly important. Clinical diagnosis usually adopts 12-lead system, together with visualization technology can more intuitively reflect the heart's signal propagation activity. In this paper, the 12-lead ECG visualization technology is used to analyze the degree of difference in ECG characteristics for different diseases. First, the GQRS-extreme method is used to identify the peaks of the ECG signal, followed by statistical analysis of the extracted ECG characteristics, and finally the information gain. The method analyzes the differences between features. For abnormal conditions of the disease, the visual analysis technology of ECG characteristics in this paper can quickly reflect the degree of lesions in different parts of the heart.

[Key words] 12 leads; disease; ECG characteristics; degree of difference; visualization

0 引言

随着社会的不断发展,人口老龄化及城镇化进程加速,我国心脑血管疾病率及死亡率加速增长,心脑血管病现患病人数2.9亿,2016年心脑血管死亡率仍居首位,高于肿瘤及其他疾病,因此心电检测尤为重要。现代检测心电的手段通常采用12导联技术,12导联的电极位置分布在四肢、前胸和腋下周围,帮助诊断心脏周围异常部位。同时由于12导联技术的发展,心电图可视化技术也随之出现,传统的可视化技术是采用面绘制的方法对体表电位进行可视化, Liberios A等^[4]通过体表电位仿真技术(BSPM)研究房性心律失常的信号传播活动,同时也出现了针对心脏三维建模的可视化技术, Ramanathan等^[5]通过ECGI对正常成人的心脏进行了电生理研究,利用ECGI技术对健康成年人心脏

心外膜电位、心电图及等时线进行无创重建,可显示正常心室激动特征,尤其可以反应心脏整体的复极特点,更加直观的观察心脏的电活动。近年的心电散点图、瀑布图技术可以直观的反映心脏的节律变化及心电特征的改变^[6-7],从而快速识别早搏、房颤、冠心病、心绞痛等。

而本文直接对不同疾病的心电特征进行可视化研究。首先选择PhysioBank开源生理数据库中5种疾病类型的数据,其次将心电特征从GQRS-极值法处理后的心电信号提取出来,分析不同疾病的特征值分布,最后,通过信息增益的方法比较不同疾病下的特征差异程度,如图1所示。心电特征差异反映了心脏不同位置的病理,当出现异常疾病类型时,本文的可视化研究可以快速诊断心脏部位的病变。

基金项目: 国家自然科学基金面上项目(61572152)。

作者简介: 王佳琪(1998-),女,本科生,主要研究方向:机器学习与深度学习、数据可视化;袁永峰(1979-),男,博士,副教授,硕士生导师,主要研究方向:计算心脏学、虚拟心脏建模与仿真、科学计算与可视化。

通讯作者: 袁永峰 Email: yongfeng.yuan@hit.edu.cn

收稿日期: 2020-04-04

1 资料

为研究 12 导联心电图的特征,从 PhysioBank 数据库中选择了共 289 440 个样本。其中,所有数据库的数据采集均为 2 min,室上性心律失常(SA)和正常窦性心律(normal)数据库采样频率(f_s)为 128 Hz,各 30 720 个样本。恶性室性心律失常(MVE)、BIDMIC 充血性心力衰竭(BIDMIC)和房颤

(AF)数据库采样频率为 250 Hz,各为 60 000 个样本,部分癫痫发作后心率震荡(PIHRO)数据库采样频率为 200 Hz,共 48 000 个样本。所有这些数据都是从波士顿贝思医院收集的,该医院是心电图信号的最大数据收集之一,并为人类受益人的工作进行记录^[9]。

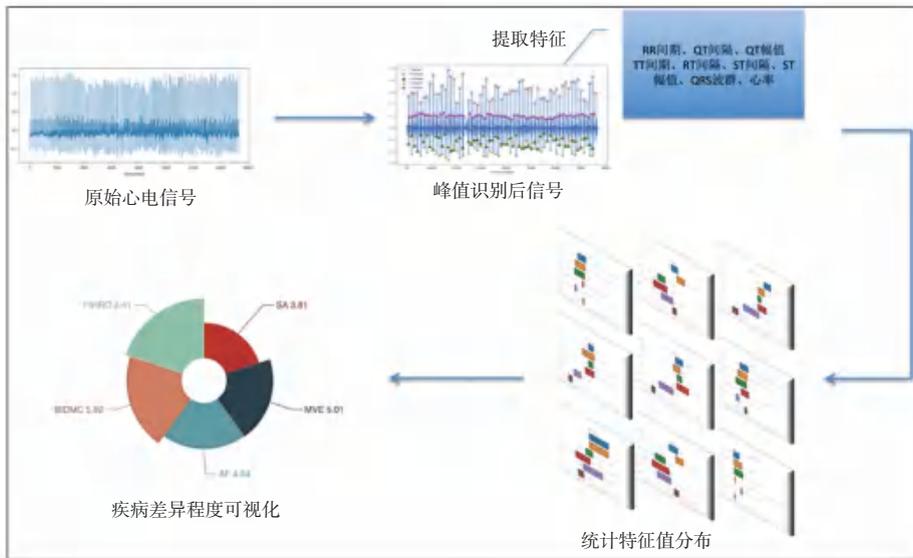


图 1 心电特征可视化研究

Fig. 1 Visualization study of ECG characteristics

2 方法

在得到不受任何干扰的心电信号、提取到大量准确的特征信息,首先对原始心电信号进行预处理。原始的 ECG 信号通常含有基线漂移、肌电、工频干扰的噪声信息,因此利用离散小波变换方法对心电信号进行去噪处理,小波变换分为小波分解和重构,离散小波分解是将噪声的信号进行多尺度的分解,选择小波函数的低通滤波和高通滤波将信号分解为低频和高频部分,小波重构是小波分解的逆过程,将细节信号和过滤后的最终逼近信号进行重组。本实验根据心电信号的特点,采用 db5 小波基函数,分解层数为 5 层,最终得到去噪后的心电信号。

在一个正常的心动周期中,窦房结形成起搏后,迅速将冲动通过传导系统传至心脏各部形成心肌整体的电活动,心肌进而形成机械性收缩,心肌激动的电流传导至体表,从而出现心电图中的各个特征波形。一个典型窦性心律正常的 ECG 波形(图 2)包括以下特征波:一个 P 波,一个 QRS 波群(Q 波,R 波,S 波),一个 T 波,以及在 50%~75%的 ECG 中可能见到的 U 波。

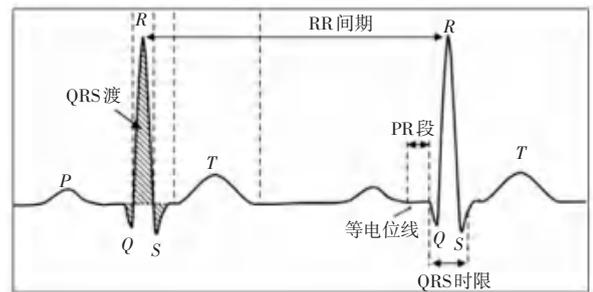


图 2 心电信号波形图

Fig. 2 ECG signal waveform

因此,想要得到心电特征,需要进一步识别心电信号的各个波峰,采用 QRS-极值法。由于 QRS 波群具有明显的特征,通常先提取 ECG 信号中的 R 波作为波形的划分依据,因此采用 QRS 算法识别,QRS 算法为 PhysioBank 提供的 WFDB(波形数据库)软件包中检测 R 波的算法,对信号进一步去噪并进行低通滤波,之后用滤波信号卷积 QRS 匹配滤波器,使用计算出的信号长度运行学习阶段,学习检测 QRS 和与主检测阶段相同的非 QRS 峰,不保存 QRS 位置,并记录 QRS 阈值,运行检测时比较 RR

间隔与 QRS 间隔从而判断 R 峰位置, Q 峰的位置在 R 峰前三分之一周期 (f_s) 样本内的第一个极值拐点, 对应的 S 峰位置在 R 峰后三分之一周期内的第一个极值拐点, T 峰在该周期与下一周期的 RR 间隔内前三分之二的样本内检测, 位置在该周期 S 峰位置后的第一个极值拐点。同时, 判断是否出现倒置情况, 因此在识别时分别求出极大值和极小值, 判断极大值与极小值距离基准电压 0 V 的幅值大小筛选 T 波是否倒置, 确定 T 波位置。

确定心电图的波峰位置后, 选择 RR 间期 (RR_

time)、Q-T 间隔 (QT_time)、Q-T 幅值 (QT_value)、S-T 间隔 (ST_time)、S-T 幅值 (ST_value)、R-T 幅值 (RT_value)、T-T 间隔 (TT_time)、QRS 间期 (QRS_time)、心率 (Heart_rate), 这些心电特征有利于患者的临床诊断、可以大幅度检测出不同疾病下的差异, 其中的心率为脉搏或心脏听诊 1min 心跳的次数, 计算方法为:

$$Heart_rate = 60 * f_s / RR_time,$$

最后将以上得到的 9 个特征对应的数据值进行统计, 如图 3 所示。

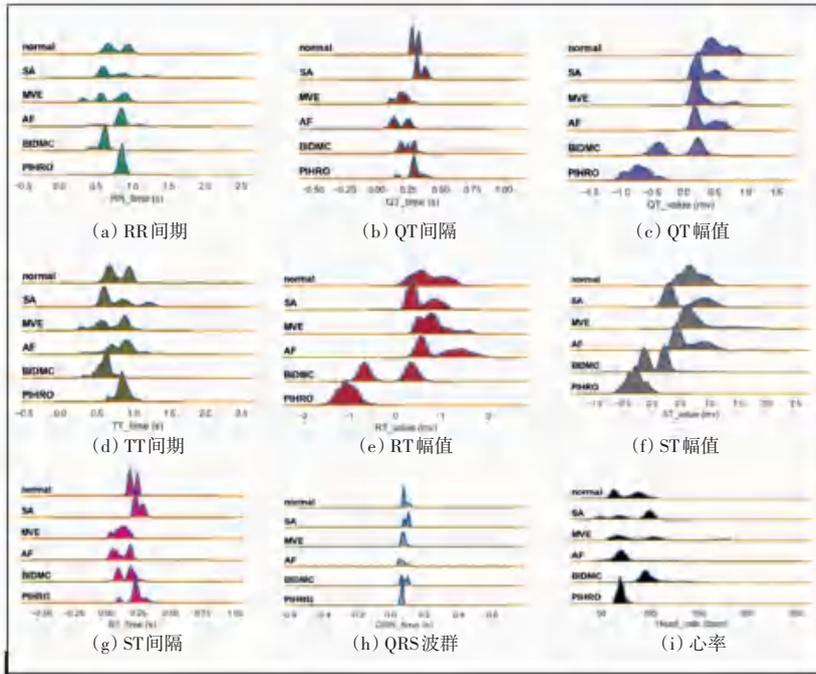


图 3 各特征值分布

Fig. 3 Distribution of eigenvalues

根据不同特征值的分布特点, 可以计算出不同疾病针对同一特征的差异程度。在此之前, 传统的差异性分析方法为相关系数、协方差矩阵、回归分析法等, Mingjing Y^[17] 也运用相关熵判断 R 峰与 RR 区间的相关性。熵表示信息在动态系统中的混乱程度, 当各特征之间不能通过单纯的数字进行差异性比较时, 则需要利用信息增益的方法。信息增益即为在一定条件下, 信息不确定性减少的程度。本研究采用信息增益的方法判断各特征之间的差异性, 比较各特征的差异性大小。收集的 Physionet 生理数据库中各类型不同特征组成的特征向量为:

$$\vec{x} = \{x_1^T, x_2^T, x_i^T, \dots, x_n^T, type\}.$$

式中, $type$ 为疾病类型, x_i 为该类型第 i 个特征值组成的特征值向量。

首先将每个特征值向量进行线性归一化处理:

$$x' = (x' - \min(x_i)) / (\max(x_i) - \min(x_i)).$$

其中, x' 为特征值向量中归一化处理后的值, 归一化后的特征值向量均为标量形式。类型值 $type$ 的熵定义式为:

$$H(type) = - \sum_{i=1}^m p_i \log p_i.$$

其中, p_i 为第 i 种疾病类别在特征向量中对应的概率。得到的熵值越大, 随机变量 $type$ 的不确定性越大。信息增益表示得知特征 X_i 的信息使得类别 Y 的信息不确定性减少程度, 信息增益为数据集 D_k 的经验熵与特征 x_i 给定条件下 D_k 的条件熵之差:

$$g(D_k, x_i) = H(D_k) - H(D_k | x_i).$$

其中, $|D_k|$ 为第 k 个疾病特征向量与正常窦性心律 (normal) 特征向量组成的特征向量矩阵。特征 x_i 对数据集的条件熵为:

$$H(D_k | x_i) = - \sum_{i=1}^c \frac{|D_i|}{|D_k|} \sum_{u=1}^2 \frac{|D_{iu}|}{|D_i|} \log_2 \frac{|D_{iu}|}{|D_i|}$$

其中, $|D_i|$ 为样本容量, 根据心电特征 x_i 的取值将 D_k 划分为 c 个子集, 记子集 D_i 中属于第 u 类别的样本集合为 D_{iu} 。将各疾病类型的所有特征进行信息增益计算, 得到的可视化可以比较不同疾病特征差异(图4)。同时在单一疾病条件下, 各特征差异程度不同, 对疾病的影响程度也不同(图5)。

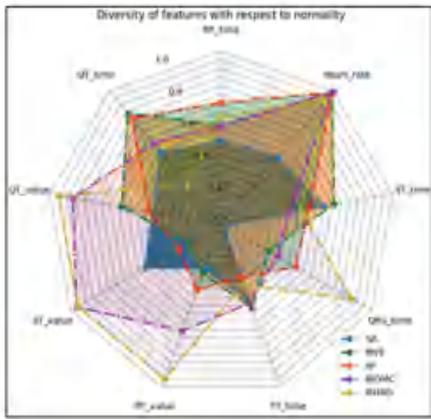


图4 疾病特征差异比较

Fig. 4 Comparison of differences in disease characteristics

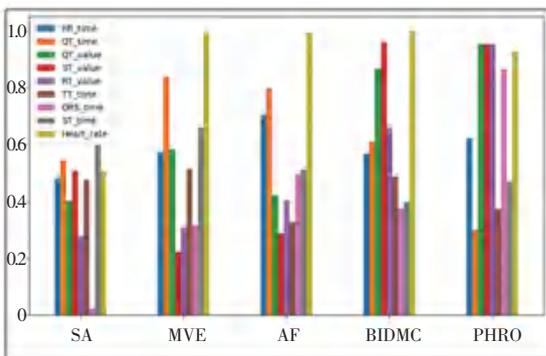


图5 不同疾病下各差异值分布

Fig. 5 Distribution of differences in different diseases

3 方法的性能

特征提取的前提是对波峰的正确检测, 波峰识别算法准确率的高低直接影响特征值提取的正确性。判断波峰的准确性通常检测灵敏度和准确率指标。表1中定义参数如下。

- * 真阳性(TP): 正确检测到的正确信号数;
- * 假阴性(FN): 没有检测到的信号数;
- * 假阳性(FP): 误报信号数;
- * 真阴性(TN): 正确检测到错误的信号数;
- * 灵敏度: $Sensitivity = TP / (TP + FN)$;
- * 准确率: $Accuracy = (TP + TN) / (TP + TN + FP + FN)$ 。

表1 各算法指标比较

识别算法	灵敏度	准确率
GQRS-极值	98.56	96.01
Pan Tompkins ^[15]	99.83	86
HistogramApproach ^[15]	99.86	99.8
Linear prediction ^[15]	96.9	93.2
Adaptive threshold ^[15]	99.55	99.70

由表1可看出, 本文采用的GQRS算法灵敏度比Linear prediction算法高, 准确率在96%以上, 识别到的心电波峰准确率较高, 为特征可视化奠定了基础。

4 结束语

本文针对不同疾病的12导联心电图进行心电特征的可视化分析。提取到的心电特征可视化统计图2.2, 可以清楚的看到不同疾病下各特征的分布特点, 特征数值在某一区间内越集中, 特征值波形越高耸, 不同疾病下特征值的区间存在差异。之后通过信息增益的方法进一步分析各特征之间的差异程度, 雷达可视化分析图2.4可以看出室上性心律失常的心电特征较其他疾病类型不明显, TT段与ST段与其他特征相比差异较大; 部分癫痫发作后心率震荡的差异程度最高, 其中QT段、ST段、RT段的幅值与QRS波群差异最大, 异常情况较其他疾病更为明显, ST段的差异可以证明该疾病大幅导致心肌梗死, QT波的差异证明心率的波动对该疾病有很大影响, 更加直观的观察不同特征对哪种疾病的影响最大, 同时统计图2.5对比了同一疾病不同特征对疾病的影响程度。对于异常心脏疾病类型情况, 本文的心电特征可视化分析可以直观快速地观察到不同心电特征的影响, 进而根据特征差异判断出心脏不同部位的病变程度。

参考文献

- [1] 胡盛寿, 高润霖, 刘力生.《中国心血管病报告2018》概要[J]. 中国循环杂志, 2019, 34(3): 209-220.
- [2] 白凤英. 12导联动态心电图诊断心律失常的临床价值分析[J]. 中西医结合心血管病杂志, 2017, 2017(33): 35.
- [3] KOCH H, RICHTER A, KURSTEN R, et al. Composition of approximated body-surface-potential-maps by utilizing a common 12-lead-ECG device [J]. IEEE Transactions on Biomedical Engineering, 2005, 52(3): 463-470.
- [4] LIBEROS A, PEDRON-TORRECILLA J, RODRIGO M, et al. Body surface potential propagation maps during macroreentrant atrial arrhythmias. A simulation study [C]// Computing in Cardiology Conference. IEEE, 2013.

(下转第58页)