

一种基于特征提取和多模板匹配的心律失常检测算法 An Arrhythmia Detection Algorithm Based on Feature Extraction and Multiple Template Matching

崔永华¹, 梁正友²

CUI Yong-hua¹, LIANG Zheng-you²

(广西大学计算机与电子信息学院, 广西南宁 530004)

(School of Computer and Electronic Information, Guangxi University, Nanning, Guangxi, 530004, China)

摘要: 基于特征提取和多模板匹配的心律失常检测算法通过样本数据库创建模板库, 对模板库进行特征提取以创建特征库, 同时建立索引与模板库关联, 然后使用特征库中的模板与待检测心电信号进行粗匹配运算, 当粗匹配的相关系数大于预设阈值后, 再利用索引加载相关模板进行细匹配运算。使用 MIT-BIH 数据库的数据验证基于特征提取和多模板匹配的心律失常检测算法的效果理想。

关键词: 检测系统 算法 特征提取 相关系数 多模板匹配

中图分类号: TP301.6, R540.7 **文献标识码:** A **文章编号:** 1002-7378(2012)01-0007-04

Abstract: The paper introduces an arrhythmia detection algorithm based on feature extraction and multiple template matching. The algorithm creates the template database from the sample database. In order to create the feature database, the algorithm extracts the features of the template database and creates the indexes associated with the corresponding template database. Then, templates of the feature database and the Electrocardiography (ECG) are used to be inspected to execute the rough matching algorithm. If the cross-correlation of the rough matching is bigger than the predefined threshold, the indexes are used to load the corresponding template and execute the fine matching algorithm. The experimental results, verified by using MIT-BIH database, show the algorithm can guarantee the detection accuracy.

Key words: detecting system, algorithm, feature database, cross-correlation, multiple template matching

在现代医学中, 心电信号 (ECG) 的自动检测和分析对于心血管疾病的分析和诊断有着十分重要的意义, 心电信号是心血管疾病的主要依据, 因此, 心电图的检测与分析得到了广泛的研究。

长期以来, 国内外的同行在心电信号的自动检测和分析领域进行了广泛而深入地研究, 并提出了一系列的算法, 如阈值检测法、小波分析法、神经网络法、模板匹配法及支持向量机聚类算法等。

收稿日期: 2012-01-30

修回日期: 2012-02-05

作者简介: 崔永华 (1978-), 男, 硕士研究生, 主要从事高性能计算和网络系统研究。

文献 [1] 将小波分析用在 QRS 波的检测上已取得比较满意的结果。文献 [2] 使用小波阈值法分析了心律失常, 但是效果不是很理想。文献 [3] 设计了模板来分析 QRS 复波并取得了良好的效果。文献 [4, 5] 首先使用模板匹配进行心电信号的检测。文献 [6] 介绍使用模式匹配来检测心肌梗死的心律失常算法。文献 [7] 对比不同模板匹配算法检测心电信号的性能。文献 [8] 使用相关分析法对异常 ECG 信号进行了特征提取。文献 [9] 使用互相关系数和多模板匹配进行心律失常的检测。文献 [10] 采用峰谷检测和互相关系数来进行室性心动过速的检测。本文提出一种

基于特征提取和多模板匹配的心律失常检测的新方法,该方法采用分级匹配的思想,首先使用特征库中的模板与待检测心电信号并进行粗匹配运算,当粗匹配的相关系数大于预设阈值后,再利用索引加载相关模板进行细匹配运算。此方法可以大大提高心电信号的检测速度和检测精度。通过使用 MIT-BIH 数据库的数据进行验证,本方法取得了非常理想的效果。

1 心律失常检测系统的总体框架

心律失常检测系统总体框架主要包括模板构造和模板匹配两大模块,每一大模块又包括许多子模块和子程序(图1)。模板构造模块的主要工作是完成模板队列的构造。该模块包含信号分解、聚类处理和特征提取3个子模块。信号分解子模块主要负责将原始心电信号按一个心电周期进行分解;聚类处理子模块主要负责将分解后的 ECG 信号进行聚类处理以创建模板库;特征提取子模块主要是对模板库进行特征提取以创建特征库。模板匹配模块主要工作是进行待检测心电信号的检测,以分辨出是正常心电信号还是心律失常心电信号。该模块包含粗匹配子模块和细匹配子模块。粗匹配子模块用来将待检测心电信号与特征库中的模板进行匹配,只有当相关系数预设阈值后,才进行下一步的细匹配操作;细匹配子模块主要用来将待检测心电信号与模板库中的模板进行匹配,当相关系数大于预设阈值时,就可以认定匹配完成。

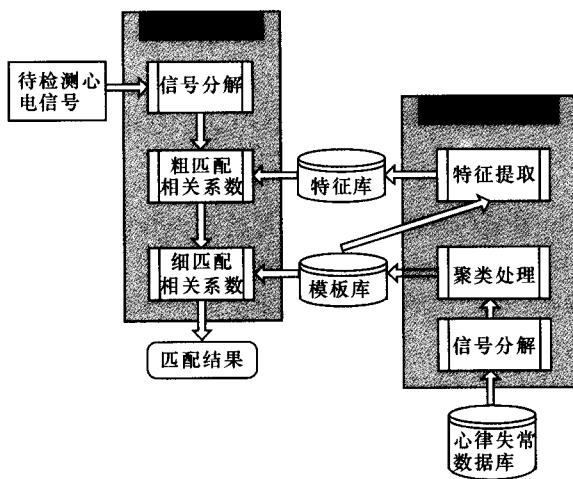


图1 心律失常检测系统的总体框架

2 模板匹配算法

2.1 构造模板

构造 ECG 模板用来与未知 ECG 信号进行匹

配。模板数据来自于 MIT-BIH 心律失常数据库,该心律失常数据库包含 48 组经过专家注解的数据集,每组数据集大小约 20MB,长约 30min,包含大约 5000 个心电周期。我们将每组数据集按照每个心电周期进行截取,全部 48 组数据集共能截取大约 240000 个心电信号,这些一个周期的心电信号我们称之为候选模板。这些候选模板数量庞大,并且含有冗余,不能直接用来与待检测心电信号进行匹配。我们需要对这些候选模板进行分组,每组包含若干个同类型的候选模板,这些被选定的候选模板称之为模板。每组所包含的模板数不能太少也不能太多,太少会影响心电信号的检测精度,太多可能会导致冗余和浪费资源。

2.1.1 构造模板库

为了构造模板库,需要对候选模板进行分组,采用 K-Medoids 聚类算法是一个较好的分组方法。但是该算法有两个方面会影响模板的正确分组。首先,初始聚类中心是随机产生的,这可能会导致局部最大值。其次,簇数量的选择可能会影响算法的效率。为克服该不足,我们对 K-Medoids 算法进行了部分初始化操作,采用指定初始中心和族数的方法,使其能够适用于候选模板的正确、快速分组。算法如下:

- (1)初始化 9 类常见心电信号模板,这 9 个模板由专家认定,作为每组的初始聚类中心。
- (2)初始化聚类分组数 $k=9$ 。
- (3)将每个数据集截取为多个只包含一个心电周期的 ECG 心电信号,也称为候选模板。
- (4)对于每个分组数 k ,指定第 k 个已初始化的心电模板作为分组的初始聚类中心。
- (5)对于每个模板 p ,计算它与各个聚类中心的相识度,将它分配给与其最相似的聚类。
- (6)对于每个模板分组,计算该分组的新的聚类中心。
- (7)如果不再有模板能改变每组的聚类结构,则分组完毕。否则重复步骤 5 到 6。

分组完毕后,选取每组 10% 个与聚类中心最相近的模板作为该组的模板队列,将其保存到模板数据库中。

2.1.2 特征库构造

生成模板库后,对模板库中的模板进行特征提取,提取每个模板的 P 波、QRS 波、T 波等波形的特征信息。特征信息的提取采用连续小波变换,函数表示如下:

函数 $f(x) \in L^2(R)$ 的连续小波变换为:

$$Wf(a, \tau) = f(t) \cdot \psi(t) =$$

$$\frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi * \left(\frac{t-\tau}{a} \right) dt$$

式中: a 为尺度因子, τ 为时移因子, $\psi_a(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t}{a}\right)$

为母小波在尺度 a 上的伸缩。提取到模板的特征信息后, 将其保存到特征库中, 并与模板库中的模板建立索引。

2.2 模板匹配算法

2.2.1 模板匹配衡量标准

使用互相关系数 (cross-correlation) 作为模板匹配相似程度的衡量标准。互相关函数定义为:

$$R_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t)y(t+\tau)dt$$

互相关系数定义为:

$$\rho_{xy}(\tau) = \frac{R_{xy}(\tau) - \mu_x \mu_y}{\sigma_x \sigma_y}, \rho_{xy}(\tau) \leq 1$$

互相关系数的取值范围是 $[-1, 1]$, 当互相关系数的绝对值等于 1 或接近 1 时, 表明待检测心电信号与模板的匹配程度极高, 可以认定待检测心电信号为模板所代表的心律失常; 当互相关系数的绝对值等于 0 或接近 0 时, 表明待检测心电信号与模板的匹配程度极低, 可以排除待检测心电信号为模板所代表的心律失常。

2.2.2 模板匹配步骤

模板匹配的任务是给定一个待检测的心电信号 X , 找出其匹配的模板。主要的匹配过程如下: (1) 对 X 进行信号分解得到其周期信号; (2) 粗匹配: 将 X 的周期信号与特征库的特征模板进行匹配, 获得比较相近的特征模板; (3) 细匹配, 根据 (2) 获得的特征模板, 在模板库中获取其对应的模板, 并与检测心电信号进行相关性计算, 当计算结果大于预设阈值时, 即可认为待检测心电信号是正常心电信号或是心律失常心电信号。模板匹配算法如下:

输入: 待检测心电信号 x 。

输出: 异常心电周期及种类。

(1) 初始化特征库中的模板 $A = \{A_1, A_2, \dots, A_n\}$, $A_1 \dots A_n$; 初始化模板库中的模板 $B = \{B_1, B_2, \dots, B_n\}$ 。

(2) 将待检测心电信号 x 按周期分为: $X = \{X_1, X_2, \dots, X_k\}$ 。

(3) 粗匹配操作: 将 X_i 与 A_j 进行模板匹配, $1 \leq i \leq k, 1 \leq j \leq n$; 即使用特征库中的模板与待检测心电信号进行匹配操作。

(4) 如果粗匹配的相关系数大于等于 0.80 时, 则根据建立的索引号加载对应的模板; 否则, 如果相关系数小于 0.80 时, 则转到 (7)。

(5) 细匹配操作: 将 X_i 与 B_j 进行模板匹配, $1 \leq i \leq k, 1 \leq j \leq n$; 即使用模板库中的模板与待检测心电信号进行匹配操作。

(6) 如果细匹配的相关系数大于等于 0.95 时, 认为该心拍的类型以及形态与匹配的模板一致, 匹配结束; 否则, 如果相关系数小于 0.95 时, 则转到 (7)。

(7) 判定是否为新模板, 如果判定结果为新模板则构造模板并加入模板库, 同时对新模板进行特征提取, 并更新特征提取后的模板队列和正常模板队列, 然后转到 (3); 否则, 直接转到 (3)。

3 算法验证

从 MIT-BIH 的心律失常数据库中选取 10 组数据进行验证, 每组数据的采样频率为 360 赫兹, 持续时间大于 30min。算法对室性早搏 (PVC) 和房性早搏 (APB) 的检测效果分别如表 1 和表 2 所示。

表 1 算法检测 PVC 的实验结果

编号	总数 (PVC)	检出数	漏检数	错检数	检测率 (%) *	平均错误率 (%) *
105	41	40	1	0	97.56	1.22
106	520	514	6	7	98.85	1.25
116	109	106	3	0	97.25	1.38
200	826	815	11	13	98.67	1.45
207	105	102	3	1	97.14	1.91
合计	1601	1577	24	21	98.50	1.41

*: 检出率 = 检出数 ÷ 总数 × 100%; 平均错误率 = 1/2 × {[错检数 ÷ (检出数 + 错检数)] + [漏检数 ÷ (检出数 + 漏检数)]} × 100%。

表 2 算法检测 APB 的实验结果

编号	总数 (APB)	检出数	漏检数	错检数	检测率 (%) *	平均错误率 (%) *
100	33	32	1	1	96.97	3.03
118	96	93	3	4	96.88	3.62
124	2	2	0	0	100.00	0
200	30	29	1	1	96.67	3.33
207	107	103	4	3	96.26	3.28
合计	268	259	9	9	96.64	3.36

*: 检出率 = 检出数 ÷ 总数 × 100%; 平均错误率 = 1/2 × {[错检数 ÷ (检出数 + 错检数)] + [漏检数 ÷ (检出数 + 漏检数)]} × 100%。

分析表 1 和表 2 的实验结果可以发现, 本文算法对室性早搏有很高的检测率, 对房性早搏有较高的检测率, 同时对这两种心律失常检测的平均错误率也比较低。漏检、错检主要是因为某些正常的心电信号受到噪声影响, 导致与模板队列的模板相似, 从而造成误匹配。

本文检测算法采取聚类算法进行模板的分类,在分类过程中对每组的聚类中心进行标准模板的初始化,由此得到精确度很高的模板队列,并且本文算法还采用两级模板匹配方法,匹配准确度有了很大的提高。相对于文献[2,3]中的算法(表3),本文算法具有更高的灵敏度和更小的平均错误率,算法更准确,并且在漏判的抑制上具有绝对的优势。

表3 3种检测方法对PVC和APD的检测结果比较

检测方法	对PVC的检测		对APB的检测	
	检测率(%)	平均错误率(%)	检测率(%)	平均错误率(%)
文献[2]	82.35	4.82	82.35	6.53
文献[3]	91.24	2.58	91.24	4.37
本文算法	98.50	1.41	96.64	3.36

总之,基于特征提取和多模板匹配的心律失常检测算法克服了传统模板匹配算法的不足,结合特征和两级模板匹配模式,提高了算法的运算效率和识别率,具有一定的使用价值。然而,该算法对于复杂的心律失常心电信号的检测识别率较低,仍然需要结合实际,进行更加深入的研究。

参考文献:

- [1] Li Cuiwei, Zheng Chongxun, Tai Changfeng. Detection of ECG characteristic points using wavelet transforms [J]. IEEE Trans on Biomed Eng, 1995, 42(1): 21-29.
- [2] 胡鹏,张永红,张菊鹏,等. 基于小波变换的心律失常判别算法[J]. 北京生物医学工程, 2003, 22(1): 25-28.
- [3] 段会龙,冯靖玮,洪玮. 基于模板匹配和特征识别相结合的心室期前收缩波形分类算法[J]. 航天医学与医学工程, 2002, 15(2): 98-102.
- [4] Silka M J, Ranville S, Jenkins J M, et al. Template matching analysis of unipolar and bipolar electrograms in the differentiation of ventricular tachycardia and fibrillation from sinus rhythm[C]. Computers in Cardiology 1992, 1992: 223-225.
- [5] Caswell S A, Kluge K S, Chiang C M J, et al. Pattern recognition of cardiac arrhythmias using two intracardiac channels [C]. Computers in Cardiology 1993, 1993: 181-184.
- [6] Bousselijot R, Kreiseler D. ECG signal analysis by pattern comparison [C]. Computers in Cardiology 1998, 1998: 349-52.
- [7] Schuckers S C, Xueyan X, Schuckers M E, et al. Ventricular arrhythmia detection using time-domain template algorithms [C]. Bioengineering Conference 1998, 1998: 21-23.
- [8] Ramli A B, Ahmad P A. Correlation analysis for abnormal ECG signal features extraction [C]. Telecommunication Technology 2003; NCTT 2003 Proceedings, 4th National Conference on 2003, 2003: 232-237.
- [9] Chin Fook Joo, Fang Qiang, Zhang Tao. A fast critical arrhythmic ECG waveform identification method using cross-correlation and multiple template matching [C]. Argentina; 32nd Annual International Conference of the IEEE EMBS Buenos Aires, 2010.
- [10] Chin Fook Joo, Fang Qiang, Cosic Irena. A computationally light-weight real-time classification method to identify different ECG signals [C]. Bioelectronics and Bioinformatics (ISBB), 2011 International Symposium on Digital Object Identifier; 10. 1109/ISBB. 2011. 6107703 Publication Year; 2011, 2011: 287 - 290.

(责任编辑:邓大玉)

IBM 科学家用 12 个原子制成世界最小磁存储器

目前的磁盘驱动器需要动用上百万个原子来存储一个比特位。传统磁盘使用铁磁材料进行数据存储,一个比特位内所有原子的自旋方向相同,用同一磁化方向来表示“0”或“1”。然而,铁磁材料在尺寸方面遇到的最大障碍是,当缩小到原子级别时,相邻比特位之间会产生交互作用,一个比特位的磁化会影响到相邻比特位。为克服这一缺陷,IBM 的科学家利用扫描隧道显微镜在原子层级对一组 12 个反铁磁材料原子进行操作,在低温下使其存储了一个比特位的信息,并保持了若干小时。利用这些原子固有的交变自旋方向,科学家展示了排列得比以往紧密得多的磁性比特位。这大大增加了磁性存储的密度。

(据科学网)