



(12)发明专利申请

(10)申请公布号 CN 110141245 A

(43)申请公布日 2019.08.20

(21)申请号 201910538632.6

(22)申请日 2019.06.20

(71)申请人 鲁东大学

地址 264025 山东省烟台市芝罘区红旗中路184号

(72)发明人 刘通 李伟 臧睦君 邹海林
柳婵娟 周树森 赵玲玲

(74)专利代理机构 长春市东师专利事务所
22202

代理人 张铁生 刘延军

(51)Int.Cl.

A61B 5/117(2016.01)

A61B 5/0402(2006.01)

A61B 5/0456(2006.01)

A61B 5/00(2006.01)

权利要求书2页 说明书5页

(54)发明名称

一种心电图特征向量提取方法

(57)摘要

一种心电图特征向量提取方法,它包括:1)导入数据;2)去除心电信号的噪声;对数组变量S内每一个数组执行小波分解重构算法,结果存储为数组变量S1;3)提取R波;4)构建特征;5)对样本进行自动识别;6)输出分类结果a;与现有技术相比,本发明的优势是:本发明截取R波顶点前后1.2秒信号进行识别,有效减小了身份识别需要的信号采集时间。本方法中对训练集采用k-means聚类,原理比较简单,实现很容易,收敛速度快。

1. 一种心电图特征向量提取方法,它包括:

1) 导入数据

从外部采集N个人的单导联心电信号,每人采集两条,N大于50,存储为数组变量S,共计2N条;

对2N条心电信号的2N个标签存储为变量label_L;

2) 去除心电信号的噪声

对数组变量S内每一个数组执行小波分解重构算法,结果存储为数组变量S1;

3) 提取R波

在数组变量S1内每一个数组中找到m个R波峰值点位置,共计 $2N*m$ 个位置,在每个位置前取99点,在每个位置后取100点,包含当前位置点共计取200点作为一条R波;

$2N*m$ 个位置提取到 $2N*m$ 条R波,依次放入 $2N*m$ 行、200列的矩阵中,记该变量为Y,作为构建特征方法的输入信号;对变量Y内 $2N*m$ 条R波的 $2N*m$ 个标签存储为变量label;

4) 构建特征

a. 将Y中的R波划分训练集和测试集,把从N个人的第一条信号中提取的 $N*m$ 条R波作为训练集,第二条信号中提取的 $N*m$ 条R波,作为测试集,分别存放在 $N*m$ 行、200列的矩阵中,记为train和test;从变量label中划分出训练集和测试集对应的标签,分别存放在长度为 $N*m$ 的数组中,训练集标签记为train_label,测试集标签记为test_label;将训练集train和测试集test分别分割为 $N*m*20$ 段,每一段长度为10,分别存放在 $N*m*20$ 行、10列的矩阵中,记为train_f和test_f;

(2) 对train_f执行k-means聚类,设置聚类簇数为50,聚类后得到大小为 $50*10$ 的矩阵,记为C1,将C1转置得到 $10*50$ 的矩阵,记为C,对train_f和C计算欧式距离,计算结果存放在 $N*m*20$ 行、50列的矩阵中,记为train_sample_f;对test_f和C计算欧式距离,计算结果存放在 $N*m*20$ 行、50列的矩阵中,记为test_sample_f;

(3) 将训练集样本train_sample_f和测试集样本test_sample_f分别按照列优先的原则变维为 $N*m$ 行、1000列的矩阵并归一化到 $[0,1]$ 之间,记为train_sample和test_sample;

train_sample即为构建好的训练集特征,test_sample即为构建好的测试集特征;

5) 对样本进行自动识别

将训练集特征train_sample进行稀疏,把稀疏后的训练集特征以及对应的标签train_label全部输入到LR分类器中进行训练,LR分类器训练结束后,保留参数,再将测试集特征test_sample进行稀疏,把稀疏后的测试集特征以及对应的标签test_label输入到训练好的LR分类器进行测试;

识别结果为判断R波所属个体的判断准确率a;

6) 输出分类结果a。

2. 根据权利要求1所述的一种心电图特征向量提取方法,其特征在于,所述小波分解重构算法为:

a. 小波分解:选择Harr函数作为小波基;

数组变量S内每一个数组分解得到近似系数向量cA1,细节系数向量cD1;对cA1分解得到近似系数向量cA2,细节系数向量cD2;对cA2分解得到近似系数向量cA3,细节系数向量cD3;对cA3分解得到近似系数向量cA4,细节系数向量cD4;对cA4分解得到近似系数向量

cA5, 细节系数向量cD5;对cA5分解得到近似系数向量cA6, 细节系数向量cD6;对cA6分解得到近似系数向量cA7, 细节系数向量cD7;对cA7分解得到近似系数向量cA8, 细节系数向量cD8;

b. 去除噪声:对需要去噪的信号cD2,采用软阈值方式去噪;

各层系数降噪所需要的阈值不同,采用STDC作为噪声强度估计;

STDC公式: $stdc2 = \text{median}(|cD2|)/0.6745$,其中 $\text{median}(M)$ 为计算M的平均值;

计算出每一层的阈值:

$thr2 = (2 * \log(\text{length}(cD2)))^{1/2} * stdc2$,其中 $\text{length}(cD2)$ 为计算细节向量cD2的长度;

对需要去噪的信号cD2,将cD2中模值小于 $3 * thr2$ 的系数置零,将模值大于 $3 * thr2$ 的系数做处理:大于 $3 * thr2$ 的系数统一减去 $3 * thr2$,小于 $-3 * thr2$ 的系数统一加 $3 * thr2$;

计算后得到去噪后的信号cd2;

对cD3、cD4、cD5、cD6、cD7、cD8也采用上述软阈值方式去除噪声,得到cd3、cd4、cd5、cd6、cd7、cd8;

c. 小波重构:对分解得到的系数向量cA8置零并赋给ca8,对cd8完成小波重构得到信号ca7;对cd7完成小波重构得到信号ca6;对cd6完成小波重构得到信号ca5;对cd5完成小波重构得到信号ca4;对ca4完成小波重构得到信号ca3;对cd3完成小波重构得到信号ca2;对cd2完成小波重构得到信号ca1;ca1即为S内一个数组执行小波分解重构算法的结果。

一种心电图特征向量提取方法

技术领域

[0001] 本发明涉及医学信号处理技术领域,更确切地说一种心电图特征向量提取方法。

背景技术

[0002] 与传统的基于密码或ID卡的认证方式相比,生物识别技术(Biometrics)具有更高的安全度与可靠度,已逐渐成为国际研究的热点,并逐渐在不同的领域发挥着举足轻重的作用。

[0003] 近年来借助于人体内蕴的心电信号ECG(Electrocardiogram)进行身份识别的方法广受关注。ECG是从人体体表采集的反映心脏心动的电位信号,人体的生理条件差异使得ECG具有许多个体特征。相较指纹、语音以及手掌,ECG作为一种活体生物信号,具备易检测、难复制的特点。利用心电信号进行身份识别的定义为:给定一条心电信号,判定该信号所属人的身份。利用心电信号进行自动身份识别的系统或装置通常基于心电信号的模式识别技术进行实现,需要实现心电信号的特征提取和识别两个必要环节。

[0004] 在已公开的利用心电信号进行身份识别的装置和方法的发明中,在提取心电图特征向量时,特征向量往往通过一段较长时间的心电信号提取,存在两方面问题,一方面用户需要较长时间的等待进行心电信号采集,另一方面较大的数据量需要耗费更多的计算成本。因此关于心跳信号的识别方法开始兴起,此类方法的主要优势在于从心电信号中分割一些独立的心跳信号来提取特征向量,提高了心电信号身份识别的实际应用效果。

发明内容

[0005] 本发明的目的是为解决用心电图进行身份识别时间长、成本高的问题,而提供一种心电图特征向量提取方法。

[0006] 一种心电图特征向量提取方法,它包括:

1) 导入数据

从外部采集N个人的单导联心电信号,每人采集两条,N大于50,存储为数组变量S,共计2N条。对2N条心电信号的2N个标签存储为变量label_L;

2) 去除心电信号的噪声

对数组变量S内每一个数组执行小波分解重构算法,结果存储为数组变量S1。

[0007] 3) 提取R波

在数组变量S1内每一个数组中找到m个R波峰值点位置,共计 $2N*m$ 个位置,在每个位置前取99点,在每个位置后取100点,包含当前位置点共计取200点作为一条R波。 $2N*m$ 个位置提取到 $2N*m$ 条R波,依次放入 $2N*m$ 行、200列的矩阵中,记该变量为Y,作为构建特征方法的输入信号;对变量Y内 $2N*m$ 条R波的 $2N*m$ 个标签存储为变量label。

[0008] 4) 构建特征

a. 将Y中的R波划分训练集和测试集,把从N个人的第一条信号中提取的 $N*m$ 条R波作为训练集,第二条信号中提取的 $N*m$ 条R波,作为测试集,分别存放在 $N*m$ 行、200列的矩阵中,记

为train和test;从变量label中划分出训练集和测试集对应的标签,分别存放在长度为 $N*m$ 的数组中,训练集标签记为train_label,测试集标签记为test_label;将训练集train和测试集test分别分割为 $N*m*20$ 段,每一段长度为10,分别存放在 $N*m*20$ 行、10列的矩阵中,记为train_f和test_f;

(2)对train_f执行k-means聚类,设置聚类簇数为50,聚类后得到大小为 $50*10$ 的矩阵,记为C1,将C1转置得到 $10*50$ 的矩阵,记为C,对train_f和C计算欧式距离,计算结果存放在 $N*m*20$ 行、50列的矩阵中,记为train_sample_f;对test_f和C计算欧式距离,计算结果存放在 $N*m*20$ 行、50列的矩阵中,记为test_sample_f;

(3)将训练集样本train_sample_f和测试集样本test_sample_f分别按照列优先的原则变维为 $N*m$ 行、1000列的矩阵并归一化到 $[0,1]$ 之间,记为train_sample和test_sample;

train_sample即为构建好的测试集特征,test_sample即为构建好的测试集特征;

5)对样本进行自动识别

将训练集特征train_sample进行稀疏,把稀疏后的训练集特征以及对应的标签train_label全部输入到LR分类器中进行训练,LR分类器训练结束后,保留参数,再将测试集特征test_sample进行稀疏,把稀疏后的测试集特征以及对应的标签test_label输入到训练好的LR分类器进行测试;

识别结果为判断R波所属个体的判断准确率a;

6)输出分类结果a;

所述小波分解重构算法为:

a.小波分解:选择Harr函数作为小波基。数组变量S内每一个数组分解得到近似系数向量cA1,细节系数向量cD1;对cA1分解得到近似系数向量cA2,细节系数向量cD2;对cA2分解得到近似系数向量cA3,细节系数向量cD3;对cA3分解得到近似系数向量cA4,细节系数向量cD4;对cA4分解得到近似系数向量cA5,细节系数向量cD5;对cA5分解得到近似系数向量cA6,细节系数向量cD6;对cA6分解得到近似系数向量cA7,细节系数向量cD7;对cA7分解得到近似系数向量cA8,细节系数向量cD8;

b.去除噪声:对需要去噪的信号cD2,采用软阈值方式去噪。各层系数降噪所需要的阈值不同,采用STDC作为噪声强度估计;

STDC公式: $stdc2 = \text{median}(|cD2|) / 0.6745$,其中median(M)为计算M的平均值;

计算出每一层的阈值:

$\text{thr2} = (2 * \log(\text{length}(cD2)))^{1/2} * stdc2$,其中length(cD2)为计算细节向量cD2的长度;

对需要去噪的信号cD2,将cD2中模值小于 $3 * \text{thr2}$ 的系数置零,将模值大于 $3 * \text{thr2}$ 的系数做处理:大于 $3 * \text{thr2}$ 的系数统一减去 $3 * \text{thr2}$,小于 $-3 * \text{thr2}$ 的系数统一加 $3 * \text{thr2}$

计算后得到去噪后的信号cd2;

对cD3、cD4、cD5、cD6、cD7、cD8也采用上述软阈值方式去除噪声,得到cd3、cd4、cd5、cd6、cd7、cd8;

c.小波重构:对分解得到的系数向量cA8置零并赋给ca8,对cd8完成小波重构得到信号ca7;对cd7完成小波重构得到信号ca6;对cd6完成小波重构得到信号ca5;对cd5完成小波重构得到信号ca4;对ca4完成小波重构得到信号ca3;对cd3完成小波重构得到信号ca2;对cd2

完成小波重构得到信号ca1;ca1即为S内一个数组执行小波分解重构算法的结果。

[0009] 与现有技术相比,本发明的优势是:本发明截取R波顶点前后1.2秒信号进行识别,有效减小了身份识别需要的信号采集时间。本方法中对训练集采用k-means聚类,原理比较简单,实现很容易,收敛速度快。

具体实施方式

[0010] 实施例1一种心电图特征向量提取方法

下面结合具体的实施方式对本发明作进一步说明。

[0011] 具体实例中使用国际通行心电图数据库ECG-ID,该数据库的数据及使用说明公开于行业内周知的physionet.org网站。ECG-ID数据集包含310录音,从90个主题获得(44个男性和46位女性)。每一个记录含有一个导联的信号,记录在30秒的持续时间,并数字化在500赫兹与12位分辨率超过标称 ± 10 mv范围。在本实例中,共使用55个人的单导联心电信号,每人采集两条通过工作在计算机上的软件系统和行业内所周知的Matlab仿真环境进行实现。

[0012] 本实施例的详细步骤如下:

一.导入数据

从外部采集55个人的单导联心电信号,每人采集两条,存储为数组变量S,所述的数组变量S存储共计110条心电信号。对应数组变量S内110条心电信号的110个标签存储为变量label_L。

[0013] 二.去除心电信号的噪声

对数组变量S内每一个数组执行小波分解重构算法,结果存储为数组变量S1,S1内每个数组对应S内一个数组执行小波分解重构算法的结果。所述小波分解重构算法为:

(1)小波分解

选择Harr函数作为小波基。对数组变量S内每一个数组分解得到近似系数向量cA1,细节系数向量cD1;对cA1分解得到近似系数向量cA2,细节系数向量cD2;对cA2分解得到近似系数向量cA3,细节系数向量cD3;对cA3分解得到近似系数向量cA4,细节系数向量cD4;对cA4分解得到近似系数向量cA5,细节系数向量cD5;对cA5分解得到近似系数向量cA6,细节系数向量cD6;对cA6分解得到近似系数向量cA7,细节系数向量cD7;对cA7分解得到近似系数向量cA8,细节系数向量cD8。

[0014] (2)去除噪声

对需要去噪的信号cD2,采用软阈值方式去噪。各层系数降噪所需要的阈值不同,采用STDC作为噪声强度估计

STDC公式: $stdc2 = \text{median}(|cD2|)/0.6745$,其中median(M)为计算M的平均值

计算出每一层的阈值:

$thr2 = (2 * \log(\text{length}(cD2)))^{1/2} * stdc2$,其中length(cD2)为计算细节向量cD2的长度

对需要去噪的信号cD2,将cD2中模小于 $3 * thr2$ 的系数全部置零,而将模大于 $3 * thr2$ 的做一个比较特殊的处理,大于 $3 * thr2$ 的系数统一减去 $3 * thr2$,小于 $-3 * thr2$ 的系数统一加 $3 * thr2$

计算后得到去噪后的信号cd2。

[0015] 对cd3、cd4、cd5、cd6、cd7、cd8也采用上述软阈值方式去除噪声。得到cd3、cd4、cd5、cd6、cd7、cd8。

[0016] (3)小波重构

对分解得到的系数向量ca8直接置零并赋给ca8,对ca8,cd8完成小波重构得到信号ca7;对ca7,cd7完成小波重构得到信号ca6;对ca6,cd6完成小波重构得到信号ca5;对ca5,cd5完成小波重构得到信号ca4;对ca4,cd4完成小波重构得到信号ca3;对ca3,cd3完成小波重构得到信号ca2;对ca2,cd2完成小波重构得到信号ca1.ca1即为S内一个数组执行小波分解重构算法的结果。

[0017] 三.提取R波

(1)在数组变量S1内每一个数组中找到13个R波峰值点位置(心电图数据库ECG-ID有个别单导联心电图信号只有13个R波峰值点),共计1430个位置。在每个位置前取99点,在每个位置后取100点,包含当前位置点共计取200点作为一条R波,1430个位置提取到1430条R波,将这些R波依次放入1430行、200列的矩阵中,命名该变量为Y。

[0018] (2)Y将会作为构建特征方法的输入信号。对应变量的Y内1430条R波的1430个标签存储为变量label。

[0019] 四.构建特征

所述构建特征方法的输入信号为Y,Y内存储着从110条心电信号中提取的1430条R波。

[0020] (1)将Y中的R波划分训练集和测试集,将从每个人第一条信号中提取的715条R波作为训练集,存放在715行、200列的矩阵中并命名为train,将从每个人第二条信号中提取的715条R波,存放在715行、200列的矩阵中并命名为test。从变量label中划分出训练集和测试集对应的标签,得到训练集标签存放在长度为715的数组中,命名为train_label。得到测试集标签存放在长度为715的数组中,命名为test_label。将训练集train分割为14300段,每一段长度为10,存放在14300行、10列的矩阵中,命名为train_f;将测试集test分割为14300段,每一段长度为10,存放在14300行、10列的矩阵中,命名为test_f。

[0021] (2)对train_f进行k-means聚类,聚类数为50,聚类后得到大小为50*10的矩阵,命名为C1。将C1转置得到10*50的矩阵,命名为C,对train_f和C计算欧式距离。计算结果存放在14300行、50列的矩阵中,命名为train_sample_f;对test_f和C计算欧式距离。计算结果存放在14300行、50列的矩阵中,命名为test_sample_f。

[0022] 在matlab中kmeans的用法为:

```
[C1,~] = kmeans(train_f,50)
```

在matlab中dist函数计算欧式距离:

```
train_sample_f=dist(train_f,C);
```

(3)将训练集样本train_sample_f按照列优先的原则变维为715行、1000列的矩阵并归一化(在matlab中用mapminmax函数实现,下同)到[0,1]之间,命名为train_sample;将测试集样本train_sample_f按照列优先的原则变维为715行、1000列的矩阵并归一化到[0,1]之间,命名为test_sample。

[0023] train_sample即为构建好的测试集特征,test_sample即为构建好的测试集特征。

[0024] 五.对样本进行自动识别

将训练集特征train_sample进行稀疏,把稀疏后的训练集特征以及对应的标签train_label全部输入到LR分类器中进行训练,LR分类器训练结束后,保留参数,再将测试集特征test_sample进行稀疏,把稀疏后的测试集特征以及对应的标签test_label输入到训练好的LR分类器进行测试。

[0025] 测试结束后网络会自动给出识别结果。识别结果为判断R波所属个体的判断准确率a。

[0026] 六、输出分类结果a。

[0027] 结果显示在表1中

表1 训练样本和测试样本数量以及识别率

训练样本	测试样本	识别率
715	715	80.32%

专利名称(译)	一种心电图特征向量提取方法		
公开(公告)号	CN110141245A	公开(公告)日	2019-08-20
申请号	CN201910538632.6	申请日	2019-06-20
[标]申请(专利权)人(译)	鲁东大学		
申请(专利权)人(译)	鲁东大学		
当前申请(专利权)人(译)	鲁东大学		
[标]发明人	刘通 李伟 臧睦君 邹海林 柳婵娟 周树森 赵玲玲		
发明人	刘通 李伟 臧睦君 邹海林 柳婵娟 周树森 赵玲玲		
IPC分类号	A61B5/117 A61B5/0402 A61B5/0456 A61B5/00		
CPC分类号	A61B5/04012 A61B5/0402 A61B5/0456 A61B5/117 A61B5/7203 A61B5/7235 A61B5/726 A61B5/7267		
代理人(译)	张铁生 刘延军		
外部链接	Espacenet SIPO		

摘要(译)

一种心电图特征向量提取方法，它包括：1) 导入数据；2) 去除心电信号的噪声；对数组变量S内每一个数组执行小波分解重构算法，结果存储为数组变量S1；3) 提取R波；4) 构建特征；5) 对样本进行自动识别；6) 输出分类结果a；与现有技术相比，本发明的优势是：本发明截取R波顶点前后1.2秒信号进行识别，有效减小了身份识别需要的信号采集时间。本方法中对训练集采用k-means聚类，原理比较简单，实现很容易，收敛速度快。