



(12)发明专利申请

(10)申请公布号 CN 109965848 A

(43)申请公布日 2019.07.05

(21)申请号 201910313610.X

(22)申请日 2019.04.18

(71)申请人 天津大学

地址 300072 天津市南开区卫津路92号

(72)发明人 魏建国 张东明

(74)专利代理机构 天津市北洋有限责任专利代

理事务所 12201

代理人 吴学颖

(51)Int.Cl.

A61B 5/00(2006.01)

A61B 5/145(2006.01)

G06K 9/00(2006.01)

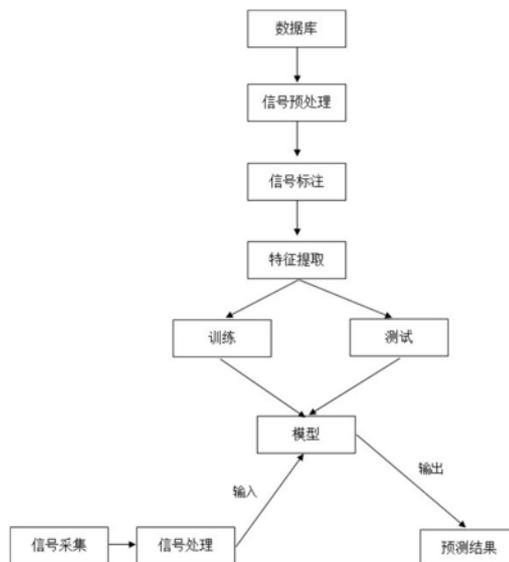
权利要求书1页 说明书5页 附图1页

(54)发明名称

一种基于血氧信号分类的睡眠呼吸暂停综合征的检测方法

(57)摘要

本发明公开了一种基于血氧信号分类的睡眠呼吸暂停综合征的检测方法:获取受试者睡眠时的血氧饱和度信号以及相对应的医生评判信息,建立数据库;对获得的血氧饱和度信号进行预处理,并进行标注,然后提取相关特征;进行“呼吸暂停-低通气”事件的分类模型训练并测试,选取其中效果最好的一个分类模型;获取受试者的睡眠血氧饱和度信号作为分类模型的输入,得到分类结果,并以此判断受试者的患病程度。本发明能够替代PSG的方便快捷、高准确率的对病人进行初步诊断,将会为OSAHS患者的早期诊断提供有力的技术支撑,具有较高的实用价值。



1. 一种基于血氧信号分类的睡眠呼吸暂停综合症的检测方法,其特征在于,包括以下步骤:

步骤一:获取受试者睡眠时的血氧饱和度信号以及相对应的医生评判信息,建立数据库;

步骤二:对获得的血氧饱和度信号进行预处理,并进行标注,然后提取相关特征;

步骤三:进行“呼吸暂停-低通气”事件的分类模型训练并测试,选取其中效果最好的一个分类模型;

步骤四:获取受试者的睡眠血氧饱和度信号作为分类模型的输入,得到分类结果,并以此判断受试者的患病程度。

2. 根据权利要求1所述的基于血氧信号分类的睡眠呼吸暂停综合症的检测方法,其特征在于,步骤二中使用MATLAB软件首先对血氧饱和度信号进行预处理,将一个受试者的整晚的睡眠血氧饱和度信号分割成1分钟的小段,采样率定为1HZ;

然后开始提取训练数据和测试数据中的每个血氧饱和度信号段的特征,提取的特征主要为:最小值、平均值、方差以及以下血氧饱和度信号段内时域上的特征:

- ①过基线的次数;
- ②线性回归的斜率、斜率的绝对值以及截距;
- ③Delta index: Δ index特征参数;
- ④近似熵;
- ⑤中央趋势测量;
- ⑥Lempel-Ziv复杂度;
- ⑦血氧饱和度指数。

一种基于血氧信号分类的睡眠呼吸暂停综合征的检测方法

技术领域

[0001] 本发明涉及信号处理、机器学习等技术领域,更具体的说,是涉及一种基于血氧信号分类的睡眠呼吸暂停综合征的检测方法。

背景技术

[0002] 睡眠对人体来说扮演着十分重要的角色,其不仅是生长发育所必需,也在维持身心健康等方面发挥着不可或缺的作用。然而,随着社会的飞速发展,工作生活压力的不断加大,睡眠问题也日益凸显。数据显示,全球约有27%的人口患有睡眠障碍性疾病,其中以阻塞性睡眠呼吸暂停低通气综合症(OSAHS)最为常见。OSAHS是指睡眠过程中口鼻呼吸气流消失或明显减弱(较基线幅度下降 $\geq 90\%$),持续时间 $\geq 10s$,并伴随着血氧饱和度的下降。目前进行OSAHS诊断的“金标准”为多导睡眠仪(PSG),但该类仪器价格昂贵、操作流程复杂,且需要长时间佩戴多种传感器,会给患者带来诸多的约束和干扰,加重患者身心负担,严重影响患者的睡眠质量,甚至无法反映真实病情。

[0003] 美国睡眠研究学会给出了睡眠呼吸暂停综合症的四个级别,其中检查者的AHI(睡眠呼吸紊乱指数,定义为每小时的睡眠呼吸暂停次数)指数大于等于30为重度患者,AHI指数在15至29区间里是中度患者,在5-14区间的检查者为轻度患者,AHI指数小于5的为正常情况。所以通过受试者睡眠时的血氧信号来计算出受试者的AHI指数,进而评估检查者的患病程度是可行的。

发明内容

[0004] 针对使用多导睡眠仪检测睡眠呼吸暂停综合征的操作流程复杂且价格昂贵问题,本发明提出一种基于血氧信号分类的睡眠呼吸暂停综合征的检测方法,能够替代PSG的方便快捷、高准确率的对病人进行初步诊断,将会为OSAHS患者的早期诊断提供有力的技术支撑,具有较高的实用价值。

[0005] 本发明的目的是通过以下技术方案实现的。

[0006] 本发明基于血氧信号分类的睡眠呼吸暂停综合征的检测方法,包括以下步骤:

[0007] 步骤一:获取受试者睡眠时的血氧饱和度信号以及相对应的医生评判信息,建立数据库;

[0008] 步骤二:对获得的血氧饱和度信号进行预处理,并进行标注,然后提取相关特征;

[0009] 步骤三:进行“呼吸暂停-低通气”事件的分类模型训练并测试,选取其中效果最好的一个分类模型;

[0010] 步骤四:获取受试者的睡眠血氧饱和度信号作为分类模型的输入,得到分类结果,并以此判断受试者的患病程度。

[0011] 步骤二中使用MATLAB软件首先对血氧饱和度信号进行预处理,将一个受试者的整晚的睡眠血氧饱和度信号分割成1分钟的小段,采样率定为1HZ;

[0012] 然后开始提取训练数据和测试数据中的每个血氧饱和度信号段的特征,提取的特

征主要为:最小值、平均值、方差以及以下血氧饱和度信号段内时域上的特征:

- [0013] ①过基线的次数;
- [0014] ②线性回归的斜率、斜率的绝对值以及截距;
- [0015] ③Delta index: Δ index特征参数;
- [0016] ④近似熵;
- [0017] ⑤中央趋势测量;
- [0018] ⑥Lempel-Ziv复杂度;
- [0019] ⑦血氧饱和度指数。

[0020] 与现有技术相比,本发明的技术方案所带来的有益效果是:

[0021] 本发明是基于血氧信号分类的睡眠呼吸暂停综合症的检测方法,由于血氧信号是直接的反映吸入的氧气量,它可以直接反映出睡眠时的氧气变化,并判断出是否发生了呼吸暂停事件。并且该检测方法仅收集了一个通道的信号(即血氧饱和度信号),比多导睡眠仪的通道信号数目大大减少,能够使受试者在一个舒适的环境下进行检测,也能够被引进到家庭环境下,而多导睡眠仪只能在医院里使用。

附图说明

[0022] 图1是基于血氧信号分类的睡眠呼吸暂停综合症的检测方法的流程图。

具体实施方式

[0023] 下面结合附图对本发明作进一步的描述。

[0024] 本发明基于血氧信号分类的睡眠呼吸暂停综合症的检测方法,如图1所示,具体实现过程如下:

[0025] 步骤一:获取受试者睡眠时的血氧饱和度信号以及相对应的医生评判信息,建立数据库。

[0026] 步骤二:对获得的血氧饱和度信号进行预处理,并进行标注,然后提取相关特征。

[0027] 本步骤中使用MATLAB软件首先对血氧饱和度信号进行预处理,将一个受试者的整晚的睡眠血氧饱和度信号分割成1分钟的小段,采样率定为1HZ。

[0028] 然后开始提取训练数据和测试数据中的每个血氧饱和度信号段的特征,提取的特征主要为:最小值、平均值、方差以及以下血氧饱和度信号段内时域上的特征:

[0029] ①过基线的次数:其中以该段内的平均值作为基线,统计该段信号内在时域上穿越基线的次数。

[0030] ②线性回归的斜率、斜率的绝对值以及截距。

[0031] ③Delta index: Δ index特征参数已经被证明能有效判断是否含有呼吸暂停事件,所以为了提高我们的发明在应用中的准确率,在我们的发明中,我们也将该参数作为我们的特征参数之一。

[0032] ④近似熵(Approximate entropy):是一种用于量化时间序列波动的规律性的非线性参数,它用一个非负数来表示一个时间序列的复杂性,反映了时间序列中新信息发生的可能性,越复杂的时间序列对应的近似熵越大。对于信号序列 $\{x(n)\} = x(1), x(2), \dots, x(N)$ (在本发明中, $N=60$), 求其近似熵的步骤为:

[0033] a) 构建 $N-m+1$ 个向量 $X(1), X(2), \dots, X(N-m+1)$, 其中:

$$[0034] X(i) = [x(i), x(i+1), \dots, x(i+m-1)], i=1 \dots N-m+1 \quad (1)$$

[0035] 其中 N 是信号长度, 本发明中为60; m 代表向量长度, 其取值为1、2。

[0036] b) 定义 $d[X(i), X(j)]$ 作为向量 $X(i)$ 和向量 $X(j)$ 的最大绝对距离, 如下:

$$[0037] d[X(i), X(j)] = \max_{k=1, 2, \dots, m} |x(i+k-1) - x(j+k-1)| \quad (2)$$

[0038] c) 对每一个向量 $X(i)$, 统计满足条件 $d[X(i), X(j)] \leq r$ ($j=1, 2, \dots, N-m+1, j \neq i$) 的 j 的总数量, 并记为 N_i^m , 然后得到:

$$[0039] C_r^m = N_i^m / (N-m+1), i=1, 2, \dots, N-m+1 \quad (3)$$

[0040] 其中 r 代表度量值, 其取值为0.1、0.15、0.2、0.25倍的标准差。

[0041] d) 计算每一个 C_r^m 的自然对数并求平均值可得:

$$[0042] \Phi_r^m = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_r^m \quad (4)$$

[0043] e) 将 m 的值增加1, 然后重复步骤a) 到d), 计算 C_r^{m+1} 和 Φ_r^{m+1} 。

[0044] f) 将此条件下的近似熵的值定为 $ApEn(m, r, N)$, 其中:

$$[0045] ApEn(m, r, N) = \Phi_r^m - \Phi_r^{m+1} \quad (5)$$

[0046] ⑤中央趋势测量(Central Tendency Measure, CTM): 是一种使用混沌建模的非线性方法, 可以总结信号变化的程度。其定义为:

$$[0047] CTM = \frac{\sum_{i=1}^{N-2} \delta(d_i)}{N-2} \quad (6)$$

[0048] 在公式(6)中 N 为信号长度, $\delta(d_i)$ 的定义为:

$$[0049] \delta(d_i) = \begin{cases} 1, & [(x(i+2) - x(i+1))^2 + (x(i+1) - x(i))^2]^{1/2} < \rho \\ 0, & \text{其他} \end{cases} \quad (7)$$

[0050] 在公式(7)中, ρ 是半径的值, 取值为0.25、0.5、0.75、1; $x(i)$ 为信号序列 $\{x(n)\} = x(1), x(2), \dots, x(N)$ (在本发明中, $N=60$)。

[0051] ⑥Lempel-Ziv复杂度: 一种用于度量随着序列长度的增加而新模式也增加的算法, 表现了有限序列和随机序列的近似程度, 序列的Lempel-Ziv复杂度越大, 就越趋近于随机序列。

[0052] ⑦血氧饱和度指数: ODI_{xy} : $x \in \{2, 3, 4, 5\}$ $y \in \{1, 3, 5\}$ 其中 x 表示至少低于基线的值, y 表示至少持续的时间(秒)。基线取值为该段数据内前20%大的值得平均值。

[0053] 步骤三: 进行“呼吸暂停-低通气”事件的分类模型训练并测试, 选取其中效果最好的一个分类模型。

[0054] 步骤四: 获取受试者的睡眠血氧饱和度信号作为分类模型的输入, 得到分类结果, 并以此判断受试者的患病程度。

[0055] 实施例:

[0056] 在本发明的步骤一中, 我们采用的数据来源是都柏林圣文森特大学医院睡眠呼吸暂停数据库(UCD database), 该数据库包含了25名(21名男性, 4名女性)被怀疑患有睡眠呼吸暂停综合征的成人受试者的完整的夜间多导睡眠图, 其生理特征如表1所示。UCD数据库

的可用性很容易验证和评估我们的方法,这25段记录中,每段都包含5.9-7.7小时的血氧饱和度(SaO_2 ,简称血氧)信号,并且包含了详细的起始时间和持续时间的注释文件,每次呼吸暂停/呼吸不足事件获得多导睡眠图是使用的Jaeger-Toennies系统(Erich Jaeger GmbH,德国)。

[0057] 表1:UCD数据库中受试者的生理特性

[0058]

	平均值	范围
年龄	50 ± 10 岁	28-68岁
身体质量指数(BMI)	$31.6 \pm 4.0 \text{kg/m}^2$	25.1-42.5 kg/m^2
呼吸紊乱指数(AHI)	24.1 ± 20.3	1.7-90.9

[0059] 在本发明的步骤二中,我们使用MATLAB软件首先对信号进行预处理,将一个受试者的整晚的睡眠血氧饱和度信号分割成1分钟的小段,采样率定为1HZ。由于训练模型和测试模型需要用到训练数据和测试数据,所以根据数据库中的注释文件对每一个1分钟的血氧饱和度信号段进行标注,如果该1分钟的血氧饱和度信号段内包含至少10秒的“呼吸暂停/呼吸不足”事件,我们就将该段对应的标签记为“呼吸暂停”,否则标记为“正常”。并且如果任何血氧饱和度信号段中血氧饱和度的值包含低于50%都要被丢弃,因为这是由于采集的机器造成的血氧饱和度下降。

[0060] 然后开始提取训练数据和测试数据中的每个血氧饱和度信号段的特征,提取的特征主要为:

[0061] 由于一个采样率为1HZ的一分钟的血氧饱和度信号段就是60个范围在0-100的数字,所以提取这60个数字的基础统计数据作为其特征也具有一定的作用,如最小值、平均值、方差等。

[0062] 其次就是统计这个一分钟的血氧饱和度信号段内时域上的特征:

[0063] ①过基线的次数:其中以该段内的平均值作为基线,统计该段信号内在时域上穿越基线的次数。

[0064] ②线性回归的斜率、斜率的绝对值以及截距:

[0065] 线性回归是利用数理统计中回归分析,来确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法,其表达式为 $y=wx+e$, e 为误差服从均值为0的正态分布。我们通过MATLAB中的`regress()`函数来获得每段血氧饱和度信号段的线性回归的斜率及其截距。

[0066] ③Delta index: Δ index特征参数已经被证明能有效判断是否含有呼吸暂停事件,所以为了提高我们的发明在应用中的准确率,在我们的发明中,我们也将该参数作为我们的特征参数之一。其提取步骤为:

[0067] (i) 首先计算每一个12秒间隔内的平均值;

[0068] (ii) 然后计算相邻的间隔窗的绝对差值;

[0069] (iii) 得到平均绝对差值,就是Delta index。

[0070] ④近似熵(Approximate entropy):是一种用于量化时间序列波动的规律性的非线性参数,它用一个非负数来表示一个时间序列的复杂性,反映了时间序列中新信息发生的可能性,越复杂的时间序列对应的近似熵越大。对于信号序列 $\{x(n)\} = x(1), x(2), \dots, x$

(N) (在本发明中, $N=60$), 按公式 (1) - (5) 求其近似熵。

[0071] ⑤CTM参数是一种非线性的特征, 该参数通过将一维时间信号投射到二维坐标系中, 并图形来表示一维时间信号(时间序列)的变化(混乱)程度。本发明中求取CTM特征参数的步骤如公式(6) - (7)所示。其中, 本发明中, CTM参数的 ρ 的取值为0.25。

[0072] ⑥Lempel-Ziv复杂度: 一种用于度量随着序列长度的增加而新模式也增加的算法, 表现了有限序列和随机序列的近似程度, 序列的Lempel-Ziv复杂度越大, 就越趋近于随机序列。

[0073] 对于Lempel-Ziv复杂度, 其计算机的实现方式是: 对于一个待求字符串S (S_1, S_2, \dots, S_n) 以及另一个字符串Q (q_1, q_2, \dots, q_n), SQ表示S和Q的级联, $SQ = (S_1, S_2, \dots, S_n, q_1, q_2, \dots, q_n)$ 。令 SQ_v 是SQ减去最后一个字符所得字符串。判断Q是否是 SQ_v 的一个子串, 如果Q是 SQ_v 的一个子串, 说明Q中的字符是可从S复制的, 这时把待求序列的下一个字符级联到Q。如果Q不是 SQ_v 的一个子串, 则表示Q是插入字符。这时把Q级联到S, $S = SQ$, 重新构造Q, 重复以上过程直到Q取待求序列的最后一位结束。每次Q级联到S, 表明出现一种新模式, 用c表示一个字符串中新模式的数量。

[0074] ⑦血氧饱和度指数: ODI_{xy} : $x \in \{2, 3, 4, 5\}$ $y \in \{1, 3, 5\}$ 其中x表示至少低于基线的值, y表示至少持续的时间(秒)。基线取值为该段数据内前20%大的值得平均值。

[0075] 在该发明的步骤三中, 我们采用了支持向量机(SVM)分类器、K-近邻(KNN)分类器以及随机森林(Random Forest)分类器。在进行测试阶段, 为了评估分类的性能与效果, 我们使用灵敏度(sensitivity)、特异度(specificity)和准确度(accuracy)来评价, 其定义如下:

$$[0076] \quad \text{sensitivity} = \frac{TP}{TP+FN} \quad (8)$$

$$[0077] \quad \text{specificity} = \frac{TN}{TN+FP} \quad (9)$$

$$[0078] \quad \text{accuracy} = \frac{TP+TN}{P+N} \quad (10)$$

[0079] 在公式(8)、(9)、(10)中P代表标注为“呼吸暂停”的血氧饱和度信号段的总数; N代表标注为“正常”的血氧饱和度信号段的总数; TP代表正确识别的标注为“呼吸暂停”的血氧饱和度信号段的总数; TN代表正确识别的标注为“正常”的血氧饱和度信号段的总数; FP代表错误识别的标注为“呼吸暂停”的血氧饱和度信号段的总数; FN代表错误识别的标注为“正常”的血氧饱和度信号段的总数。

[0080] 根据步骤三, 我们可以得到一个分类效果比较好的分类模型, 在步骤四中, 受试者可以通过市场上流行的测量血氧饱和度的血氧仪, 然后将整晚的血氧饱和度信号进行预处理和分割, 就可以输入到在步骤三时得到的模型, 最后获得模型的预测结果, 然后就算得到AHI指数, 进而可以对该名受试者的患病程度进行评估。

[0081] 尽管上面结合附图对本发明进行了描述, 但本发明并不局限于上述, 上述的具体实施方式仅仅是示意性的, 而不是限制性的, 本领域的普通技术人员在本发明的启示下, 还可做出很多形式, 这些均属于本发明的保护之内。

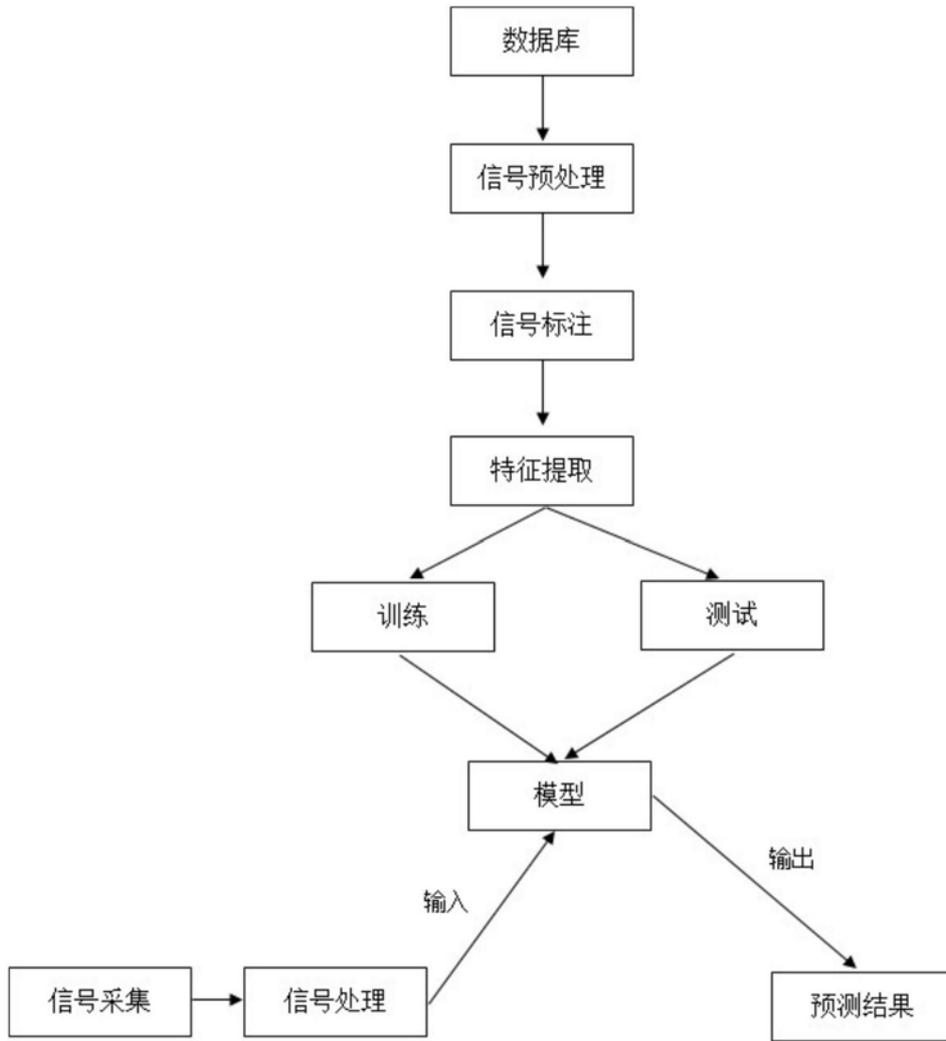


图1

专利名称(译)	一种基于血氧信号分类的睡眠呼吸暂停综合征的检测方法		
公开(公告)号	CN109965848A	公开(公告)日	2019-07-05
申请号	CN201910313610.X	申请日	2019-04-18
[标]申请(专利权)人(译)	天津大学		
申请(专利权)人(译)	天津大学		
当前申请(专利权)人(译)	天津大学		
[标]发明人	魏建国 张东明		
发明人	魏建国 张东明		
IPC分类号	A61B5/00 A61B5/145 G06K9/00		
CPC分类号	A61B5/14542 A61B5/4818 A61B5/7267 G06K9/0053 G06K9/0055		
外部链接	Espacenet SIPO		

摘要(译)

本发明公开了一种基于血氧信号分类的睡眠呼吸暂停综合征的检测方法：获取受试者睡眠时的血氧饱和度信号以及相对应的医生评判信息，建立数据库；对获得的血氧饱和度信号进行预处理，并进行标注，然后提取相关特征；进行“呼吸暂停-低通气”事件的分类模型训练并测试，选取其中效果最好的一个分类模型；获取受试者的睡眠血氧饱和度信号作为分类模型的输入，得到分类结果，并以此判断受试者的患病程度。本发明能够替代PSG的方便快捷、高准确率的对病人进行初步诊断，将会为OSAHS患者的早期诊断提供有力的技术支撑，具有较高的实用价值。

