



# (12)发明专利申请

(10)申请公布号 CN 106295225 A

(43)申请公布日 2017.01.04

(21)申请号 201610728659.8

(22)申请日 2016.08.26

(71)申请人 复旦大学

地址 200433 上海市杨浦区邯郸路220号

(72)发明人 李翔 董昭

(74)专利代理机构 上海正旦专利代理有限公司

31200

代理人 陆飞 王洁平

(51)Int. Cl.

G06F 19/00(2011.01)

G06K 9/00(2006.01)

G06K 9/62(2006.01)

A61B 5/00(2006.01)

A61B 5/0452(2006.01)

A61B 5/0456(2006.01)

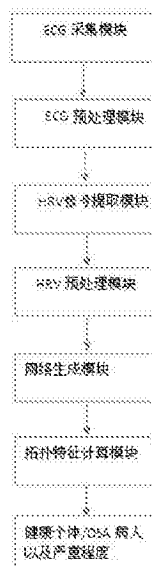
权利要求书2页 说明书6页 附图4页

## (54)发明名称

基于互信息网络的睡眠呼吸暂停综合征的检测系统

## (57)摘要

本发明公开了一种基于互信息网络的睡眠呼吸暂停综合征的检测系统。本发明系统由ECG采集模块,ECG预处理模块,HRV信号提取模块,HRV预处理模块,网络生成模块和拓扑特征计算模块构成。该系统所需信号为单导联ECG信号,采集方便。本发明的检测系统操作方法简单易行,准确性高。



1. 一种基于互信息网络的睡眠呼吸暂停综合症的检测系统,其特征在于,其由ECG采集模块,ECG预处理模块,HRV信号提取模块,HRV预处理模块,网络生成模块和拓扑特征计算模块构成;其中:

ECG采集模块,负责采集患者睡眠的ECG信号;

ECG预处理模块,对ECG信号滤波,去除噪声以及基线漂移;

HRV信号提取模块,通过检测ECG信号的R波,获取心跳间隔信号;

HRV预处理模块,滤除HRV信号中的异常值;

网络生成模块,将HRV切割成N个片段,通过互信息方法计算片段之间的相似性,并将互信息矩阵转换成无权无向网络;其中:互信息的计算方法为:

设两个随机变量(X,Y)的联合分布为 $P(x,y)$ ,边缘分布为 $p(x)$ 、 $p(y)$ ,

$$I(x,y) = H(x) + H(y) - H(x,y)$$

其中: $H(x) = \sum_{i=1}^n p(x_i) \log_2 p(x_i)$

$H(x,y) = \sum_x \sum_y p(x,y) \log_2 p(x,y)$ ;

当互信息矩阵内的元素大于阈值 $\tau$ ,则对应的网络节点i与j是相连的,否则不连,如此构建无权无向网络;

拓扑特征计算模块,计算生成的无权无向网络的拓扑特征指标,利用拓扑特征指标检测OSA病人;其中:拓扑特征指标包括平均度k、局部聚类系数C、全局聚类系数T、全局效率E和模块性Q;其计算方法如下:

平均度 $k$ 为网络中每个节点i度值 $k_i$ 的平均值,

$$k = \frac{1}{N} \sum k_i$$

其中:N为网络中节点的数目;

节点i的聚类系数定义为

$$C_i = \frac{2E_i}{k_i(k_i - 1)}$$

其中: $E_i$ 为节点i的边数;

网络的局部聚类系数 $C$ 定义为所有节点聚类系数的均值,

$$C = \frac{1}{N} \sum C_i$$

全局聚类系数T定义为

$$T = \frac{3 * \text{三角形数目}}{\text{三元组个数}}$$

全局效率E定义为

$$E = \frac{1}{N(N-1)} \sum \frac{1}{L_{i,j}}$$

其中: $L_{i,j}$ 为节点i与节点j之间最短路径。

2. 根据权利要求1所述的检测系统,其特征在于,采用带宽滤波0.5-40Hz滤除ECG信号中噪声以及基线漂移。

3. 根据权利要求1所述的检测系统,其特征在于,采用最大值方法检测ECG的R波,获取HRV信号。

4. 根据权利要求1所述的检测系统,其特征在于,通过滑动平均滤波器对HRV信号进行预处理。

5. 根据权利要求1所述的检测系统,其特征在于, 阈值 $\epsilon$ 的取值在0.5~1之间。

## 基于互信息网络的睡眠呼吸暂停综合征的检测系统

### 技术领域

[0001] 本发明属于医学检测技术领域,具体的说,涉及一种基于互信息网络的用于睡眠呼吸暂停综合征的检测系统。

### 背景技术

[0002] 睡眠呼吸暂停通常分为阻塞性睡眠呼吸暂停(obstructive sleep apnea, OSA)、中枢性睡眠呼吸暂停(central sleep apnea, CSA)和混合性睡眠呼吸暂停(mixed sleep apnea, MSA)三种类型。其中,OSA指睡眠中因上气道阻塞引起的呼吸暂停和通气不足的症状,表现为口鼻腔气流停止而胸腹呼吸动作尚存在。CSA指由于中枢神经系统的呼吸中枢功能障碍或支配呼吸肌的神经或呼吸肌病变,导致气道无阻塞但发生呼吸暂停的症状,表现为口鼻腔气流和胸腹呼吸动作同时停止。MSA是OSA与CSA并存。

[0003] 阻塞性睡眠呼吸暂停综合征(obstructive sleep apnea syndrome, OSAS)是指睡眠状态下呼吸暂停和低通气。目前我国约有4000万人患有该类疾病。患有该病症的人容易产生白天困倦、性格暴躁、注意力不集中等症状,易引发高血压、冠心病、脑血栓等疾病。因此准确及时的检测OSA病人非常有意义。

[0004] 睡眠呼吸暂停综合征是 $AH1 \geq 5$ 次/小时(睡眠呼吸暂停/低通气指数)的一种睡眠呼吸疾病。 $0 < AH1 < 5$ :正常; $5 < AH1 \leq 20$ 属于轻度OSA; $20 < AH1 \leq 40$ 属于中度OSA; $AH1 > 40$ 属于重度OSA。报告表明9%的中年男性及4%的中年女性患有此病。OSA会导致白天嗜睡,头昏,头疼记忆力衰退,反应迟钝。长期患有此病可引起高血压、头昏、冠心病、心衰、中风等多种疾病。

[0005] 通常OSA的检测的标准是多导睡眠图监测(PSG),检测的信号有:眼电图,肌电图,心电图,脑电图,血氧饱和度,胸部和腹部呼吸幅度图,鼻、口通气量等十多个信号。PSG的检测需要在睡眠实验室睡1到2天,睡眠实验室造价昂贵,且需要专业技术人员,大部分OSA患者无法到睡眠实验室监测确诊。

[0006] 目前OSA检测的创新点主要是寻求PSG的替代方法,如仅采用心电图(ECG)、口鼻气流、鼾声中的一个信号或几个信号完成对OSA严重程度的判定。

[0007] 诸多研究通过从单导联的ECG信号中提取信息来检测OSA病人,并判断发病的严重程度。睡眠呼吸障碍(OSA)发病时,呼吸阻滞,导致了心率也发生明显的变化。呼吸暂停时,心动过缓;呼吸恢复,心动过速,心率变异性(HRV)呈现为锯齿波。这种生理的现象称之为心率周期性变化(Cyclic variation of heart rate, CVHR)。研究人员通常采用了诸多信号处理的方法来检测OSA病人。他们通过从时域频域提取信息导入分类器并通过分类器检测OSA病人。

[0008] HRV分析可以分为时域分析法、频域分析法、时频分析法及非线性分析。HRV的时域分析可以分为统计法和图解法。《黄宛临床心电图学》(2009年1月出版)给出了最具代表性的时域指标。

[0009] (1)SDNN:全部正常窦性心搏间期的标准差,单位ms;

- (2)SDANN:全程按5分钟分成连续的时间段,先计算每5分钟到NN间期平均值,再计算所有平均值的标准差,单位:ms;
- (3) rMSSD:全程相邻NN间期之差的均方根值,单位:ms;
- (4) SDNN Index: 全程按5分钟分成连续的时间段,先计算每5分钟的NN间期标准差,再计算这些标准差的平均值,单位:ms;
- (5) SDSD:全部相邻NN间期的之差的标准差,单位:ms;
- (6) NN50:全部NN间期中,相邻的NN间期之差大于50ms的心搏数,单位:个;
- (7) PNN50:NN50除以总NN间期个数、乘以100,单位:%。

[0010] 图解法的指标及其定义为:

- (1)HRV三角指数:以1/128s为采样间隔绘制某段;
- (2)TINN:使用最小法,求出全部NN间期的直方图近似三角形底边的宽度,单位ms。

[0011] 时域分析的优势在于简单易行。

[0012] Pearson相关系数与互信息(mutual information,MI)是评估两个时间序列之间相似度的最简单最常用的方法。与Pearson相关系数相比,互信息优势在于还能够用于非线性信号分析。

## 发明内容

[0013] 为了克服现有技术的不足,本发明的目的在于提供一种基于互信息网络的睡眠呼吸暂停综合症的检测系统。本发明的检测系统中采用互信息的方法衡量HRV片段之间的相似度。

[0014] 与传统的检测装置不同,我们采用复杂网络来检测OSA病人,即通过互信息(MI)网络的方法刻画了这种变化:两个OSA片段的相似性明显大于两个非OSA片段的相似性。采用互信息计算HRV片段之间的相似性以此生成互信息矩阵,然后设定合适的阈值将互信息矩阵转换成网络,用网络特征值检测OSA病人。本发明的技术方案具体介绍如下。

[0015] 本发明提供一种基于互信息网络的睡眠呼吸暂停综合症的检测系统,其由ECG采集模块,ECG预处理模块,HRV信号提取模块,HRV预处理模块,网络生成模块和拓扑特征计算模块构成;其中:

ECG采集模块,负责采集患者睡眠的ECG信号;

ECG预处理模块,对ECG信号滤波,去除噪声以及基线漂移;

HRV信号提取模块,通过检测ECG信号的R波,获取心跳间隔信号;

HRV预处理模块,滤除HRV信号中的异常值;

网络生成模块,通过将HRV切割成N个片段,通过互信息方法计算片段之间的相似性,并将互信息矩阵转换成无权无向网络;其中:互信息的计算方法为:设两个随机变量(X,Y)的联合分布 $p(x,y)$ ,边际分布 $p(x)$ , $p(y)$ ,互信息 $I(x,y)$ 可以表示为

$$I(x,y) = H(x) + H(y) - H(x,y)$$

$$\text{其中: } H(x) = \sum_{i=1}^n p(x_i) \log_2 p(x_i)$$

$$H(x,y) = \sum_x \sum_y p(x,y) \log_2 p(x,y);$$

当互信息矩阵内的元素大于阈值 $\tau$ ,则对应的网络节点i与j是相连的,否则不连,如此

构建无权无向网络；

拓扑特征计算模块，计算生成的无权无向网络的拓扑特征指标，并利用拓扑特征指标检测OSA病人；其中：拓扑特征指标包括平均度 $k$ 、局部聚类系数 $C$ 、全局聚类系数 $T$ 、全局效率 $E$ 和模块性 $Q$ ；其计算方法如下：

平均度 $k$ 为网络的每个节点 $i$ 度值 $k_i$ 的平均值，

$$k = \frac{1}{N} \sum k_i$$

其中： $N$ 为网络节点数目；

节点的聚类系数定义为，

$$C_i = \frac{2E_i}{k_i(k_i - 1)}$$

其中： $E_i$ 为节点 $i$ 的边数；

网络的局部聚类系数 $C$ 定义为所有节点聚类系数的均值，

$$C = \frac{1}{N} \sum C_i$$

全局聚类系数 $T$ 定义为

$$T = \frac{3 \cdot \text{三角形数目}}{\text{三元组个数}}$$

全局效率 $E$ 定义为

$$E = \frac{1}{N(N-1)} \sum \frac{1}{L_{i,j}}$$

其中： $L_{i,j}$ 为节点 $i$ 与节点 $j$ 之间最短路径。

[0016] 本发明中，采用带宽滤波0.5-40Hz滤除ECG信号中噪声以及基线漂移。

[0017] 本发明中，采用最大值方法检测ECG的R波，获取HRV信号。

[0018] 本发明中，通过滑动平均滤波器对HRV信号进行预处理。

[0019] 本发明中，阈值 $\tau$ 的取值在0.5~1之间。

[0020] 本发明的有益效果在于：该系统操作时的优势分析的信号为RR间隔序列，容易获取，简单易行。

## 附图说明

[0021] 图1 是基于互信息网络的睡眠呼吸暂停综合症的检测流程图。

[0022] 图2 是睡眠呼吸障碍发病时的HRV片段与正常HRV片段的图示。

[0023] 图3 是OSA个体/健康个体的HRV片段之间的互信息矩阵图示。

[0024] 图4 是R波检测示意图。

[0025] 图5 是全局聚类系数 $T$ 的ROC曲线图(最佳阈值由原点标记)。

[0026] 图6 是全局聚类系数 $T$ 与局部聚类系数 $C$ 联立在训练集(a)、测试集(b)的区分效果示意图。

[0027] 图7 是AH1与全局聚类系数T的相关性图示。

### 具体实施方式

[0028] 数据库介绍:我们采用MIT-Physionet Apnea数据库,该数据库分为训练集和测试集。训练集和测试集中分别有35个长度约为8小时的夜间ECG信号,抽样频率为100Hz。个体年龄为27到63岁,体重为53公斤到135公斤。

[0029] 训练集共有35个个体,按照AH1=5作为标准划分健康个体与OSA个体,可以将35个个体分为22个OSA病人(AH1 $\geq$ 5),13个健康个体(AH1<5)。同理,35个测试集样本可以分为23个OSA病人,12个健个体。

将每个长度为8小时的ECG分别通过如下模块,流程图见图1。

[0030] 1. ECG预处理模块:采用带宽滤波(0.5-40Hz)滤除ECG信号中工频噪声,基线漂移。ECG信号处理后,通过三次样条内插法,重新采样,采样频率为500Hz。

[0031] 2. HRV信号提取模块:采用最大值方法检测ECG的R波,获取HRV信号。(见图4)。

[0032] 3. HRV预处理模块:①去除异常值:通过滑动平均滤波器,滑动窗长5-beat, mRR为窗内RR信号的平均值,如果RR>1.2\*mRR或RR<0.8\*mRR,则该RR值为异常值由mRR替代。②信号重采样:采用三次样条内插(cubic-spline),采样频率为4Hz。

[0033] 4. 网络生成模块:将预处理后的RR序列分割成长度为5分钟的片段,在这里共有70个片段。计算片段之间互信息。互信息的计算方法为:设两个随机变量(x,y)的联合分布为p(x,y),边际分布为p(x),p(y);互信息I(x,y)可以通过如下公式计算:

$$I(x,y) = H(x) + H(y) - H(x,y)$$

$$\text{其中: } H(x) = \sum_{i=1}^n p(x_i) \log_2 p(x_i)$$

$$H(x,y) = \sum_x \sum_y p(x,y) \log_2 p(x,y)$$

这里利用了睡眠呼吸阻塞发病时HRV信号呈规律的锯齿波(见图2),这与正常状况下的HRV呈鲜明的对比(见图2)。这种发病时HRV的锯齿波导致了两个睡眠呼吸障碍的HRV片段的互信息大于两个正常HRV片段的互信息。

[0034] 将长度为350分钟的HRV分割成70个片段,每个片段长度为5分钟。计算HRV片段两两之间互信息获得70\*70的互信息矩阵。如果互信息矩阵内的元素大于阈值 $\tau$ ,则对应的网络节点i与j是相连的,否则不连。如此构成一个无权无向网络。

[0035] 两个睡眠呼吸障碍发病时HRV片段的互信息明显大于两个正常的HRV片段之间的互信息,而OSA病人含有大量的睡眠呼吸发病的HRV片段导致了所生成的互信息网络的平均值远大于正常个体的HRV片段所生成的互信息网络。(见图3)。

[0036] 5. 拓扑特征计算模块,计算网络拓扑特征值,选取区分效果最好的网络特征值区分OSA个体与健康个体。为了获取最好的区分效果,我们遍历了加权网络转换到无权网络的阈值,从0到1,每次增加0.01,并计算了生成的无权无向的网络拓扑指标,即平均度k(average degree)、局部聚类系数C(clustering coefficient)、全局聚类系数T(transitivity)、全局效率E(global efficiency)、模块性Q(Modularity)这五个参数,它们的具体计算方法如下:

平均度 $k$ 为网络中每个节点i度值 $k_i$ 的平均值

$$k = \frac{1}{N} \sum k_i$$

这里N为网络中节点数目。

[0037] 节点的聚类系数定义为

$$C_i = \frac{2E_i}{k_i(k_i - 1)}$$

网络的局部聚类系数C定义为所有节点聚类系数的均值，

$$C = \frac{1}{N} \sum C_i$$

全局聚类系数定义为

$$T = \frac{3 \times \text{三角形数目}}{\text{三元组个数}}$$

全局效率E定义为

$$E = \frac{1}{N(N-1)} \sum \frac{1}{L_{i,j}}$$

这里 $L_{i,j}$ 为节点i与节点j之间最短路径，N为节点的数目。

[0038] 我们发现当阈值 $\tau$ 取[0.5 1]之间区分效果最好。

[0039] 表1为 $\tau = 0.8$ ，来自训练集的12个健康个体与23个OSA个体的平均度k、全局效率E、全局聚类系数T、局部聚类系数C、模块性Q。很明显，OSA病人对应的互信息网络平均值、全局效率、全局聚类系数、局部聚类系数、模块性都明显大于健康个体。

[0040] 表1：训练集中所有健康个体与OSA病人的网络拓扑特征值(平均值±方差)

网络拓扑特征	OSA病人	健康个体	P值
平均度 k	46.5 ± 14.7	9.78 ± 5.06	2.4E-8
全局效率 E	0.8287 ± 0.1193	0.46 ± 0.14	5.1E-6
局部聚类系数 C	0.8531 ± 0.11	0.4 ± 0.05	1.5E-5
全局聚类系数 T	0.84 ± 0.12	0.39 ± 0.15	6.0E-6
模块 Q	0.2 ± 0.1	0.05 ± 0.02	6.0E-6

首先我们讨论单个阈值对OSA病人与健康个体的区分效果。

[0041] 单个参数需要设定阈值来确定最优的区分效果，这里我们采用受试者工作特征曲线，即ROC曲线。ROC曲线很容易找到特征值在不同数值时对疾病的识别能力。以参数全局聚类系数T为例，图5展示了相应的ROC曲线，并用圆圈标记了最佳阈值。

[0042] 表2展示了当 $\tau = 0.8$ 时，k、Q、T、E在训练集中所选取的阈值以及以此阈值对OSA病人和健康个体的区分。可以看出，在训练集中，平均度k的区分效果最为理想，当平均度k=19时，对OSA病人与健康个体区分效果最为理想，准确性Ac、敏感性Se、特异性Sp都达到100%。但在测试集中，平均度k的区分效果有所降低。

[0043] 表2：单个网络特征值的区分效果(阈值为0.8)。

网络特征值	阈值	训练集			测试集		
		敏感性	特异性	准确性	敏感性	特异性	准确性
平均度 $k$	19	100	100	100	95.65	91.67	94.29
模块 $Q$	0.1	100	92.31	97.14	95.65	100	97.14
局部聚类系数 $C$	0.68	86.36	92.31	88.57	95.65	91.67	94.29
全局聚类系数 $T$	0.61	95.65	100	97.14	95.65	100	97.14
全局效率 $E$	0.61	95.65	100	97.14	95.65	91.67	94.29

[0044] 在测试集中,模块性 $Q$ 与全局聚类系数均能达到最好的区分效果,准确性 $A_c$ 为97.14%,敏感性 $Se$ 为100%,特异性 $Sp$ 为95.65%。

[0045] 以上我们讨论了单个参数的区分能力,为了提高性能,我们还测试了两个参数联立的区分效果。图6展示了 $C$ 与 $T$ 联立在训练集、测试集中的区分效果。可以看出OSA病人与健康个体完美的区分。区分的准确性 $A_c$ 为100%、敏感性 $Se$ 为100%、特异性 $Sp$ 为100%。

[0046] 我们还发现 $AHI$ 与全局聚类系数 $T$ 有明显的相关性(见图7),Pearson相关系数为0.77,因此可以用全局聚类系数 $T$ 来判断OSA发病的严重程度。



图 1

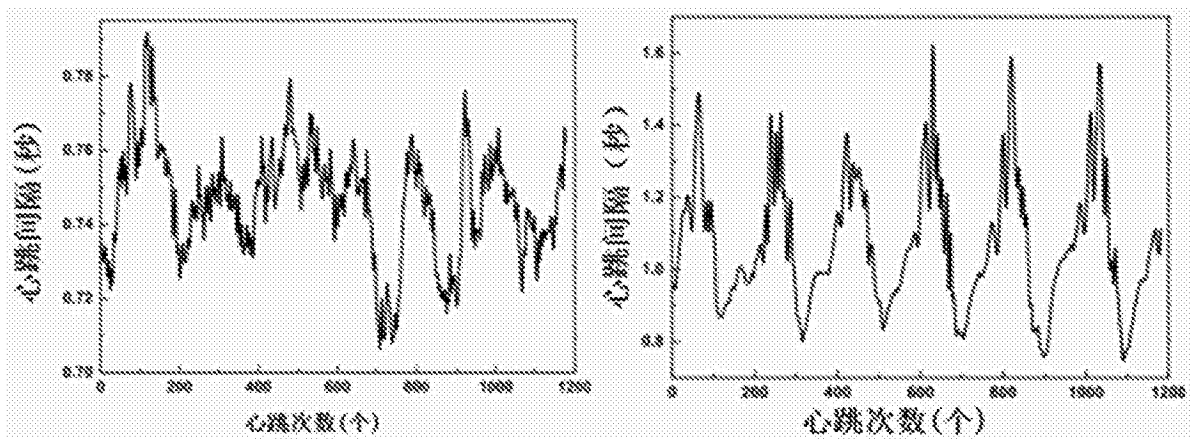


图 2

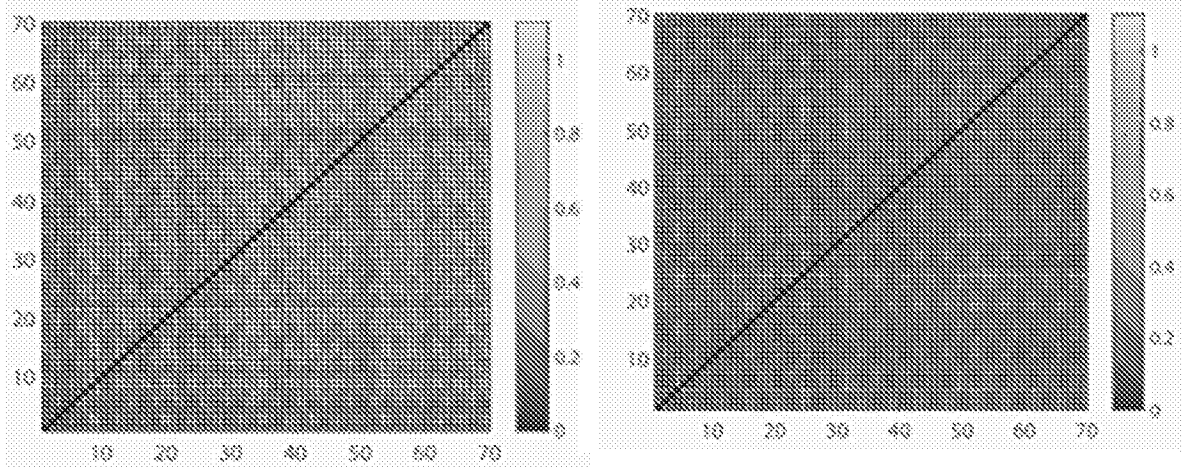


图 3

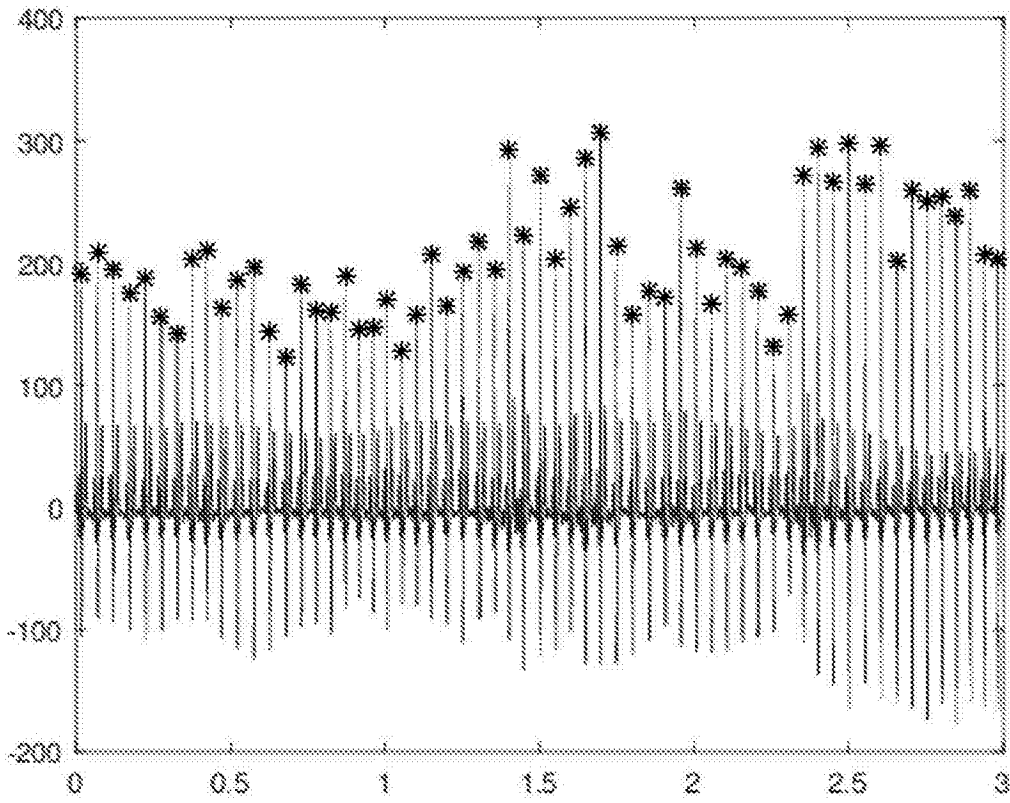


图 4

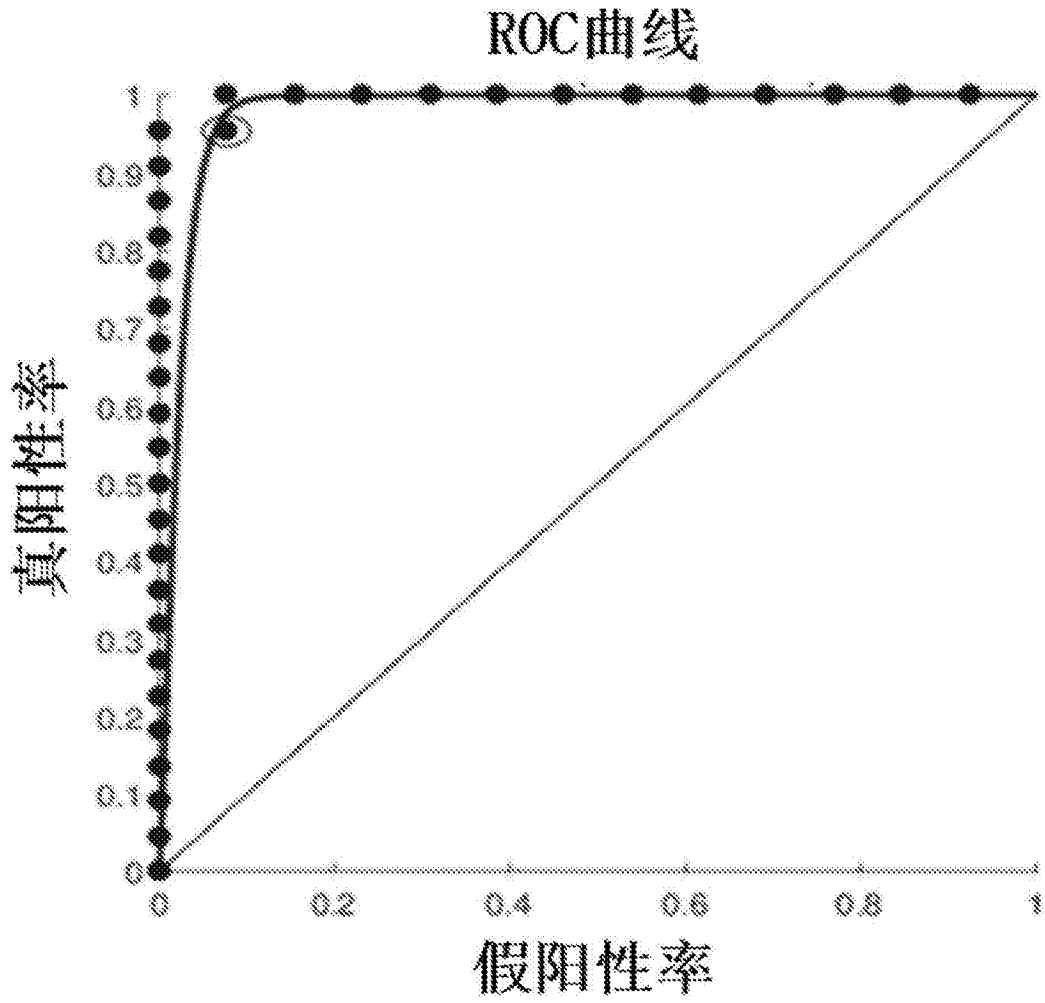


图 5

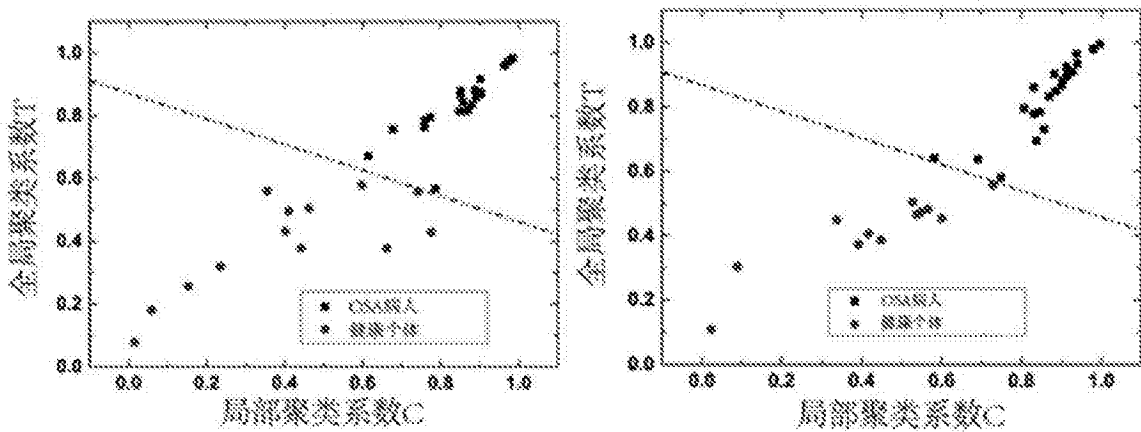


图 6

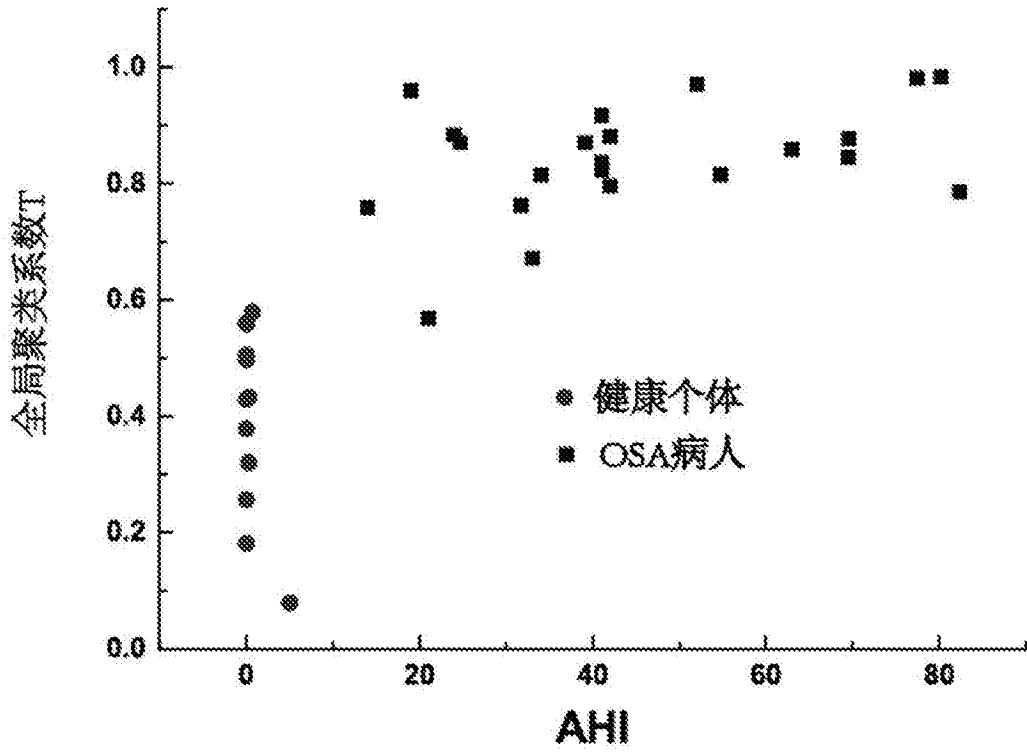


图 7

专利名称(译)	基于互信息网络的睡眠呼吸暂停综合征的检测系统		
公开(公告)号	<a href="#">CN106295225A</a>	公开(公告)日	2017-01-04
申请号	CN201610728659.8	申请日	2016-08-26
[标]申请(专利权)人(译)	复旦大学		
申请(专利权)人(译)	复旦大学		
当前申请(专利权)人(译)	复旦大学		
[标]发明人	李翔 董昭		
发明人	李翔 董昭		
IPC分类号	G06F19/00 G06K9/00 G06K9/62 A61B5/00 A61B5/0452 A61B5/0456		
CPC分类号	A61B5/0452 A61B5/0456 A61B5/4818 A61B5/72 G06K9/0051 G06K9/622		
代理人(译)	陆飞 王洁平		
外部链接	<a href="#">Espacenet</a> <a href="#">SIPO</a>		

摘要(译)

本发明公开了一种基于互信息网络的睡眠呼吸暂停综合征的检测系统。本发明系统由ECG采集模块，ECG预处理模块，HRV信号提取模块，HRV预处理模块，网络生成模块和拓扑特征计算模块构成。该系统所需信号为单导联ECG信号，采集方便。本发明的检测系统操作方法简单可行，准确性高。

