



(12) 发明专利

(10) 授权公告号 CN 113314143 B

(45) 授权公告日 2024. 01. 30

(21) 申请号 202110629654.0

(22) 申请日 2021.06.07

(65) 同一申请的已公布的文献号
申请公布号 CN 113314143 A

(43) 申请公布日 2021.08.27

(73) 专利权人 南京优博一创智能科技有限公司
地址 210009 江苏省南京市鼓楼区幕府东路199号A7幢204室

(72) 发明人 竹东翔 程齐明

(74) 专利代理机构 北京知果之信知识产权代理有限公司 11541
专利代理师 卜荣丽 李志刚

(51) Int. Cl.

G10L 25/03 (2013.01)

G10L 25/30 (2013.01)

G10L 25/66 (2013.01)

A61B 5/08 (2006.01)

(56) 对比文件

CN 102499637 A, 2012.06.20

CN 104688229 A, 2015.06.10

CN 104739412 A, 2015.07.01

CN 106691382 A, 2017.05.24

CN 109431470 A, 2019.03.08

CN 111166297 A, 2020.05.19

CN 111345782 A, 2020.06.30

CN 111696575 A, 2020.09.22

CN 111938649 A, 2020.11.17

CN 111938650 A, 2020.11.17

CN 112472066 A, 2021.03.12

US 2012004749 A1, 2012.01.05

US 2018256069 A1, 2018.09.13

US 2020383633 A1, 2020.12.10

孙井鹏; 胡晰远; 彭思龙; 马彦. 鼾声检测研究综述. 世界睡眠医学杂志. 2020, (第03期), 全文.

审查员 丁虎

权利要求书3页 说明书11页 附图2页

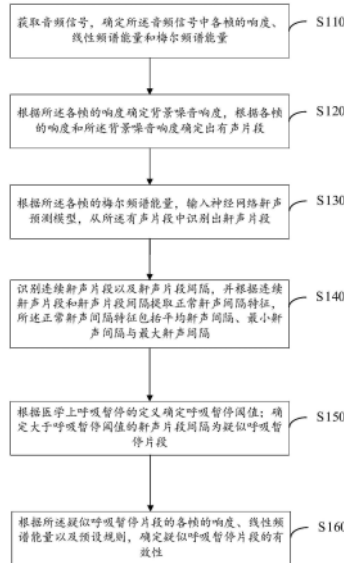
(54) 发明名称

呼吸暂停的判断方法、装置和电子设备

(57) 摘要

本申请公开了一种呼吸暂停的判断方法、装置和电子设备,其中方法包括:获取音频信号;根据各帧的响度确定背景噪音响度,根据各帧的响度和背景噪音响度确定出有声片段;基于神经网络鼾声预测模型,从有声片段中识别出鼾声片段;识别连续鼾声片段以及鼾声片段间隔,并根据连续鼾声片段和鼾声片段间隔提取正常鼾声间隔特征;确定大于呼吸暂停阈值的鼾声片段间隔为疑似呼吸暂停片段;根据疑似呼吸暂停片段的各帧的响度、线性频谱能量以及预设规则,确定疑似呼吸暂停片段的有效性。本申请的通过对鼾声间隔的监测,确定用户是否在顺畅呼吸,该方法逻辑严谨、计算量小、识别准确度高。

CN 113314143 B



1. 一种呼吸暂停的判断方法,其特征在于,包括:
 - 获取音频信号,确定所述音频信号中各帧的响度、线性频谱能量和梅尔频谱能量;
 - 根据所述各帧的响度确定背景噪音响度,根据各帧的响度和所述背景噪音响度确定出有声片段;
 - 根据所述各帧的梅尔频谱能量,输入神经网络鼾声预测模型,从所述有声片段中识别出鼾声片段;
 - 识别连续鼾声片段以及鼾声片段间隔,并根据连续鼾声片段和鼾声片段间隔提取正常鼾声间隔特征,所述正常鼾声间隔特征包括平均鼾声间隔、最小鼾声间隔与最大鼾声间隔;
 - 根据医学上呼吸暂停的定义确定呼吸暂停阈值;确定大于呼吸暂停阈值的鼾声片段间隔为疑似呼吸暂停片段;
 - 根据所述疑似呼吸暂停片段的各帧的响度、线性频谱能量以及预设规则,确定疑似呼吸暂停片段的有效性。
2. 根据权利要求1所述的方法,其特征在于,所述获取音频信号,确定所述音频信号中各帧的响度、线性频谱能量和梅尔频谱能量包括:
 - 以预设频率采集音频信号;
 - 对所述音频信号进行分帧、移帧、短时傅里叶变换获得各帧的线性频谱能量,对所述线性频谱能量进行梅尔频谱变换得到梅尔频谱能量;
 - 和,
 - 根据采样获得的麦克声压信号数据确定各帧的响度。
3. 根据权利要求1所述的方法,其特征在于,所述根据所述各帧的响度确定背景噪音响度包括:
 - 按照预设时长在所述音频信号中截取背景噪音样本片段;
 - 确定组成所述背景噪音样本片段的各帧的响度;
 - 根据所述组成所述背景噪音样本片段的各帧的响度确定所述背景噪音样本片段的响度均值和响度方差;
 - 将所述响度方差与预设平稳噪声上限阈值比较,在所述响度方差小于预设平稳噪声上限阈值的情况下,将所述背景噪音样本片段的响度均值作为所述背景噪音响度。
4. 根据权利要求1所述的方法,其特征在于,所述根据各帧的响度和所述背景噪音响度确定出有声片段包括:
 - 在一帧的响度大于所述背景噪音响度与波动响度的加和情况下,确定该帧为所述有声片段的起始帧;
 - 在预设数量的连续帧的响度小于所述背景噪音响度与波动响度的加和情况下,确定所述连续帧的第一帧为所述有声片段的截止帧;
 - 将所述起始帧与所述截止帧之间的各帧作为有声片段。
5. 根据权利要求1所述的方法,其特征在于,提取所述各帧的梅尔频谱能量,输入神经网络鼾声预测模型,从所述有声片段中识别出鼾声片段包括:
 - 将组成所述有声片段的各帧的梅尔频谱能量作为输入值输入到所述神经网络鼾声预测模型,得到各帧的鼾声概率值,将高于鼾声概率阈值的帧标记为鼾声帧;
 - 将包含有鼾声帧的有声片段标记为鼾声片段。

6. 根据权利要求5所述的方法,其特征在于,所述根据所述各帧的梅尔频谱能量,输入神经网络鼾声预测模型,从所述有声片段中识别出鼾声片段还包括:

从所述有声片段确定出连续疑似鼾声片段;

在所述连续疑似鼾声片段满足预设鼾声规则的条件下,将连续疑似鼾声片段标记为鼾声片段,其中,所述预设鼾声规则为:相邻的连续疑似鼾声片段的起始帧之间的间隔时间处于第一预设时长范围。

7. 根据权利要求6所述的方法,其特征在于,所述根据所述各帧的梅尔频谱能量,输入神经网络鼾声预测模型,从所述有声片段中识别出鼾声片段还包括:

根据所述有声片段和所述鼾声片段,确定正常相邻的连续疑似鼾声片段;

确定所述正常相邻的连续疑似鼾声片段的起始帧之间的间隔的平均值、最小值与最大值,作为个性化呼吸特征数据进行保存。

8. 根据权利要求1所述的方法,其特征在于,所述根据所述疑似呼吸暂停片段的各帧的响度、线性频谱能量以及预设规则,确定疑似呼吸暂停片段的有效性包括:

在所述疑似呼吸暂停片段的各帧的响度、线性频谱能量满足第一预设规则的情况下,确定所述疑似呼吸暂停片段为疑似身体在床上移动片段,其中,所述第一预设规则为:组成所述疑似身体在床上移动片段的各帧的响度大于背景噪音响度与第一预设响度阈值的加和,且线性频谱能量在整个频段能量分布均匀;

在所述疑似身体在床上移动片段的时长大于第一预设时长阈值的情况下,确定所述疑似呼吸暂停片段的标记为无效。

9. 根据权利要求1所述的方法,其特征在于,所述根据所述疑似呼吸暂停片段的各帧的响度、线性频谱能量以及预设规则,确定疑似呼吸暂停片段的有效性还包括:

在所述疑似呼吸暂停片段的各帧的响度、线性频谱能量满足第二预设规则的情况下,确定所述疑似呼吸暂停片段为疑似呼吸片段,其中,所述第二预设规则为:组成疑似呼吸片段的各帧的响度大于背景噪音响度与第二预设响度阈值的加和,或0-1K各线性频点频谱能量之和大于背景噪音0-1K各线性频点频谱能量和叠加第一预设0-1K各线性频点频谱能量加和阈值;

在所述疑似呼吸片段的时长小于第二时长阈值的情况下,同时相邻的疑似呼吸片段特征的起始帧之间的间隔时间处于第二预设时长范围,确定所述疑似呼吸暂停片段标记为无效。

10. 一种呼吸暂停的判断装置,其特征在于,包括:

获取单元,用于获取音频信号,确定所述音频信号中各帧的响度、线性频谱能量和梅尔频谱能量;

第一识别单元,用于根据所述各帧的响度确定背景噪音响度,根据各帧的响度和所述背景噪音响度确定出有声片段;

第二识别单元,用于根据所述各帧的梅尔频谱能量,输入神经网络鼾声预测模型,从所述有声片段中识别出鼾声片段;

第三识别单元,用于识别连续鼾声片段以及鼾声片段间隔,并根据连续鼾声片段和鼾声片段间隔提取正常鼾声间隔特征,所述正常鼾声间隔特征包括平均鼾声间隔、最小鼾声间隔与最大鼾声间隔;

第四识别单元,用于根据医学上呼吸暂停的定义确定呼吸暂停阈值;确定大于呼吸暂停阈值的鼾声片段间隔为疑似呼吸暂停片段;

判断单元,用于根据所述疑似呼吸暂停片段的各帧的响度、线性频谱能量以及预设规则,确定疑似呼吸暂停片段的有效性。

呼吸暂停的判断方法、装置和电子设备

技术领域

[0001] 本申请涉及人工智能技术领域,尤其涉及呼吸暂停的判断方法、装置和电子设备。

背景技术

[0002] 呼吸暂停综合症是一种严重危害健康的睡眠障碍,医学上通过鼻息气流来探测,佩戴设备会影响用户的睡眠状态,用户面临监测的心理门槛,这大大影响了筛查的普及率。

[0003] 目前市场上有基于压电原理、电磁波/雷达波监测呼吸,面临的问题在于非神经中枢性呼吸暂停的用户在呼吸暂停情况下,同样有胸廓的运动,影响了识别的精确率。智能手表通过监测血氧来发现呼吸暂停,面临的困难在于连续血氧监测的精确度与耗电的设计矛盾,用户夜晚佩戴感觉不舒适。而这些智能设备面临的另外一个巨大问题在于,硬件本身的购买成本将大部分用户拒之门外,呼吸暂停综合症初筛市场需要方便便宜的方案。

[0004] 采用声音进行呼吸暂停综合症的初筛是一种可行高效的手段,一方面,用户无佩戴不适的心理问题,另外一方面,智能手机的普及化大大降低用户对监测成本的担忧。但如何准确识别鼾声和呼吸声,是技术的关键。

[0005] 目前很多相关产品/专利聚焦于呼吸暂停或低通气发生前后频谱特征进行辨识,但是普遍存在计算量大、识别准确度低等问题。

发明内容

[0006] 本申请实施例提供了一种呼吸暂停的判断方法、装置和电子设备,以解决或至少部分解决上述问题。

[0007] 根据本申请的第一方面,提供了一种呼吸暂停的判断方法,包括:

[0008] 获取音频信号,确定所述音频信号中各帧的响度、线性频谱能量和梅尔频谱能量;

[0009] 根据所述各帧的响度确定背景噪音响度,根据各帧的响度和所述背景噪音响度确定出有声片段;

[0010] 根据所述各帧的梅尔频谱能量,输入神经网络鼾声预测模型,从所述有声片段中识别出鼾声片段;

[0011] 识别连续鼾声片段以及鼾声片段间隔,并根据连续鼾声片段和鼾声片段间隔提取正常鼾声间隔特征,所述正常鼾声间隔特征包括平均鼾声间隔、最小鼾声间隔与最大鼾声间隔;

[0012] 根据医学上呼吸暂停的定义确定呼吸暂停阈值;确定大于呼吸暂停阈值的鼾声片段间隔为疑似呼吸暂停片段;

[0013] 根据所述疑似呼吸暂停片段的各帧的响度、线性频谱能量以及预设规则,确定疑似呼吸暂停片段的有效性。

[0014] 根据本申请的第二方面,提供了一种呼吸暂停的判断装置,包括:

[0015] 获取单元,用于获取音频信号,确定所述音频信号中各帧的响度、线性频谱能量和梅尔频谱能量;

[0016] 第一识别单元,用于根据所述各帧的响度确定背景噪音响度,根据各帧的响度和所述背景噪音响度确定出有声片段;

[0017] 第二识别单元,用于根据所述各帧的梅尔频谱能量,输入神经网络鼾声预测模型,从所述有声片段中识别出鼾声片段;

[0018] 第三识别单元,用于识别连续鼾声片段以及鼾声片段间隔,并根据连续鼾声片段和鼾声片段间隔提取正常鼾声间隔特征,所述正常鼾声间隔特征包括平均鼾声间隔、最小鼾声间隔与最大鼾声间隔;

[0019] 第四识别单元,用于根据医学上呼吸暂停的定义确定呼吸暂停阈值;确定大于呼吸暂停阈值的鼾声片段间隔为疑似呼吸暂停片段;

[0020] 判断单元,用于根据所述疑似呼吸暂停片段的各帧的响度、线性频谱能量以及预设规则,确定疑似呼吸暂停片段的有效性。

[0021] 根据本申请的另一方面,提供了一种电子设备,包括:处理器;以及被安排成存储计算机可执行指令的存储器,所述可执行指令在被执行时使所述处理器执行上述任一的方法。

[0022] 本申请实施例采用的上述至少一个技术方案能够达到以下有益效果:

[0023] 首先根据组成音频信号各帧的响度确定当前环境的背景噪音响度,在背景噪音响度以上的片段被标记为有声片段,进一步的,基于神经网络鼾声预测模型,从有声片段中识别出鼾声片段,在鼾声片段中,将鼾声片段间隔超过一定阈值的片段识别为疑似呼吸暂停片段,根据间隔段各帧的响度、线性频谱能量和预设规则判断呼吸是否停止。本申请的通过神经网络鼾声预测模型准确识别鼾声片段,进一步的在鼾声片段中识别出鼾声间隔,通过对鼾声间隔的监测,确定用户是否在顺畅呼吸,为监控用户的健康状态提供了坚实的基础,该方法逻辑严谨、计算量小、识别准确度高。

附图说明

[0024] 此处所说明的附图用来提供对本申请的进一步理解,构成本申请的一部分,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。在附图中:

[0025] 图1为根据本申请的一个实施例的呼吸暂停的判断方法的流程示意图;

[0026] 图2为根据本申请的一个实施例的呼吸暂停的判断装置的结构示意图;

[0027] 图3为本申请实施例中一种电子设备的结构示意图。

具体实施方式

[0028] 为使本申请的目的、技术方案和优点更加清楚,下面将结合本申请具体实施例及相应的附图对本申请技术方案进行清楚、完整地描述。显然,所描述的实施例仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0029] 以下结合附图,详细说明本申请各实施例提供的技术方案。

[0030] 图1示出了根据本申请的一个实施例的呼吸暂停的判断方法的流程示意图,从图1可以看出,该方法至少包括步骤S110~S160:

[0031] 需要说明的是,本申请主要针对鼾声过程中出现的呼吸暂停进行判断,这是由于

人们发生呼吸障碍通常是在打鼾的过程中,这是由于打鼾时气道被阻碍,容易发生呼吸暂停。

[0032] 步骤S110:获取音频信号,确定音频信号中各帧的线性频谱能量、梅尔频谱能量和响度。

[0033] 音频信号可以但不限于通过智能终端收集用户睡觉时声音信号。如以16000Hz采集智能终端的麦克的音频信号,得到音频信号的时间序列,以16000Hz、以16位表示一个采集点信号大小、单声道为例,音频信号为(2,4,100,120,140,60,-60,-130,...),每个点的间隔时间=1/16000秒。

[0034] 对于音频信号中各帧的线性频谱能量、梅尔频谱能量的获得可根据下述步骤:以预设频率采集音频信号中各采样点的声压幅值;对所述音频信号进行分帧、移帧、傅里叶变换和梅尔频率变换,确定组成各帧的多个线性频点和对应的梅尔频率及对应的线性频谱能量、梅尔频谱能量。

[0035] 将音频信号的采样时间序列截取连续N个数据做为一组,这一组数据就叫做一帧数据,如每次取出连续512个数据,这叫做一帧数据,这个过程即为分帧,每次取出的数据的数量可根据计算量设置,通常为512或1024。具体在一帧数据中,可以获得的哪些频点的数据,这与频点分辨率有关,如以频点分辨率为16000,每次取出512个数据为例,由于 $16000/512=31.25\text{Hz}$,即在0-8000Hz频域中,只能获得 $31.25*N$ 的频点的信息, $N=1-256$ 的整数。

[0036] 在本申请中,为了提高检测精度,在分帧的过程中,每次取值并不是从前一帧数据的尾部开始,而是从前一帧数据的中间位置开始,在本实施例中,第二帧是从第一帧的中间位置开始的,即257个频点开始,取512个频点对应的幅值。为了方便处理,为每帧音频信号设置帧号,顺序递增。

[0037] 组成一帧音频信号中的各采样点对应的声压幅值进行短时傅里叶变换,并根据时间的先后顺序组合,就组成了一帧音频信号的频谱能量。即每一帧的频谱能量可以使用一维数组来表示($a_1, a_2, a_3, \dots, a_{256}$),分别对应31.25Hz,62.5Hz,93.75Hz,...,8000Hz的频谱能量幅值,然后使用梅尔滤波器组获得梅尔频谱能量,梅尔频谱与线性频谱的关系可参考现有技术。

[0038] 进一步的,根据各采样点的声压幅值可确定出各帧的响度,具体计算公式可参考现有技术。

[0039] 步骤S120:根据各帧的响度确定背景噪音响度;根据各帧的响度和背景噪音响度确定出有声片段。

[0040] 人们睡觉的环境通常存在背景噪音,背景噪音通常是连续不断的,并且响度平稳,因此,首先要排除背景噪音的影响。由于不同的用户所处的环境是不同的,因此背景噪音的大小也不相同,为了准确排除背景噪音,首先对背景噪音响度进行计算。

[0041] 具体的,可根据各帧的响度计算得到背景噪音响度,如将连续5秒内音频信号中所有帧的响度排列为1维数组,计算其均值和方差,如果方差小于预设平稳上限阈值,则作为背景噪音,响度的均值作为背景噪音响度。

[0042] 承上,能够根据各帧的响度确定背景噪音响度,如,在某一帧的响度大于背景噪音响度则将该帧标记为有声帧,某一帧的响度小于等于背景噪音响度则将该帧标记为背景帧,将有声帧或者连续的有声帧作为有声片段。

[0043] 需要说明的是,上述说明仅作为示例,为了提高检测精度,可将有声片段的检测规则制定的更加严格,本申请不做限制。

[0044] 步骤S130:根据各帧的梅尔频谱能量,输入神经网络鼾声预测模型,从所述有声片段中识别出鼾声片段。

[0045] 有声片段可能是梦呓,可能是鼾声,也可能是其他。因此,需要在有声片段中识别出鼾声片段。

[0046] 对于鼾声的识别可通过神经网络鼾声预测模型进行,该模型为二分类模型,将各帧的梅尔频谱能量作为输入值,输入到神经网络鼾声预测模型,输出结果是有声片段为鼾声片段或非鼾声片段。

[0047] 在本申请中,神经网络鼾声预测模型是基于多层的神经网络建立的,如将1层全连接层、1层长短期记忆网络层(LSTM,Long Short-Term Memory)、1层全连接层和1层逻辑回归层(softmax)结合起来建立神经网络鼾声预测模型。

[0048] 步骤S140,识别连续鼾声片段以及鼾声片段间隔,并根据连续鼾声片段和鼾声片段间隔提取正常鼾声间隔特征,所述正常鼾声间隔特征包括平均鼾声间隔、最小鼾声间隔与最大鼾声间隔。确定所述有声片段中的连续鼾声片段,非连续鼾声片段之间的部分为鼾声片段间隔,每个人的睡觉呼吸习惯不同,打鼾具有一定的规律性,根据连续鼾声片段,以及连续鼾声片段之间的间隔,进行特征提取,可获得一个用户的正常鼾声间隔特征,正常鼾声间隔特征包括平均鼾声间隔、最小鼾声间隔与最大鼾声间隔。

[0049] 步骤S150,根据医学上呼吸暂停的定义确定呼吸暂停阈值;确定大于呼吸暂停阈值的鼾声片段间隔为疑似呼吸暂停片段。

[0050] 根据医学上呼吸暂停的定义,呼吸暂停阈值在10s。

[0051] 进一步的,可确定大于呼吸暂停阈值的鼾声片段间隔为疑似呼吸暂停片段如鼾声片段间隔超过10秒以上的区间段,将所述鼾声间隔超过10秒以上的区间段作为疑似呼吸暂停片段。

[0052] 在人们连续打鼾时,说明人们并没有发生呼吸障碍,只有在鼾声与鼾声间隔超过10秒以上的情况下,才有可能发生了呼吸暂停。因此,将鼾声片段切割为连续鼾声帧和鼾声间隔超过10秒以上的区间段,包括但不限于根据组成鼾声片段的各帧的响度和各帧的编号进行判断,如设定一预设鼾声阈值,在某帧的响度大于该预设鼾声阈值时,认为该帧为鼾声帧,若多帧鼾声帧的帧号连续,则确定该多帧鼾声帧为连续鼾声段。响度小于预设鼾声阈值的一帧或者连续多帧就组成了鼾声间隔段,即一声鼾声与另一声鼾声之间的间隔,该间隔超过10秒以上即为疑似呼吸暂停片段。

[0053] 步骤S160,根据所述疑似呼吸暂停片段的各帧的响度、线性频谱能量以及预设规则,确定疑似呼吸暂停片段的有效性。

[0054] 在确定疑似呼吸暂停片段,对其有效性进行判断,若无效,则认为没有发生呼吸暂停,若有效,则认为发生了呼吸暂停。

[0055] 在本申请的一些实施例,呼吸暂停片段的有效性可以使用以下3个条件,但不限于这3个条件:

[0056] 首先使用响度和0-1K线性频点频谱能量和两个条件,发现响度或0-1K频谱能量小幅波动的有声片段。如设定背景噪音响度加更小的响度阈值作为响度波动阈值;找出5秒内

0-1K线性频点频谱能量和的方差小于预设阈值作为背景噪音0-1K线性频点频谱能量和,叠加预设阈值作为0-1K线性频点频谱能量和平稳背景噪音上限阈值;超过2个条件之一则标识为有声帧;连续3个疑似的有声帧组成有声片段;

[0057] 第一,如果发现疑似呼吸暂停片段中存在身体在床上移动的声音,则该疑似呼吸暂停片段无效。在本申请中,将有声片段起始帧的线性频点的频谱能量组成1维数组,后续有声片段中的有声帧的线性频点频谱能量不断叠加,最终除以有声帧的数量,得到有声片段平均单帧的线性频点的频谱能量。如果0-8K各个线性频点的频谱能量起伏小于白噪音起伏预设阈值,同时有声片段的时间长度超过2秒以上,识别为身体在床上移动的声音;

[0058] 第二,有声片段的时间间隔符合连续鼾声统计得到的呼吸特征,即间隔满足在最小和最大呼吸间隔范围内,同时有声片段的时间长度在呼吸预设长度范围内。如果满足条件,则疑似呼吸暂停段无效;

[0059] 第三,预设规则:疑似呼吸暂停的时长达到或超过10秒,否则无效。

[0060] 需要说明的是,上述说明仅作为示例,为了提高检测精度,可将疑似呼吸暂停片段有效性的检测规则制定的更加严格,本申请不做限制。

[0061] 从图1所示的方法可以看出,本申请首先根据组成音频信号的各帧的响度确定当前环境的背景噪音响度,在背景噪音响度以上的片段被标记为有声片段,进一步的,基于神经网络鼾声预测模型,从有声片段中识别出鼾声片段,在鼾声片段间隔中,识别出疑似呼吸暂停片段,根据各帧的响度、线性频谱能量和预设规则判断呼吸是否停止。本申请的通过神经网络鼾声预测模型准确识别鼾声片段,进一步的在鼾声片段中识别出鼾声间隔,通过对鼾声间隔的监测,确定用户是否在顺畅呼吸,为监控用户的健康状态提供了坚实的基础,该方法逻辑严谨、计算量小、识别准确度高。

[0062] 在本申请的一些实施例中,在上述方法中,根据各帧的响度确定背景噪音响度包括:按照预设时长在音频信号中截取背景噪音样本片段;确定组成背景噪音样本片段的各帧的响度;根据组成背景噪音样本片段的各帧的响度确定背景噪音样本片段的响度均值和响度方差。将响度方差与预设背景噪音方差上限阈值比较,在响度方差小于预设背景噪音方差上限阈值的情况下,将所背景噪音样本片段的响度均值作为背景噪音响度。在对背景噪音响度进行确定的时候,如果以音频信号为全样本进行,计算量会非常大,由于人们梦呓或打鼾的声音通常不会持续睡觉的全过程,因此,可在音频信号中截取当前时刻向前预设时长的数据作为背景噪音样本片段,如此可以适应背景噪音不断变换的环境,黎明的小区噪音比深夜的小区噪音是很常见的现象。

[0063] 具体的,按照预设时长在音频信号中截取背景噪音样本片段,截取最近的音频部分,如截取最近连续5s时长内的音频信号,作为背景噪音样本片段。

[0064] 对于背景噪音的确定可参考现有技术,也可以采用本申请推荐的方法,具体的,确定组成背景噪音样本片段的各帧的响度,根据各帧的响度,确定出背景噪音样本片段整体的响度均值和响度方差,统计中的方差是每个样本值与全体样本值的平均数之差的平方值的平均数,方差能够表征平稳度。可根据方差的定义与各帧的响度计算出背景噪音样本片段整体的响度方差。

[0065] 在响度方差小于预设背景噪音方差上限阈值的情况下,则认为该背景噪音样本片段中没有鼾声和梦呓等有声片段,而只有背景噪音,进一步的,将该背景噪音样本片段的响

度均值作为背景噪音响度。均值和方差都是通常意义上的定义,响度均值和响度方差计算方法请参考现有技术。

[0066] 在响度方差大于于预设背景噪音方差上限阈值的情况下,则认为该背景噪音样本片段有除了背景噪音以外,其他声音的存在,则舍弃该背景噪音样本片段,重新截取。

[0067] 在本申请的一些实施例中,在上述方法中,根据各帧的响度和背景噪音响度确定出有声片段包括:在一帧的响度大于背景噪音响度与波动响度的加和情况下,确定该帧为有声片段的起始帧;在预设数量的连续帧的响度小于背景噪音响度与波动响度的加和情况下,确定连续帧的第一帧为有声片段的截止帧;将起始帧与截止帧之间的各帧作为有声片段。

[0068] 对于有声片段的识别,可以通过有声片段的起始帧和截止帧来实现,为了提高检测精度,本实施例中,在背景噪音响度的基础上增加一个波动响度,该波动响度如可以为4-6分贝,在某一帧前面的一帧或多帧不符合有声片段的认定规则的情况下,在该帧的响度大于背景噪音响度与波动响度的加和时,则认为该帧为有声片段的起始帧;对于截止帧的判断,可在预设数量的连续帧的响度小于背景噪音响度与波动响度的加和情况下,确定连续帧的第一帧为有声片段的截止帧,举例来讲,连续3帧的响度均小于背景噪音响度与波动响度的加和,则认为有声片段截止,将连续3帧的第一帧作为有声片段的截止帧。最后将起始帧与截止帧之间的所有帧作为有声片段。在本申请的一些实施例中,在上述方法中,根据各帧的梅尔频谱能量,输入神经网络鼾声预测模型,从有声片段中识别出鼾声片段包括:将组成有声片段的各帧的梅尔频谱能力作为输入值输入到神经网络鼾声预测模型,得到各帧的鼾声概率值;在鼾声概率值大于于预设概率阈值的情况下,将帧设为鼾声帧,包含鼾声帧的有声片段标记为鼾声片段。

[0069] 在对鼾声片段的识别时,可通过神经网络鼾声预测模型,输入值为组成有声片段的各帧的梅尔频谱能量,输出值为有声片段的各帧的鼾声概率值。

[0070] 将鼾声概率值与预设概率阈值进行比较,如将预设概率阈值设为0.35,在鼾声概率值大于等于预设概率阈值的情况下,将帧标记为鼾声帧,包含鼾声帧的有声片段标记为疑似鼾声片段。

[0071] 对于没有被标记为疑似鼾声片段的有声片段,即未包含鼾声帧的有声片段,将这部分有声片段虽然没有被神经网络鼾声预测模型标记为鼾声片段,但是为了提高检测的准确性,本实施例结合鼾声的节律特征和均值线性频谱能量相似特征对鼾声进一步检测提取。首先,从疑似鼾声片段确定出连续疑似鼾声片段,具体的可以根据相邻鼾声是否符合呼吸特征进行识别,如果前一鼾声片段的起始帧的帧号与后一鼾声片段的起始帧的帧号不超过预设阈值内,如10秒,则可认为是连续疑似鼾声片段。

[0072] 在连续疑似鼾声片段满足预设鼾声规则的条件下,将连续疑似鼾声片段标记为鼾声片段,其中,所述预设鼾声规则为:相邻疑似鼾声起始帧相距在呼吸预设阈值内,如2-7秒。

[0073] 对于未识别为疑似鼾声的有声段,为了提高检测的准确性,采用预设鼾声规则对鼾声进一步检测提取。

[0074] 预设鼾声规则可以理解为两个子鼾声规则,记为鼾声规则一和鼾声规则二,鼾声规则一表征鼾声的节律特征,具体内容为若连续有声片段的起始帧间距帧的时间处于预设

时长范围,这个范围是人呼吸的正常范围,预设时长范围可以为但不限于2-7s,且连续3个时间间隔保持稳定,如时间间隔的方差小于1秒,则满足鼾声规则一,否则,则不满足。

[0075] 鼾声规则二表征线性频谱能量相似特征,选择最近4个连续有声片段作为样本,将每个有声片段起始帧到截止帧每一帧的线性频谱能量叠加,再除以起始帧到截止帧的帧数量,得到有声片段的帧均值线性频谱能量,再计算线性频谱能量在线性频点的局部极大值,并将对应的频点放到特征频点数组中,如果4个连续有声片段的特征频点都保持一致,则满足鼾声规则二,否则,则不满足。

[0076] 在本申请的一些实施例中,所述根据所述疑似呼吸暂停片段的各帧的响度、线性频谱能量以及预设规则,确定疑似呼吸暂停片段的有效性包括:

[0077] 根据所述疑似呼吸暂停片段的各帧的响度和线性频谱能量,确定疑似动作片段,其中,组成所述疑似动作片段的各帧的响度大于第一预设响度阈值,同时帧均值线性频谱能量在全频段中保持平稳,在所述疑似动作片段的时长大于第一预设时长阈值的情况下,确定所述疑似呼吸暂停片段标记为无效。

[0078] 人们在打鼾的过程中,也会发生转身、移动等动作,通常发生动作的时候打鼾就会停止,因此,在对疑似呼吸暂停片段的有效性进行检测的时候,最好排除动作的影响。具体的,将有声片段起始帧的线性频点的频谱能量组成1维数组,叠加有声片段中的所有帧的线性频点频谱能量,最后除以有声帧的数量,得到有声片段平均单帧的线性频点的频谱能量。如果0-8K各个线性频点的频谱能量起伏小于白噪音起伏预设阈值,同时有声片段的时间长度超过2秒以上,识别为身体在床上移动,而非呼吸暂停。

[0079] 在本申请的一些实施例中,所述根据所述疑似呼吸暂停片段的各帧的响度以及预设规则,确定疑似呼吸暂停片段的有效性还包括:根据所述疑似呼吸暂停片段的各帧的响度,确定疑似呼吸片段,其中组成疑似呼吸片段的各帧的响度大于第二预设响度阈值的,有声片段的时间间隔符合连续鼾声统计得到的呼吸特征,即间隔满足在最小和最大呼吸间隔范围内,同时有声片段的时间长度在呼吸预设长度范围内。如果满足条件,确定所述疑似呼吸暂停片段标记为无效。

[0080] 进一步,人们在打鼾的过程中,也可能由于发生动作、环境变化等原因突然停止打鼾而进入正常呼吸状态,这种情况下,人们会均匀呼吸,且呼吸声音很小,因此需要排除人们进入正常呼吸状态的影响。

[0081] 同理,也可以通过设定一个合适的第二预设响度阈值实现,这里需要说明的是,第二预设响度阈值的值是比较小的,如4分贝,且第二预设响度阈值小于第一预设响度阈值。响度大于该第二预设响度阈值的各帧组成了疑似呼吸片段,进一步的,在所述疑似呼吸片段的时长小于第二时长阈值的情况下,确定所述疑似呼吸暂停片段标记为无效。

[0082] 第二时长阈值可以根据用户的打鼾的规律进行计算,如统计鼾声片段中,非连续鼾声帧,即鼾声间隔的时长,得出最小时长和最大时长,根据最小时长和最大时长加权平均得出。

[0083] 图2示出了根据本申请一个实施例的呼吸暂停的判断装置,从图2可以看出,该装置200包括:

[0084] 获取单元210,用于获取音频信号,确定所述音频信号中各帧的响度、线性频谱能量和梅尔频谱能量;

[0085] 第一识别单元220,用于根据所述各帧的响度确定背景噪音响度,根据各帧的响度和所述背景噪音响度确定出有声片段;

[0086] 第二识别单元230,用于根据所述各帧的梅尔频谱能量,输入神经网络鼾声预测模型,从所述有声片段中识别出鼾声片段;

[0087] 第三识别单元240,用于识别连续鼾声片段以及鼾声片段间隔,并根据连续鼾声片段和鼾声片段间隔提取正常鼾声间隔特征,所述正常鼾声间隔特征包括平均鼾声间隔、最小鼾声间隔与最大鼾声间隔;

[0088] 第四识别单元250,用于根据医学上呼吸暂停的定义确定呼吸暂停阈值;确定大于呼吸暂停阈值的鼾声片段间隔为疑似呼吸暂停片段;

[0089] 判断单元260,用于根据所述疑似呼吸暂停片段的各帧的响度、线性频谱能量以及预设规则,确定疑似呼吸暂停片段的有效性。

[0090] 在本申请的一些实施例中,在上述装置中,获取单元210,用于以预设频率采集音频信号;对所述音频信号进行分帧、移帧、短时傅里叶变换获得各帧的线性频谱能量,对所述线性频谱能量进行梅尔频谱变换得到梅尔频谱能量;和,根据采样获得的麦克声压信号数据确定各帧的响度。。

[0091] 在本申请的一些实施例中,在上述装置中,用于按照预设时长在所述音频信号中截取背景噪音样本片段;确定组成所述背景噪音样本片段的各帧的响度;根据所述组成所述背景噪音样本片段的各帧的响度确定所述背景噪音样本片段的响度均值和响度方差;将所述响度方差与预设鼾声阈值比较,在所述响度方差小于预设鼾声阈值的情况下,将所述背景噪音样本片段的响度均值作为所述背景噪音响度。

[0092] 在本申请的一些实施例中,在上述装置中,第一识别单元220,用于在一帧的响度大于所述背景噪音响度与波动响度的加和情况下,确定该帧为所述有声片段的起始帧;在预设数量的连续帧的响度大于所述背景噪音响度与波动响度的加和情况下,确定所述连续帧的第一帧为所述有声片段的截止帧;将所述起始帧与所述截止帧之间的各帧的幅值谱作为有声片段。

[0093] 在本申请的一些实施例中,在上述装置中,第二识别单元230,用于将组成所述有声片段的各帧的梅尔频谱能量作为输入值输入到所述神经网络鼾声预测模型,得到各有声帧的鼾声概率值;在所述鼾声概率值大于等于预设概率阈值的情况下,将所述有声帧标记为鼾声帧,包含鼾声帧的有声片段标记为鼾声片段。

[0094] 在本申请的一些实施例中,在上述装置中,第二识别单元230,还用于从所述鼾声片段确定出连续疑似鼾声片段;在所述连续疑似鼾声片段满足预设鼾声规则的条件下,将连续疑似鼾声片段标记为鼾声片段,其中,所述预设鼾声规则为:所述连续疑似鼾声片段中相邻鼾声片段的起始帧间隔的时间处于预设时长范围;另外,为了提高检测的准确性,结合鼾声的节律特征和均值线性频谱能量相似特征对连续有声段中对鼾声进一步检测提取,除了前面提及的预设鼾声规则外,组成所述连续有声片段的各个有声片段的单帧均值线性频谱能量的局部极大值完全一致,即满足均值线性频谱能量相似特征,也标记为鼾声片段。

[0095] 在本申请的一些实施例中,在上述装置中,判定单元250,用于在所述疑似呼吸暂停片段的各帧的响度、线性频谱能量满足第一预设规则的情况下,确定所述疑似呼吸暂停片段为疑似身体在床上移动片段,其中,所述第一预设规则为:组成所述疑似身体在床上移

动片段的各帧的响度大于背景噪音响度与第一预设响度阈值的加和,且线性频谱能量在整个频段能量分布均匀;在所述疑似身体在床上移动片段的时长大于第一预设时长阈值的情况下,确定所述疑似呼吸暂停片段的标记为无效。

[0096] 在本申请的一些实施例中,在上述装置中,判定单元250,用于在所述疑似呼吸暂停片段的各帧的响度、线性频谱能量满足第二预设规则的情况下,确定所述疑似呼吸暂停片段为疑似呼吸片段,其中,所述第二预设规则为:组成疑似呼吸片段的各帧的响度大于背景噪音响度与第二预设响度阈值的加和,或0-1K各线性频点频谱能量之和大于背景噪音0-1K各线性频点频谱能量和叠加第一预设0-1K各线性频点频谱能量加和阈值;

[0097] 在所述疑似呼吸片段的时长小于第二时长阈值的情况下,同时相邻的疑似呼吸片段特征的起始帧之间的间隔时间处于第二预设时长范围,确定所述疑似呼吸暂停片段标记为无效。

[0098] 能够理解,上述装置,能够实现前述实施例中提供的方法的各个步骤,关于方法的相关阐释均适用于装置,此处不再赘述。

[0099] 图3是本申请的一个实施例电子设备的结构示意图。请参考图3,在硬件层面,该电子设备包括处理器,可选地还包括内部总线、网络接口、存储器。其中,存储器可能包含内存,例如高速随机存取存储器(Random-Access Memory, RAM),也可能还包括非易失性存储器(non-volatile memory),例如至少1个磁盘存储器等。当然,该电子设备还可能包括其他业务所需要的硬件。

[0100] 处理器、网络接口和存储器可以通过内部总线相互连接,该内部总线可以是ISA (Industry Standard Architecture,工业标准体系结构)总线、PCI (Peripheral Component Interconnect,外设部件互连标准)总线或EISA (Extended Industry Standard Architecture,扩展工业标准结构)总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示,图3中仅用一个双向箭头表示,但并不表示仅有一根总线或一种类型的总线。

[0101] 存储器,用于存放程序。具体地,程序可以包括程序代码,所述程序代码包括计算机操作指令。存储器可以包括内存和非易失性存储器,并向处理器提供指令和数据。

[0102] 处理器从非易失性存储器中读取对应的计算机程序到内存中然后运行,在逻辑层面上形成呼吸暂停的判断装置。处理器,执行存储器所存放的程序。

[0103] 上述如本申请图3所示实施例揭示的呼吸暂停的判断装置执行的方法可以应用于处理器中,或者由处理器实现。处理器可能是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器,包括中央处理器(Central Processing Unit, CPU)、网络处理器(Network Processor, NP)等;还可以是数字信号处理器(Digital Signal Processor, DSP)、专用集成电路(Application Specific Integrated Circuit, ASIC)、现场可编程门阵列(Field-Programmable Gate Array, FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本申请实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,

闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器,处理器读取存储器中的信号,结合其硬件完成上述方法的步骤。

[0104] 该电子设备还可执行图3中呼吸暂停的判断装置执行的方法,并实现呼吸暂停的判断装置在图3所示实施例的功能,本申请实施例在此不再赘述。

[0105] 本申请实施例还提出了一种计算机可读存储介质,该计算机可读存储介质存储一个或多个程序,该一个或多个程序包括指令,该指令当被包括多个应用程序的电子设备执行时,能够使该电子设备执行图3所示实施例中呼吸暂停的判断装置执行的方法。

[0106] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0107] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0108] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0109] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0110] 在一个典型的配置中,计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

[0111] 内存可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

[0112] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信号存储。信号可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括,但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带,磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信号。按照本文中的界定,计算

机可读介质不包括暂存电脑可读媒体(transitory media),如调制的数据信号和载波。

[0113] 还需要说明的是,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

[0114] 本领域技术人员应明白,本申请的实施例可提供为方法、系统或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0115] 以上所述仅为本申请的实施例而已,并不用于限制本申请。对于本领域技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等,均应包含在本申请的权利要求范围之内。

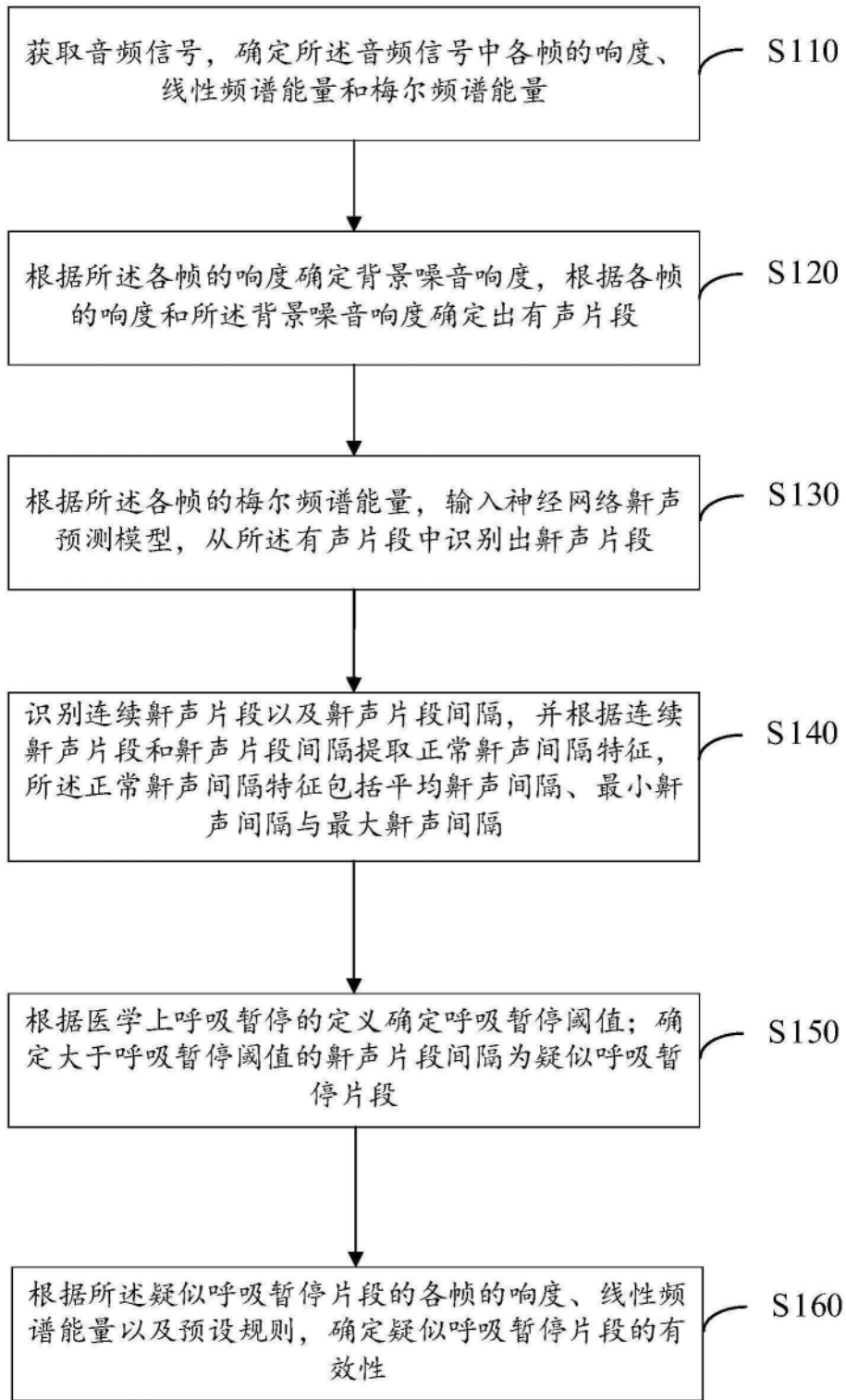


图1

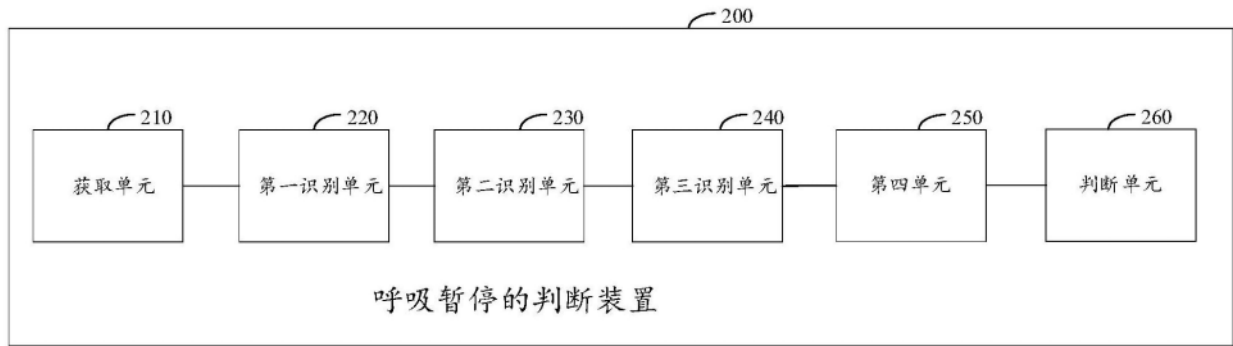


图2

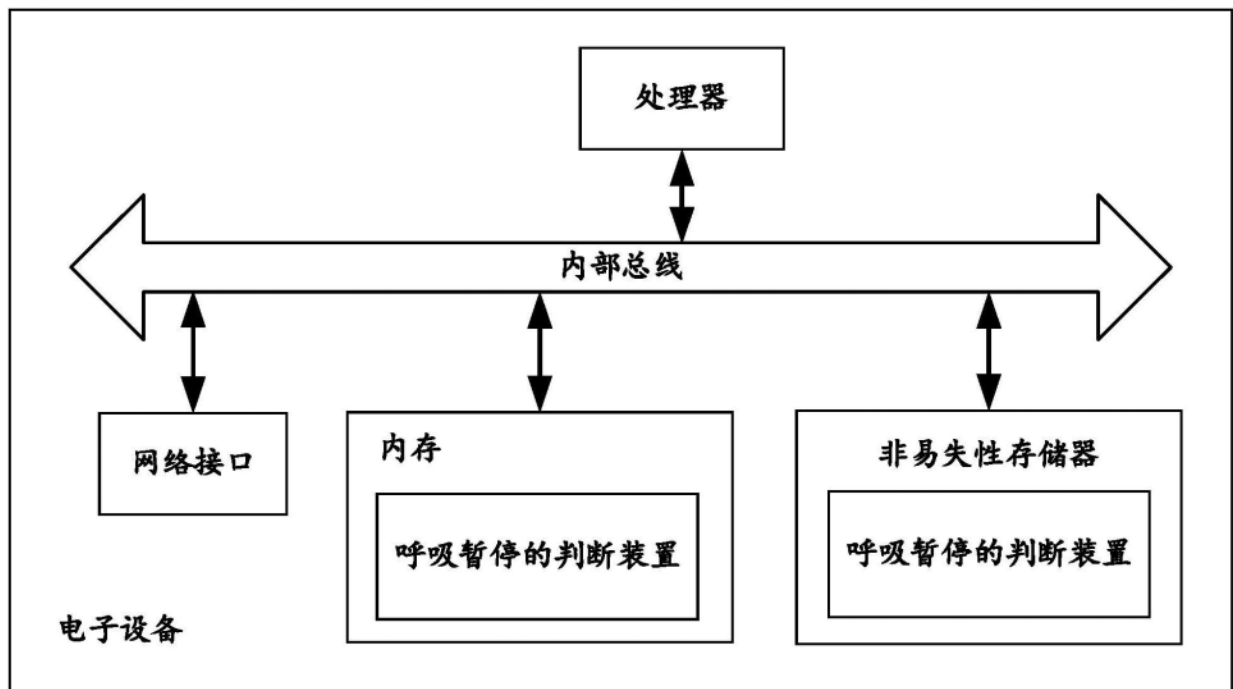


图3