



(12) 发明专利

(10) 授权公告号 CN 111161748 B

(45) 授权公告日 2022. 09. 23

(21) 申请号 202010105947.4

(22) 申请日 2020.02.20

(65) 同一申请的已公布的文献号

申请公布号 CN 111161748 A

(43) 申请公布日 2020.05.15

(73) 专利权人 百度在线网络技术(北京)有限公司

地址 100085 北京市海淀区上地十街10号
百度大厦三层

(72) 发明人 吴俊楠 徐杨飞 宁峻 贡昱洲
周楠

(74) 专利代理机构 北京银龙知识产权代理有限公司 11243

专利代理师 许静 黄灿

(51) Int.Cl.

G10L 19/02 (2013.01)

G10L 25/51 (2013.01)

审查员 赵忆睿

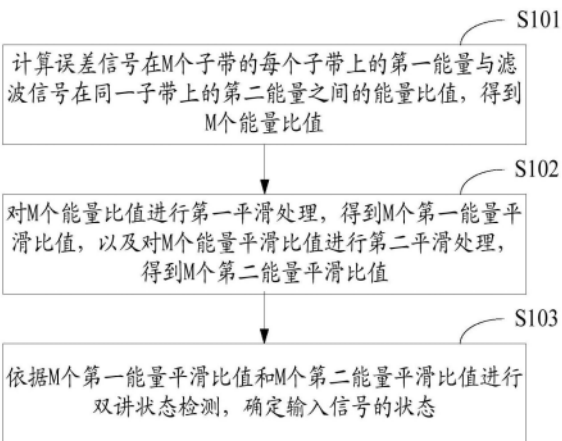
权利要求书3页 说明书15页 附图4页

(54) 发明名称

一种双讲状态检测方法、装置以及电子设备

(57) 摘要

本申请公开了一种双讲状态检测方法、装置以及电子设备,涉及语音处理技术领域。具体实现方案为:计算误差信号在M个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值,得到M个能量比值,其中,误差信号为麦克风采集的输入信号与滤波信号之间的差,滤波信号为对参考信号进行滤波处理后的信号,M为正整数;对M个能量比值进行第一平滑处理,得到M个第一能量平滑比值,以及对M个第一能量平滑比值进行第二平滑处理,得到M个第二能量平滑比值;依据M个第一能量平滑比值和M个第二能量平滑比值进行双讲状态检测,确定输入信号的状态。可提高状态检测准确性。



1. 一种双讲状态检测方法,其特征在于,所述方法包括:

计算误差信号在M个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值,得到M个能量比值,其中,所述误差信号为麦克风采集的输入信号与滤波信号之间的差,所述滤波信号为对参考信号进行滤波处理后的信号,M为正整数;

对所述M个能量比值进行第一平滑处理,得到M个第一能量平滑比值,以及对所述M个第一能量平滑比值进行第二平滑处理,得到M个第二能量平滑比值;

依据所述M个第一能量平滑比值和所述M个第二能量平滑比值进行双讲状态检测,确定所述输入信号的状态;

所述依据所述M个第一能量平滑比值和所述M个第二能量平滑比值进行双讲状态检测,确定所述输入信号的状态,包括:

根据所述M个第二能量平滑比值,预测所述M个第二能量平滑比值的噪声;

根据所述M个第一能量平滑比值以及所述M个第二能量平滑比值的噪声,得到所述M个能量比值的信噪比;

根据所述M个能量比值的信噪比以及预设门限值进行双讲状态检测,确定所述输入信号的状态。

2. 根据权利要求1所述的方法,其特征在于,所述滤波信号包括N帧滤波子信号,所述误差信号包括N帧误差子信号,N为正整数;

其中,目标子带上的所述能量比值包括所述目标子带上的N个子能量比,所述目标子带上的N个子能量比包括每一帧误差子信号在所述目标子带上的第一能量和对应帧滤波子信号在所述目标子带上的第二能量的能量比,所述目标子带上的所述第一能量平滑比值包括所述N个子能量比进行所述第一平滑处理后得到的N个第一子能量平滑比值,所述目标子带上的所述第二能量平滑比值包括所述N个第一子能量平滑比值进行所述第二平滑处理后得到的N个第二子能量平滑比值,所述目标子带为所述M个子带中的任一子带。

3. 根据权利要求2所述的方法,其特征在于,通过以下公式对所述M个能量比值进行第一平滑处理:

$$\text{ratio_short_sm}(n+1,k) = \alpha * \text{ratio_short_sm}(n,k) + (1-\alpha) * \text{ratio}(n+1,k);$$

其中, α 为第一平滑因子,所述 $\text{ratio}(n+1,k)$ 为第 $n+1$ 帧误差子信号在第 k 个子带上的第一能量与第 $n+1$ 帧滤波子信号在第 k 个子带上的第二能量的子能量比,所述 $\text{ratio_short_sm}(n+1,k)$ 为对所述 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对所述 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第 n 帧误差子信号在第 k 个子带上的第一能量与第 n 帧滤波子信号在第 k 个子带上的第二能量的子能量比, n 为正整数,且 $n \leq N-1$, k 为正整数,且 $k \leq M$, $\text{ratio_short_sm}(1,k)$ 等于 $\text{ratio}(1,k)$, $\text{ratio}(1,k)$ 为第1帧误差子信号在第 k 个子带上的第一能量与第1帧滤波子信号在第 k 个子带上的第二能量的子能量比。

4. 根据权利要求2所述的方法,其特征在于,通过以下公式对所述M个第一能量平滑比值进行第二平滑处理:

$$\text{ratio_long_sm}(n+1,k) = \beta * \text{ratio_long_sm}(n,k) + (1-\beta) * \text{ratio_short_sm}(n+1,k);$$

其中, β 为第二平滑因子,所述 $\text{ratio_long_sm}(n+1,k)$ 为对所述 $\text{ratio_short_sm}(n+1,k)$ 进行第二平滑处理后得到的第二子能量平滑比值,所述 $\text{ratio_short_sm}(n+1,k)$ 为对所

述 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值,所述 $\text{ratio}(n+1,k)$ 为第 $n+1$ 帧误差子信号在第 k 个子带上的第一能量与第 $n+1$ 帧滤波子信号在第 k 个子带上的第二能量的子能量比,所述 $\text{ratio_long_sm}(n,k)$ 为对所述 $\text{ratio_short_sm}(n,k)$ 进行第二平滑处理后得到的第二子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对所述 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第 n 帧误差子信号在第 k 个子带上的第一能量与第 n 帧滤波子信号在第 k 个子带上的第二能量的子能量比, n 为正整数,且 $n \leq N-1$, k 为正整数,且 $k \leq M$, $\text{ratio_long_sm}(1,k)$ 等于 $\text{ratio_short_sm}(1,k)$ 。

5. 根据权利要求1所述的方法,其特征在于,所述根据所述 M 个能量比值的信噪比以及预设门限值进行双讲状态检测,确定所述输入信号的状态,包括:

计算所述 M 个能量比值的信噪比中 I 个能量比值信噪比的总和, I 为正整数,且 $I \leq M$;

在所述总和大于所述预设门限的情况下,确定所述输入信号为双讲状态,或者,在所述总和小于或等于所述预设门限的情况下,确定所述输入信号为单讲状态。

6. 一种双讲状态检测装置,其特征在于,所述装置包括:

第一计算模块,用于计算误差信号在 M 个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值,得到 M 个能量比值,其中,所述误差信号为麦克风采集的输入信号与滤波信号之间的差,所述滤波信号为对参考信号进行滤波处理后的信号, M 为正整数;

平滑模块,用于对所述 M 个能量比值进行第一平滑处理,得到 M 个第一能量平滑比值,以及对所述 M 个第一能量平滑比值进行第二平滑处理,得到 M 个第二能量平滑比值;

状态检测模块,用于依据所述 M 个第一能量平滑比值和所述 M 个第二能量平滑比值进行双讲状态检测,确定所述输入信号的状态;

所述状态检测模块,包括:

噪声预测模块,用于根据所述 M 个第二能量平滑比值,预测所述 M 个第二能量平滑比值的噪声;

信噪比获取模块,用于根据所述 M 个第一能量平滑比值以及所述 M 个第二能量平滑比值的噪声,得到所述 M 个能量比值的信噪比;

状态确定模块,用于根据所述 M 个能量比值的信噪比以及预设门限值进行双讲状态检测,确定所述输入信号的状态。

7. 根据权利要求6所述的装置,其特征在于,所述滤波信号包括 N 帧滤波子信号,所述误差信号包括 N 帧误差子信号, N 为正整数;

其中,目标子带上的所述能量比值包括所述目标子带上的 N 个子能量比,所述目标子带上的 N 个子能量比包括每一帧误差子信号在所述目标子带上的第一能量和对应帧滤波子信号在所述目标子带上的第二能量的能量比,所述目标子带上的所述第一能量平滑比值包括所述 N 个子能量比进行所述第一平滑处理后得到的 N 个第一子能量平滑比值,所述目标子带上的所述第二能量平滑比值包括所述 N 个第一子能量平滑比值进行所述第二平滑处理后得到的 N 个第二子能量平滑比值,所述目标子带为所述 M 个子带中的任一子带。

8. 根据权利要求7所述的装置,其特征在于,所述平滑模块通过以下公式对所述 M 个能量比值进行第一平滑处理:

$$\text{ratio_short_sm}(n+1,k) = \alpha * \text{ratio_short_sm}(n,k) + (1-\alpha) * \text{ratio}(n+1,k);$$

其中, α 为第一平滑因子,所述 $\text{ratio}(n+1,k)$ 为第 $n+1$ 帧误差子信号在第 k 个子带上的第一能量与第 $n+1$ 帧滤波子信号在第 k 个子带上的第二能量的子能量比,所述 $\text{ratio_short_sm}(n+1,k)$ 为对所述 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对所述 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第 n 帧误差子信号在第 k 个子带上的第一能量与第 n 帧滤波子信号在第 k 个子带上的第二能量的子能量比, n 为正整数,且 $n \leq N-1$, k 为正整数,且 $k \leq M$, $\text{ratio_short_sm}(1,k)$ 等于 $\text{ratio}(1,k)$, $\text{ratio}(1,k)$ 为第1帧误差子信号在第 k 个子带上的第一能量与第1帧滤波子信号在第 k 个子带上的第二能量的子能量比。

9. 根据权利要求7所述的装置,其特征在于,所述平滑模块通过以下公式对所述 M 个第一能量平滑比值进行第二平滑处理:

$$\text{ratio_long_sm}(n+1,k) = \beta * \text{ratio_long_sm}(n,k) + (1-\beta) * \text{ratio_short_sm}(n+1,k);$$

其中, β 为第二平滑因子,所述 $\text{ratio_long_sm}(n+1,k)$ 为对所述 $\text{ratio_short_sm}(n+1,k)$ 进行第二平滑处理后得到的第二子能量平滑比值,所述 $\text{ratio_short_sm}(n+1,k)$ 为对所述 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值,所述 $\text{ratio}(n+1,k)$ 为第 $n+1$ 帧误差子信号在第 k 个子带上的第一能量与第 $n+1$ 帧滤波子信号在第 k 个子带上的第二能量的子能量比,所述 $\text{ratio_long_sm}(n,k)$ 为对所述 $\text{ratio_short_sm}(n,k)$ 进行第二平滑处理后得到的第二子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对所述 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第 n 帧误差子信号在第 k 个子带上的第一能量与第 n 帧滤波子信号在第 k 个子带上的第二能量的子能量比, n 为正整数,且 $n \leq N-1$, k 为正整数,且 $k \leq M$, $\text{ratio_long_sm}(1,k)$ 等于 $\text{ratio_short_sm}(1,k)$ 。

10. 根据权利要求6所述的装置,其特征在于,所述状态确定模块,确定所述输入信号的状态,包括:

求和模块,用于计算所述 M 个能量比值的信噪比中 I 个能量比值信噪比的总和, I 为正整数,且 $I \leq M$;

在所述总和大于所述预设门限的情况下,确定所述输入信号为双讲状态,或者,在所述总和小于或等于所述预设门限的情况下,确定所述输入信号为单讲状态。

11. 一种电子设备,其特征在于,包括:

至少一个处理器;以及

与所述至少一个处理器通信连接的存储器;其中,

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行如权利要求1至5中任一项所述的方法。

12. 一种存储有计算机指令的非瞬时计算机可读存储介质,其特征在于,所述计算机指令用于使所述计算机执行如权利要求1至5中任一项所述的方法。

一种双讲状态检测方法、装置以及电子设备

技术领域

[0001] 本申请涉及计算机技术中的语音技术领域,尤其涉及一种双讲状态检测方法、装置以及电子设备。

背景技术

[0002] 在一些电子设备中,例如可进行语音交互的智能音箱等,回声消除(AEC)起到关键作用,回声消除的效果直接智能音箱的语音交互性能。又例如,利用电子设备进行通话过程中通过回声消除可提升通话质量。在AEC算法中,为了更好的消除远端参考信号,并且保护近端语音信号,通常会有一个双讲检测模块。双讲检测模块用于检测当前信号的状态:单讲状态和双讲状态。单讲状态指仅有远端参考信号,双讲状态是指既有近端说话人的语音信号,又有远端参考信号。

[0003] 现有的双讲状态检测方法采用全带能量比的方法,即将麦克风采集的信号经线性滤波器处理之后的残余信号能量和扬声器的信号能量进行比较,若比值大于设定的检测门限则判断为双讲(double talk)状态,否则为单讲(single talk)状态。

[0004] 然而,通过上述双讲状态检测方法进行状态检测时,当参考信号能量较大时,或近端信号的能量较小时(例如说话人声音很小),容易出现误判,即容易导致状态检测的准确性较差。

发明内容

[0005] 本申请提供一种双讲状态检测方法、装置和电子设备,以解决问题双讲状态检测准确性较差的问题。

[0006] 第一方面,本申请一个实施例提供一种双讲状态检测方法,包括:

[0007] 计算误差信号在M个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值,得到M个能量比值,其中,所述误差信号为麦克风采集的输入信号与滤波信号之间的差,所述滤波信号为对参考信号进行滤波处理后的信号,M为正整数;

[0008] 对所述M个能量比值进行第一平滑处理,得到M个第一能量平滑比值,以及对所述M个第一能量平滑比值进行第二平滑处理,得到M个第二能量平滑比值;

[0009] 依据所述M个第一能量平滑比值和所述M个第二能量平滑比值进行双讲状态检测,确定所述输入信号的状态。

[0010] 在本申请实施例提供的双讲状态检测方法中,可对误差信号在M个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值进行第一平滑处理,使得到的M个第一能量平滑比值更加稳定,从而可减少由于误差信号在个别子带上存在异常导致M能量比值中存在异常能量比值对状态检测的准确性的影响,而且可对第一平滑处理后得到的M个第一能量平滑比值进行第二平滑处理,使得到的M个第二能量平滑比值更加稳定,从而可减少M个第一能量平滑比值中存在个别异常比值对状态检测的准确性的影响,因此依据所述M个第一能量平滑比值和所述M个第二能量平滑比值进行双讲状态检测,可提

高状态检测准确性。

[0011] 可选的,所述依据所述M个第一能量平滑比值和所述M个第二能量平滑比值进行双讲状态检测,确定所述输入信号的状态,包括:

[0012] 根据所述M个第二能量平滑比值,预测所述M个第二能量平滑比值的噪声;

[0013] 根据所述M个第一能量平滑比值以及所述M个第二能量平滑比值的噪声,得到所述M个能量比值的信噪比;

[0014] 根据所述M个能量比值的信噪比以及预设门限值进行双讲状态检测,确定所述输入信号的状态。

[0015] 在本实施例中,利用M个第一能量平滑比值以及预测的M个第二能量平滑比值的噪声进行能量比值的信噪比计算,利用能量比值的信噪比进行双讲状态检测,以提高检测准确性。

[0016] 可选的,所述滤波信号包括N帧滤波子信号,所述误差信号包括N帧误差子信号,N为正整数;

[0017] 其中,目标子带上的所述能量比值包括所述目标子带上的N个子能量比,所述目标子带上的N个子能量比包括每一帧误差子信号在所述目标子带上的第一能量和对应帧滤波子信号在所述目标子带上的第二能量的能量比,所述目标子带上的所述第一能量平滑比值包括所述N个子能量比进行所述第一平滑处理后得到的N个第一子能量平滑比值,所述目标子带上的所述第二能量平滑比值包括所述N个第一子能量平滑比值进行所述第二平滑处理后得到的N个第二子能量平滑比值,所述目标子带为所述M个子带中的任一子带。

[0018] 在本实施例中,对误差信号以及滤波信号进行分帧处理,计算N帧误差子信号中每一帧误差子信号在目标子带上的第一能量和与其对应帧滤波子信号的第二能量的能量比,得到目标子带上的N个子能量比,目标子带上的能量比值包括目标子带上的N个子能量比,由于有M个子带,从而,可得到M个子带中每个子带的N个子能量比,通过M个子带中每个子带的N个子能量比作为双讲状态检测的基础,可提高状态检测的准确性。

[0019] 可选的,通过以下公式对所述M个能量比值进行第一平滑处理:

[0020] $\text{ratio_short_sm}(n+1,k) = \alpha * \text{ratio_short_sm}(n,k) + (1-\alpha) * \text{ratio}(n+1,k)$;

[0021] 其中, α 为第一平滑因子,所述 $\text{ratio}(n+1,k)$ 为第 $n+1$ 帧误差子信号在第 k 个子带上的第一能量与第 $n+1$ 帧滤波子信号在第 k 个子带上的第二能量的子能量比,所述 $\text{ratio_short_sm}(n+1,k)$ 为对所述 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对所述 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第 n 帧误差子信号在第 k 个子带上的第一能量与第 n 帧滤波子信号在第 k 个子带上的第二能量的子能量比, n 为正整数,且 $n \leq N-1$, k 为正整数,且 $k \leq M$, $\text{ratio_short_sm}(1,k)$ 等于 $\text{ratio}(1,k)$, $\text{ratio}(1,k)$ 为第1帧误差子信号在第 k 个子带上的第一能量与第1帧滤波子信号在第 k 个子带上的第二能量的子能量比。

[0022] 通过上述第一平滑处理,使得到的M个第一能量平滑比值相对M个能量比值更加稳定,依此来进行双讲状态检测,提高检测准确性。

[0023] 可选的,通过以下公式对所述M个第一能量平滑比值进行第二平滑处理:

[0024] $\text{ratio_long_sm}(n+1,k) = \beta * \text{ratio_long_sm}(n,k) + (1-\beta) * \text{ratio_short_sm}(n+1,k)$;

[0025] 其中, β 为第二平滑因子,所述 $\text{ratio_long_sm}(n+1,k)$ 为对所述 $\text{ratio_short_sm}(n+1,k)$ 进行第二平滑处理后得到的第二子能量平滑比值,所述 $\text{ratio_short_sm}(n+1,k)$ 为对所述 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值,所述 $\text{ratio}(n+1,k)$ 为第 $n+1$ 帧误差子信号在第 k 个子带上的第一能量与第 $n+1$ 帧滤波子信号在第 k 个子带上的第二能量的子能量比,所述 $\text{ratio_long_sm}(n,k)$ 为对所述 $\text{ratio_short_sm}(n,k)$ 进行第二平滑处理后得到的第二子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对所述 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第 n 帧误差子信号在第 k 个子带上的第一能量与第 n 帧滤波子信号在第 k 个子带上的第二能量的子能量比, n 为正整数,且 $n \leq N-1$, k 为正整数,且 $k \leq M$, $\text{ratio_long_sm}(1,k)$ 等于 $\text{ratio_short_sm}(1,k)$ 。

[0026] 通过上述第二平滑处理,使得到的 M 个第二能量平滑比值相对于 M 个第一能量平滑值更加稳定,依此来进行双讲状态检测,提高检测准确性。

[0027] 可选的,所述根据所述 M 个能量比值的信噪比以及预设门限值进行双讲状态检测,确定所述输入信号的状态,包括:

[0028] 计算所述 M 个能量比值的信噪比中 I 个能量比值信噪比的总和, I 为正整数,且 $I \leq M$;

[0029] 在所述总和大于所述预设门限的情况下,确定所述输入信号为双讲状态,或者,在所述总和小于或等于所述预设门限的情况下,确定所述输入信号为单讲状态。

[0030] 在本实施例中,以 M 个能量比值中对应子带的子频率范围在第一预设范围内的能量比值(即 I 个能量比值)的信噪比之和进行双讲状态检测。如此,不但能够减少计算量而且能提高检测准确性。

[0031] 第二方面,本申请一个实施例提供一种双讲状态检测装置,所述装置包括:

[0032] 第一计算模块,用于计算误差信号在 M 个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值,得到 M 个能量比值,其中,所述误差信号为麦克风采集的输入信号与滤波信号之间的差,所述滤波信号为对参考信号进行滤波处理后的信号, M 为正整数;

[0033] 平滑模块,用于对所述 M 个能量比值进行第一平滑处理,得到 M 个第一能量平滑比值,以及对所述 M 个第一能量平滑比值进行第二平滑处理,得到 M 个第二能量平滑比值;

[0034] 状态检测模块,用于依据所述 M 个第一能量平滑比值和所述 M 个第二能量平滑比值进行双讲状态检测,确定所述输入信号的状态。

[0035] 可选的,所述状态检测模块,包括:

[0036] 噪声预测模块,用于根据所述 M 个第二能量平滑比值,预测所述 M 个第二能量平滑比值的噪声;

[0037] 信噪比获取模块,用于根据所述 M 个第一能量平滑比值以及所述 M 个第二能量平滑比值的噪声,得到所述 M 个能量比值的信噪比;

[0038] 状态确定模块,用于根据所述 M 个能量比值的信噪比以及预设门限值进行双讲状态检测,确定所述输入信号的状态。

[0039] 可选的,所述滤波信号包括 N 帧滤波子信号,所述误差信号包括 N 帧误差子信号, N 为正整数;

[0040] 其中,目标子带上的所述能量比值包括所述目标子带上的 N 个子能量比,所述目标

子带上的N个子能量比包括每一帧误差子信号在所述目标子带上的第一能量和对应帧滤波子信号在所述目标子带上的第二能量的能量比,所述目标子带上的所述第一能量平滑比值包括所述N个子能量比进行所述第一平滑处理后得到的N个第一子能量平滑比值,所述目标子带上的所述第二能量平滑比值包括所述N个第一子能量平滑比值进行所述第二平滑处理后得到的N个第二子能量平滑比值,所述目标子带为所述M个子带中的任一子带。

[0041] 可选的,所述平滑模块通过以下公式对所述M个能量比值进行第一平滑处理:

[0042] $\text{ratio_short_sm}(n+1,k) = \alpha * \text{ratio_short_sm}(n,k) + (1-\alpha) * \text{ratio}(n+1,k)$;

[0043] 其中, α 为第一平滑因子,所述 $\text{ratio}(n+1,k)$ 为第 $n+1$ 帧误差子信号在第 k 个子带上的第一能量与第 $n+1$ 帧滤波子信号在第 k 个子带上的第二能量的子能量比,所述 $\text{ratio_short_sm}(n+1,k)$ 为对所述 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对所述 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第 n 帧误差子信号在第 k 个子带上的第一能量与第 n 帧滤波子信号在第 k 个子带上的第二能量的子能量比, n 为正整数,且 $n \leq N-1$, k 为正整数,且 $k \leq M$, $\text{ratio_short_sm}(1,k)$ 等于 $\text{ratio}(1,k)$, $\text{ratio}(1,k)$ 为第1帧误差子信号在第 k 个子带上的第一能量与第1帧滤波子信号在第 k 个子带上的第二能量的子能量比。

[0044] 可选的,所述平滑模块通过以下公式对所述M个第一能量平滑比值进行第二平滑处理:

[0045] $\text{ratio_long_sm}(n+1,k) = \beta * \text{ratio_long_sm}(n,k) + (1-\beta) * \text{ratio_short_sm}(n+1,k)$;

[0046] 其中, β 为第二平滑因子,所述 $\text{ratio_long_sm}(n+1,k)$ 为对所述 $\text{ratio_short_sm}(n+1,k)$ 进行第二平滑处理后得到的第二子能量平滑比值,所述 $\text{ratio_short_sm}(n+1,k)$ 为对所述 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值,所述 $\text{ratio}(n+1,k)$ 为第 $n+1$ 帧误差子信号在第 k 个子带上的第一能量与第 $n+1$ 帧滤波子信号在第 k 个子带上的第二能量的子能量比,所述 $\text{ratio_long_sm}(n,k)$ 为对所述 $\text{ratio_short_sm}(n,k)$ 进行第二平滑处理后得到的第二子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对所述 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第 n 帧误差子信号在第 k 个子带上的第一能量与第 n 帧滤波子信号在第 k 个子带上的第二能量的子能量比, n 为正整数,且 $n \leq N-1$, k 为正整数,且 $k \leq M$, $\text{ratio_long_sm}(1,k)$ 等于 $\text{ratio_short_sm}(1,k)$ 。

[0047] 可选的,所述状态确定模块,确定所述输入信号的状态,包括:

[0048] 求和模块,用于计算所述M个能量比值的信噪比中 I 个能量比值信噪比的总和, I 为正整数,且 $I \leq M$;

[0049] 在所述总和大于所述预设门限的情况下,确定所述输入信号为双讲状态,或者,在所述总和小于或等于所述预设门限的情况下,确定所述输入信号为单讲状态。

[0050] 第三方面,本申请一个实施例还提供一种电子设备,包括:

[0051] 至少一个处理器;以及

[0052] 与所述至少一个处理器通信连接的存储器;其中,

[0053] 所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行本申请各实施例提供的方法。

[0054] 第四方面,本申请一个实施例还提供一种存储有计算机指令的非瞬时计算机可读

存储介质,所述计算机指令用于使所述计算机执行本申请各实施例提供的方法。

附图说明

- [0055] 附图用于更好地理解本方案,不构成对本申请的限定。其中:
- [0056] 图1是本申请提供的一个实施例的双讲状态检测方法的流程示意图之一;
- [0057] 图2是本申请提供的一个实施例的双讲状态检测方法的流程示意图之二;
- [0058] 图3是本申请提供的一个实施例的AEC算法的电路图;
- [0059] 图4是本申请提供的一个实施例的双讲状态检测装置的结构图;
- [0060] 图5是用来实现本申请实施例的双讲状态检测方法的电子设备的框图。

具体实施方式

[0061] 以下结合附图对本申请的示范性实施例做出说明,其中包括本申请实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本申请的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0062] 如图1所示,根据本申请的实施例,本申请提供一种双讲状态检测方法,可应用于电子设备,例如,电子设备可以包括智能音箱等,上述方法包括:

[0063] 步骤S101:计算误差信号在M个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值,得到M个能量比值。

[0064] 其中,误差信号为麦克风采集的输入信号与滤波信号之间的差,滤波信号为对参考信号进行滤波处理后的信号,M为正整数。作为一个示例,可通过线性滤波器对参考信号进行滤波处理。需要说明的是,上述参考信号可以理解为传输至电子设备的扬声器进行播放的信号。电子设备的麦克风可采集输入信号,若电子设备的扬声器在播放,则麦克风采集的输入信号中可包括对扬声器播放的信号的采集的信号以及近端语音信号(例如,可能有用户在进行说话等)。

[0065] 在频域中,可将一段频率分割成多个子带,每个子带对应一定的频率范围,例如,对于0-100Hz的频率段,可将其分割成5个子带,每个子带对应子频率范围大小为20Hz,5个子带对应的子频率范围分别为0-20Hz、20-40Hz、40-60Hz、60-80Hz、80-100Hz。在本实施例中,可预先设置一个频率段,该频率段可与电子设备可处理的信号的频率范围有关。例如,对于智能音箱,作为一个示例,该频率段可以为3000Hz。且可预先对预先设置的频率段进行频率分割,得到M个子带。

[0066] M个子带中的任意一个子带上的误差信号为该子带上的输入信号与该子带上的滤波信号的差。对于任意一个子带,可计算误差信号在该子带上的第一能量与滤波信号在该子带上的第二能量之间的能量比值,也就是说,每个能量比值为误差信号和滤波信号在同一子带上的能量之比。即误差信号在每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值,可以理解为在每个相同子带上,误差信号的第一能量与滤波信号的第二能量之间的能量之比,得到的M个能量比值包括每个相同子带上,误差信号的第一能量与滤波信号的第二能量之间的能量之比,即每个能量比值有对应的子带,M个能量比值可以理解为在M个子带上的能量比值。例如,M个子带包括第1个子带和第2个子带,即M为2,则

计算误差信号在第1个子带上的第一能量和滤波信号在第1个子带上的第二能量之间的能量比值,以及误差信号在第2个子带上的第一能量和滤波信号在第2个子带上的第二能量之间的能量比值,如此,得到2个能量比值。

[0067] 步骤S102:对M个能量比值进行第一平滑处理,得到M个第一能量平滑比值,以及对M个第一能量平滑比值进行第二平滑处理,得到M个第二能量平滑比值。

[0068] 得到M个能量比之后,可先对进行第一平滑处理,得到M个第一能量平滑比值,然后再对M个第一能量平滑比值进行第二平滑处理,得到M个第二能量平滑比值,通过对M个能量比值进行第一平滑处理,可使得到的M个第一能量平滑比值相对M个能量比值更加稳定,减少由于误差信号在个别子带上存在异常导致M能量比值中存在异常能量比值对状态检测的准确性的影响,再对M个第一能量平滑比值进行第二平滑处理,可使得到M个第二能量平滑比值相对M个第一能量平滑比值更加稳定,减少M个第一能量平滑比值中存在个别异常比值对状态检测的准确性的影响。作为一个示例,第一平滑处理的平滑参数和第二平滑处理的平滑参数不同。

[0069] 步骤S103:依据M个第一能量平滑比值和M个第二能量平滑比值进行双讲状态检测,确定输入信号的状态。

[0070] 在本实施例中,将M个第一能量平滑比值和M个第二能量平滑比值作为双讲状态检测的依据,以提高状态检测的准确性。

[0071] 在本申请实施例提供的双讲状态检测方法中,可对误差信号在M个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值进行第一平滑处理,使得到的M个第一能量平滑比值更加稳定,从而可减少由于误差信号在个别子带上存在异常导致M能量比值中存在异常能量比值对状态检测的准确性的影响,而且可对第一平滑处理后得到的M个第一能量平滑比值进行第二平滑处理,使得到的M个第二能量平滑比值更加稳定,从而可减少M个第一能量平滑比值中存在个别异常比值对状态检测的准确性的影响,因此依据所述M个第一能量平滑比值和所述M个第二能量平滑比值进行双讲状态检测,可提高状态检测准确性。

[0072] 在一个实施例中,对M个能量比值进行第一平滑处理,得到M个第一能量平滑比值,以及对M个第一能量平滑比值进行第二平滑处理,得到M个第二能量平滑比值,包括:在参考信号的信噪比大于预设信噪比的情况下,对M个能量比值进行第一平滑处理,得到M个第一能量平滑比值,以及对M个第一能量平滑比值进行第二平滑处理,得到M个第二能量平滑比值。

[0073] 即在获得参考信号后,可先计算参考信号的信噪比,例如,可对参考信号的噪声进行估计(例如,可采用最小值跟踪法对噪声进行估计),即对参考信号的噪声进行跟踪,得到参考信号的噪声,然后计算参考信号与参考信号的噪声的比例,得到参考信号的信噪比。为了减少参考信号的信噪比较低时,噪声的存在会使能量比值偏大降低状态检测的准确性的情况,在本实施例的平滑处理过程中,可先比较参考信号的信噪比与预设信噪比,在参考信号的信噪比大于预设信噪比的情况下,对M个能量比值进行第一平滑处理,得到M个第一能量平滑比值。在参考信号比较强(即参考信号的信噪比较强)的情况下,为避免后续跟踪到特别小的噪声,还需对M个第一能量平滑比值进行第二平滑处理,得到M个第二能量平滑比值,以提高状态检测准确性。

[0074] 作为一个示例,在确定输入信号的状态后,还可以包括:根据输入信号的状态,对线性滤波器的滤波参数进行调整。

[0075] 线性滤波器即为对参考信号进行滤波得到滤波信号的滤波器,例如,若输入信号的状态为双讲状态,维持线性滤波器的滤波参数不变,若输入信号的状态为单讲状态,对线性滤波器的滤波参数进行更新。其中,对线性滤波器的滤波参数进行更新的方式有多种,在本申请中不做限定。

[0076] 如图2所示,在一个实施例中,依据M个第一能量平滑比值和M个第二能量平滑比值进行双讲状态检测,确定输入信号的状态的步骤103,包括:

[0077] 步骤S1031:根据M个第二能量平滑比值,预测M个第二能量平滑比值的噪声。

[0078] M个第二能量平滑比值为对M个第一能量平滑比值进行第二平滑处理后的结果,如此,根据进行第二平滑处理后得到的M个第二能量平滑进行噪声预测,可提高噪声预测的准确性,从而提高状态检测的准确性。

[0079] 需要说明的是,上述预测M个第二能量平滑比值的噪声的方式有多种,作为一个示例,可通过最小值跟踪算法预测M个第二能量平滑比值的噪声,即通过最小值跟踪算法对第二能量平滑比值进行最小值跟踪,预测第二能量平滑比值的最小值,将其作为第二能量平滑比值的噪声。

[0080] 步骤S1032:根据M个第一能量平滑比值以及M个第二能量平滑比值的噪声,得到M个能量比值的信噪比。

[0081] 第一能量平滑比值是对能量比值进行第一平滑处理后的结果,第一能量平滑比值与能量比值是一一对应的,第二能量平滑比值是对第一能量平滑比值进行第二平滑处理后的结果,即第二能量平滑比值与第一能量平滑比值是一一对应的,可以将每一个第一能量平滑比值与和其对应的第二能量平滑比值之间的比值作为对应的能量比值的信噪比,如此,可得到M各能量比值的信噪比。

[0082] 步骤S1033:根据M个能量比值的信噪比以及预设门限值进行双讲状态检测,确定输入信号的状态。

[0083] 得到M个能量比值的信噪比之后,即可根据M个能量比值的信噪比以及预设门限值进行双讲状态检测,从而确定输入信号的状态。

[0084] 在本实施例中,利用M个第一能量平滑比值以及预测的M个第二能量平滑比值的噪声进行能量比值的信噪比计算,利用能量比值的信噪比进行双讲状态检测,以提高检测准确性。

[0085] 在一个实施例中,计算误差信号在M个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值之前,还包括:获取参考信号以及麦克风采集的输入信号;对参考信号进行滤波处理后,得到滤波信号;获取输入信号与滤波信号之间的误差信号;计算误差信号在M个子带上的M个第一能量以及滤波信号在M个子带上的M个第二能量。

[0086] 即在计算能量比之前,需要进行能量的获取。由于有M个子带,则需要计算误差信号在M个子带中每个子带上的第一能量,以及计算滤波信号在M个子带中每个子带上的第二能量,如此,得到M个第一能量和M个第二能量,为后续进行能量比值的计算提供依据。

[0087] 在一个实施例中,滤波信号包括N帧滤波子信号,误差信号包括N帧误差子信号,N

为正整数；

[0088] 其中，目标子带上的能量比值包括目标子带上的N个子能量比，目标子带上的N个子能量比包括每一帧误差子信号在目标子带上的第一能量和对应帧滤波子信号在目标子带上的第二能量的能量比，目标子带上的第一能量平滑比值包括N个子能量比进行第一平滑处理后得到的N个第一子能量平滑比值，目标子带上的第二能量平滑比值包括N个第一子能量平滑比值进行第二平滑处理后得到的N个第二子能量平滑比值，目标子带为M个子带中的任一子带。

[0089] 可以理解，输入信号包括N帧输入子信号，参考信号包括N帧参考子信号，误差信号根据输入信号和参考信号确定，从而，误差子信号根据输入子信号和参考子信号确定，输入子信号与误差子信号一一对应，参考子信号与误差子信号一一对应。滤波信号由参考信号确定，从而，滤波子信号由参考子信号确定，滤波子信号和参考子信号一一对应。如此，误差子信号与滤波子信号一一对应，目标子带上的每一帧误差子信号，有其对应的第一能量，目标子带上的每一帧滤波子信号，有其对应的第二能量。通过计算在目标子带上，N帧误差子信号中每一帧误差子信号的第一能量和与其对应帧滤波子信号的第二能量的能量比，从而可得到N个子能量比。

[0090] 在本实施例中，对误差信号以及滤波信号进行分帧处理，计算N帧误差子信号中每一帧误差子信号在目标子带上的第一能量和与其对应帧滤波子信号的第二能量的能量比，得到目标子带上的N个子能量比，目标子带上的能量比值包括目标子带上的N个子能量比，由于有M个子带，从而，可得到M个子带中每个子带的N个子能量比，通过M个子带中每个子带的N个子能量比作为双讲状态检测的基础，可提高状态检测的准确性。

[0091] 在一个实施例中，通过以下公式对所述M个能量比值进行第一平滑处理：

[0092] $\text{ratio_short_sm}(n+1,k) = \alpha * \text{ratio_short_sm}(n,k) + (1-\alpha) * \text{ratio}(n+1,k)$ ；

[0093] 其中， α 为第一平滑因子，所述 $\text{ratio}(n+1,k)$ 为第 $n+1$ 帧误差子信号在第 k 个子带上的第一能量与第 $n+1$ 帧滤波子信号在第 k 个子带上的第二能量的子能量比，所述 $\text{ratio_short_sm}(n+1,k)$ 为对所述 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值， $\text{ratio_short_sm}(n,k)$ 为对所述 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值， $\text{ratio}(n,k)$ 为第 n 帧误差子信号在第 k 个子带上的第一能量与第 n 帧滤波子信号在第 k 个子带上的第二能量的子能量比， n 为正整数，且 $n \leq N-1$ ， k 为正整数，且 $k \leq M$ ， $\text{ratio_short_sm}(1,k)$ 等于 $\text{ratio}(1,k)$ ， $\text{ratio}(1,k)$ 为第1帧误差子信号在第 k 个子带上的第一能量与第1帧滤波子信号在第 k 个子带上的第二能量的子能量比。

[0094] 上述误差子信号的帧顺序为在误差信号中误差子信号的时间先后顺序，上述滤波子信号的帧顺序为在滤波信号中滤波子信号的时间先后顺序， $\text{ratio_short_sm}(1,k)$ 等于 $\text{ratio}(1,k)$ ，可以理解为，时间最先得到的第1帧误差子信号在第 k 个子带上的第一能量与第1帧滤波子信号在第 k 个子带上的第二能量的子能量比，进行第一平滑后的结果与平滑前的结果相同，也就是说对最先采集的第1帧误差子信号在第 k 个子带上的第一能量与第1帧滤波子信号在第 k 个子带上的第二能量的子能量比不进行平滑。通过上述第一平滑处理，使得到的M个第一能量平滑比值相对M个能量比值更加稳定，依此来进行双讲状态检测，提高检测准确性。

[0095] 作为一个示例， $0 \leq \alpha \leq 1$ 。其中，在上述第一平滑处理的公式中，第一平滑因子 α 的

取值可根据公式右边的 $\text{ratio_short_sm}(n,k)$ 和 $\text{ratio}(n+1,k)$ 确定,随着 n 或/和 k 的变化, $\text{ratio_short_sm}(n,k)$ 和 $\text{ratio}(n+1,k)$ 随之变化,则 α 可能发生变化。例如, $\text{ratio}(n,k)$ 进行第一平滑处理完毕,得到了 $\text{ratio_short_sm}(n,k)$,当前需对 $\text{ratio}(n+1,k)$ 进行第一平滑处理,在 $\text{ratio_short_sm}(n,k)$ 小于 $\text{ratio}(n+1,k)$ 的情况下,说明当前存在近端语音的可能性比较大,需要尽快的跟踪到当前 $\text{ratio}(n+1,k)$ 上,因此第一平滑因子选择一个比较小的值,反之则选择比较大的值。例如,预先设置有第一预设值和第二预设值,第一预设值小于第二预设值,作为一个示例,第一预设值可以取0.5,第二预设值可以取0.7,在 $\text{ratio_short_sm}(n,k)$ 小于 $\text{ratio}(n+1,k)$ 的情况下,可选取第一预设值作为第一平滑因子的值,否则选取第二预设值作为第一平滑因子的值。

[0096] 在一个实施例中,通过以下公式对所述 M 个第一能量平滑比值进行第二平滑处理:

[0097] $\text{ratio_long_sm}(n+1,k) = \beta * \text{ratio_long_sm}(n,k) + (1-\beta) * \text{ratio_short_sm}(n+1,k)$;

[0098] 其中, β 为第二平滑因子,所述 $\text{ratio_long_sm}(n+1,k)$ 为对所述 $\text{ratio_short_sm}(n+1,k)$ 进行第二平滑处理后得到的第二子能量平滑比值,所述 $\text{ratio_short_sm}(n+1,k)$ 为对所述 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值,所述 $\text{ratio}(n+1,k)$ 为第 $n+1$ 帧误差子信号在第 k 个子带上的第一能量与第 $n+1$ 帧滤波子信号在第 k 个子带上的第二能量的子能量比,所述 $\text{ratio_long_sm}(n,k)$ 为对所述 $\text{ratio_short_sm}(n,k)$ 进行第二平滑处理后得到的第二子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对所述 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第 n 帧误差子信号在第 k 个子带上的第一能量与第 n 帧滤波子信号在第 k 个子带上的第二能量的子能量比, n 为正整数,且 $n \leq N-1$, k 为正整数,且 $k \leq M$, $\text{ratio_long_sm}(1,k)$ 等于 $\text{ratio_short_sm}(1,k)$ 。

[0099] 上述误差子信号的帧顺序为在误差信号中误差子信号的时间先后顺序,上述滤波子信号的帧顺序为在滤波信号中滤波子信号的时间先后顺序, $\text{ratio_long_sm}(1,k)$ 等于 $\text{ratio_short_sm}(1,k)$,可以理解为,时间最先得到的第1帧误差子信号在第 k 个子带上的第一能量与第1帧滤波子信号在第 k 个子带上的第二能量的子能量比进行第一平滑处理后的结果,进行第二平滑后的结果与第一平滑处理后的结果相同,也就是说对最先采集的第1帧误差子信号在第 k 个子带上的第一能量与第1帧滤波子信号在第 k 个子带上的第二能量的子能量比进行第一平滑处理后不进行第二平滑处理。通过上述第二平滑处理,使得到的 M 个第二能量平滑比值相对于 M 个第一能量平滑值更加稳定,依此来进行双讲状态检测,提高检测准确性。作为一个示例, $0 \leq \beta \leq 1$,例如, β 可以取0.99。另外, $\beta > \alpha$ 。

[0100] 在一个实施例中,根据 M 个能量比值的信噪比以及预设门限值进行双讲状态检测,确定输入信号的状态,包括:

[0101] 计算 M 个能量比值的信噪比中 I 个能量比值的信噪比的总和, I 为正整数,且 $I \leq M$;

[0102] 在总和大于预设门限的情况下,确定输入信号为双讲状态,或者,在总和小于或等于预设门限的情况下,确定输入信号为单讲状态。

[0103] 能量比值与子带一一对应,而子带对应一段子频率范围,作为一个示例,上述 I 个能量比值可以为 M 个能量比值中对应子带的子频率范围属于第一频率范围内的能量比值,例如,预先设置一个频率段为3000Hz,分成的 M 个子带,由于近端语音信号的能量主要在低频段(2000Hz以下),则第一频率范围可以为0-2000Hz。或者,也可以根据电子设备的麦克风

接收的输入信号的频率分布选择第一频率范围,例如,选取麦克风采集的输入信号中信噪比较大(例如,大于预设阈值)的频带作为第一频率范围。在本实施例中,以M个能量比值中对应子带的子频率范围在第一预设范围内的能量比值(即I个能量比值)的信噪比之和进行双讲状态检测。如此,不但能够减少计算量而且能提高检测准确性。

[0104] 在一个实施例中,输入信号包括N帧输入子信号,输入子信号与误差信号一一对应,对于分帧处理的情况,上述总和可以包括每一帧误差子信号(或每一帧滤波子信号)下的M个能量比值的信噪比中I个能量比值的信噪比之和,即总和包括N个信噪比之和,如此,在所述总和大于所述预设门限的情况下,确定所述输入信号为双讲状态,包括:确定所述N个信噪比之和中大于预设门限的第一目标信噪比对应的第一目标帧输入子信号(属于输入信号中的输入子信号)为双讲状态,确定所述N个信噪比之和中小于或等于预设门限的第二目标信噪比对应的第二目标帧输入子信号(属于输入信号中的输入子信号)为单讲状态。

[0105] 下面一个具体实施例对上述双讲状态检测方法的过程加以具体说明。

[0106] 如图3所示,为AEC算法的电路图,在AEC算法中,为了更好的消除远端信号,并且保护近端语音信号s,通常会有一个双讲检测模块。双讲检测模块用于检测输入信号y的状态:近端状态、单讲状态和双讲状态。近端状态是指只有近端说话人的近端语音信号,单讲状态指仅有远端信号,双讲状态是指既有近端说话人的近端语音信号,又有远端信号。只有分辨出输入信号的状态之后,才能在后续处理中针对不同的状态进行不同的处理,例如,在近端状态下,不需要进行回声消除,双讲状态下回声消除的等级较低,单讲状态下回声消除等级更高,从而在只有远端信号的时候将其尽可能的消除,在有近端信号的时候减少语音信号的失真。

[0107] 如图3所示,实现AEC算法的电路图中包括线性滤波器、双讲检测模块及非线性处理模块,另外,在麦克风与差分电路之间还连接有第一子带变换模块,对输入信号进行子带分解,而且,在参考信号x与线性滤波器之间还连接有第二子带变换模块,对参考信号进行子带分解。在线性滤波器模块对参考信号进行线性滤波后,麦克风的输入信号y与滤波信号 $\hat{x}$ 通过差分电路进行差分处理得到误差信号e,麦克风接收到输入信号中的远端信号已经被消除了大部分,而近端语音信号被保留,非线性处理模块对误差信号e再进行进一步的回声消除。此时在各子带上分别计算误差信号与滤波信号之间的能量比值ratio(n,k)。其中,

中, $\text{ratio}(n,k) = \frac{|E(n,k)|^2}{|X(n,k)|^2}$, E(n,k) 为误差信号中第n帧误差子信号在第k个子带上的频域信号, $|E(n,k)|^2$ 为误差信号中第n帧误差子信号在第k个子带上的第一能量, X(n,k) 为滤波信号中第n帧滤波子信号在第k个子带上的频域信号, $|X(n,k)|^2$ 为滤波信号中第n帧滤波子信号在第k个子带上的第二能量。

号, $|E(n,k)|^2$ 为误差信号中第n帧误差子信号在第k个子带上的第一能量, X(n,k) 为滤波信号中第n帧滤波子信号在第k个子带上的频域信号, $|X(n,k)|^2$ 为滤波信号中第n帧滤波子信号在第k个子带上的第二能量。

[0108] 当近端存在语音信号时ratio(n,k)较大,在仅有远端信号时ratio(n,k)较小,若直接利用该ratio(n,k)进行状态判决,则判决门限在参考信号能量不同的条件下难以统一,并且个别子带上的异常点容易影响最终的判决结果。

[0109] 在参考信号的信噪比较大时对ratio(n,k)进行短时平滑(即第一平滑处理)和长时平滑(第二平滑处理),要求参考信号的信噪比较大是为了防止参考信号的信噪比较低时,噪音信号的存在会使得ratio(n,k)偏大,从而影响判决结果。

[0110] 本实施例对上述双讲状态检测方法的过程如下:

[0111] 计算误差信号在M个子带上的M个第一能量以及滤波信号在M个子带上的M个第二能量。

[0112] 计算子能量比 $\text{ratio}(n,k)$,其中, n 取 $1,2,\dots,N-1$, k 取 $1,2,\dots,M$;

[0113] 对 $\text{ratio}(n,k)$ 进行短时平滑,得到 $\text{ratio_short_sm}(n,k)$ 。

[0114] 对 $\text{ratio_short_sm}(n,k)$ 进行长时平滑,得到 $\text{ratio_long_sm}(n,k)$ 。

[0115] 通过最小值跟踪算法跟踪 $\text{ratio_short_sm}(n,k)$ 的噪声 $\text{ratio_long_min}(n,k)$ 。

[0116] 计算 $\text{ratio}(n,k)$ 的信噪比 $\text{ratio_snr}(n,k)$,其中,

$$\text{ratio_snr}(n,k) = \frac{\text{ratio_short_sm}(n,k)}{\text{ratio_long_min}(n,k)}。$$

[0117] 子带的子频率范围在第一预设频率范围的 $\text{ratio_snr}(n,k)$ 进行求和,即

$$\text{RATIO_SNR} = \sum_{k=k_1}^{k_2} \text{ratio_snr}(n,k) ;$$

[0118] 其中 k_1 、 k_2 分别对应的子频率范围在第一预设频率范围内, k_1 为正整数,且 $k_1 \leq M$, k_2 为正整数,且 $k_2 \leq M$, $k_1 \leq k_2$, RATIO_SNR 为第 n 帧误差信号下 k_2-k_1+1 个能量比值的信噪比之和,也可以为第 n 帧输入子信号下 k_2-k_1+1 个能量比值的信噪比之和,即 I 为 k_2-k_1+1 。

[0119] 将 RATIO_SNR 与判决门限(即预设门限) T 进行比较,若 $\text{RATIO_SNR} > T$ 则判断为双讲状态,否则为单讲状态。

[0120] 由于最终与判决门限比较的是误差信号与滤波信号(对应参考信号)的能量比值的信噪比,这样即可以实现双讲状态检测的自适应调整,无论误差信号和滤波信号的能量大小,其信噪比均能保持一个水平,即当近端语音存在时信噪比高,否则信噪比低。同时选择低频段(或根据实际设备选择其他频段)也能避免个别频点的异常点的影响,从而使双讲检测准确度更高。

[0121] 请参考图4,本申请提供一种实施例的双讲状态检测装置400,装置包括:

[0122] 第一计算模块401,用于计算误差信号在M个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值,得到M个能量比值,其中,误差信号为麦克风采集的输入信号与滤波信号之间的差,滤波信号为对参考信号进行滤波处理后的信号, M 为正整数;

[0123] 平滑模块402,用于对M个能量比值进行第一平滑处理,得到M个第一能量平滑比值,以及对M个第一能量平滑比值进行第二平滑处理,得到M个第二能量平滑比值;

[0124] 状态检测模块403,用于依据M个第一能量平滑比值和M个第二能量平滑比值进行双讲状态检测,确定输入信号的状态。

[0125] 在一个实施例中,状态检测模块,包括:

[0126] 噪声预测模块,用于根据M个第二能量平滑比值,预测M个第二能量平滑比值的噪声;

[0127] 信噪比获取模块,用于根据M个第一能量平滑比值以及M个第二能量平滑比值的噪声,得到M个能量比值的信噪比;

[0128] 状态确定模块,用于根据M个能量比值的信噪比以及预设门限值进行双讲状态检

测,确定输入信号的状态。

[0129] 在一个实施例中,滤波信号包括N帧滤波信号,误差信号包括N帧误差信号,N为正整数;

[0130] 其中,目标子带上的能量比值包括目标子带上的N个子能量比,目标子带上的N个子能量比包括每一帧误差信号在目标子带上的第一能量和对应帧滤波信号在目标子带上的第二能量的能量比,目标子带上的第一能量平滑比值包括N个子能量比进行第一平滑处理后得到的N个第一子能量平滑比值,目标子带上的第二能量平滑比值包括N个第一子能量平滑比值进行第二平滑处理后得到的N个第二子能量平滑比值,目标子带为M个子带中的任一子带。

[0131] 在一个实施例中,平滑模块通过以下公式对M个能量比值进行第一平滑处理:

[0132] $\text{ratio_short_sm}(n+1,k) = \alpha * \text{ratio_short_sm}(n,k) + (1-\alpha) * \text{ratio}(n+1,k)$;

[0133] 其中, α 为第一平滑因子, $\text{ratio}(n+1,k)$ 为第n+1帧误差信号在第k个子带上的第一能量与第n+1帧滤波信号在第k个子带上的第二能量的子能量比, $\text{ratio_short_sm}(n+1,k)$ 为对 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第n帧误差信号在第k个子带上的第一能量与第n帧滤波信号在第k个子带上的第二能量的子能量比,n为正整数,且 $n \leq N-1$,k为正整数,且 $k \leq M$, $\text{ratio_short_sm}(1,k)$ 等于 $\text{ratio}(1,k)$, $\text{ratio}(1,k)$ 为第1帧误差信号在第k个子带上的第一能量与第1帧滤波信号在第k个子带上的第二能量的子能量比。

[0134] 在一个实施例中,平滑模块通过以下公式对M个第一能量平滑比值进行第二平滑处理:

[0135] $\text{ratio_long_sm}(n+1,k) = \beta * \text{ratio_long_sm}(n,k) + (1-\beta) * \text{ratio_short_sm}(n+1,k)$;

[0136] 其中, β 为第二平滑因子, $\text{ratio_long_sm}(n+1,k)$ 为对 $\text{ratio_short_sm}(n+1,k)$ 进行第二平滑处理后得到的第二子能量平滑比值, $\text{ratio_short_sm}(n+1,k)$ 为对 $\text{ratio}(n+1,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n+1,k)$ 为第n+1帧误差信号在第k个子带上的第一能量与第n+1帧滤波信号在第k个子带上的第二能量的子能量比, $\text{ratio_long_sm}(n,k)$ 为对 $\text{ratio_short_sm}(n,k)$ 进行第二平滑处理后得到的第二子能量平滑比值, $\text{ratio_short_sm}(n,k)$ 为对 $\text{ratio}(n,k)$ 进行第一平滑处理后得到的第一子能量平滑比值, $\text{ratio}(n,k)$ 为第n帧误差信号在第k个子带上的第一能量与第n帧滤波信号在第k个子带上的第二能量的子能量比,n为正整数,且 $n \leq N-1$,k为正整数,且 $k \leq M$, $\text{ratio_long_sm}(1,k)$ 等于 $\text{ratio_short_sm}(1,k)$ 。

[0137] 在一个实施例中,状态确定模块,确定输入信号的状态,包括:

[0138] 求和模块,用于计算M个能量比值的信噪比中I个能量比值信噪比的总和,I为正整数,且 $I \leq M$;

[0139] 在总和大于预设门限的情况下,确定输入信号为双讲状态,或者,在总和小于或等于预设门限的情况下,确定输入信号为单讲状态。

[0140] 上述各实施例的双讲状态检测装置为实现上述各实施例的双讲状态检测方法的装置,技术特征对应,技术效果对应,在此不再赘述。

[0141] 根据本申请的实施例,本申请还提供了一种电子设备和一种可读存储介质。

[0142] 如图5所示,是根据本申请实施例的双讲状态检测方法的电子设备的框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅作为示例,并且不意在限制本文中描述的和/或者要求的本申请的实现。

[0143] 如图5所示,该电子设备包括:一个或多个处理器501、存储器502,以及用于连接各部件的接口,包括高速接口和低速接口。各个部件利用不同的总线互相连接,并且可以被安装在公共主板上或者根据需要以其它方式安装。处理器可以对在电子设备内执行的指令进行处理,包括存储在存储器中或者存储器上以在外部输入/输出装置(诸如,耦合至接口的显示设备)上显示GUI的图形信息的指令。在其它实施方式中,若需要,可以将多个处理器和/或多条总线与多个存储器和多个存储器一起使用。同样,可以连接多个电子设备,各个设备提供部分必要的操作(例如,作为服务器阵列、一组刀片式服务器、或者多处理器系统)。图5中以一个处理器501为例。

[0144] 存储器502即为本申请所提供的非瞬时计算机可读存储介质。其中,所述存储器存储有可由至少一个处理器执行的指令,以使所述至少一个处理器执行本申请所提供的双讲状态检测方法。本申请的非瞬时计算机可读存储介质存储计算机指令,该计算机指令用于使计算机执行本申请所提供的双讲状态检测方法。

[0145] 存储器502作为一种非瞬时计算机可读存储介质,可用于存储非瞬时软件程序、非瞬时计算机可执行程序以及模块,如本申请实施例中的双讲状态检测方法对应的程序指令/模块(例如,附图4所示的第一计算模块401、平滑模块402、状态检测模块403)。处理器501通过运行存储在存储器502中的非瞬时软件程序、指令以及模块,从而执行服务器的各种功能应用以及数据处理,即实现上述方法实施例中的双讲状态检测方法。

[0146] 存储器502可以包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需要的应用程序;存储数据区可存储根据键盘显示的电子设备的使用所创建的数据等。此外,存储器502可以包括高速随机存取存储器,还可以包括非瞬时存储器,例如至少一个磁盘存储器件、闪存器件、或其他非瞬时固态存储器件。在一些实施例中,存储器502可选包括相对于处理器501远程设置的存储器,这些远程存储器可以通过网络连接至键盘显示的电子设备。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0147] 双讲状态检测方法的电子设备还可以包括:输入装置503和输出装置504。处理器501、存储器502、输入装置503和输出装置504可以通过总线或者其他方式连接,图5中以通过总线连接为例。

[0148] 输入装置503可接收输入的数字或字符信息,以及产生与键盘显示的电子设备的用户设置以及功能控制有关的键信号输入,例如触摸屏、小键盘、鼠标、轨迹板、触摸板、指示杆、一个或者多个鼠标按钮、轨迹球、操纵杆等输入装置。输出装置504可以包括显示设备、辅助照明装置(例如,LED)和触觉反馈装置(例如,振动电机)等。该显示设备可以包括但不限于,液晶显示器(LCD)、发光二极管(LED)显示器和等离子体显示器。在一些实施方式

中,显示设备可以是触摸屏。

[0149] 此处描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、专用ASIC(专用集成电路)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0150] 这些计算程序(也称作程序、软件、软件应用、或者代码)包括可编程处理器的机器指令,并且可以利用过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。如本文使用的,术语“机器可读介质”和“计算机可读介质”指的是用于将机器指令和/或数据提供给可编程处理器的任何计算机程序产品、设备、和/或装置(例如,磁盘、光盘、存储器、可编程逻辑装置(PLD)),包括,接收作为机器可读信号的机器指令的机器可读介质。术语“机器可读信号”指的是用于将机器指令和/或数据提供给可编程处理器的任何信号。

[0151] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0152] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0153] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务端关系的计算机程序来产生客户端和服务端的关系。

[0154] 根据本申请实施例的技术方案,在本申请实施例提供的双讲状态检测方法中,可对误差信号在M个子带的每个子带上的第一能量与滤波信号在同一子带上的第二能量之间的能量比值进行第一平滑处理,使得到的M个第一能量平滑比值更加稳定,从而可减少由于误差信号在个别子带上存在异常导致M能量比值中存在异常能量比值对状态检测的准确性的影响,而且可对第一平滑处理后得到的M个第一能量平滑比值进行第二平滑处理,使得到的M个第二能量平滑比值更加稳定,从而可减少M个第一能量平滑比值中存在个别异常比值对状态检测的准确性的影响,因此依据所述M个第一能量平滑比值和所述M个第二能量平滑比值进行双讲状态检测,可提高状态检测准确性。

[0155] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例

如,本发申请中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本申请公开的技术方案所期望的结果,本文在此不进行限制。

[0156] 上述具体实施方式,并不构成对本申请保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本申请的精神和原则之内所作的修改、等同替换和改进等,均应包含在本申请保护范围之内。

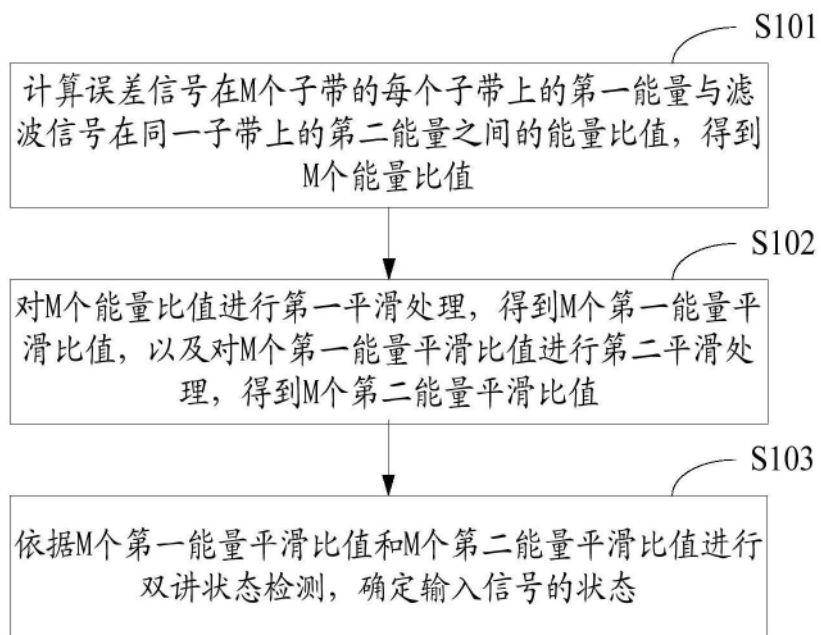


图1

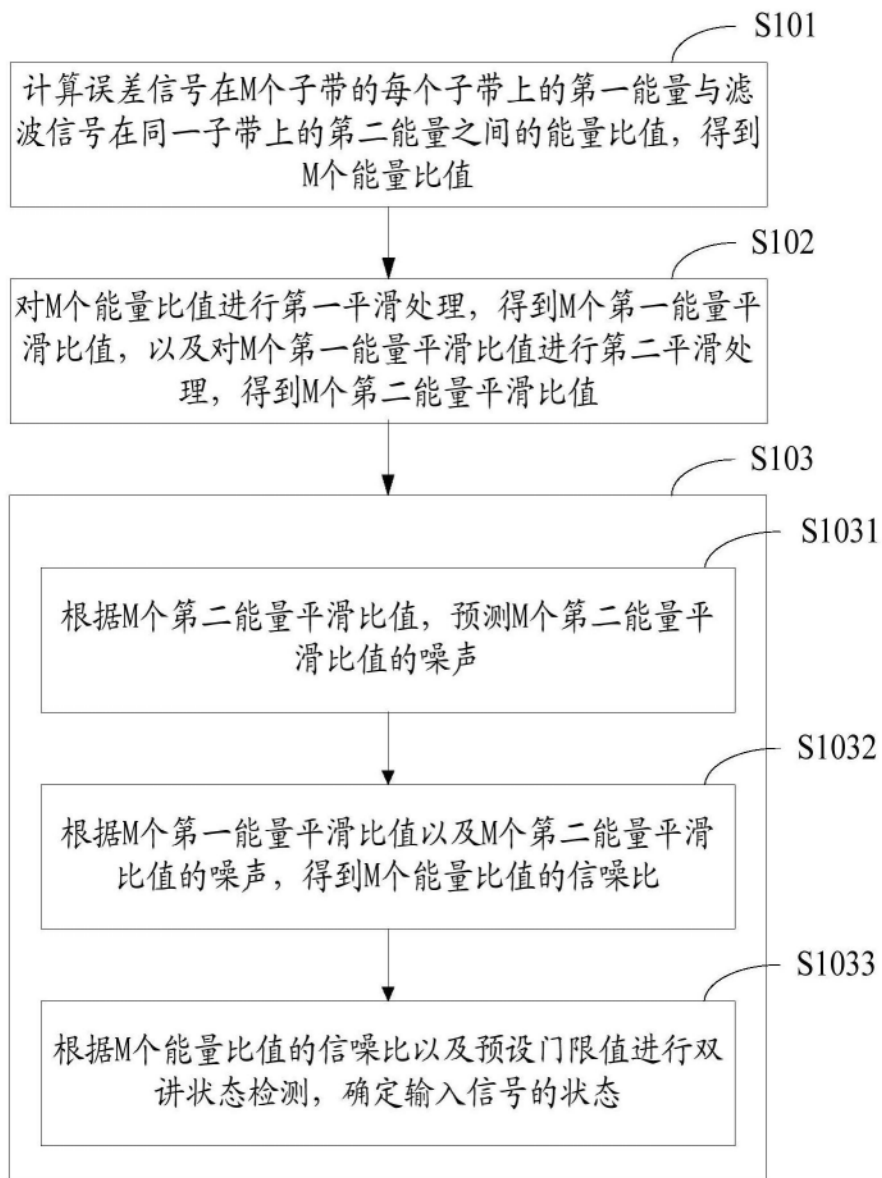


图2

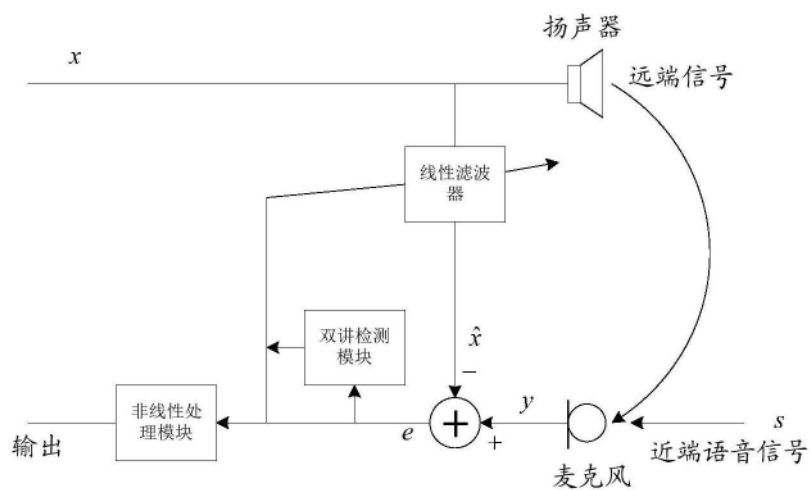


图3

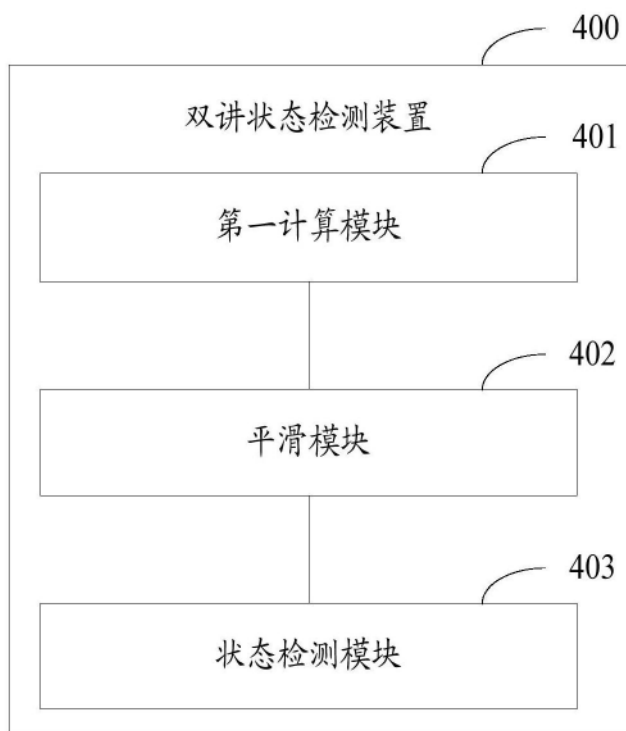


图4

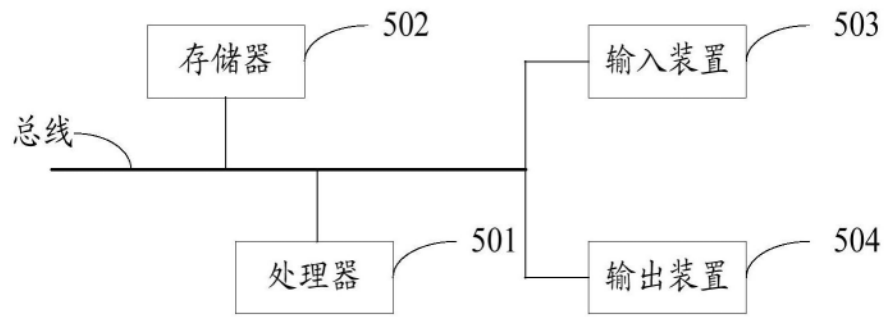


图5