



(12) 发明专利申请

(10) 申请公布号 CN 112802487 A

(43) 申请公布日 2021.05.14

(21) 申请号 202011615540.2

(22) 申请日 2020.12.31

(71) 申请人 北京梧桐车联科技有限责任公司
地址 100089 北京市海淀区海淀大街8号A座5层B区

(72) 发明人 王倩 隋园 来杏杏

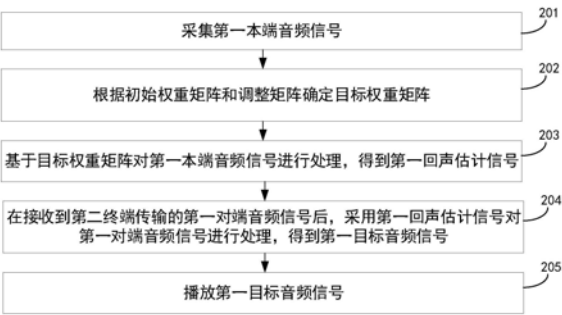
(74) 专利代理机构 北京三高永信知识产权代理有限公司 11138
代理人 谢冬寒

(51) Int.Cl.
G10L 21/0208 (2013.01)
G10L 21/0216 (2013.01)
H04M 9/08 (2006.01)

权利要求书3页 说明书16页 附图5页

(54) 发明名称
回声处理方法、装置及系统

(57) 摘要
本申请公开了一种回声处理方法、装置及系统,属于音频处理技术领域。该第一终端可以基于初始权重矩阵和调整矩阵确定目标权重矩阵,并基于通过该目标权重矩阵和第一本端音频信号得到的第一回声信号,对第一对端音频信号进行处理,得到第一目标音频信号,并播放该第一目标音频信号。由于调整数值与步长因子正相关,且该步长因子是根据第二本端音频信号确定,因此提高了确定的调整数值的可靠性,进而提高了确定的第一回声估计信号的准确度,从而有效提高了回声处理的效果。



1. 一种回声处理方法,其特征在于,应用于第一终端,所述方法包括:

采集第一本端音频信号;

根据初始权重矩阵和调整矩阵确定目标权重矩阵,其中,所述初始权重矩阵为在所述第一本端音频信号之前采集到的第二本端音频信号对应的权重矩阵,所述调整矩阵中的调整数值与所述第二本端音频信号的步长因子矩阵中的步长因子正相关,所述目标权重矩阵中的目标权重系数与所述初始权重矩阵中的初始权重系数和所述调整数值均正相关,所述步长因子根据所述第二本端音频信号确定;

基于所述目标权重矩阵对所述第一本端音频信号进行处理,得到第一回声估计信号;

在接收到第二终端传输的第一对端音频信号后,采用所述第一回声估计信号对所述第一对端音频信号进行处理,得到第一目标音频信号;

播放所述第一目标音频信号。

2. 根据权利要求1所述的方法,其特征在于,所述第二本端音频信号包括L个音频帧,每个所述音频帧包括s个频点;所述初始权重矩阵包括与所述第二本端音频信号中L×s个频点一一对应的L×s个初始权重系数,所述调整矩阵包括与所述L×s个频点一一对应的L×s个调整数值,L和s均为大于或等于1的整数;在根据初始权重矩阵和调整矩阵确定目标权重矩阵之前,所述方法还包括:

基于所述初始权重矩阵对所述第二本端音频信号进行处理,得到第二回声估计信号;

在接收到所述第二终端传输的第二对端音频信号后,采用所述第二回声估计信号对所述第二对端音频信号进行处理,得到第二目标音频信号;

基于所述第二本端音频信号的频域自相关系数,以及所述第二本端音频信号和所述第二目标音频信号的频域互相关系数,确定每个所述音频帧的音频泄漏系数,所述音频泄漏系数与所述频域自相关系数负相关,且与所述频域互相关系数正相关;

对于所述L×s个频点中的每个频点,基于所述频点所属的音频帧的音频泄漏系数,所述第二目标音频信号在所述频点的第一频谱值,以及所述第二回声估计信号在所述频点的第二频谱值,确定所述频点对应的所述步长因子,所述步长因子与所述频点所属的音频帧的音频泄漏系数和所述第二频谱值均正相关,且与所述第一频谱值负相关;

采用块稀疏成比例仿射投影算法对所述第二本端音频信号在所述频点的第三频谱值以及所述第二目标音频信号进行处理,得到所述频点对应的更新步长;

基于所述频点对应的步长因子和所述频点对应的更新步长,确定所述频点对应的调整数值;

所述根据初始权重矩阵和调整矩阵确定目标权重矩阵,包括:

将所述第二本端音频信号中每个所述频点对应的初始权重系数与所述频点对应的调整数值之和确定为所述第一本端音频信号中所述频点的目标权重系数。

3. 根据权利要求2所述的方法,其特征在于,所述第二本端音频信号中第i个音频帧的

$$\text{音频泄露系数} \eta(i) \text{ 满足: } \eta(i) = \frac{\sum_{k=(i-1) \times s}^{i \times s - 1} R_{EX}(k, i)}{\sum_{k=(i-1) \times s}^{i \times s - 1} R_{XX}(k, i)};$$

其中,所述i为大于0,且小于或等于L的整数,所述k为大于或等于0,且小于L×s的整

数;所述 $R_{ex}(k, i)$ 为所述第二本端音频信号中第 k 个频点和所述第二目标音频信号中第 k 个频点的频域互相关系数;

所述第 k 个频点对应的步长因子 $\mu_{out}(k)$ 满足: $\mu_{out}(k) = \min\{\eta(i) \times \frac{|y(k)|^2}{|e(k)|^2}, \mu_{max}\}$;

其中, $\min$ 表示取最小值, 所述 μ_{max} 为大于0且小于1的数, 所述 $y(k)$ 为所述第二回声估计信号中第 k 个频点的第二频谱值, 所述 $e(k)$ 为所述第二目标音频信号中第 k 个频点第一频谱值。

4. 根据权利要求2所述的方法, 其特征在于, 所述第二本端音频信号包括 P 段第二子音频信号, 每段所述第二子音频信号包括 n 个频点, 所述 $L \times s$ 个频点中第 k 个频点对应的更新步长 $H(k)$ 满足:

$$H(k) = \frac{\mu \times G \times X_0^H(k) \times \text{sgn}[e]}{\sqrt{[G \times X_0^H(k) \times \text{sgn}[e]]^T \times (G \times X_0^H(k) \times \text{sgn}[e]) + \beta \times I}};$$

其中, 所述 k 为大于或等于0, 且小于 $L \times s$ 的整数, 所述 μ 为大于0的数, G 为对角矩阵, 所述对角矩阵中对角线上的第 t 个元素为 $g_t(t) \times I_{1 \times n}$, 且 $g_t(t)$ 与 $w(t)$ 正相关, 所述 $w(t)$ 为所述第二本端音频信号中第 t 个频点对应的初始权重系数, 所述 $I_{1 \times n}$ 为 $1 \times n$ 的单位矩阵, 所述 $X_0^H(k)$ 为对所述第二本端音频信号中第 k 个频点的频谱值的共轭转置, 所述 $\text{sgn}[e]$ 为行向量, 所述行向量中第 t 个元素为 $\text{sgn}(e(t))$, 所述 $e(t)$ 为所述第二目标音频信号中第 t 个频点的第一频谱值, 所述 sgn 为符号函数, 所述 t 的取值范围为 k 至 $1 \times n - 1$, 所述第 k 个频点属于第1段所述第二子音频信号, 所述1为大于或等于0, 且小于 P 的正整数, 所述 β 为不等于0的整数, 所述 $n = \frac{L}{P} \times s$, 所述 P 为大于0且小于 $L \times s$ 的整数, n 为不大于 P 的整数。

5. 根据权利要求4所述的方法, 其特征在于, 所述 $L \times s$ 个频点中第 k 个频点的调整数值 $D(k)$ 满足: $D(k) = F(d(k))$, 所述 $d(k)$ 满足: $d(k) = \begin{bmatrix} I_{1 \times n} & 0_{1 \times n} \\ 0_{1 \times n} & 0_{1 \times n} \end{bmatrix} \times F^{-1}[\mu_{out}(k) \times H(k)]$;

其中, 所述 k 为大于或等于0, 且小于 $L \times s$ 的整数, F 表示傅立叶变换, 所述 F^{-1} 表示逆傅立叶变换。

6. 根据权利要求4所述的方法, 其特征在于, 所述 $L \times s$ 个频点中第 k 个频点的调整数值 $D(k)$ 满足: $D(k) = F(d(k))$, 所述 $d(k)$ 满足: $d(k) = \begin{bmatrix} I_{1 \times n} & 0_{1 \times n} \\ 0_{1 \times n} & 0_{1 \times n} \end{bmatrix} \times f \times F^{-1}[\mu_{out}(k) \times H(k)]$;

所述 f 为窗函数, F 表示傅立叶变换, 所述 F^{-1} 表示逆傅立叶变换。

7. 根据权利要求1至6任一所述的方法, 其特征在于, 所述基于所述目标权重矩阵对所述第一本端音频信号进行处理, 得到第一回声估计信号, 包括:

对所述第一本端音频信号进行分段处理, 得到 P 段第一子音频信号;

对所述目标权重矩阵进行分段处理, 得到 P 段目标权重子矩阵;

基于每段所述目标权重子矩阵对对应的一段所述第一子音频信号进行处理, 得到第一回声估计信号;

其中, 每段所述第一子音频信号包括 n 个频点, 每段所述目标权重子矩阵包括 n 个权重

系数,所述P为大于0且小于 $L \times s$ 的整数,n为不大于P的整数。

8.根据权利要求7所述的方法,其特征在于,所述第一回声估计信号y满足:

$$y = [0_{1 \times n}, I_{1 \times n}] \times F^{-1} \left[\sum_{j=0}^{P-1} X_j \times W_j \right];$$

其中,所述 $I_{1 \times n}$ 为 $1 \times n$ 的单位矩阵,所述 X_j 为第j段所述第一子音频信号的第四频谱值,所述 $X_j = \text{diag} \left(F \begin{bmatrix} x_j \\ x_{j+1} \end{bmatrix} \right)$,所述 x_j 为第j段所述第一子音频信号的信号值,所述 x_{j+1} 为第j+1段

所述第一子音频信号的信号值,diag表示对角矩阵;

所述 W_j 为第j段所述目标权重子矩阵的目标权重系数,所述j为大于或等于0,且小于所述P的整数,所述 $0_{1 \times n}$ 为 $1 \times n$ 的全0向量。

9.一种回声处理装置,其特征在于,应用于第一终端,所述装置包括:

采集模块,用于采集第一本端音频信号;

第一确定模块,用于根据初始权重矩阵和调整矩阵确定目标权重矩阵,其中,所述初始权重矩阵为在所述第一本端音频信号之前采集到的第二本端音频信号对应的权重矩阵,所述调整矩阵中的调整数值与所述第二本端音频信号的步长因子矩阵中的步长因子正相关,所述目标权重矩阵中的目标权重系数与所述初始权重矩阵中的初始权重系数和所述调整数值均正相关,所述步长因子根据所述第二本端音频信号确定;

第一处理模块,用于基于所述目标权重矩阵对所述第一本端音频信号进行处理,得到第一回声估计信号;

第二处理模块,用于在接收到第二终端传输的第一对端音频信号后,采用所述第一回声估计信号对所述第一对端音频信号进行处理,得到第一目标音频信号;

播放模块,用于播放所述第一目标音频信号。

10.一种回声处理系统,其特征在于,所述系统包括:第一终端和第二终端,所述第一终端包括如权利要求9所述的回声处理装置;

所述第二终端与所述第一终端建立有通信连接,所述第二终端用于向所述第一终端发送音频信号。

11.一种计算机可读存储介质,其特征在于,所述计算机可读存储介质中存储有指令,所述指令由处理器加载并执行以实现如权利要求1至8任一所述的回声处理方法。

回声处理方法、装置及系统

技术领域

[0001] 本公开涉及音频处理技术领域,特别涉及一种回声处理方法、装置及系统。

背景技术

[0002] 目前,本端用户在通过本端设备与对端用户进行语音通话的过程中,本端设备可以将接收到的本端用户的语音传输至对端设备。该语音经过对端设备播放后,可能再次被对端设备的麦克风拾取从而形成回声。该回声会再次通过对端设备传输至本端设备,从而导致本端用户听到回声。

发明内容

[0003] 本公开实施例提供了一种回声处理方法、装置及系统,可以解决相关技术中本端用户在通过本端设备与对端用户进行语音通话的过程中,本端用户会听到回声的问题。所述技术方案如下:

[0004] 一方面,提供了一种回声处理方法,应用于第一终端,所述方法包括:

[0005] 采集第一本端音频信号;

[0006] 根据初始权重矩阵和调整矩阵确定目标权重矩阵,其中,所述初始权重矩阵为在所述第一本端音频信号之前采集到的第二本端音频信号对应的权重矩阵,所述调整矩阵中的调整数值与所述第二本端音频信号的步长因子矩阵中的步长因子正相关,所述目标权重矩阵中的目标权重系数与所述初始权重矩阵中的初始权重系数和所述调整数值均正相关,所述步长因子根据所述第二本端音频信号确定;

[0007] 基于所述目标权重矩阵对所述第一本端音频信号进行处理,得到第一回声估计信号;

[0008] 在接收到第二终端传输的第一对端音频信号后,采用所述第一回声估计信号对所述第一对端音频信号进行处理,得到第一目标音频信号;

[0009] 播放所述第一目标音频信号。

[0010] 可选的,所述第二本端音频信号包括 L 个音频帧,每个所述音频帧包括 s 个频点;所述初始权重矩阵包括与所述第二本端音频信号中 $L \times s$ 个频点一一对应的 $L \times s$ 个初始权重系数,所述调整矩阵包括与所述 $L \times s$ 个频点一一对应的 $L \times s$ 个调整数值, L 和 s 均为大于或等于1的整数;在根据初始权重矩阵和调整矩阵确定目标权重矩阵之前,所述方法还包括:

[0011] 基于所述初始权重矩阵对所述第二本端音频信号进行处理,得到第二回声估计信号;

[0012] 在接收到所述第二终端传输的第二对端音频信号后,采用所述第二回声估计信号对所述第二对端音频信号进行处理,得到第二目标音频信号;

[0013] 基于所述第二本端音频信号的频域自相关系数,以及所述第二本端音频信号和所述第二目标音频信号的频域互相关系数,确定每个所述音频帧的音频泄漏系数,所述音频泄漏系数与所述频域自相关系数负相关,且与所述频域互相关系数正相关;

[0014] 对于所述 $L \times s$ 个频点中的每个频点,基于所述频点所属的音频帧的音频泄露系数,所述第二目标音频信号在所述频点的第一频谱值,以及所述第二回声估计信号在所述频点的第二频谱值,确定所述频点对应的所述步长因子,所述步长因子与所述频点所属的音频帧的音频泄露系数和所述第二频谱值均正相关,且与所述第一频谱值负相关;

[0015] 采用块稀疏成比例仿射投影算法对所述第二本端音频信号在所述频点的第三频谱值以及所述第二目标音频信号进行处理,得到所述频点对应的更新步长;

[0016] 基于所述频点对应的步长因子和所述频点对应的更新步长,确定所述频点对应的调整数值;

[0017] 所述根据初始权重矩阵和调整矩阵确定目标权重矩阵,包括:

[0018] 将所述第二本端音频信号中每个所述频点对应的初始权重系数与所述频点对应的调整数值之和确定为所述第一本端音频信号中所述频点的目标权重系数。

[0019] 可选的,所述第二本端音频信号中第 i 个音频帧的音频泄露系数 $\eta(i)$ 满足:

$$[0020] \quad \eta(i) = \frac{\sum_{k=(i-1) \times s}^{i \times s - 1} R_{EX}(k, i)}{\sum_{k=(i-1) \times s}^{i \times s - 1} R_{XX}(k, i)};$$

[0021] 其中,所述 i 为大于0,且小于或等于 L 的整数,所述 k 为大于或等于0,且小于 $L \times s$ 的整数;所述 $R_{EX}(k, i)$ 为所述第二本端音频信号中第 k 个频点和所述第二目标音频信号中第 k 个频点的频域互相关系数;

[0022] 所述第 k 个频点对应的步长因子 $\mu_{out}(k)$ 满足: $\mu_{out}(k) = \min\{\eta(i) \times \frac{|y(k)|^2}{|e(k)|^2}, \mu_{max}\};$

[0023] 其中, $\min$ 表示取最小值,所述 μ_{max} 为大于0且小于1的数,所述 $y(k)$ 为所述第二回声估计信号中第 k 个频点的第二频谱值,所述 $e(k)$ 为所述第二目标音频信号中第 k 个频点第一频谱值。

[0024] 可选的,所述第二本端音频信号包括 P 段第二子音频信号,每段所述第二子音频信号包括 n 个频点,所述 $L \times s$ 个频点中第 k 个频点对应的更新步长 $H(k)$ 满足:

$$[0025] \quad H(k) = \frac{\mu \times G \times X^H(k) \times \text{sgn}[e]}{\sqrt{[G \times X^H(k) \times \text{sgn}[e]]^T} \times (G \times X^H(k) \times \text{sgn}[e]) + \beta \times I};$$

[0026] 其中,所述 k 为大于或等于0,且小于 $L \times s$ 的整数,所述 μ 为大于0的数, G 为对角矩阵,所述对角矩阵中对角线上的第 t 个元素为 $g_t(t) \times I_{1 \times n}$,且 $g_t(t)$ 与 $w(t)$ 正相关,所述 $w(t)$ 为所述第二本端音频信号中第 t 个频点对应的初始权重系数,所述 $I_{1 \times n}$ 为 $1 \times n$ 的单位矩阵,所述 $X^H(k)$ 为对所述第二本端音频信号中第 k 个频点的频谱值的共轭转置,所述 $\text{sgn}[e]$ 为行向量,所述行向量中第 t 个元素为 $\text{sgn}(e(t))$,所述 $e(t)$ 为所述第二目标音频信号中第 t 个频点的第一频谱值,所述 sgn 为符号函数,所述 t 的取值范围为 k 至 $1 \times n - 1$,所述第 k 个频点属于第1段所述第二子音频信号,所述 1 为大于或等于0,且小于 P 的正整数,所述 β 为不等于0的整数,所述 $n = \frac{L}{P} \times s$,所述 P 为大于0且小于 $L \times s$ 的整数, n 为不大于 P 的整数。

[0027] 可选的,所述 $L \times s$ 个频点中第 k 个频点的调整数值 $D(k)$ 满足: $D(k) = F(d(k))$,所述

$d(k)$ 满足: $d(k) = \begin{bmatrix} I_{l \times n} & 0_{l \times n} \\ 0_{l \times n} & 0_{l \times n} \end{bmatrix} \times F^{-1}[\mu_{out}(k) \times H(k)];$

[0028] 其中,所述 k 为大于或等于0,且小于 $L \times s$ 的整数, F 表示傅立叶变换,所述 F^{-1} 表示逆傅立叶变换。

[0029] 可选的,所述 $L \times s$ 个频点中第 k 个频点的调整数值 $D(k)$ 满足: $D(k) = F(d(k))$,所述

$d(k)$ 满足: $d(k) = \begin{bmatrix} I_{l \times n} & 0_{l \times n} \\ 0_{l \times n} & 0_{l \times n} \end{bmatrix} \times f \times F^{-1}[\mu_{out}(k) \times H(k)];$

[0030] 所述 f 为窗函数, F 表示傅立叶变换,所述 F^{-1} 表示逆傅立叶变换。

[0031] 可选的,所述基于所述目标权重矩阵对所述第一本端音频信号进行处理,得到第一回声估计信号,包括:

[0032] 对所述第一本端音频信号进行分段处理,得到 P 段第一子音频信号;

[0033] 对所述目标权重矩阵进行分段处理,得到 P 段目标权重子矩阵;

[0034] 基于每段所述目标权重子矩阵对对应的一段所述第一子音频信号进行处理,得到第一回声估计信号;

[0035] 其中,每段所述第一子音频信号包括 n 个频点,每段所述目标权重子矩阵包括 n 个权重系数,所述 P 为大于0且小于 $L \times s$ 的整数, n 为不大于 P 的整数。

[0036] 可选的,所述第一回声估计信号 y 满足: $y = [0_{l \times n}, I_{l \times n}] \times F^{-1} \left[\sum_{j=0}^{P-1} X_j \times W_j \right];$

[0037] 其中,所述 $I_{l \times n}$ 为 $l \times n$ 的单位矩阵,所述 X_j 为第 j 段所述第一子音频信号的第四频谱值,所述 $X_j = \text{diag} \left(F \begin{bmatrix} x_j \\ x_{j+1} \end{bmatrix} \right)$,所述 x_j 为第 j 段所述第一子音频信号的信号值,所述 x_{j+1} 为第

$j+1$ 段所述第一子音频信号的信号值, diag 表示对角矩阵;

[0038] 所述 W_j 为第 j 段所述目标权重子矩阵的目标权重系数,所述 j 为大于或等于0,且小于所述 P 的整数,所述 $0_{l \times n}$ 为 $l \times n$ 的全0向量。

[0039] 另一方面,提供了一种回声处理装置,应用于第一终端,所述装置包括:

[0040] 采集模块,用于采集第一本端音频信号;

[0041] 第一确定模块,用于根据初始权重矩阵和调整矩阵确定目标权重矩阵,其中,所述初始权重矩阵为在所述第一本端音频信号之前采集到的第二本端音频信号对应的权重矩阵,所述调整矩阵中的调整数值与所述第二本端音频信号的步长因子矩阵中的步长因子正相关,所述目标权重矩阵中的目标权重系数与所述初始权重矩阵中的初始权重系数和所述调整数值均正相关,所述步长因子根据所述第二本端音频信号确定;

[0042] 第一处理模块,用于基于所述目标权重矩阵对所述第一本端音频信号进行处理,得到第一回声估计信号;

[0043] 第二处理模块,用于在接收到第二终端传输的第一对端音频信号后,采用所述第一回声估计信号对所述第一对端音频信号进行处理,得到第一目标音频信号;

[0044] 播放模块,用于播放所述第一目标音频信号。

[0045] 再一方面,提供了一种回声处理系统,所述系统包括:第一终端和第二终端,所述

第一终端包括如上述方面所述的回声处理装置；

[0046] 所述第二终端与所述第一终端建立有通信连接,所述第二终端用于向所述第一终端发送音频信号。

[0047] 又一方面,提供了一种计算机可读存储介质,所述计算机可读存储介质中存储有指令,所述指令由处理器加载并执行以实现如上述方面所述的回声处理方法。

[0048] 再一方面,提供了一种终端,包括:存储器,处理器及存储在所述存储器上的计算机程序,所述处理器执行所述计算机程序时实现如上述方面所述的回声处理方法。

[0049] 再一方面,提供了一种包含指令的计算机程序产品,当所述计算机程序产品在所述计算机上运行时,使得所述计算机执行上述方面所述的回声处理方法。

[0050] 本公开实施例提供的技术方案带来的有益效果至少包括:

[0051] 本公开实施例提供了一种回声处理方法、装置及系统,该第一终端可以基于初始权重矩阵和调整矩阵确定目标权重矩阵,并基于通过该目标权重矩阵和第一本端音频信号得到的第一回声信号,对第一对端音频信号进行处理,得到第一目标音频信号,并播放该第一目标音频信号。由于调整数值与步长因子正相关,且该步长因子是根据第二本端音频信号确定,因此提高了确定的调整数值的可靠性,进而提高了确定的第一回声估计信号的准确度,从而有效提高了回声处理的效果。

附图说明

[0052] 为了更清楚地说明本公开实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本公开的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0053] 图1是本公开实施例提供的一种回声处理方法所涉及的实施环境的示意图;

[0054] 图2是本公开实施例提供的一种回声处理方法的流程图;

[0055] 图3是本公开实施例提供的一种回声处理方法的流程图;

[0056] 图4是本公开实施例提供的一种回声处理方法的示意图;

[0057] 图5是本公开实施例提供的一种回声处理装置的框图;

[0058] 图6是本公开实施例提供的另一种回声处理装置的框图。

具体实施方式

[0059] 为使本公开的目的、技术方案和优点更加清楚,下面将结合附图对本公开实施方式作进一步地详细描述。

[0060] 图1是本公开实施例提供的一种回声处理方法所涉及的实施环境的示意图。如图1所示,该实施环境可以包括:第一终端10、服务器20和第二终端30。该第一终端10和第二终端30可以均为个人电脑、笔记本电脑、平板电脑或手机等设备。该服务器20可以是一台服务器,或者由若干台服务器组成的服务器集群,或者是一个云计算服务中心。该第一终端10和第二终端30可以通过该服务器20交互音频信号。该服务器20与该第一终端10和第二终端30之间均可以通过有线网络或无线网络建立连接。

[0061] 图2是本公开实施例提供的一种回声处理方法的流程图。该方法可以应用于图1所

示的第一终端10中。如图2所示,该方法可以包括:

[0062] 步骤201、采集第一本端音频信号。

[0063] 在本公开实施例中,第一终端的用户通过该第一终端与第二终端的用户进行语音通话的过程中,该第一终端可以实时采集本端音频信号。其中,该第一本端音频信号可以为第一终端在第一时长内采集的本端音频信号。可选的,该第一本端音频信号可以为第一终端的用户在第一时长内发出的声音信号。该第一时长可以为第一终端预先存储的固定时长。

[0064] 步骤202、根据初始权重矩阵和调整矩阵确定目标权重矩阵。

[0065] 第一终端在采集第一本端音频信号之前,还可以采集时长为第一时长的第二本端音频信号。该第一终端在采集第一本端音频信号之后,可以根据初始权重矩阵和调整矩阵确定目标权重矩阵。

[0066] 其中,该初始权重矩阵可以为在该第一本端音频信号之前采集到的第二本端音频信号对应的权重矩阵,该调整矩阵中的调整数值与该第二本端音频信号的步长因子矩阵中的步长因子正相关,该目标权重矩阵中的目标权重系数均与初始权重矩阵中的初始权重系数和调整数值正相关。该步长因子矩阵中的步长因子是根据第二本端音频信号确定,即不同的第二本端音频信号可以确定出不同的步长因子,也即是该步长因子为可变步长。可选的,该初始权重矩阵为该第一终端中预先存储的固定权重矩阵。该初始权重矩阵和目标权重矩阵为该第一终端中的自适应滤波器的权重矩阵。

[0067] 步骤203、基于目标权重矩阵对第一本端音频信号进行处理,得到第一回声估计信号。

[0068] 第一终端在确定目标权重矩阵之后,可以基于该目标权重矩阵对该第一本端音频信号进行处理,从而得到第一回声估计信号。

[0069] 步骤204、在接收到第二终端传输的第一对端音频信号后,采用第一回声估计信号对第一对端音频信号进行处理,得到第一目标音频信号。

[0070] 第一终端的用户在通过第一终端与第二终端的用户进行语音通话的过程中,第二终端也可以采集在第二时长内的对端语音信号,并将该采集到的在第二时长内的第一对端音频信号通过服务器传输至第一终端。第一终端在接收到该第一对端音频信号后,可以采用第一回声估计信号对第一对端音频信号进行处理,从而得到第一目标音频信号。其中,该第二时长为第二终端存储的固定时长,该第二时长与第一时长可以相同,也可以不同,本公开实施例对此不做限定。

[0071] 可选的,第一终端可以将第一对端音频信号和第一回声估计信号相减得到该第一目标音频信号,即该第一目标音频信号为该第一对端音频信号中除该第一回声估计信号之外的音频信号。

[0072] 由于第一本端音频信号经过服务器传输至第二终端,并由该第二终端播放后,可能再次被第二终端的麦克风拾取从而形成回声信号。该回声信号会和第二终端的用户的语音信号一同传输至第一终端。因此,第一终端接收到的第一对端音频信号实际上包括该回声信号和第二终端的用户的语音信号。

[0073] 在本公开实施例中,该第一回声估计信号为第一终端基于初始权重矩阵和调整矩阵所确定的目标权重矩阵所模拟实际的回声路径估计的回声信号,该第一终端确定的第一

回声估计信号与实际的回声信号越接近,则该第一目标音频信号中的回声信号越少,第一终端对该回声信号的处理的效果越好。

[0074] 步骤205、播放第一目标音频信号。

[0075] 该第一终端在确定第一目标音频信号后,可以播放该第一目标音频信号。

[0076] 综上所述,本公开实施例提供了一种回声处理方法,该第一终端可以基于初始权重矩阵和调整矩阵确定目标权重矩阵,并基于通过该目标权重矩阵和第一本端音频信号得到的第一回声信号,对第一对端音频信号进行处理,得到第一目标音频信号,并播放该第一目标音频信号。由于该调整数值与步长因子正相关,且该步长因子是根据第二本端音频信号确定,因此提高了确定的调整数值的可靠性,进而提高了确定的第一回声估计信号的准确度,从而有效提高了回声处理的效果。

[0077] 图3是本公开实施例提供的另一种回声处理方法的流程图。该方法可以应用于图1所示的第一终端10中。如图3所示,该方法可以包括:

[0078] 步骤301、采集第一本端音频信号。

[0079] 在本公开实施例中,第一终端的用户通过该第一终端与第二终端的用户进行语音通话的过程中,该第一终端可以实时采集本端音频信号。其中,该第一本端音频信号可以为第一终端在第一时长内采集的本端音频信号。可选的,该第一本端音频信号可以为第一终端的用户在第一时长内发出的声音信号。该第一时长可以为第一终端预先存储的固定时长,即该第一终端能够以第一时长为单位采集音频信号。

[0080] 步骤302、基于初始权重矩阵对第二本端音频信号进行处理,得到第二回声估计信号。

[0081] 第一终端在采集第一本端音频信号之前,还可以采集时长为第一时长的第二本端音频信号。则第一终端在采集到第二本端音频信号之后,可以根据初始权重矩阵对该第二本端音频信号进行处理,得到第二回声估计信号。其中,该初始权重矩阵为该第一终端中预先存储的固定权重矩阵,该初始权重矩阵为该第一终端中的自适应滤波器的系数矩阵。

[0082] 其中,该第二本端音频信号可以包括L个音频帧,每个音频帧包括s个频点。该初始权重矩阵可以包括与第二本端音频信号中 $L \times s$ 个频点一一对应的 $L \times s$ 个初始权重系数。该L和s均为大于或等于1的整数。

[0083] 参考图4,第一终端在接收到时长为第一时长T的第二本端音频信号 $X_0(T)$ 之后,可以对该第二本端音频信号 $X_0(T)$ 进行分段处理,从而得到P段第二子音频信号。

[0084] 参考图4,第一终端在得到P段第二子音频信号之后,可以对每段第二子音频信号依次进行加窗处理和傅立叶变换,从而得到第二本端音频信号中每个频点的第三频谱值。示例的,第一终端可以对每段第二子音频信号加窗函数f1。

[0085] 之后第一终端可以对每个频点进行延迟处理。对于每个频点,第一终端可以将该第k个频点的第三频谱值与对应的一个初始权重系数 $w(k)$ 相乘,得到第k个频点的第三频谱值 $y_0(k)$,该 $y_0(k) = X_0(k) \times w_0(k)$ 。之后第一终端可以将该 $L \times s$ 个第二频谱值相加,得到在频域上的第二回声估计信号 Y_0 。之后第一终端可以对 Y_0 进行逆傅立叶变换。并对进行逆傅立叶变换之后的 Y_0 进行加窗处理,得到第二回声估计信号 y_0 。示例的,第一终端可以对进行逆傅立叶变换之后的 Y_0 加窗函数f2。 k 为大于或等于0,且小于 $L \times s$ 的整数。可选的,该窗函数f1和f2均可以为汉明窗函数、矩形窗函数或者三角窗函数等。

[0086] 步骤303、在接收到第二终端传输的第二对端音频信号后,采用第二回声估计信号对第二对端音频信号进行处理,得到第二目标音频信号。

[0087] 第一终端的用户在通过第一终端与第二终端的用户进行语音通话的过程中,第二终端也可以采集在第二时长内的对端语音信号,并将该采集到的在第二时长内的第二对端音频信号通过服务器传输至第一终端。第一终端在接收到该第二对端音频信号后,可以采用该第二回声估计信号对第二对端音频信号进行处理,得到第二目标音频信号。其中,该第二时长为第二终端存储的固定时长,并且该第二时长与该第一时长可以相同,也可以不同,本公开实施例对此不做限定。

[0088] 可选的,参考图4,第一终端可以将第二对端音频信号d和第一回声估计信号 y_0 相减得到该第二目标音频信号E,即该第二目标音频信号为该第二对端音频信号中除该第二回声估计信号之外的音频信号。同时,该第一终端可以对该第二目标音频信号E进行傅立叶

变换,得到在频域上的第二目标音频信号 $e = F \begin{bmatrix} 0_{1 \times n} \\ E \end{bmatrix}$,该 $0_{1 \times n}$ 为 $1 \times n$ 的权0向量,该F表示傅立叶变换。

[0089] 由于第二本端音频信号经过服务器传输至第二终端,并由该第二终端播放后,可能再次被第二终端的麦克风拾取从而形成回声信号。该回声信号会和第二终端的用户的语音信号一同传输至第一终端。因此,第一终端接收到的第二对端音频信号实际上包括该回声信号和第二终端的用户的语音信号。

[0090] 在本公开实施例中,该第二回声估计信号为第一终端基于初始权重矩阵模拟实际的回声路径所估计的回声信号,该第一终端确定的第二回声估计信号与实际的回声信号越接近,则该第二目标音频信号中的回声信号越少,第一终端对该回声信号的处理的效果越好。

[0091] 步骤304、基于第二本端音频信号的频域自相关系数,以及第二本端音频信号和第二目标音频信号的频域互相关系数,确定每个音频帧的音频泄漏系数。

[0092] 第一终端在确定第二目标音频信号后,可以播放该第二目标音频信号。同时,第一终端可以分别确定该第二本端音频信号的频域自相关系数,以及该第二本端音频信号和第二目标音频信号的频域互相关系数。之后,第一终端可以基于该频域自相关系数和频域互相关系数,确定每个音频帧的音频泄漏系数。

[0093] 其中,该音频泄漏系数与该自相关系数负相关,且与该互相关系数正相关。也即是,该自相关系数越大,该音频泄露系数越小,该自相关系数越小,该音频泄露系数越大;该互相关系数越大,该音频泄露系数越大,该互相关系数越小,该音频泄露系数越小。

[0094] 可选的,该第一终端可以确定该第二本端音频信号中第k个频点的频域自相关系数 $R_{XX}(k, i)$ 满足: $R_{XX}(k, i) = (1 - \alpha(i)) \times R_{XX}(k, i - 1) + \alpha(i) \times (P_X(k))^2$ 。

[0095] 其中,该i为大于0,且小于或等于L的整数,该第k个频点属于该第i个音频帧。该 $P_X(k)$ 为第二本端音频信号中第k个频点的功率谱。该 $\alpha(i) = \alpha_0 \times \min \left(\frac{\delta_X^2(i)}{\delta_E^2(i)}, 1 \right)$,该 $\delta_X^2(i)$ 为第二本端音频信号中第i帧音频信号的方差,该 $\delta_E^2(i)$ 为第二目标音频信号中第i帧音频信号的方差,该 α_0 为大于0,且小于0.01的数。

[0096] 该第一终端可以确定第二本端音频信号中第k个频点和第二目标音频信号中第k个频点的频域互相关系数 $R_{EX}(k, i)$ 满足： $R_{EX}(k, i) = (1 - \alpha(i)) \times R_{EX}(k, i-1) + \alpha(i) \times P_X(k) \times P_E(k)$ ，其中，该 $P_E(k)$ 为第二目标音频信号中第k个频点的功率谱。

[0097] 之后，参考图4，第一终端基于频域自相关系数 $R_{XX}(k, i)$ 和频域互相关系数 $R_{EX}(k, i)$ 可以确定第二本端音频信号中第i个音频帧的音频泄露系数 $\eta(i)$ 满足：

$$\eta(i) = \frac{\sum_{k=(i-1) \times s}^{i \times s - 1} R_{EX}(k, i)}{\sum_{k=(i-1) \times s}^{i \times s - 1} R_{XX}(k, i)}。$$

[0098] 步骤305、对于 $L \times s$ 个频点中的每个频点，基于频点所属的音频帧的音频泄露系数，第二目标音频信号在频点的第一频谱值，以及第二回声估计信号在频点的第二频谱值，确定频点对应的步长因子。

[0099] 第一终端在确定每个音频帧的音频泄露系数之后，可以分别确定该频点所属的音频帧的音频泄露系数，该第二目标音频信号在该频点的第一频谱值，以及第二回声估计信号在该频点的第二频谱值。之后第一终端可以基于该确定的频点所属的音频帧的音频泄露系数，该第二目标音频信号在该频点的第一频谱值，以及第二回声估计信号在该频点的第二频谱值，确定该频点对应的步长因子，从而得到步长因子矩阵，该步长因子矩阵可以包括 $L \times s$ 个步长因子。

[0100] 其中，该步长因子与该频点所属的音频帧的音频泄露系数和第二频谱值均正相关，且与该第一频谱值负相关，也即是音频泄露系数越大，该步长因子越大，该音频泄露系数越小，该步长因子越小；该第二频谱值越大，该步长因子越大，该第二频谱值越小，该步长因子越小；该第一频谱值越大，该步长因子越小，该第二频谱值越大，该步长因子越小。

[0101] 可选的，参考图4，第一终端基于第k个频点所属的第i帧音频帧的泄露系数 $\eta(i)$ ，第二目标音频信号中第k个频点的第一频谱值 $e(k)$ 以及第二回声估计信号中第k个频点的第二频谱值 $y(k)$ ，可以确定该第k个频点对应的步长因子 $\mu_{out}(k)$ 。该

$$\mu_{out}(k) = \min\left\{\eta(i) \times \frac{|y(k)|^2}{|e(k)|^2}, \mu_{max}\right\}，由此第一终端可以确定出 $L \times s$ 个频点中的每个频点的步长$$

因子，得到步长因子矩阵。其中，该min表示取最小值，该 μ_{max} 为大于0且小于1的数。

[0102] 步骤306、采用块稀疏成比例仿射投影算法对第二本端音频信号在频点的第三频谱值以及第二目标音频信号进行处理，得到频点对应的更新步长。

[0103] 第一终端还可以采用块稀疏成比例仿射投影算法对第二本端音频信号在第k个频点的第三频谱值以及第二目标音频信号进行处理，从而得到该第k个频点对应的更新步长。

[0104] 在本公开实施例中，该第二本端音频信号可以包括P段第二子音频信号，该每段第二子音频信号可以包括n个频点，该 $L \times s$ 个频点中第k个频点对应的更新步长 $H(k)$ 满足：

$$H(k) = \frac{\mu \times G \times X_0^H(k) \times \text{sgn}[c]}{\sqrt{[G \times X_0^H(k) \times \text{sgn}[c]]^T \times (G \times X_0^H(k) \times \text{sgn}[c])} + \beta \times I}。$$

[0105] 其中，该 $n = \frac{L}{P} \times s$ ，该P为大于0且小于 $L \times s$ 的整数，该n为不大于P的整数。该 μ 为大

于0的数, G 为对角矩阵, 该对角矩阵中对角线上的第 t 个元素为 $g_t(t) \times I_{1 \times n}$, 且 $g_t(t)$ 与 $w(t)$ 正相关, 该 $w(t)$ 为第二本端音频信号中第 t 个频点对应的初始权重系数。该 $I_{1 \times n}$ 为 $1 \times n$ 的单位矩阵, 参考图4, $X_0^H(k)$ 为对第二本端音频信号中第 k 个频点的频谱值的共轭转置。该 $\text{sgn}[e]$ 为行向量, 该行向量中第 t 个元素为 $\text{sgn}(e(t))$ 。该 $e(t)$ 为该第二目标音频信号中第 t 个频点的第一频谱值, 该 t 的取值范围为 k 至 $1 \times n - 1$, 该第 k 个频点属于第1段第二子音频信号。该 l 为大于或等于0, 且小于 P 的正整数, 该 β 为不等于0的整数。 $[G \times X_0^H(k) \times \text{sgn}[e]]^T$ 为该 $G \times X_0^H(k) \times \text{sgn}[e]$ 的转置矩阵。

[0106] 其中, 该 sgn 为符号函数, 若该 $e(t)$ 大于0, 则该 $\text{sgn}(e(t))$ 为1, 若该 $e(t)$ 小于0, 则该 $\text{sgn}(e(t))$ 为-1。若该 $e(t)$ 等于0, 则该 $\text{sgn}(e(t))$ 为0。

[0107] 步骤307、基于频点对应的步长因子和频点对应的更新步长, 确定频点对应的调整数值。

[0108] 在本公开实施例中, 第一终端在确定每个频点的步长因子和该频点对应的更新步长之后, 可以基于该每个频点对应的步长因子和该频点对应的更新步长确定该频点对应的调整数值, 由此可以得到调整矩阵。该调整矩阵可以包括与 $L \times s$ 个频点一一对应的 $L \times s$ 个调整数值。

[0109] 其中, 该 $L \times s$ 个频点中第 k 个频点的调整数值 $D(k)$ 满足: $D(k) = F(d(k))$ 。

[0110] 在本公开实施例一种可选的实现方式中, 参考图4, 该 $d(k)$ 满足:

$$d(k) = \begin{bmatrix} I_{1 \times n} & 0_{1 \times n} \\ 0_{1 \times n} & 0_{1 \times n} \end{bmatrix} \times F^{-1}[\mu_{\text{out}}(k) \times H(k)]$$
。其中, 该 F 表示傅立叶变换, 即该 $F(d(k))$ 表示对该 $d(k)$ 进行傅立叶变换, 该 F^{-1} 表示逆傅立叶变换, 即该 $F^{-1}[\mu_{\text{out}}(k) \times H(k)]$ 表示对该 $\mu_{\text{out}}(k) \times H(k)$ 进行逆傅立叶变换。

[0111] 在本公开实施例另一种可选的实现方式中, 该 $d(k)$ 满足:

$$d(k) = \begin{bmatrix} I_{1 \times n} & 0_{1 \times n} \\ 0_{1 \times n} & 0_{1 \times n} \end{bmatrix} \times f \times F^{-1}[\mu_{\text{out}}(k) \times H(k)]$$
。其中, 该 f 为窗函数。可选的, 该窗函数可以为汉明窗函数、矩形窗函数或者三角窗函数等, 本公开实施例对此不做限定。

[0112] 步骤308、将第二本端音频信号中每个频点对应的初始权重系数与频点对应的调整数值之和确定为第一本端音频信号中频点的目标权重系数。

[0113] 第一终端在确定每个频点对应的调整数值之后, 参考图4, 该第一终端可以将该第二本端音频信号中每个频点对应的初始权重系数与频点对应的调整数值之和确定为第一本端音频信号中频点的目标权重系数, 由此得到目标权重矩阵。

[0114] 其中, 该调整矩阵与该第二本端音频信号的步长因子矩阵正相关, 该目标权重矩阵均与初始权重矩阵和调整矩阵正相关。该目标权重矩阵为该第一终端中的自适应滤波器的系数矩阵。

[0115] 步骤309、对第一本端音频信号进行分段处理, 得到 P 段第一子音频信号。

[0116] 第一终端在确定目标权重矩阵之后, 可以对接收到的该第一本端音频信号进行分段处理, 从而得到 P 段第一子音频信号, 该每段第一子音频信号在频域上可以包括 n 个频点。

[0117] 示例的, 若第一本端音频信号的时长为 T , 该第一本端音频信号可以表示为 $X(T)$,

该 $X(T) = \begin{bmatrix} x_0 \\ x_1 \\ \dots \\ x_{P-1} \end{bmatrix}$, 即该第一本端音频信号包括 x_0 至 x_{P-1} 共 P 段第一子音频信号。

[0118] 第一终端在得到 P 段第一子音频信号之后, 可以对每段第一子音频信号进行傅立叶变换, 从得到每段第一子音频信号中每个频点的第四频谱值。该第一终端还可以对每段第一子音频信号进行延迟处理, 以对每段第一子音频信号进行修正。

[0119] 步骤310、对目标权重矩阵进行分段处理, 得到 P 段目标权重子矩阵。

[0120] 第一终端在确定目标权重矩阵后, 还可以对目标权重矩阵进行分段处理, 从而得到 P 段目标权重子矩阵, 该每段目标权重子矩阵包括与 n 个频点一一对应的 n 个权重系数。

[0121] 示例的, 该目标权重矩阵可以表示为 W , 该 $W = \begin{bmatrix} W_0 \\ W_1 \\ \dots \\ W_{P-1} \end{bmatrix}$, 即该目标权重矩阵可以包括

W_0 至 W_{P-1} 共 P 段目标权重子矩阵。第一终端中的自适应滤波器的阶数可以为 N 阶, 该第一终端可以将该自适应滤波器划分为 P 个自适应滤波器块, 该每个目标权重子矩阵即为该自适应滤波器块的系数。

[0122] 步骤311、基于每段目标权重子矩阵对对应的一段第一子音频信号进行处理, 得到第一回声估计信号。

[0123] 第一终端在得到 P 段第一子音频信号和 P 段目标权重子矩阵之后, 可以基于第 j 段目标权重子矩阵对第 j 段第一子音频信号进行处理, 从而得到第一回声估计信号。

[0124] 该第一回声估计信号 y 满足: $y = [0_{1 \times n}, I_{1 \times n}] \times F^{-1} \left[\sum_{j=0}^{P-1} X_j \times W_j \right]$ 。

[0125] 其中, 该 j 大于或等于 0, 且小于 P 。该 $0_{1 \times n}$ 为 $1 \times n$ 的全 0 向量。该 X_j 为第 j 段第一子音频信号和第 $j+1$ 段第一子音频信号的第四频谱值, 该 $X_j = \text{diag} \left(F \begin{bmatrix} x_j \\ x_{j+1} \end{bmatrix} \right)$, 该 x_j 为第 j 段第一

子音频信号的信号值, 该 x_{j+1} 为第 $j+1$ 段第一子音频信号的信号值。该 diag 表示对角矩阵, 该 W_j 为第 j 段目标权重子矩阵的目标权重系数。

[0126] 在本公开实施例中, 第一终端在得到 P 段第一子音频信号之后, 还可以对该每段第一子音频信号进行加窗处理。即第 j 段 $x_j = f1 \times x(T - j \times M)$, 则该第一估计回声信号 y 满足:

$y = [0_{1 \times n}, I_{1 \times n}] \times f2 \times F^{-1} \left[\sum_{j=0}^{P-1} X_j \times W_j \right]$, 其中, 该 $f1$ 和 $f2$ 均为窗函数, 该 $M = \frac{L}{2 \times P}$ 。

[0127] 第一终端在对第一本端音频信号进行分段处理后, 通过对每段第一子音频信号进行加窗处理, 确保了相邻两段第一子音频信号的连续性, 从而确保了最终播放的第一目标音频信号在听感上的连续性。

[0128] 步骤312、在接收到第二终端传输的第一对端音频信号后, 采用第一回声估计信号对第一对端音频信号进行处理, 得到第一目标音频信号。

[0129] 第一终端的用户在通过第一终端与第二终端的用户进行语音通话的过程中,第二终端也可以采集在第二时长内的对端语音信号,并将该采集到的在第二时长内的第一对端音频信号通过服务器传输至第一终端。第一终端在接收到该第一对端音频信号后,可以采用第一回声估计信号对第一对端音频信号进行处理,得到第一目标音频信号。

[0130] 可选的,第一终端可以将第一对端音频信号第一回声估计信号相减得到该第一目标音频信号,即该第一目标音频信号为该第一对端音频信号中除该第一回声估计信号之外的音频信号。

[0131] 由于第一本端音频信号经过服务器传输至第二终端,并由该第二终端播放后,可能再次被第二终端的麦克风拾取从而形成回声信号。该回声信号会和第二终端的用户的语音信号一同传输至第一终端。因此,第一终端接收到的第一对端音频信号实际上包括该回声信号和第二终端的用户的语音信号。

[0132] 在本公开实施例中,该第一回声估计信号为第一终端基于初始权重矩阵和调整矩阵确定的目标权重矩阵模拟实际的回声路径所估计的回声信号,该第一终端确定的第一回声估计信号与实际的回声信号越接近,则该第一目标音频信号中的回声信号越少,第一终端对该回声信号的处理的效果越好。

[0133] 步骤313、播放第一目标音频信号。

[0134] 该第一终端在确定第一目标音频信号后,可以播放该第一目标音频信号。

[0135] 本公开实施例提供的回声处理方法,该方法由于步长因子矩阵可以不断进行更新,即该步长因子矩阵中的步长为可变步长。因此该回声处理方法能够在长时延及自适应滤波器阶数较高,且系统稀疏性未知条件下,对与输入的本端音频信号有较好的收敛速度,较好的跟踪性能,并且能够避免非高斯噪声对于自适应滤波器滤波的干扰。使得对本端音频信号的滤波效果更好,从而更好的消除回声。

[0136] 以下对调整数值的推导过程进行说明,首先引入一个对角矩阵 G ,该对角矩阵中对角线上的第 t 个元素为 $g_t(t) \times I_{1 \times n}$,基于成比例仿射投影算法中的权重迭代公式为:

$Z1(T) = Z0(T) + \mu \frac{G \times e(T) \times X(T)}{X^T(T) \times X(T) + \delta \times I}$ 。为了对非高斯噪声有良好的抑制作用,引入块稀疏

成比例仿射投影算法将该成比例仿射投影算法的范数约束推到: $\min\{\|Z1(T) - Z0(T)\|^2 \times G0^{-1}\}$,该约束条件为: $s.t. V(T) - X^T(T) \times Z1(T) = 0$ 。使用拉格朗日乘子法可以得到该成比例仿射投影算法的代价函数为: $J = (\|Z1(T) - Z0(T)\|)^T G^{-1} (\|Z1(T) - Z0(T)\|) + \lambda1 (V(T) - X^T(T) \times Z1(T))$,对 $\lambda1$ 求解可以得到权重更新公式为:

$$Z1(T) = Z0(T) + \mu \frac{G \times X(T) \times \text{sgn}(e)}{\sqrt{(G \times X(T) \times \text{sgn}(e))^T \times (G \times X(T) \times \text{sgn}(e)) + \delta \times I}}。$$

[0137] 其中,该 $Z1(T)$ 为第一终端在 T 时长内采集到的第一本端信号对应的权重系数,该 $Z0(T)$ 为第一终端在该第一本端信号之前采集的第二本端信号对应的权重系数,该 $V(T)$ 为第一终端在第一本端信号之前接收到的对端信号,该 $e(t)$ 为基于第二本端信号和对端信号所确定的回声估计信号。该 $X(T)$ 为第一本端信号,该 $\|\cdot\|$ 表示的是范数。

[0138] 以下对步长因子的推导过程进行说明。在时域上归一化最小均方(normalized least mean square, NLMS)算法的权重公式为:

$$Z1(T) = Z0(T) + \mu \frac{(V(T) - Z0(T) \times X(T)) \times X(T)}{\|X(T)\|^2} \quad \text{。其中,该权重变化量}$$

$$\Delta(n1) = \Delta(n0) + \mu \frac{(V(T) - \Delta(n0) \times X(T)) \times X(T)}{\|X(T)\|^2}, \text{该 } \Delta(n0) = Z1(T) - Z0(T)。$$

[0139] 该每个时刻的权重失调可以表示为

$$\Delta(n1) = \left\| \Delta(n0) + \mu \times \frac{(V(T) - \Delta(n0) \times X(T)) \times X(T)}{\|X(T)\|^2} \right\|^2, \text{通过进一步推导,该}$$

$$\Delta(n1) = \Delta(n0) \times \left\| 1 - \frac{2 \times \mu}{V} + \frac{\mu^2}{V} + \frac{2 \times \mu \times [(V(T) - \Delta(n0) \times X(T)) \times X(T)]^2}{\Delta(n0) \times \|\Delta(n0)\|^2} \right\|^2 \quad \text{其中,该V为不等于0的数。}$$

[0140] 通过对该 μ 求导可以得到:

$$\mu_o(T) = \frac{1}{1 + \frac{[(V(T) - \Delta(n0)) \times X(T)]^2}{\Delta(n0) \times \|\Delta(n0)\|^2}}, \text{由于该 } \frac{\Delta(n0)}{V \times \|\Delta(n0)\|^2}$$

为残留回声方差,因此该 $\sigma_e^2 = \sigma_s^2 + \sigma_r^2$,因此该

$$\mu_o(T) = \frac{1}{1 + \frac{\sigma_s^2}{\sigma_r^2}} = \frac{\sigma_r^2}{\sigma_e^2} \quad \text{。其中,该 } \sigma_e^2 \text{ 为基于第二}$$

本端信号和对端信号所确定的回声估计信号的方差,该 σ_s^2 为对端信号的方差, σ_r^2 为误差信号的残留回声方差,该误差信号对端信号中除回声估计信号以外的信号。因此该 $\mu_o(T)$ 频域

表达式可以为 $\mu_o(j,i) = \frac{\sigma_s^2(j,i)}{\sigma_e^2(j,i)}$ 。由于残留回声无法直接估计,因此可以近似估计该 $\sigma_s^2(j,i) = \eta(i) \times \sigma_y^2(j,i)$ 。其中,该 $\eta(i)$ 为第i帧音频信号的泄露系数,该 $\sigma_y^2(j,i)$ 为第j段回声估计信号的频谱值。因此可以推导出第k个频点的步长因子满足: $\min\{\eta(i) \times \frac{|y(k)|^2}{|e(k)|^2}, \mu_{\max}\}$ 。

[0141] 将该步长因子带入上述权重更新公式可以得到该第一权重更新公式在频域上的

表达式为: $Z1(k) = Z0(k) + \mu_{out}(k) \frac{\mu \times G \times X(k) \times \text{sgn}(e)}{\sqrt{[G \times X(k) \times \text{sgn}(e)]^T \times (G \times X(k) \times \text{sgn}(e)) + \delta \times I}},$

[0142] 第二权重迭代更新公式为: $Z1(k) = Z0(k) + F \begin{bmatrix} I_{1 \times n} & 0_{1 \times n} \\ 0_{1 \times n} & 0_{1 \times n} \end{bmatrix} F^{-1} [\mu_f X^H(k)],$ 该 μ_f 为 $1 \times n$ 的

行向量,该 μ_f 中第h个数值为 $\frac{\mu_1}{P_h}$,该 μ_1 固定数值,该 P_h 为第二本端信号中第h个频点的信号功率。该h为小于n的数,则结合上述第一权重更新公式和第二权重更新公式可以确定上述调整数值。

[0143] 需要说明的是,本公开实施例提供的回声处理方法步骤的先后顺序可以进行适当

调整,例如,步骤309和步骤310可以根据情况删除。任何熟悉本技术领域的技术人员在本公开揭露的技术范围内,可轻易想到变化的方法,都应涵盖在本公开的保护范围之内,因此不再赘述。

[0144] 综上所述,本公开实施例提供了一种回声处理方法,该第一终端可以基于初始权重矩阵和调整矩阵确定目标权重矩阵,并基于通过该目标权重矩阵和第一本端音频信号得到的第一回声信号,对第一对端音频信号进行处理,得到第一目标音频信号,并播放该第一目标音频信号。由于该调整数值与步长因子正相关,且该步长因子是根据第二本端音频信号确定,因此提高了确定的调整数值的可靠性,进而提高了确定的第一回声估计信号的准确度,从而有效提高了回声处理的效果。

[0145] 图5是本公开实施例提供的一种回声处理装置的框图。该装置可以应用于图1所示的第一终端中。如图5所示,该装置可以包括:

[0146] 采集模块501,用于采集第一本端音频信号。

[0147] 第一确定模块502,用于根据初始权重矩阵和调整矩阵确定目标权重矩阵,其中,初始权重矩阵为在第一本端音频信号之前采集到的第二本端音频信号对应的权重矩阵,调整矩阵中的调整数值与第二本端音频信号的步长因子矩阵中的步长因子正相关,目标权重矩阵中的目标权重系数与初始权重矩阵中的初始权重系数和调整数值均正相关,步长因子根据第二本端音频信号确定。

[0148] 第一处理模块503,用于基于目标权重矩阵对第一本端音频信号进行处理,得到第一回声估计信号。

[0149] 第二处理模块504,用于在接收到第二终端传输的第一对端音频信号后,采用第一回声估计信号对第一对端音频信号进行处理,得到第一目标音频信号。

[0150] 播放模块505,用于播放第一目标音频信号。

[0151] 综上所述,本公开实施例提供了一种回声处理装置,该装置可以基于初始权重矩阵和调整矩阵确定目标权重矩阵,并基于通过该目标权重矩阵和第一本端音频信号得到的第一回声信号,对第一对端音频信号进行处理,得到第一目标音频信号,并播放该第一目标音频信号。由于该调整数值与步长因子正相关,且该步长因子是根据第二本端音频信号确定,因此提高了确定的调整数值的可靠性,进而提高了确定的第一回声估计信号的准确度,从而有效提高了回声处理的效果。

[0152] 可选的,第二本端音频信号包括 L 个音频帧,每个音频帧包括 s 个频点;初始权重矩阵包括与第二本端音频信号中 $L \times s$ 个频点一一对应的 $L \times s$ 个初始权重系数,调整矩阵包括与 $L \times s$ 个频点一一对应的 $L \times s$ 个调整数值, L 和 s 均为大于或等于1的整数;在根据初始权重矩阵和调整矩阵确定目标权重矩阵之前,如图6所示,该装置还可以包括:

[0153] 第三处理模块506,用于基于初始权重矩阵对第二本端音频信号进行处理,得到第二回声估计信号。

[0154] 第四处理模块507,用于在接收到第二终端传输的第二对端音频信号后,采用第二回声估计信号对第二对端音频信号进行处理,得到第二目标音频信号。

[0155] 第二确定模块508,用于基于第二本端音频信号的频域自相关系数,以及第二本端音频信号和第二目标音频信号的频域互相关系数,确定每个音频帧的音频泄漏系数,音频泄漏系数与频域自相关系数负相关,且与频域互相关系数正相关。

[0156] 第三确定模块509,用于对于 $L \times s$ 个频点中的每个频点,基于频点所属的音频帧的音频泄露系数,第二目标音频信号在频点的第一频谱值,以及第二回声估计信号在频点的第二频谱值,确定频点对应的步长因子,步长因子与频点所属的音频帧的音频泄露系数和第二频谱值均正相关,且与第一频谱值负相关。

[0157] 第五处理模块510,用于采用块稀疏成比例仿射投影算法对第二本端音频信号在频点的第三频谱值以及第二目标音频信号进行处理,得到频点对应的更新步长。

[0158] 第四确定模块511,用于基于频点对应的步长因子和频点对应的更新步长,确定频点对应的调整数值。

[0159] 第一确定模块502用于:

[0160] 将第二本端音频信号中每个频点对应的初始权重系数与频点对应的调整数值之和确定为第一本端音频信号中频点的目标权重系数。

[0161] 可选的,第二本端音频信号中第 i 个音频帧的音频泄露系数 $\eta(i)$ 满足:

$$[0162] \quad \eta(i) = \frac{\sum_{k=(i-1)s}^{i \times s - 1} R_{EX}(k, i)}{\sum_{k=(i-1)s}^{i \times s - 1} R_{XX}(k, i)};$$

[0163] 其中, i 为大于0,且小于或等于 L 的整数, k 为大于或等于0,且小于 $L \times s$ 的整数; $R_{EX}(k, i)$ 为第二本端音频信号中第 k 个频点和第二目标音频信号中第 k 个频点的频域互相关系数;

[0164] 第 k 个频点对应的步长因子 $\mu_{out}(k)$ 满足: $\mu_{out}(k) = \min\{\eta(i) \times \frac{|y(k)|^2}{|e(k)|^2}, \mu_{max}\};$

[0165] 其中, $\min$ 表示取最小值, μ_{max} 为大于0且小于1的数, $y(k)$ 为第二回声估计信号中第 k 个频点的第二频谱值, $e(k)$ 为第二目标音频信号中第 k 个频点第一频谱值。

[0166] 可选的,第二本端音频信号包括 P 段第二子音频信号,每段第二子音频信号包括 n 个频点, $L \times s$ 个频点中第 k 个频点对应的更新步长 $H(k)$ 满足:

$$[0167] \quad H(k) = \frac{\mu \times G \times X^H(k) \times \text{sgn}[e]}{\sqrt{[G \times X^H(k) \times \text{sgn}[e]]^T} \times (G \times X^H(k) \times \text{sgn}[e]) + \beta \times I};$$

[0168] 其中, k 为大于或等于0,且小于 $L \times s$ 的整数, μ 为大于0的数, G 为对角矩阵,对角矩阵中对角线上的第 t 个元素为 $g_t(t) \times I_{1 \times n}$,且 $g_t(t)$ 与 $w(t)$ 正相关, $w(t)$ 为第二本端音频信号中第 t 个频点对应的初始权重系数, $I_{1 \times n}$ 为 $1 \times n$ 的单位矩阵, $X^H(k)$ 为对第二本端音频信号中第 k 个频点的频谱值的共轭转置, $\text{sgn}[e]$ 为行向量,行向量中第 t 个元素为 $\text{sgn}(e(t))$, $e(t)$ 为第二目标音频信号中第 t 个频点的第一频谱值, sgn 为符号函数, t 的取值范围为 k 至 $1 \times n - 1$,第 k 个频点属于第 l 段第二子音频信号, l 为大于或等于0,且小于 P 的正整数, β 为不等于0的整数, $n = \frac{L}{P} \times s$, P 为大于0且小于 $L \times s$ 的整数, n 为不大于 P 的整数。

[0169] 可选的, $L \times s$ 个频点中第 k 个频点的调整数值 $D(k)$ 满足: $D(k) = F(d(k))$, $d(k)$ 满

$$\text{足: } d(k) = \begin{bmatrix} I_{1 \times n} & 0_{1 \times n} \\ 0_{1 \times n} & 0_{1 \times n} \end{bmatrix} \times F^{-1}[\mu_{\text{out}}(k) \times H(k)];$$

[0170] 其中, k 为大于或等于 0, 且小于 $L \times s$ 的整数, F 表示傅立叶变换, F^{-1} 表示逆傅立叶变换。

[0171] 可选的, $L \times s$ 个频点中第 k 个频点的调整数值 $D(k)$ 满足: $D(k) = F(d(k))$, $d(k)$ 满

$$\text{足: } d(k) = \begin{bmatrix} I_{1 \times n} & 0_{1 \times n} \\ 0_{1 \times n} & 0_{1 \times n} \end{bmatrix} \times f \times F^{-1}[\mu_{\text{out}}(k) \times H(k)];$$

[0172] f 为窗函数, F 表示傅立叶变换, F^{-1} 表示逆傅立叶变换。

[0173] 可选的, 第一处理模块 503 用于:

[0174] 对第一本端音频信号进行分段处理, 得到 P 段第一子音频信号;

[0175] 对目标权重矩阵进行分段处理, 得到 P 段目标权重子矩阵;

[0176] 基于每段目标权重子矩阵对对应的一段第一子音频信号进行处理, 得到第一回声估计信号;

[0177] 其中, 每段第一子音频信号包括 n 个频点, 每段目标权重子矩阵包括 n 个权重系数, P 为大于 0 且小于 $L \times s$ 的整数, n 为不大于 P 的整数。

$$[0178] \quad \text{可选的, 第一回声估计信号 } y \text{ 满足: } y = [0_{1 \times n}, I_{1 \times n}] \times F^{-1} \left[\sum_{j=0}^{P-1} X_j \times W_j \right];$$

[0179] 其中, $I_{1 \times n}$ 为 $1 \times n$ 的单位矩阵, X_j 为第 j 段第一子音频信号的第四频谱值,

$$X_j = \text{diag} \left(F \begin{bmatrix} x_j \\ x_{j+1} \end{bmatrix} \right), x_j \text{ 为第 } j \text{ 段第一子音频信号的信号值, } x_{j+1} \text{ 为第 } j+1 \text{ 段第一子音频信号}$$

的信号值, diag 表示对角矩阵;

[0180] W_j 为第 j 段目标权重子矩阵的系数, j 大于或等于 0, 且小于 P , $0_{1 \times n}$ 为 $1 \times n$ 的全 0 向量。

[0181] 综上所述, 本公开实施例提供了一种回声处理装置, 该装置可以基于初始权重矩阵和调整矩阵确定目标权重矩阵, 并基于通过该目标权重矩阵和第一本端音频信号得到的第一回声信号, 对第一对端音频信号进行处理, 得到第一目标音频信号, 并播放该第一目标音频信号。由于该调整数值与步长因子正相关, 且该步长因子是根据第二本端音频信号确定, 因此提高了确定的调整数值的可靠性, 进而提高了确定的第一回声估计信号的准确度, 从而有效提高了回声处理的效果。

[0182] 图 1 是本公开实施例提供的一种回声处理系统, 如图 1 所示, 该系统可以包括第一终端 10 和第二终端 30。该第一终端可以包括图 5 或图 6 所示的回声处理装置。该第二终端 30 与该第一终端 10 建立有通信连接, 该第二终端 30 用于向第一终端 10 发送音频信号。

[0183] 本公开实施例提供了一种计算机可读存储介质, 该计算机可读存储介质中存储有指令, 该指令由处理器加载并执行以实现上述方法实施例 (例如如图 2 或图 3 所示的实施例)。

[0184] 本公开实施例提供了一种终端, 包括: 存储器, 处理器及存储在该存储器上的计算机程序, 该处理器执行该计算机程序时实现上述方法实施例 (例如如图 2 或图 3 所示的实施例)。

[0185] 本公开实施例提供了一种包含指令的计算机程序产品,当该计算机程序产品在计算机上运行时,使得计算机实现上述方法实施例(例如图2或图3所示的实施例)。

[0186] 在本申请实施例中,术语“第一”、“第二”、“第三”、“第四”和“第五”仅用于描述目的,而不能理解为指示或暗示相对重要性。本申请实施例中术语“至少一个”的含义是指一个或多个以上。本申请实施例中术语“多个”的含义是指两个或两个以上。

[0187] 以上所述仅为本公开的可选实施例,并不用以限制本公开,凡在本公开的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本公开的保护范围之内。

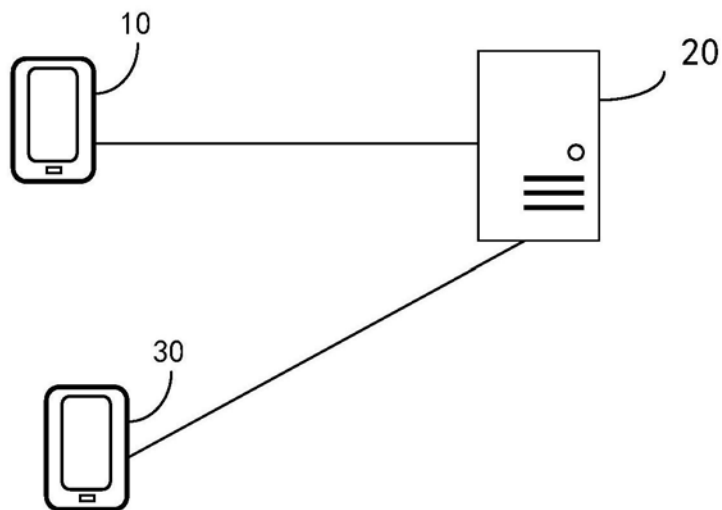


图1

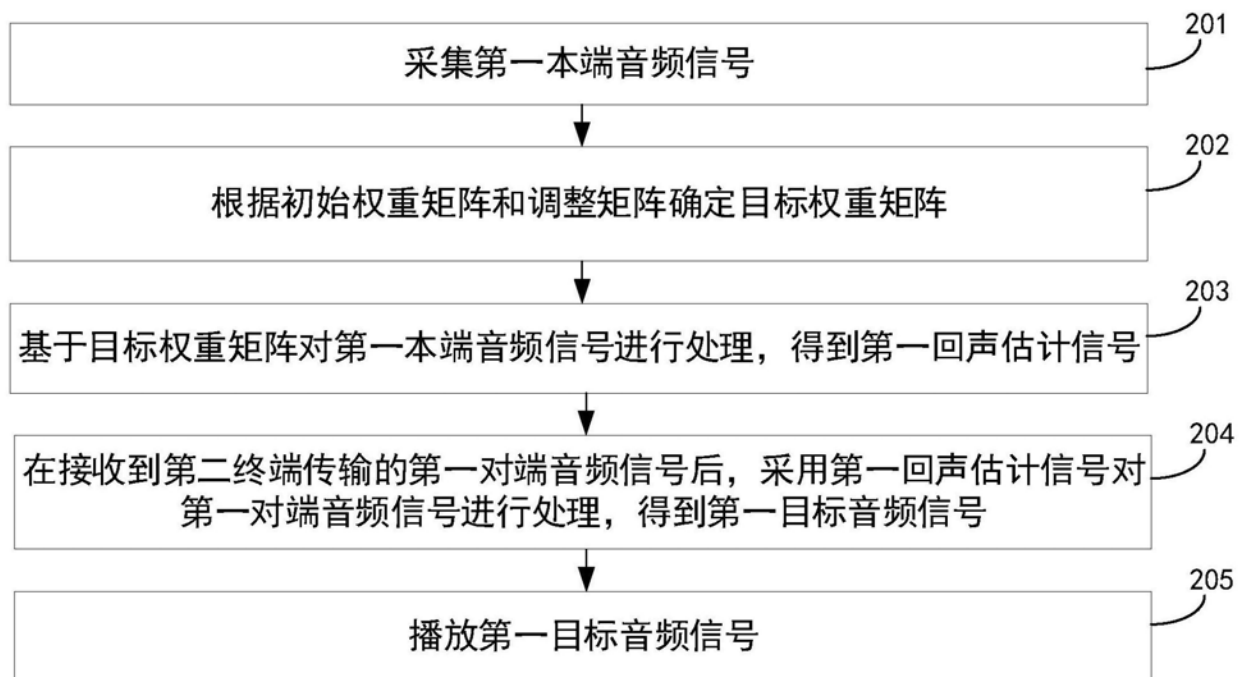


图2

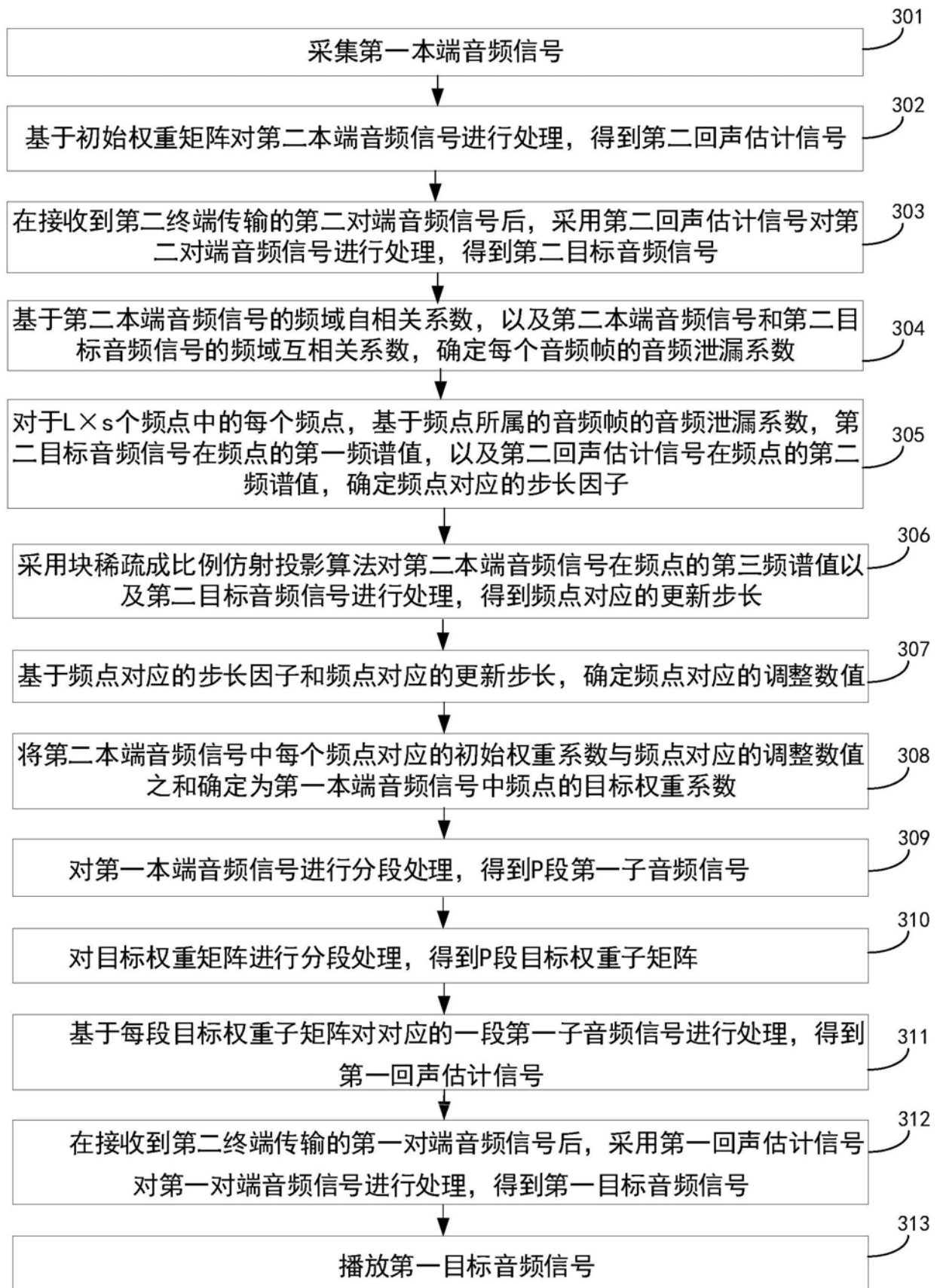


图3

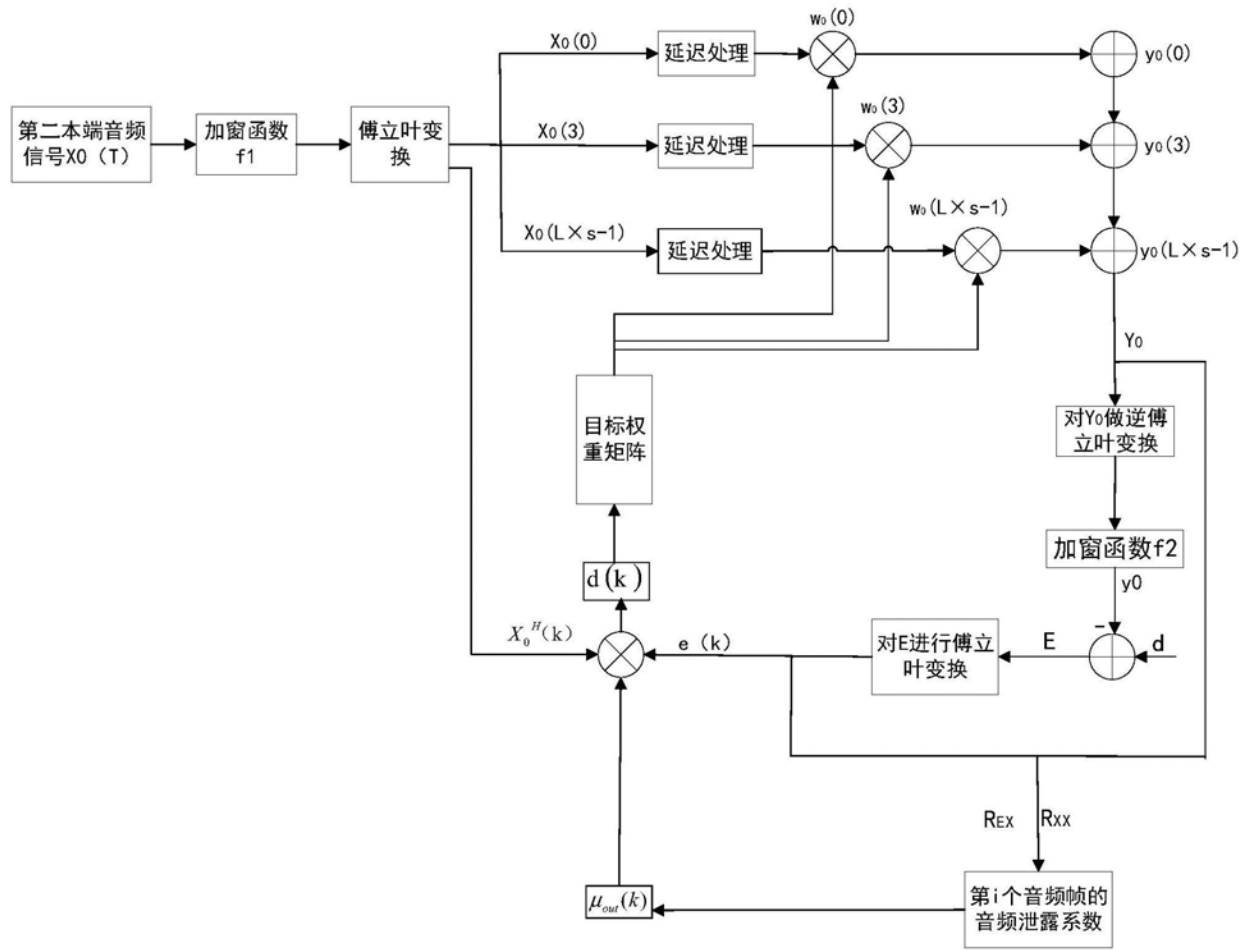


图4

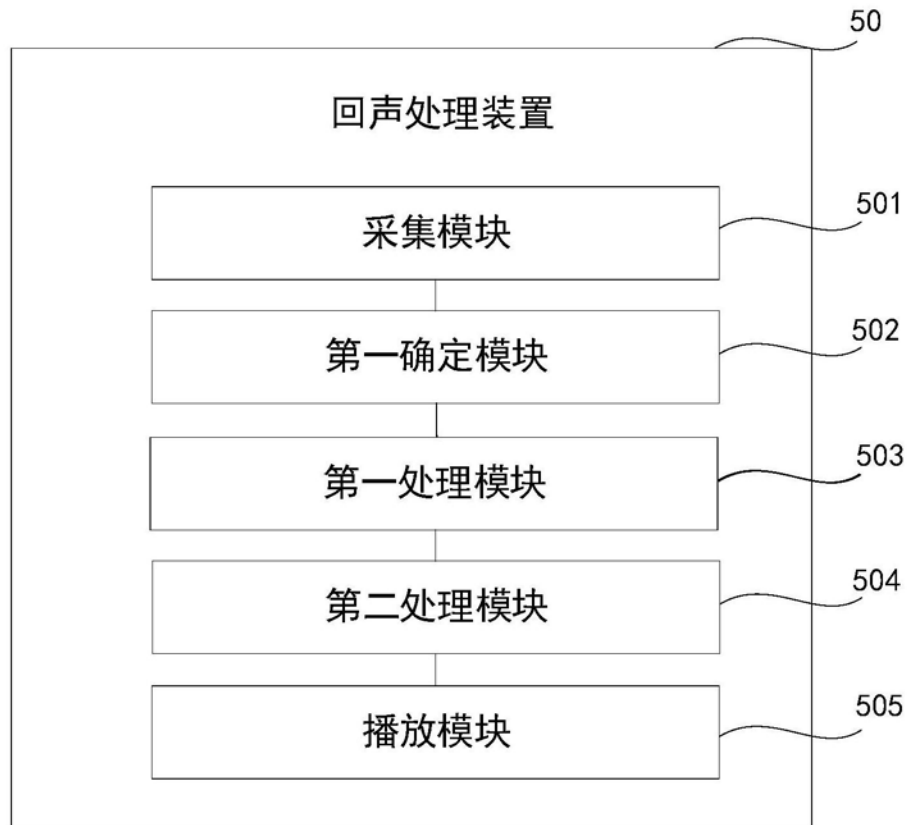


图5

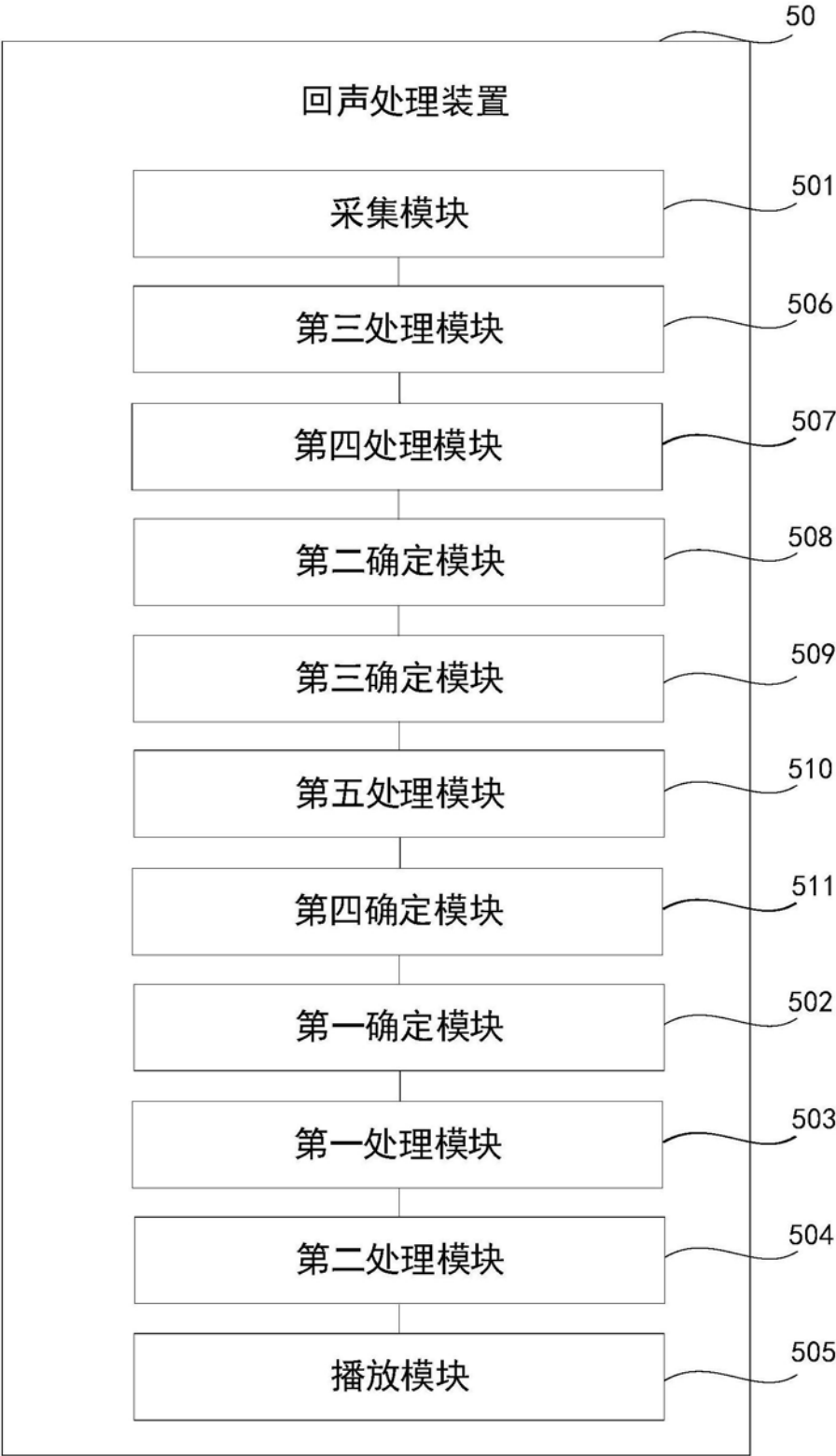


图6