

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G10L 21/02 (2006.01)

H04R 3/02 (2006.01)

H03H 21/00 (2006.01)



[12] 发明专利说明书

专利号 ZL 200610144055.5

[45] 授权公告日 2009年8月5日

[11] 授权公告号 CN 100524466C

[22] 申请日 2006.11.24

[21] 申请号 200610144055.5

[73] 专利权人 北京中星微电子有限公司

地址 100083 北京市海淀区学院路35号
世宁大厦15层

[72] 发明人 张晨

[56] 参考文献

EP0731593A2 1996.9.11

CN1388956A 2003.1.1

EP0462254A1 1991.12.27

US5680450A1 1997.10.21

审查员 邓薇

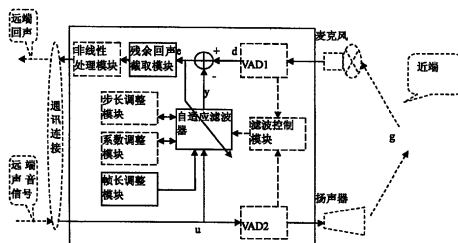
权利要求书 10 页 说明书 27 页 附图 4 页

[54] 发明名称

一种麦克风回声消除装置及回声消除方法

[57] 摘要

本发明公开了一种麦克风回声消除装置及方法，用于消除扬声器与麦克风之间的声学回路产生的回声，其装置中包括帧长调整模块，用于调整数据帧长，使当前帧以及之前的若干数据联合以组成一个长度与频域自适应滤波器系数长度相等的数据帧，以进行自适应滤波。使用本发明后，能够在保证反馈时间的前提下，使频域自适应滤波器快速收敛，达到快速消除回声的目的。可广泛应用于消除扬声器与麦克风之间的声学回路产生的噪声中。



1、一种麦克风回声消除装置，用于消除扬声器与麦克风之间的声学回路所产生的回声，包括：

数据采集合并模块，用于采集即将输出到扬声器的来自远端的语音信号 u ，每次采集的数据帧长为 M ，当前帧即第 k 帧数据记为 $\bar{u}(k)$ ，并与上一帧即第 $k-1$ 帧数据 $\bar{u}(k-1)$ 联合组成一个长度为 $2M$ 的大帧 $\bar{u}(k)$ ：

$$\bar{u}(k) = \underbrace{u(kM-M), \dots, u(kM-1)}_{\text{第 } k-1 \text{ 帧}}, \underbrace{u(kM), \dots, u(kM+M-1)}_{\text{第 } k \text{ 帧}},$$

其中 $u(kM-M)$ 为原始的第 $k-1$ 帧数据中的第 1 个数据；

$u(kM-1)$ 为原始的第 $k-1$ 帧数据中的第 M 个数据；

$u(kM)$ 为原始的第 k 帧数据中的第 1 个数据；

$u(kM+M-1)$ 为原始的第 k 帧数据中的第 M 个数据；

频域自适应滤波器，其当前的滤波器频域系数记为

$$W(k) = FFT \begin{bmatrix} w(k) \\ 0 \end{bmatrix}, \text{ 长度为 } 2M, \text{ 其中 } w(k) \text{ 为滤波器的时域系数,}$$

长度为 M ；所述频域自适应滤波器用于将所述 $\bar{u}(k)$ 进行 FFT 处理，

转换到频域，得到 $U(k) = FFT[\bar{u}(k)]$ ；并以当前的滤波器系数

$W(k)$ 对所述 $U(k)$ 进行滤波处理，然后对滤波处理结果进行 IFFT 处

理，得到回声 $\bar{d}(k)$ 的一帧预测值 $\bar{y}(k) = IFFT[U(k) * W(k)]$ ，结果取其

后 M 个点；

减法器，用麦克风采集到的长度为 M 的所述回声 $\bar{d}(k)$ 减去所述预测值 $\bar{y}(k)$ ，得到残余回声 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ；

所述频域自适应滤波器中还包括语音相关性检测单元,用于在频域中计算所述残余回声 $\bar{e}(k)$ 与所述来自远端的声音信号 $\bar{u}(k)$ 的语音相关性,得到语音相关性参数 $\bar{\phi}(k) = IFFT[U^H(k) * E(k)]$, 其中 $U^H(k)$ 为对所述 $U(k)$ 的共轭值, $E(k) = FFT \begin{bmatrix} 0 \\ \bar{e}(k) \end{bmatrix}$, 对 $\bar{\phi}(k)$ 取其结果的前 M 个点;

所述频域自适应滤波器中还包括系数更新单元,用于根据所述语音相关性,结合所述自适应滤波器的自适应步长 μ 对所述频域自适应滤波器的系数 $W(k)$ 进行更新,得到

$$W(k+1) = W(k) + \mu FFT \begin{bmatrix} \bar{\phi}(k) \\ 0 \end{bmatrix};$$

所述频域自适应滤波器每进行一次自适应滤波,其系数 $W(k)$ 即被更新一次,下一次自适应滤波时,所述自适应滤波器以更新后的系数 $W(k+1)$ 对下一个组合的大帧数据进行频域滤波;

其特征在于,所述麦克风回声消除装置还包括一个帧长调整模块,用于设置所述 u 的数据帧长 L , 其中 L 为小于 M 的值;

相应地,所述数据采集合并模块,用于将 u 当前的第 k 帧数据的 L 个数据及其之前紧临的 $2M-L$ 个连续的数据联合起来组成一个长度为 $2M$ 的大帧;

相应地,所述频域自适应滤波器对所述 $2M$ 的大帧进行自适应滤波;每帧长度为 L 的数据滤波处理完毕后,更新所述滤波器的频域滤波系数;

以及相应地,所述麦克风回声消除装置还包括一个残余回声截

取模块，用于截取所述残余回声 $\bar{e}(k)$ 每帧结果的前 L 个信号，得到最终的残余回声 e 。

2、根据权利要求 1 所述的回声消除装置，其特征在于，所述帧长调整模块将帧长由 M 调整为 $L = M/n$ ， n 为大于 1 的整数；相应地，所述数据采集合并模块将 u 的当前帧以及之前紧临的 $2n-1$ 个数据帧联合成为一个长度为 $2M$ 的大帧。

3、根据权利要求 1 或 2 所述的回声消除装置，其特征在于，还包括有声检测模块和滤波控制模块，

所述有声检测模块，包括两个有声检测单元，分别用于检测麦克风输入端及扬声器输出端的有声情况，并将检测结果输出给滤波控制模块；

所述滤波控制模块，用于根据所述有声检测模块的输出结果控制所述频域自适应滤波器的工作，

若所述麦克风输入端有声检测结果为无声，则不进行自适应滤波，也不进行系数更新，直接令输出 $\bar{e}(k) = \bar{d}(k)$ ，完成此帧处理；

若麦克风输入端检测结果为有声，那么再看扬声器输出端的检测结果，如果扬声器输出端检测无声，自适应滤波正常进行，但是不进行系数更新，输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，完成此帧处理；

如果麦克风输入端及扬声器输出端的检测结果均为有声，那么自适应滤波器处于正常工作状态，即进行自适应滤波，也进行系数更新，得到输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，以及更新后的滤波器系数 $W(k+1)$ ，完成此帧处理。

4、根据权利要求 3 所述的回声消除装置，其特征在于，所述有声检测模块通过将麦克风输入端及扬声器输出端的声音信号的短时平均幅值与噪声电平相比较来判断是否有声，具体为：

如果 $MicSignal_avg > NoiseFloor$ ，则判断麦克风线路有声，否则无声；

所述 $MicSignal_avg = \frac{1}{M} \sum_0^{M-1} |\bar{d}(k)|$ 为麦克风输入信号的短时平均幅值，其中 $\bar{d}(k)$ 为麦克风采集到的一帧长度为 M 的回声， M 为帧长， $NoiseFloor$ 是估计的噪声电平；

如果 $SpkSignal_avg > NoiseFloor$ ，则判断扬声器线路有声，否则无声；

所述 $SpkSignal_avg = \frac{1}{L} \sum_0^{L-1} |\bar{u}(k)|$ ，为输入到扬声器的信号的短时平均幅值， $\bar{u}(k)$ 为输入到扬声器的信号， L 为帧长。

5、根据权利要求 1 或 2 所述的回声消除装置，其特征在于，还包括步长调整模块，用于检测自适应滤波器的系数更新步长 μ ，并在所述 μ 大于所设定的最大系数更新步长阈值时，减小 μ 值。

6、根据权利要求 5 所述的回声消除装置，其特征在于，在检测到自适应滤波器系数的更新步长恢复正常时，将系数更新步长恢复到初始值。

7、根据权利要求 1 或 2 所述的回声消除装置，其特征在于，还包括系数调整模块，用于在检测到自适应滤波器的系数 $W(k)$ 大于所设定的系数阈值时，减小所述滤波器系数 $W(k)$ 。

8、根据权利要求 1 或 2 所述的回声消除装置，其特征在于，还包括：非线性处理模块，用于抑制回声中的非线性成分。

9、根据权利要求 8 所述的回声消除装置，其特征在于，所述非线性处理模块在 $E(e) > NLPfloor$ 时，令

$$e' = \begin{cases} e - NLPfloor, & (\text{在 } e > NLPfloor \text{ 时}); \\ e + NLPfloor, & (\text{在 } e < NLPfloor \text{ 时}); \\ 0, & (\text{其它}) \end{cases}$$

其中， e 为残差信号，也为所述非线性处理模块的输入，而所述非线性处理模块的输出为 e' ， $E(e)$ 为残差信号的短时平均幅值，

$NLPfloor$ 为判决电平。

10、 根据权利要求9所述的回声消除装置，其特征在于，当 $E(e) \leq NLPfloor$ 时， e' 直接用舒适噪声代替。

11、 根据权利要求8所述的回声消除装置，其特征在于，还包括：

扬声器有声检测模块，用于检测扬声器输出端的有声情况；

非线性处理控制模块，用于根据扬声器有声检测模块的输出结果开启或关闭所述非线性处理模块；

所述扬声器有声检测模块检测到扬声器输出端有声时，即 $SpkSignal_avg > NoiseFloor$ 时，

并且在扬声器输出端信号比残差信号大 α 倍时，即 $SpkSignal_avg / E[e] > \alpha$ 时，才启动非线性处理模块；

如果所述两个条件有一个不满足，都将关闭非线性处理模块；

其中： $SpkSignal_avg$ 为扬声器输出信号短时平均幅值， $Noisefloor$ 是估计的噪声电平， $E(e)$ 为 e 的短时平均幅值。

12、一种麦克风回声消除方法，利用频域自适应滤波方法消除来自远端的声音信号 u 经过扬声器与麦克风之间的声学回路产生的回声 d ，最终得到残余回声 e ，其时域滤波器系数为 $w(k)$ ，长度为

M ，其对应频域滤波器系数为：
$$W(k) = FFT \begin{bmatrix} w(k) \\ 0 \end{bmatrix}$$
，长度为

$2M$ ，采用重叠保留法；

其特征在于，

- 1) 设置每次采集的信号 u 的数据帧长 L ；
- 2) 以所述设置的帧长 L 采集一帧信号 $\bar{u}'(k)$ ， $\bar{u}'(k)$ 表示第 k 帧信号；

- 3) 将当前帧 $\bar{u}'(k)$ 与之前的 $2M-L$ 个数据合并成一长度为 $2M$ 的大帧 $\bar{u}(k)$;
 - 4) 将所述 $\bar{u}(k)$ 转换到频域, 采用重叠保留法, 以滤波器系数 $W(k)$ 对所述频域的 $\bar{u}(k)$ 进行滤波, 将结果转换到时域, 得到所述回声的时域的预测值 $\bar{y}(k)$;
 - 5) 采集回声 $\bar{d}(k)$ 并减去 $\bar{y}(k)$, 得到第 k 帧的最小残余回声信号 $\bar{e}(k)$;
 - 6) 根据所述 $\bar{e}(k)$ 与 $\bar{u}(k)$ 的相关性, 更新所述滤波器系数 $W(k)$, 得到 $W(k+1)$;
 - 7) 执行步骤 2), 采集下一帧信号, 对其进行数据合并后, 以所述更新的滤波器系数进行频域自适应滤波, 直至数据输入完毕。
- 13、 根据权利要求 12 所述的方法, 其特征在于,

所述频域自适应滤波算法包括如下步骤:

- 1) 帧长调整步骤, 将 u 的帧长由 M 调整为小于 M 的正整数值 L ;
- 2) 采集 u 的第 k 帧信号, 帧长为 L , 记为 $\bar{u}'(k)$;
- 3) 将所述 $\bar{u}'(k)$ 中的 L 个数据, 以及之前紧临的 $2M-L$ 个数据联合起来组成一个长度为 $2M$ 的大帧 $\bar{u}(k)$:

$$\bar{u}(k) = \underbrace{u(kL-2M+L) \cdots u(kL-2), u(kL-1)}_{2M-L \text{ 个}}, \underbrace{u(kL), \cdots, u(kL+L-1)}_{L \text{ 个}}$$

$u(kL-2M+L)$ 为原始的第 k 帧前的第 $2M-L$ 个数据,

$u(kL-2)$ 为原始的第 k 帧前的第 2 个数据,

$u(kL-1)$ 为原始的第 k 帧的前一个数据,

$u(kL)$ 为原始的第 k 帧中的第 1 个数据,

$u(kL+L-1)$ 为原始的第 k 帧中的第 L 个数据;

- 4) 将 $\bar{u}(k)$ 做 FFT 处理, 转换到频域得到: $U(k) = FFT[\bar{u}(k)]$;

- 5) 采用重叠保留法, 以当前的滤波器系数 $W(k)$ 对所述 $U(k)$ 进行滤波, 即频域上的相乘, 然后对其结果进行 IFFT 处理后取其结果的

后 M 个数据，记为 $\bar{y}(k)$ ，即有：

$$\bar{y}(k) = [y(kM), y(kM+1), \dots, y(kM+M-1)] = \text{IFFT}[U(k) * W(k)] ;$$

6) 所述 u 被扬声器播放后，经过扬声器与麦克风之间的声学回路，然后被麦克风采集到长度为 M 的回声信号以 $\bar{d}(k)$ 表示，即：

$$\bar{d}(k) = [d(kM), d(kM+1), \dots, d(kM+M-1)] ,$$

所述 $\bar{d}(k)$ 与所述 $\bar{y}(k)$ 相减后得到最小残余回声信号 $\bar{e}(k)$ 为：

$$\bar{e}(k) = [e(kM), e(kM+1), \dots, e(kM+M-1)] = \bar{d}(k) - \bar{y}(k) ;$$

7) 截取所述 $\bar{e}(k)$ 结果的 L 个信号作为最终的残余回声输出；

8) 在所述长度为 M 的未经截取的 $\bar{e}(k)$ 前补 M 个零，并进行 FFT 处

理，得到：
$$E(k) = \text{FFT} \begin{bmatrix} 0 \\ \bar{e}(k) \end{bmatrix} ;$$

同时对步骤 4) 所述 $U(k)$ 取共轭得到 $U^H(k)$ ，然后与所述 $E(k)$ 进行点乘，其结果进行 IFFT 运算，同时根据重叠保留法，得到：

$$\bar{\phi}(k) = \text{IFFT}[U^H(k) * E(k)] ,$$

在上式中，需删除结果的最后一帧，只取 IFFT 结果的前 M 个点；

9) 在所述 $\bar{\phi}(k)$ 后补 M 个零，然后进行 FFT 处理，其结果与自适应步长 μ 相乘，得到的乘积与滤波器系数 $W(k)$ 相加，可得到滤波器

系数频域形式的更新值如下：
$$W(k+1) = W(k) + \mu \text{FFT} \begin{bmatrix} \bar{\phi}(k) \\ 0 \end{bmatrix} ,$$
 下一

次自适应滤波即采用该更新后的滤波器系数 $W(k+1)$ 进行滤波；

10) 执行所述频域自适应滤波算法中的步骤 2)，直至来自远端的声音信号输入完毕。

14、 根据权利要求 12 或 13 所述的方法，其特征在于，所述 L 值为 M/n ，n 为大于 1 的整数。

15、 根据权利要求 12 所述的方法，其特征在于，在所述步骤

1) 之前还包括有声检测步骤和滤波控制步骤，包括：

有声检测步骤，检测麦克风输入端及扬声器输出端的有声情况；

滤波控制步骤，根据所述有声检测步骤的结果控制滤波器的工作；

具体为：

若麦克风输入端检测结果为无声，则不进行自适应滤波，也不进行系数更新，直接令输出 $\bar{e}(k) = \bar{d}(k)$ ，完成此帧处理；

若麦克风输入端检测结果为有声，那么再看扬声器输出端的检测结果，如果扬声器输出端检测无声，则自适应滤波正常进行，但是不进行系数更新，输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，完成此帧处理；

如果麦克风输入端及扬声器输出端的检测结果均为有声，则自适应滤波器处于正常工作状态，既进行自适应滤波，也进行系数更新，输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，完成此帧处理；

其中， $\bar{d}(k)$ 为麦克风收到的回声， $\bar{y}(k)$ 为所述自适应滤波器输出的对 $\bar{d}(k)$ 的预测值， $\bar{e}(k)$ 为残余回声。

16、 根据权利要求 15 所述的方法，其特征在于，所述有声检测是通过将麦克风输入端及扬声器输出端的声音信号的短时平均幅值与噪声电平相比较来判断是否有声，具体为：

如果 $MicSignal_avg > NoiseFloor$ ，则判断麦克风线路有声，否则无声；

所述 $MicSignal_avg = \frac{1}{M} \sum_0^{M-1} |\bar{d}(k)|$ ，为麦克风输入信号的短时平均幅值， $\bar{d}(k)$ 为麦克风输入信号，即收到的回声信号，M 为一帧语音信号的长度， $NoiseFloor$ 为估计的噪声电平；

如果 $SpkSignal_avg > NoiseFloor$ ，则判断扬声器线路有声，否则无声；

所述 $SpkSignal_avg = \frac{1}{L} \sum_0^{L-1} |\bar{u}(k)|$ ，为扬声器输出信号短时平均

幅值， $\bar{u}(k)$ 为扬声器输出信号， L 为一帧语音信号的长度。

17、 根据权利要求 12 或 13 所述的方法，其特征在于，还包括步
长调整步骤，用于在检测到自适应滤波器的系数更新步长大于所设
定的最大系数更新步长阈值时，减小自适应滤波器的系数更新步
长。

18、 根据权利要求 17 所述的方法，其特征在于，在检测到自适
应滤波器系数的更新步长恢复正常时，将系数更新步长恢复到初始
值。

19、 根据权利要求 12 或 13 所述的方法，其特征在于，还包括系
数调整步骤，用于在检测到自适应滤波器的系数大于所设定的系数
阈值时，减小滤波器的系数。

20、 根据权利要求 12 所述的方法，其特征在于，在所述步骤
5) 之后还包括非线性处理步骤：

首先计算最小化残差信号的短时平均幅值 $E(e)$ ；

然后判断 $E(e)$ 是否大于事先设定的非线性处理阈值
 $NLPfloor$ ，如果判断结果为是，则使用下式计算最小化残余噪声
 $e'(n)$ ：

$$e' = \begin{cases} e - NLPfloor, & (\text{在 } e > NLPfloor \text{ 时}); \\ e + NLPfloor, & (\text{在 } e < NLPfloor \text{ 时}); \\ 0, & (\text{其它}) \end{cases}$$

其中， e 为残差信号，为所述非线性处理模块的输入，而非线
性处理模块的输出为 e' ， $E(e)$ 为残差信号的短时平均幅值，
 $NLPfloor$ 为判决电平。

21、 根据权利要求 20 所述的方法，其特征在于，如果
 $E(e) \leq NLPfloor$ ， e' 直接用舒适噪声代替。

22、 根据权利要求 20 所述的方法，其特征在于，在非线形处理

步骤之前，还包括非线性处理开关控制步骤，具体为：

检测扬声器输出端的有声情况；

根据所述检测结果开启或关闭所述非线性处理步骤，具体为：

当检测到扬声器输出端有声，即 $SpkSignal_avg > NoiseFloor$ ，
并且扬声器输出端信号比残差信号大 α 倍时，即
 $SpkSignal_avg / E[e] > \alpha$ ，才启动非线性处理模块；

如果所述两个条件有一个不满足，都将关闭非线性处理模块；

其中： $SpkSignal_avg$ 为扬声器输出信号短时平均幅值， $NoiseFloor$ 是估计的噪声电平， $E(e)$ 为残差信号的短时平均幅值， α 为预设的倍数值。

一种麦克风回声消除装置及回声消除方法

技术领域

本发明涉及回声消除领域，特别涉及一种利用频域自适应滤波器的麦克风回声消除装置和方法，用于消除扬声器与麦克风之间的声学回路产生的回声。

背景技术

回声的产生是由于扬声器和麦克风之间存在声学回路。如图1所示，来自远端的声音信号，经过通讯连接，到达近端记为信号 u ，通过近端的扬声器放出，经过扬声器与麦克风之间的声学回路 g 后被麦克风采集到参考信号 d ，然后经过通讯连接再传回远端。此时，远端的说话人就能听到自己的回声，即远端回声。从而会严重地影响通话质量。

由于从扬声器到麦克风的声学回路 g 是未知并且时变的，因而在回声消除方案中，广泛采用自适应滤波的方法。如图1所示，为利用自适应滤波方式进行回声消除的基本原理图。其中自适应滤波器以最小化残余回声 e 为目标，通过自适应地调整自身的滤波器系数对来自远端的声音信号 u 进行滤波处理，来跟踪扬声器到麦克风的声学反馈回路 g ，产生麦克风接收到的回声 d 的预测值 y 。当滤波器准确的跟踪到 g 时， y 非常接近 d ，使 $e = d - y$ 趋向于 0，从而实现消除回声的作用。

在自适应滤波过程中，自适应滤波器需要跟踪未知的反馈回路，也就是说要模拟一个未知装置。当未知的反馈回路 g 有较大延时的时候，即该未知装置有较高的阶数，那么自适应滤波器如果要得到最佳的模拟效果，也至少需要有相同的阶数才够用。由于时域的自适应滤波的过程就是输入信号与自适应滤波器卷积的过程，因而随着自适应滤波器阶数的增加，算法的复杂度会急剧增加，当反馈回路延时较大时，并不实用。用子带自适应滤波可以降低运算复

杂度，不过会带来信号混叠的问题。

而时域的卷积等于频域的相乘，并且通过借助 FFT 的快速算法，频域的自适应滤波算法在当滤波器阶数较高时，可以减少算法复杂度，提高运算效率，是一种非常实用的滤波方式。

现有技术中的频域自适应滤波算法一般如下所述。

首先对下文中用到的一些信号标记做出说明。在进行频域自适应滤波时，对输入信号是以帧为单位进行处理的，在本文中以“ $\bar{x}(k)$ ”的形式表示信号 x 的当前帧信号，即第 k 帧信号。如以 $\bar{u}'(k)$ 表示第 k 帧来自远端的并且即将输出给扬声器的声音信号，以 $\bar{u}(k)$ 表示经过合并长度为 $2M$ 的声音信号，而 $\bar{d}(k)$ 表示麦克风采集到的第 k 帧回声信号等等。另外，用 $w(k)$ 表示处理 $\bar{u}(k)$ 时的时域滤波器系数，其相应频域滤波器系数用 $W(k)$ 表示。FFT 表示快速傅立叶变换，IFFT 表示快速傅立叶反变换。

应用频域自适应滤波器的回声消除装置中一般包括以下各部分。

(1) 数据采集合并模块，用于采集即将输出到扬声器的来自远端的声音信号 u ，每次采集的数据帧长为 M ，当前帧即第 k 帧数据记为 $u'(k)$ ，并与上一帧即第 $k-1$ 帧数据 $u'(k-1)$ 联合组成一个长度为 $2M$ 的大帧 $\bar{u}(k)$ ：

$$\bar{u}(k) = \underbrace{u(kM - M), \dots, u(kM - 1)}_{\text{第 } k-1 \text{ 帧}}, \underbrace{u(kM), \dots, u(kM + M - 1)}_{\text{第 } k \text{ 帧}}$$

。

(2) 频域自适应滤波器，假设我们采用自适应滤波器的阶数是 M ，滤波器的时域系数记为 $w(k)$ ，采用重叠保留法，为避免混叠，将 M 阶的滤波器扩展 M 个 0，组成 $N=2M$ 个系数的滤波器，经

FFT 处理后得到滤波器的频域系数为： $W(k) = FFT \begin{bmatrix} w(k) \\ 0 \end{bmatrix}$ ，长度

为 $2M$ 。

所述频域自适应滤波器用于将所述 $\bar{u}(k)$ 进行 FFT 处理，转换到频域，得到 $U(k) = FFT[\bar{u}(k)]$ ；并以当前的滤波器系数 $W(k)$ 对所述 $U(k)$ 进行滤波处理，然后对滤波处理结果进行 IFFT 处理，得到所述回声 $\bar{d}(k)$ 的一帧预测值 $\bar{y}(k) = IFFT[U(k) * W(k)]$ ，结果取其

后 M 个点。

(3) 减法器，用麦克风采集到的所述回声 $\bar{d}(k)$ 减去所述预测值 $\bar{y}(k)$ ，得到残余回声 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，所采集的 $\bar{d}(k)$ 长度也为 M 。

(4) 所述频域自适应滤波器中还包括语音相关性检测单元，用于在频域中计算所述残余回声 $\bar{e}(k)$ 与所述来自远端的声音信号 $\bar{u}(k)$ 的语音相关性，得到语音相关性参数

$\bar{\phi}(k) = IFFT[U^H(k) * E(k)]$ ，其中 $U^H(k)$ 为对所述 $U(k)$ 的共轭值，

$E(k) = FFT \begin{bmatrix} 0 \\ \bar{e}(k) \end{bmatrix}$ ，对 $\bar{\phi}(k)$ 取其结果的前 M 个点。

(5) 所述频域自适应滤波器中还包括系数更新单元，用于根据所述语音相关性，结合所述自适应滤波器的自适应步长 μ 对所述频域自适应滤波器的系数 $W(k)$ 进行更新，得到

$W(k+1) = W(k) + \mu FFT \begin{bmatrix} \bar{\phi}(k) \\ 0 \end{bmatrix}$ 。

所述频域自适应滤波器每进行一次自适应滤波，其系数 $W(k)$

即被更新一次，下一次自适应滤波时，所述自适应滤波器以更新后的系数 $W(k+1)$ 作为当前的 $W(k)$ 对下一个组合的大帧数据进行频域滤波。

如图 2 所示，即为现有技术中利用频域自适应滤波方法进行回声消除的方法示意图，其中细箭头代表时域信号处理，粗箭头代表频域信号处理。因为采用频域自适应滤波方法，信号要分帧处理，所以图 1 中所示的 u 、 y 、 d 和 e 信号，在图 2 中分别对应着 $\bar{u}(k)$ 、 $\bar{y}(k)$ 、 $\bar{d}(k)$ 和 $\bar{e}(k)$ ，分别表示各个信号的第 k 帧信号；另外，以 $\bar{u}(k)$ 表示将 $u'(k)$ 的数据合并后得到的长度为 $2M$ 的大帧。我们知道，长序列截短后分块处理再合并，需要采用重叠相加法或者重叠保留法避免混叠，在此以重叠保留法说明。

首先假设时域自适应滤波器的阶数是 M ，其系数记为 $w(k)$ ，因为采用重叠保留法，为避免混叠，将 M 阶的滤波器扩展 M 个 0，经 FFT 处理后得到滤波器的频域系数向量为：

$$W(k) = FFT \begin{bmatrix} w(k) \\ 0 \end{bmatrix} \quad (1.1)$$

从上式 (1.1) 可以看出，频域自适应滤波器系数 $W(k)$ 的长度 N 是时域系数向量长度 M 的 2 倍。对于频域自适应滤波算法，自适应滤波和滤波器系数更新都是在频域中完成的，所以将不出现时域滤波器的形式。值得注意的是以后我们提到的 FFT 或者 IFFT 处理长度都为 N 点。

频域自适应滤波处理的步骤如下：

- 1) 采集一帧来自远端的声音信号 $u'(k)$ ，其帧长为 M 。
- 2) 对输入信号 $u'(k)$ 进行连接两帧的处理，即将 $u'(k)$ 与上一帧第 $k-1$ 帧的数据合并为一个大的帧，得到下式：

$$\bar{u}(k) = \underbrace{u(kM - M), \dots, u(kM - 1)}_{\text{上一帧(记为k-1)}} \underbrace{u(kM), \dots, u(kM + M - 1)}_{\text{当前帧(记为k)}} \quad (1.2)$$

其中 $\bar{u}(k)$ 为第 k 个合并后的大帧，长度为 $N=2M$ ；

$u(kM - M)$ 为原始的第 $k-1$ 帧数据中的第 1 个数据；

$u(kM - 1)$ 为原始的第 $k-1$ 帧数据中的第 M 个数据；

$u(kM)$ 为原始的第 k 帧数据中的第 1 个数据；

$u(kM + M - 1)$ 为原始的第 k 帧数据中的第 M 个数据。

3) 将 $\bar{u}(k)$ 做 FFT 处理，转换到频域得到：

$$U(k) = FFT[\bar{u}(k)] \quad (1.3)$$

4) 对输入信号进行滤波，即频域上的相乘，然后进行 IFFT 处理，转换到时域，并取其结果的最后一帧，即后 M 个数据，即有回声信号的预测值：

$$\bar{y}(k) = [y(kM), y(kM + 1), \dots, y(kM + M - 1)] = IFFT[U(k) * W(k)] \quad (1.4)$$

5) 采集回声信号用 $\bar{d}(k)$ 表示，即：

$$\bar{d}(k) = [d(kM), d(kM + 1), \dots, d(kM + M - 1)] \quad (1.5)$$

则残余回声信号为回声信号与其预测值的差值：

$$\begin{aligned} \bar{e}(k) &= [e(kM), e(kM + 1), \dots, e(kM + M - 1)] \\ &= \bar{d}(k) - \bar{y}(k) \end{aligned} \quad (1.6)$$

6) 在所述残余回声信号前补 M 个零，并进行 FFT 处理，得到频域的残余回声信号为：

$$E(k) = FFT \begin{bmatrix} 0 \\ \bar{e}(k) \end{bmatrix} \quad (1.7)$$

利用 $E(k)$ 和 $U(k)$ 来计算自适应滤波器系数的更新量。首先将 $U(k)$ 取共轭得到 $U^H(k)$ 。在频域中，自适应滤波器系数矢量的更新量是通过计算误差信号和输入信号的相关性确定的，由于线性相关

从形式上看相当于一个逆的线性卷积，所以，借助于时域的卷积在频域上有FFT的快速算法，有：

$$\bar{\phi}(k) = \text{IFFT}[U^H(k) * E(k)] \quad (1.8)$$

根据重叠保留法，在上式中，需删除结果的最后一帧，即只取IFFT结果的前M个点。

7) 最后我们利用 $\bar{\phi}(k)$ 来更新自适应滤波器系数。注意：频域的滤波器系数是将时域系数后面补零，然后进行FFT处理生成的，所以在此相应地，将 $\bar{\phi}(k)$ 后补M个零，然后进行FFT处理，其结果与自适应步长 μ 相乘，得到的乘积与更新前的滤波器系数 $W(k)$ 相加，可得到滤波器系数更新的频域形式如下：

$$W(k+1) = W(k) + \mu \text{FFT} \begin{bmatrix} \phi(k) \\ 0 \end{bmatrix}$$

(1.9)

下一次自适应滤波即采用所述 $W(k+1)$ 作为更新后的当前滤波器系数 $W(k)$ 进行滤波。

8) 步骤1)至7)循环进行，直至数据处理完毕。

由上述频域自适应滤波算法的步骤可以看出，每过帧长为M的一帧信号，才更新一次频域自适应滤波器的滤波系数，所以收敛速度较慢，尤其是当反馈回路特性变化较快时，效果并不理想。

发明内容

为了解决上述现有技术的缺陷，本发明提供一种回声消除装置及回声消除方法，使得频域自适应滤波器的系数能够高效而稳定的工作，从而达到有效消除回声的目的。

为解决上述问题，本发明提供了一种麦克风回声消除装置，用于消除扬声器与麦克风之间的声学回路所产生的回声，包括：

数据采集合并模块，用于采集即将输出到扬声器的来自远端的声音信号 u ，每次采集的数据帧长为M，当前帧即第k帧数据记

为 $\bar{u}'(k)$ ，并与上一帧即第 $k-1$ 帧数据 $\bar{u}'(k-1)$ 联合组成一个长度为 $2M$ 的大帧 $\bar{u}(k)$ ：

$$\bar{u}(k) = \underbrace{u(kM - M), \dots, u(kM - 1)}_{\text{第 } k-1 \text{ 帧}}, \underbrace{u(kM), \dots, u(kM + M - 1)}_{\text{第 } k \text{ 帧}} ;$$

频域自适应滤波器，其当前的滤波器频域系数记为

$$W(k) = FFT \begin{bmatrix} w(k) \\ 0 \end{bmatrix} , \text{ 长度为 } 2M, \text{ 其中 } w(k) \text{ 为滤波器的时域系数,}$$

长度为 M ；所述频域自适应滤波器用于将所述 $\bar{u}(k)$ 进行 FFT 处理，转换到频域，得到 $U(k) = FFT[\bar{u}(k)]$ ；并以当前的滤波器系数 $W(k)$ 对所述 $U(k)$ 进行滤波处理，然后对滤波处理结果进行 IFFT 处理，得到所述回声 $\bar{d}(k)$ 的一帧预测值 $\bar{y}(k) = IFFT[U(k) * W(k)]$ ，结果取其后 M 个点；

减法器，用麦克风采集到的长度为 M 的所述回声 $\bar{d}(k)$ 减去所述预测值 $\bar{y}(k)$ ，得到残余回声 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ；

所述频域自适应滤波器中还包括语音相关性检测单元，用于在频域中计算所述残余回声 $\bar{e}(k)$ 与所述来自远端的声音信号 $\bar{u}(k)$ 的语音相关性，得到语音相关性参数 $\bar{\phi}(k) = IFFT[U^H(k) * E(k)]$ ，其中 $U^H(k)$ 为对所述 $U(k)$ 的共轭值， $E(k) = FFT \begin{bmatrix} 0 \\ \bar{e}(k) \end{bmatrix}$ ，对 $\bar{\phi}(k)$ 取其结果的前 M 个点；

所述频域自适应滤波器中还包括系数更新单元，用于根据所述语音相关性，结合所述自适应滤波器的自适应步长 μ 对所述频

域自适应滤波器的系数 $W(k)$ 进行更新, 得到

$$W(k+1) = W(k) + \mu FFT \begin{bmatrix} \tilde{\phi}(k) \\ 0 \end{bmatrix} ;$$

所述频域自适应滤波器每进行一次自适应滤波, 其系数 $W(k)$ 即被更新一次, 下一次自适应滤波时, 所述自适应滤波器以更新后的系数 $W(k+1)$ 对下一个组合的大帧数据进行频域滤波;

还包括一个帧长调整模块, 用于设置所述 u 的数据帧长为小于 M 的值 L ;

相应地, 所述数据采集合并模块, 用于将 u 当前的第 k 帧数据的 L 个数据及其之前紧临的 $2M-L$ 个连续的数据联合起来组成一个长度为 $2M$ 的大帧;

相应地, 所述频域自适应滤波器对所述 $2M$ 的大帧进行自适应滤波; 每帧长度为 L 的数据滤波处理完毕后, 更新所述滤波器的频域滤波系数;

以及相应地, 还包括一个残余回声截取模块, 用于截取所述残余回声 $\bar{e}(k)$ 每帧结果的前 L 个信号, 得到最终的残余回声 e 。

优选的, 所述帧长调整模块将帧长由 M 调整为 $L = M/n$, n 为大于 1 的整数; 相应地, 所述数据采集合并模块将 u 的当前帧以及之前紧临的 $2n-1$ 个数据帧联合成为一个长度为 $2M$ 的大帧。

优选的, 还包括有声检测模块和滤波控制模块,

所述有声检测模块, 包括两个有声检测单元, 分别用于检测麦克风输入端及扬声器输出端的有声情况, 并将检测结果输出给滤波控制模块;

所述滤波控制模块, 用于根据所述有声检测模块的输出结果控制所述频域自适应滤波器的工作,

若所述麦克风输入端有声检测结果为无声，则不进行自适应滤波，也不进行系数更新，直接令输出 $\bar{e}(k) = \bar{d}(k)$ ，完成此帧处理；

若麦克风输入端检测结果为有声，那么再看扬声器输出端的检测结果，如果扬声器输出端检测无声，自适应滤波正常进行，但是不进行系数更新，输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，完成此帧处理；

如果麦克风输入端及扬声器输出端的检测结果均为有声，那么自适应滤波器处于正常工作状态，即进行自适应滤波，也进行系数更新，得到输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，以及更新后的滤波器系数 $W(k+1)$ ，完成此帧处理。

优选的，所述有声检测模块通过将麦克风输入端及扬声器输出端的语音信号的短时平均幅值与噪声电平相比较来判断是否有声，具体为：

如果 $MicSignal_avg > NoiseFloor$ ，则判断麦克风线路有声，否则无声；

所述 $MicSignal_avg = \frac{1}{M} \sum_0^{M-1} |\bar{d}(k)|$ 为麦克风输入信号的短时平均幅值，其中 $\bar{d}(k)$ 为麦克风采集到的一帧长度为 M 的声音信号， M 为帧长， $NoiseFloor$ 是估计的噪声电平；

如果 $SpkSignal_avg > NoiseFloor$ ，则判断扬声器线路有声，否则无声；

所述 $SpkSignal_avg = \frac{1}{L} \sum_0^{L-1} |\bar{u}(k)|$ ，为输入到扬声器的信号的短时平均幅值， $\bar{u}(k)$ 为输入到扬声器的信号， L 为帧长。

优选的，还包括步长调整模块，用于检测自适应滤波器的系数更新步长 μ ，并在所述 μ 大于设定最大系数更新步长阈值时，减小 μ 值。

优选的，在检测到自适应滤波器系数的更新步长恢复正常时，将系数更新步长恢复到初始值。

优选的，还包括系数调整模块，用于在检测到自适应滤波器的系数 $W(k)$ 大于设定系数阈值时，减小所述滤波器系数 $W(k)$ 。

优选的，还包括：非线性处理模块，用于抑制回声中的非线性成分。

优选的，所述非线性处理模块在 $E(e) > NLPfloor$ 时，令

$$e' = \begin{cases} e - NLPfloor, & (\text{在 } e > NLPfloor \text{ 时}); \\ e + NLPfloor, & (\text{在 } e < NLPfloor \text{ 时}); \\ 0, & (\text{其它}) \end{cases}$$

其中， e 为残差信号，也为所述非线性处理模块的输入，而所述非线性处理模块的输出为 e' ， $E(e)$ 为残差信号的短时平均幅值， $NLPfloor$ 为判决电平。

优选的，当 $E(e) \leq NLPfloor$ 时， e' 直接用舒适噪声代替。

优选的，还包括：

扬声器有声检测模块，用于检测扬声器输出端的有声情况；

非线性处理控制模块，用于根据扬声器有声检测模块的输出结果开启或关闭所述非线性处理模块；

所述扬声器有声检测模块检测到扬声器输出端有声时，即 $SpkSignal_avg > NoiseFloor$ 时，

并且在扬声器输出端信号比残差信号大 α 倍时，即 $SpkSignal_avg / E[e] > \alpha$ 时，才启动非线性处理模块；

如果所述两个条件有一个不满足，都将关闭 NLP 处理；

其中： $SpkSignal_avg$ 为扬声器输出信号短时平均幅值， $Noisefloor$ 是估计的噪声电平， $E(e)$ 为 e 的短时平均幅值。

本发明还提供了一种麦克风回声消除方法，其利用频域自适应滤波方法消除来自远端的声音信号 u 经过扬声器与麦克风之间的声学回路产生的回声 d ，最终得到残余回声 e ，其时域滤波器系数为 $w(k)$ ，长度为 M ，其对应频域滤波器系数为：

$$W(k) = FFT \begin{bmatrix} w(k) \\ 0 \end{bmatrix} \quad , \text{ 长度为 } 2M, \text{ 采用重叠保留法; 其包括以下步骤,}$$

骤，

- 1) 设置每次采集的信号 u 的数据帧长 L ;
- 2) 以所述设置的帧长 L 采集一帧信号 $\bar{u}'(k)$ ， $\bar{u}'(k)$ 表示第 k 帧信号;
- 3) 将当前帧 $\bar{u}'(k)$ 与之前的 $2M-L$ 个数据合并成一长度为 $2M$ 的大帧 $\bar{u}(k)$;
- 4) 将所述 $\bar{u}(k)$ 转换到频域，采用重叠保留法，以滤波器系数 $W(k)$ 对所述频域的 $\bar{u}(k)$ 进行滤波，将结果转换到时域，得到所述回声的时域的预测值 $\bar{y}(k)$;
- 5) 采集回声 $\bar{d}(k)$ 并减去 $\bar{y}(k)$ ，得到第 k 帧的最小残余回声信号 $\bar{e}(k)$;
- 6) 根据所述 $\bar{e}(k)$ 与 $\bar{u}(k)$ 的相关性，更新所述滤波器系数 $W(k)$ ，得到 $W(k+1)$;
- 7) 执行步骤 2)，采集下一帧信号，对其进行数据合并后，以所述更新的滤波器系数进行频域自适应滤波，直至数据输入完毕。

优选的，所述频域自适应滤波算法包括如下步骤：

- 1) 帧长调整步骤，将 u 的帧长由 M 调整为小于 M 的正整数值 L ;
- 2) 采集 u 的第 k 帧信号，帧长为 L ，记为 $\bar{u}'(k)$;
- 3) 将所述 $\bar{u}'(k)$ 中的 L 个数据，以及之前紧临的 $2M-L$ 个数据联合

起来组成一个长度为 $2M$ 的大帧 $\bar{u}(k)$:

$$\bar{u}(k) = \underbrace{u(kL-2M+L) \cdots u(kL-2), u(kL-1), u(kL)}_{2M-L \text{ 个}}, \underbrace{\cdots, u(kL+L-1)}_{L \text{ 个}}$$

$u(kL-2M+L)$ 为原始的第 k 帧前的第 $2M-L$ 个数据,

$u(kL-2)$ 为原始的第 k 帧前的第 2 个数据,

$u(kL-1)$ 为原始的第 k 帧的前一个数据,

$u(kL)$ 为原始的第 k 帧中的第 1 个数据,

$u(kL+L-1)$ 为原始的第 k 帧中的第 L 个数据;

4) 将 $\bar{u}(k)$ 做 FFT 处理, 转换到频域得到: $U(k) = FFT[\bar{u}(k)]$;

5) 采用重叠保留法, 以当前的滤波器系数 $W(k)$ 对所述 $U(k)$ 进行滤波, 即频域上的相乘, 然后对其结果进行 IFFT 处理后取其结果的后 M 个数据, 记为 $\bar{y}(k)$, 即有:

$$\bar{y}(k) = [y(kM), y(kM+1), \dots, y(kM+M-1)] = IFFT[U(k) * W(k)] \quad ;$$

6) 所述 u 被扬声器播放后, 经过扬声器与麦克风之间的声学回路, 然后被麦克风采集到长度为 M 的回声信号以 $\bar{d}(k)$ 表示, 即:

$$\bar{d}(k) = [d(kM), d(kM+1), \dots, d(kM+M-1)] \quad ,$$

所述 $\bar{d}(k)$ 与步骤 5) 中所述 $\bar{y}(k)$ 相减后得到误差信号 $\bar{e}(k)$ 为:

$$\begin{aligned} \bar{e}(k) &= [e(kM), e(kM+1), \dots, e(kM+M-1)] \quad , \\ &= \bar{d}(k) - \bar{y}(k) \end{aligned}$$

7) 截取所述 $\bar{e}(k)$ 结果的 L 个信号作为最终的残余回声输出;

8) 在所述长度为 M 的未经截取的 $\bar{e}(k)$ 前补 M 个零, 并进行 FFT

处理, 得到: $E(k) = FFT \begin{bmatrix} 0 \\ \bar{e}(k) \end{bmatrix}$;

同时对步骤 4) 所述 $U(k)$ 取共轭得到 $U^H(k)$, 然后与所述 $E(k)$ 进行点乘, 其结果进行 IFFT 运算, 同时根据重叠保留法, 得到:

$$\bar{\phi}(k) = IFFT[U^H(k) * E(k)] \quad ,$$

在上式中, 需删除结果的后一帧, 只取 IFFT 结果的前 M 个点;

9) 在所述 $\bar{\phi}(k)$ 后补 M 个零, 然后进行 FFT 处理, 其结果与自适应步长 μ 相乘, 得到的乘积与滤波器系数 $W(k)$ 相加, 可得到滤波

器系数频域形式的更新值如下：
$$W(k+1) = W(k) + \mu FFT \begin{bmatrix} \bar{\phi}(k) \\ 0 \end{bmatrix}$$
，下

一次自适应滤波即采用该更新后的滤波器系数 $W(k+1)$ 进行滤波；

10) 执行步骤2)，直至来自远端的声音信号输入完毕。

优选的，所述 L 值为 M/n ， n 为大于1的整数。

优选的，在所述步骤1)之前还包括有声检测步骤和滤波控制步骤，包括：

有声检测步骤，检测麦克风输入端及扬声器输出端的有声情况；

滤波控制步骤，根据所述有声检测步骤的结果控制滤波器的工作；

具体为：

若麦克风输入端检测结果为无声，则不进行自适应滤波，也不进行系数更新，直接令输出 $\bar{e}(k) = \bar{d}(k)$ ，完成此帧处理；

若麦克风输入端检测结果为有声，那么再看扬声器输出端的检测结果，如果扬声器输出端检测无声，则自适应滤波正常进行，但是不进行系数更新，输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，完成此帧处理；

如果麦克风输入端及扬声器输出端的检测结果均为有声，则自适应滤波器处于正常工作状态，既进行自适应滤波，也进行系数更新，输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，完成此帧处理；

其中， $\bar{d}(k)$ 为麦克风收到的回声， $\bar{y}(k)$ 为所述自适应滤波器输出的对 $\bar{d}(k)$ 的预测值， $\bar{e}(k)$ 为残余回声。

优选的，所述有声检测是通过将麦克风输入端及扬声器输出端的声音信号的短时平均幅值与噪声电平相比较来判断是否有声，具体为：

如果 $MicSignal_avg > NoiseFloor$ ，则判断麦克风线路有声，否则

无声;

所述 $MicSignal_avg = \frac{1}{M} \sum_0^{M-1} |\bar{d}(k)|$, 为麦克风输入信号的短时平均幅值, $\bar{d}(k)$ 为麦克风输入信号, 即收到的回声信号, M 为一帧语音信号的长度, $NoiseFloor$ 为估计的噪声电平;

如果 $SpkSignal_avg > NoiseFloor$, 则判断扬声器线路有声, 否则无声;

所述 $SpkSignal_avg = \frac{1}{L} \sum_0^{L-1} |\bar{u}(k)|$, 为扬声器输出信号短时平均幅值, $\bar{u}(k)$ 为扬声器输出信号, L 为一帧语音信号的长度。

优选的, 还包括步长调整步骤, 用于在检测到自适应滤波器的系数更新步长大于设定最大系数更新步长阈值时, 减小自适应滤波器的系数更新步长。

优选的, 在检测到自适应滤波器系数的更新步长恢复正常时, 将系数更新步长恢复到初始值。

优选的, 还包括系数调整步骤, 用于在检测到自适应滤波器的系数大于设定系数阈值时, 减小滤波器的系数。

优选的, 还包括非线性处理步骤:

首先计算最小化残差信号的短时平均幅值 $E(e)$;

然后判断 $E(e)$ 是否大于事先设定的非线性处理阈值 $NLPfloor$, 如果判断结果为是, 则使用下式计算最小化残余噪声 $e'(n)$:

$$e' = \begin{cases} e - NLPfloor, & (\text{在 } e > NLPfloor \text{ 时}); \\ e + NLPfloor, & (\text{在 } e < NLPfloor \text{ 时}); \\ 0, & (\text{其它}) \end{cases}$$

其中, e 为残差信号, 为所述非线性处理模块的输入, 而非线

性处理模块的输出为 e' ， $E(e)$ 为残差信号的短时平均幅值， $NLPfloor$ 为判决电平。

优选的，如果 $E(e) \leq NLPfloor$ ， e' 直接用舒适噪声代替。

优选的，还包括非线性处理开关控制步骤，具体为：

检测扬声器输出端的有声情况；

根据所述检测结果开启或关闭所述非线性处理步骤，具体为：

当检测到扬声器输出端有声，即 $SpkSignal_avg > NoiseFloor$ ，并且扬声器输出端信号比残差信号大 α 倍时，即 $SpkSignal_avg / E[e] > \alpha$ ，才启动非线性处理模块；

如果所述两个条件有一个不满足，都将关闭 NLP 处理；

其中： $SpkSignal_avg$ 为扬声器输出信号短时平均幅值， $NoiseFloor$ 是估计的噪声电平， $E(e)$ 为残差信号的短时平均幅值， α 为预设的倍数。

本发明在自适应频域滤波器中所增加的帧长调整模块，令一次处理的来自远端的声音信号的帧长小于自适应滤波器的时域系数长度，然后将多于一帧的信号联合成为一个大帧，进行自适应滤波。一方面这使得自适应滤波器的长度保持原有的足够长度，可以满足反馈回路的延时要求；另一方面，提高了自适应滤波器系数的更新频率，使自适应滤波器能够高效地工作。另外，本发明公开的滤波控制模块可以使自适应滤波器在麦克风输入线路或者扬声器输出线路无声的特殊情况下也不会错误收敛，保证了其正常工作；而步长调整模块和系数调整模块使得自适应滤波器在发散的情况下能够恢复正常工作状态；非线性处理模块可以消除反馈回路中的非线性失真。因此，利用本发明的回声消除装置，使得自适应滤波器能够高效而稳定地工作，从而达到有效消除回声的目的。

附图说明

图 1 为利用自适应滤波方式进行回声消除的装置结构基本原理图；

图 2 为现有技术中利用频域自适应滤波方式进行回声消除的方法示意图;

图 3 为本发明装置中语音检测模块及滤波控制模块的结构示意图;

图 4 为本发明数据合并单元的示意图;

图 5 为本发明非线性处理模块非线性处理前后的回声与判决电平关系示意图。

具体实施方式

下面结合附图详细说明本发明的回声消除装置和方法。

自适应滤波器为了能有效地跟踪反馈回路，其系数长度必须大于反馈延时的采样点数。例如对于 8K 采样率的信号，如果时域自适应滤波器系数长度 $M=1024$ ，那么滤波器能跟踪和模拟的反馈回路的最大反馈延时就为： $1024/8000=128\text{ms}$ 。

在背景技术中所述的频域自适应滤波方法中，频域滤波器系数的长度是 $2M$ ，对应时域系数长度为 M ，每次新来数据帧长也为 M 。也就是说自适应滤波器时域系数长度与新来的数据帧长是一样的，即可以令自适应滤波器系数长度为 1024 ，那么一次处理的数据帧长也为 1024 。这样的话，一秒种只进行大约 8 次滤波和系数更新。对于反馈回路变化较快的环境，这种更新频率有时是不够用的。

因此，如图 3 所示，本发明在频域自适应滤波的基础上，增加了帧长调整模块，用于调整数据帧的长度为 L 。注意，一次调整之后，帧长是相对固定的，而不是每采集一帧数据都进行帧长调整。比如：频域滤波器系数的长度是 $2M$ ，对应时域滤波器系数长度为 M ，每次新来数据帧长 L 可以为滤波器时域系数长度的一半，即 $L=M/2$ (M 为偶数)。则对输入信号 $\bar{u}(k)$ 的处理也由原来的两帧联合变成了四帧联合。经过这样的改进，一方面，自适应滤波器的长度仍为 $2M$ ，足够长，可以满足反馈回路的延时要求；另一方面，每过 $M/2$ 的帧长即更新一次自适应滤波器系数，也兼顾了自适应滤波器系数的更新频率。不过采用这种方式付出的代价是增加了算法的

复杂度。因为每帧数据量为 L ，所以最后在输出残余回声时，要加入一残余回声截取模块，用于截取所得残余回声的前 L 个数据作为最终结果输出。

上面的例子中 $L=M/2$ ，在实际使用中，也可以为 $M/3$ 、 $M/4$ 、 $M/8$ 等情况，这样可以使自适应滤波器的系数更新频率更高。只是相应地，需要改变残余回声截取模块所截取的数据长度。

除此之外，每个数据帧的长度 L 还可以是小于 M 的任意数，例如： M 为 1024，则 L 可以为 1000，900，650 等任意小于 1024 的值。只是在数据帧联合时要保证联合后的大帧长度为 $2M$ 。这个问题可以通过如下方式解决：如图 4 所示，使用一个长度为 $2M$ 的先进先出缓存器来存储输入数据，每接收完一帧新的数据 $\bar{u}'(k)$ ，就将 $\bar{u}'(k)$ 与之前的 $2M-L$ 个数据联合成为一个大帧 $\bar{u}(k)$ 来进行一次自适应滤波处理。

自适应滤波可以自动跟踪反馈回路，不过对于一些特殊情况，自适应滤波器容易错误跟踪，比如麦克风和扬声器线路同时无声的情况。在这种情况下，自适应滤波器的输入信号和参考信号都很小，这时，自适应滤波器就容易错误收敛。

为了防止滤波器错误收敛，如图 3 所示，本发明提出在回声消除装置中可以加入有声检测模块和滤波控制模块。

有声检测模块，即 VAD (Voice Activity Detector) 模块，可以包括位于麦克风输入端和扬声器输出端的两个有声检测单元 VAD1 和 VAD2。VAD 检测可以通过将信号短时平均幅值与噪声电平比较来进行判决。信号的短时平均幅值可以通过计算一帧信号的平均幅值得到。

$$\text{对于麦克风输入端: } \text{MicSignal_avg} = \frac{1}{M} \sum_{k=0}^{M-1} |\bar{d}(k)|$$

(2.1)

式中： MicSignal_avg 为麦克风输入信号短时平均幅值， $\bar{d}(k)$ 为麦克风输入信号， M 为一帧语音信号的长度。

如果 $MicSignal_avg > NoiseFloor$ ，则判决麦克风线路有声，否则无声。其中， $NoiseFloor$ 是估计的噪声电平。

$$\text{同理，对于扬声器输出端： } SpkSignal_avg = \frac{1}{L} \sum_{k=0}^{L-1} |\bar{u}'(k)|$$

(2.2)

式中： $SpkSignal_avg$ 为扬声器输出信号短时平均幅值， $\bar{u}'(k)$ 为输入给扬声器的声音信号， L 为一帧语音信号的长度。

如果 $SpkSignal_avg > NoiseFloor$ ，则判决扬声器线路有声，否则无声。

根据上述有声检测单元的输出结果，滤波控制模块对滤波器的工作进行整体控制，具体为：

如果 VAD1 检测无声，则不进行自适应滤波，也不进行滤波器系数更新，直接令输出 $\bar{e}(k) = \bar{d}(k)$ ，完成此帧处理；如果 VAD1 检测有声，那么再看 VAD2 检测结果，如果 VAD2 检测无声，自适应滤波正常进行，但是不进行滤波器系数更新，输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ 完成此帧处理；如果 VAD1 和 VAD2 都检测有声，那么自适应滤波器处于正常工作状态，即进行自适应滤波，也进行滤波器系数更新，输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ 完成此帧处理。

实验表明，加入滤波控制后，自适应滤波器在麦克风输入线路或者扬声器输出线路无声的特殊情况下也不会错误收敛，保证了其正常工作。

另外，对于自适应滤波而言，如果麦克风采集到的参考信号 $\bar{d}(k)$ 完全由扬声器发出的声音产生，那么自适应滤波器就容易跟踪反馈回路，也能稳定的工作。但是麦克风采集的信号一般来说不仅包含扬声器发出的声音，还包括近端的声音信号，并且这种声音信号有时还占主要成分。因此这种信号将干扰自适应滤波器正确地跟踪反馈回路，从而有可能导致自适应滤波器的错误跟踪甚至系数发散。

滤波器错误跟踪，系数开始发散，表现在系数更新上时，往往

在这个时候自适应滤波器的系数更新量比较大。因此，如图3所示本发明可以增加一个步长调整模块，当检测到系数更新量比较大时，就判定此时自适应滤波器处于非正常工作状态，将其系数更新步长减小，从而可以有效抑制滤波器的错误跟踪，避免了系数发散。等到检测到系数更新量恢复正常时，就判定此时自适应滤波器恢复了正常工作状态，于是可以将其系数更新步长进行调整，如恢复到初始值。这样能够加快自适应滤波器的收敛速度。

具体来说，对于频域自适应算法中的NLMS算法，

如前面所述，系数更新如下式所示：

$$W(k+1) = W(k) + \mu FFT \begin{bmatrix} \phi(k) \\ 0 \end{bmatrix} \quad (2.3)$$

$$\text{令 } \Phi(k) = FFT \begin{bmatrix} \phi(k) \\ 0 \end{bmatrix}$$

(2.4)

$$\text{则, } W(k+1) = W(k) + \mu \cdot \Phi(k) \quad (2.5)$$

式中 $W(k)$ 为频域自适应滤波器系数，是一个 N 维复数矢量， μ 为系数更新步长， $\Phi(k)$ 也是一个 N 维复数矢量， N 为 FFT 点数。即：

$$\Phi(k) = [\Phi_0(k), \Phi_1(k), \dots, \Phi_{N-1}(k)]^T$$

(2.6)

由此可得系数更新量为：

$$\mu \cdot \Phi(k) = [\mu \cdot \Phi_0(k), \mu \cdot \Phi_1(k), \dots, \mu \cdot \Phi_{N-1}(k)]^T \quad (2.7)$$

前面提到步长调整的关键是检测系数更新量的大小。系数更新量的大小，可以用复数的模来量度。即：

$$[\mu \cdot \|\Phi_0(k)\|, \mu \cdot \|\Phi_1(k)\|, \dots, \mu \cdot \|\Phi_{N-1}(k)\|]^T \quad (2.8)$$

在本发明中，步长调整的方法可以为：

对于 $\mu \cdot \|\Phi_i(k)\|$ ， $i=0, 1, \dots, N-1$ ，

若 $\mu \cdot \|\Phi_i(k)\| > \text{MaxStepSize}$ ， MaxStepSize 为最大步长阈值，则判定此时自适应滤波器处于非正常工作状态，于是调整步长，所述调整可以是按步长比例缩小，如缩小10倍。即 $\mu = 0.1\mu$ 。

试验表明,加入步长调整模块后,虽然频域自适应滤波器的收敛速度在一定程度上有所减慢,但是系数不易发散,自适应滤波器稳定性大大增强。

上述滤波控制模块、步长调整模块在一定程度上已经保证了自适应滤波器的稳定工作。然而,一些突发事件,或者意料之外的情况仍然有可能导致自适应滤波器的发散,发散的滤波器会使扬声器发出很大的噪声。所以本发明提出了一种应对特殊情况的策略,如图3所示,可以加入系数调整模块,作为确保自适应滤波器稳定工作的最后一道防线。

系数调整模块工作的原理很简单,即当自适应滤波器发散时,其系数往往比较大,于是系数调整的任务就是在每次系数更新后,检查一下系数的大小,如果大于设定的阈值,就认为滤波器发散了。具体来说,对于频域NLMS算法,如前所述,系数更新如下式所示:

$$W(k+1) = W(k) + \mu FFT \begin{bmatrix} \phi(k) \\ 0 \end{bmatrix} \quad (2.9)$$

式中 $W(k)$ 为频域自适应滤波器系数,是一个 N 维复数矢量, N 为 FFT 点数。即: $W(k) = [W_0(k), W_1(k), \dots, W_{N-1}(k)]^T$

$$(2.10)$$

系数的大小,用复数的模来量度。即:

$$[\|W_0(k)\|, \|W_1(k)\|, \dots, \|W_{N-1}(k)\|]^T \quad (2.11)$$

对于 $\|W_i(k)\|$, $i=0, 1, \dots, N-1$,

若 $\|W_i(k)\| > MaxParam$, 其中 $MaxParam$ 为最大系数阈值,则判定此时频域自适应滤波器已发散,于是调整自适应滤波器的系数,所述调整可以是将自适应滤波器系数减小,如可以置零,即:

$W(k) = 0$ 。系数置零后,自适应滤波器将重新开始收敛,这样就可以把滤波器从发散状态抢救回来。所述阈值 $MaxParam$ 需要根据反馈回路的增益仔细选择,其值太大,系数监控就不灵敏,不能有效地识别发散状态;其值太小,容易误判,导致自适应滤波器频繁

重启，无法正常工作。

另外，还可以增加一非线性处理模块，即 NLP (Non-Linear Processor) 模块。这是因为一般的扬声器都有5%-10%的非线性失真，而自适应滤波只能跟踪线性系统，因而反馈回路中信号的非线性失真是无法预测和消除的。因此在自适应滤波后可以增加 NLP 处理模块以消除非线性失真。

因为 NLP 处理只是针对扬声器的非线性失真而进行的，所以可以在不需要使用时关闭该模块，这就需要增加一非线性处理控制模块和扬声器有声检测模块，用于对非线性处理模块的开启与关闭进行控制，其中扬声器有声检测模块，可以使用前文中所述有声检测模块中的VAD2。

具体控制原理为：当 (1) $SpkSignal_avg > NoiseFloor$ ，即VAD2检测到扬声器有声时；且 (2) $SpkSignal_avg / E[e] > \alpha$ ，即扬声器信号比残差信号大 α 倍时；才启动 NLP 处理。如果 (1)、(2) 有任何一个条件不满足，NLP 模块都将关闭。

其中条件 (1) 说明，当扬声器无声时，不可能产生回声，也就不必使用 NLP 处理；条件 (2) 说明，当近端有声时， $E[e]$ 会比较大，使得条件 (2) 不成立，从而关闭 NLP 处理，让近端信号无失真的传输出去。

式中： $SpkSignal_avg$ 为扬声器输出信号短时平均幅值， $NoiseFloor$ 是估计的噪声电平， $E[e]$ 为残差信号的短时平均幅值， α 值在本实施例中可以取 2。所述短时平均幅值可以为为一帧信号中各信号绝对值和的均值。

本方案中的 NLP 处理可以采用中心削波的方法来抑制残余回声。如图 5 所示，为中心削波的 NLP 处理示意图。其作用可以用下式表示：当 $E[e] > NLPfloor$ 时，

$$e' = \begin{cases} e - NLPfloor, & \text{if } (e > NLPfloor) \\ e + NLPfloor, & \text{elseif } (e < -NLPfloor) \\ 0, & \text{else} \end{cases} \quad (3.1)$$

式中， e 和 e' 为经过 NLP 模块前后的残余回声。 $E[e]$ 为短时平

均幅值，NLPfloor为判决电平，其值需要仔细挑选，太小不足以有效抑制残余回声，太大会严重影响近端声音质量。

另外，当 $E[e] \leq \text{NLPfloor}$ 时， e' 可以用舒适噪声代替。之所以用舒适噪声代替 e' ，是因为如果将 e' 直接置零，则在NLP启动和关闭切换时，会引入噪声，同时会给人一种半双工的错觉。舒适噪声可以用模拟高斯随机信号产生。

下面说明本发明利用频域自适应滤波方式进行麦克风回声消除的方法。

首先说明下文中用到的一些基本概念，频域滤波器系数为：

$$W(k) = \text{FFT} \begin{bmatrix} w(k) \\ 0 \end{bmatrix}, \text{ 长度为 } 2M, \text{ 其中 } w(k) \text{ 为对应的长度为 } M \text{ 的}$$

时域自适应滤波器系数，采用重叠保留法。

在背景技术中所述回声消除的方法基础上，本发明提出加入帧长调整步骤，其作用是调整数据帧的长度。首先详细说明该步骤，

在本发明中，是将其长度调整为小于 M 的任意正整数 L 。比如：频域滤波器系数的长度是 $2M$ ，对应时域滤波器系数长度为 M ，可以将每次新来数据帧的长度调整为时域滤波器系数长度的一半，即 $L=M/2$ （ M 为偶数）。则与背景技术中相比，对输入信号 $\bar{u}'(k)$ 的处理也由原来的两帧联合变成了四帧联合。经过这样的改进，一方面，自适应滤波器的长度够长，可以满足反馈回路的延时要求；另一方面，也兼顾了自适应滤波器系数的更新频率。

上面的例子中 $L=M/2$ ，在实际使用中，也可以为 $M/4$ 、 $M/3$ 、 $M/8$ 等情况，这样可以使自适应滤波器的系数更新频率更高。只是相应地，需要改变残余回声截取模块所截取的数据长度。在实际使用中，数据帧的长度 L 可以是小于 M 的任意数，例如： M 为1024，则 L 可以为1000，900，650等任意小于1024的值。不过采用这种方式付出的代价是增加了算法的复杂度。注意，一次调整之后，直

至所有数据处理完毕，帧长是相对固定的，而不是每采集一帧数据都进行帧长调整。最后，因为每帧数据量为 L ，所以最后在输出残余回声时，要加入一残余回声截取步骤，用于截取所得残余回声的前 L 个或 L 个数据作为最终结果输出。

下面以 $M=1024$ 为例完整说明本发明加入了上述帧长调整步骤及残余回声截取步骤的一种利用频域自适应滤波方式的麦克风回声消除方法，包括以下步骤。

- 1) 帧长调整步骤，将帧长调整为小于 M 的正整数值 L ；在本实施例中，令 L 为 800。
- 2) 采集一帧即将输出到扬声器的第 k 帧远端声音信号 $\bar{u}'(k)$ ，帧长度为 800。
- 3) 将当前帧 $\bar{u}'(k)$ 中的 800 个数据，以及之前的 $2M-L=2048-800=1248$ 个数据联合起来组成一个长度为 $2M$ 的大帧 $\bar{u}(k)$ 。如图 4 所示，新采集的当前帧 $\bar{u}'(k)$ 的 800 个数据与之前的 1248 个数据组成了一个长度为 2048 的大帧 $\bar{u}(k)$ 。

$$\bar{u}(k) = \underbrace{u(800k-1248) \cdots u(800k-2)}_{1248\text{个}}, \underbrace{u(800k-1), u(800k), \cdots, u(800k+799)}_{800\text{个}}$$

$u(800k-1248)$ 为原始的第 k 帧前的第 1248 个数据，

$u(800k-2)$ 为原始的第 k 帧前的第 2 个数据，

$u(800k-1)$ 为原始的第 k 帧前的前一个数据，

$u(800k)$ 为原始的第 k 帧数据中的第 1 个数据，

$u(800k+799)$ 为原始的第 k 帧数据中的第 800 个数据。

最初采集到第一、二帧信号时，可以不做处理，等待第三帧数据到来后再与第一帧数据中的后 448 个数据以及第二帧中的 800 个数据联合组成一个长度为 2048 的大帧 $\bar{u}(k)$ 进行一次自适应滤波处理。之后的数据都是每来一帧新数据 $\bar{u}'(k)$ ，即进行数据合并然后进行一次自适应滤波处理。

- 4) 将 $\bar{u}(k)$ 做 FFT 处理，转换到频域得到：

$$U(k) = FFT[\bar{u}(k)]。$$

5) 采用重叠保留法, 以当前的滤波器系数 $W(k)$ 对所述 $U(k)$ 进行滤波, 即频域上的相乘, 然后对其结果进行IFFT处理后取其结果的后 M 个数据, 即后1024个数据, 记为 $\bar{y}(k)$, 即有:

$$\bar{y}(k) = [y(kM), y(kM+1), \dots, y(kM+M-1)] = \text{IFFT}[U(k) * W(k)]。$$

6) 所述远端声音信号 $\bar{u}'(k)$ 被扬声器播放后, 经过扬声器与麦克风之间的声学回路, 然后被麦克风采集到长度为 M 的回声信号以 $\bar{d}(k)$ 表示, 即:

$$\bar{d}(k) = [d(kM), d(kM+1), \dots, d(kM+M-1)]，$$

所述 $\bar{d}(k)$ 与步骤 5) 中所述 $\bar{y}(k)$ 相减后得到误差信号 $\bar{e}(k)$ 为:

$$\begin{aligned} \bar{e}(k) &= [e(kM), e(kM+1), \dots, e(kM+M-1)]； \\ &= \bar{d}(k) - \bar{y}(k) \end{aligned}$$

7) 截取所述 $\bar{e}(k)$ 结果的前 L 个信号作为最终的残余回声输出;

8) 在所述长度为 M 的未经截取的 $\bar{e}(k)$ 前补 M 个零, 并进行FFT处理, 得到:

$$E(k) = \text{FFT} \begin{bmatrix} 0 \\ \bar{e}(k) \end{bmatrix}；$$

;

同时对步骤 4) 所述 $U(k)$ 取共轭得到 $U^H(k)$, 然后与所述 $E(k)$ 进行点乘, 其结果进行IFFT运算, 同时根据重叠保留法, 得到:

$$\bar{\phi}(k) = \text{IFFT}[U^H(k) * E(k)]，$$

在上式中, 需删除结果的最后一帧, 只取IFFT结果的前 M 个点;

9) 在 $\bar{\phi}(k)$ 后补 M 个零, 然后进行FFT处理, 其结果与自适应步长 μ 相乘, 得到的乘积与滤波器系数 $W(k)$ 相加, 可得到滤波器系数

频域形式的更新值如下: $W(k+1) = W(k) + \mu \text{FFT} \begin{bmatrix} \bar{\phi}(k) \\ 0 \end{bmatrix}$, 下一次自

适应滤波即采用该更新后的滤波器系数 $W(k+1)$ 作为当前的 $W(k)$ 进行滤波;

10) 执行步骤 2), 直至来自远端的声音信号输入完毕, 则整个过程结束。

在上面所述的具体实施例中，所述L取值为800，在实际使用中，也可以为小于M的其它整数值，如600、500等。另外，L值可以为M/n，即 $1024/n$ ，n为大于1的整数，且 $1024/n$ 也为整数。如可以为 $1024/2$ ，则在数据合并时只需联合4个数据帧即可得到长度为2048的一个大帧。在这种情况下，每过 $1024/2$ 个数据即可更新滤波器系数一次，加快了滤波器系数的收敛速度，提高了效率。

在所述步骤1)之前还可以包括有声检测步骤和滤波控制步骤，用于对滤波器的工作进行整体的控制，包括：

有声检测步骤，检测麦克风输入端及扬声器输出端的有声情况；

滤波控制步骤，根据所述有声检测步骤的结果控制滤波器的工作，具体为：

若麦克风输入端检测结果为无声，则不进行自适应滤波，也不进行系数更新，直接令输出 $\bar{e}(k) = \bar{d}(k)$ ，完成此帧处理；

若麦克风输入端检测结果为有声，那么再看扬声器输出端的检测结果，如果扬声器输出端检测无声，则自适应滤波正常进行，但是不进行系数更新，输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，完成此帧处理；

如果麦克风输入端及扬声器输出端的检测结果均为有声，则自适应滤波器处于正常工作状态，既进行自适应滤波，也进行系数更新，输出 $\bar{e}(k) = \bar{d}(k) - \bar{y}(k)$ ，完成此帧处理。

其中， $\bar{d}(k)$ 为麦克风收到的回声， $\bar{y}(k)$ 为所述自适应滤波器输出的对 $\bar{d}(k)$ 的预测值， $\bar{e}(k)$ 为残余回声。

所述有声检测是通过将麦克风输入端及扬声器输出端的语音信号的短时平均幅值与噪声电平相比较来判断是否有声，具体为：

如果 $MicSignal_avg > NoiseFloor$ ，则判断麦克风线路有声，否则无声；

所述 $MicSignal_avg = \frac{1}{M} \sum_0^{M-1} |\bar{d}(k)|$ ，为麦克风输入信号的短时平均幅值， $\bar{d}(k)$ 为麦克风输入信号，即收到的回声信号， M 为一帧语音信号的长度， $NoiseFloor$ 为估计的噪声电平；

如果 $SpkSignal_avg > NoiseFloor$ ，则判断扬声器线路有声，否则无声；

所述 $SpkSignal_avg = \frac{1}{L} \sum_0^{L-1} |\bar{u}(k)|$ ，为扬声器输出信号短时平均幅值， $\bar{u}(k)$ 为扬声器输出信号， L 为一帧语音信号的长度。

还包括步长调整步骤，用于在检测到自适应滤波器的系数更新步长大于设定最大系数更新步长阈值时，减小自适应滤波器的系数更新步长。减小自适应滤波器的系数更新步长，可以为使系数更新步长以一定比例缩小。

在检测到自适应滤波器系数的更新步长恢复正常时，将系数更新步长恢复到初始值。

另外，还包括系数调整步骤，用于在检测到自适应滤波器的系数大于设定系数阈值时，减小滤波器的系数。以有效地防止滤波器系数发散。

还包括非线性处理步骤：首先计算最小化残差信号的短时平均幅值 $E(e)$ ；然后判断 $E(e)$ 是否大于事先设定的非线性处理阈值 $NLPfloor$ ，如果判断结果为是，则使用下式计算最小化残余噪声 $e'(n)$ ：

$$e' = \begin{cases} e - NLPfloor, & (\text{在 } e > NLPfloor \text{ 时}); \\ e + NLPfloor, & (\text{在 } e < NLPfloor \text{ 时}); \\ 0, & (\text{其它}) \end{cases}$$

其中， e 为残差信号，为所述非线性处理模块的输入，而非线性处理模块的输出为 e' ， $E(e)$ 为残差信号的短时平均幅值，

NLP_{floor} 为判决电平。

如果 $E(e) \leq NLP_{floor}$, e' 直接用舒适噪声代替。

在所述步骤 7) 后还可以包括非线性处理开关控制步骤, 具体为: 检测扬声器输出端的有声情况; 根据所述检测结果开启或关闭所述非线性处理步骤。

开启或关闭的方法具体为: 当检测到扬声器输出端有声, 即 $SpkSignal_avg > NoiseFloor$, 并且扬声器输出端信号比残差信号大 α 倍时, 即 $SpkSignal_avg / E[e] > \alpha$, 如 α 为 6 时才启动非线性处理模块;

如果所述两个条件有一个不满足, 都将关闭 NLP 处理;
其中: $SpkSignal_avg$ 为扬声器输出信号短时平均幅值, $NoiseFloor$ 是估计的噪声电平, $E(e)$ 为残差信号的短时平均幅值。

使用本发明的技术方案, 可以使频域滤波器进行高效稳定的工作, 经实验得到的具体性能指标为:

回声压缩: 50-60dB;

收敛时间: <1s;

支持的反馈回路延迟时间: 可调, 如在采样率为 8K, 滤波器长度 1024 时, 可支持 128ms 延迟。

以上所述仅为本发明的较佳实施例而已, 并不用以限制本发明, 凡在本发明的精神和原则之内, 所作的任何修改、等同替换等, 均应包含在本发明的保护范围之内。

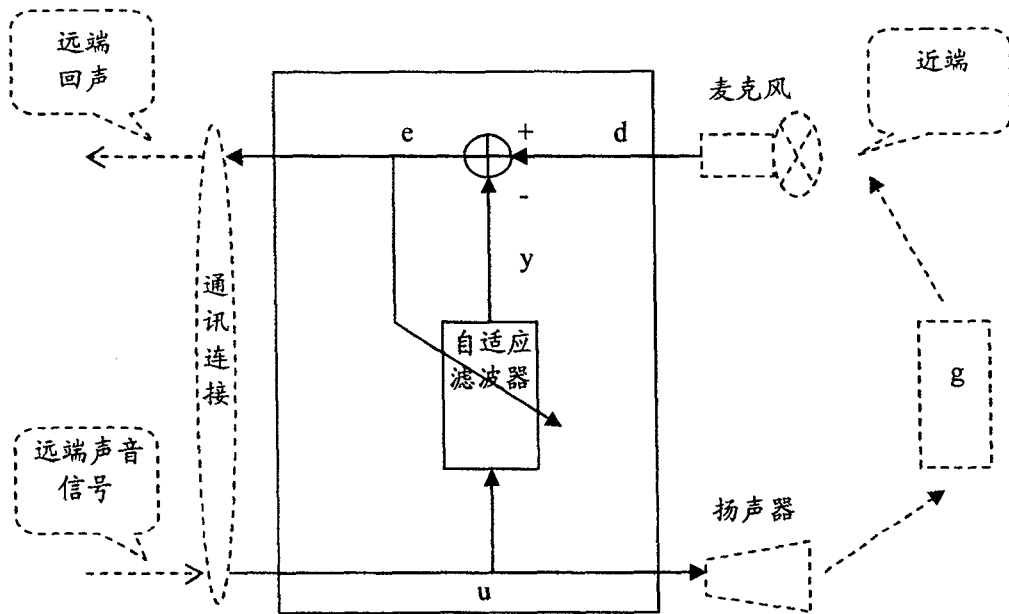


图 1

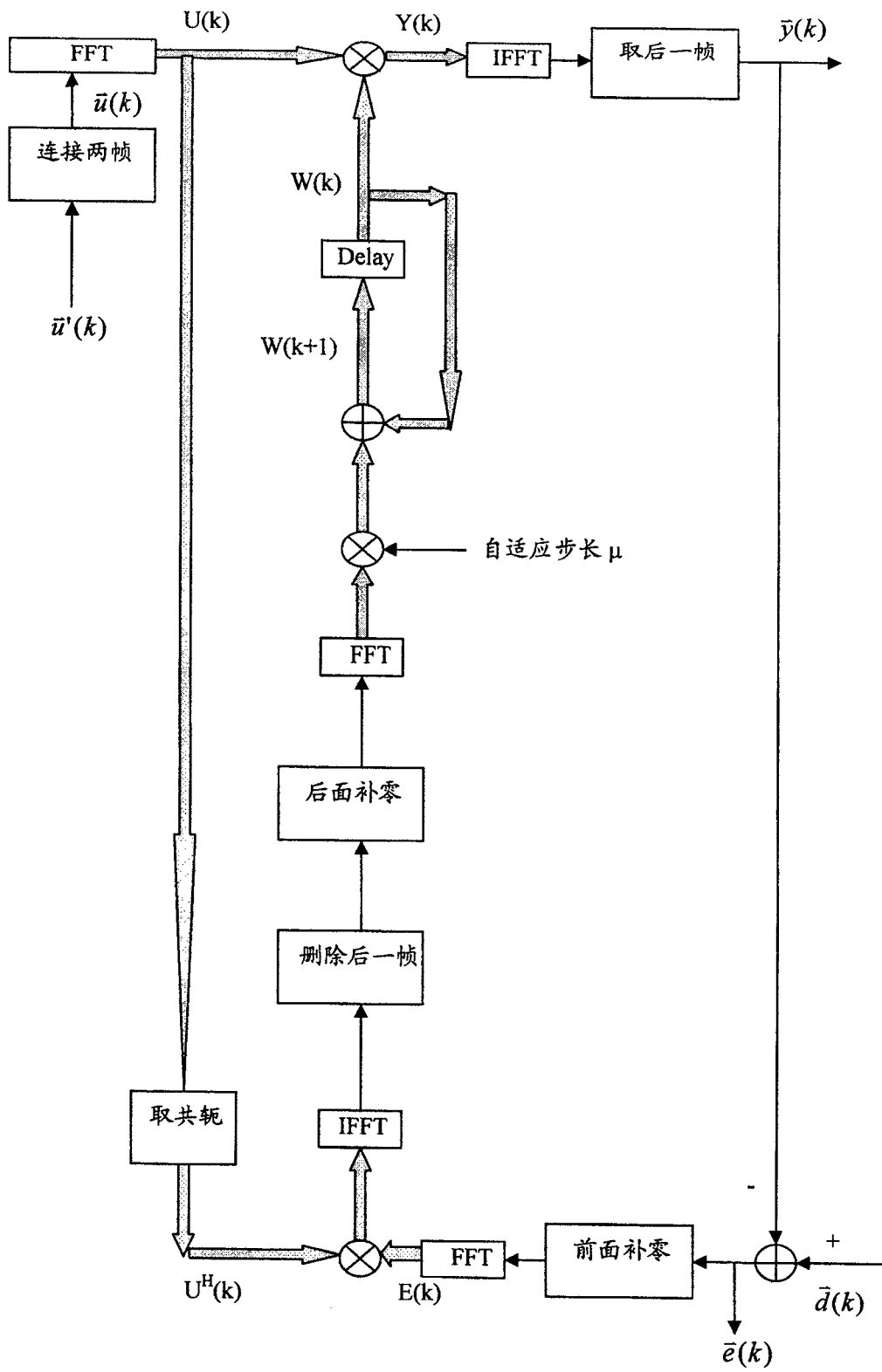


图 2

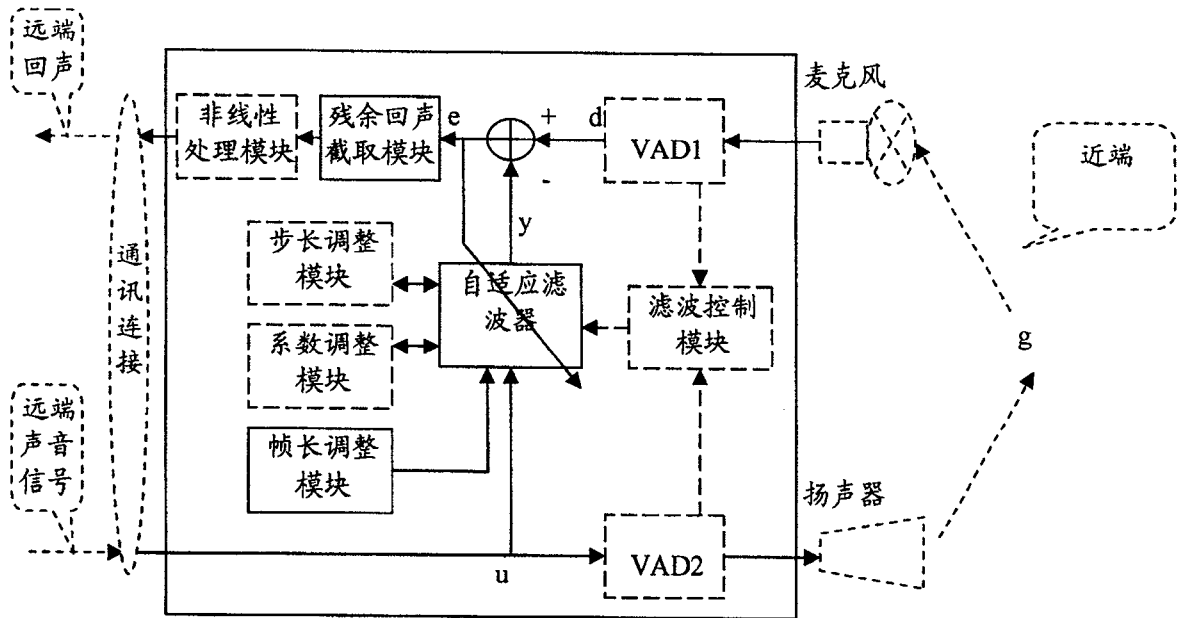


图 3

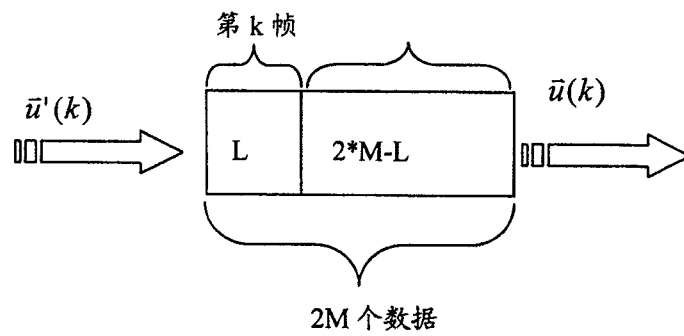


图 4

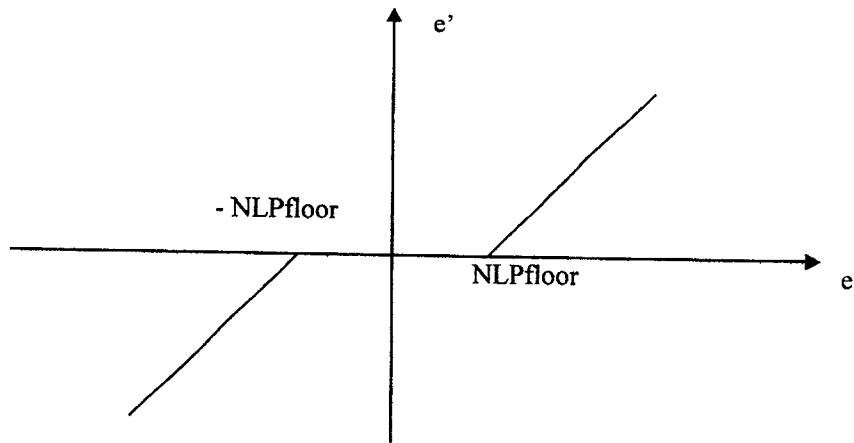


图 5