



(12) 发明专利

(10) 授权公告号 CN 101416237 B

(45) 授权公告日 2012. 05. 30

- (21) 申请号 200680054124. 1
- (22) 申请日 2006. 05. 01
- (85) PCT申请进入国家阶段日
2008. 10. 06
- (86) PCT申请的申请数据
PCT/US2006/016741 2006. 05. 01
- (87) PCT申请的公布数据
W02007/130026 EN 2007. 11. 15
- (73) 专利权人 日本电信电话株式会社
地址 日本东京都
专利权人 佐治亚科技研究公司
- (72) 发明人 中谷智广 庄炳煌
- (74) 专利代理机构 中国专利代理(香港)有限公司
72001
代理人 卢江 刘春元
- (51) Int. Cl.
G10L 21/02 (2006. 01)
- (56) 对比文件
US 5774562 A, 1998. 06. 30, 全文.
US 2004/0213415 A1, 2004. 10. 28, 全文.
CN 1542737 A, 2004. 11. 03, 全文.
US 2005/0010410 A1, 2005. 01. 13, 全文.
US 6944590 B2, 2005. 09. 13, 全文.

Tomohiro Nakatani, et al..HARMONICITY BASED DEREVERBERATION WITH MAXIMUM A POSTERIORI ESTIMATION. 《2005 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics》. 2005, 94-97.

Tetsuya Takiguchi, et al..ACOUSTIC MODEL ADAPTATION USING FIRST ORDER PREDICTION FOR REVERBERANT SPEECH. 《IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. (ICASSP '04). 》. 2004, 869-872.

审查员 王馨宁

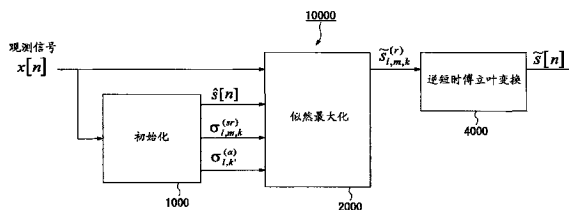
权利要求书 8 页 说明书 37 页 附图 23 页

(54) 发明名称

基于源和室内声学的概率模型的语音去混响方法和设备

(57) 摘要

通过接受用于初始化(1000)的观测信号以及执行包括傅立叶变换(4000)的似然最大化(2000)来实现语音去混响。



l ($= 1$ 到 L): 长时傅立叶频谱(LTFS)的时间索引
 m ($= 0$ 到 $M-1$): 短时傅立叶频谱(STFS)的时间子索引
 k ($= 1$ 到 $K^{(r)}$): STFS的频率索引
 k' ($= 1$ 到 K): LTFS的频率索引

1. 一种语音去混响设备,包括:

确定最大化似然函数的源信号估计的似然最大化单元,其中参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定,其中,所述表示源信号不确定性的第一方差是在给定源信号的情况下初始源信号估计的概率密度函数的方差以及所述表示声学环境不确定性的第二方差是在给定源信号的情况下观测信号和室内传递函数的逆滤波器的联合概率密度函数的方差,以及

其中,所述似然函数基于根据未知参数、丢失数据的第一随机变量、和观测数据的第二随机变量所评估的概率密度函数被定义,其中参考源信号估计来定义所述未知参数,丢失数据的第一随机变量表示室内传递函数的逆滤波器,以及参考观测信号和初始源信号估计来定义观测数据的第二随机变量。

2. 根据权利要求1所述的语音去混响设备,其中所述似然最大化单元使用迭代优化算法来确定源信号估计。

3. 根据权利要求2所述的语音去混响设备,其中所述迭代优化算法是期望最大化算法。

4. 根据权利要求1所述的语音去混响设备,其中所述似然最大化单元进一步包括:

逆滤波器估计单元,其参考初始源信号估计和更新的源信号估计之一、观测信号以及第二方差来计算逆滤波器估计;

滤波单元,其将逆滤波器估计应用于观测信号并产生滤波信号;

源信号估计和收敛检验单元,其参考初始源信号估计、第一方差、第二方差、和滤波信号来计算源信号估计,所述源信号估计和收敛检验单元进一步确定是否源信号估计的收敛被获得,如果源信号估计的收敛被获得,所述源信号估计和收敛检验单元进一步输出源信号估计作为去混响的信号;以及

更新单元,其将所述源信号估计更新为更新的源信号估计,如果源信号估计的收敛未被获得,所述更新单元则进一步提供更新的源信号估计给逆滤波器估计单元,以及所述更新单元在初始更新步骤中进一步提供所述初始源信号估计给逆滤波器估计单元。

5. 根据权利要求4所述的语音去混响设备,其中所述似然最大化单元进一步包括:

第一长时傅立叶变换单元,其执行波形观测信号到变换后的观测信号的第一长时傅立叶变换,该第一长时傅立叶变换单元进一步将变换后的观测信号作为观测信号提供给逆滤波器估计单元和滤波单元;

LTFS到STFS变换单元,其执行滤波信号到变换后的滤波信号的LTFS到STFS变换,该LTFS到STFS变换单元进一步将变换后的滤波信号作为滤波信号提供给源信号估计和收敛检验单元;

STFS到LTFS变换单元,其执行源信号估计到变换后的源信号估计的STFS到LTFS变换,如果源信号估计的收敛未被获得,该STFS到LTFS变换单元进一步将变换后的源信号估计作为源信号估计提供给更新单元;

第二长时傅立叶变换单元,其执行波形初始源信号估计到第一变换初始源信号估计的第二长时傅立叶变换,该第二长时傅立叶变换单元进一步将第一变换初始源信号估计作为初始源信号估计提供给更新单元;以及

短时傅立叶变换单元,其执行波形初始源信号估计到第二变换初始源信号估计的短时

傅立叶变换,该短时傅立叶变换单元进一步将第二变换初始源信号估计作为初始源信号估计提供给源信号估计和收敛检验单元。

6. 根据权利要求 1 所述的语音去混响设备,进一步包括:

执行源信号估计到波形源信号估计的逆短时傅立叶变换的逆短时傅立叶变换单元。

7. 根据权利要求 1 所述的语音去混响设备,进一步包括:

基于观测信号产生初始源信号估计、第一方差、和第二方差的初始化单元。

8. 根据权利要求 7 所述的语音去混响设备,其中所述初始化单元进一步包括:

基频估计单元,其估计来自变换后的信号的每个短时间帧的基频和发声量度,其中通过观测信号的短时傅立叶变换来给定所述变换后的信号;以及

源信号不确定性确定单元,其基于所述基频和发声量度确定第一方差。

9. 根据权利要求 1 所述的语音去混响设备,进一步包括:

初始化单元,其基于观测信号产生初始源信号估计、第一方差、和第二方差;和

收敛检验单元,其从似然最大化单元接收源信号估计,该收敛检验单元确定是否源信号估计的收敛被获得,如果获得源信号估计的收敛,则所述收敛检验单元进一步输出源信号估计作为去混响的信号,并且如果未获得源信号估计的收敛,则所述收敛检验单元进一步提供源信号估计给初始化单元以使初始化单元能够基于该源信号估计产生初始源信号估计、第一方差、和第二方差。

10. 根据权利要求 9 所述的语音去混响设备,其中所述初始化单元进一步包括:

第二短时傅立叶变换单元,其执行观测信号到第一变换观测信号的第二短时傅立叶变换;

第一选择单元,其执行第一选择操作以产生第一选择输出以及第二选择操作以产生第二选择输出,第一和第二选择操作是相互独立的,当所述第一选择单元接收第一变换观测信号的输入而不接收源信号估计的任何输入时,第一选择操作用于选择第一变换观测信号作为第一选择输出,而当所述第一选择单元接收第一变换观测信号和源信号估计的输入时,第一选择操作用于选择第一变换观测信号和源信号估计之一作为第一选择输出,当所述第一选择单元接收第一变换观测信号的输入而不接收源信号估计的任何输入时,第二选择操作用于选择第一变换观测信号作为第二选择输出,而当所述第一选择单元接收第一变换观测信号和源信号估计的输入时,第二选择操作用于选择第一变换观测信号和源信号估计之一作为第二选择输出;

基频估计单元,其接收第二选择输出并估计来自第二选择输出的每个短时间帧的基频和发声量度;以及

自适应谐波滤波单元,其接收第一选择输出、基频和发声量度,该自适应谐波滤波单元基于基频和发声量度增强第一选择输出的谐波结构以产生初始源信号估计。

11. 根据权利要求 9 所述的语音去混响设备,其中所述初始化单元进一步包括:

第三短时傅立叶变换单元,其执行观测信号到第二变换观测信号的第三短时傅立叶变换;

第二选择单元,其执行第三选择操作以产生第三选择输出,当所述第二选择单元接收第二变换观测信号的输入而不接收源信号估计的任何输入时,第三选择操作用于选择第二变换观测信号作为第三选择输出,而当所述第二选择单元接收第二变换观测信号和源信号

估计的输入时,第三选择操作用于选择第二变换观测信号和源信号估计之一作为第三选择输出;

基频估计单元,其接收第三选择输出并估计来自第三选择输出的每个短时间帧的基频和发声量度;以及

源信号不确定性确定单元,其基于基频和发声量度确定第一方差。

12. 根据权利要求 9 所述的语音去混响设备,进一步包括:

逆短时傅立叶变换单元,如果获得源信号估计的收敛,则所述逆短时傅立叶变换单元执行源信号估计到波形源信号估计的逆短时傅立叶变换。

13. 一种语音去混响设备,包括:

确定最大化似然函数的逆滤波器估计的似然最大化单元,其中参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定,其中,所述表示源信号不确定性的第一方差是在给定源信号的情况下初始源信号估计的概率密度函数的方差以及所述表示声学环境不确定性的第二方差是在给定源信号的情况下观测信号和室内传递函数的逆滤波器的联合概率密度函数的方差,以及

其中,所述似然函数基于根据第一未知参数、第二未知参数、以及观测数据的第一随机变量所评估的概率密度函数被定义,其中参考源信号估计来定义第一未知参数,参考室内传递函数的逆滤波器来定义第二未知参数,参考观测信号和初始源信号估计来定义观测数据的第一随机变量,逆滤波器估计是室内传递函数的逆滤波器的估计。

14. 根据权利要求 13 所述的语音去混响设备,其中所述似然最大化单元使用迭代优化算法来确定逆滤波器估计。

15. 根据权利要求 13 所述的语音去混响设备,进一步包括:

将逆滤波器估计应用于观测信号并产生源信号估计的逆滤波器应用单元。

16. 根据权利要求 15 所述的语音去混响设备,其中所述逆滤波器应用单元进一步包括:

第一逆长时傅立叶变换单元,其执行逆滤波器估计到变换后的逆滤波器估计的第一逆长时傅立叶变换;以及

卷积单元,其接收变换后的逆滤波器估计和观测信号,以及将观测信号与变换后的逆滤波器估计进行卷积以产生源信号估计。

17. 根据权利要求 15 所述的语音去混响设备,其中所述逆滤波器应用单元进一步包括:

第一长时傅立叶变换单元,其执行观测信号到变换后的观测信号的第一长时傅立叶变换;

第一滤波单元,其将逆滤波器估计应用于变换后的观测信号,以及产生滤波源信号估计;和

第二逆长时傅立叶变换单元,其执行滤波源信号估计到源信号估计的第二逆长时傅立叶变换。

18. 根据权利要求 13 所述的语音去混响设备,其中所述似然最大化单元进一步包括:

逆滤波器估计单元,其参考初始源信号估计和更新的源信号估计之一、观测信号以及第二方差计算逆滤波器估计;

收敛检验单元,其确定是否逆滤波器估计的收敛被获得,如果获得逆滤波器估计的收敛,该收敛检验单元进一步输出逆滤波器估计作为将要对所述观测信号进行去混响的滤波器,

滤波单元,如果未获得逆滤波器估计的收敛,该滤波单元从所述收敛检验单元接收逆滤波器估计,该滤波单元进一步将逆滤波器估计应用于观测信号并产生滤波信号;

源信号估计单元,其参考初始源信号估计、第一方差、第二方差、和滤波信号计算源信号估计;

更新单元,其将源信号估计更新为更新的源信号估计,该更新单元进一步在初始更新步骤中提供初始源信号估计给逆滤波器估计单元,该更新单元进一步在除初始更新步骤之外的更新步骤中提供更新的源信号估计给逆滤波器估计单元。

19. 根据权利要求 18 所述的语音去混响设备,其中所述似然最大化单元进一步包括:

第二长时傅立叶变换单元,其执行波形观测信号到变换后的观测信号的第二长时傅立叶变换,该第二长时傅立叶变换单元进一步将变换后的观测信号作为观测信号提供给逆滤波器估计单元和滤波单元;

LTFS 到 STFS 变换单元,其执行滤波信号到变换后的滤波信号的 LTFS 到 STFS 变换,该 LTFS 到 STFS 变换单元进一步将变换后的滤波信号作为滤波信号提供给源信号估计单元;

STFS 到 LTFS 变换单元,其执行源信号估计到变换后的源信号估计的 STFS 到 LTFS 变换,该 STFS 到 LTFS 变换单元进一步将变换后的源信号估计作为源信号估计提供给更新单元;

第三长时傅立叶变换单元,其执行波形初始源信号估计到第一变换初始源信号估计的第三长时傅立叶变换,该第三长时傅立叶变换单元进一步将第一变换初始源信号估计作为初始源信号估计提供给更新单元;以及

短时傅立叶变换单元,其执行波形初始源信号估计到第二变换初始源信号估计的短时傅立叶变换,该短时傅立叶变换单元进一步将第二变换初始源信号估计作为初始源信号估计提供给源信号估计单元。

20. 根据权利要求 13 所述的语音去混响设备,进一步包括:

基于观测信号产生初始源信号估计、第一方差、和第二方差的初始化单元。

21. 根据权利要求 20 所述的语音去混响设备,其中所述初始化单元进一步包括:

基频估计单元,其估计来自变换后的信号的每个短时间帧的基频和发声量度,其中通过观测信号的短时傅立叶变换来给定所述变换后的信号;以及

源信号不确定性确定单元,其基于基频和发声量度确定第一方差。

22. 一种语音去混响方法,包括:

确定最大化似然函数的源信号估计,其中参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定,其中,所述表示源信号不确定性的第一方差是在给定源信号的情况下初始源信号估计的概率密度函数的方差以及所述表示声学环境不确定性的第二方差是在给定源信号的情况下观测信号和室内传递函数的逆滤波器的联合概率密度函数的方差,以及

其中,所述似然函数基于根据未知参数、丢失数据的第一随机变量、和观测数据第二随机变量所评估的概率密度函数被定义,其中参考源信号估计来定义未知参数,丢失数据

的第一随机变量表示室内传递函数的逆滤波器,参考观测信号和初始源信号估计来定义观测数据的第二随机变量。

23. 根据权利要求 22 所述的语音去混响方法,其中使用迭代优化算法来确定源信号估计。

24. 根据权利要求 23 所述的语音去混响方法,其中所述迭代优化算法是期望最大化算法。

25. 根据权利要求 22 所述的语音去混响方法,其中确定源信号估计进一步包括:
参考初始源信号估计和更新的源信号估计之一、观测信号以及第二方差计算逆滤波器估计;

将逆滤波器估计应用于观测信号,以产生滤波信号;

参考初始源信号估计、第一方差、第二方差、和滤波信号计算源信号估计;

确定是否源信号估计的收敛被获得;

如果获得源信号估计的收敛,则输出源信号估计作为去混响的信号;以及

如果未获得源信号估计的收敛,则将源信号估计更新为更新的源信号估计。

26. 根据权利要求 25 所述的语音去混响方法,其中确定源信号估计进一步包括:
执行波形观测信号到变换后的观测信号的第一长时傅立叶变换,并进一步将变换后的观测信号作为观测信号提供以计算逆滤波器估计和产生滤波信号;

执行滤波信号到变换后的滤波信号的 LTFS 到 STFS 变换,并且进一步将变换后的滤波信号作为滤波信号提供以计算源信号估计和确定是否源信号估计的收敛被获得;

如果未获得源信号估计的收敛,执行源信号估计到变换后的源信号估计的 STFS 到 LTFS 变换并进一步将变换后的源信号估计作为源信号估计提供以将源信号估计更新为更新的源信号估计;

执行波形初始源信号估计到第一变换初始源信号估计的第二长时傅立叶变换,并进一步将第一变换初始源信号估计作为初始源信号估计提供以在初始更新步骤中被提供用于计算逆滤波器估计;以及

执行波形初始源信号估计到第二变换初始源信号估计的短时傅立叶变换,并进一步将第二变换初始源信号估计作为初始源信号估计提供以计算源信号估计和确定是否源信号估计的收敛被获得。

27. 根据权利要求 22 所述的语音去混响方法,进一步包括:

执行源信号估计到波形源信号估计的逆短时傅立叶变换。

28. 根据权利要求 22 所述的语音去混响方法,进一步包括:

基于观测信号产生初始源信号估计、第一方差、和第二方差。

29. 根据权利要求 28 所述的语音去混响方法,其中产生初始源信号估计、第一方差、和第二方差进一步包括:

估计来自变换后的信号的每个短时间帧的基频和发声量度,其中通过观测信号的短时傅立叶变换来给定所述变换后的信号;以及

基于所述基频和发声量度确定第一方差。

30. 根据权利要求 22 所述的语音去混响方法,进一步包括:

基于观测信号产生初始源信号估计、第一方差、和第二方差;

确定是否源信号估计的收敛被获得；

如果获得源信号估计的收敛,则输出源信号估计作为去混响的信号;以及

如果未获得源信号估计的收敛,则返回到产生初始源信号估计、第一方差、和第二方差。

31. 根据权利要求 30 所述的语音去混响方法,其中产生初始源信号估计、第一方差、和第二方差进一步包括:

执行观测信号到第一变换观测信号的第二短时傅立叶变换;

执行第一选择操作以产生第一选择输出,当接收第一变换观测信号的输入而不接收源信号估计的任何输入时,第一选择操作用于选择第一变换观测信号作为第一选择输出,当接收第一变换观测信号和源信号估计的输入时,第一选择操作用于选择第一变换观测信号和源信号估计之一作为第一选择输出;

执行第二选择操作以产生第二选择输出,当接收第一变换观测信号的输入而不接收源信号估计的任何输入时,第二选择操作用于选择第一变换观测信号作为第二选择输出,当接收第一变换观测信号和源信号估计的输入时,第二选择操作用于选择第一变换观测信号和源信号估计之一作为第二选择输出;

估计来自第二选择输出的每个短时间帧的基频和发声量度;以及

基于基频和发声量度增强第一选择输出的谐波结构以产生初始源信号估计。

32. 根据权利要求 30 所述的语音去混响方法,其中产生初始源信号估计、第一方差、和第二方差进一步包括:

执行观测信号到第二变换观测信号的第三短时傅立叶变换;

执行第三选择操作以产生第三选择输出,当接收第二变换观测信号的输入而不接收源信号估计的任何输入时,第三选择操作用于选择第二变换观测信号作为第三选择输出,当接收第二变换观测信号和源信号估计的输入时,第三选择操作用于选择第二变换观测信号和源信号估计之一作为第三选择输出;

估计来自第三选择输出的每个短时间帧的基频和发声量度;以及

基于所述基频和发声量度确定第一方差。

33. 根据权利要求 30 所述的语音去混响方法,进一步包括:

如果获得源信号估计的收敛,则执行源信号估计到波形源信号估计的逆短时傅立叶变换。

34. 一种语音去混响方法,包括:

确定最大化似然函数的逆滤波器估计,其中参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定,其中,所述表示源信号不确定性的第一方差是在给定源信号的情况下初始源信号估计的概率密度函数的方差以及所述表示声学环境不确定性的第二方差是在给定源信号的情况下观测信号和室内传递函数的逆滤波器的联合概率密度函数的方差,以及

其中,所述似然函数基于根据第一未知参数、第二未知参数、和观测数据的第一随机变量所评估的概率密度函数被定义,其中参考源信号估计来定义第一未知参数,参考室内传递函数的逆滤波器来定义第二未知参数,以及参考观测信号和初始源信号估计来定义观测数据的第一随机变量,所述逆滤波器估计是室内传递函数的逆滤波器的估计。

35. 根据权利要求 34 所述的语音去混响方法,其中使用迭代优化算法来确定逆滤波器估计。

36. 根据权利要求 34 所述的语音去混响方法,进一步包括:
将逆滤波器估计应用于观测信号以产生源信号估计。

37. 根据权利要求 36 所述的语音去混响方法,其中将逆滤波器估计应用于观测信号进一步包括:

执行逆滤波器估计到变换后的逆滤波器估计的第一逆长时傅立叶变换;以及
将观测信号与变换后的逆滤波器估计进行卷积以产生源信号估计。

38. 根据权利要求 36 所述的语音去混响方法,其中将逆滤波器估计应用于观测信号进一步包括:

执行观测信号到变换后的观测信号的第一长时傅立叶变换;
将逆滤波器估计应用于变换后的观测信号以产生滤波源信号估计;以及
执行滤波源信号估计到源信号估计的第二逆长时傅立叶变换。

39. 根据权利要求 34 所述的语音去混响方法,其中确定逆滤波器估计进一步包括:
参考初始源信号估计和更新的源信号估计之一、观测信号以及第二方差计算逆滤波器估计;

确定是否逆滤波器估计的收敛被获得;

如果获得逆滤波器估计的收敛,则输出逆滤波器估计作为将要对所述观测信号进行去混响的滤波器;

如果未获得逆滤波器估计的收敛,则将逆滤波器估计应用于观测信号以产生滤波信号;

参考初始源信号估计、第一方差、第二方差、和滤波信号计算源信号估计;以及
将源信号估计更新为更新的源信号估计。

40. 根据权利要求 39 所述的语音去混响方法,其中确定逆滤波器估计进一步包括:
执行波形观测信号到变换后的观测信号的第二长时傅立叶变换并进一步将变换后的观测信号作为观测信号提供以计算逆滤波器估计和产生滤波信号;

执行滤波信号到变换后的滤波信号的 LTFS 到 STFS 变换,并进一步将变换后的滤波信号作为滤波信号提供以计算源信号估计;

执行源信号估计到变换后的源信号估计的 STFS 到 LTFS 变换,并进一步将变换后的源信号估计作为源信号估计提供以将源信号估计更新为更新的源信号估计;

执行波形初始源信号估计到第一变换初始源信号估计的第三长时傅立叶变换,并进一步将第一变换初始源信号估计作为初始源信号估计提供以在初始更新步骤中被提供用于计算逆滤波器估计;以及

执行波形初始源信号估计到第二变换初始源信号估计的短时傅立叶变换,并进一步将第二变换初始源信号估计作为初始源信号估计提供以计算源信号估计。

41. 根据权利要求 34 所述的语音去混响方法,进一步包括:
基于观测信号产生初始源信号估计、第一方差、和第二方差。

42. 根据权利要求 41 所述的语音去混响方法,其中产生初始源信号估计、第一方差、和第二方差进一步包括:

估计来自变换后的信号的每个短时间帧的基频和发声量度,其中通过观测信号的短时傅立叶变换来给定所述变换后的信号;以及
基于所述基频和发声量度确定第一方差。

基于源和室内声学的概率模型的语音去混响方法和设备

技术领域

[0001] 本发明通常涉及用于语音去混响的方法和设备。更具体而言,本发明涉及基于源和室内声学的概率模型的语音去混响方法和设备。

背景技术

[0002] 以下在本申请中将被引用或标识的所有专利、专利申请、专利公开、科学文献等等将在此以其整体被包含作为参考以便更全面地描述本发明所属技术领域的状态。

[0003] 在普通室内通过远距离麦克风捕获的语音信号不可避免地包含混响,这对于语音信号的感知质量和清晰度具有有害的影响并且使自动语音识别 (ASR) 系统的性能降低。当混响时间长于 0.5 秒时,甚至当使用已经在匹配的混响条件下被训练的声学模型时,不能够提高识别性能。这在 B. Kingsbury 和 N. Morgan 的“Recognizing reverberant speech with rasta-plp”(Proc. 1997 IEEE International Conference Acoustic Speech and Signal Processing (ICASSP-97), Vol. 2, 第 1259-1262 页, 1997 年) 中被公开。语音信号的去混响无论对于高质量记录和重放或对于自动语音识别 (ASR) 都是必需的。

[0004] 尽管语音信号的盲去混响仍是有挑战性的问题,但是近来已经提出若干技术。已经提出将观测信号去相关同时在信号的短时间段内保持相关性的技术。这由 B. W. Gillespie 和 L. E. Atlas 的“Strategies for improving audible quality and speech recognition accuracy of reverberant speech”(Proc. 2003 IEEE International Conference Acoustics, Speech and Signal Processing (ICASSP-2003), Vol. 1, 第 676-679 页, 2003 年) 公开。这还由 H. Buchner、R. Aichner、和 W. Kellermann 的“Trinicon: a versatile framework for multichannel blind signal processing”(Proc. of the 2004 IEEE International Conference Acoustics, Speech and Signal Processing (ICASSP-2004), Vol. III, 第 889-892 页, 2004 年 5 月) 公开。

[0005] 已经提出用于估计和均衡室内的声响应中的极点的方法。这由 T. Hikichi 和 M. Miyoshi 的“Blind algorithm for calculating common poles based on linear prediction”(Proc. of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004), Vol. IV, 第 89-92 页, 2004 年 5 月) 公开。这还由 J. R. Hopgood 和 P. J. W. Rayner 的“Blind single channel deconvolution using nonstationary signal processing”(IEEE Transactions Speech and Audio processing, vol. 11, no. 5, 第 467-488 页, 2003 年 9 月) 公开。

[0006] 而且,已经提出基于语音信号的本质特征的两种方法,即在下文中被称为 HERB 的基于调和性的去混响、和在下文中被称为 SBD 的基于稀疏性的去混响。HERB 由 T. Nakatani 和 M. Miyoshi 的“Blind dereverberation of single channel speech signal based on harmonic structure”(Proc. ICASSP-2003, vol. 1, 第 92-95 页, 2003 年 4 月) 公开。首次出版第 2004-274234 号的日本未审查专利申请公开了用于 HERB 的传统技术的一个实例。SBD 由 K. Kinoshita、T. Nakatani 和 M. Miyoshi 的“Efficient

blind dereverberation framework for automatic speech recognition” (Proc. Interspeech-2005, 2005 年 9 月) 公开。

[0007] 这些方法广泛使用源信号的初始估计中的各语音特征。然后,初始源信号估计和观测到的混响信号被一起用于估计用于去混响的逆滤波器,这允许进一步改进源信号估计。为了获得初始源信号估计,HERB 利用自适应谐波滤波器,以及 SBD 利用基于最小统计的频谱相减。在实验上已经显示:如果信号足够长,这些方法大大地提高观测到的混响信号的 ASR 性能。

[0008] 鉴于上述内容,本领域普通技术人员从本公开内容将会明白存在对用于语音去混响的改进设备和 / 或方法的需要。本发明解决本领域中的这种需要以及其他需要,这对于本领域普通技术人员来说根据本公开内容将会变得显而易见。

发明内容

[0009] 相应地,本发明的主要目的是提供语音去混响设备。

[0010] 本发明的另一个目的是提供语音去混响方法。

[0011] 本发明的再一个目的是提供将通过计算机执行的用于执行语音去混响方法的程序。

[0012] 本发明的又一个目的是提供存储将通过计算机执行的用于执行语音去混响方法的程序的存储介质。

[0013] 根据本发明的第一方面,语音去混响设备包括确定最大化似然函数的源信号估计的似然最大化单元。参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定。

[0014] 所述似然函数可以优选地基于根据未知参数、丢失数据的第一随机变量、和观测数据第二随机变量所评估的概率密度函数被定义。参考源信号估计来定义未知参数。丢失数据的第一随机变量表示室内传递函数的逆滤波器。参考观测信号和初始源信号估计来定义观测数据第二随机变量。

[0015] 上述似然最大化单元可以优选地使用迭代优化算法来确定源信号估计。迭代优化算法可以优选地是期望最大化算法。

[0016] 似然最大化单元可以进一步包括但不限于逆滤波器估计单元、滤波单元、源信号估计和收敛检验单元、以及更新单元。所述逆滤波器估计单元参考观测信号、第二方差、以及初始源信号估计和更新的源信号估计之一计算逆滤波器估计。滤波单元将逆滤波器估计应用于观测信号,并且产生滤波信号。源信号估计和收敛检验单元参考初始源信号估计、第一方差、第二方差、和滤波信号计算源信号估计。源信号估计和收敛检验单元进一步确定是否源信号估计的收敛被获得。如果源信号估计的收敛被获得,所述源信号估计和收敛检验单元进一步输出源信号估计作为去混响的信号。所述更新单元将源信号估计更新为更新的源信号估计。如果源信号估计的收敛未被获得,则所述更新单元进一步提供更新的源信号估计给逆滤波器估计单元。该更新单元进一步在初始更新步骤中提供初始源信号估计给逆滤波器估计单元。

[0017] 似然最大化单元可以进一步包括但不限于第一长时傅立叶变换单元、LTFS 到 STFS 变换单元、STFS 到 LTFS 变换单元、第二长时傅立叶变换单元、和短时傅立叶变换单元。

第一长时傅立叶变换单元执行波形观测信号到变换后的观测信号的第一长时傅立叶变换。第一长时傅立叶变换单元进一步将变换后的观测信号作为观测信号提供给逆滤波器估计单元和滤波单元。LTFS到STFS变换单元执行滤波信号到变换后的滤波信号的LTFS到STFS变换。LTFS到STFS变换单元进一步将变换后的滤波信号作为滤波信号提供给源信号估计和收敛检验单元。STFS到LTFS变换单元执行源信号估计到变换后的源信号估计的STFS到LTFS变换。如果源信号估计的收敛未被获得,则STFS到LTFS变换单元进一步将变换后的源信号估计作为源信号估计提供给更新单元。第二长时傅立叶变换单元执行波形初始源信号估计到第一变换初始源信号估计的第二长时傅立叶变换。第二长时傅立叶变换进一步将第一变换初始源信号估计作为初始源信号估计提供给更新单元。短时傅立叶变换单元执行波形初始源信号估计到第二变换初始源信号估计的短时傅立叶变换。短时傅立叶变换单元进一步将第二变换初始源信号估计作为初始源信号估计提供给源信号估计和收敛检验单元。

[0018] 所述语音去混响设备可以进一步包括但不限于执行源信号估计到波形源信号估计的逆短时傅立叶变换的逆短时傅立叶变换单元。

[0019] 所述语音去混响设备可以进一步包括但不限于基于观测信号产生初始源信号估计、第一方差、和第二方差的初始化单元。在这种情况下,初始化单元可以进一步包括但不限于基频估计单元和源信号不确定性确定单元。基频估计单元估计来自变换后的信号的每个短时间帧的基频和发声量度(voicing measure),其中通过观测信号的短时傅立叶变换来给定所述变换后的信号。源信号不确定性确定单元基于基频和发声量度确定第一方差。

[0020] 所述语音去混响设备可以进一步包括但不限于初始化单元、和收敛检验单元。初始化单元基于观测信号产生初始源信号估计、第一方差、和第二方差。收敛检验单元从似然最大化单元接收源信号估计。收敛检验单元确定是否源信号估计的收敛被获得。如果获得源信号估计的收敛,则所述收敛检验单元进一步输出源信号估计作为去混响的信号。如果未获得源信号估计的收敛,则收敛检验单元进一步提供源信号估计给初始化单元以使初始化单元能够基于该源信号估计产生初始源信号估计、第一方差、和第二方差。

[0021] 在最后描述的情况下,初始化单元可以进一步包括但不限于第二短时傅立叶变换单元、第一选择单元、基频估计单元、和自适应谐波滤波单元。第二短时傅立叶变换单元执行观测信号到第一变换观测信号的第二短时傅立叶变换。第一选择单元执行第一选择操作以产生第一选择输出以及第二选择操作以产生第二选择输出。第一和第二选择操作是相互独立的。当第一选择单元接收第一变换观测信号的输入而不接收源信号估计的任何输入时,第一选择操作用于选择第一变换观测信号作为第一选择输出。当第一选择单元接收第一变换观测信号和源信号估计的输入时,第一选择操作还用于选择第一变换观测信号和源信号估计之一作为第一选择输出。当第一选择单元接收第一变换观测信号的输入而不接收源信号估计的任何输入时,第二选择操作用于选择第一变换观测信号作为第二选择输出。当第一选择单元接收第一变换观测信号和源信号估计的输入时,第二选择操作还用于选择第一变换观测信号和源信号估计之一作为第二选择输出。基频估计单元接收第二选择输出。基频估计单元还估计来自第二选择输出的每个短时间帧的基频和发声量度。自适应谐波滤波单元接收第一选择输出、基频和发声量度。自适应谐波滤波单元基于基频和发声量度增强第一选择输出的谐波结构以产生初始源信号估计。

[0022] 初始化单元可以进一步包括但不限于第三短时傅立叶变换单元、第二选择单元、基频估计单元、和源信号不确定性确定单元。第三短时傅立叶变换单元执行观测信号到第二变换观测信号的第三短时傅立叶变换。第二选择单元执行第三选择操作以产生第三选择输出。当第二选择单元接收第二变换观测信号的输入而不接收源信号估计的任何输入时，第三选择操作用于选择第二变换观测信号作为第三选择输出。当第二选择单元接收第二变换观测信号和源信号估计的输入时，第三选择操作还用于选择第二变换观测信号和源信号估计之一作为第三选择输出。基频估计单元接收第三选择输出。基频估计单元估计来自第三选择输出的每个短时间帧的基频和发声量度。源信号不确定性确定单元基于基频和发声量度确定第一方差。

[0023] 所述语音去混响设备可以进一步包括但不限于逆短时傅立叶变换单元，如果获得源信号估计的收敛，则该逆短时傅立叶变换单元执行源信号估计到波形源信号估计的逆短时傅立叶变换。

[0024] 根据本发明的第二方面，语音去混响设备包括确定最大化似然函数的逆滤波器估计的似然最大化单元。参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定。

[0025] 似然函数可以优选地基于根据第一未知参数、第二未知参数、以及观测数据的第一随机变量所评估的概率密度函数被定义。参考源信号估计来定义第一未知参数。参考室内传递函数的逆滤波器来定义第二未知参数。参考观测信号和初始源信号估计来定义观测数据的第一随机变量。逆滤波器估计是室内传递函数的逆滤波器的估计。

[0026] 似然最大化单元可以优选地使用迭代优化算法来确定逆滤波器估计。

[0027] 所述语音去混响设备可以进一步包括但不限于将逆滤波器估计应用于观测信号并产生源信号估计的逆滤波器应用单元。

[0028] 该逆滤波器应用单元可以进一步包括但不限于第一逆长时傅立叶变换单元、和卷积单元。第一逆长时傅立叶变换单元执行逆滤波器估计到变换后的逆滤波器估计的第一逆长时傅立叶变换。卷积单元接收变换后的逆滤波器估计和观测信号。卷积单元将观测信号与变换后的逆滤波器估计进行卷积以产生源信号估计。

[0029] 该逆滤波器应用单元可以进一步包括但不限于第一长时傅立叶变换单元、第一滤波单元、和第二逆长时傅立叶变换单元。第一长时傅立叶变换单元执行观测信号到变换后的观测信号的第一长时傅立叶变换。第一滤波单元将逆滤波器估计应用于变换后的观测信号。第一滤波单元产生滤波源信号估计。第二逆长时傅立叶变换单元执行滤波源信号估计到源信号估计的第二逆长时傅立叶变换。

[0030] 似然最大化单元可以进一步包括但不限于逆滤波器估计单元、收敛检验单元、滤波单元、源信号估计单元、和更新单元。逆滤波器估计单元参考观测信号、第二方差、以及初始源信号估计和更新的源信号估计之一计算逆滤波器估计。收敛检验单元确定是否逆滤波器估计的收敛被获得。如果获得源信号估计的收敛，收敛检验单元进一步输出逆滤波器估计作为将对所述观测信号进行去混响的滤波器。如果未获得源信号估计的收敛，滤波单元从收敛检验单元接收逆滤波器估计。滤波单元进一步将逆滤波器估计应用于观测信号。滤波单元进一步产生滤波信号。源信号估计单元参考初始源信号估计、第一方差、第二方差、和滤波信号计算源信号估计。更新单元将源信号估计更新为更新的源信号估计。更新单元

进一步在初始更新步骤中提供初始源信号估计给逆滤波估计单元。更新单元进一步在除初始更新步骤之外的更新步骤中提供更新的源信号估计给逆滤波器估计单元。

[0031] 似然最大化单元可以进一步包括但不限于第二长时傅立叶变换单元、LTFS 到 STFS 变换单元、STFS 到 LTFS 变换单元、第三长时傅立叶变换单元、和短时傅立叶变换单元。第二长时傅立叶变换单元执行波形观测信号到变换后的观测信号的第二长时傅立叶变换。第二长时傅立叶变换单元进一步将变换后的观测信号作为观测信号提供给逆滤波器估计单元和滤波单元。LTFS 到 STFS 变换单元执行滤波信号到变换后的滤波信号的 LTFS 到 STFS 变换。LTFS 到 STFS 变换单元进一步将变换后的滤波信号作为滤波信号提供给源信号估计单元。STFS 到 LTFS 变换单元执行源信号估计到变换后的源信号估计的 STFS 到 LTFS 变换。STFS 到 LTFS 变换单元进一步将变换后的源信号估计作为源信号估计提供给更新单元。第三长时傅立叶变换单元执行波形初始源信号估计到第一变换初始源信号估计的第三长时傅立叶变换。第三长时傅立叶变换单元进一步将第一变换初始源信号估计作为初始源信号估计提供给更新单元。短时傅立叶变换单元执行波形初始源信号估计到第二变换初始源信号估计的短时傅立叶变换。短时傅立叶变换单元进一步将第二变换初始源信号估计作为初始源信号估计提供给源信号估计单元。

[0032] 所述语音去混响设备可以进一步包括但不限于基于观测信号产生初始源信号估计、第一方差、和第二方差的初始化单元。

[0033] 所述初始化单元可以进一步包括但不限于基频估计单元、和源信号不确定性确定单元。基频估计单元估计来自变换后的信号的每个短时间帧的基频和发声量度,其中通过观测信号的短时傅立叶变换来给定所述变换后的信号。源信号不确定性确定单元基于基频和发声量度确定第一方差。

[0034] 根据本发明的第三方面,语音去混响方法包括确定最大化似然函数的源信号估计。参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定。

[0035] 所述似然函数可以优选地基于根据未知参数、丢失数据的第一随机变量、和观测数据第二随机变量所评估的概率密度函数被定义。参考源信号估计来定义未知参数。丢失数据的第一随机变量表示室内传递函数的逆滤波器。参考观测信号和初始源信号估计来定义观测数据第二随机变量。

[0036] 可以优选地使用迭代优化算法来确定源信号估计。该迭代优化算法可以优选地是期望最大化算法。

[0037] 用于确定源信号估计的过程可以进一步包括但不限于以下过程。参考观测信号、第二方差、以及初始源信号估计和更新的源信号估计之一计算逆滤波器估计。将该逆滤波器估计应用于观测信号,以产生滤波信号。参考初始源信号估计、第一方差、第二方差、和滤波信号来计算所述源信号估计。确定是否源信号估计的收敛被获得。如果获得源信号估计的收敛,则输出源信号估计作为去混响的信号。如果未获得源信号估计的收敛,则将源信号估计更新为更新的源信号估计。

[0038] 用于确定源信号估计的过程可以进一步包括但不限于以下过程。执行第一长时傅立叶变换以将波形观测信号变换为变换后的观测信号。执行 LTFS 到 STFS 变换以将滤波信号变换为变换后的滤波信号。如果未获得源信号估计的收敛,则执行 STFS 到 LTFS 变换以

将源信号估计变换为变换后的源信号估计。执行第二长时傅立叶变换以将波形初始源信号估计变换为第一变换初始源信号估计。执行短时傅立叶变换以将波形初始源信号估计变换为第二变换初始源信号估计。

[0039] 所述语音去混响方法可以进一步包括但不限于执行源信号估计到波形源信号估计的逆短时傅立叶变换。

[0040] 所述语音去混响方法可以进一步包括但不限于基于观测信号产生初始源信号估计、第一方差、和第二方差。

[0041] 在最后描述的情况下，产生初始源信号估计、第一方差、和第二方差可以进一步包括但不限于以下过程。估计来自变换后的信号的每个短时间帧的基频和发声量度，其中通过观测信号的短时傅立叶变换来给定所述变换后的信号。基于所述基频和发声量度确定第一方差。

[0042] 所述语音去混响方法可以进一步包括但不限于以下过程。基于观测信号产生初始源信号估计、第一方差、和第二方差。确定是否源信号估计的收敛被获得。如果获得源信号估计的收敛，则输出源信号估计作为去混响的信号。如果未获得源信号估计的收敛，则该过程将返回产生初始源信号估计、第一方差、和第二方差。

[0043] 在最后描述的情况下，产生初始源信号估计、第一方差、和第二方差可以进一步包括但不限于以下过程。执行第二短时傅立叶变换以将观测信号变换为第一变换观测信号。执行第一选择操作以产生第一选择输出。当接收第一变换观测信号的输入而不接收源信号估计的任何输入时，第一选择操作用于选择第一变换观测信号作为第一选择输出。当接收第一变换观测信号和源信号估计的输入时，第一选择操作用于选择第一变换观测信号和源信号估计之一作为第一选择输出。执行第二选择操作以产生第二选择输出。当接收第一变换观测信号的输入而不接收源信号估计的任何输入时，第二选择操作用于选择第一变换观测信号作为第二选择输出。当接收第一变换观测信号和源信号估计的输入时，第二选择操作用于选择第一变换观测信号和源信号估计之一作为第二选择输出。估计来自第二选择输出的每个短时间帧的基频和发声量度。基于基频和发声量度增强第一选择输出的谐波结构以产生初始源信号估计。

[0044] 产生初始源信号估计、第一方差、和第二方差可以进一步包括但不限于以下过程。执行第三短时傅立叶变换以将观测信号变换为第二变换观测信号。执行第三选择操作以产生第三选择输出。当接收第二变换观测信号的输入而不接收源信号估计的任何输入时，第三选择操作用于选择第二变换观测信号作为第三选择输出。当接收第二变换观测信号和源信号估计的输入时，第三选择操作用于选择第二变换观测信号和源信号估计之一作为第三选择输出。估计来自第三选择输出的每个短时间帧的基频和发声量度。基于所述基频和发声量度确定第一方差。

[0045] 所述语音去混响方法可以进一步包括但不限于：如果获得源信号估计的收敛，则执行源信号估计到波形源信号估计的逆短时傅立叶变换。

[0046] 根据本发明的第四方面，语音去混响方法包括确定最大化似然函数的逆滤波器估计。参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定。

[0047] 所述似然函数可以优选地基于根据第一未知参数、第二未知参数、观测数据的第

一随机变量所评估的概率密度函数被定义。参考源信号估计来定义第一未知参数。参考室内传递函数的逆滤波器来定义第二未知参数。参考观测信号和初始源信号估计来定义观测数据的第一随机变量。逆滤波器估计是室内传递函数的逆滤波器的估计。

[0048] 可以优选地使用迭代优化算法来确定逆滤波器估计。

[0049] 所述语音去混响方法可以进一步包括但不限于将逆滤波器估计应用于观测信号以产生源信号估计。

[0050] 在一种情况下,最后描述的用于将逆滤波器估计应用于观测信号的过程可以进一步包括但不限于以下过程。执行第一逆长时傅立叶变换以将逆滤波器估计变换为变换后的逆滤波器估计。将观测信号与变换后的逆滤波器估计进行卷积以产生源信号估计。

[0051] 在另一种情况下,最后描述的用于将逆滤波器估计应用于观测信号的过程可以进一步包括但不限于以下过程。执行第一长时傅立叶变换以将观测信号变换为变换后的观测信号。将逆滤波器估计应用于变换后的观测信号以产生滤波源信号估计。执行第二逆长时傅立叶变换以将滤波源信号估计变换为源信号估计。

[0052] 在再一种情况下,确定逆滤波器估计可以进一步包括但不限于以下过程。参考观测信号、第二方差、以及初始源信号估计和更新的源信号估计之一计算逆滤波器估计。确定是否逆滤波器估计的收敛被获得。如果获得源信号估计的收敛,则输出逆滤波器估计作为将对所述观测信号进行去混响的滤波器。如果未获得源信号估计的收敛,则将逆滤波器估计应用于观测信号以产生滤波信号。参考初始源信号估计、第一方差、第二方差、和滤波信号计算源信号估计。将源信号估计更新为更新的源信号估计。

[0053] 在最后描述的情况下,用于确定逆滤波器估计的过程可以进一步包括但不限于以下过程。执行第二长时傅立叶变换以将波形观测信号变换为变换后的观测信号。执行 LTFS 到 STFS 变换以将滤波信号变换为变换后的滤波信号。执行 STFS 到 LTFS 变换以将源信号估计变换为变换后的源信号估计。执行第三长时傅立叶变换以将波形初始源信号估计变换为第一变换初始源信号估计。执行短时傅立叶变换以将波形初始源信号估计变换为第二变换初始源信号估计。

[0054] 所述语音去混响方法可以进一步包括但不限于基于观测信号产生初始源信号估计、第一方差、和第二方差。

[0055] 在一种情况下,最后描述的用于产生初始源信号估计、第一方差、和第二方差的过程可以进一步包括但不限于以下过程。估计来自变换后的信号的每个短时间帧的基频和发声量度,其中通过观测信号的短时傅立叶变换来给定所述变换后的信号。基于所述基频和发声量度确定第一方差。

[0056] 根据本发明的第五方面,要通过计算机执行的用于执行语音去混响方法的程序,所述方法包括确定最大化似然函数的源信号估计。参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定。

[0057] 根据本发明的第六方面,要通过计算机执行的用于执行语音去混响方法的程序,所述方法包括确定最大化似然函数的逆滤波器估计。参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定。

[0058] 根据本发明的第七方面,存储介质存储要通过计算机执行的用于执行语音去混响方法的程序,所述方法包括确定最大化似然函数的源信号估计。参考观测信号、初始源信号

估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定。

[0059] 根据本发明的第八方面,存储介质存储要通过计算机执行的用于执行语音去混响方法的程序,所述方法包括确定最大化似然函数的逆滤波器估计。参考观测信号、初始源信号估计、表示源信号不确定性的第一方差、和表示声学环境不确定性的第二方差进行所述确定。

[0060] 根据以下结合附图进行的详细描述,本发明的这些和其他目的、特征、方面和优点对于本领域普通技术人员来说将会变得显而易见,所述附图图示本发明的实施例。

附图说明

[0061] 现在参考形成本原始公开内容的一部分的附图:

[0062] 图 1 是图示在本发明的第一实施例中基于源和室内声学的概率模型的语音去混响设备的方框图;

[0063] 图 2 是图示包括在图 1 中所示的语音去混响设备中的似然最大化单元的配置的方框图;

[0064] 图 3A 是图示包括在图 2 中所示的似然最大化单元中的 STFS 到 LTFS 变换单元的配置的方框图;

[0065] 图 3B 是图示包括在图 2 中所示的似然最大化单元中的 LTFS 到 STFS 变换单元的配置的方框图;

[0066] 图 4A 是图示包括在图 2 中所示的似然最大化单元中的长时傅立叶变换单元的配置的方框图;

[0067] 图 4B 是图示包括在图 3B 中所示的 LTFS 到 STFS 变换单元中的逆长时傅立叶变换单元的配置的方框图;

[0068] 图 5A 是图示包括在图 3B 中所示的 LTFS 到 STFS 变换单元中的短时傅立叶变换单元的配置的方框图;

[0069] 图 5B 是图示包括在图 3A 中所示的 STFS 到 LTFS 变换单元中的逆短时傅立叶变换单元的配置的方框图;

[0070] 图 6 是图示包括在图 1 中所示的初始化单元中的初始源信号估计单元的配置的方框图;

[0071] 图 7 是图示包括在图 1 中所示的初始化单元中的源信号不确定性确定单元的配置的方框图;

[0072] 图 8 是图示包括在图 1 中所示的初始化单元中的声学环境不确定性确定单元的配置的方框图;

[0073] 图 9 是图示根据本发明第二实施例的另一种语音去混响设备的配置的方框图;

[0074] 图 10 是图示包括在图 9 中所示的初始化单元中的修改的初始源信号估计单元的配置的方框图;

[0075] 图 11 是图示包括在图 9 中所示的初始化单元中的修改的源信号不确定性确定单元的配置的方框图;

[0076] 图 12 是图示根据本发明第三实施例的又一种语音去混响设备的配置的方框图;

[0077] 图 13 是图示包括在图 12 中所示的语音去混响设备中的似然最大化单元的配置的方框图；

[0078] 图 14 是图示包括在图 12 中所示的语音去混响设备中的逆滤波器应用单元的配置的方框图；

[0079] 图 15 是图示包括在图 12 中所示的语音去混响设备中的另一个逆滤波器应用单元的配置的方框图；

[0080] 图 16A 图示当女人说话时在 $RT60 = 1.0$ 秒的能量衰减曲线；

[0081] 图 16B 图示当女人说话时在 $RT60 = 0.5$ 秒的能量衰减曲线；

[0082] 图 16C 图示当女人说话时在 $RT60 = 0.2$ 秒的能量衰减曲线；

[0083] 图 16D 图示当女人说话时在 $RT60 = 0.1$ 秒的能量衰减曲线；

[0084] 图 16E 图示当男人说话时在 $RT60 = 1.0$ 秒的能量衰减曲线；

[0085] 图 16F 图示当男人说话时在 $RT60 = 0.5$ 秒的能量衰减曲线；

[0086] 图 16G 图示当男人说话时在 $RT60 = 0.2$ 秒的能量衰减曲线；以及

[0087] 图 16H 图示当男人说话时在 $RT60 = 0.1$ 秒的能量衰减曲线。

具体实施方式

[0088] 根据本发明的一个方面，提供单声道语音去混响方法，其中源信号和室内声学的特征通过概率密度函数 (pdf) 表示以及通过最大化基于概率密度函数 (pdf) 定义的似然函数来估计源信号。基于两种基本语音信号特征、即调和性和稀疏性，两种类型的概率密度函数 (pdf) 被引入用于源信号，同时基于逆滤波操作来定义室内声学的概率密度函数 (pdf)。期望最大化 (EM) 算法被用于有效地解决最大似然问题。所产生的算法通过经由期望最大化 (EM) 迭代将其源信号特征与室内声学特征结合来精心制作仅仅基于其源信号特征所给定的初始源信号估计。用去混响脉冲响应的能量衰减曲线来显示本方法的有效性。

[0089] 尽管上述的 HERB 和 SBD 在获得去混响滤波器中有效地利用语音信号特征，但是它们并不提供能够优化其性能的分析构架。根据本发明的一个方面，上述的 HERB 和 SBD 被重新用公式表示为最大似然 (ML) 估计问题，其中源信号被确定为在给定观测信号的情况下最大化似然函数的信号。为此目的，两个概率密度函数 (pdf) 被引入用于初始源信号估计和去混响滤波器，以便基于期望最大化 (EM) 算法最大化似然函数。实验结果显示：在给定相同数量的观测信号的情况下在去混响脉冲响应的能量衰减曲线方面能够进一步提高 HERB 和 SBD 的性能。以下描述将针对在本发明的一个方面中使用的傅立叶频谱。

[0090] 短时傅立叶频谱和长时傅立叶频谱

[0091] 本发明的一个方面将合并有关考虑源特性的语音信号特征和有关考虑混响效果的室内声学特征的信息。大约几十毫秒的短时间帧的相继应用对于分析这样的随时间变化的语音特征可能是有用的，而通常可能需要大约几千毫秒的相对长的时间帧来计算室内声学特征。本发明的一个方面将引入基于这两种分析帧的两种类型的傅立叶频谱、即在下文中被称为“STFS”的短时傅立叶频谱和在下文中被称为“LTFS”的长时傅立叶频谱。STFS 和 LTFS 中的各频率分量通过具有尾标“(r)”的符号被表示为 $s_{l,m,k}^{(r)}$ 以及通过另一个不具有尾标的符号被表示为 $s_{l,k'}$ ，其中 $s_{l,k'}$ 的 l 是用于 LTFS 的长时间帧的索引， k' 是用于 LTFS 的频率索引， $s_{l,m,k}^{(r)}$ 的 l 是包括用于 STFS 的短时间帧的长时间帧的索引， $s_{l,m,k}^{(r)}$ 的 m 是包括在长

时间帧中的短时间帧的索引,以及 $s_{l,m,k}^{(r)}$ 的 k 是用于 STFS 的频率索引。短时间帧能够被视为长时间帧的组成部分。所以,STFS 中的频率分量具有两个下标 l 和 m 。两个频谱被如下定义:

$$[0092] \quad s_{l,m,k}^{(r)} = 1/K^{(r)} \sum_{n=0}^{K^{(r)}-1} g^{(r)}[n]s[t_{l,m}+n]e^{-j2\pi kn/K^{(r)}},$$

$$[0093] \quad s_{l,k} = 1/K \sum_{n=0}^{K-1} g[n]s[t_l+n]e^{-j2\pi kn/K},$$

(1)

[0094] 其中 $s[n]$ 是数字化波形信号, $g^{(r)}[n]$ 和 $g[n]$ 、 $K^{(r)}$ 和 K 、以及 $t_{l,m}$ 和 t_l 分别是窗函数、离散傅立叶变换(DFT)点的数量、和用于STFS和LTFS的时间索引。设置 $t_{l,m}$ 和 t_l 之间的关系为对于 $m=0$ 到 $M-1$, $t_{l,m} = t_l + m\tau$,其中 τ 是相继的短时间帧之间的帧移。此外,引入以下归一化条件:

$$[0095] \quad K = \kappa K^{(r)},$$

$$[0096] \quad g[n] = \kappa \sum_{m=0}^{M-1} g^{(r)}[n - m\tau].$$

(2)

[0097] 其中 κ 是整数常量。利用该条件,在STFS, $s_{l,m,k}^{(r)}$ 和LTFS, $s_{l,k'}$ 之间保持以下方程,其中 $k' = \kappa k$:

$$[0098] \quad s_{l,k'} = \sum_{m=0}^{M-1} s_{l,m,k}^{(r)} \eta^{-m},$$

(3)

[0099] 其中 $\eta = e^{j2\pi\kappa\tau/K^{(r)}}$ 。定义逆操作,通过 $LS_{m,k}\{\cdot\}$ 来表示,该逆操作将由 $\{s_{l,k'}\}_1$ 所表示的在长时间帧 l 的一组LTFS仓 $s_{l,k'}$ ($k' = 1-K$)变换为在短时间帧 m 和频率索引 k 的STFS仓(bin):

$$[0100] \quad s_{l,m,k}^{(r)} = LS_{m,k}\{\{s_{l,k'}\}_1\}.$$

(4)

[0101] 通过级联逆长时傅立叶变换和短时傅立叶变换能够实现这种变换。显然, $LS_{m,k}\{\cdot\}$ 是线性操作符。

[0102] 信号的三种类型的表示、即波形数字化信号、短时傅立叶频谱(STFS)和长时傅立叶频谱(LTFS)包含相同的信息,并且能够使用已知的变换在无任何主要信息丢失的情况下从一种类型被变换到另一种类型。

[0103] 源和室内声学的概率模型

[0104] 定义以下术语：

[0105] $x_{l,m,k}^{(r)}$ ：观测到的混响信号的 STFS

[0106] $s_{l,m,k}^{(r)}$ ：未知源信号的 STFS

[0107] $\hat{s}_{l,m,k}^{(r)}$ ：初始源信号估计的 STFS

[0108] $w_{k'}$ ：未知逆滤波器的 LTFS ($k' = \kappa k$)

[0109] (5)

[0110] 假设 $x_{l,m,k}^{(r)}$ 、 $s_{l,m,k}^{(r)}$ 、 $\hat{s}_{l,m,k}^{(r)}$ 和 $w_{k'}$ 分别是随机过程 $X_{l,m,k}^{(r)}$ 、 $S_{l,m,k}^{(r)}$ 、 $\hat{S}_{l,m,k}^{(r)}$ 和 $W_{k'}$ 的实现，以及 $\hat{s}_{l,m,k}^{(r)}$ 是根据观测信号基于语音信号的特征、诸如调和性和稀疏性而给定的。

[0111] 在以下描述的本发明的一个实施例中， $s_{l,m,k}^{(r)}$ 或 $s_{l,k'}$ 被作为未知参数

[0112] 处理， $w_{k'}$ 被作为丢失数据的第一随机变量处理，以及 $x_{l,m,k}^{(r)}$ 或 $X_{l,k'}$ 被作为第二随机变量的一部分处理，和 $\hat{s}_{l,m,k}^{(r)}$ 或 $\hat{s}_{l,k'}$ 被作为第二随机变量的另一部分处理。

[0113] 假设：针对一定的持续时间给定 $x_{l,m,k}^{(r)}$ 和 $\hat{s}_{l,m,k}^{(r)}$ ，以及给定 $z_k^{(r)} = \{\{x_{l,m,k}^{(r)}\}_k \{\hat{s}_{l,m,k}^{(r)}\}_k\}$ ，其中 $\{\cdot\}_k$ 表示在频率索引 k 的 STFS 仓的时间系列。

[0114] 在此情况下，假设语音通过估计最大化在每个频率索引 k 定义的似然函数的源信号能够去混响：

$$\begin{aligned} [0115] \quad \theta_k &= \arg \max_{\Theta_k} \log p\{z_k^{(r)} | \Theta_k\} \\ [0116] \quad &= \arg \max_{\Theta_k} \log \int p\{w_{k'}, z_k^{(r)} | \Theta_k\} dw_{k'}, \end{aligned}$$

(6)

[0117] 其中 $\Theta_k = \{S_{l,m,k}^{(r)}\}_k$ 和 $\theta_k = \{s_{l,m,k}^{(r)}\}_k$ 以及 $k' = \kappa k$ 是 LTFS 仓的频率索引。 θ_k 的上述方程中的积分是对 $w_{k'}$ 的实部和虚部的简单二重积分。未被观测的逆滤波器被作为上述似然函数中的丢失数据处理并通过积分被排斥 (marginalize)。为了分析该函数，进一步假设：在给定 $\{S_{l,m,k}^{(r)}\}_k$ 的情况下， $\{\hat{s}_{l,m,k}^{(r)}\}_k$ 以及 $\{X_{l,m,k}^{(r)}\}_k$ 和 $w_{k'}$ 联合事件在统计上是独立的。在此情况下，上述方程 (6) 中的 $p\{w_{k'}, z_k | \Theta_k\}$ 能够被如下分成两个函数：

$$p\{w_{k'}, z_k | \Theta_k\} = p\{w_{k'}, \{x_{l,m,k}^{(r)}\}_k | \Theta_k\} p\{\{\hat{s}_{l,m,k}^{(r)}\}_k | \Theta_k\}.$$

[0118]

(7)

[0119] 前者是与室内声学有关的概率密度函数 (pdf)，即在给定源信号的情况下观测信号和逆滤波器的联合概率密度函数 (pdf)。后者是与由初始估计提供的信息有关的另一

个概率密度函数 (pdf), 即在给定源信号的情况下初始源信号估计的概率密度函数 (pdf)。第二分量能够被解释为在给定真实源信号的情况下语音特征的概率存在。它们将分别在下文中被称为“声学概率密度函数 (声学 pdf)”以及“源概率密度函数 (源 pdf)”。理想地, 逆传递函数 w_k 将 $x_{l, k'}$ 变换成 $s_{l, k'}$, 即 $w_k x_{l, k'} = s_{l, k'}$ 。但是, 在实际的声学环境中, 该方程由于诸如不足的逆滤波器长度和室内传递函数的波动这样的原因而可能包含一定误差 $\varepsilon_{l, k'}^{(a)} = w_k x_{l, k'} - s_{l, k'}$ 。所以, 声学 pdf 能够被认为是这种误差的概率密度函数 (pdf), 如 $p\{w_k, \{x_{l, m, k}^{(r)}\}_k | \Theta_k\} = p\{\{\varepsilon_{l, k'}^{(a)}\}_k | \Theta_k\}$ 。类似地, 源概率密度函数 (源 pdf) 能够被认为是误差 $\varepsilon_{l, m, k}^{(sr)} = s_{l, m, k}^{(r)} - s_{l, m, k}^{(sr)}$ 或者源信号和基于特征的信号之差的另一个概率密度函数 (pdf), 如 $p\{\{s_{l, m, k}^{(sr)}\}_k | \Theta_k\} = p\{\{\varepsilon_{l, m, k}^{(sr)}\}_k | \Theta_k\}$ 。为了简化起见, 假设这些误差是在给定 $\{s_{l, m, k}^{(sr)}\}_k$ 的情况下顺序地独立的随机过程。假设上述两个误差过程的实部和虚部以相同的方差相互独立并且能够单独地通过具有零平均值的高斯随机过程被建模。利用这些假设, 误差概率密度函数 (误差 pdf) 被表示为:

[0120]

$$p\{\{\varepsilon_{l, k'}^{(a)}\}_k | \Theta_k\} = \prod_l b_{l, k'}^{(a)} \exp\left\{-\frac{|\varepsilon_{l, k'}^{(a)}|^2}{2\sigma_{l, k'}^{(a)}}\right\},$$

[0121]

$$p\{\{\varepsilon_{l, m, k}^{(sr)}\}_k | \Theta_k\} = \prod_l \prod_m b_{l, m, k}^{(sr)} \exp\left\{-\frac{|\varepsilon_{l, m, k}^{(sr)}|^2}{2\sigma_{l, m, k}^{(sr)}}\right\},$$

(8)

[0122] 其中 $\sigma_{l, k'}^{(a)}$ 和 $\sigma_{l, m, k}^{(sr)}$ 分别是两个概率密度函数 (pdf) 的方差, 在下文中被称为声学环境不确定性和源信号不确定性。假设这两个值是基于语音信号和室内声学的特征被给定的。

[0123] EM 算法的说明

[0124] 期望最大化 (EM) 算法是找到最大化包括丢失数据的给定似然函数的一组参数的优化方法。这通过 A. P. Dempster, N. M. Laird, 和 D. B. Rubin 在“maximum likelihood from incorporate data via the EM algorithm”(Journal of the Royal Statistical Society, Series B, 39(1):1-38, 1977 年) 中被公开。通常, 似然函数被表示为:

[0125]

$$\mathcal{L}(\Theta) = p\{X = x | \Theta\},$$

[0126]

$$= \int p\{X = x, Y = y | \Theta\} dy,$$

(9)

[0127] 其中 $p\{\cdot | \Theta\}$ 表示在参数集 Θ 被给定并且 X 和 Y 是随机变量的条件下随机变量的概率密度函数 (pdf)。 $X = x$ 意味着 x 作为关于 X 的观测数据被给定。在上述似然函数

中,假设 Y 未被观测到,被称为丢失数据,因此利用 Y 排斥概率密度函数 (pdf)。通过找到最大化似然函数的参数集 $\Theta = \theta$ 的实现能够求解最大似然问题。

[0128] 根据期望最大化 (EM) 算法,利用辅助函数 $Q\{\Theta | \theta\}$ 的期望步骤 (E-step) 和最大化步骤 (M-step) 被分别定义为:

[0129] • E-step:

$$[0130] \quad Q\{\Theta | \theta\} = E_{\theta} \{\log p\{X = x, Y = y | \Theta\} | \Theta = \theta\},$$

$$[0131] \quad = \int p\{X = x, Y = y | \Theta = \theta\} \log p\{X = x, Y = y | \Theta\} d\gamma.$$

$$[0132] \quad \bullet \text{ M-step: } \tilde{\theta} = \arg \max_{\Theta} Q\{\Theta | \theta\}, \quad (10)$$

[0133] 其中在上述方程 (10) 的标记为“E-step”的上方方程中的 $E_{\theta} \{\cdot | \theta\}$ 是在 $\Theta = \theta$ 为固定的条件下的期望函数,其被更具体地定义为所述方程在 E-step 中的第二行。似然函数 $L\{\Theta\}$ 被显示为通过经由期望步骤 (E-step) 和最大化步骤 (M-step) 的一次迭代利用 $\Theta = \tilde{\theta}$ 更新 $\Theta = \theta$ 进行增加,其中在期望步骤 (E-step) 中计算 $Q\{\Theta | \theta\}$,而在最大化步骤 (M-step) 中获得最大化 $Q\{\Theta | \theta\}$ 的 $\Theta = \tilde{\theta}$ 。最大似然问题的解通过重复所述迭代获得。

[0134] 基于 EM 算法的解

[0135] 求解 θ_k 的上述方程 (6) 的一种有效方式是使用上述期望最大化 (EM) 算法。利用这种方法,为了语音去混响,利用辅助函数 $Q(\Theta_k | \theta_k)$ 的期望步骤 (E-step) 和最大化步骤 (M-step) 被分别定义为:

$$[0136] \quad Q(\Theta_k | \theta_k) = E_{\theta} \{\log p\{W_k, Z_k^{(r)} = z_k^{(r)} | \Theta_k\} | \Theta_k = \theta_k\}$$

$$[0137] \quad = \int p\{W_k = w_k, Z_k^{(r)} = z_k^{(r)} | \Theta_k = \theta_k\} \log p\{W_k = w_k, Z_k^{(r)} = z_k^{(r)} | \Theta_k\} d\gamma, \quad (11)$$

$$[0138] \quad \tilde{\theta}_k = \arg \max_{\Theta_k} Q(\Theta_k | \theta_k)$$

[0139] 其中假设 $z_k^{(r)}$ 是以下随机过程的实现:

$$[0140] \quad Z_k^{(r)} = \left\{ \{x_{l,m,k}^{(r)}\}_k, \{s_{l,m,k}^{(r)}\}_k \right\}.$$

[0141] 根据 EM 算法,对数似然 $\log p\{z_k^{(r)} | \theta_k\}$ 通过利用经由 EM 迭代获得的 $\tilde{\theta}_k$ 更新 θ_k 而增加,并且通过重复所述迭代收敛到平稳点解。

[0142] 解:

[0143] 取代直接计算 E-step 和 M-step,由于 $Q(\Theta_k | \theta_k) - Q(\theta_k | \theta_k)$ 在相同的 Θ_k 具有其最大值如 $Q(\Theta_k | \theta_k)$,因此它被分析。在 $Q(\Theta_k | \theta_k) - Q(\theta_k | \theta_k)$ 的某种排列之后并且仅仅提取包含 Θ_k 的项,从而获得以下函数。

[0144]

$$Q_e\{\Theta_k | \theta_k\}.$$

[0145]

$$= \sum_l \left\{ \frac{-|\tilde{w}_{k'} x_{l,k'} - S_{l,k'}|^2}{2\sigma_{l,k'}^{(a)}} + \sum_m \frac{-|\tilde{s}_{l,m,k}^{(r)} - S_{l,m,k}^{(r)}|^2}{2\sigma_{l,m,k}^{(sr)}} \right\}.$$

[0146] 其中

[0147]

$$\tilde{w}_{k'} = \frac{\sum_l s_{l,k'} x_{l,k'}^* / \sigma_{l,k'}^{(a)}}{\sum_l x_{l,k'} x_{l,k'}^* / \sigma_{l,k'}^{(a)}} \tag{12}$$

[0148] 其中 $a \neq r$ 意味着复数共轭。应该注意： Θ_k 最大化 $Q_{\Theta} \{\Theta_k | \theta_k\}$ 并且也最大化 $Q(\Theta_k | \theta_k)$ ， Θ_k 产生 $Q_{\Theta} \{\Theta_k | \theta_k\} > Q_{\Theta} \{\theta_k | \theta_k\}$ 并且也产生 $Q(\Theta_k | \theta_k) > Q(\theta_k | \theta_k)$ 。最大化 $Q_{\Theta} \{\Theta_k | \theta_k\}$ 的 Θ_k 能够通过以 $S_{l,m,k}^{(r)}$ 对它求微分、将它设置成零、以及求解所产生的联立方程而获得。但是，获得该解的计算成本相当高，这是因为需要针对每个 l 和 k 求解具有 M 个未知变量的该方程。

[0149] 替代地，为了以更有效的方式最大化上述方程 (12) 的 $Q_{\Theta} \{\Theta_k | \theta_k\}$ 引入以下假设。通过基于上述方程 (3) 对构成 LTFS 仓的 STFS 仓的幂求和，能够近似 LTFS 仓的幂，即：

[0150]

$$|s_{l,k'}|^2 \simeq \sum_{m=0}^{M-1} |s_{l,m,k}^{(r)}|^2. \tag{13}$$

[0151] 利用这种假设，通过上述方程 (12) 给定的 $Q_{\Theta} \{\Theta_k | \theta_k\}$ 能够被重写为：

$$Q_{\Theta} \{\Theta_k | \theta_k\} = \sum_l \sum_m \frac{-|LS_{m,k} \{ \{\tilde{w}_{k'} x_{l,k'}\}_l \} - S_{l,m,k}^{(r)}|^2}{2\sigma_{l,k'}^{(a)}} \tag{14}$$

$$+ \sum_l \sum_m \frac{-|\tilde{s}_{l,m,k}^{(r)} - S_{l,m,k}^{(r)}|^2}{2\sigma_{l,m,k}^{(sr)}}.$$

[0154] 通过对上述方程求微分并将它设置成零，能够如下针对通过上述方程 (11) 的 M -step 所给定的 $\tilde{\theta}_k$ 获得封闭形式解：

$$\tilde{s}_{l,m,k}^{(r)} = \frac{\sigma_{l,m,k}^{(sr)} LS_{m,k} \{ \{\tilde{w}_{k'} x_{l,k'}\}_l \} + \sigma_{l,k'}^{(a)} \tilde{s}_{l,m,k}^{(r)}}{\sigma_{l,k'}^{(a)} + \sigma_{l,m,k}^{(sr)}}. \tag{15}$$

[0156] 讨论：

[0157] 利用这种方法，通过依次重复地计算由上述方程 (12) 给定的 $\tilde{w}_{k'}$ 以及由上述方程

(15) 给定的 $\tilde{s}_{l,m,k}^{(r)}$ 来实现去混响。

[0158] 上述方程 (12) 中的 \tilde{w}_k 对应于在给定初始源信号估计为 $s_{l,k}$ 和观测信号为 $x_{l,k'}$ 的情况下通过传统 HERB 和 SBD 方法获得的去混响滤波器。

[0159] 上述方程 (15) 通过初始源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 和通过将 $x_{l,k'}$ 乘以 \tilde{w}_k 获得的源估计的加权平均来更新所述源估计。根据源信号不确定性和声学环境不确定性来确定权重。换言之，一次 EM 迭代通过集成基于源和室内声学特性获得的两种类型的源估计来精心制作源估计。

[0160] 从不同的观点来看，能够将通过上述方程 (12) 计算的逆滤波器估计 $w_k = \tilde{w}_k$ 视为最大化如下在 θ_k 是固定的条件下所定义的似然函数的逆滤波器估计，

$$\begin{aligned} [0161] \quad L\{w_k, \theta_k\} &= p\{w_k, z_k^{(r)} | \theta_k\} \\ [0162] \quad &= p\{w_k, \{x_{l,m,k}^{(r)}\}_k | \theta_k\} p\{\{s_{l,m,k}^{(r)}\}_k | \theta_k\} \end{aligned} \quad (16)$$

[0163] 其中采用与上述方程 (8) 相同的定义用于上述似然函数中的概率密度函数 (pdf)。另外，通过上述方程 (15) 计算的源信号估计 $\theta_k = \tilde{\theta}_k$ 也在逆滤波器估计 \tilde{w}_k 是固定的条件下最大化上述似然函数。所以，通过分别重复地计算上述方程 (12) 和 (15)，能够获得最大化上述似然函数的逆滤波器估计 \tilde{w}_k 和源信号估计 $\tilde{\theta}_k$ 。换言之，通过这种迭代优化算法，能够计算最大化上述似然函数的逆滤波器估计 \tilde{w}_k 。

[0164] 现在将参考附图描述本发明的所选择的实施例。对于本领域普通技术人员来说根据本公开内容将显而易见的是，本发明的实施例的以下描述仅仅被提供用于说明而不是用于限制如通过所附的权利要求及其等同物定义的本发明的目的。

[0165] 第一实施例：

[0166] 图 1 是图示根据本发明的第一实施例的、基于源和室内声学的概率模型的语音去混响设备的方框图。能够通过一组功能单元来实现语音去混响设备 10000，该组功能单元协作用于接收观测信号 $x[n]$ 的输入并产生波形信号 $\tilde{s}[n]$ 的输出。功能单元中的每一个都可以包括被构造和 / 或被编程用于执行预定功能的硬件和 / 或软件。术语“被适配”和“被配置”被用于描述被构造和 / 或被编程用于执行一个或多个期望功能的硬件和 / 或软件。语音去混响设备 10000 能够例如通过计算机或处理器来实现。语音去混响设备 10000 执行语音去混响操作。语音去混响方法能够通过要由计算机执行的程序来实现。

[0167] 语音去混响设备 10000 可以典型地包括初始化单元 1000、似然最大化单元 2000 和逆短时傅立叶变换单元 4000。初始化单元 1000 可以被适配用于接收能够是数字化波形信号的观测信号 $x[n]$ ，其中 n 是样本索引。数字化波形信号 $x[n]$ 可以包含具有未知混响程度的语音信号。通过诸如一个或多个麦克风的设备能够捕获该语音信号。初始化单元 1000 可以被适配用于从观测信号提取与源信号和声学环境有关的初始源信号估计和不确定性。初始化单元 1000 还可以被适配用于用公式表达初始源信号估计、源信号不确定性和声学环境不确定性的表示。针对所有索引 l 、 m 、 k 和 k' ，这些表示被列举为是数字化波形初始源信号估计的 $\tilde{s}[n]$ 、是表示源信号不确定性的方差或离差的 $\sigma_{l,m,k}^{(r)}$ 、以及是表示声学环境不确

定性的方差或离差的 $\sigma_{l,m,k}^{(s)}$ 。即,初始化单元 1000 可以被适配用于接收数字化波形信号 $x[n]$ 的输入作为观测信号并产生数字化波形初始源信号估计 $\hat{s}[n]$ 、表示源信号不确定性的方差或离差 $\sigma_{l,m,k}^{(s)}$ 、以及表示声学环境不确定性的方差或离差 $\sigma_{l,m,k}^{(a)}$ 。

[0168] 似然最大化单元 2000 可以与初始化单元 1000 协作。即,似然最大化单元 2000 可以被适配用于从初始化单元 1000 接收数字化波形初始源信号估计 $\hat{s}[n]$ 的输入、源信号不确定性 $\sigma_{l,m,k}^{(s)}$ 和声学环境不确定性 $\sigma_{l,m,k}^{(a)}$ 。似然最大化单元 2000 还可以被适配用于接收数字化波形观测信号 $x[n]$ 的另一个输入作为观测信号。 $\hat{s}[n]$ 是数字化波形初始源信号估计。 $\sigma_{l,m,k}^{(s)}$ 是表示源信号不确定性的第一方差。 $\sigma_{l,m,k}^{(a)}$ 是表示声学环境不确定性的第二方差。似然最大化单元 2000 还可以被适配用于确定最大化似然函数的源信号估计 θ_k ,其中参考数字化波形观测信号 $x[n]$ 、数字化波形初始源信号估计 $\hat{s}[n]$ 、表示源信号不确定性的第一方差 $\sigma_{l,m,k}^{(s)}$ 、以及表示声学环境不确定性的第二方差 $\sigma_{l,m,k}^{(a)}$ 来进行所述确定。通常,可以基于概率密度函数来定义似然函数,该概率密度函数根据参考源信号估计所定义的未知参数、表示室内传递函数的逆滤波器的丢失数据的第一随机变量、和参考观测信号和初始源信号估计所定义的观测数据第二随机变量进行评估。使用迭代优化算法来执行源信号估计 θ_k 的确定。

[0169] 迭代优化算法的典型实例可以包括但不限于上述期望最大化算法。在一个实例中,似然最大化单元 2000 可以被适配用于针对所有 k 搜索源信号 $\theta_k = \{\hat{s}_{l,m,k}^{(s)}\}$,并估计最大化如下定义的似然函数的源信号:

[0170]

$$L\{\theta_k\} = \log p\{z_k^{(r)} | \Theta_k = \theta_k\}$$

[0171] 其中 $z_k^{(r)} = \{\{x_{l,m,k}^{(r)}\}_k, \{\hat{s}_{l,m,k}^{(r)}\}_k\}$ 是此刻的短时观测 $x_{l,m,k}^{(r)}$ 和初始源信号估计 $\hat{s}_{l,m,k}^{(r)}$ 的联合事件。已经参考上述方程 (6) 描述了该函数的细节。因此,似然最大化单元 2000 可以被适配用于确定并输出最大化所述似然函数的源信号估计 $\hat{s}_{l,m,k}^{(r)}$ 。

[0172] 逆短时傅立叶变换单元 4000 可以与似然最大化单元 2000 协作。即,逆短时傅立叶变换单元 4000 可以被适配用于从似然最大化单元 2000 接收最大化似然函数的源信号估计 $\hat{s}_{l,m,k}^{(r)}$ 的输入。逆短时傅立叶变换单元 4000 还可以被适配用于将源信号估计 $\hat{s}_{l,m,k}^{(r)}$ 变换成数字化波形信号 $\tilde{s}[n]$ 并输出该数字化波形信号 $\tilde{s}[n]$ 。

[0173] 似然最大化单元 2000 能够通过相互协作以确定并输出最大化似然函数的源信号估计 $\hat{s}_{l,m,k}^{(r)}$ 的一组子功能单元来实现。图 2 是图示图 1 中所示的似然最大化单元 2000 的配置的方框图。在一种情况下,似然最大化单元 2000 还可以包括长时傅立叶变换单元 2100、更新单元 2200、STFS 到 LTFS 变换单元 2300、逆滤波器估计单元 2400、滤波单元 2500、LTFS 到 STFS 变换单元 2600、源信号估计和收敛检验单元 2700、短时傅立叶变换单元 2800、以及长时傅立叶变换单元 2900。这些单元协作以继续执行迭代操作,直到已经确定最大化似然函数的源信号估计为止。

[0174] 长时傅立叶变换单元 2100 被适配用于从初始化单元 1000 接收数字化波形观测信号 $x[n]$ 作为观测信号。长时傅立叶变换单元 2100 还被适配用于执行数字化波形观测信号 $x[n]$ 到作为长期傅立叶频谱 (LTFS) 的变换后的观测信号 $x_{1,k}$ 的长时傅立叶变换。

[0175] 短时傅立叶变换单元 2800 被适配用于从初始化单元 1000 接收数字化初始源信号估计 $\hat{s}[n]$ 。短时傅立叶变换单元 2800 被适配用于执行数字化波形初始源信号估计 $\hat{s}[n]$ 到初始源信号估计 $\hat{s}_{1,m,k}^{(n)}$ 的短时傅立叶变换。

[0176] 长时傅立叶变换单元 2900 被适配用于从初始化单元 1000 接收数字化波形初始源信号估计 $\hat{s}[n]$ 。长时傅立叶变换单元 2900 被适配用于执行数字化波形初始源信号估计 $\hat{s}[n]$ 到初始源信号估计 $\hat{s}_{1,k}$ 的长时傅立叶变换。

[0177] 更新单元 2200 与长时傅立叶变换单元 2900 和 STFS 到 LTFS 变换单元 2300 协作。更新单元 2200 被适配用于在迭代的初始步骤中从长时傅立叶变换单元 2900 接收初始源信号估计 $\hat{s}_{1,k}$ 并且还适配用于用从长时傅里叶变换单元 2900 接收的源信号估计 $\{\hat{s}_{1,k}\}_k$ 代替更新的源信号估计 θ_k 。更新单元 2200 此外被适配用于发送更新的源信号估计 θ_k 给逆滤波器估计单元 2400。更新单元 2200 还被适配用于在迭代的随后步骤中从 STFS 到 LTFS 变换单元 2300 接收源信号估计 $\hat{s}_{1,k}$ ，并且用于用从 STFS 到 LTFS 变换单元 2300 接收的源信号估计 $\{\hat{s}_{1,k}\}_k$ 代替更新的源信号估计 θ_k 。更新单元 2200 还被适配用于发送更新的源信号估计 θ_k 给逆滤波器估计单元 2400。

[0178] 逆滤波器估计单元 2400 与长时傅立叶变换单元 2100、更新单元 2200 和初始化单元 1000 协作。逆滤波器估计单元 2400 被适配用于从长时傅立叶变换单元 2100 接收观测信号 $x_{1,k}$ 。逆滤波器估计单元 2400 还被适配用于从更新单元 2200 接收更新的源信号估计 θ_k 。逆滤波器估计单元 2400 还被适配用于从初始化单元 1000 接收表示声学环境不确定性的第二方差 $\sigma_{1,k}^{(a)}$ 。逆滤波器估计单元 2400 进一步被适配用于根据上述方程 (12) 基于观测信号 $x_{1,k}$ 、更新的源信号估计 θ_k 和表示声学环境不确定性的第二方差 $\sigma_{1,k}^{(a)}$ 来计算逆滤波器估计 \tilde{w}_k 。逆滤波器估计单元 2400 进一步被适配用于输出逆滤波器估计 \tilde{w}_k 。

[0179] 滤波单元 2500 与长时傅立叶变换单元 2100 和逆滤波器估计单元 2400 协作。滤波单元 2500 被适配用于从长时傅立叶变换单元 2100 接收观测信号 $x_{1,k}$ 。滤波单元 2500 还被适配用于从逆滤波器估计单元 2400 接收逆滤波器估计 \tilde{w}_k 。滤波单元 2500 还被适配用于将观测信号 $x_{1,k}$ 应用于逆滤波器估计 \tilde{w}_k 以产生滤波源信号估计 $\bar{s}_{1,k}$ 。将观测信号 $x_{1,k}$ 应用于逆滤波器估计 \tilde{w}_k 的滤波过程的典型实例可以包括但不限于计算观测信号 $x_{1,k}$ 和逆滤波器估计 \tilde{w}_k 的乘积 $\tilde{w}_k x_{1,k}$ 。在这种情况下，通过观测信号 $x_{1,k}$ 和逆滤波器估计 \tilde{w}_k 的乘积 $\tilde{w}_k x_{1,k}$ 来给定滤波源信号估计 $\bar{s}_{1,k}$ 。

[0180] LTFS 到 STFS 变换单元 2600 与滤波单元 2500 协作。LTFS 到 STFS 变换单元 2600 被适配用于从滤波单元 2500 接收滤波源信号估计 $\bar{s}_{1,k}$ 。LTFS 到 STFS 变换单元 2600 进一步

被适配用于执行滤波源信号估计 $\bar{s}_{l,m,k}$ 到变换后的滤波源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的 LTFS 到 STFS 变换。当滤波过程是计算观测信号 $x_{l,k}$ 和逆滤波器估计 \bar{w}_k 的乘积 $\bar{w}_k x_{l,k}$ 时, LTFS 到 STFS 变换单元 2600 进一步被适配用于执行乘积 $\bar{w}_k x_{l,k}$ 到变换后的信号 $LS_{m,k} \{ \bar{w}_k x_{l,k} \}$ 的 LTFS 到 STFS 变换。在这种情况下, 乘积 $\bar{w}_k x_{l,k}$ 表示滤波源信号估计 $\bar{s}_{l,k}$, 以及变换后的信号 $LS_{m,k} \{ \bar{w}_k x_{l,k} \}$ 表示变换后的滤波源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 。

[0181] 源信号估计和收敛检验单元 2700 与 LTFS 到 STFS 变换单元 2600、短时傅立叶变换单元 2800、以及初始化单元 1000 协作。源信号估计和收敛检验单元 2700 被适配用于从 LTFS 到 STFS 变换单元 2600 接收变换后的滤波源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 。源信号估计和收敛检验单元 2700 还被适配用于从初始化单元 1000 接收表示源信号不确定性的第一方差 $\sigma_{l,k}^{(s)}$ 和表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 。源信号估计和收敛检验单元 2700 还被适配用于从短时傅立叶变换单元 2800 接收初始源信号估计 $\bar{s}_{l,m,k}^{(i)}$ 。源信号估计和收敛检验单元 2700 进一步被适配用于基于变换后的滤波源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 、表示源信号不确定性的第一方差 $\sigma_{l,m,k}^{(s)}$ 、表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 和初始源信号估计 $\bar{s}_{l,m,k}^{(i)}$ 估计源信号 $\bar{s}_{l,m,k}^{(r)}$, 其中根据上述方程 (15) 进行所述估计。

[0182] 源信号估计和收敛检验单元 2700 此外被适配用于例如通过比较当前已经被估计的源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的当前值和之前已经被估计的源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的前一值以及检验是否当前值偏离前一值小于一定的预定量来确定迭代过程的收敛状态。如果源信号估计和收敛检验单元 2700 确认源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的当前值偏离其前一值小于一定的预定量, 那么源信号估计和收敛检验单元 2700 识别出已经获得源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的收敛。如果源信号估计和收敛检验单元 2700 确认源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的当前值偏离其前一值不小于一定的预定量, 那么源信号估计和收敛检验单元 2700 识别出还未获得源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的收敛。

[0183] 作为修改有可能的是当迭代次数达到一定的预定值时终止迭代过程。即, 源信号估计和收敛检验单元 2700 已经确认迭代次数达到一定的预定值, 于是源信号估计和收敛检验单元 2700 识别出已经获得源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的收敛。如果源信号估计和收敛检验单元 2700 已经确认已经获得源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的收敛, 那么源信号估计和收敛检验单元 2700 向逆短时傅立叶变换单元 4000 提供源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 作为第一输出。如果源信号估计和收敛检验单元 2700 已经确认还未获得源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的收敛, 那么源信号估计和收敛检验单元 2700 向 STFS 到 LTFS 变换单元 2300 提供源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 作为第二输出。

[0184] STFS 到 LTFS 变换单元 2300 与源信号估计和收敛检验单元 2700 协作。STFS 到 LTFS 变换单元 2300 被适配用于从源信号估计和收敛检验单元 2700 接收源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 。STFS 到 LTFS 变换单元 2300 被适配用于执行源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 到变换后的源信号估计 $\bar{s}_{l,k}^{(r)}$ 的

STFS 到 LTFS 变换。

[0185] 在迭代操作的随后步骤中,更新单元 2200 从 STFS 到 LTFS 变换单元 2300 接收所述源信号估计 $\hat{s}_{l,k}$, 并用源信号估计 θ_k 代替 $\hat{s}_{l,k}$, 以及发送更新的源信号估计 θ_k 给逆滤波器估计单元 2400。

[0186] 上述迭代过程将被继续,直到源信号估计和收敛检验单元 2700 已经确认源信号估计 $\hat{s}_{l,k}^{(n)}$ 的收敛已经被获得。在迭代的初始步骤中,

[0187] 更新的源信号估计 $\theta_{k'}$ 是从长时傅立叶变换单元 2900 提供的 $\{\hat{s}_{l,k}\}_k$ 。在迭代的第二或随后步骤中,更新的源信号估计 $\theta_{k'}$ 是 $\{\hat{s}_{l,k}\}_k$ 。

[0188] 如果源信号估计和收敛检验单元 2700 已经确认已经获得源信号估计 $\hat{s}_{l,k}^{(n)}$ 的收敛,那么源信号估计和收敛检验单元 2700 向逆短时傅立叶变换单元 4000 提供源信号估计 $\hat{s}_{l,k}^{(n)}$ 作为第一输出。逆短时傅立叶变换单元 4000 可以被适配用于将源信号估计 $\hat{s}_{l,k}^{(n)}$ 变换成数字化波形信号 $\hat{s}[n]$ 并输出该数字化波形信号 $\hat{s}[n]$ 。

[0189] 将参考图 2 描述似然最大化单元 2000 的操作。

[0190] 在迭代的初始步骤中,将数字化波形观测信号 $x[n]$ 从初始化单元 1000 提供给长时傅立叶变换单元 2100。由长时傅立叶变换单元 2100 执行长时傅立叶变换以便数字化波形观测信号 $x[n]$ 被变换成作为长期傅立叶频谱 (LTFS) 的变换后的观测信号 $x_{l,k'}$ 。数字化波形初始源信号估计 $\hat{s}[n]$ 从初始化单元 1000 被提供给短时傅立叶变换单元 2800 和长时傅立叶变换单元 2900。由短时傅立叶变换单元 2800 执行短时傅立叶变换以便将数字化波形初始源信号估计 $\hat{s}[n]$ 变换成初始源信号估计 $\hat{s}_{l,k}^{(r)}$ 。由长时傅立叶变换单元 2900 执行长时傅立叶变换以便将数字化波形初始源信号估计 $\hat{s}[n]$ 变换成初始源信号估计 $\hat{s}_{l,k}$ 。

[0191] 从长时傅立叶变换单元 2900 提供初始源信号估计 $\hat{s}_{l,k}$ 给更新单元 2200。通过更新单元 2200 用源信号估计 $\theta_{k'}$ 代替初始源信号估计 $\{\hat{s}_{l,k}\}_k$ 。然后从更新单元 2200 提供初始源信号估计 $\theta_{k'} = \{\hat{s}_{l,k}\}_k$ 给逆滤波器估计单元 2400。观测信号 $x_{l,k'}$ 从长时傅立叶变换单元 2100 被提供给逆滤波器估计单元 2400。表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 从初始化单元 1000 被提供给逆滤波器估计单元 2400。逆滤波器估计单元 2400 基于观测信号 $x_{l,k'}$ 初始源信号估计 $\theta_{k'}$ 、以及表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 计算逆滤波器估计 $\tilde{w}_{l,k}$, 其中根据上述方程 (12) 进行所述计算。

[0192] 逆滤波器估计 $\tilde{w}_{l,k}$ 从逆滤波器估计单元 2400 被提供给滤波单元 2500。观测信号 $x_{l,k'}$ 进一步从长时傅立叶变换单元 2100 被提供给滤波单元 2500。滤波单元 2500 将逆滤波器估计 $\tilde{w}_{l,k}$ 应用于观测信号 $x_{l,k'}$ 以产生滤波源信号估计 $\hat{s}_{l,k}$ 。将观测信号 $x_{l,k'}$ 应用于逆滤波器估计 $\tilde{w}_{l,k}$ 的滤波过程的典型实例可以用于计算观测信号 $x_{l,k'}$ 和逆滤波器估计 $\tilde{w}_{l,k}$ 的乘积

$\tilde{w}_{k'} x_{1,k'}$ 。在这种情况下,通过观测信号 $x_{1,k'}$ 和逆滤波器估计 $\tilde{w}_{k'}$ 的乘积 $\tilde{w}_{k'} x_{1,k'}$ 来给定滤波源信号估计 $\bar{s}_{1,k'}$ 。

[0193] 从滤波单元 2500 提供滤波源信号估计 $\bar{s}_{1,k'}$ 给 LTFS 到 STFS 变换单元 2600。由 LTFS 到 STFS 变换单元 2600 执行 LTFS 到 STFS 变换以便滤波源信号估计 $\bar{s}_{1,k'}$ 被变换成变换后的滤波源信号估计 $\bar{s}_{1,m,k}^{(r)}$ 。当滤波过程用于计算观测信号 $x_{1,k'}$ 和逆滤波器估计 $\tilde{w}_{k'}$ 的乘积 $\tilde{w}_{k'} x_{1,k'}$ 时,乘积 $\tilde{w}_{k'} x_{1,k'}$ 被变换成变换后的信号 $LS_{m,k} \{ \{ \tilde{w}_{k'} x_{1,k'} \} \}$ 。

[0194] 从 LTFS 到 STFS 变换单元 2600 提供变换后的滤波源信号估计 $\bar{s}_{1,m,k}^{(r)}$ 给源信号估计和收敛检验单元 2700。从初始化单元 1000 提供表示源信号不确定性的第一方差 $\sigma_{1,m,k}^{(s)}$ 和表示声学环境不确定性的第二方差 $\sigma_{1,m,k}^{(a)}$ 给源信号估计和收敛检验单元 2700。从短时傅立叶变换单元 2800 提供初始源信号估计 $\hat{s}_{1,m,k}^{(i)}$ 给源信号估计和收敛检验单元 2700。源信号估计和收敛检验单元 2700 基于变换后的滤波源信号估计 $\bar{s}_{1,m,k}^{(r)}$ 、表示源信号不确定性的第一方差 $\sigma_{1,m,k}^{(s)}$ 、表示声学环境不确定性的第二方差 $\sigma_{1,m,k}^{(a)}$ 和初始源信号估计 $\hat{s}_{1,m,k}^{(i)}$ 计算源信号估计 $\bar{s}_{1,m,k}^{(r)}$, 其中根据上述方程 (15) 进行所述估计。

[0195] 在迭代的初始步骤中,从源信号估计和收敛检验单元 2700 提供源信号估计 $\bar{s}_{1,m,k}^{(r)}$ 给 STFS 到 LTFS 变换单元 2300 以便将源信号估计 $\bar{s}_{1,m,k}^{(r)}$ 变换成变换后的源信号估计 $\tilde{s}_{1,k}$ 。从 STFS 到 LTFS 变换单元 2300 提供变换后的源信号估计 $\tilde{s}_{1,k}$ 给更新单元 2200。更新单元 2200 用源信号估计 $\theta_{k'}$ 代替变换后的源信号估计 $\tilde{s}_{1,k}$ 。从更新单元 2200 提供更新的源信号估计 $\theta_{k'}$ 给逆滤波器估计单元 2400。

[0196] 在迭代的第二或随后步骤中,从更新单元 2200 提供源信号估计 $\theta_{k'} = \{ \tilde{s}_{1,k} \}$ 给逆滤波器估计单元 2400。观测信号 $x_{1,k'}$ 也从长时傅立叶变换单元 2100 被提供给逆滤波器估计单元 2400。从初始化单元 1000 提供表示声学环境不确定性的第二方差 $\sigma_{1,m,k}^{(a)}$ 给逆滤波器估计单元 2400。逆滤波器估计单元 2400 基于观测信号 $x_{1,k'}$ 、更新的源信号估计 $\theta_{k'} = \{ \tilde{s}_{1,k} \}$ 和表示声学环境不确定性的第二方差 $\sigma_{1,m,k}^{(a)}$ 计算更新的逆滤波器估计 $\tilde{w}_{k'}$, 其中根据上述方程 (12) 进行所述计算。

[0197] 从逆滤波器估计单元 2400 提供更新的逆滤波器估计 $\tilde{w}_{k'}$ 给滤波单元 2500。观测信号 $x_{1,k'}$ 进一步被从长时傅立叶变换单元 2100 提供给滤波单元 2500。滤波单元 2500 将观测信号 $x_{1,k'}$ 应用于更新的逆滤波器估计 $\tilde{w}_{k'}$ 以产生滤波源信号估计 $\bar{s}_{1,k}$ 。

[0198] 从滤波单元 2500 提供更新的滤波源信号估计 $\bar{s}_{1,k}$ 给 LTFS 到 STFS 变换单元 2600。LTFS 到 STFS 变换单元 2600 执行 LTFS 到 STFS 变换以便将更新的滤波源信号估计 $\bar{s}_{1,k}$ 变换成变换后的滤波源信号估计 $\bar{s}_{1,m,k}^{(r)}$ 。

[0199] 从 LTFS 到 STFS 变换单元 2600 提供更新的滤波源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 给源信号估计和收敛检验单元 2700。还从初始化单元 1000 提供表示源信号不确定性的第一方差 $\sigma_{l,m,k}^{(r)}$ 和表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 给源信号估计和收敛检验单元 2700。从短时傅立叶变换单元 2800 提供更新的初始源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 给源信号估计和收敛检验单元 2700。源信号估计和收敛检验单元 2700 基于变换后的滤波源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 、表示源信号不确定性的第一方差 $\sigma_{l,m,k}^{(r)}$ 和表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 和初始源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 计算源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ ，其中根据上述方程 (15) 进行所述估计。比较当前已经被估计的源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的当前值与之前已经被估计的源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的前一值。源信号估计和收敛检验单元 2700 验证是否当前值偏离前一值小于一定的预定量。

[0200] 如果源信号估计和收敛检验单元 2700 确认源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的当前值偏离其前一值小于一定的预定量，那么源信号估计和收敛检验单元 2700 识别出已经获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛。源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 作为第一输出从源信号估计和收敛检验单元 2700 被提供给逆短时傅立叶变换单元 4000。逆短时傅立叶变换单元 4000 将源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 变换成数字化波形源信号估计 $\tilde{s}[n]$ 。

[0201] 如果源信号估计和收敛检验单元 2700 确认源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的当前值不偏离其前一值小于一定的预定量，那么源信号估计和收敛检验单元 2700 识别出还未获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛。从源信号估计和收敛检验单元 2700 提供源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 给 STFS 到 LTFS 变换单元 2300 以便将源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 变换成变换后的源信号估计 $\tilde{s}_{l,k}^{(r)}$ 。从 STFS 到 LTFS 变换单元 2300 提供变换后的源信号估计 $\tilde{s}_{l,k}^{(r)}$ 给更新单元 2200。更新单元 2200 用源信号估计 θ_k 代替变换后的源信号估计 $\tilde{s}_{l,k}^{(r)}$ 。从更新单元 2200 提供更新的源信号估计 θ_k 给逆滤波器估计单元 2400。

[0202] 作为修改有可能的是当迭代次数达到一定的预定值时终止迭代过程。即，已经通过源信号估计和收敛检验单元 2700 确认迭代次数达到一定的预定值，那么源信号估计和收敛检验单元 2700 识别出已经获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛。如果已经通过源信号估计和收敛检验单元 2700 确认已经获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛，那么源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 作为第一输出从源信号估计和收敛检验单元 2700 被提供给逆短时傅立叶变换单元 4000。如果已经通过源信号估计和收敛检验单元 2700 确认还未获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛，那么源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 作为第二输出从源信号估计和收敛检验单元 2700 被提供给 STFS 到 LTFS 变换单元 2300 以便然后将源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 变换成变换后的源信号估计 $\tilde{s}_{l,k}^{(r)}$ 。进一步用源信号估计 θ_k 代替变换后的源信号估计 $\tilde{s}_{l,k}^{(r)}$ 。

[0203] 上述迭代过程将被继续,直到通过源信号估计和收敛检验单元 2700 已经确认已经获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛。在迭代的初始步骤中,更新的源信号估计 θ_k 是从长时傅立叶变换单元 2900 提供的。在迭代的第二或随后步骤中,更新的源信号估计 θ_k 是 $\{\tilde{s}_{l,m,k}\}$ 。

[0204] 如果已经通过源信号估计和收敛检验单元 2700 确认已经获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛,那么源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 作为第一输出从源信号估计和收敛检验单元 2700 被提供给逆短时傅立叶变换单元 4000。逆短时傅立叶变换单元 4000 将源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 变换成数字化波形源信号估计 $\tilde{s}[n]$ 并输出该数字化波形源信号估计 $\tilde{s}[n]$ 。

[0205] 图 3A 是图示在图 2 中所示的 STFS 到 LTFS 变换单元 2300 的配置的方框图。STFS 到 LTFS 变换单元 2300 可以包括逆短时傅立叶变换单元 2310 和长时傅立叶变换单元 2320。逆短时傅立叶变换单元 2310 与源信号估计和收敛检验单元 2700 协作。逆短时傅立叶变换单元 2310 被适配用于从源信号估计和收敛检验单元 2700 接收源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。逆短时傅立叶变换单元 2310 进一步被适配用于将源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 变换成作为输出的数字化波形源信号估计 $\tilde{s}[n]$ 。

[0206] 长时傅立叶变换单元 2320 与逆短时傅立叶变换单元 2310 协作。长时傅立叶变换单元 2320 被适配用于从逆短时傅立叶变换单元 2310 接收数字化波形源信号估计 $\tilde{s}[n]$ 。长时傅立叶变换单元 2320 进一步被适配用于将数字化波形源信号估计 $\tilde{s}[n]$ 变换成作为输出的变换后的源信号估计 $\tilde{s}_{l,m,k}$ 。

[0207] 图 3B 是图示在图 2 中所示的 LTFS 到 STFS 变换单元 2600 的配置的方框图。LTFS 到 STFS 变换单元 2600 可以包括逆长时傅立叶变换单元 2610 和短时傅立叶变换单元 2620。逆长时傅立叶变换单元 2610 与滤波单元 2500 协作。逆长时傅立叶变换单元 2610 被适配用于从滤波单元 2500 接收滤波源信号估计 $\tilde{s}_{l,k}$ 。逆长时傅立叶变换单元 2610 进一步被适配用于将滤波源信号估计 $\tilde{s}_{l,k}$ 变换成作为输出的数字化波形滤波源信号估计 $\tilde{s}[n]$ 。

[0208] 短时傅立叶变换单元 2620 与逆长时傅里叶变换单元 2610 协作。短时傅立叶变换单元 2620 被适配用于从逆长时傅立叶变换单元 2610 接收数字化波形滤波源信号估计 $\tilde{s}[n]$ 。短时傅里叶变换单元 2620 进一步被适配用于将数字化波形滤波源信号估计 $\tilde{s}[n]$ 变换成作为输出的变换后的滤波源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。

[0209] 图 4A 是图示图 2 中所示的长时傅立叶变换单元 2100 的配置的方框图。长时傅立叶变换单元 2100 可以包括开窗单元 2110 和离散傅立叶变换单元 2120。开窗单元 2100 被适配用于接收数字化波形观测信号 $x[n]$ 。开窗单元 2100 进一步被适配用于如以下所给定的那样将分析窗函数 $g[n]$ 重复地应用于数字化波形观测信号 $x[n]$ ：

$$[0210] \quad x_1[n] = g[n]x[n_1+n].$$

[0211] 其中 n_1 是长时间帧 1 所开始的样本索引。开窗单元 2110 被适配用于针对所有 1 产生分段波形观测信号 $x_1[n]$ 。

[0212] 离散傅立叶变换单元 2120 与开窗单元 2110 协作。离散傅立叶变换单元 2120 被适配用于从开窗单元 2110 接收分段波形观测信号 $x_1[n]$ 。离散傅立叶变换单元 2120 进一步被适配用于执行分段波形信号 $x_1[n]$ 中的每一个到变换后的观测信号 $x_{1,k}$ 的如下给定的 K 点离散傅立叶变换：

[0213]

$$x_{1,k} = 1/K \sum_{n=0}^{K-1} x_1[n] e^{-j2\pi kn/K}$$

[0214] 图 4B 是图示图 3B 中所示的逆长时傅立叶变换单元 2610 的配置的方框图。逆长时傅立叶变换单元 2610 可以包括逆离散傅立叶变换单元 2612 以及重叠相加合成单元 2614。逆离散傅立叶变换单元 2612 与滤波单元 2500 协作。逆离散傅立叶变换单元 2612 被适配用于接收滤波源信号估计 $\bar{s}_{1,k}$ 。逆离散傅立叶变换单元 2612 进一步被适配用于应用滤波源信号估计 $\bar{s}_{1,k}$ 的每个帧到作为如下所给定的输出的分段波形滤波源信号估计 $\bar{s}_1[n]$ 的相应逆离散傅立叶变换：

[0215]
$$\bar{s}_1[n] = \sum_{k=0}^{K-1} \bar{s}_{1,k} e^{j2\pi kn/K}$$

[0216] 重叠相加合成单元 2614 与逆离散傅立叶变换单元 2612 协作。重叠相加合成单元 2614 被适配用于从逆离散傅立叶变换单元 2612 接收分段波形滤波源信号估计 $\bar{s}_1[n]$ 。重叠相加合成单元 2614 进一步被适配用于根据重叠相加合成技术利用重叠相加合成窗 $g_s[n]$ 针对所有 l 连接或合成分段波形滤波源信号估计 $\bar{s}_1[n]$ ，以便获得如下给定的数字化波形滤波源信号估计 $\bar{s}[n]$ 。

[0217]
$$\bar{s}[n] = \sum_k g_s[n-n_l] \bar{s}_1[n-n_l]$$

[0218] 图 5A 是图示在图 3B 中所示的短时傅立叶变换单元 2620 的配置的方框图。短时傅立叶变换单元 2620 可以包括开窗单元 2622 和离散傅立叶变换单元 2624。开窗单元 2622 与逆长时傅立叶变换单元 2610 协作。开窗单元 2622 被适配用于从逆长时傅立叶变换单元 2610 接收数字化波形滤波源信号估计 $\bar{s}[n]$ 。开窗单元 2622 进一步被适配用于将分析窗函数 $g^{(\tau)}[n]$ 以窗偏移 τ 重复地应用于数字化波形滤波源信号估计 $\bar{s}[n]$ 以便产生如下给定的分段滤波源信号估计 $\bar{s}_{l,m}[n]$ 。

[0219]
$$\bar{s}_{l,m}[n] = g^{(\tau)}[n] \bar{s}[n_{l,m} + n]$$

[0220] 其中 $n_{l,m}$ 是时间帧所开始的样本索引。针对所有，开窗单元 2622 产生分段波形滤波源信号估计 $\bar{s}_{l,m}[n]$ 。

[0221] 离散傅立叶变换单元 2624 与开窗单元 2622 协作。离散傅立叶变换单元 2624 被

适配用于从开窗单元 2622 接收分段波形滤波源信号估计 $\tilde{s}_{l,m}[n]$ 。离散傅立叶变换单元 2624 进一步被适配用于执行分段波形滤波源信号估计 $\tilde{s}_{l,m}[n]$ 中的每一个到如下给定的变换后的滤波源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的 $K^{(r)}$ 点离散傅立叶变换。

[0222]

$$\tilde{s}_{l,m,k}^{(r)} = 1/K^{(r)} \sum_{n=0}^{K^{(r)}-1} \tilde{s}_{l,m}[n] e^{-j2\pi n l k^{(r)}}$$

[0223] 图 5B 是图示在图 3A 中所示的逆短时傅立叶变换单元 2310 的配置的方框图。逆短时傅立叶变换单元 2310 可以包括逆离散傅立叶变换单元 2312 和重叠相加合成单元 2314。逆离散傅立叶变换单元 2312 与源信号估计和收敛检验单元 2700 协作。逆离散傅立叶变换单元 2312 被适配用于从源信号估计和收敛检验单元 2700 接收源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。逆离散傅立叶变换单元 2312 进一步被适配用于将相应的逆离散傅立叶变换应用于源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的每个帧并产生如下给定的分段波形源信号估计 $\tilde{s}_{l,m}[n]$ 。

[0224]

$$\tilde{s}_{l,m}[n] = \sum_{k=0}^{K^{(r)}-1} \tilde{s}_{l,m,k}^{(r)} e^{j2\pi n l k^{(r)}}$$

[0225] 重叠相加合成单元 2314 与逆离散傅立叶变换单元 2312 协作。重叠相加合成单元 2314 被适配用于从逆离散傅立叶变换单元 2312 接收分段波形源信号估计 $\tilde{s}_{l,m}[n]$ 。重叠相加合成单元 2314 进一步被适配用于针对所有 l 和 m 基于重叠相加合成技术利用合成窗 $g_s^{(r)}[n]$ 连接或合成分段波形源信号估计 $\tilde{s}_{l,m}[n]$ 以便获得如下给定的数字化波形源信号估计 $\tilde{s}[n]$ 。

[0226]
$$\tilde{s}[n] = \sum_{l,m} g_s^{(r)}[n - n_{l,m}] \tilde{s}_{l,m}[n - n_{l,m}]$$

[0227] 初始化单元 1000 被适配用于执行三个操作,即初始源信号估计、源信号不确定性确定和声学环境不确定性确定。如上所述,初始化单元 1000 被适配用于接收数字化波形观测信号 $x[n]$ 并产生表示源信号不确定性的第一方差 $\sigma_{l,m,k}^{(s)}$ 、表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 以及数字化波形初始源信号估计 $\tilde{s}[n]$ 。详细地,初始化单元 1000 被适配用于执行初始源信号估计,该估计根据数字化波形观测信号 $x[n]$ 产生数字化波形初始源信号估计 $\tilde{s}[n]$ 。初始化单元 1000 进一步被适配用于执行源信号不确定性确定,该确定根据数字化波形观测信号 $x[n]$ 产生表示源信号不确定性的第一方差 $\sigma_{l,m,k}^{(s)}$ 。初始化单元 1000 此外被适配用于执行声学环境不确定性确定,该确定根据数字化波形观测信号 $x[n]$ 产生表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 。

[0228] 初始化单元 1000 可以包括三个功能子单元,即执行初始源信号估计的初始源信号估计单元 1100、执行源信号不确定性确定的源信号不确定性确定单元 1200、以及执行声学环境不确定性确定的声学环境不确定性确定单元 1300。图 6 是图示包括在图 1 中所示的初始化单元 1000 中的初始源信号估计单元 1100 的配置的方框图。图 7 是图示包括在图 1 中所示的初始化单元 1000 中的源信号不确定性确定单元 1200 的配置的方框图。图 8 是图示包括在图 1 中所示的初始化单元 1000 中的声学环境不确定性确定单元 1300 的配置的方框图。

[0229] 参考图 6,初始源信号估计单元 1100 可以进一步包括短时傅立叶变换单元 1110、基频估计单元 1120 和自适应谐波滤波单元 1130。短时傅立叶变换单元 1110 被适配用于接收数字化波形观测信号 $x[n]$ 。短时傅立叶变换单元 1110 被适配用于执行数字化波形观测信号 $x[n]$ 到作为输出的变换后的观测信号 $x_{l,m,k}^{(r)}$ 的短时傅立叶变换。

[0230] 基频估计单元 1120 与短时傅立叶变换单元 1110 协作。基频估计单元 1120 被适配用于从短时傅立叶变换单元 1110 接收变换后的观测信号 $x_{l,m,k}^{(r)}$ 。基频估计单元 1120 进一步被适配用于估计来自变换后的观测信号 $x_{l,m,k}^{(r)}$ 的每个短时间帧的基频 $f_{1,m}$ 和发声量度 $v_{1,m}$ 。

[0231] 自适应谐波滤波单元 1130 与短时傅立叶变换单元 1110 和基频估计单元 1120 协作。自适应谐波滤波单元 1130 被适配用于从短时傅立叶变换单元 1110 接收变换后的观测信号 $x_{l,m,k}^{(r)}$ 。自适应谐波滤波单元 1130 还被适配用于从基频估计单元 1120 接收基频 $f_{1,m}$ 和发声量度 $v_{1,m}$ 。自适应谐波滤波单元 1130 还被适配用于基于基频 $f_{1,m}$ 和发声量度 $v_{1,m}$ 增强 $x_{l,m,k}^{(r)}$ 的谐波结构以便谐波结构的增强产生所得到的数字化波形初始源信号估计 $\hat{s}[n]$ 作为输出。本实例的过程流程在由 Tomohiro Nakatani、Masato Miyoshi 和 Keisuke Kinoshita 在 Speech Enhancement (Benesty, J. Makino, S., 和 Chen, J. Eds), Chapter 11, 第 247-270 页, 2005 年春) 中发表的“Single Microphone Blind Dereverberation”中被详细公开。

[0232] 参考图 7,源信号不确定性确定单元 1200 可以进一步包括短时傅立叶变换单元 1110、基频估计单元 1120 和源信号不确定性确定子单元 1140。短时傅立叶变换单元 1110 被适配用于接收数字化波形观测信号 $x[n]$ 。短时傅立叶变换单元 1110 被适配用于执行数字化波形观测信号 $x[n]$ 到作为输出的变换后的观测信号 $x_{l,m,k}^{(r)}$ 的短时傅立叶变换。

[0233] 基频估计单元 1120 与短时傅立叶变换单元 1110 协作。基频估计单元 1120 被适配用于从短时傅立叶变换单元 1110 接收变换后的观测信号 $x_{l,m,k}^{(r)}$ 。基频估计单元 1120 进一步被适配用于估计来自变换后的观测信号 $x_{l,m,k}^{(r)}$ 的每个短时间帧的基频 $f_{1,m}$ 和发声量度 $v_{1,m}$ 。

[0234] 源信号不确定性确定子单元 1140 与基频估计单元 1120 协作。源信号不确定性确定子单元 1140 被适配用于从基频估计单元 1120 接收基频 $f_{1,m}$ 和发声量度 $v_{1,m}$ 。源信号不确定性确定子单元 1140 进一步被适配用于根据基频 $f_{1,m}$ 和发声量度 $v_{1,m}$ 确定表示源信号不确定性的第一方差 $\sigma_{l,m,k}^{(r)}$ 。如下给定表示源信号不确定性的第一方差 $\sigma_{l,m,k}^{(r)}$ 。

[0235]

$$\sigma_{l,m,k}^{(r)} = \begin{cases} G \left\{ \frac{v_{l,m} - \delta}{\max_{l,m} \{v_{l,m}\} - \delta} \right\} & \text{if } v_{l,m} > \delta \text{ 并且 } k \text{ 是谐波频率} \\ \infty & \text{if } v_{l,m} > \delta \text{ 并且 } k \text{ 不是谐波频率} \\ G \left\{ \frac{v_{l,m} - \delta}{\min_{l,m} \{v_{l,m}\} - \delta} \right\} & \text{if } v_{l,m} \leq \delta \end{cases} \quad (17)$$

[0236] 其中 $G\{u\}$ 是被定义为例如具有某些正的常量“a”和“b”的 $G\{u\} = e^{-a(u-b)}$ 的归一化函数, 并且谐波频率意味着用于基频及其倍数之一的频率索引。

[0237] 参考图 8, 声学环境不确定性确定单元 1300 可以包括声学环境不确定性确定子单元 1150。声学环境不确定性确定子单元 1150 被适配用于接收数字化波形观测信号 $x[n]$ 。声学环境不确定性确定子单元 1150 进一步被适配用于产生表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 。在一种典型情况下, 对于所有 l 和 k' 来说, 第二方差 $\sigma_{l,k}^{(a)}$ 能够是常量, 即如在图 8 中所示, $\sigma_{l,k} = 1$ 。

[0238] 所述混响信号能够通过修改的包括执行反馈过程的反馈回路的语音去混响设备 20000 更有效地去混响。根据反馈过程的流程, 通过利用反馈回路迭代相同的处理流程, 能够提高源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的质量。尽管只有数字化波形观测信号 $x[n]$ 在初始步骤中被用作流程的输入, 但是已经被包含在前一步骤中的源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 也在随后的步骤中被用作输入。更优选的是使用源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 而不是使用观测信号 $x[n]$ 来对源概率密度函数 (源 pdf) 的参数 $\tilde{s}_{l,m,k}^{(r)}$ 和 $\sigma_{l,m,k}^{(r)}$ 进行估计。

[0239] 第二实施例:

[0240] 图 9 是图示根据本发明的第二实施例进一步包括反馈回路的另一语音去混响设备的配置的方框图。修改的语音去混响设备 20000 可以包括初始化单元 1000、似然最大化单元 2000、收敛检验单元 3000、和逆短时傅立叶变换单元 4000。初始化单元 1000、似然最大化单元 2000 和逆短时傅立叶变换单元 4000 的配置和操作如上所述。在本实施例中, 在似然最大化单元 2000 和逆短时傅立叶变换单元 4000 之间附加地引入收敛检验单元 3000 以便收敛检验单元 3000 检验已经从似然最大化单元 2000 输出的源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛。如果收敛检验单元 3000 识别出已经获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛, 那么收敛检验单元 3000 发送源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 给逆短时傅立叶变换单元 4000。如果收敛检验单元 3000 识别出还未获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛, 那么收敛检验单元 3000 发送源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 给初始化单元 1000。以下描述将聚焦于第二实施例与第一实施例的差异。

[0241] 收敛检验单元 3000 与初始化单元 1000 和似然最大化单元 2000 协作。收敛检验单元 3000 被适配用于从似然最大化单元 2000 接收源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。收敛检验单元 3000 进一步被适配用于例如通过验证是否源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的当前更新的值偏离源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的前一值小于一定的预定量来确定迭代过程的收敛状态。如果收敛检验单元 3000 确认源信

号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的当前更新的值偏离源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的前一值小于一定的预定量,那么收敛检验单元 3000 识别出已经获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛。如果收敛检验单元 3000 确认源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的当前更新的值不偏离源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的前一值小于一定的预定量,那么收敛检验单元 3000 识别出还未获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛。

[0242] 作为修改有可能的是当反馈或迭代次数达到一定的预定值时终止所述反馈过程。当收敛检验单元 3000 已经确认源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛已经被获得时,那么收敛检验单元 3000 发送源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 给逆短时傅立叶变换单元 4000。如果收敛检验单元 3000 已经确认还未获得源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的收敛,那么收敛检验单元 3000 将源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 作为输出提供给初始化单元 1000 以执行上述迭代的进一步的步骤。

[0243] 收敛检验单元 3000 提供反馈回路给初始化单元 1000。即,初始化单元 1000 与收敛检验单元 1000 协作。因此,初始化单元 1000 需要适配于反馈回路。根据第一实施例,初始化单元 1000 包括初始源信号估计单元 1100、源信号不确定性确定单元 1200、和声学环境不确定性确定单元 1300。根据第二实施例,修改的初始化单元 1000 包括修改的初始源信号估计单元 1400、修改的源信号不确定性确定单元 1500、和声学环境不确定性确定单元 1300。以下描述将聚焦于修改的初始源信号估计单元 1400、和修改的源信号不确定性确定单元 1500。

[0244] 图 10 是图示包括在图 9 中所示的初始化单元 1000 中的修改的初始源信号估计单元 1400 的配置的方框图。修改的初始源信号估计单元 1400 可以进一步包括短时傅立叶变换单元 1110、基频估计单元 1120、自适应谐波滤波单元 1130、和信号切换单元 1160。信号切换单元 1160 的增加能够提高数字化波形初始源信号估计 $\hat{s}[n]$ 的精度。

[0245] 短时傅立叶变换单元 1110 被适配用于接收数字化波形观测信号 $x[n]$ 。短时傅立叶变换单元 1110 被适配用于执行数字化波形观测信号 $x[n]$ 到作为输出的变换后的观测信号 $x_{l,m,k}^{(r)}$ 的短时傅立叶变换。信号切换单元 1160 与短时傅立叶变换单元 1110 和收敛检验单元 3000 协作。信号切换单元 1160 被适配用于从短时傅立叶变换单元 1110 接收变换后的观测信号 $x_{l,m,k}^{(r)}$ 。信号切换单元 1160 被适配用于从收敛检验单元 3000 接收源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。信号切换单元 1160 被适配用于执行第一选择操作以产生第一输出。信号切换单元 1160 还被适配用于执行第二选择操作以产生第二输出。第一和第二选择操作是相互独立的。第一选择操作是用于选择变换后的观测信号 $x_{l,m,k}^{(r)}$ 和源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 之一。在一种情况下,第一选择操作可以用于在除了一个或多个限制步骤之外的迭代的所有步骤中选择变换后的观测信号 $x_{l,m,k}^{(r)}$ 。例如,第一选择操作可以是用于在除了迭代的最后一个或两个步骤之外的迭代的所有步骤中选择变换后的观测信号 $x_{l,m,k}^{(r)}$ 以及仅仅在最后一个或两个步骤中选择源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。在一种情况下,第二选择操作可以是用于在除初始步骤之外的迭代的所有步骤中选择源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。在迭代的初始步骤中,信号切换单元 1160 仅接收变换后的观测信号 $x_{l,m,k}^{(r)}$ 并选择变换后的观测信号 $x_{l,m,k}^{(r)}$ 。更优选的是就基频 $f_{1,m}$ 和发声量度 $v_{1,m}$ 两者的估计而言使用源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 而不使用变换后的观测信号 $x_{l,m,k}^{(r)}$ 。

[0246] 信号切换单元 1160 执行第一选择操作并产生第一输出。信号切换单元 1160 执行第二选择操作并产生第二输出。

[0247] 基频估计单元 1120 与信号切换单元 1160 协作。基频估计单元 1120 被适配用于从信号切换单元 1160 接收第二输出。即,基频估计单元 1120 被适配用于在迭代的初始或第一步骤中从信号切换单元 1160 接收变换后的观测信号 $x_{l,m,k}^{(r)}$ 以及在迭代的第二或随后步骤中从信号切换单元 1160 接收源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。基频估计单元 1120 进一步被适配用于基于变换后的观测信号 $x_{l,m,k}^{(r)}$ 或源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 估计每个短时间帧的基频 $f_{l,m}$ 及其发声量度 $v_{l,m}$ 。

[0248] 自适应谐波滤波单元 1130 与信号切换单元 1160 和基频估计单元 1120 协作。自适应谐波滤波单元 1130 被适配用于从信号切换单元 1160 接收第一输出以及还从基频估计单元 1120 接收基频 $f_{l,m}$ 和发声量度 $v_{l,m}$ 。即,自适应谐波滤波单元 1130 被适配用于在除迭代的最后一个或两个步骤之外的迭代的所有步骤中从信号切换单元 1160 接收变换后的观测信号 $x_{l,m,k}^{(r)}$ 。自适应谐波滤波单元 1130 还被适配用于在迭代的最后一个或两个步骤中从信号切换单元 1160 接收源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。自适应谐波滤波单元 1130 还被适配用于在迭代的所有步骤中从基频估计单元 1120 接收基频 $f_{l,m}$ 和发声量度 $v_{l,m}$ 。自适应谐波滤波单元 1130 还被适配用于基于基频 $f_{l,m}$ 和发声量度 $v_{l,m}$ 增强观测信号 $x_{l,m,k}^{(r)}$ 或源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的谐波结构。增强操作产生估计精度被提高的数字化波形初始源信号估计 $\hat{s}[n]$ 。

[0249] 如上所述,更优选的是就基频 $f_{l,m}$ 和发声量度 $v_{l,m}$ 两者的估计而言基频估计单元 1120 使用源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 而不是使用观测信号 $x_{l,m,k}^{(r)}$ 。因此在迭代的第二或随后步骤中提供源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 而不是观测信号 $x_{l,m,k}^{(r)}$ 给基频估计单元 1120 能够改善数字化波形初始源信号估计 $\hat{s}[n]$ 的估计。

[0250] 在某些情况下,可能更适合于将自适应谐波滤波器应用于源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 而不是观测信号 $x_{l,m,k}^{(r)}$ 以便获得数字化波形初始源信号估计 $\hat{s}[n]$ 的较佳估计。去混响步骤的一次迭代可能将一定的特殊失真添加到源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 中以及当将自适应谐波滤波器应用于源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 时,该失真被直接继承到数字化波形初始源信号估计 $\hat{s}[n]$ 。另外,这种失真可能通过迭代的去混响步骤被累积到源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 中。为了避免失真的这种累积,有效的是信号切换单元 1160 被适配用于除了在迭代结束之前的最后一个步骤或最后几个步骤中之外将观测信号 $x_{l,m,k}^{(r)}$ 提供给自适应谐波滤波单元 1130,在所述迭代中使源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 的估计精确。

[0251] 图 11 是图示包括在图 9 中所示的初始化单元 1000 中的修改的源信号不确定性确定单元 1500 的配置的方框图。修改的源信号不确定性确定单元 1500 可以进一步包括短时傅立叶变换单元 1112、基频估计单元 1122、源信号不确定性确定子单元 1140 和信号切换单元 1162。信号切换单元 1162 的增加能够改善源信号不确定性 $\sigma_{l,m,k}^{(r)}$ 的估计。根据第二实施例,似然最大化单元 2000 的配置与在第一实施例中描述的配置相同。

[0252] 短时傅立叶变换单元 1112 被适配用于接收数字化波形观测信号 $x[n]$ 。短时傅立

叶变换单元 1112 被适配用于执行数字化波形观测信号 $x[n]$ 到作为输出的变换后的观测信号 $x_{l,m,k}^{(r)}$ 的短时傅立叶变换。信号切换单元 1162 与短时傅立叶变换单元 1110 和收敛检验单元 3000 协作。信号切换单元 1162 被适配用于从短时傅立叶变换单元 1110 接收变换后的观测信号 $x_{l,m,k}^{(r)}$ 。信号切换单元 1162 被适配用于从收敛检验单元 3000 接收源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。信号切换单元 1162 被适配用于执行第一选择操作以产生第一输出。第一选择操作是用于选择变换后的观测信号 $x_{l,m,k}^{(r)}$ 和源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 之一。在一种情况下,第一选择操作可以是用于在除了迭代的初始步骤之外的迭代的所有步骤中选择源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。在迭代的初始步骤中,信号切换单元 1162 仅接收变换后的观测信号 $x_{l,m,k}^{(r)}$ 并选择变换后的观测信号 $x_{l,m,k}^{(r)}$ 。更优选的是就基频 $f_{1,m}$ 和发声量度 $v_{1,m}$ 两者的估计而言使用源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 而不使用变换后的观测信号 $x_{l,m,k}^{(r)}$ 。

[0253] 基频估计单元 1122 与信号切换单元 1162 协作。基频估计单元 1122 被适配用于从信号切换单元 1162 接收第一输出。即,基频估计单元 1122 被适配用于在迭代的初始步骤中接收变换后的观测信号 $x_{l,m,k}^{(r)}$ 以及在除迭代的初始步骤之外的迭代的所有步骤中接收源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。基频估计单元 1122 进一步被适配用于估计每个短时间帧的基频 $f_{1,m}$ 及其发声量度 $v_{1,m}$ 。参考变换后的观测信号 $x_{l,m,k}^{(r)}$ 或源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 进行所述估计。

[0254] 源信号不确定性确定单元 1140 与基频估计单元 1122 协作。源信号不确定性确定单元 1140 被适配用于从基频估计单元 1122 接收基频 $f_{1,m}$ 和发声量度 $v_{1,m}$ 。源信号不确定性确定单元 1140 进一步被适配用于确定源信号不确定性 $\sigma_{l,m,k}^{(s)}$ 。如上所述,更优选的是就基频 $f_{1,m}$ 和发声量度 $v_{1,m}$ 两者的估计而言使用源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 而不使用观测信号 $x_{l,m,k}^{(r)}$ 。

[0255] 第三实施例:

[0256] 图 12 是图示根据本发明的第三实施例基于源和室内声学的概率模型的语音去混响设备的配置的方框图。语音去混响设备 30000 能够通过一组功能单元来实现,这些功能单元协作用于接收观测信号 $x[n]$ 的输入并产生数字化波形源信号估计 $\tilde{s}[n]$ 或滤波源信号估计 $\tilde{s}[n]$ 的输出。语音去混响设备 30000 能够例如通过计算机或处理器来实现。语音去混响设备 30000 执行语音去混响操作。语音去混响方法能够通过要由计算机执行的程序来实现。

[0257] 语音去混响设备 30000 典型地可以包括上述初始化单元 1000、上述似然最大化单元 2000-1 和逆滤波器应用单元 5000。初始化单元 1000 可以被适配用于接收数字化波形观测信号 $x[n]$ 。数字化波形观测信号 $x[n]$ 可以包含具有未知混响程度的语音信号。通过诸如一个或多个麦克风的设备能够捕获该语音信号。初始化单元 1000 可以被适配用于从观测信号提取与源信号和声学环境有关的初始源信号估计和不确定性。初始化单元 1000 还可以被适配用于用公式表达初始源信号估计、源信号不确定性和声学环境不确定性的表示。针对所有索引 l, m, k 和 k' , 这些表示被列举为是数字化波形初始源信号估计的 $\tilde{s}[n]$ 、是表示源信号不确定性的方差或离差的 $\sigma_{l,m,k}^{(s)}$ 、以及是表示声学环境不确定性的方差或离差的 $\sigma_{l,m,k}^{(a)}$ 。即,初始化单元 1000 可以被适配用于接收数字化波形信号 $x[n]$ 的输入作为观测信号并产生数字化波形初始源信号估计 $\tilde{s}[n]$ 、表示源信号不确定性的方差或离差 $\sigma_{l,m,k}^{(s)}$ 、以及表示

声学环境不确定性的方差或离差 $\sigma_{i,x}^{(s)}$ 。

[0258] 似然最大化单元 2000-1 可以与初始化单元 1000 协作。即,似然最大化单元 2000-1 可以被适配用于从初始化单元 1000 接收数字化波形初始源信号估计 $\hat{s}[n]$ 、源信号不确定性 $\sigma_{i,m,t}^{(r)}$ 、和声学环境不确定性 $\sigma_{i,x}^{(s)}$ 的输入。似然最大化单元 2000-1 还可以被适配用于接收数字化波形观测信号 $x[n]$ 的另一个输入作为观测信号。 $\hat{s}[n]$ 是数字化波形初始源信号估计。 $\sigma_{i,m,t}^{(r)}$ 是表示源信号不确定性的第一方差。 $\sigma_{i,x}^{(s)}$ 是表示声学环境不确定性的第二方差。似然最大化单元 2000-1 还可以被适配用于确定最大化似然函数的逆滤波器估计 \tilde{w}_k ,其中参考数字化波形观测信号 $x[n]$ 、数字化波形初始源信号估计 $\hat{s}[n]$ 、表示源信号不确定性的第一方差 $\sigma_{i,m,t}^{(r)}$ 、以及表示声学环境不确定性的第二方差 $\sigma_{i,x}^{(s)}$ 进行所述确定。通常,可以基于概率密度函数定义似然函数,该概率密度函数根据第一未知参数、第二未知参数、和观测数据的第一随机变量进行评估。参考源信号估计定义所述第一未知参数。参考室内传递函数的逆滤波器定义第二未知参数。参考观测信号和初始源信号估计定义观测数据的第一随机变量。逆滤波器估计是室内传递函数的逆滤波器的估计。使用迭代优化算法来执行逆滤波器估计 \tilde{w}_k 的确定。

[0259] 迭代优化算法可以在不使用上述期望最大化算法的情况下进行组织。例如,能够获得逆滤波器估计 \tilde{w}_k 和源信号估计 $\hat{\theta}_k$ 作为最大化如下定义的似然函数的逆滤波器估计和源信号估计:

$$\begin{aligned}
 [0260] \quad L\{w_k, \theta_k\} &= p\{w_k, z_k^{(v)} | \theta_k\} \\
 [0261] \quad &= p\{w_k, \{x_{i,m,t}^{(v)}\}_k | \theta_k\} p\{\{\hat{s}_{i,m,t}^{(r)}\}_k | \theta_k\} \quad (16)
 \end{aligned}$$

[0262] 能够通过紧接着的迭代算法最大化该似然函数。

[0263] 第一步骤是设置初始值为 $\theta_k = \hat{\theta}_k$ 。

[0264] 第二步骤是计算在 θ_k 固定的条件下最大化似然函数的逆滤波器估计 $w_k = \tilde{w}_k$ 。

[0265] 第三步骤是计算在 w_k 固定的条件下最大化似然函数的源信号估计 $\theta_k = \hat{\theta}_k$ 。

[0266] 第四步骤是重复上述第二和第三步骤,直到迭代的收敛被确认为止。

[0267] 当与上述方程(8)相同的定义被用于上述似然函数中的概率密度函数(pdf)时,容易地显示出能够分别通过上述方程(12)和(15)获得上述第二步骤中的逆滤波器估计 \tilde{w}_k 和上述第三步骤中的源信号估计 $\hat{\theta}_k$ 。通过检验是否逆滤波器估计 \tilde{w}_k 的当前获得的值和逆滤波器估计 \tilde{w}_k 的之前获得的值之差小于预定的阈值,可以完成第四步骤中的上述收敛确认。最后,该观测信号可以通过将在上述第二步骤中获得的逆滤波器估计 \tilde{w}_k 应用于观测信号来去混响。

[0268] 逆滤波器应用单元 5000 可以与似然最大化单元 2000-1 协作。即,逆滤波器应用

单元 5000 可以被适配用于从似然最大化单元 2000-1 接收最大化似然函数 (16) 的逆滤波器估计 \tilde{w}_k 的输入。逆滤波器应用单元 5000 还可以被适配用于接收数字化波形观测信号 $x[n]$ 。逆滤波器应用单元 5000 还可以被适配用于将逆滤波器估计 \tilde{w}_k 应用于数字化波形观测信号 $x[n]$ 以便产生恢复后的数字化波形源信号估计 $\hat{s}[n]$ 或滤波数字化波形源信号估计 $\hat{s}[n]$ 。

[0269] 在某一情况下,逆滤波器应用单元 5000 可以被适配用于将长时傅立叶变换应用于数字化波形观测信号 $x[n]$ 以产生变换后的观测信号 $x_{1,k'}$ 。逆滤波器应用单元 5000 可以进一步被适配用于将每个帧中的变换后的观测信号 $x_{1,k'}$ 乘以逆滤波器估计 \tilde{w}_k 以产生滤波源信号估计 $\hat{s}_{1,k'} = \tilde{w}_k x_{1,k'}$ 。逆滤波器应用单元 5000 可以进一步被适配用于将逆长时傅立叶变换应用于滤波源信号估计 $\hat{s}_{1,k'} = \tilde{w}_k x_{1,k'}$ 以产生滤波数字化波形源信号估计 $\hat{s}[n]$ 。

[0270] 在另一情况下,逆滤波器应用单元 5000 可以被适配用于将逆长时傅立叶变换应用于逆滤波器估计 \tilde{w}_k 以产生数字化波形逆滤波器估计 $\tilde{w}[n]$ 。逆滤波器应用单元 5000 可以被适配用于将数字化波形观测信号 $x[n]$ 与数字化波形逆滤波器估计 $\tilde{w}[n]$ 进行卷积以产生恢复后的数字化波形源信号估计 $\hat{s}[n] = \sum_m x[n-m]\tilde{w}[m]$ 。

[0271] 似然最大化单元 2000-1 能够通过相互协作以确定并输出最大化所述似然函数的逆滤波器估计 \tilde{w}_k 的一组子功能单元来实现。图 13 是图示在图 12 中所示的似然最大化单元 2000-1 的配置的方框图。在一种情况下,似然最大化单元 2000-1 可以进一步包括上述长时傅立叶变换单元 2100、上述更新单元 2200、上述 STFS 到 LTFS 变换单元 2300、上述逆滤波器估计单元 2400、上述滤波单元 2500、LTFS 到 STFS 变换单元 2600、源信号估计单元 2710、收敛检验单元 2720、上述短时傅立叶变换单元 2800、以及上述长时傅立叶变换单元 2900。这些单元协作以继续执行迭代操作,直到最大化似然函数的逆滤波器估计已经被确定为止。

[0272] 长时傅立叶变换单元 2100 被适配用于从初始化单元 1000 接收数字化波形观测信号 $x[n]$ 作为观测信号。长时傅立叶变换单元 2100 还被适配用于执行数字化波形观测信号 $x[n]$ 到作为长期傅立叶频谱 (LTFS) 的变换后的观测信号 $x_{1,k'}$ 的长时傅立叶变换。

[0273] 短时傅立叶变换单元 2800 被适配用于从初始化单元 1000 接收数字化波形初始源信号估计 $\hat{s}[n]$ 。短时傅立叶变换单元 2800 被适配用于执行数字化波形初始源信号估计 $\hat{s}[n]$ 到初始源信号估计 $\hat{s}_{1,m,k}$ 的短时傅立叶变换。

[0274] 长时傅立叶变换单元 2900 被适配用于从初始化单元 1000 接收数字化波形初始源信号估计 $\hat{s}[n]$ 。长时傅立叶变换单元 2900 被适配用于执行数字化波形初始源信号估计 $\hat{s}[n]$ 到初始源信号估计 $\hat{s}_{1,k}$ 的长时傅立叶变换。

[0275] 更新单元 2200 与长时傅立叶变换单元 2900 和 STFS 到 LTFS 变换单元 2300 协作。更新单元 2200 被适配用于在迭代的初始步骤中从长时傅立叶变换单元 2900 接收初始源信号估计 $\hat{s}_{1,k}$ 以及进一步被适配用于用源信号估计 θ_k 代替 $\hat{s}_{1,k}$ 。更新单元 2200 此外被适配用于发送更新的源信号估计 θ_k 给逆滤波器估计单元 2400。更新单元 2200 还被适配用于

在迭代的随后步骤中从 STFS 到 LTFS 变换单元 2300 接收源信号估计 $\tilde{s}_{1,k}$ ，以及用源信号估计 $\theta_{k'}$ 代替 $\{\tilde{s}_{1,k}\}$ 。更新单元 2200 还被适配用于发送更新的源信号估计 $\theta_{k'}$ 给逆滤波器估计单元 2400。

[0276] 逆滤波器估计单元 2400 与长时傅立叶变换单元 2100、更新单元 2200 和初始化单元 1000 协作。逆滤波器估计单元 2400 被适配用于从长时傅立叶变换单元 2100 接收观测信号 $x_{1,k'}$ 。逆滤波器估计单元 2400 还被适配用于从更新单元 2200 接收更新的源信号估计 $\theta_{k'}$ 。逆滤波器估计单元 2400 还被适配用于从初始化单元 1000 接收表示声学环境不确定性的第二方差 $\sigma_{1,k}^{(a)}$ 。逆滤波器估计单元 2400 还被适配用于根据上述方程 (12) 基于观测信号 $x_{1,k'}$ 、更新的源信号估计 $\theta_{k'}$ 、和表示声学环境不确定性的第二方差 $\sigma_{1,k}^{(a)}$ 计算逆滤波器估计 \tilde{w}_k 。逆滤波器估计单元 2400 进一步被适配用于输出逆滤波器估计 \tilde{w}_k 。

[0277] 收敛检验单元 2720 与逆滤波器估计单元 2400 协作。收敛检验单元 2720 被适配用于从逆滤波器估计单元 2400 接收逆滤波器估计 \tilde{w}_k 。收敛检验单元 2720 被适配用于例如通过比较当前已经被估计的逆滤波器估计 \tilde{w}_k 的当前值与之前已经被估计的逆滤波器估计 \tilde{w}_k 的前一值以及检验是否所述当前值偏离前一值小于一定的预定量来确定迭代过程的收敛状态。如果收敛检验单元 2720 确认逆滤波器估计 \tilde{w}_k 的当前值偏离其前一值小于一定的预定量，那么收敛检验单元 2720 识别出已经获得逆滤波器估计 \tilde{w}_k 的收敛。如果收敛检验单元 2720 确认逆滤波器估计 \tilde{w}_k 的当前值偏离其前一值不小于一定的预定量，那么收敛检验单元 2720 识别出还未获得逆滤波器估计 \tilde{w}_k 的收敛。

[0278] 作为修改有可能的是当迭代次数达到一定的预定值时终止所述迭代过程。即，收敛检验单元 2720 已经确认迭代次数达到一定的预定值，那么收敛检验单元 2720 识别出已经获得逆滤波器估计 \tilde{w}_k 的收敛。如果收敛检验单元 2720 已经确认已经获得逆滤波器估计 \tilde{w}_k 的收敛，那么收敛检验单元 2720 将逆滤波器估计 \tilde{w}_k 作为第一输出提供给逆滤波器应用单元 5000。如果收敛检验单元 2720 已经确定还未获得逆滤波器估计 \tilde{w}_k 的收敛，那么收敛检验单元 2720 将逆滤波器估计 \tilde{w}_k 作为第二输出提供给滤波单元 2500。

[0279] 滤波单元 2500 与长时傅立叶变换单元 2100 和收敛检验单元 2720 协作。滤波单元 2500 被适配用于从长时傅立叶变换单元 2100 接收观测信号 $x_{1,k'}$ 。滤波单元 2500 还被适配用于从收敛检验单元 2720 接收逆滤波器估计 \tilde{w}_k 。滤波单元 2500 还被适配用于将观测信号 $x_{1,k'}$ 应用于逆滤波器估计 \tilde{w}_k 以产生滤波源信号估计 $\tilde{s}_{1,k}$ 。

[0280] 用于将观测信号 $x_{1,k'}$ 应用于逆滤波器估计 \tilde{w}_k 的滤波过程的典型实例可以包括但不限于计算观测信号 $x_{1,k'}$ 和逆滤波器估计 \tilde{w}_k 的乘积 $\tilde{w}_k x_{1,k'}$ 。在这种情况下，通过观测信号 $x_{1,k'}$ 和逆滤波器估计 \tilde{w}_k 的乘积 $\tilde{w}_k x_{1,k'}$ 来给定滤波源信号估计 $\tilde{s}_{1,k}$ 。

[0281] LTFS 到 STFS 变换单元 2600 与滤波单元 2500 协作。LTFS 到 STFS 变换单元 2600 被适配用于从滤波单元 2500 接收滤波源信号估计 $\tilde{s}_{1,k}$ 。LTFS 到 STFS 变换单元 2600 进一步

被适配用于执行滤波源信号估计 $\bar{s}_{l,m,k}$ 到变换后的滤波源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的 LTFS 到 STFS 变换。当滤波过程是计算观测信号 $x_{l,k}$ 和逆滤波器估计 \tilde{w}_k 的乘积 $\tilde{w}_k x_{l,k}$ 时, LTFS 到 STFS 变换单元 2600 进一步被适配用于执行乘积 $\tilde{w}_k x_{l,k}$ 到变换后的信号 $LS_{m,k} \{ \{ \tilde{w}_k x_{l,k} \}_l \}$ 的 LTFS 到 STFS 变换。在这种情况下, 乘积 $\tilde{w}_k x_{l,k}$ 表示滤波源信号估计 $\bar{s}_{l,m,k}$ 以及变换后的信号 $LS_{m,k} \{ \{ \tilde{w}_k x_{l,k} \}_l \}$ 表示变换后的滤波源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 。

[0282] 源信号估计单元 2710 与 LTFS 到 STFS 变换单元 2600、短时傅立叶变换单元 2800、以及初始化单元 1000 协作。源信号估计单元 2710 被适配用于从 LTFS 到 STFS 变换单元 2600 接收变换后的滤波源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 。源信号估计单元 2710 还被适配用于从初始化单元 1000 接收表示源信号不确定性的第一方差 $\sigma_{l,m,k}^{(m)}$ 以及表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 。源信号估计单元 2710 还被适配用于从短时傅立叶变换单元 2800 接收初始源信号估计 $\hat{s}_{l,m,k}^{(r)}$ 。源信号估计单元 2710 还被适配用于基于变换后的滤波源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 、表示源信号不确定性的第一方差 $\sigma_{l,m,k}^{(m)}$ 、表示声学环境不确定性的第二方差 $\sigma_{l,k}^{(a)}$ 以及初始源信号估计 $\hat{s}_{l,m,k}^{(r)}$ 估计源信号 $\tilde{s}_{l,m,k}^{(r)}$, 其中根据上述方程 (15) 进行所述估计。

[0283] STFS 到 LTFS 变换单元 2300 与源信号估计单元 2710 协作。STFS 到 LTFS 变换单元 2300 被适配用于从源信号估计单元 2710 接收源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 。STFS 到 LTFS 变换单元 2300 被适配用于执行源信号估计 $\tilde{s}_{l,m,k}^{(r)}$ 到变换后的源信号估计 $\bar{s}_{l,m,k}^{(r)}$ 的 STFS 到 LTFS 变换。

[0284] 在迭代操作的随后步骤中, 更新单元 2200 从 STFS 到 LTFS 变换单元 2300 接收源信号估计 $\bar{s}_{l,m,k}^{(r)}$, 并且用源信号估计 θ_k 代替 $\{s_{l,m,k}\}_k$ 以及发送更新的源信号估计 θ_k 给逆滤波器估计单元 2400。在迭代的初始步骤中, 更新的源信号估计 θ_k 是从长时傅立叶变换单元 2900 提供的 $\{s_{l,m,k}\}_k$ 。在迭代的第二或随后步骤中, 更新的源信号估计 θ_k 是 $\{s_{l,m,k}\}_k$ 。

[0285] 将参考图 13 描述似然最大化单元 2000-1 的操作。

[0286] 在迭代的初始步骤中, 提供数字化波形观测信号 $x[n]$ 给长时傅立叶变换单元 2100。长时傅立叶变换单元 2100 执行长时傅立叶变换以便将数字化波形观测信号 $x[n]$ 变换成变换后的观测信号 $x_{l,k}$ 作为长时傅立叶频谱 (LTFS)。将数字化波形初始源信号估计 $\hat{s}[n]$ 从初始化单元 1000 提供给短时傅立叶变换单元 2800 和长时傅立叶变换单元 2900。短时傅立叶变换单元 2800 执行短时傅立叶变换以便将数字化波形初始源信号估计 $\hat{s}[n]$ 变换成初始源信号估计 $\hat{s}_{l,m,k}^{(r)}$ 。长时傅立叶变换单元 2900 执行长时傅立叶变换以便将数字化波形初始源信号估计 $\hat{s}[n]$ 变换成初始源信号估计 $\hat{s}_{l,m,k}^{(r)}$ 。

[0287] 从长时傅立叶变换单元 2900 提供初始源信号估计 $\hat{s}_{l,m,k}^{(r)}$ 给更新单元 2200。更新单元 2200 用源信号估计 θ_k 代替初始源信号估计 $\{s_{l,m,k}\}_k$ 。然后将初始源信号估计 $\theta_k = \{s_{l,m,k}\}_k$ 从更新单元 2200 提供给逆滤波器估计单元 2400。从长时傅立叶变换单元 2100 提供观测信号 $x_{l,k}$ 给逆滤波器估计单元 2400。从初始化单元 1000 提供表示声学环境不确定性的第二方差

$\sigma_{i,k}^{(q)}$ 给逆滤波器估计单元 2400。逆滤波器估计单元 2400 基于观测信号 $x_{1,k}$ 、初始源信号估计 θ_k 、和表示声学环境不确定性的第二方差 $\sigma_{i,k}^{(q)}$ 计算逆滤波器估计 \tilde{w}_k ，其中根据上述方程 (12) 进行所述计算。

[0288] 从逆滤波器估计单元 2400 提供逆滤波器估计 \tilde{w}_k 给收敛检验单元 2720。收敛检验单元 2720 进行迭代过程的收敛状态的确定。例如，通过比较当前已经被估计的逆滤波器估计 \tilde{w}_k 的当前值与之前已经被估计的逆滤波器估计 \tilde{w}_k 的前一值来进行所述确定。收敛检验单元 2720 检验是否当前值偏离前一值小于一定的预定量。如果收敛检验单元 2720 确认逆滤波器估计 \tilde{w}_k 的当前值偏离其前一值小于一定的预定量，那么收敛检验单元 2720 识别出已经获得逆滤波器估计 \tilde{w}_k 的收敛。如果收敛检验单元 2720 确认逆滤波器估计 \tilde{w}_k 的当前值偏离其前一值不小于一定的预定量，那么收敛检验单元 2720 识别出还未获得逆滤波器估计 \tilde{w}_k 的收敛。

[0289] 如果已经获得逆滤波器估计 \tilde{w}_k 的收敛，那么从收敛检验单元 2720 提供逆滤波器估计 \tilde{w}_k 给逆滤波器应用单元 5000。如果还未获得逆滤波器估计 \tilde{w}_k 的收敛，那么从收敛检验单元 2720 提供逆滤波器估计 \tilde{w}_k 给滤波单元 2500。观测信号 $x_{1,k}$ 进一步从长时傅立叶变换单元 2100 被提供给滤波单元 2500。滤波单元 2500 将逆滤波器估计 \tilde{w}_k 应用于观测信号 $x_{1,k}$ 以产生滤波源信号估计 $\bar{s}_{i,k}$ 。用于将观测信号 $x_{1,k}$ 应用于逆滤波器估计 \tilde{w}_k 的滤波过程的典型实例可以是计算观测信号 $x_{1,k}$ 和逆滤波器估计 \tilde{w}_k 的乘积 $\tilde{w}_k x_{1,k}$ 。在这种情况下，通过观测信号 $x_{1,k}$ 和逆滤波器估计 \tilde{w}_k 的乘积 $\tilde{w}_k x_{1,k}$ 来给定滤波源信号估计 $\bar{s}_{i,k}$ 。

[0290] 从滤波单元 2500 提供滤波源信号估计 $\bar{s}_{i,k}$ 给 LTFS 到 STFS 变换单元 2600。LTFS 到 STFS 变换单元 2600 执行 LTFS 到 STFS 变换以便将滤波源信号估计 $\bar{s}_{i,k}$ 变换成变换后的滤波源信号估计 $\bar{s}_{i,m,k}^{(r)}$ 。当所述滤波过程是计算观测信号 $x_{1,k}$ 和逆滤波器估计 \tilde{w}_k 的乘积 $\tilde{w}_k x_{1,k}$ 时，该乘积 $\tilde{w}_k x_{1,k}$ 被变换成变换后的信号 $LS_{m,k} \{ \{ \tilde{w}_k x_{1,k} \}_r \}$ 。

[0291] 从 LTFS 到 STFS 变换单元 2600 提供变换后的滤波源信号估计 $\bar{s}_{i,m,k}^{(r)}$ 给源信号估计单元 2710。从初始化单元 1000 提供表示源信号不确定性的第一方差 $\sigma_{i,m,k}^{(r)}$ 和表示声学环境不确定性的第二方差 $\sigma_{i,k}^{(q)}$ 给源信号估计单元 2710。从短时傅立叶变换单元 2800 提供初始源信号估计 $\hat{s}_{i,m,k}^{(r)}$ 给源信号估计单元 2710。源信号估计单元 2710 基于变换后的滤波源信号估计 $\bar{s}_{i,m,k}^{(r)}$ 、表示源信号不确定性的第一方差 $\sigma_{i,m,k}^{(r)}$ 、和表示声学环境不确定性的第二方差 $\sigma_{i,k}^{(q)}$ 以及初始源信号估计 $\hat{s}_{i,m,k}^{(r)}$ 计算源信号估计 $\tilde{s}_{i,m,k}^{(r)}$ ，其中根据上述方程 (15) 进行所述估计。

[0292] 从源信号估计单元 2710 提供源信号估计 $\tilde{s}_{i,m,k}^{(r)}$ 给 STFS 到 LTFS 变换单元 2300 以便将源信号估计 $\tilde{s}_{i,m,k}^{(r)}$ 变换成变换后的源信号估计 $\tilde{s}_{i,k}$ 。从 STFS 到 LTFS 变换单元 2300 提供所述变换后的源信号估计 $\tilde{s}_{i,k}$ 给更新单元 2200。更新单元 2200 用源信号估计 θ_k 代替变换后的源信号估计 $\tilde{s}_{i,k}$ 。从更新单元 2200 提供更新的源信号估计 θ_k 给逆滤波器估计单元 2400。

[0293] 在迭代的第二或随后步骤中,从更新单元 2200 提供源信号估计 $\hat{\theta}_k = \{\hat{s}_{i,k}\}_k$ 给逆滤波器估计单元 2400。还将观测信号 $x_{1,k}$ 从长时傅立叶变换单元 2100 提供给逆滤波器估计单元 2400。从初始化单元 1000 提供表示声学环境不确定性的第二方差 $\sigma_{i,k}^{(2)}$ 给逆滤波器估计单元 2400。逆滤波器估计单元 2400 基于观测信号 $x_{1,k}$ 、更新的源信号估计 $\hat{\theta}_k = \{\hat{s}_{i,k}\}_k$ 和表示声学环境不确定性的第二方差 $\sigma_{i,k}^{(2)}$ 计算更新的逆滤波器估计 \tilde{w}_k , 其中根据上述方程 (12) 进行所述计算。

[0294] 从逆滤波器估计单元 2400 提供更新的逆滤波器估计 \tilde{w}_k 给收敛检验单元 2720。收敛检验单元 2720 进行对迭代过程的收敛状态的确定。

[0295] 上述迭代过程将被继续,直到收敛检验单元 2720 已经确认逆滤波器估计 \tilde{w}_k 的收敛已经被获得为止。

[0296] 图 14 是图示在图 12 中所示的逆滤波器应用单元 5000 的配置的方框图。逆滤波器应用单元 5000 的典型实例可以包括但不限于逆长时傅立叶变换单元 5100 和卷积单元 5200。逆长时傅立叶变换单元 5100 与似然最大化单元 2000-1 协作。逆长时傅立叶变换单元 5100 被适配用于从似然最大化单元 2000-1 接收逆滤波器估计 \tilde{w}_k 。长时傅立叶变换单元 5100 进一步被适配用于执行逆滤波器估计 \tilde{w}_k 到数字化波形逆滤波器估计 $\tilde{w}[n]$ 的逆长时傅立叶变换。

[0297] 卷积单元 5200 与逆长时傅立叶变换单元 5100 协作。卷积单元 5200 被适配用于从逆长时傅立叶变换单元 5100 接收数字化波形逆滤波器估计 $\tilde{w}[n]$ 。卷积单元 5200 还被适配用于接收数字化波形观测信号 $x[n]$ 。卷积单元 5200 还被适配用于执行卷积过程以将数字化波形观测信号 $x[n]$ 与数字化波形逆滤波器估计 $\tilde{w}[n]$ 进行卷积以产生恢复后的数字化波形源信号估计 $\hat{s}[n] = \sum_m x[n-m]\tilde{w}[m]$ 作为去混响的信号。

[0298] 图 15 是图示在图 12 中所示的逆滤波器应用单元 5000 的配置的方框图。逆滤波器应用单元 5000 的典型实例可以包括但不限于长时傅立叶变换单元 5300、滤波单元 5400、以及逆长时傅立叶变换单元 5500。长时傅立叶变换单元 5300 被适配用于接收数字化波形观测信号 $x[n]$ 。长时傅立叶变换单元 5300 被适配用于执行数字化波形观测信号 $x[n]$ 到变换后的观测信号 $x_{1,k}$ 长时傅立叶变换。

[0299] 滤波单元 5400 与长时傅立叶变换单元 5300 和似然最大化单元 2000-1 协作。滤波单元 5400 被适配用于从长时傅立叶变换单元 5300 接收变换后的观测信号 $x_{1,k}$ 。滤波单元 5400 还被适配用于从似然最大化单元 2000-1 接收逆滤波器估计 \tilde{w}_k 。滤波单元 5400 进一步被适配用于将逆滤波器估计 \tilde{w}_k 应用于变换后的观测信号 $x_{1,k}$ 以产生滤波源信号估计 $\hat{s}_{i,k} = \tilde{w}_k x_{1,k}$ 。将逆滤波器估计 \tilde{w}_k 应用于变换后的观测信号 $x_{1,k}$ 可以通过将每个帧中的变换后的观测信号 $x_{1,k}$ 乘以逆滤波器估计 \tilde{w}_k 来进行。

[0300] 逆长时傅立叶变换单元 5500 与滤波单元 5400 协作。逆长时傅立叶变换单元 5500 被适配用于从滤波单元 5400 接收滤波源信号估计 $\hat{s}_{i,k}$ 。逆长时傅立叶变换单元 5500 被适配用于执行滤波源信号估计 $\hat{s}_{i,k}$ 到作为去混响的信号的滤波数字化波形源信号估计 $\hat{s}[n]$ 的

逆长时傅立叶变换。

[0301] 试验：

[0302] 借助于确认本方法的性能执行了简单的试验。如由 Tomohiro Nakatani 和 Masato Miyoshi 的“Blind dereverberation of singlechannel speech signal based on harmonic structure”(Proc. ICASSP-2003, vol. 1, 第 92-95 页, 2003 年 4 月) 详细公开的, 以 0.1 秒、0.2 秒、0.5 秒、和 1.0 秒的 RT60 时间采用了单词话语 (word utterances) 的相同的源信号以及相同的脉冲响应。通过将源信号与脉冲响应进行卷积来合成观测信号。准备了与用于 HERB 和 SBD 相同的两种类型的初始源信号估计, 即 $\hat{s}_{l,m,k}^{(r)} = H\{x_{l,m,k}^{(r)}\}$ 和 $\hat{s}_{l,m,k}^{(r)} = N\{x_{l,m,k}^{(r)}\}$, 其中 $H\{\cdot\}$ 和 $N\{\cdot\}$ 分别是用于 HERB 的谐波滤波器和用于 SBD 的噪声减小滤波器。相对于发声量度 $v_{l,m}$ 确定源信号不确定性, 其中该发声量度和 HERB 一起被用于判定观测信号的每个短时间帧的语音状态。根据该量度, 当对于固定的阈值 δ , $v_{l,m} > \delta$ 时, 帧被确定为有声 (voiced)。特别是, $\sigma_{l,m,k}^{(sr)}$ 在试验中被确定为:

[0303]

$$\sigma_{l,m,k}^{(sr)} = \begin{cases} G\left\{\frac{v_{l,m}-\delta}{\max_i\{v_{l,m}\}-\delta}\right\} & \text{if } v_{l,m} > \delta \text{ 并且 } k \text{ 是谐波频率,} \\ \infty & \text{if } v_{l,m} > \delta \text{ 并且 } k \text{ 不是谐波频率,} \\ G\left\{\frac{v_{l,m}-\delta}{\min_i\{v_{l,m}\}-\delta}\right\} & \text{if } v_{l,m} \leq \delta, \end{cases} \quad (17)$$

[0304] 其中 $G\{u\}$ 是被定义成 $G\{u\} = e^{-160(u-0.95)}$ 的非线性归一化函数。另一方面, $\sigma_{l,k}^{(s)}$ 被设置成常量值 1。结果, 当 $G\{u\}$ 中的 u 从 0 移动到 1 时, 上述方程 (15) 中的 $\hat{s}_{l,m,k}^{(r)}$ 的权重变成从 0 到 1 变化的 S 形函数。对于每个试验, EM 步骤被迭代了四次。另外, 还引入了具有反馈回路的重复估计方案。采用对应于 42ms 的 $K^{(r)} = 504$ 、对应于 10.9s 的 $K = 130800$ 、对应于 1ms 的 $\tau = 12$ 、以及 12kHz 采样频率作为分析条件。

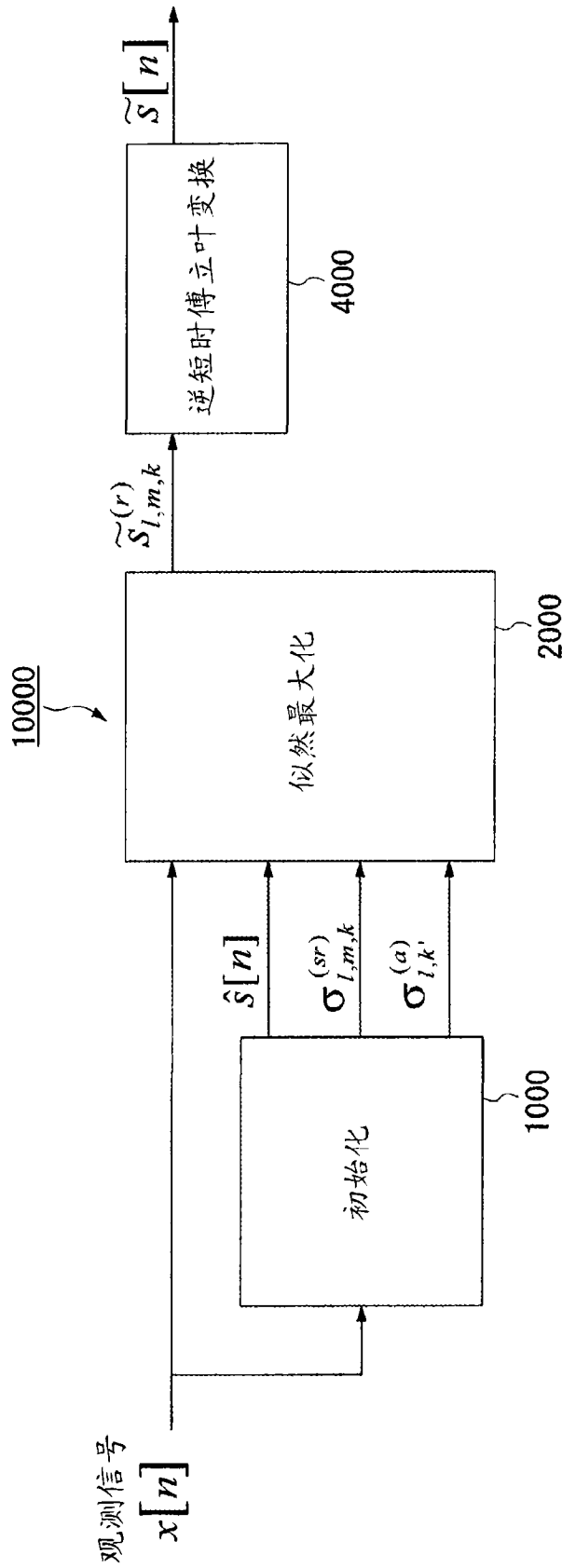
[0305] 能量衰减曲线：

[0306] 图 12A 到 12H 显示使用女人和男人所说的 100 单词观测信号在利用和不利用 EM 算法的情况下室内脉冲响应和通过 HERB 和 SBD 去混响的脉冲响应的能量衰减曲线。图 12A 图示当女人说话时在 RT60 = 1.0 秒的能量衰减曲线。图 12B 图示当女人说话时在 RT60 = 0.5 秒的能量衰减曲线。图 12C 图示当女人说话时在 RT60 = 0.2 秒的能量衰减曲线。图 12D 图示当女人说话时在 RT60 = 0.1 秒的能量衰减曲线。图 12E 图示当男人说话时在 RT60 = 1.0 秒的能量衰减曲线。图 12F 图示当男人说话时在 RT60 = 0.5 秒的能量衰减曲线。图 12G 图示当男人说话时在 RT60 = 0.2 秒的能量衰减曲线。图 12H 图示当男人说话时在 RT60 = 0.1 秒的能量衰减曲线。图 12A 到 12H 清楚地表明 EM 算法能够利用 HERB 和 SBD 有效地减小混响能量。

[0307] 相应地, 如上所述, 本发明的一个方面针对新的去混响方法, 其中源信号和室内声学的特征借助高斯概率密度函数 (pdf) 来表示, 以及源信号被估计为最大化基于这些概率密度函数 (pdf) 所定义的似然函数的信号。迭代优化算法被用于有效地解决这种优化问

题。试验结果显示出就去混响脉冲响应的能量衰减曲线而言,本方法能够大大地提高基于语音信号特征的两种去混响方法、即 HERB 和 SBD 的性能。由于 HERB 和 SBD 在提高在混响环境中捕获的语音信号的 ASR 性能方面是有效的,所以本方法能够利用较少的观测信号提高性能。

[0308] 尽管以上已经描述和图示了本发明的优选实施例,但是应该理解这些优选实施例是本发明的示例并且不应被认为是限制性的。在不背离本发明的精神或范围的情况下能够进行添加、省略、替换、和其他修改。因此,本发明不应该被认为是受上述描述限制,而是仅仅受所附的权利要求书的范围限制。



l ($=1$ 到 L): 长时傅立叶频谱 (LTFS) 的时间索引
 m ($=0$ 到 $M-1$): 短时傅立叶频谱 (STFS) 的时间子索引
 k ($=1$ 到 $K^{(r)}$): STFS 的频率索引
 k' ($=1$ 到 K): LTFS 的频率索引

图 1

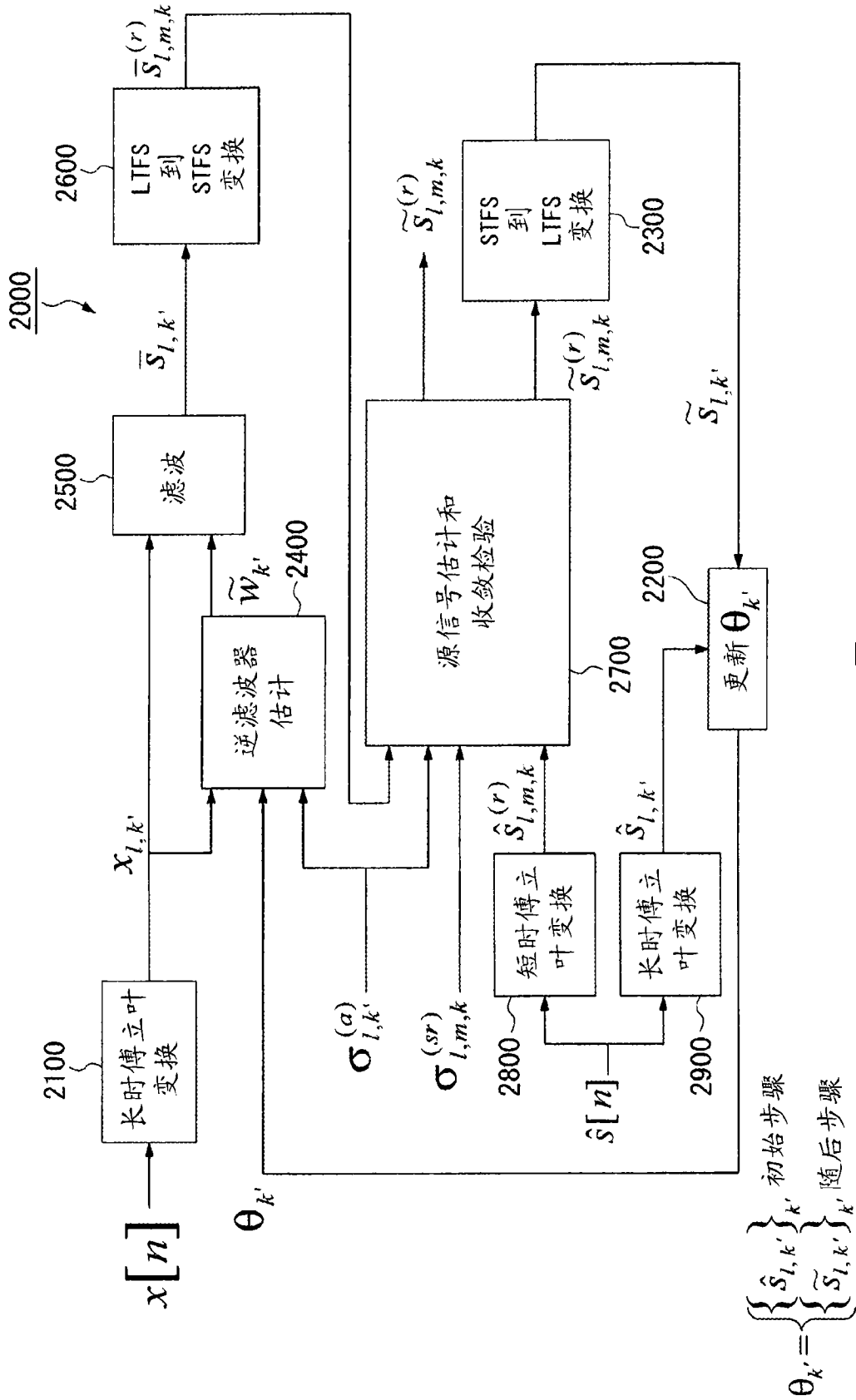


图 2

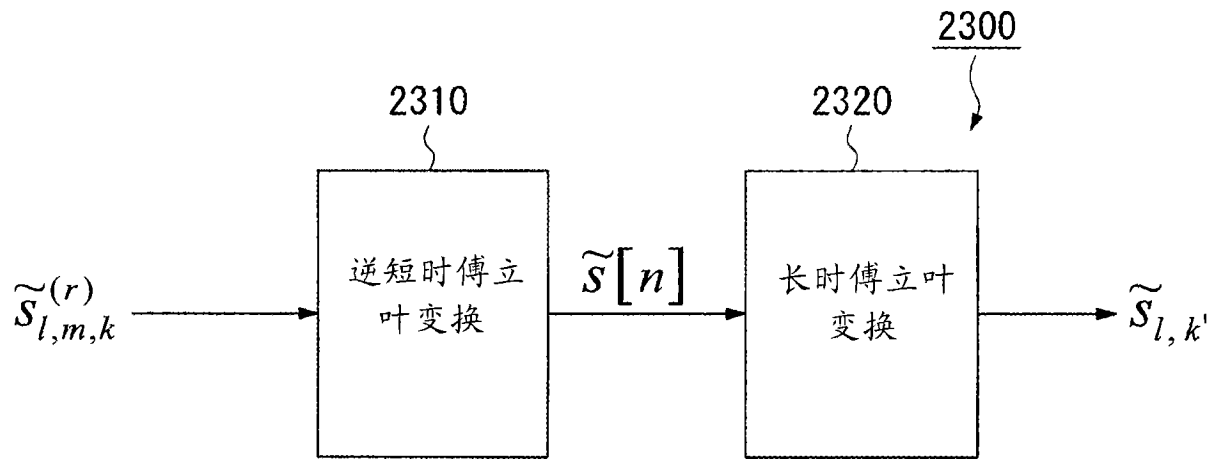


图 3A

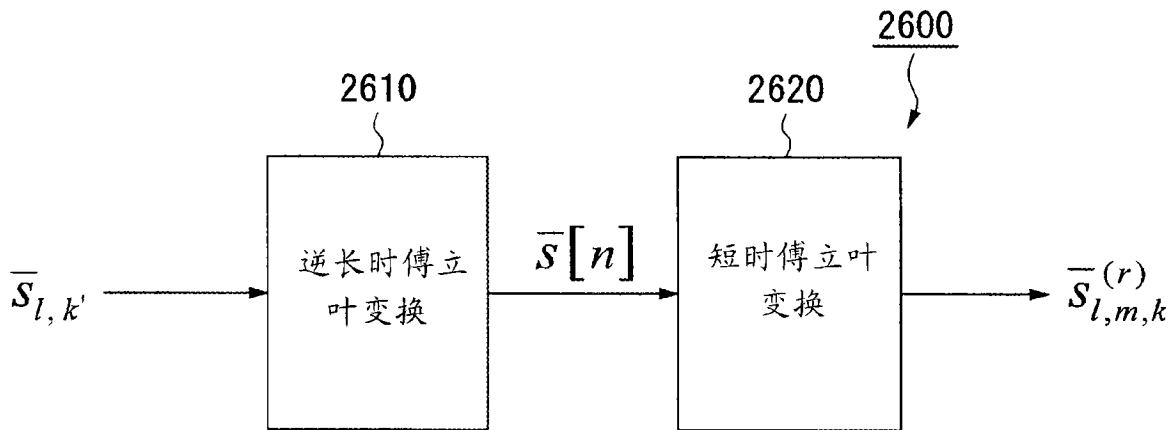


图 3B

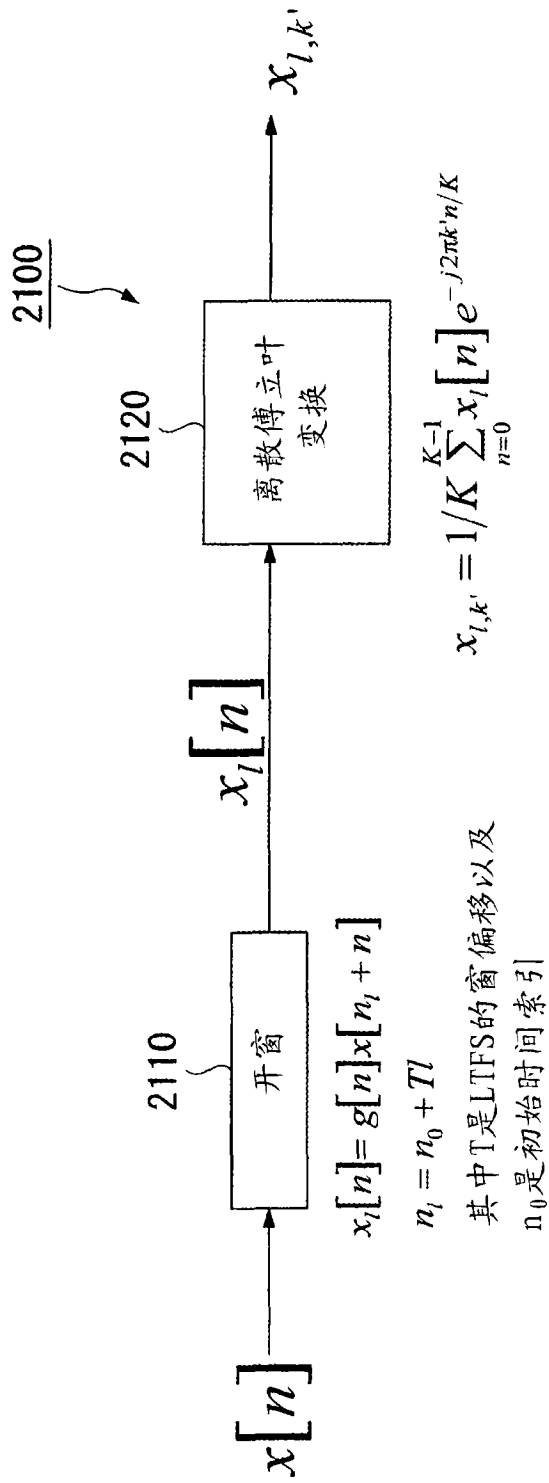


图 4A

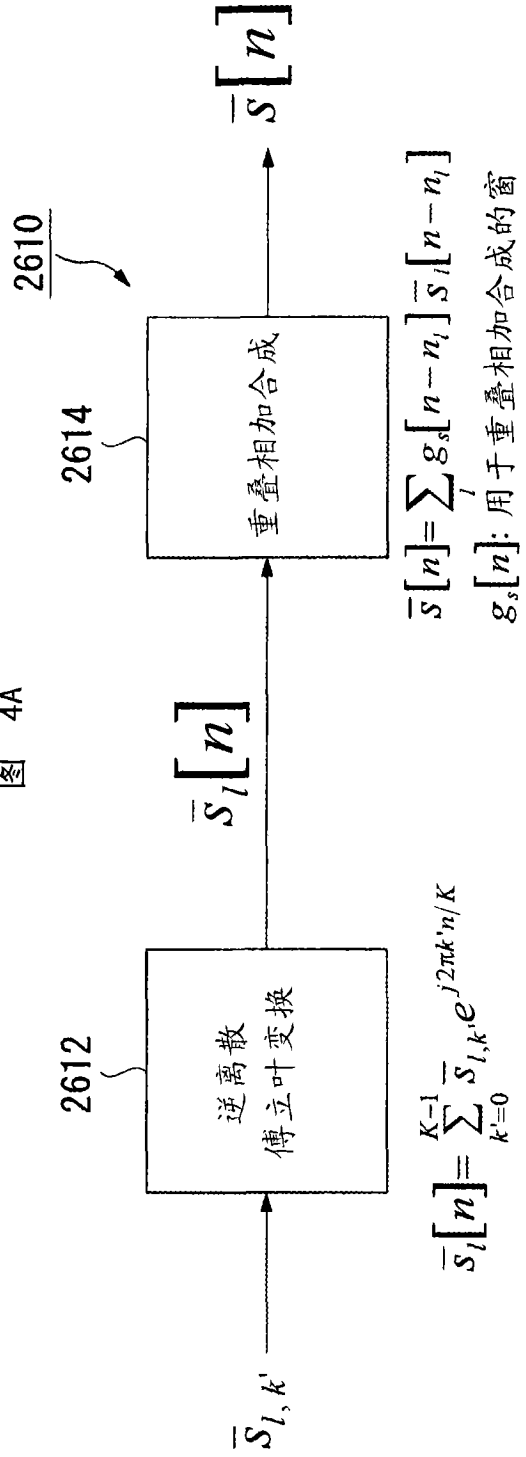


图 4B

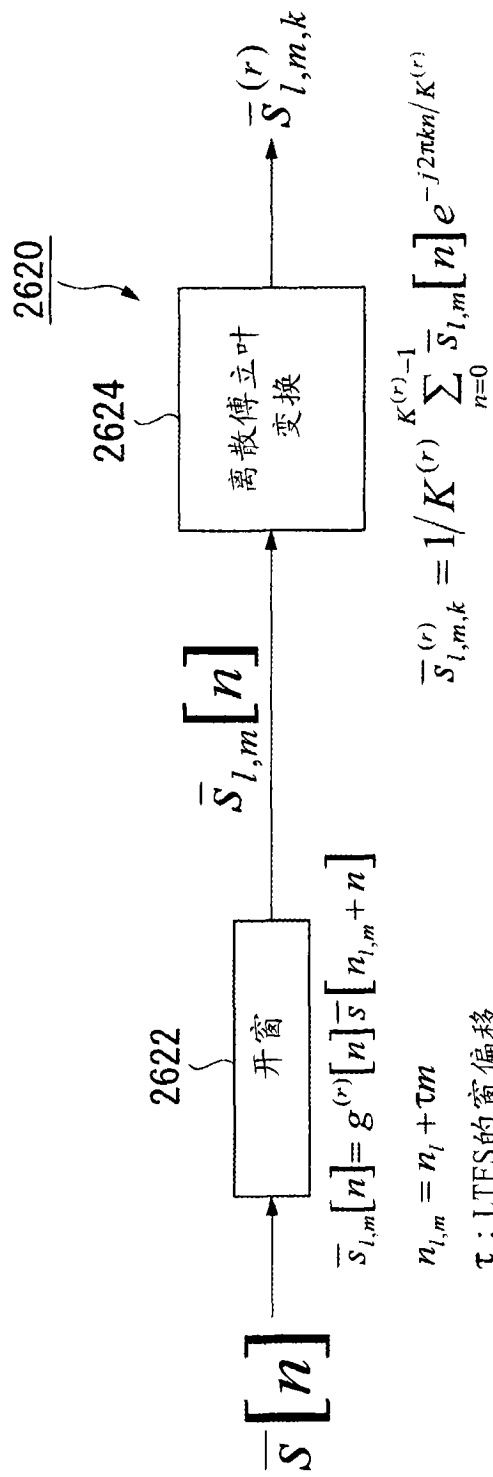


图 5A

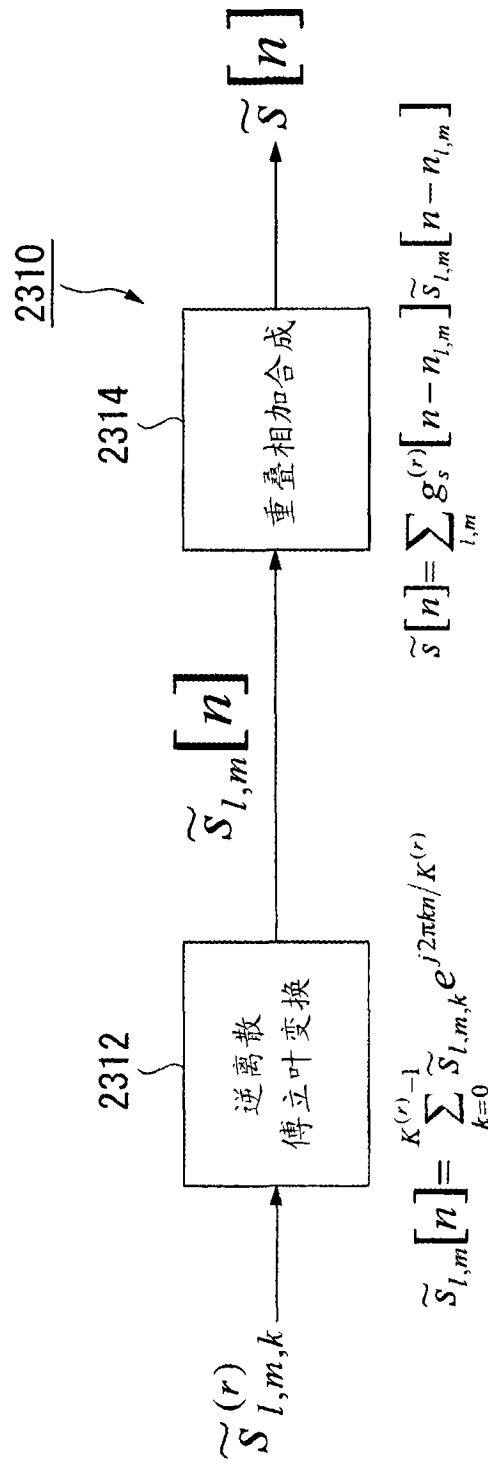
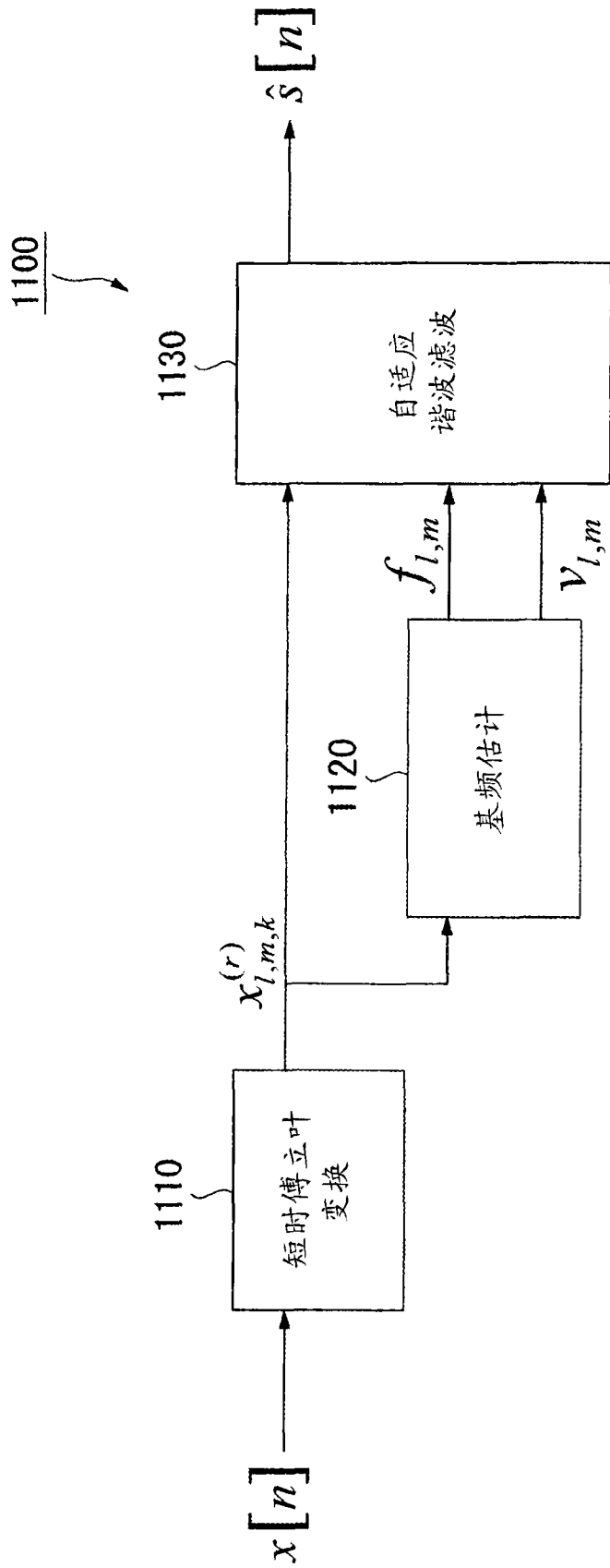


图 5B $g_s^{(r)}[n]$: 用于重叠相加合成的窗



$f_{l,m}$: 通过1和m索引的STFS的基频

$v_{l,m}$: 通过1和m索引的STFS的发声量度

图 6

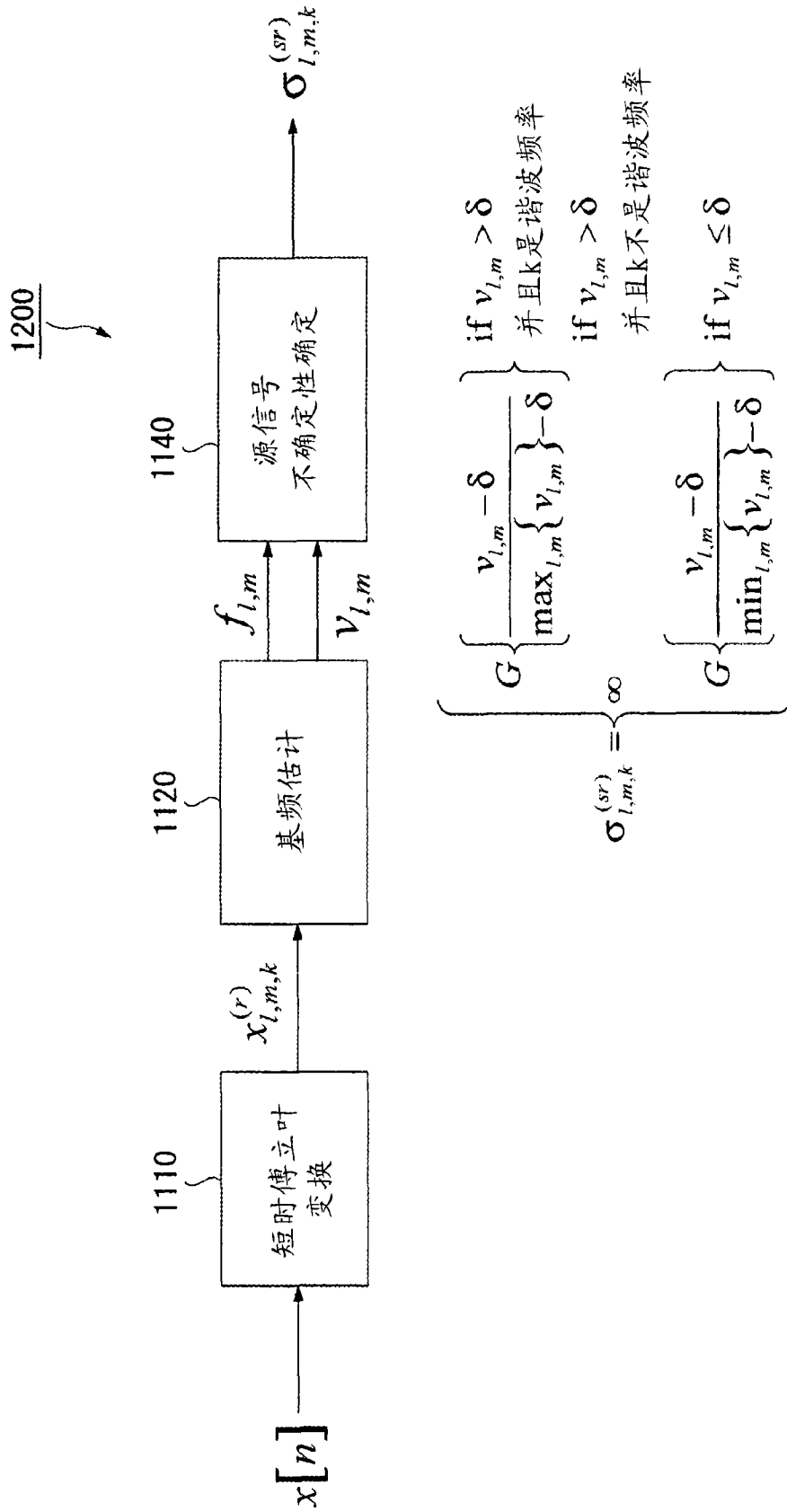


图 7

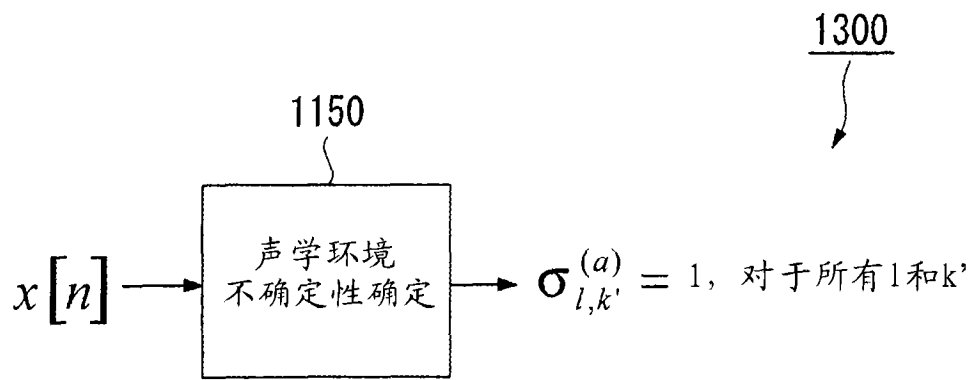


图 8

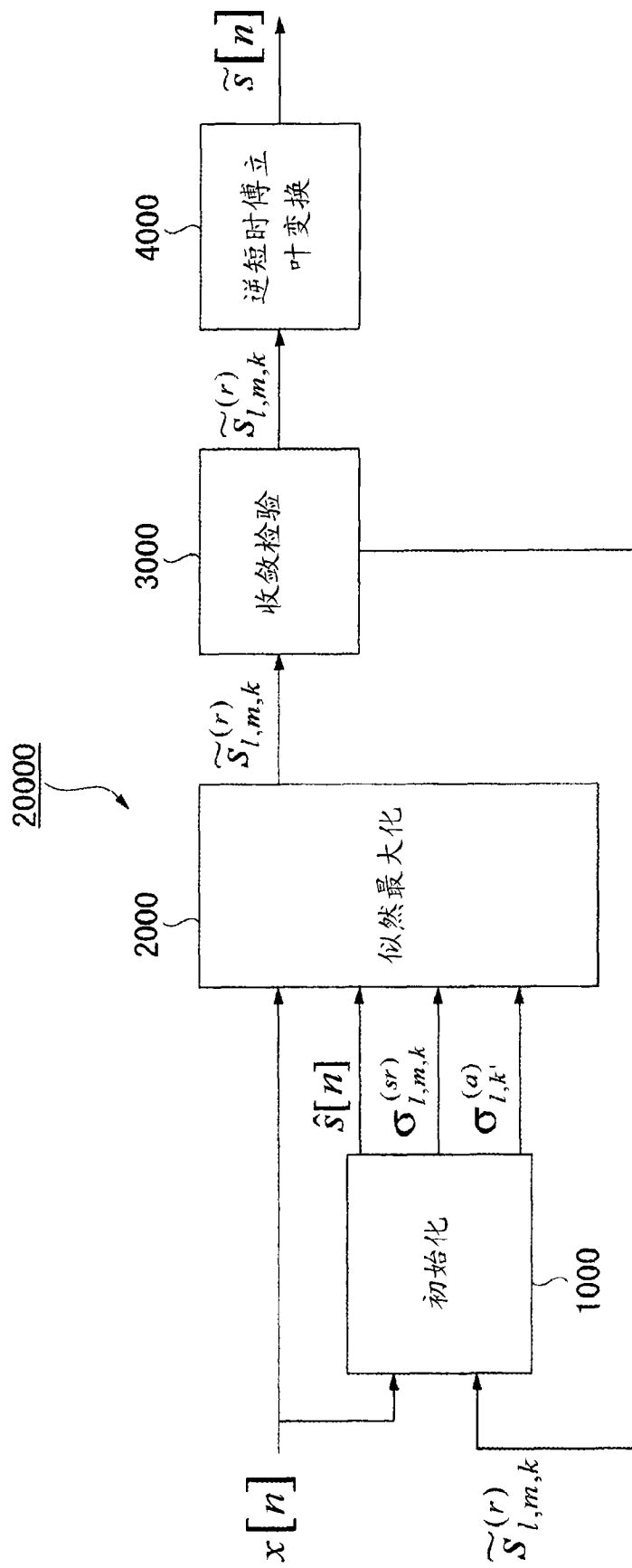


图 9

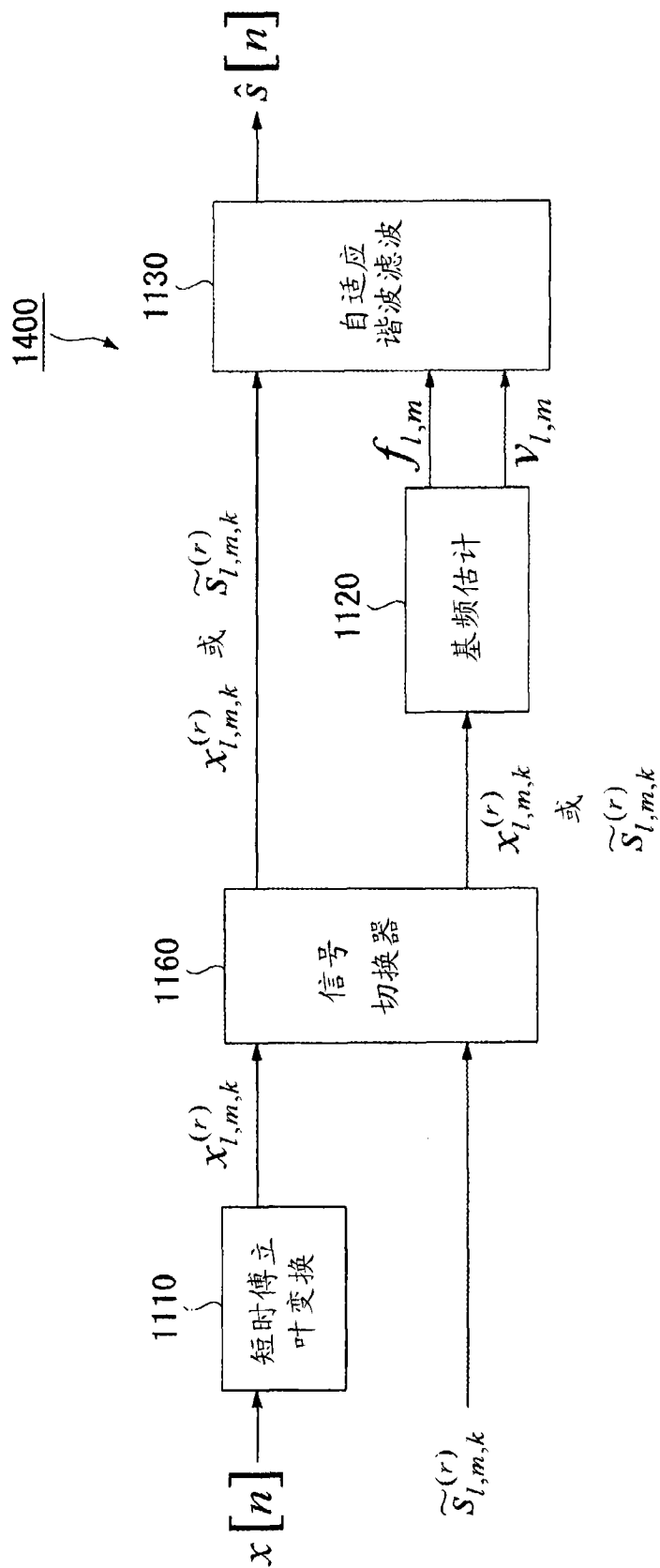


图 10

1500

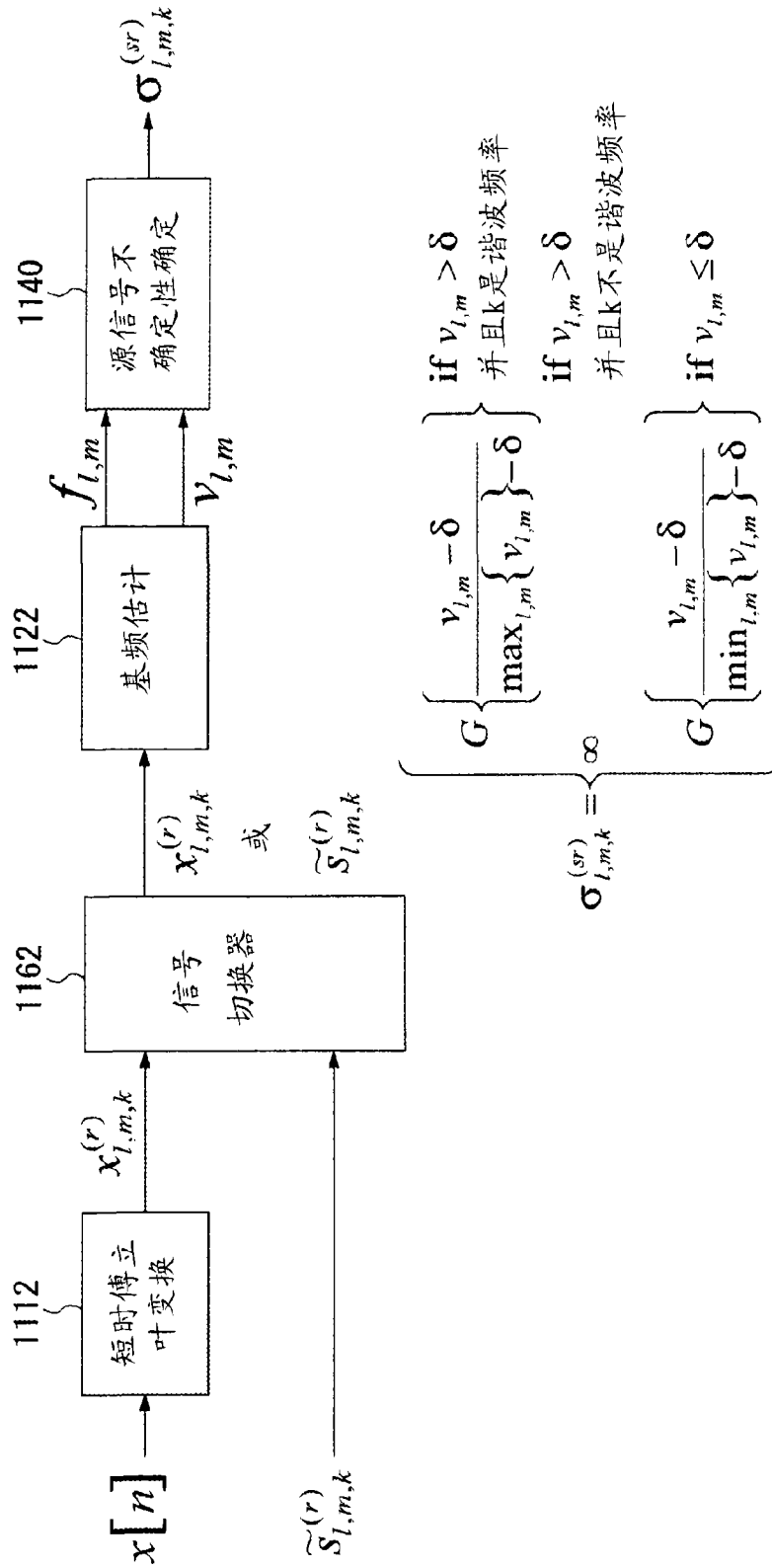


图 11

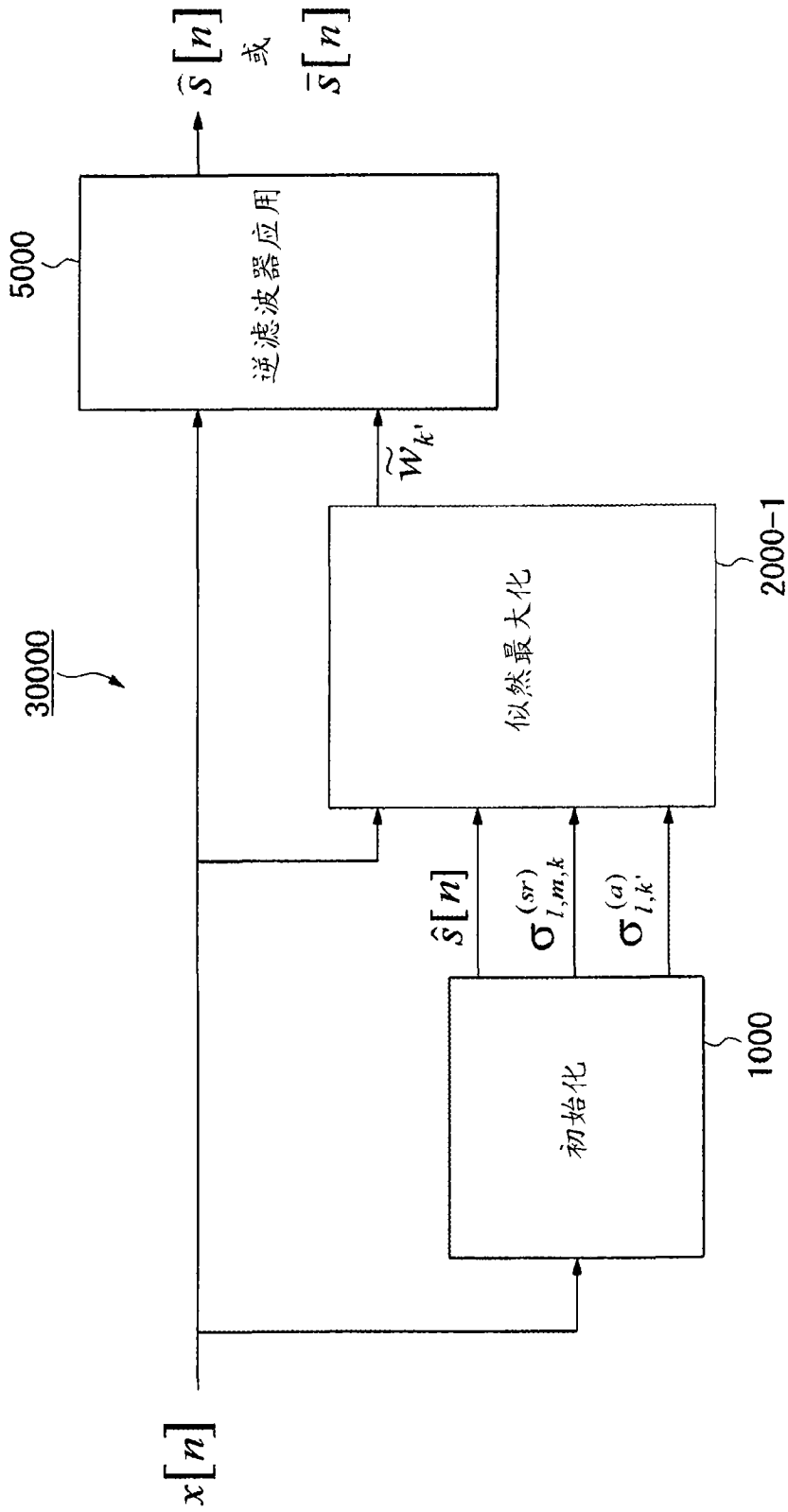


图 12

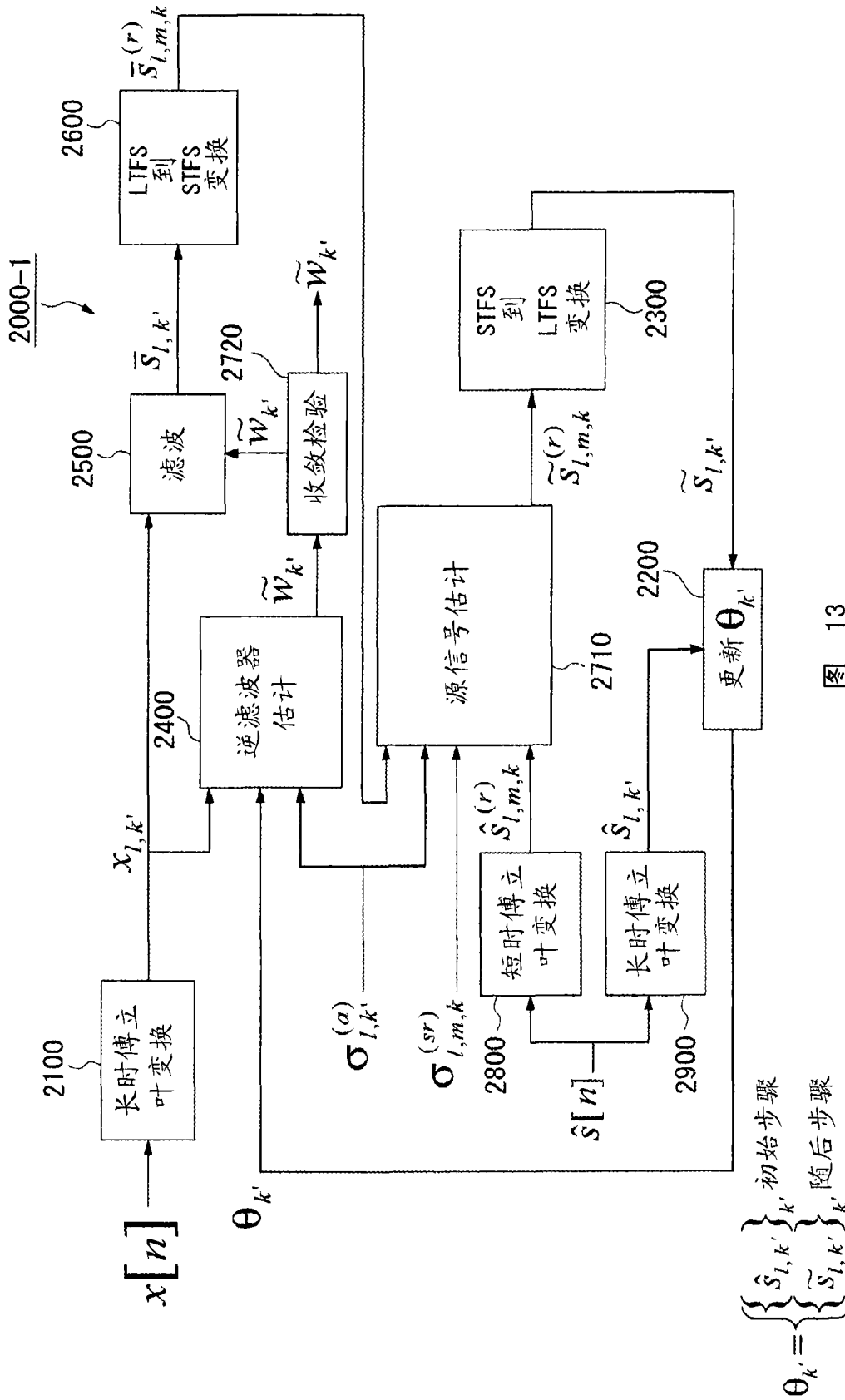


图 13

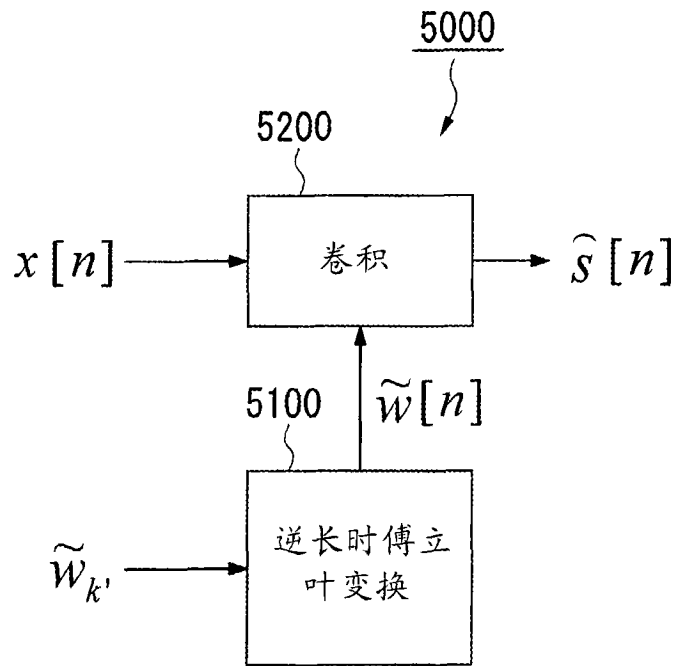


图 14

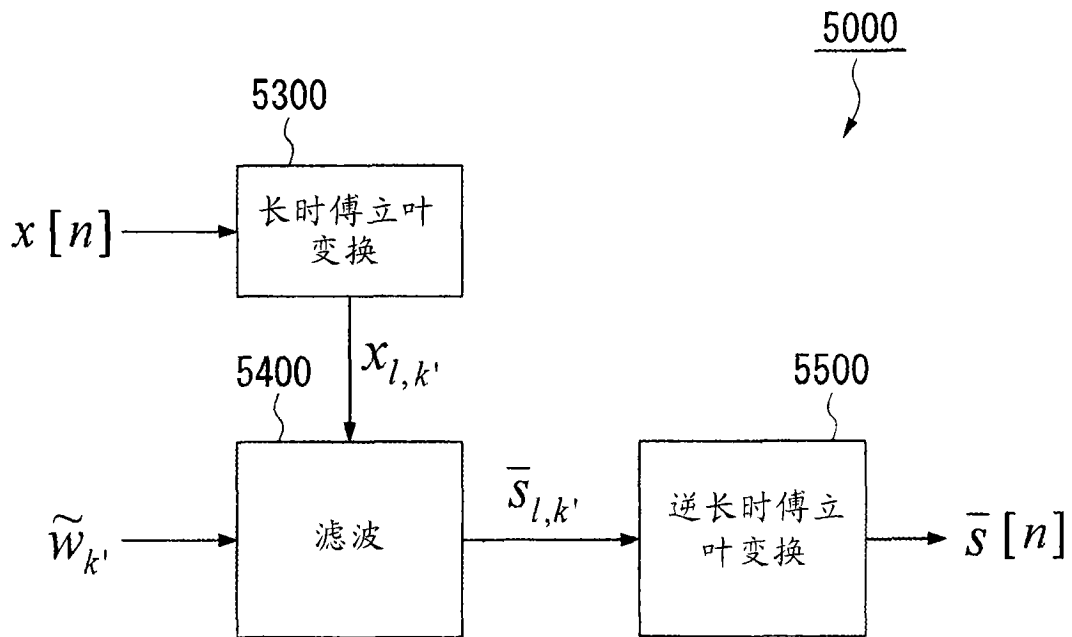


图 15

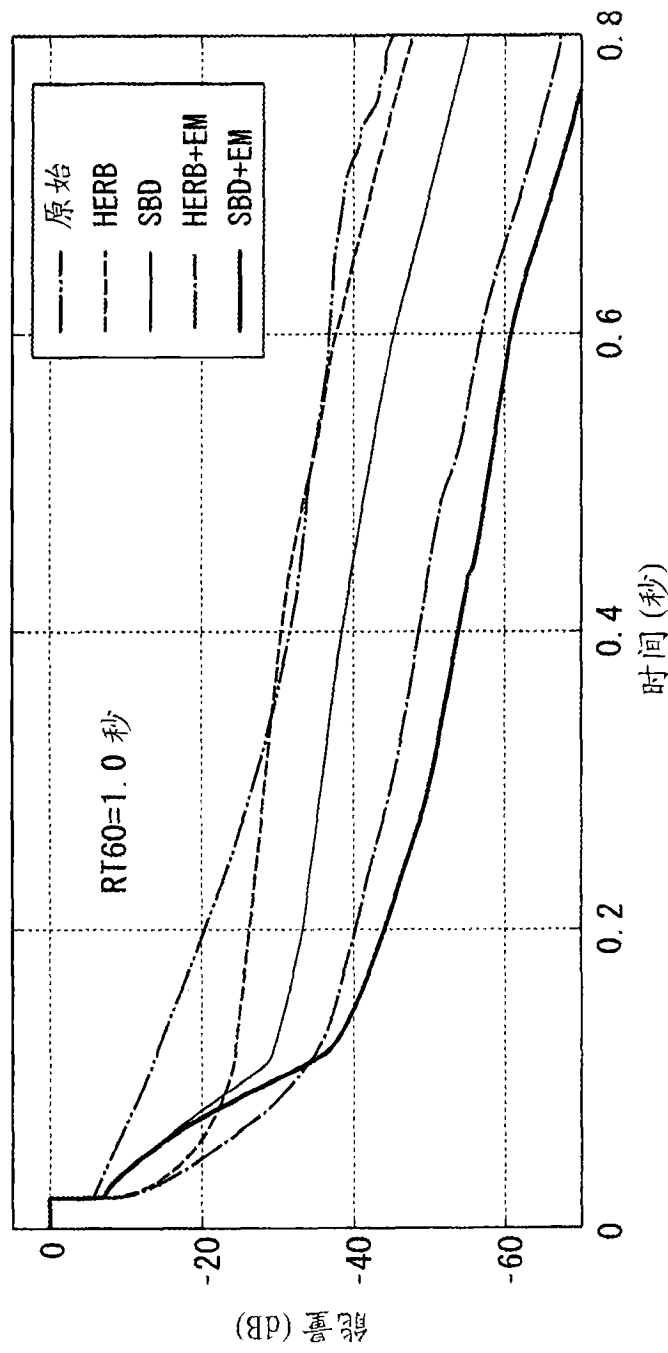


图 16A

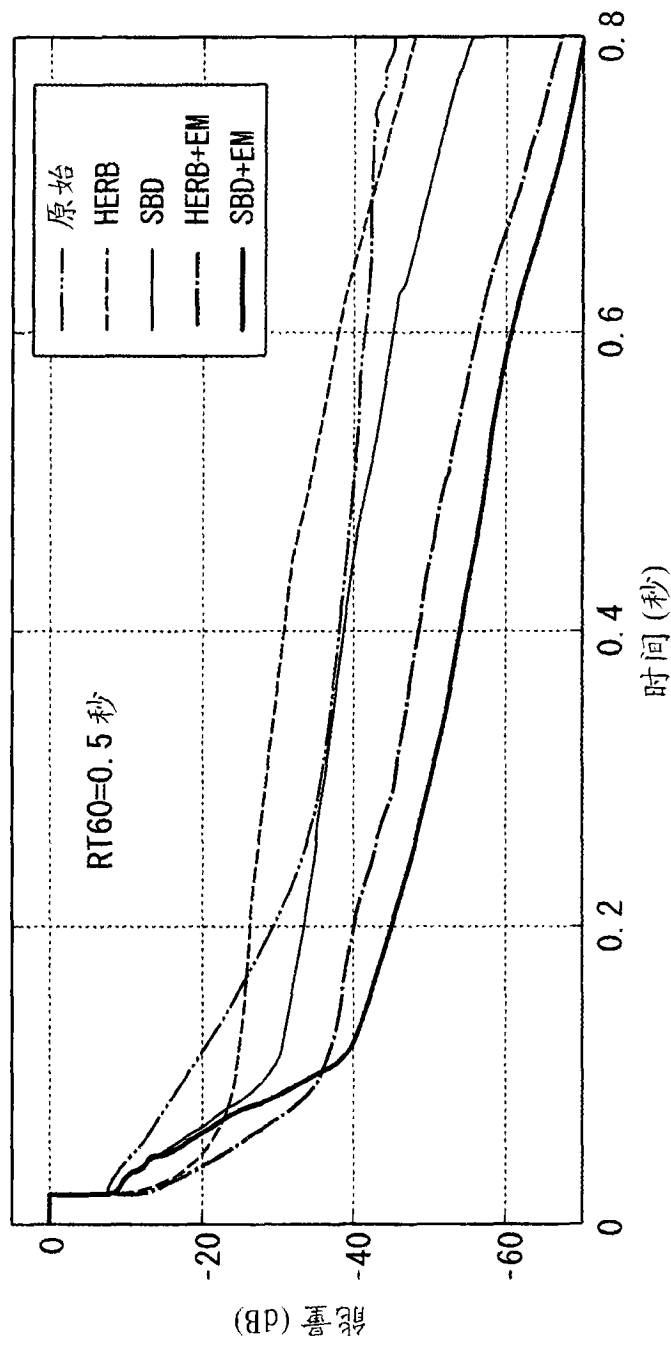


图 16B

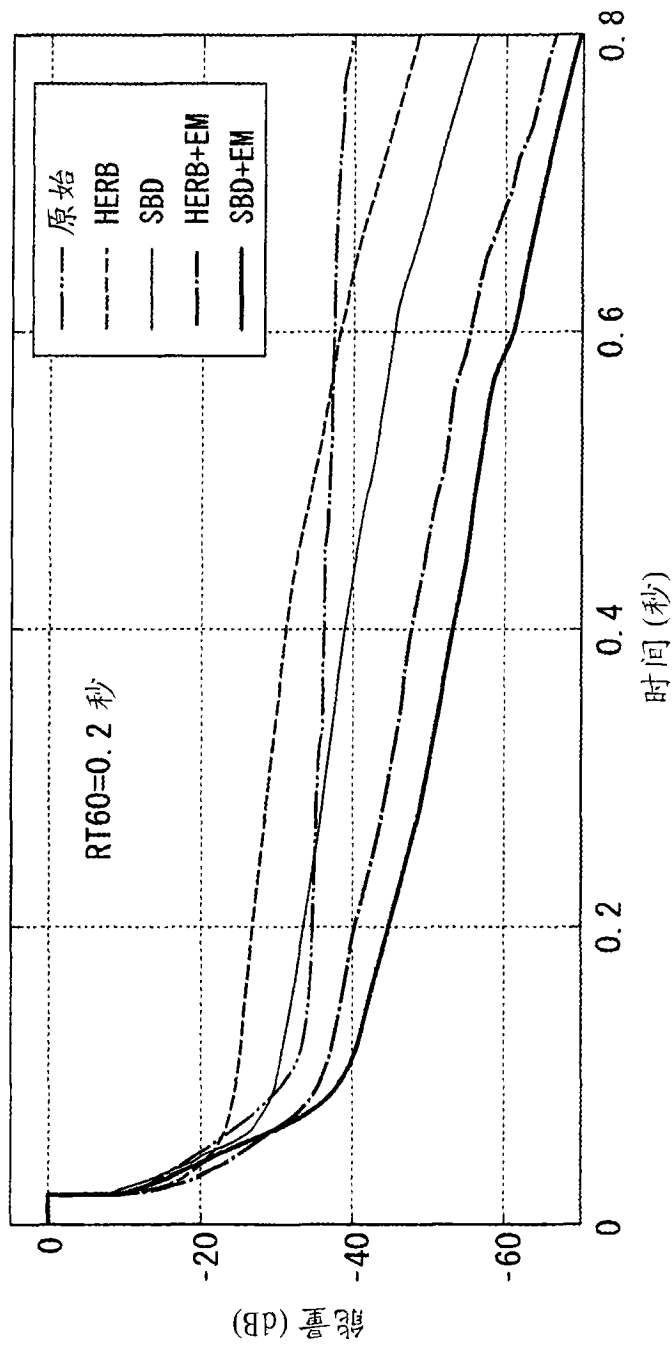


图 16C

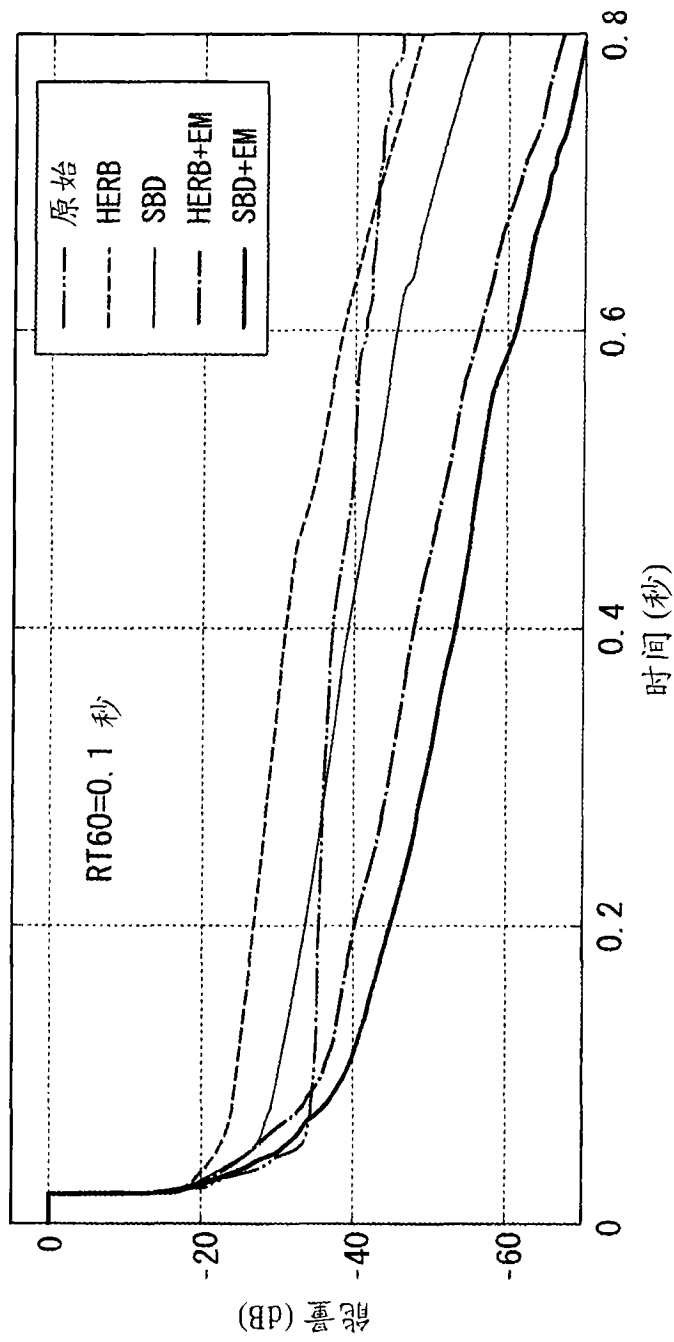


图 16D

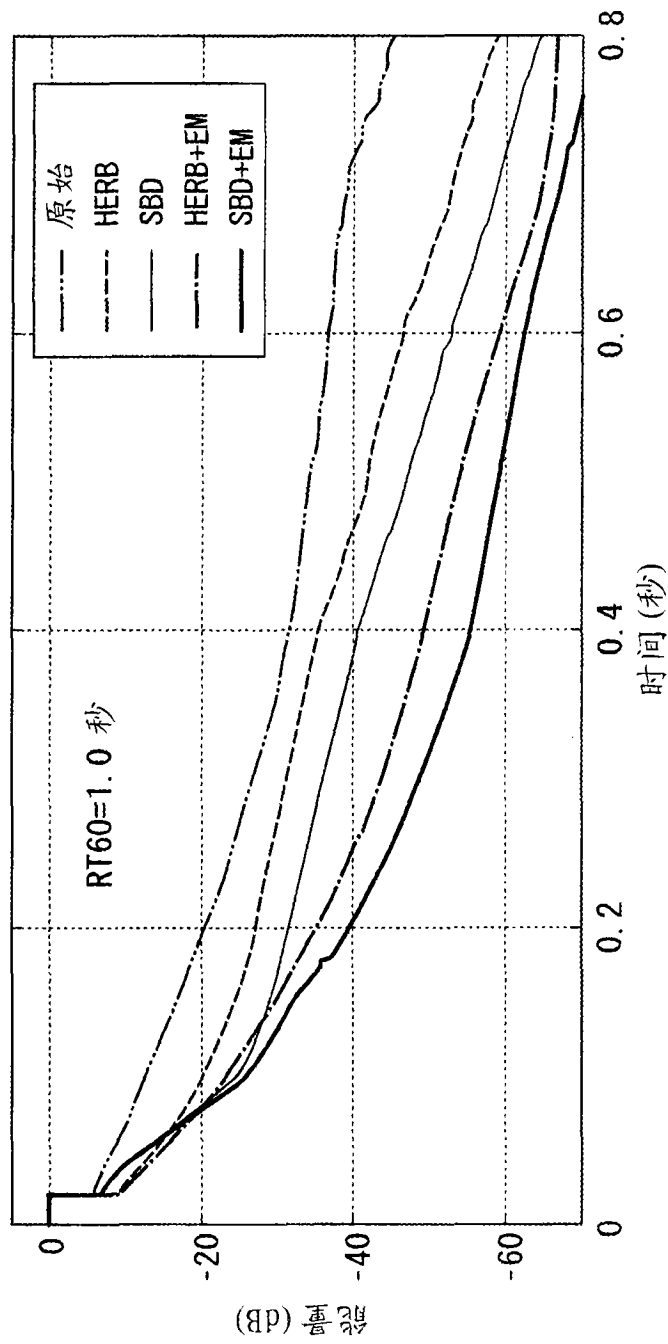


图 16E

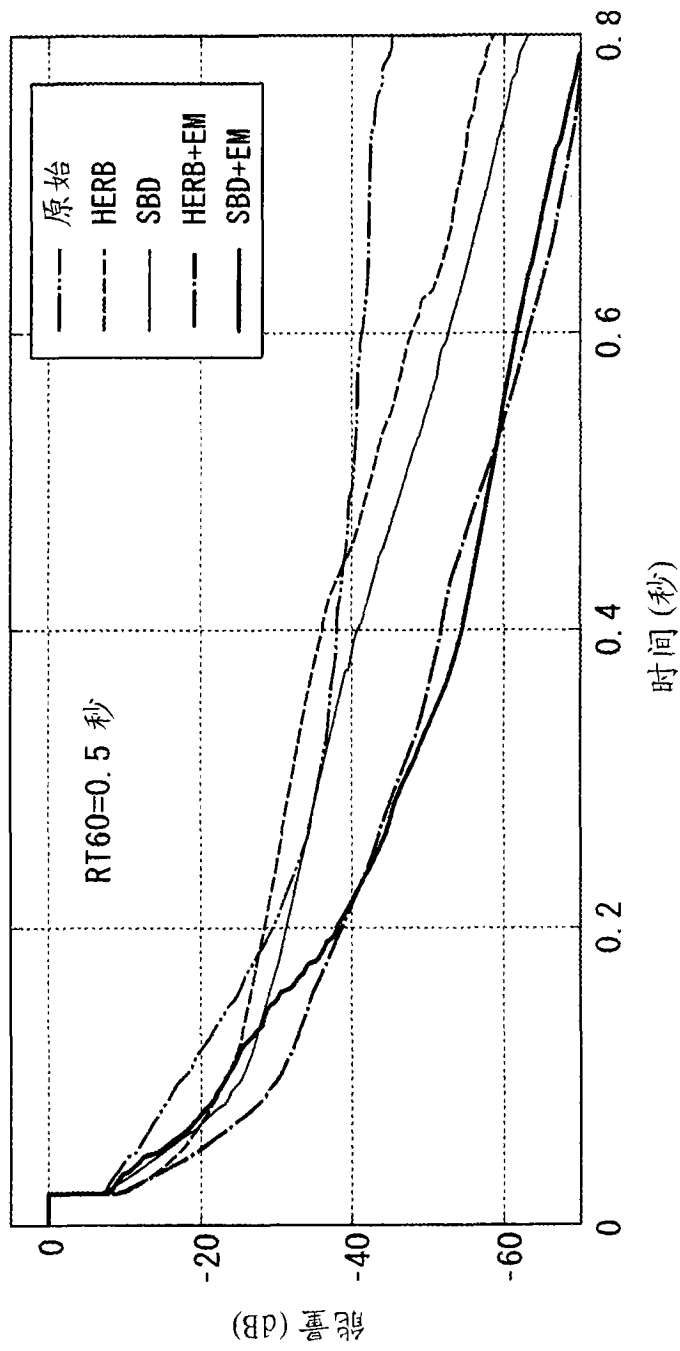


图 16F

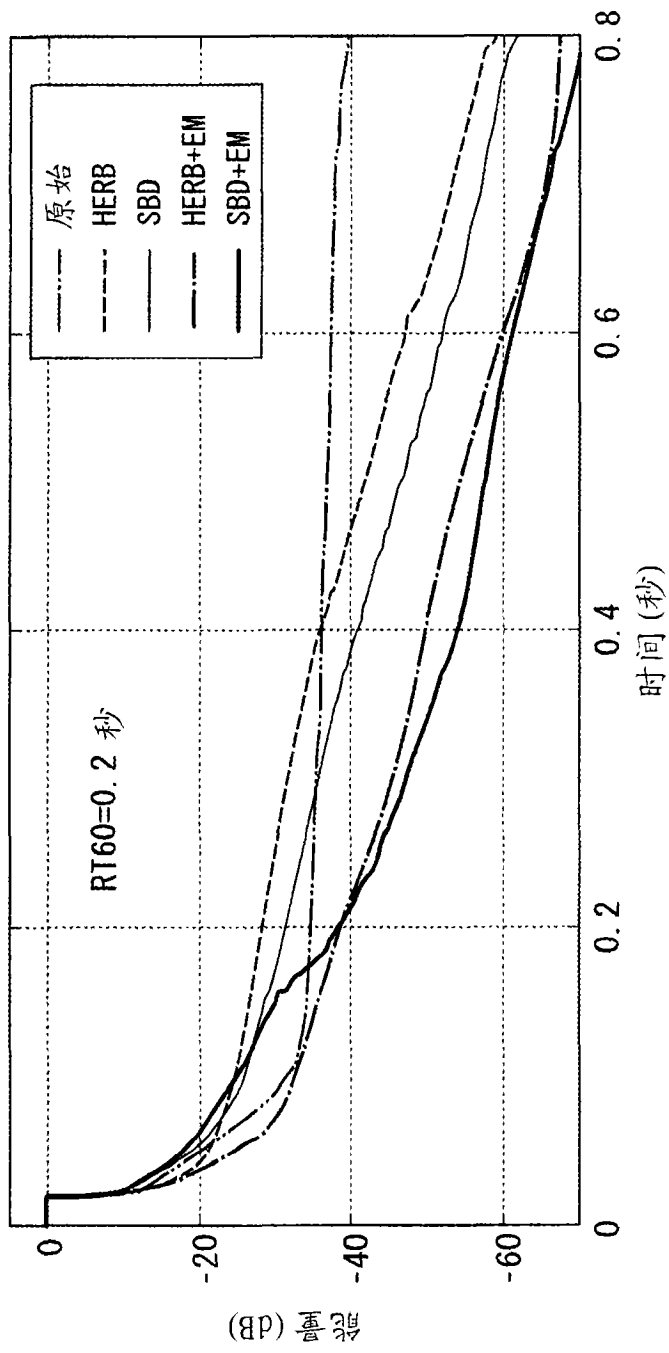


图 16G

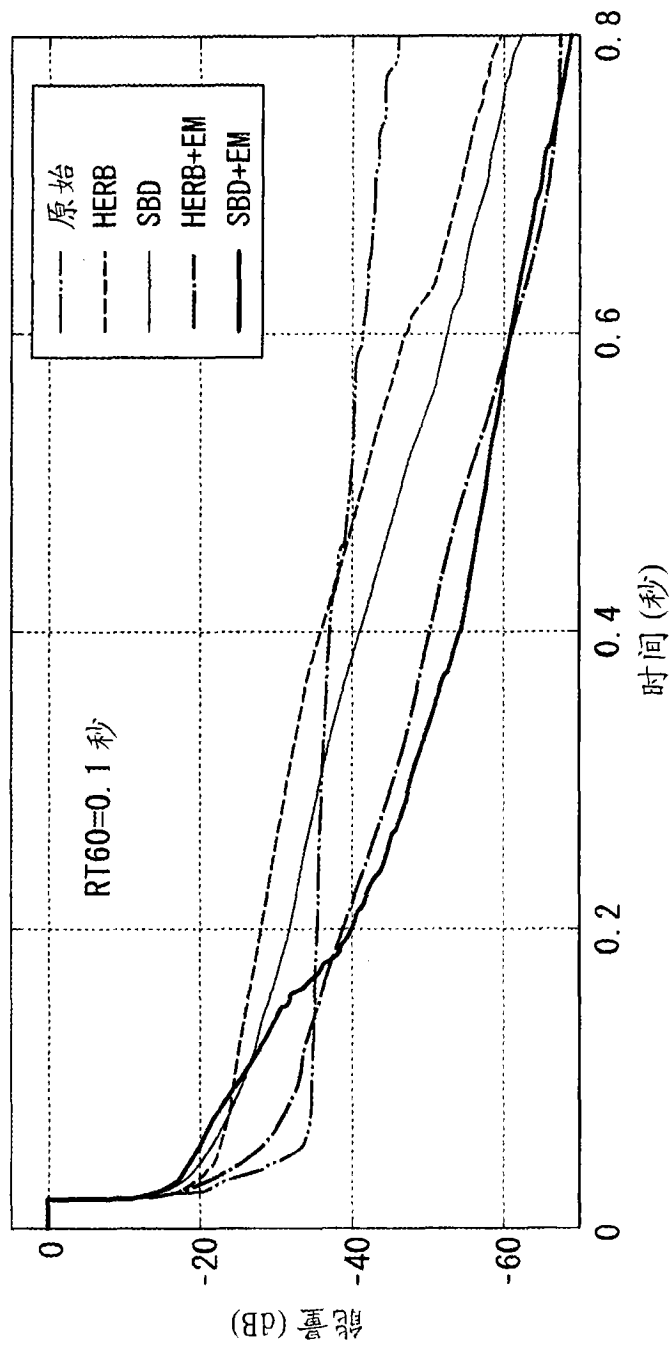


图 16H