



(12) 发明专利

(10) 授权公告号 CN 103354937 B

(45) 授权公告日 2015. 07. 29

(21) 申请号 201280008279. 7

(51) Int. Cl.

(22) 申请日 2012. 02. 08

G10L 21/0208(2013. 01)

(30) 优先权数据

61/441, 611 2011. 02. 10 US

(56) 对比文件

DE 4405723 A1, 1995. 08. 24,

CN 1140869 A, 1997. 01. 22,

(85) PCT国际申请进入国家阶段日

2013. 08. 09

US 2006/0270467 A1, 2006. 11. 30,

US 6351731 B1, 2002. 02. 26,

(86) PCT国际申请的申请数据

PCT/US2012/024372 2012. 02. 08

CN 1149945 A, 1997. 05. 14,

(87) PCT国际申请的公布数据

W02012/109385 EN 2012. 08. 16

审查员 陈红红

(73) 专利权人 杜比实验室特许公司

地址 美国加利福尼亚

(72) 发明人 G·N·迪金斯

(74) 专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 欧阳帆

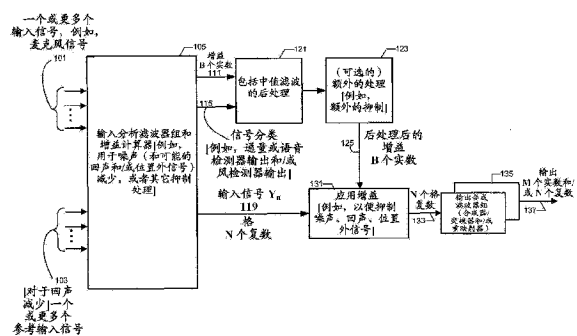
权利要求书6页 说明书27页 附图8页

(54) 发明名称

包括噪声抑制增益的中值滤波的后处理

(57) 摘要

本发明涉及对用于应用于音频信号的分带的增益进行后处理的方法、用于分带的经后处理的增益的设备、以及包括在运行时实现方法的指令的有形计算机可读存储介质。分带的增益通过对一个或多个输入音频信号进行输入处理来确定。该方法包括对分带的增益进行后处理以产生分带的经后处理的增益，产生对于特定频带的特定的经后处理的增益，包括使用来自用于与所述特定频带相邻的频带的增益值的增益值的中值滤波。后处理的至少一个特性依赖于一个或多个输入音频信号的分类。



1. 一种对原始的分带的增益 (111) 应用后处理来改善用于应用于一个或多个音频信号 (119) 的原始的分带的增益 (111) 以便实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或多个的方法 (310), 所述方法 (310) 包括如下步骤:

对原始的分带的增益 (111) 应用后处理 (315) 来产生用于应用于一个或多个音频信号 (119) 的分带的经后处理的增益 (125), 原始的分带的增益 (111) 通过对一个或多个输入音频信号 (101) 进行输入处理以便在多个频带处产生原始的分带的增益 (111) 来确定, 所述多个频带中的一些频带包括多于一个频率格, 原始的分带的增益 (111) 用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或多个,

其中对于特定频带的经后处理的增益 (125) 的产生至少包括使用用于与所述特定频带相邻的频带的原始的增益值的中值滤波 (315), 由此产生经中值滤波的增益,

其中后处理根据一个或多个特性, 所述一个或多个特性包括用于中值滤波 (315) 的宽度以及结束条件, 以及

其中中值滤波的宽度以及中值滤波的结束条件中的至少一个依赖于一个或多个输入音频信号 (101) 的信号分类 (115)。

2. 一种处理一个或多个输入音频信号 (101) 的方法, 所述方法包括如下步骤:

对一个或多个输入音频信号 (101) 进行输入处理, 以便在多个频带处产生原始的分带的增益 (111), 所述多个频带中的一些频带包括多于一个频率格, 原始的分带的增益 (111) 用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或多个; 以及

将后处理 (315) 应用于原始的分带的增益 (111), 以便产生用于应用于一个或多个输入音频信号 (101) 的分带的经后处理的增益 (125),

其中对于特定频带的经后处理的增益 (125) 的产生至少包括使用用于与所述特定频带相邻的频带的原始的增益值的中值滤波, 由此产生经中值滤波的增益,

其中后处理根据一个或多个特性, 所述一个或多个特性包括用于中值滤波的宽度以及结束条件, 以及

其中中值滤波的宽度以及中值滤波的结束条件中的至少一个依赖于一个或多个输入音频信号 (101) 的信号分类 (115)。

3. 如权利要求 2 所述的方法, 其中后处理还包括经中值滤波的增益的频带到频带平滑和跨越时间平滑 (317) 中的至少一个。

4. 如权利要求 3 所述的方法, 其中频带到频带平滑和跨越时间平滑 (317) 中的至少一个依赖于信号分类 (115)。

5. 如权利要求 1 到 4 中任何一个所述的方法, 其中信号分类 (115) 包括一个或多个输入音频信号 (101) 是否分类为风。

6. 如权利要求 1 到 4 中任何一个所述的方法, 其中用于中值滤波的宽度依赖于信号分类。

7. 如权利要求 1 到 4 中任何一个所述的方法, 其中信号分类 (115) 包括一个或多个输入音频信号 (101) 是否分类为语音。

8. 如权利要求 1 到 4 中任何一个所述的方法,其中信号分类 (115) 包括一个或更多个输入音频信号 (101) 是否分类为噪声。

9. 如权利要求 1 到 4 中任何一个所述的方法,其中频带为感知尺度或对数尺度。

10. 如权利要求 1 到 4 中任何一个所述的方法,其中输入处理要确定用于减少噪声的原始的分带的增益 (111)。

11. 如权利要求 1 到 4 中任何一个所述的方法,其中输入处理要根据多于一个输入音频信号 (101) 来确定用于减少噪声和位置外信号的原始的分带的增益 (111)。

12. 如权利要求 1 到 4 中任何一个所述的方法,其中输入处理要根据一个或更多个输入音频信号 (101) 以及一个或更多个参考信号 (103) 来确定原始的分带的增益 (111),所确定的增益 (111) 用于减少噪声和回声。

13. 如权利要求 1 到 4 中任何一个所述的方法,其中输入处理要确定用于基于感知域的调平、基于感知域的动态范围控制、和基于感知域的动态均衡中的一个或更多个的原始的分带的增益 (111)。

14. 一种用于处理一个或更多个输入音频信号 (101) 的设备,所述设备包括:

用于对一个或更多个输入音频信号 (101) 进行输入处理,以便在多个频带处产生原始的分带的增益 (111) 的部件,所述多个频带中的一些频带包括多于一个频率格,原始的分带的增益 (111) 用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个;以及

用于将后处理 (315) 应用于原始的分带的增益 (111),以便产生用于应用于一个或更多个输入音频信号 (101) 的分带的经后处理的增益 (125) 的部件,

其中用于应用后处理来生成对于特定频带的经后处理的增益 (125) 的部件至少包括用于使用用于与所述特定频带相邻的频带的原始的增益值的中值滤波,由此产生经中值滤波的增益的部件,

其中后处理根据一个或更多个特性,所述一个或更多个特性包括用于中值滤波的宽度以及结束条件,以及

其中中值滤波的宽度以及中值滤波的结束条件中的至少一个依赖于一个或更多个输入音频信号 (101) 的信号分类 (115)。

15. 如权利要求 14 所述的设备,其中后处理还包括经中值滤波的增益的频带到频带平滑和跨越时间平滑 (317) 中的至少一个。

16. 如权利要求 15 所述的设备,其中频带到频带平滑和跨越时间平滑 (317) 中的至少一个依赖于信号分类 (115)。

17. 如权利要求 14 到 16 中任何一个所述的设备,其中信号分类 (115) 包括一个或更多个输入音频信号 (101) 是否分类为风。

18. 如权利要求 14 到 16 中任何一个所述的设备,其中用于中值滤波的宽度依赖于信号分类。

19. 如权利要求 14 到 16 中任何一个所述的设备,其中信号分类 (115) 包括一个或更多个输入音频信号 (101) 是否分类为语音。

20. 如权利要求 14 到 16 中任何一个所述的设备,其中信号分类 (115) 包括一个或更多个输入音频信号 (101) 是否分类为噪声。

21. 如权利要求 14 到 16 中任何一个所述的设备,其中频带为感知尺度或对数尺度。

22. 如权利要求 14 到 16 中任何一个所述的设备,其中用于输入处理的部件要确定用于减少噪声的原始的分带的增益 (111)。

23. 如权利要求 14 到 16 中任何一个所述的设备,其中用于输入处理的部件要根据多于一个输入音频信号 (101) 来确定用于减少噪声和位置外信号的原始的分带的增益 (111)。

24. 如权利要求 14 到 16 中任何一个所述的设备,其中用于输入处理的部件要根据一个或更多个输入音频信号 (101) 以及一个或更多个参考信号 (103) 来确定原始的分带的增益 (111),所确定的增益 (111) 用于减少噪声和回声。

25. 如权利要求 14 到 16 中任何一个所述的设备,其中用于输入处理的部件要确定用于基于感知域的调平、基于感知域的动态范围控制、和基于感知域的动态均衡中的一个或更多个的原始的分带的增益 (111)。

26. 一种用于对原始的分带的增益 (111) 应用后处理来改善用于应用于一个或更多个音频信号 (119) 的原始的分带的增益 (111) 以便实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个的设备 (121),所述设备 (121) 包括:

用于对原始的分带的增益 (111) 应用后处理 (315) 来产生用于应用于一个或更多个音频信号 (119) 的分带的经后处理的增益 (125) 的部件,原始的分带的增益 (111) 通过对一个或更多个输入音频信号 (101) 进行输入处理以便在多个频带处产生原始的分带的增益 (111) 来确定,所述多个频带中的一些频带包括多于一个频率格,原始的分带的增益 (111) 用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个,

其中用于应用后处理来产生对于特定频带的经后处理的增益 (125) 的部件至少包括用于使用用于与所述特定频带相邻的频带的原始的增益值的中值滤波 (315),由此产生经中值滤波的增益的部件,

其中后处理根据一个或更多个特性,所述一个或更多个特性包括用于中值滤波 (315) 的宽度以及结束条件,以及

其中中值滤波的宽度以及中值滤波的结束条件中的至少一个依赖于一个或更多个输入音频信号 (101) 的信号分类 (115)。

27. 一种对用于应用于音频信号 (119) 的原始的分带的增益 (111) 进行后处理的设备 (121),所述设备 (121) 包括:

后处理器 (121),接受通过由输入处理器 (105) 对一个或更多个输入音频信号 (101) 进行输入处理来确定的原始的分带的增益 (111),所述后处理器 (121) 对原始的分带的增益 (111) 应用后处理以便产生用于应用于一个或更多个输入音频信号 (101) 的分带的经后处理的增益 (125),所述输入处理用于在多个频带处产生原始的分带的增益 (111),所述多个频带中的一些频带包括多于一个频率格,原始的分带的增益 (111) 用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个,

其中分带的经后处理的增益 (125) 用于应用于一个或更多个输入音频信号 (101),

其中后处理器 (121) 包括用于实现原始的分带的增益 (111) 的中值滤波的中值滤波器

(305), 由此产生经中值滤波的增益,

其中通过后处理器 (121) 进行的用于特定频带的经后处理的增益 (125) 的产生包括使用用于与特定频带相邻的频带的原始的增益值的中值滤波,

其中后处理根据一个或更多个特性, 所述一个或更多个特性包括用于中值滤波的宽度以及结束条件, 以及

其中中值滤波的宽度以及中值滤波的结束条件中的至少一个依赖于一个或更多个输入音频信号 (101) 的信号分类 (115)。

28. 一种处理一个或更多个输入音频信号的设备, 所述设备包括:

输入处理器 (105), 接受一个或更多个输入音频信号 (101) 并且对一个或更多个输入音频信号 (101) 进行输入处理, 以便在多个频带处产生原始的分带的增益 (111), 所述多个频带中的一些频带包括多于一个频率格, 原始的分带的增益 (111) 用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个;

后处理器 (121), 该后处理器接受原始的分带的增益 (111) 并且将后处理应用于原始的分带的增益 (111) 以便产生用于应用于一个或更多个输入信号 (101) 的经后处理的原始的分带的增益 (125),

其中后处理器 (121) 包括用于实现原始的分带的增益 (111) 的中值滤波的中值滤波器 (305), 由此产生经中值滤波的增益,

其中通过后处理器 (121) 进行的用于特定频带的经后处理的增益 (125) 的产生包括使用用于与特定频带相邻的频带的原始的增益值的中值滤波,

其中后处理根据一个或更多个特性, 所述一个或更多个特性包括用于中值滤波的宽度以及结束条件, 以及

其中中值滤波的宽度以及中值滤波的结束条件中的至少一个依赖于一个或更多个输入音频信号 (101) 的信号分类 (115)。

29. 如权利要求 27 或 28 所述的设备, 其中后处理器 (121) 包括用于使经中值滤波的增益平滑的平滑滤波器 (307), 该平滑包括频带到频带平滑和跨越时间平滑中的至少一个。

30. 如权利要求 27 到 28 中任何一个所述的设备, 还包括用于产生一个或更多个输入音频信号 (101) 的信号分类 (115) 的信号分类器 (423), 其中中值滤波的宽度依赖于一个或更多个输入音频信号 (101) 的信号分类 (115)。

31. 如权利要求 30 所述的设备, 其中信号分类器 (423) 包括语音活动检测器, 使得信号分类 (115) 包括一个或更多个输入音频信号 (101) 是否分类为语音。

32. 如权利要求 27 到 28 中任何一个所述的设备, 其中中值滤波的宽度依赖于一个或更多个输入音频信号 (101) 的谱通量。

33. 如权利要求 27 到 28 中任何一个所述的设备, 其中用于特定频带的中值滤波的宽度依赖于特定频带。

34. 如权利要求 27 到 28 中任何一个所述的设备, 其中频带为感知尺度或对数尺度。

35. 如权利要求 27 到 28 中任何一个所述的设备, 其中中值滤波依赖于一个或更多个输入音频信号的一个或更多个分类。

36. 如权利要求 27 到 28 中任何一个所述的设备, 其中通过输入处理确定的原始的分带

的增益 (111) 用于减少噪声。

37. 如权利要求 27 到 28 中任何一个所述的设备, 其中通过输入处理确定的原始的分带的增益 (111) 是根据多于一个输入音频信号 (101) 确定的并且用于减少噪声和位置外信号。

38. 如权利要求 27 到 28 中任何一个所述的设备, 其中输入处理器 (105) 还接受一个或更多个参考信号 (103), 并且其中通过输入处理确定的原始的分带的增益 (111) 是根据一个或更多个输入音频信号 (101) 以及一个或更多个参考信号 (103) 确定的, 并且用于减少噪声和回声。

39. 如权利要求 27 到 28 中任何一个所述的设备, 其中通过输入处理确定的原始的分带的增益 (111) 用于基于感知域的调平、基于感知域的动态范围控制、和基于感知域的动态均衡中的一个或更多个。

40. 一种用于对原始的分带的增益 (111) 进行后处理以便产生用于应用于音频信号的分带的经后处理的增益 (125) 的系统 (121), 所述系统包括:

用于对原始的分带的增益 (111) 进行后处理来产生分带的经后处理的增益 (125) 的装置 (121), 原始的分带的增益 (111) 由用于通过对一个或更多个输入音频信号 (101) 进行输入处理以便在多个频带处产生原始的分带的增益 (111) 的装置 (105) 来确定, 所述多个频带中的一些频带包括多于一个频率格, 原始的分带的增益 (111) 用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个,

其中分带的经后处理的增益 (125) 被应用于一个或更多个音频信号以便实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个,

其中对于特定频带的经后处理的增益 (125) 的产生至少包括使用用于与所述特定频带相邻的频带的原始的增益值的中值滤波, 由此产生经中值滤波的增益,

其中由用于后处理的装置 (121) 应用的后处理根据一个或更多个特性, 所述一个或更多个特性包括用于中值滤波的宽度和结束条件, 以及

其中中值滤波的宽度以及中值滤波的结束条件中的至少一个依赖于一个或更多个输入音频信号 (101) 的信号分类 (115)。

41. 一种用于处理一个或更多个输入音频信号 (101) 的系统, 所述系统包括:

用于对一个或更多个输入音频信号 (101) 进行输入处理以便在多个频带处产生原始的分带的增益 (111) 的装置 (105), 所述多个频带中的一些频带包括多于一个频率格, 原始的分带的增益 (111) 用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个;

用于对原始的分带的增益 (111) 进行后处理以便产生用于应用于一个或更多个输入音频信号 (101) 的分带的经后处理的增益 (125) 的装置 (121),

其中对于特定频带的经后处理的增益 (125) 的产生至少包括使用用于与所述特定频带相邻的频带的原始的增益值的中值滤波, 由此产生经中值滤波的增益,

通过用于后处理的装置 (121) 应用的后处理根据一个或更多个特性, 所述一个或更多个特性包括用于中值滤波的宽度以及结束条件,

其中中值滤波的宽度以及中值滤波的结束条件中的至少一个依赖于一个或更多个输入音频信号 (101) 的信号分类 (115)。

42. 如权利要求 40 或 41 所述的系统, 其中后处理还包括频带到频带平滑和跨越时间平滑中的至少一个。

43. 如权利要求 42 所述的系统, 其中频带到频带平滑和跨越时间平滑中的至少一个依赖于信号分类 (115)。

44. 如权利要求 40 到 41 中任何一个所述的系统, 其中信号分类 (115) 包括一个或更多个输入音频信号是否分类为风。

45. 如权利要求 40 到 41 中任何一个所述的系统, 其中用于中值滤波的宽度依赖于信号分类。

46. 如权利要求 40 到 41 中任何一个所述的系统, 其中信号分类 (115) 包括一个或更多个输入音频信号 (101) 是否分类为语音。

47. 如权利要求 40 到 41 中任何一个所述的系统, 其中信号分类 (115) 包括一个或更多个输入音频信号 (101) 是否分类为噪声。

48. 如权利要求 40 到 41 中任何一个所述的系统, 其中频带为感知尺度或对数尺度。

49. 如权利要求 40 到 41 中任何一个所述的系统, 其中输入处理要确定用于减少噪声的原始的分带的增益 (111)。

50. 如权利要求 40 到 41 中任何一个所述的系统, 其中输入处理要根据多于一个输入音频信号 (101) 确定用于减少噪声和位置外信号的原始的分带的增益 (111)。

51. 如权利要求 40 到 41 中任何一个所述的系统, 其中输入处理要根据一个或更多个输入音频信号 (101) 以及一个或更多个参考信号 (103) 确定原始的分带的增益 (111), 所确定的增益 (111) 用于减少噪声和回声。

52. 如权利要求 40 到 41 中任何一个所述的系统, 其中输入处理要确定用于基于感知域的调平、基于感知域的动态范围控制、和基于感知域的动态均衡中的一个或更多个的原始的分带的增益 (111)。

包括噪声抑制增益的中值滤波的后处理

[0001] 相关专利申请

[0002] 本申请要求 2011 年 2 月 10 日提交的美国临时申请 No. 61/441,611 的优先权,其内容全部通过参考被并入于此。

[0003] 本申请与同时提交的题为“COMBINED SUPPRESSION OF NOISE, ECHO, and OUT-OF-LOCATION SIGNALS”的国际申请 No. PCT/US2012/024370 有关,该国际申请 No. PCT/US2012/024370 具有案卷参考号(Docket/Ref. No.) D09110AW001,并且也要求 2011 年 2 月 10 日提交的美国临时申请 No. 61/441,611 的优先权。这个国际申请 No. PCT/US2012/024370 的内容全部通过参考被并入于此。

[0004] 本申请与下面的均于 2011 年 2 月 10 日提交的美国临时专利申请有关:

[0005] •美国临时专利申请 No. 61/441,396,题为“VECTOR NOISE CANCELLATION”,发明人为 Jon C. Taenzer,案卷:60175-0060,客户参考号 No. A09070USP1。

[0006] •美国临时专利申请 No. 61/441,397,题为“VECTOR NOISE CANCELLATION”,发明人为 Jon C. Taenzer 和 Steven H. Puthuff,案卷:60175-0087,客户参考号 No. A09071USP1。

[0007] •美国临时专利申请 No. 61/441,528,题为“MULTI-CHANNEL WIND NOISE SUPPRESSION SYSTEM AND METHOD”,发明人为 Jon C. Taenzer,案卷:435372-000052,客户参考号 No. D10025USP1。

[0008] •美国临时专利申请 No. 61/441,551,题为“SYSTEM AND METHOD FOR WIND DETECTION AND SUPPRESSION”,发明人为 Glenn N. Dickins 和 Leif Jonas Samuelsson,案卷:435372-000053,客户参考号 No: D10089USP1,这个临时专利申请 No. 61/441,551 在此被称为“Wind Detection/Suppression Application”。

[0009] •美国临时专利申请 No. 61/441,633,题为“SPATIAL ADAPTATION FOR MULTI-MICROPHONE SOUND CAPTURE”,发明人为 Leif Jonas Samuelsson,案卷:435372-000054,客户参考号 No. D11008USP1。

技术领域

[0010] 本公开通常涉及信号处理,特别地涉及音频信号的信号处理。

背景技术

[0011] 声信号处理现今可适用于提高诸如来自麦克风的语音信号的质量。作为一个示例,许多装置(诸如电话听筒(handsets))在存在回声(echo)的源(例如,扬声器)的情况下操作。此外,来自麦克风的信号可能在嘈杂的环境中(例如,在汽车中或在存在其它噪声的情况下)出现。此外,可能有来自干扰位置的声音,例如,别人进行的位置外的(out-of-location)谈话或位置外的干扰、风等。因此声信号处理是发明的一个重要领域。

[0012] 已知处理系统用于抑制噪声、抑制回声和增加空间选择性(spatial selectivity)中的一个或多个。声噪声减少系统典型地包括用于确定抑制概率指示器(indicator)例如作为一组噪声减少增益的增益计算模块和噪声估计器,该组增益在例如

一组频带上被确定,并且在变换到频域并且分带(banding)到该组频带之后被应用于(嘈杂的)输入音频信号以便衰减噪声分量。声噪声减少系统可以包括一个麦克风输入,或多个麦克风输入并且向下混合(downmix)(例如波束成形(beamform))以产生一个输入音频信号。声噪声减少系统可以还包括回声减少,并且可以还包括位置外信号减少。

[0013] 已知存在音乐的噪声,并且该音乐的噪声可能由于对带中的一些中的增益造成的随时间的短期错误而出现。这种增益中错误可以被认为是统计的离群值(outlier),即,跨越(across)一组带的增益的值统计上位于预期的范围外,因此似乎“被孤立”。

[0014] 这种统计的离群值可能出现在输入音频信号被变换和分带的其它类型的处理中。这种其它类型的处理包括基于感知域(perceptual domain)的调平(leveling)、基于感知域的动态范围控制、和考虑了依赖于音频信号的再现电平的音频的感知的变化的基于感知域的动态均衡。参见例如国际申请 PCT/US2004/016964,其作为 W02004111994 公开。如下是可能的,即用于调平和 / 或动态均衡的对于每个带确定的增益包括统计的离群值,例如,孤立的值,并且这种离群值可能导致假象,诸如音乐的噪声。

[0015] 对增益(例如,噪声减少增益)进行中值滤波、或者对跨越频带的增益进行调平和 / 或动态均衡可以减少音乐的噪声假象。DE4405723A1 描述了可能在谱相减的背景中出现的“音乐音调”的减少的中值滤波的使用。

[0016] 增益值可能在频率之间显著变化,并且在这种情况下,沿着频带运行相对宽的中值滤波具有破坏时间包络的连续性的风险,这是对于许多信号而言固有的性质并且对于感知也是关键的。虽然对于离群值提供更大的抗扰性,但是更长的中值滤波器可能减少处理的频谱选择性,并且可能地在跨越频率和时间的增益值中引入更大的不连续性或跳变。

[0017] 在这个背景技术部分中描述的方法是可以被实行的方法,但不一定是先前已经被设想或者实行的方法。因此,除非另有指明,否则不应该仅仅由于它们包括在该部分中而假设在该部分中描述的任何方法有资格作为现有技术。类似地,除非另有指明,否则不应该基于该部分而假设与一个或更多个方案联系的问题已经在任何现有技术中被认识到。

附图说明

[0018] 图 1 示出包括本发明的一个实施例的一组一个或更多个输入音频信号(例如,来自不同位置的麦克风的麦克风信号 101)的处理的一个示例。

[0019] 图 2 图解表示地示出本发明的中值滤波器的一个实施例的频率覆盖范围和原始的(raw)分带的(banded)增益的组。

[0020] 图 3A 示出根据本发明实施例的包括中值滤波器的后处理器的简化框图。

[0021] 图 3B 示出根据本发明实施例的包括中值滤波的后处理的方法的简化流程图。

[0022] 图 4 示出设备实施例的一个示例,所述设备实施例被配置为确定用于抑制噪声、以及在某些版本中同时抑制回声、以及在某些版本中同时抑制位置外信号的一组经后处理的增益。

[0023] 图 5 更详细地示出设备实施例的一个示例。

[0024] 图 6 示出包括空间敏感的语音活动(activity)检测器和风活动检测器的增益计算元件的示例实施例。

[0025] 图 7 示出操作处理设备以便抑制噪声和位置外信号以及在某些实施例中抑制回

声的方法的实施例的流程图。

[0026] 图 8 示出处理设备实施例的简化框图,所述处理设备实施例用于处理一个或更多个音频输入以便确定一组原始的增益、以便对原始的增益进行后处理(包括对所确定的原始的增益进行中值滤波)、以及以便产生已经通过应用经后处理的增益被修改的音频输出。

具体实施方式

[0027] 概述

[0028] 本发明的实施例包括方法、设备、以及在一个或更多个计算机可读的有形的介质中编码以便实现该方法的逻辑。

[0029] 一个实施例包括对原始的分带的增益应用后处理以便改善用于应用于一个或更多个输入音频信号的原始的分带的增益的方法。通过对一个或更多个输入音频信号进行输入处理来确定在包括一个或更多个频率格(bin)的多个频带处的原始的分带的增益。原始的分带的增益用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个。该方法包括对原始的分带的增益应用后处理以便产生分带的经后处理的增益。用于特定频带的特定的经后处理的增益的产生至少包括使用用于与该特定频带相邻的频带的增益值的中值滤波。后处理根据一个或更多个特性,包括用于中值滤波的宽度以及结束条件(end condition)。后处理的至少一个特性依赖于一个或更多个输入音频信号的信号分类。

[0030] 一个实施例包括处理一个或更多个输入音频信号的方法。该方法包括对一个或更多个输入音频信号进行输入处理以便确定用于应用于音频信号的原始的分带的增益,该原始的分带的增益在包括一个或更多个频率格的多个频带处。原始的分带的增益用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个。该方法还包括对原始的分带的增益应用后处理以便产生分带的经后处理的增益。用于特定频带的特定的经后处理的增益的产生至少包括使用用于与该特定频带相邻的频带的增益值的中值滤波。后处理根据一个或更多个特性,包括用于中值滤波的宽度以及结束条件。后处理的至少一个特性依赖于一个或更多个输入音频信号的信号分类。

[0031] 一个实施例包括用于对用于应用于一个或更多个输入音频信号的原始的分带的增益后处理的设备。通过对一个或更多个输入音频信号进行输入处理来确定在包括一个或更多个频率格的多个频带处的原始的分带的增益。原始的分带的增益用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个。该设备包括后处理器,该后处理器接受原始的分带的增益并且将后处理应用于原始的分带的增益,以便产生用于应用于一个或更多个输入信号的分带的经后处理的增益。后处理器包括用于实现原始的分带的增益的中值滤波的中值滤波器。通过后处理器进行的用于特定频带的特定的经后处理的增益的产生包括使用用于与该特定频带相邻的频带的增益值的中值滤波。后处理根据一个或更多个特性,包括用于中值滤波的宽度以及结束条件。后处理的至少一个特性依赖于一个或更多个输入音频信号的信号分类。

[0032] 一个实施例包括用于处理一个或更多个输入音频信号的设备。该设备包括输入处

理器,该输入处理器接受一个或更多个输入音频信号并且对一个或更多个输入音频信号进行输入处理以便产生在包括一个或更多个频率格的多个频带处的原始的分带的增益。原始的分带的增益用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个。该设备还包括后处理器,该后处理器接受原始的分带的增益并且将后处理应用于原始的分带的增益以便产生用于应用于一个或更多个输入信号的分带的经后处理的增益。后处理器包括用于实现原始的分带的增益的中值滤波的中值滤波器。通过后处理器进行的用于特定频带的特定的经后处理的增益的产生包括使用用于与该特定频带相邻的频带的增益值的中值滤波。后处理根据一个或更多个特性,包括用于中值滤波的宽度以及结束条件。后处理的至少一个特性依赖于一个或更多个输入音频信号的信号分类。

[0033] 一个实施例包括用于对原始的分带的增益进行后处理以便产生用于应用于音频信号的分带的经后处理的增益的系统。该系统包括用于对原始的分带的增益进行后处理以便产生分带的经后处理的增益的装置,原始的分带的增益是由用于对一个或更多个输入音频信号进行输入处理以便产生在包括一个或更多个频率格的多个频带处的原始的分带的增益的装置确定的。原始的分带的增益用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个。分带的经后处理的增益用于应用于一个或更多个音频信号。产生用于特定频带的特定的经后处理的增益至少包括使用用于与该特定频带相邻的频带的增益值的中值滤波。后处理根据一个或更多个特性,包括用于中值滤波的宽度以及结束条件。后处理的至少一个特性依赖于一个或更多个输入音频信号的信号分类。

[0034] 一个实施例包括用于处理一个或更多个输入音频信号的系统。该系统包括用于对一个或更多个输入音频信号进行输入处理以便产生在包括一个或更多个频率格的多个频带处的原始的分带的增益的装置。原始的分带的增益用于实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个。该系统还包括用于对原始的分带的增益后处理以便产生用于应用于一个或更多个输入音频信号的分带的经后处理的增益以便实现减少噪声、减少位置外信号、减少回声、基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个的装置。通过用于后处理的装置进行的用于特定频带的特定的经后处理的增益的产生至少包括使用用于与特定频带相邻的频带的增益值的中值滤波。通过用于后处理的装置应用的后处理根据一个或更多个特性,包括用于中值滤波的宽度和结束条件。后处理的至少一个特性依赖于一个或更多个输入音频信号的信号分类。

[0035] 在一些实施例中,后处理包括频带到频带平滑和跨越时间平滑(smoothing across time)中的至少一个。

[0036] 在一些实施例中,中值滤波的宽度和结束条件中的一个或两个依赖于一个或更多个输入音频信号的信号分类。在一些实施例中,分类包括输入音频信号是否可能是语音。在一些实施例中,分类包括输入音频信号是否可能是风。

[0037] 在一些实施例中,频带为感知尺度或对数尺度(on perceptual or logarithmic scale)。

[0038] 在一些实施例中,根据一个或更多个输入音频信号确定的原始的分带的增益用于

减少噪声。在一些实施例中,原始的分带的增益是根据多于一个输入音频信号确定的并且用于减少噪声和位置外信号。在一些实施例中,原始的分带的增益是根据一个或更多个输入音频信号和一个或更多个参考信号确定的,并且用于减少噪声和回声。

[0039] 一个实施例包括有形的计算机可读的存储介质,其包括指令,该指令在由处理系统的一个或更多个处理器运行时使得处理硬件实现如在此描述的对用于应用于音频信号的原始的分带的增益进行后处理的方法。

[0040] 一个实施例包括程序逻辑,该程序逻辑在由至少一个处理器运行时使得实现如在此描述的方法。

[0041] 特定的实施例可以提供这些方面、特征或优点中的所有、一些或没有。特定的实施例可以提供一个或更多个其它方面、特征或优点,其中的一个或更多个对于本领域技术人员而言可以根据在本申请中的附图、说明书和权利要求而容易明白。

[0042] 一些示例实施例

[0043] 本发明的一个方面是一个或更多个输入音频信号的处理,包括输入处理,以便产生用于噪声减少或者用于其它形式的输入信号改善的原始的增益。该处理包括应用后处理于原始的增益,包括用于增益平滑的原始的增益的中值滤波。本发明的另一个方面是后处理,包括通过输入处理确定的例如用于噪声减少或者用于其它输入处理的原始的增益的中值滤波。中值滤波器利用预定义数量的原始的增益值的中值(例如,通过特定的原始的增益值和预定义的一组相邻原始的增益值的中值)代替特定的原始的增益值。中值滤波器具有一个或更多个特性,例如,确定中值的值的数量以及结束条件。该特性中的至少一个可以是与数据相关的。因此,在在此描述的一些示例中,可以存在用于一种类型的数据(例如,可能是噪声的数据)的第一中值滤波器、以及用于另一种类型的数据(例如,可能是语音的数据)的不同的中值滤波器。

[0044] 图 1 示出包括本发明的一个实施例的一组一个或更多个输入音频信号(例如,来自不同位置的麦克风的麦克风信号 101)的处理的一个示例。该处理利用多个(例如 M 个)样本的时间帧。在简单的实施例中,仅仅存在一个输入,例如,一个麦克风,并且在另一实施例中,存在 P 表示的多个输入,例如,麦克风信号 101。输入处理器 105 接受采样的一个或更多个输入音频信号 101 并且形成用于多个(B 个)频带的一个或更多个输入音频信号 101 的分带的瞬时频域振幅度量(metric)119。在其中存在多于一个的输入音频信号的一些实施例中,度量 119 是从输入音频信号向下混合的。振幅度量表示谱内容(spectral content)。在此描述的许多实施例中,谱内容为功率谱方面的内容。然而,本发明不限于处理功率谱值。相反地,可以使用任何谱振幅相关的度量。例如,如果直接使用振幅谱,则这种谱内容有时被称为谱包络。因此,有时在说明书中使用短语“功率(或者其它振幅度量)谱”。

[0045] 在一个噪声减少实施例中,输入处理器 105 确定用于应用于瞬时振幅度量 119 的一组原始的分带的增益 111。在一个实施例中,输入处理还包括确定一个或更多个输入音频信号的信号分类,例如,如由语音活动检测器(VAD)确定的一个或更多个输入音频信号是否可能是语音的指示、和 / 或如由风活动检测器(WAD)确定的一个或更多个输入音频信号是否可能是风的指示、和 / 或如例如由超过阈值的谱通量指示的信号能量正快速地变化的指示。

[0046] 本发明的实施例的一个特征包括将后处理应用于原始的增益以便提高输出的质

量。在一个实施例中,后处理包括通过输入处理确定的原始的增益的中值滤波。中值滤波器考虑一组原始的增益并且输出作为该组原始的增益的中值的增益。对于每个帧确定一组 B 个原始的增益,使得存在在 B 个频带上的 B 个原始的增益的组的时间序列。在一个实施例中,中值滤波器跨越频率延伸。

[0047] 图 2 图解表示地示出多组原始的分带的增益,一个组用于当前时间、往回一个帧、往回两个帧、往回三个帧等中的每一个,并且还示出包括当前帧中的以频带 b_c 为中心的五个原始的增益值的示例中值滤波器的覆盖范围。对于滤波器宽度,我们意指频带域中的滤波器的宽度。

[0048] 返回到图 1,后处理产生被应用于瞬时功率(或者其它振幅度量) 119 的一组经后处理的增益 125 以便产生输出,例如作为多个经处理的频率格 133。输出合成滤波器组 135 (或者用于后续的编码,变换器/重映射器)将这些频率格转换为期望的输出 137。

[0049] 输入处理元件 105 包括输入分析滤波器组以及原始的增益计算器。用于一个输入音频信号 101 的情况的输入分析滤波器组包括用于将帧的样本变换成频率格的变换器以及用于形成大多数包括多个频率格的频带的分带元件。用于多个输入音频信号 101 的情况的输入分析滤波器组包括用于将每个输入音频信号的帧的样本变换成频率格的变换器、用于向下混合多个成单个信号的向下混合器(例如波束成形器)、以及用于形成大多数包括多个频率格的频带的分带元件。

[0050] 在一个实施例中,变换器实行短时间傅里叶变换(STFT)。为了计算效率,变换器使用由快速傅里叶变换(FFT)实行的离散有限长度傅里叶变换(DFT)。其它实施例使用不同的变换。

[0051] 在一个实施例中, B 个带处于其间隔单调地不减少的频率处。合理的数量(例如,90%)的频带包括来自多于一个频率格的贡献,并且在特定实施例中,每个频带包括来自两个或更多个频率格的贡献。在一些实施例中,带以对数状方式单调地增大。在一些实施例中,带为心理声学尺度(psycho-acoustic scale),即,频带利用与心理声学关键间隔(critical spacing)有关的缩放比例(scaling)被分隔,这种分带在此称为“感知分带(perceptually-banding)”。在特定实施例中,带间隔约为 1ERB 或者 0.5Bark,或者等同的带具有大约 10% 的中心频率处的频率分离。频率间隔的合理范围是从 5 ~ 20% 或者大约 0.5-2ERB。

[0052] 在其中输入处理包括噪声减少的一些实施例中,输入处理,多个输入音频信号由输入处理器接受,并且输入处理包括减少位置外信号。包括减少位置外信号的输入处理的一个示例被描述在同时提交的题为“COMBINED SUPPRESSION OF NOISE, ECHO, and OUT-OF-LOCATION SIGNALS”的国际申请 No. PCT/US2012/024370 中,该国际申请也要求 2011 年 2 月 10 日提交的发明人为 Dickins 等的题为“COMBINED SUPPRESSION OF NOISE, ECHO, AND OUT-OF-LOCATION SIGNALS”的美国临时申请 No. 61/441,611 的优先权,这个国际申请和美国临时申请的内容通过参考被并入于此。结果得到的原始的分带的增益实现同时的回声减少和噪声减少。

[0053] 在其中输入处理包括噪声减少的一些实施例中,输入处理还包括回声减少。在同时提交的国际申请 No. PCT/US_/____ 中描述了包括回声减少的输入处理的一个示例。对于其中输入处理包括回声减少的那些实施例,一个或更多个参考信号还被包括并且用来获得

回声的某个特性的估计值,例如,回声的功率(或者其它振幅度量)谱的估计值。结果得到的原始的分带的增益实现同时的回声减少和噪声减少。

[0054] 在包括噪声减少和回声减少的一些实施例中,经后处理的增益由元件 123 接受,该元件 123 修改增益以便包括额外的回声抑制。结果是一组经后处理的增益 125,其被应用于一个或更多个输入信号,例如,被用来在频域中处理输入音频信号(如果存在一个输入)、或者输入音频信号的向下混合(如果存在(例如,来自不同位置的麦克风的)多个输入音频信号),作为频率格。

[0055] 增益应用模块 131 接受分带的经后处理的增益 125 并且应用这种增益于一个或更多个输入音频信号。在一个实施例中,分带的增益被内插并且应用于输入音频信号(如果一个)或者向下混合的输入音频信号(如果存在多于一个的输入音频信号)的频率格数据(表示为 Y_n , $n=0, 1, \dots, N-1$, 其中 N 是频率格的数量)。 Y_n , $n=0, 1, \dots, N-1$ 是输入音频信号样本 Y_m , $m=1, \dots, M$ 的帧的频率格。经处理的数据 133 然后可以被输出合成滤波器组 135 转换回到样本域,以便制作 M 个信号样本 137 的帧。在一些实施例中,另外或者作为替代,信号 133 经受变换或者重映射例如到准备好根据某个编码方法编码的形式。

[0056] 下面更详细地描述与包括用于减少噪声(以及可能地回声和位置外信号)的输入处理的 PCT/US2012/024370 类似的系统的示例实施例。

[0057] 本发明当然不限于在国际申请 No. PCT/US2012/024370、U. S. 61/441,611 中描述的输入处理和增益计算,或者甚至不限于噪声减少。

[0058] 虽然在一个实施例中输入处理要减少噪声(并且可能地回声和位置外信号),在其它实施例中,输入处理可以另外或者主要要实现基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或更多个,其考虑依赖于音频信号的再现电平的音频的感知的变化,如例如在共同拥有的 W02004111994 中描述的。按照 W02004111994 计算的原始的分带的增益被后处理(包括中值滤波),以便确定用于应用于(经变换的)输入的经后处理的增益 125。

[0059] 包括中值滤波的后处理

[0060] 图 3A 示出包括根据本发明实施例的中值滤波器 305 的后处理器 121 的简化框图。后处理器 121 接受原始的增益 111 以及,在其中后处理根据信号分类改变的实施例中,接受一个或更多个信号分类指示器 115 (例如, VAD 和 WAD 中的一个或更多个)的输出。虽然未被包括在所有实施例中,但是后处理器的一些实施例包括用于确保增益不低于预定义的可能依赖于频率的值的的最小增益处理器 303。此外虽然未被包括在所有实施例中,但是后处理器的一些实施例包括平滑滤波器 307,其在中值滤波之后处理增益以便平滑频带到频带变化和/或平滑时间变化。图 3B 示出后处理 310 的方法的简化流程图,其在 311 中包括接受原始的增益,以及在其中后处理根据信号分类改变的实施例中,接受一个或更多个信号分类指示器 115。后处理包括根据本发明实施例的中值滤波 315。本发明人已经发现中值滤波是有效的非线性平滑技术,其在与仅仅使用平滑方法相比时对于消除不期望的离群值很适用。一些实施例包括在步骤 313 中确保增益不低于预定义的最小值,其可以是频带相关的。一些实施例还包括,在步骤 317 中,带到带和/或时间平滑,例如,使用例如加权移动平均(weighted moving average)的线性平滑。

[0061] 因此,在本发明的一些实施例中,原始的分带的增益值的中值滤波器 315 的特征

在于：1) 包括以确定中值的原始的分带的增益的数量，2) 被包括的原始的分带的增益的频带位置；3) 结束条件，即，用来延伸原始的分带的增益以便允许在频带和时间的边缘处计算中值的条件；以及 4) 中值滤波器的特征如何受信号分类的影响，例如，语音的存在和风的的存在中的一个或多个。

[0062] 一些实施例包括用于控制基于信号分类的频率和 / 或时间之上的中值滤波特性中的一个或多个的机制。例如，在包括语音活动检测的一个实施例中，中值滤波特性中的一个或多个根据输入是否由 VAD 确认为语音而变化。在包括风活动检测的一个实施例中，中值滤波特性中的一个或多个根据输入是否由 WAD 确认为风而变化。

[0063] 不同的边缘条件的示例包括 (a) 对于边缘的内部值的外插；(b) 使用最小增益值来延伸边缘处的原始的分带的增益，(c) 使用零增益值来延伸边缘处的原始的分带的增益 (d) 复制中心滤波器位置值以便延伸边缘处的原始的分带的增益，以及 (e) 使用最大增益值来延伸边缘处的原始的分带的增益。

[0064] 额外的后处理

[0065] 虽然未被包括在所有实施例中，但是在一些实施例中后处理器 121 包括实现步骤 313 以便确保增益不低于预定义的最小增益值的最小增益处理器 303。在一些实施例中，最小增益处理器以频带相关的方式确保最小值。在一些实施例中，防止最小值的方式依赖于活动分类 115，例如，是否语音。

[0066] 在一个实施例中，由增益 $Gain'_{b,S}$ 表示来自输入处理的所计算的原始的增益，在最小处理器之后的由 $Gain'_{b,RAW}$ 表示的增益的一些可替代值 (alternative) 为

$$[0067] \quad Gain'_{b,RAW} = Gain'_{b,MIN} + (1 - Gain'_{b,MIN}) \cdot Gain'_{b,S}$$

$$[0068] \quad Gain'_{b,RAW} = Gain'_{b,MIN} + Gain'_{b,S}$$

$$[0069] \quad Gain'_{b,RAW} = \begin{cases} Gain'_{b,MIN} & Gain'_{b,S} < Gain'_{b,MIN} \\ Gain'_{b,S} & \text{otherwise} \end{cases}$$

[0070] 作为一个示例，在后处理器 121 和步骤 310 的一些实施例中，最大抑制深度或者最小增益的范围可以从 -80dB 到 -5dB 的范围并且是频率相关的。在一个实施例中抑制深度在 200Hz 以下的低频处大约是 -20dB，在 1kHz 处变为大约 -10dB 并且在大约 4kHz 的上语音频率处放宽到仅仅 -6dB。此外，在一个实施例中，如果 VAD 确定信号是语音，则增益 $Gain'_{b,MIN}$ 增大，例如，以频带相关的方式 (或者在另一实施例中，通过对于每个带 b 相同的量)。在一个实施例中，最小值的增大的量在中频带 (例如，500Hz 到 2kHz 之间的带) 中更大。

[0071] 此外，虽然未被包括在所有实施例中，但是在一些实施例中后处理器 121 包括平滑滤波器 307，例如，线性平滑滤波器，其实现频带到频带平滑和时间平滑中的一个或者两者。在一些实施例中，这种平滑化根据信号分类 115 而变化。

[0072] 平滑化 317 的一个实施例使用具有固定核 (kernel) 的加权移动平均。一个示例使用用于加权移动平均的高斯加权核的二项近似。作为一个示例，5 点二项式平滑器具有核

$\frac{1}{16} [1 \ 4 \ 6 \ 4 \ 1]$ 实际上，当然，因子 1/16 可以省去，在需要时在一个点或另一个中实

现缩放比例 16。作为另一个示例,3 点二项式平滑器具有核 $\frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$ 。已知许多其它

加权移动平均滤波器,并且任何这种滤波器可以适当地被修改以用于增益的带到带平滑。

[0073] 在一个实施例中,通过信号分类控制带到带平滑。在一个实施例中,VAD(例如,空间选择性的 VAD)被包括,并且如果 VAD 确定存在语音,则在检测到噪声时平滑的程度增大。在一个示例实施例中,在 VAD 指示检测到语音的情况下 5 点带到带加权平均平滑被执行,否则,在 VAD 确定不存在语音时,不执行平滑。

[0074] 在一些实施例中,还包括增益的时间平滑。在一些实施例中,B 个带中的每一个的增益被如下一阶平滑滤波器平滑:

$$[0075] \quad Gain_{b,Smoothed} = \alpha_b Gain_b + (1 - \alpha_b) Gain_{b,SmoothedPrev}$$

[0076] 其中 $Gain_b$ 是当前的时间帧增益, $Gain_{b,Smoothed}$ 是经时间平滑的增益,并且 $Gain_{b,SmoothedPrev}$ 是来自前一 M 样本帧的 $Gain_{b,Smoothed}$ 。 α_b 是可以与频带相关的并且典型地在 20 到 500ms 范围内的时间常数。在一个实施例中使用 50ms 的值。在一个实施例中,

通过当前帧的信号分类来控制时间平滑的量。在包括增益的一阶时间平滑的特定的实施例中,当前帧的信号分类被用来控制用于滤波每个带中的随时间的增益的一阶时间常数的值。在包括 VAD 的情况下,一个实施例在检测到语音的情况下停止时间平滑。

[0077] 本发明人发现重要的是在语音的开始(onset)处中止积极的(aggressive)平滑。因此优选的是后处理的参数通过立即信号分类器(VAD, WAD)值来控制,该立即信号分类器(VAD, WAD)值具有低等待时间(latency)并且能够实现从噪声到语音(或者其它期望信号)模式中的后处理的快速转变。在检测到语音之后(即,在拖出(trail out)处)恢复(reinstate)更积极的后处理的速度已经被认为是次要的,因为它更少程度地影响语音的可识度。

[0078] 语音活动控制的示例

[0079] 在一个实施例中,通过信号分类控制带到带中值滤波。在一个实施例中,包括 VAD, 并且如果 VAD 确定很可能不存在语音,则实现具有 5 点带到带和 3 点时间中值滤波的 7 点 T 形中值滤波器,其具有包括在边缘处延伸最小增益值或者零值以便计算中值的边缘处理。如果 VAD 确定很可能存在语音,则在第一版本(version)中,5 点 T 形时间-频率中值滤波在当前的时间帧中利用三个频带被实现,并且使用两个先前时间帧,并且在第二实施例中,使用三点少存储仅频带中值滤波器,具有在边缘处外插以便计算中值的边界值。在一个这种组的实施例中,中值是中值,使得中值滤波器是中值滤波器。

[0080] 风活动控制的示例

[0081] 本发明的一个特征在于,后处理,例如,中值滤波依赖于信号的分类,并且在一些实施例中,一个这种分类是是否存在风。在一些实施例中,包括 WAD, 并且如果 WAD 确定不存在风,并且 VAD 指示不存在语音,则更少原始的增益值被包括在中值滤波器内。当包括 WAD 和 VAD 时,如果 WAD 确定可能不存在风并且 VAD 确定语音是可能的,则中值滤波应该更短,例如通过使用 3 点带到带中值滤波器,其具有外插在边缘处施加的边界值。如果 WAD 指示风是不太可能的,并且 VAD 指示语音也是不太可能的,则可以使用更多中值滤波,例如,实现 5

点带到带中值滤波,其具有包括在边缘处延伸最小增益值或者零值以便计算中值的边缘处理。如果 WAD 指示风是可能,并且 VAD 指示语音是不太可能的,则可以使用甚至更多的中值滤波,例如,可以实现 7 点带到带中值滤波,其具有包括在边缘处延伸最小增益值或者零值以便计算中值的边缘处理。

[0082] 示例声噪声减少系统

[0083] 声噪声减少系统典型地包括用于确定一组噪声减少增益的噪声估计器和原始增益计算模块,该组增益例如在一组频带上被确定并且在变换到频域并且分带到该组频带之后被应用于(嘈杂的)输入音频信号以便衰减噪声分量。声噪声减少系统可以包括一个麦克风,或来自不同位置的麦克风的多个输入并且向下混合(例如波束成形)以产生一个输入音频信号。声噪声减少系统可以还包括回声减少,并且可以还包括位置外信号减少。

[0084] 图 4 示出设备的一个示例,所述设备被配置为确定用于抑制噪声、以及在一些版本中同时抑制回声、以及在一些版本中同时抑制位置外信号的一组后处理的增益。

[0085] 这种系统例如被描述在国际申请 PCT/US2012/024370 中和在 US61/441,611 中。该输入包括一组一个或更多个输入音频信号 101,例如,来自不同位置的麦克风的信号,每个信号对于每个帧有多组 M 个样本。当包括空间信息时,存在两个或更多个输入音频信号,例如,来自空间分开的麦克风的信号。当包括回声抑制时,也接受一个或更多个参考信号 103,例如,在多帧的 M 样本中。这些可以是例如来自一个或更多个扬声器的一个或更多个信号,或者,在另一实施例中,被用来驱动一个或更多个扬声器的一个或更多个信号。第一输入处理级 403 确定 P'_{ω} 表示的分带的信号功率(或者其它振幅度量)谱 413,以及 Y'_{ω} 表示的瞬时功率 417 的分带的测量。当包括多于一个输入音频信号时,瞬时的分带的测量 417 和谱 413 中的每一个具有在被向下混合器(例如,波束成形器)向下混合之后的输入。当包括回声抑制时,第一输入处理级 403 也确定 E'_{ω} 表示的回声 415 的分带的功率谱估计值,该确定根据使用具有一组自适应确定的滤波器系数的滤波器的回声的先前计算的功率谱估计值。在包括位置外信号抑制的那些版本中,第一输入处理级 403 也确定具有分带的位置概率指示器 419 形式的空间特征 419,其可用于将信号空间地分离成来源于期望位置的分量和未来源于期望方向的那些分量。

[0086] 来自第一级 403 的量被用在第二级 405 中,该第二级 405 确定原始增益,并且对该原始增益进行后处理,包括本发明的实施例的中值滤波,以便确定分带的经后处理的增益 125。第二级 405 的实施例包括用于确定 E'_{ω} 表示的噪声功率(或者其它振幅度量)谱的测量的噪声功率(或者其它振幅度量)谱计算器 421、以及用于确定信号分类 115 的信号分类器 423,例如,语音活动检测器(VAD)、风活动检测器和功率通量计算器中的一个或更多个。图 4 示出包括 VAD 的信号分类器 423。

[0087] 图 5 更详细地示出图 4 的元件的一个实施例 500,并且包括用于噪声、回声和位置外的噪声抑制的示例实施例的抑制器 131 和用于产生输出信号 137 的输出合成器(或者变换器或者重映射器) 135,该抑制器 131 应用经后处理的增益 125。

[0088] 比较图 4 和图 5,图 4 的第一级处理器 403 包括图 5 的元件 503, 505, 507, 509, 511, 513, 515, 517, 521, 523, 525 和 527。更详细地,一个或更多个输入的一个或更多个帧 101 由一个或更多个输入变换器 503 变换,以便确定经变换的输入信号格,频率格的数量由 N 表示。在多于一个输入音频信号的情况下,这些频域信号由波束形成器 507 波束成形以便形

成 $Y_n, n=1, \dots, N$ 表示的输入频率格数据, 并且输入频率格数据 Y_n 由谱分带元件 509 分带成 B 个谱带, 在一个实施例中, 感知分隔的谱带, 以便产生功率 $X'_{b}, b=1, \dots, B$ 的瞬时分带的测量。在包括位置外的抑制以及多于一个输入音频信号的版本中, 来自输入变换器 503 的频域信号由分带的空间特征计算器接受以便确定分带的位置概率指示器, 每个在 0 与 1 之间。在包括回声抑制的版本中, 如果存在多于一个的参考信号, 称为 Q 个参考信号, 信号由组合器 511 (在一个实施例中, 加法器 (summer)) 组合以便产生经组合的参考输入。输入变换器 513 和谱分带器 (bander) 515 将参考转换成对于 B 个带的由 $X'_{b}, b=1, \dots, B$ 表示的分带的参考谱内容。 L 抽头线性预测滤波器 517 使用 $L \times B$ 滤波器更新系数 528 来预测分带的回声谱内容 $E'_{b}, b=1, \dots, B$ 。信号谱计算器 521 计算 (向下混合的) 功率 (或者其它振幅度量) 谱 $P'_{b}, b=1, \dots, B$ 的测量值。在一些实施例中, Y'_{b} 被用作到 P'_{b} 的足够好的近似。

[0089] 用于滤波器 517 的 $L \times B$ 滤波器系数由自适应滤波器更新器 (updater) 527 确定, 并且由噪声计算器 523 根据瞬时功率 Y'_{b} 和来自信号谱计算器 521 的测量值来确定, 该自适应滤波器更新器 527 使用当前的分带的回声谱内容 E'_{b} 、(向下混合的) 功率 (或者其它振幅度量) 谱 P'_{b} 的测量值、 $N'_{b}, b=1, \dots, B$ 表示的分带的噪声功率 (或者其它振幅度量) 谱 524。更新由 S 表示的如由语音活动检测器 (VAD) 525 使用 P'_{b} (或者 y'_{b})、 N'_{b} 和 E'_{b} 确定的语音活动信号来触发。当 S 超过阈值时, 该信号被假设为语音。回声更新语音活动检测器 525 和滤波器更新器 527 中推导的 VAD 用于控制回声预测的修改的特定目的。具有该目的的 VAD 或者检测器通常被称为双说话检测器。在一个实施例中, 更新器 527 的回声滤波器系数更新被门控, 在与预期的噪声和当前输入功率相比预期的回声显著时出现更新, 如由 VAD 525 确定并且由本地信号活动 S 的低值指示。

[0090] 根据图 4 和图 5 的第一级 403 的元件在一些实施例中如何操作的细节如下。在一个实施例中, 输入变换器 503、511 确定短时间傅里叶变换 (STFT)。在另一实施例中, 以下变换和逆对被用于元件 503 和 511 中以及输出合成元件 135 中的前向变换。

$$[0091] \quad X_{2n} = \frac{1}{\sqrt{N}} \sum_{n'=0}^{N-1} e^{\frac{-i\pi n'}{2N}} (u_{n'} x_{n'} - i u_{N+n'} x_{N+n'}) e^{\frac{-i2\pi n n'}{N}} \quad n=0 \dots N/2-1$$

$$[0092] \quad X_{2n+1} = \frac{1}{\sqrt{N}} \sum_{n'=0}^{N-1} e^{\frac{i\pi n'}{2N}} (u_{n'} x_{n'} + i u_{N+n'} x_{N+n'}) e^{\frac{-i2\pi n n'}{N}} \quad n=0 \dots N/2-1$$

[0093]

$$y_n = v_n \operatorname{real} \left[\frac{1}{\sqrt{N}} e^{\frac{i\pi n}{4N}} \left(\sum_{n'=0}^{N/2-1} X_{n'} e^{\frac{i4\pi n n'}{N}} + \sum_{n'=N/2}^{N-1} X_{N-n'-1} e^{\frac{i4\pi n n'}{N}} \right) \right] \quad n=0 \dots N-1$$

[0094]

$$y_{N+n} = -v_{N+n} \operatorname{imag} \left[\frac{1}{\sqrt{N}} e^{\frac{i\pi n}{4N}} \left(\sum_{n'=0}^{N/2-1} X_{n'} e^{\frac{i4\pi n n'}{N}} + \sum_{n'=N/2}^{N-1} X_{N-n'-1} e^{\frac{i4\pi n n'}{N}} \right) \right] \quad n=0 \dots N-1$$

[0095] 其中 $i^2=-1$, u_n 和 v_n 是适当的窗函数, x_n 表示最后 $2N$ 个输入样本, 具有 x_{N-1} 表示最近的样本, X_n 以增大频率顺序表示 N 个复数值的频率格。逆变换或者合成以最后二个等式方

式表示。 y_n 表示由用于设计的窗口的在适当时在交迭、相加和丢弃之前的独立逆变换得到的 $2N$ 个输出样本。应当注意,这个变换具有有效的实现方式作为块乘(block multiply)和 FFT。注意,上述变换的表达式中的 x_n 和 X_n 的使用是为了方便起见。在这个公开内容的其它部分中, $X_n, n=0, \dots, N-1$ 表示代表参考信号的特征的信号的频率格,并且 $Y_n, n=0, \dots, N-1$ 表示向下混合的输入音频信号的频率格。

[0096] 在一个实施例中,在一个实施例中用于上述变换的窗函数 u_n 和 v_n 是正弦窗家族,其一个建议的实施例是

$$[0097] \quad u_n = v_n = \sin \left(\frac{n + \frac{1}{2}}{2N} \pi \right) \quad n = 0 \dots 2N - 1$$

[0098] 本领域技术人员应该清楚的是,分析和合成窗口(也被称为原型滤波器)可以具有大于或小于在此给出的示例的长度。

[0099] 虽然本发明在具有任何向下混合的信号的情况下工作,但是在一些实施例中,向下混合器是设计成实现朝向期望位置的某个空间选择性的波束形成器 507。在一个实施例中,波束形成器 507 是线性时间不变处理,即,一般由用于每个输入声道的一组复数值的依赖于频率的增益定义的被动式的(passive)波束形成器。对于双麦克风阵列的示例,期望声源位于阵列的广侧(broad side),即,在中垂线处,一个实施例将确定两个输入声道的简单的和的被动式的波束形成器 107 用于波束形成器 507。在一些版本中,波束形成器 507 通过一组复数值的权重对该组输入(作为频率格)加权。在一个实施例中,根据最大比率组合(MRC)确定波束形成器 107 的波束形成权重。在另一实施例中,波束形成器 507 使用通过使用迫零(zero-forcing)确定的权重。这种方法在本领域中是公知的。

[0100] 谱分带元件 509 和 514 的分带可以通过如下描述

$$[0101] \quad Y'_b = W_b \sum_{n=0}^{N-1} w_{b,n} |Y_n|^2$$

[0102] 其中 Y'_b 是向下混合的、例如波束成形的信号的分带的瞬时功率, W_b 是规格化(normalization)增益并且 $w_{b,n}$ 是来自分带矩阵的元素。

[0103] 在一个实施例中的信号谱计算器 521 通过如下平滑处理描述

$$[0104] \quad P'_b = \alpha_{P,b} (Y'_b + Y'_{\min}) + (1 - \alpha_{P,b}) P'_{b\text{PREV}},$$

[0105] 其中 $P'_{b\text{PREV}}$ 是先前确定的、例如最近确定的信号功率(或其它频域振幅度量)估计值, $\alpha_{P,b}$ 是时间信号估计值时间常数,并且 Y'_{\min} 是偏移(offset)。发现对于信号估计值时间常数 $\alpha_{P,b}$ 的合适的范围在 20 到 200ms 之间。在一个实施例中,偏移 Y'_{\min} 被添加以避免零电平功率谱(或其它振幅度量谱)估计值。 Y'_{\min} 可以被测量,或可以基于先验知识被选择。 Y'_{\min} 例如可以与听觉极限或器件噪声阈值有关。

[0106] 在一个实施例中,自适应滤波器 517 包括通过使用由如下描述的 L 抽头自适应滤

波器来确定对于带 b 由 T'_b 表示的瞬时的回声功率谱(或其它振幅度量谱)

$$[0107] \quad T'_b = \sum_{l=0}^{L-1} F_{b,l} X'_{b,l},$$

[0108] 其中当前帧为 $X'_b = X'_{b,0}$ 其中 $X'_{b,0}, \dots, X'_{b,1}, \dots, X'_{b,L-1}$ 是 L 个最近的帧的(组合的)分带的参考信号 X'_b , 包括当前帧 $X'_b = X'_{b,0}$, 其中用于给定带 b 的 L 个滤波器系数分别由 $F_{b,0}, \dots, F_{b,1}, \dots, F_{b,L-1}$ 表示。

[0109] 一个实施例包括来自回声预测滤波器 517 的瞬时的回声的时间平滑以便确定回声谱估计值 E'_b 。在一个实施例中,一阶时间平滑滤波器被如下使用

$$[0110] \quad E'_b = T'_b \text{ for } T'_b \geq E'_{b\text{Prev}}, \text{ 以及}$$

$$[0111] \quad E'_b = \alpha_{E,b} T'_b + (1 - \alpha_{E,b}) E'_{b\text{Prev}}, \text{ 对于 } T'_b < E'_{b\text{Prev}}$$

[0112] 其中 $E'_{b\text{Prev}}$ 是先前确定的回声谱估计值,例如,最近确定的回声谱估计值,或者其它先前确定的估计值,并且 $\alpha_{E,b}$ 是一阶平滑时间常数。

[0113] 在一个实施例中,噪声功率谱计算器 523 使用具有指数生长的最小值跟随器:

$$[0114] \quad N'_b = \min(P'_b, (1 + \alpha_{N,b}) N'_{b\text{Prev}}) \text{ 在 } E'_b \text{ 小于 } N'_{b\text{Prev}} \text{ 时}$$

$$[0115] \quad N'_b = N'_{b\text{Prev}}, \text{ 其它情况,}$$

[0116] 其中 $\alpha_{N,b}$ 是指定最小值跟随器可以增加以跟踪噪声的任何增加的随时间的速率的参数。在一个实施例中,准则 E'_b 小于 $N'_{b\text{Prev}}$ 是在 $E'_b < N'_{b\text{Prev}}/2$ 的情况下,即在(经平滑的)回声谱估计值 E'_b 小于 N'_b 的先前值小 3dB 的情况下,在该情况下噪声估计值跟随生长或当前功率。否则, $N'_b = N'_{b\text{Prev}}$, 即, N'_b 被保持在 N'_b 的先前值。参数 $\alpha_{N,b}$ 最好用最小值跟随器将跟踪的随时间速率表示。那个速率可以用 dB/sec 表示,其然后提供用于确定 b 的值的机制。范围是 1 到 30dB/sec。在一个实施例中使用 20dB/sec 的值。

[0117] 在其它实施例中,可以使用用于噪声估计的不同的近似。这种不同的近似的示例包括但不限于,确定信号观察的窗口之上的最小值的迭代法,例如 1 和 10 秒的窗口。另外或者可替代最小值,这种不同的近似可能还确定在它被分类为可能是噪声或者语音不太可能的时间期间的信号的变化和意思。

[0118] 在一个实施例中,最小值跟随器的一个或多个泄漏率参数由如由语音活动检测(VAD)确定存在的语音的概率控制。在一个实施例中,VAD 元件 525 确定 S 表示的整个信号活动电平为

$$[0119] \quad S = \sum_{b=1}^B \frac{\max(0, Y'_b - \beta_N N'_b - \beta_E E'_b)}{Y'_b + Y'_{\text{sens}}}$$

[0120] 其中 $\beta_N, \beta_E > 1$ 是分别用于噪声和回声的裕度并且 Y'_{sens} 是可设置的灵敏度偏移。这些参数一般可以在各带之间变化。在一个实施例中, β_N, β_E 的值在 1 与 4 之间。在特定的实施例中, β_N, β_E 各是 2。 Y'_{sens} 被设置为在通过对典型的组件进行实验获得的预期麦克风和系统噪声电平附近。可替代地, 可以使用听觉极限来确定用于 Y_{sens} 的值。

[0121] 在一个实施例中, 更新器 527 的回声滤波器系数更新被门控, 如下。如果本地信号活动水平低, 例如, 在预定义的阈值 S_{thresh} 以下, 即, 如果 $S < S_{thresh}$, 则自适应滤波器系数被更新为:

$$[0122] \quad F_{b,l} = F_{b,l} + \mu \frac{(\max(0, Y'_b - \gamma_N N'_b) - T'_b) X'_{b,l}}{\sum_{l'=0}^{L-1} (X'_{b,l'}^2 + X'_{sens}{}^2)} \quad \text{如果 } S < S_{thresh},$$

[0123] 其中 γ_N 是调整以便确保噪声与回声估计值之间的稳定性的调整参数。 γ_N 的代表值是 1.4 (+3dB)。可以使用 1 到 4 的数值范围。 μ 是影响回声估计值的稳定性与收敛速度的调整参数。 0 与 1 之间的值可能在不同的实施例中有用。在一个实施例中, $\mu = 0.1$ 与帧尺寸 M 无关。 X'_{sens} 被设定为避免对于小参考信号的不稳定的修改。在一个实施例中 X'_{sens} 与听觉极限有关。 S_{thresh} 的值的选择不依赖于带的数量。 S_{thresh} 在 1 与 B 之间, 并且对于具有 24 个带到 8kHz 的一个实施例, 发现合适的范围在 2 与 8 之间, 特定的实施例使用 4 的值。

[0124] 本发明的实施例使用具有根据带 b 中的一个或多个空间特征确定的一个或多个测量值形式的空间信息, 其与特定的带 b 具有从所关心的空间区域入射的这种能量的概率单调一致 (monotonic with)。这种量被称为空间概率指示器。在一个实施例中, 一个或多个空间概率指示器是输入音频信号的一个或多个分带的加权的协方差矩阵的函数。假设 P 输入变换的输出 $X_{p,n}, p=1, \dots, P$, 具有 N 个频率格, $n=0, \dots, N-1$, 我们构造一组加权协方差矩阵以便通过对于格 n 的跨越 P 个输入的输入向量与其共轭转置的积求和, 并且通过具有元素 $w_{b,n}$ 的分带矩阵 W_b 加权来对应

$$[0125] \quad \mathbf{R}'_b = \sum_{n=0}^{N-1} w_{b,n} [X_{1,n} \quad \dots \quad X_{P,n}]^H [X_{1,n} \quad \dots \quad X_{P,n}].$$

[0126] $W_{b,n}$ 提供每个格如何被加权以用于贡献于带的指示。在一些实施例中, 一个或多个协方差矩阵被随时间平滑。在一些实施例中, 分带矩阵包括用于加权移动平均的时间依赖的加权, 表示为具有元素 $w'_{b,n,l}$ 的 $W_{b,l}$, 其中 l 表示时间帧, 使得, 在 L 个时间帧之上,

$$[0127] \quad \mathbf{R}'_b = \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} w_{b,n,l} [X_{1,n} \quad \dots \quad X_{P,n}]^H [X_{1,n} \quad \dots \quad X_{P,n}].$$

[0128] 在两个输入的情况下, $P=2$, 定义

$$[0129] \quad \mathbf{R}'_b = \begin{bmatrix} R'_{b11} & R'_{b12} \\ R'_{b21} & R'_{b22} \end{bmatrix},$$

[0130] 使得每个带协方差矩阵 R_b 是 2×2 Hermetian 正定矩阵, $R'_{b21} = \overline{R'_{b12}}$, 其中上划线被用来指示复共轭。

[0131] 通过空间特征“比率”表示与分带的幅度 $\frac{R'_{b11}}{R'_{b22}}$ 的比率单调一致的量。在一个实施例中, 使用对数关系:

$$[0132] \quad \text{Ratio}'_b = 10 \log_{10} \frac{R'_{b11} + \sigma}{R'_{b22} + \sigma}$$

[0133] 其中 σ 是添加以避免奇异(singularities)的小偏移。 σ 可以被认为是对于 R'_{b11} 的最小的预期值。在一个实施例中, 它是用于麦克风和有关的电子设备的带 b 内的噪声功率(或者其它频域振幅度量)的确定或者估计(先验(priori))值。也就是说, 使用任何预处理的最小灵敏度。

[0134] 由空间特征相位表示与 $\tan^{-1} R'_{b21}$ 单调一致的量。

[0135] $\text{Phase}'_b = \tan^{-1} R'_{b21}$

[0136] 由空间特征“相干(coherence)”表示与 $\frac{R'_{b21} R'_{b12}}{R'_{b11} R'_{b22}}$ 单调一致的量。在一些实施

例中, 相干的有关的测量值可以被使用, 诸如 $\frac{2R'_{b21} R'_{b12}}{R'_{b11} R'_{b11} + R'_{b22} R'_{b22}}$ 或与调节、排序或协方差矩阵的本征值扩展有关的值。在一个实施例中, 相干特征为

$$[0137] \quad \text{Coherence}'_b = \sqrt{\frac{R'_{b21} R'_{b12} + \sigma^2}{R'_{b11} R'_{b22} + \sigma^2}}$$

[0138] 如上面定义的具有偏移 σ 。

[0139] 噪声、回声和位置外信号抑制的一些实施例的一个特征在于, 基于期望信号特征的先验预期或者当前估计值(例如, 表示空间位置、从统计的数据收集的目标值), 每个带中的每个空间特征可以被用来产生用于对于带 b 的特征的概率指示器。

[0140] 在一个实施例中, 对于期望位置的预期空间特征的分布被建模为高斯分布, 其呈现捕获由每个空间特征和带导出的对于概率指示器的感兴趣区的稳健的方式。

[0141] 三个空间概率指示器与这三个空间特征有关, 并且是 RPI'_b 表示的比率概率指示器、 PPI'_b 表示的相位概率指示器以及 CPI'_b 表示的相干概率指示器, 其中

$$[0142] \quad RPI'_b = f_{R_b} \left(\text{Ratio}'_b - \text{Ratio}_{\text{target}_b} \right) = f_{R_b} \left(\Delta \text{Ratio}'_b \right),$$

[0143] 其中 $\Delta \text{Ratio}'_b = \text{Ratio}'_b - \text{Ratio}_{\text{target}_b}$, 并且 $\text{Ratio}_{\text{target}_b}$ 根据在前估计值或者对所使用的设备(例如耳机)进行的例如根据诸如图 9A 所示的数据获得的实验来确定。

[0144] 函数 $f_{R_b}(\Delta Ratio')$ 是平滑函数。在一个实施例中,比率概率指示器函数为

$$[0145] \quad f_{R_b}(\Delta Ratio') = \exp\left[-\frac{\Delta Ratio'_b}{Width_{Ratio,b}}\right]^2,$$

[0146] 其中 $Width_{Ratio,b}$ 是以对数为单位(例如 dB)表示的宽度调整参数。

$Width_{Ratio,b}$ 相关于但是不要求根据实际数据来确定。它被设置为覆盖正常和嘈杂情形中的空间特征的预期变化,但是也仅仅需要与在整体系统的背景中实现期望抑制所需的那样窄。

[0147] 对于相位概率指示器,

$$[0148] \quad PPI'_b = f_{P_b}(Phase'_b - Phase_{target_b}) = f_{R_b}(\Delta Phase'_b),$$

[0149] 其中 $\Delta Phase'_b = Phase'_b - Phase_{target_b}$ 并且 $Phase_{target_b}$ 根据在前估计值或者对所使用的设备(例如耳机)进行的例如根据数据获得的实验来确定。

[0150] 函数 $f_{P_b}(\Delta Phase')$ 是平滑函数。在一个实施例中,

$$[0151] \quad f_{R_b}(\Delta Phase'_b) = \exp\left[-\frac{\Delta Phase'_b}{Width_{Phase,b}}\right]^2$$

[0152] 其中 $Width_{Phase,b}$ 是以相位单位表示的宽度调整参数。在一个实施例中, $Width_{Phase,b}$ 相关于但是不要求根据实际数据来确定。

[0153] 对于相干概率指示器,不使用目标(target),并且在一个实施例中,

$$[0154] \quad CPI'_b = \left(\frac{R'_{b21}R'_{b12} + \sigma^2}{R'_{b11}R'_{b22} + \sigma^2} \right)^{CFactor_b}$$

[0155] 其中 $CFactor_b$ 是可以为 0.1 ~ 10 范围内的恒定值的调整参数;在一个实施例中,发现 0.25 的值是有效的。

[0156] 图 6 示出原始增益的元件 529 中的计算的一个示例,并且包括空间灵敏的语音活动检测器(VAD) 621 以及风活动检测器(WAD) 623。噪声减少的可替代的版本可以不包括 WAD,或者空间灵敏的 VAD,并且还可以不包括回声抑制或者其它减少。此外,图 6 中示出的实施例包括额外的回声抑制,其可以不被包括在更简单的版本内。

[0157] 在一个实施例中,空间概率指示器被用来确定被称为波束(beam)增益的由 $BeamGain'_b$ 表示的统计量,其可以被用来根据总功率例如使用波束外谱计算器 603 估计波束内和波束外功率,并且此外,可以被用来由空间抑制增益计算器 611 确定波束外抑制增益。按照惯例并且在此呈现的实施例中,概率指示器被缩放(scale)使得波束增益具有最大值 1。

[0158] 在一个实施例中,波束增益是

$$[0159] \text{BeamGain}'_b = \text{BeamGain}_{\min} + (1 - \text{BeamGain}_{\min}) \text{RPI}'_b \cdot \text{PPI}'_b \cdot \text{CPI}'_b.$$

[0160] 一些实施例使用 0.01 到 0.3 (-40dB 到 -10dB) 的 BeamGain_{\min} 。一个实施例使用 0.1 的 BeamGain_{\min} 。

[0161] 波束内和波束外功率为:

$$[0162] \text{Power}'_{b, \text{InBeam}} = \text{BeamGain}'_b Y'_b$$

$$[0163] \text{Power}'_{b, \text{OutOfBeam}} = (1 - \text{BeamGain}'_b) Y'_b.$$

[0164] 注意 $\text{Power}'_{b, \text{InBeam}}$ 和 $\text{Power}'_{b, \text{OutOfBeam}}$ 是用于抑制的统计测量值。

[0165] 在元件 603 的一个版本中,

$$[0166] \text{Power}'_{b, \text{OutOfBeam}} = [0.1 + 0.9(1 - \text{BeamGain}'_b)] Y'_b.$$

[0167] 增益计算的一个版本使用空间选择性的噪声功率谱计算器 605, 其确定噪声功率 (或者振幅的其它度量) 谱的估计值。本发明的一个实施例使用泄漏的 (leaky) 最小值跟随器, 其具有通过至少一个泄漏率 (leak rate) 参数确定的跟踪速度。泄漏率参数不必与在回声系数更新中使用的用于非空间选择性的噪声估计值的相同。由 $N'_{b,s}$ 表示空间选择性的噪声谱估计值。在一个实施例中,

$$[0168] N'_{b,s} = \min(\text{Power}'_{b, \text{OutOfBeam}}, (1 + \alpha_b) N'_{b, \text{SPrev}}),$$

[0169] 其中 $N'_{b, \text{SPrev}}$ 已经被确定, 即, $N'_{b,s}$ 的先前值。泄漏率参数 α_b 以 dB/s 表示, 使得对于以 T 表示的帧时间, $(1 + \alpha_b) 1/T$ 在语音概率低时在 1.2 与 4 之间, 而在语音概率高时为 1。 α_b 的标称值为 3dB/s 使得 $(1 + \alpha_b) 1/T = 1.4$ 。

[0170] 在一些实施例中, 为了避免将偏置加到噪声估计值, 使用回声门控 (gating), 即,

[0171]

$$N'_{b,s} = \min(\text{Power}'_{b, \text{OutOfBeam}}, (1 + \alpha_b) N'_{b, \text{SPrev}}) \text{ if } N'_{b, \text{SPrev}} > 2E'_b, \text{ else}$$

$$[0172] N'_{b,s} = N'_{b, \text{SPrev}}.$$

[0173] 也就是说, 噪声估计值仅仅在先前噪声估计值表明噪声水平例如大于当前回声预测的两倍的情况下才被更新。否则回声会使噪声估计值偏离。

[0174] 图 4、5 和 6 中示出的降噪器的一个特征包括同时抑制: 1) 基于空间选择性的噪声估计值的噪声, 和 2) 波束外信号。增益计算器 529 包括元件 613, 其用于基于噪声功率 (或者其它频域振幅度量) 谱的空间选择性的估计值、并且进一步基于特定带中的瞬时的分带的输入功率 Y'_b , 计算被表示为用于中间信号 (例如, 频率格 Y_n) 的增益的概率指示器 614。为简单起见这个概率指示器被称为增益, 由 Gain_N 表示。然而应当注意这个增益 Gain_N 不直接应用, 而是在增益组合器 615 中与额外增益 (即, 额外概率指示器) 相组合, 以便实现用于应用于实现单个抑制的单个增益。

[0175] 元件 613 被示出具有回声抑制, 而在一些版本中不包括回声抑制。

[0176] 通过如下给出被认为在计算复杂性和效果方面是有效的表达式

$$[0177] \quad Gain'_N = \left(\frac{\max(0, Y'_b - \beta'_N N_{b,S})}{Y'_b} \right)^{GainExp}$$

[0178] 其中 Y'_b 是瞬时的分带的功率 (或其它频域振幅度量), $N'_{b,S}$ 是分带的空间选择性的 (波束外的) 噪声估计值, 并且 β'_N 是缩放参数, 典型地在 1 到 4 的范围。在一个版本中, $\beta'_N=1.5$ 。参数 GainExp 是积极性 (aggressiveness) 的控制或者抑制增益从抑制到传输的转变速率。这个指数一般采用 0.25 到 4 的范围内的值。在一个版本中, GainExp=2。

[0179] 添加回声抑制

[0180] 用于噪声减少的输入处理的一些实施例不仅包括噪声抑制, 而且包括同时抑制回声。在增益计算器 529 的一些实施例中, 元件 613 包括回声抑制并且在增益计算器 529 中, 用于抑制回声的概率指示器被表示为增益 $Gain'_{b,N+E}$ 表示的增益。在还包括回声抑制的情况下上述噪声抑制增益表达式变为

$$[0181] \quad Gain'_{b,N+E} = \left(\frac{\max(0, Y'_b - \beta'_N N'_{b,S} - \beta'_E E'_b)}{Y'_b} \right)^{GainExp_b} \quad (\text{"Gain 1"})$$

[0182] 其中 Y'_b 也是瞬时的分带的功率, $N'_{b,S}, E'_b$ 是分带的空间选择性的噪声以及分带的回声估计值, 并且 β'_N, β'_E 是在 1 到 4 范围内的缩放比例参数, 用于允许噪声和回声估计值中的误差以及相应地偏移增益曲线。同样, 它们的目的和幅度与 VAD 功能中使用的常数类似, 但是它们不一定是相同的值。在一个实施例中合适的调整值是 $\beta'_N=1.5$, $\beta'_E=1.4$, GainExp_b 对于 b 的所有值为 2。

[0183] 对于在此描述的 $Gain'_{N+E}$ 的几个表达式在分子和分母两者中具有瞬时的分带的输入功率 (或者其它频域振幅度量) Y'_b 。这在如在此描述的适当地设计分带而具有对数状或者感知分隔的频带时适用。在本发明的可替代实施例中, 分母使用估计的分带的功率谱 (或者其它振幅度量谱) P'_b , 使得对于 $Gain'_{b,N+E}$ 的上述表达式变为:

$$[0184] \quad Gain'_{b,N+E} = \left(\frac{\max(0, Y'_b - \beta'_N N'_{b,S} - \beta'_E E'_b)}{P'_b} \right)^{GainExp} \quad (\text{"Gain 1}_{MOD})$$

[0185] 额外的回声抑制的单独控制

[0186] 上面的抑制增益表达式可以被概括为瞬时的输入功率与预期不期望的信号功率 (为简单起见有时称为“噪声”) 的比率的域上的函数。在这些增益表达式中, 不期望的信号功率是估计的 (位置敏感的) 噪声功率以及预测或估计的回声功率的和。以这种方式将噪声和回声组合在一起提供了单个概率指示器, 其具有使得同时的衰减不期望的噪声和不期望的回声两者的抑制增益的形式。

[0187] 在一些情况下, 例如, 在回声可以实现基本上高于噪声的电平的电平的情况下, 这种抑制可能不引起足够的回声衰减。例如, 在一些应用中, 可以存在对仅仅温和的减少环境噪声的需要, 虽然通常要求任何回声抑制在能听度以下。为了实现这种预期效果, 在一个实施例中, 概率指示器或增益的额外的缩放比例被使用, 这种额外的缩放比例仅仅基于输入

音频信号与回声功率的比率。

[0188] 由 $f_A(\cdot)$, $f_B(\cdot)$ 表示一对抑制增益函数, 每个函数对于抑制增益具有期望特性, 例如, 如上所述, 包括例如为平滑。作为一个示例, $f_A(\cdot)$, $f_B(\cdot)$ 中的每一个具有 S 形的

函数特性。在一些实施例中, 并非被定义为 $f_A\left(\frac{Y'_b}{N'_{b,S} + E'_b}\right)$ 的增益表达, 而替代地可

以使用一对概率指示器, 例如, 增益 $f_A\left(\frac{Y'_b}{N'_{b,S}}\right)$, $f_B\left(\frac{Y'_b}{E'_b}\right)$ 以及根据 $f_A\left(\frac{Y'_b}{N'_{b,S}}\right)$ 和

$f_B\left(\frac{Y'_b}{E'_b}\right)$ 来确定组合的增益因子, 其允许独立控制响应于噪声和回声信号功率的深度和

积极性。在又一个实施例中, $f_A\left(\frac{Y'_b}{N'_{b,S} + E'_b}\right)$ 可以被应用于噪声和回声抑制两者, 并且

$f_B\left(\frac{Y'_b}{E'_b}\right)$ 可以被应用于额外的回声抑制。

[0189] 在一个实施例中, 两个函数 $f_A\left(\frac{Y'_b}{N'_{b,S}}\right)$, $f_B\left(\frac{Y'_b}{E'_b}\right)$, 或者在另一个实施例

中, 两个函数 $f_A\left(\frac{Y'_b}{N'_{b,S} + E'_b}\right)$, $f_B\left(\frac{Y'_b}{E'_b}\right)$ 被组合为乘积以便实现组合的概率指示器,

作为抑制增益。

[0190] 组合用于同时抑制位置外信号的抑制增益

[0191] 在一个实施例中, 用于波束内信号的表示为波束增益 612 的称为空间抑制增益的并且 $\text{Gain}'_{b,S}$ 表示的抑制概率指示器由元件 529 (图 5) 中的空间抑制增益计算器 611 确定为

[0192] $\text{Gain}'_{b,S} = \text{BeamGain}'_b = \text{BeamGain}_{\min}(1 - \text{BeamGain}_{\min}) \text{RPI}'_b \cdot \text{CPI}'_b$

[0193] 空间抑制增益 612 在增益组合器 615 中与其它抑制增益组合以便形成表示为抑制增益的总体概率指示器。用于同时抑制噪声、回声和波束外信号的总体概率指示器(表示为增益 $\text{Gain}'_{b,RAW}$) 在一个实施例中为增益的乘积:

[0194] $\text{Gain}'_{b,RAW} = \text{Gain}'_{b,S} \cdot \text{Gain}'_{b,N+E}$

[0195] 在可替代实施例中, 应用额外的平滑。在增益元件 615 的一个示例实施例中:

[0196] $\text{Gain}'_{b,RAW} = 0.1 + 0.9 \text{Gain}'_{b,S} \cdot \text{Gain}'_{b,N+E}$

[0197] 其中最小增益 0.1 和 0.9=(1-0.1) 因子可以根据不同实施例而变化, 以便实现对

于增益的不同最小值,具有 0.001 ~ 0.3(-60dB ~ -10dB) 的建议范围。

[0198] 对于 $Gain'_{b,RAW}$ 的上述表达式相等地抑制噪声和回声。如上所述,可以期望的是,不完全消除噪声,而是完全消除回声。在增益确定的一个这种实施例中,

[0199]

$$Gain'_{b,RAW} = 0.1 + 0.9Gain'_{b,S} \cdot f_A\left(\frac{Y'_b}{N'_{b,S} + E'_b}\right) \cdot f_B\left(\frac{Y'_b}{E'_b}\right),$$

[0200] 其中 $f_A\left(\frac{Y'_b}{N'_{b,S} + E'_b}\right)$ 实现对噪声和回声二者的(相对)适度的抑制,而

$f_B\left(\frac{Y'_b}{E'_b}\right)$ 更多抑制回声。在不同实施例中, $f_A(\cdot)$ 仅抑制噪声,而 $f_B(\cdot)$ 抑制回声。

[0201] 在又一个实施例中,

[0202] $Gain'_{b,RAW} = 0.1 + 0.9Gain'_{b,S} \cdot Gain'_{b,N+E}$,

[0203] 其中

$$Gain'_{b,E+B} = \left(0.1 + 0.9f_A\left(\frac{Y'_b}{N'_{b,S} + E'_b}\right)\right) \cdot \left(0.1 + 0.9f_B\left(\frac{Y'_b}{E'_b}\right)\right).$$

[0205] 在一些实施例中,这个噪声和回声抑制增益与空间特征概率指示器或用于形成原始的组合的增益的增益组合,并且随后通过后处理器 625 和通过后处理步骤被后处理以确保稳定性和其它期望性质。

[0206] 在另一实施例中,专用于回声抑制的增益函数 $f_B\left(\frac{Y'_b}{E'_b}\right)$ 被应用作为通过后处理

器 625 进行后处理之后的增益。增益计算器 529 的一些实施例包括额外的回声抑制增益的确定器以及额外的回声抑制增益与经后处理的增益的组合器 627, 以便导致应用整个 B 增益。本发明人发现这种实施例可以提供更特定的和更深的回声的衰减,这是因为回声概率

指示器或增益 $f_B\left(\frac{Y'_b}{E'_b}\right)$ 不经历后处理施加的平滑和连续性。

[0207] 图 7 示出操作处理设备 100 以便抑制噪声和位置外信号以及在一些实施例中回声的方法 700 的流程图,具有数量 $P \geq 1$ 的信号输入 101,例如,来自不同位置的麦克风。在包括回声抑制的实施例中,方法 700 包括处理 $Q \geq 1$ 参考输入 102,例如,要被呈现在 Q 个扬声器上的 Q 个输入,或者根据 Q 个扬声器获得的信号。

[0208] 在一个实施例中,方法 700 包括:在处理设备中接受 701 多个采样的输入音频信号 101,以及形成 703、707、709 用于多个频带的输入音频信号 101 的向下混合的分带的瞬时频

域振幅度量 417, 该形成包括变换 703 成对于一组频率格的复数值的频域值。在一个实施例中, 该形成包括在 703 中将输入音频信号变换为频率格, 向下混合(例如, 波束形成) 707 频率数据, 以及在 709 中分带。在 711 中, 该方法包括计算信号的功率(或者其它振幅度量) 谱。在可替代实施例中, 向下混合可以在变换之前, 使得单个向下混合信号被变换。在可替代实施例中, 系统可以利用分带的回声参考的估计值, 或者实现的系统内的另一个处理组件或者源提供的回声参考的频域谱的相似表示。

[0209] 该方法包括在 705 中确定分带的空间特征, 例如, 来自多个采样的输入音频信号的位置概率指示器 419。

[0210] 在包括同时的回声抑制的实施例中, 该方法包括接受 713 一个或更多个参考信号以及在 715 和 717 中形成一个或更多个参考信号的分带的频域振幅度量表示。在一个实施例中的该表示为和。此外在包括回声抑制的实施例中, 该方法包括在 721 中使用自适应确定的回声滤波器系数来预测回声 415 的分带的频域振幅度量表示。在一个实施例中的预测还包括语音活动检测—VAD—使用向下混合的信号 413 的分带的谱振幅度量、噪声的分带的谱振幅度量的估计值以及先前预测的回声谱内容 415。根据语音活动检测的结果更新或者不更新系数。更新使用噪声的分带的谱振幅度量的估计值、先前预测的回声谱内容 415、以及向下混合的信号 413 的分带的谱振幅度量的估计值。向下混合的信号的的分带的谱振幅度量的估计值在一个实施例中为输入音频信号的向下混合的分带的瞬时频域振幅度量 417, 但是在其它实施例中, 使用信号谱估计。

[0211] 在一些实施例中, 方法 700 包括:a) 在 723 中计算原始的抑制增益, 其包括使用两个或更多个空间特征 419 确定的位置外信号增益、以及使用空间选择性的噪声谱内容确定的噪声抑制增益; 以及 b) 对于每个带组合原始的抑制抑制增益与第一组合的增益。在一些实施例中噪声抑制增益包括回声的抑制、以及其计算 723 还使用预测的回声谱内容 415。

[0212] 在一些实施例中, 方法 700 还包括在 725 中实现使用两个或更多个空间特征 419 确定的空间选择性的语音活动检测, 以便产生信号分类, 例如是否语音。在一些实施例中, 使用风检测使得信号分类还包括信号是否为风。

[0213] 方法 700 还包括对带的第一组合增益实现后处理以便产生对于每个带的经后处理的增益 125。在一些实施例中, 后处理包括例如, 以带相关的方式确保最小增益。本发明的实施例的一个特征在于, 后处理包括实现组合增益的中值滤波, 例如, 以便确保不存在离群值增益。后处理的一些实施例包括通过实现时间和 / 或带到带平滑来确保平滑。

[0214] 在一些实施例中, 后处理 725 根据信号分类, 例如, 是否语音, 或者是否风, 以及在一些实施例中, 中值滤波的特性根据信号分类而变化, 例如, 是否语音或者是否风。

[0215] 在其中包括回声抑制的一个实施例中, 方法包括在 726 中计算额外的回声抑制增益。在一个实施例中, 额外的回声抑制增益被包括在被用作对于每个带的最终增益的第一组合增益内, 以及在另一实施例中, 额外的回声抑制增益与对第一组合增益应用后处理的结果组合以便产生对于每个带的最终增益。

[0216] 方法包括在 727 中应用最终增益, 包括对于格数据内插增益以便实现对向下混合的信号的格数据的抑制以便形成经抑制的信号数据 133, 并且在 729 中应用 a) 输出合成和变换以产生输出样本以及 b) 输出重映射以产生输出频率格中的一个或者两者。

[0217] 典型地, $P \geq 2$ 并且 $Q \geq 1$ 。然而, 在本申请中公开的方法、系统和设备可以向下缩

小(scale down),以便对于 $P=1, Q \geq 1$ 以及 $P \geq 2, Q=0$ 的更简单的情况仍然有效。在本申请中公开的方法和设备甚至对于 $P=1, Q=0$ 也相当地适用。虽然这最后的示例是本发明的减少和或许琐碎的实施例,但是请注意提出的构架缩放的能力是有利的,并且此外实际上可能要求更低信号操作情况,如果输入音频信号或者参考信号中的一个或多个例如由于传感器或者麦克风的失效而变为毁坏或者不可用的话。

[0218] 虽然本公开呈现完整的噪声降低方法(图 7)、包括抑制的所有方面(包括同时的回声、噪声和空间位置外的抑制)或者呈现为包括在由处理系统(参见下面描述的图 8)的一个或多个处理器运行时使得包括处理系统的处理设备实现诸如图 7 的方法的指令的计算机可读的存储介质的系统或者设备(图 5、图 6),但是注意,示例实施例还提供对于更简单应用和情形的可缩放的解决方案。此外,噪声减少仅仅是确定可以通过本发明的实施例中描述的包括中值滤波的后处理方法后处理的增益的输入处理的一个示例。

[0219] 基于处理系统的设备

[0220] 图 8 示出用于处理例如来自麦克风(未示出)的音频输入 101 中的一个或多个的一个处理设备实施例 800 的简化框图。处理设备 800 要确定一组增益,后处理该增益(包括中值滤波所确定的增益),以及产生已经通过增益的应用修改的音频输出 137。一个版本实现基于感知域的调平、基于感知域的动态范围控制以及基于感知域的动态均衡中的一个或多个,其考虑了依赖于音频信号的再现电平的音频的感知的变化。另一个版本实现噪声减少。

[0221] 一个噪声减少版本包括回声减少,以及在这种版本中,处理设备还接受一个或多个参考信号 103,,例如,来自一个或多个扬声器(未示出)或者来自到一个或多个这种扬声器的一个或多个供给(feed)。在一个这种噪声减少版本中,处理设备 800 要产生已经通过抑制在一个实施例中噪声以及位置外信号,以及在另一实施例中还有回声而被修改的音频输出 137,如根据本发明的一个或多个特征指定的。设备例如可以实现图 6 中示出的系统,以及其任何可替代方案,并且可以在操作时实现包括在此描述的方法的任何变化的图 7 的方法。这种设备可以被包括例如在耳机(headset)(诸如蓝牙耳机)中。音频输入 101、一个或多个参考输入 103 以及音频输出 137 被假设为具有采样数据的 M 个样本的帧的形式。在模拟输入的情况下,将存在包括模数转换器和量化器的数字化仪(digitizer)。对于音频回放,将存在去量化器和数模转换器。可能被包括在完整的音频处理系统(例如,耳机装置)内的这种和其它元件被省去,并且本领域技术人员将清楚如何包括这种元件。

[0222] 图 8 中示出的实施例包括处理系统 803,其在操作中被配置为实现在此描述的抑制方法。处理系统 803 包括至少一个处理器 805,其可以是数字信号处理装置的一个或多个处理单元,或者更通用的处理装置的 CPU。处理系统 803 还包括典型地包括一个或多个存储元件的存储子系统 807。处理系统的元件例如通过图 8 中未示出的总线子系统或者其它一些互连机制耦接。处理系统 803 的一些元件可以被集成为单个电路,使用本领域技术人员公知的技术。

[0223] 存储子系统 807 包括在由一个或多个处理器 805 运行时使得实现在此描述的方法的指令 811。

[0224] 在一些实施例中,存储子系统 807 被配置为存储一个或多个调整参数 813,该调整参数可以被用来改变由处理系统 803 实现的一些处理步骤。

[0225] 图 8 中示出的系统可以被并入专业化装置(诸如耳机,例如,无线蓝牙耳机)中。该系统还可以是通用计算机的一部分,例如,被配置为处理音频信号的个人计算机。

[0226] 具有可设置的灵敏度的语音活动检测

[0227] 在本发明的一些实施例中,通过如 VAD 确定的信号分类控制后处理,例如,中值滤波。本发明不限于任何特定的类型 VAD,并且许多 VAD 在本领域中已知。当应用于抑制时,本发明人已经发现在抑制系统的不同的部件由不同的 VAD 控制时抑制工作得最好,每种 VAD 针对它被用在其中的抑制器的功能被定制设计,而不是对于所有使用具有“最佳的”VAD。因此,在针对噪声减少的输入处理的一些版本中,使用多个 VAD,每个 VAD 由分离地控制灵敏度和选择性的小集合的调整参数控制,包括空间选择性,这种参数根据其中使用 VAD 的抑制元件被调整。多个 VAD 中的每一个是根据 Y'_b 确定语音活动的指示的通用的 VAD 的例示。通用的 VAD 由一组参数控制并且使用噪声谱内容的估计值、回声的分带的频域振幅度量表示以及分带的空间特征。该组参数包括噪声谱内容的估计值是否是空间选择性的。特定的例示确定的语音活动的指示的类型由参数的选择控制。

[0228] 一般的空间选择性的 VAD 结构的一个实施例(用于计算可以针对各种函数被调整的语音活动的通用的 VAD)是

[0229]

$$S = \sum_{b=1}^B (\text{BeamGain}'_b)^{\text{BeamGainExp}} \left(\frac{\max(0, Y'_b - \beta_N \cdot (N'_b \vee N'_{b,S}) - \beta_E E'_b)}{Y'_b + Y'_{b,sens}} \right),$$

[0230] 其中 $\text{BeamGain}'_b = \text{BeamGain}_{\min} + (1 - \text{BeamGain}_{\min}) \text{RPI}_b \cdot \text{PPI}_b \cdot \text{CPI}_b$, BeamGainExp 是对于更大值增加 VAD 的空间选择性的积极性的参数,并且对于非空间选择性的 VAD 是 0, $N'_b \vee N'_{b,S}$ 表示总噪声功率(或者其它频域振幅度量)估计值 N'_b 或者使用波束外功率(或者其它频域振幅度量)确定的空间选择性的噪声估计值 $N'_{b,S}$, $\beta_N, \beta_E, 1$ 是分别对于噪声和回声的裕度以及 Y'_{sens} 是可设置的灵敏度偏移。 β_N, β_E 的值在 1 和 4 之间。 BeamGainExp 在期望空间选择性时在 0.5 到 2.0 之间,并且对于空间选择性的 VAD 的一个实施例是 1.5,例如,在本发明一些实施例中用于控制后处理。 $\text{PPI}'_b, \text{PPI}'_b$ 和 CPI'_b 如上所述是三个空间概率指示器,即比率概率指示器,相位概率指示器以及相干概率指示器。

[0231] 上述表达式还控制通用的语音活动检测方法的的操作。

[0232] 对于任何给出的用于产生语音指示器值 S 的一组参数,可以通过考虑如指示存在语音的测试 $S > S_{\text{thresh}}$ 来获得二元判定(binary decision)或者分类器。还应该明白的是值 S 可以被用作瞬时的语音电平连续指示器。此外,可以使用合适的“悬挂(hang over)”或者检测的事件之后的语音的继续指示的时段获得用于操作(诸如传输控制或者控制后处理)的改善的有用的通用的 VAD。这种悬挂时段可以从 0 变化到 500ms,并且在一个实施例中使用 200ms 的值。在悬挂时段期间,减少激活阈值(例如以 2/3 的因子减少)可以是有用的。一旦说话爆发(talk burst)已经开始,这产生对语音的增加的灵敏度以及稳定性。

[0233] 对于用于控制一个或多个后处理操作的空间选择性的语音活动检测,例如对于空间选择性的 VAD,上述表达式中的噪声是使用功率(或者其它频域振幅度量)的波束外估计值确定的 $N'_{b,S}$ 。 Y'_{sens} 被设定为通过对典型的组件进行实验获得的预期的麦克风以及系统噪声水平周围。

[0234] 原则

[0235] 除非具体陈述,否则,明白在整个说明书的讨论中使用诸如“产生”、“处理”、“计算”、“运算”、“确定”等之类的术语可以没有限制地指的是操纵和 / 或变换表示为物理量(诸如电子学量)的数据为类似地表示为物理量的其它数据的硬件(例如,电子电路,计算机或计算系统或者类似的电子计算装置)的动作和 / 或处理。

[0236] 以类似的方式,术语“处理器”可以指的是处理例如来自寄存器和 / 或存储器的电子数据以便将该电子数据变换为例如可以被存储在寄存器和 / 或存储器中的其它电子数据的任何装置或者装置的部分。“计算机”或“计算机器”或“计算平台”可以包括一个或更多个处理器。

[0237] 注意,在描述包括若干要素(例如,若干步骤)的方法时,除非具体地陈述,否则没有暗示这种要素(例如,这种步骤)的顺序。

[0238] 在本申请中描述的方法论在一些实施例中是可由接受在一个或更多个计算机可读介质上编码的指令、逻辑的一个或更多个处理器执行的。在由一个或更多个处理器执行时,指令使得实现在本申请中描述的至少一个方法。包括能够执行指定要采取的动作的一组指令(顺序的或者其它)的任何处理器。因此,一个示例是包括一个或更多个处理器的典型的处理系统。每个处理器可以包括 CPU 或类似的元件、图形处理单元(GPU)、现场可编程门阵列、专用集成电路和 / 或可编程的 DSP 单元中的一个或更多个。处理系统还包括具有至少一个存储介质的存储子系统,其可以包括嵌入半导体装置中的存储器,或者包括主 RAM 和 / 或静态 RAM、和 / 或 ROM、以及还包括高速缓冲存储器的分离的存储器子系统。存储子系统还可以包括一个或更多个其它存储装置,诸如磁和 / 或光学和 / 或另外的固态存储装置。可以包括总线子系统以用于在组件之间进行通信。处理系统还可以是具有通过网络(例如,经由网络接口装置或无线网络接口装置)耦接的处理器器的分布式处理系统。如果处理系统要求显示器,则可以包括这种显示器,例如,液晶显示器(LCD)、有机发光显示器(OLED)、或阴极射线管(CRT)显示器。如果要求手动的数据输入,则处理系统还包括诸如字母数字的输入单元(诸如键盘)、指向控制装置(诸如鼠标)等中的一个或更多个之类的输入装置。如果根据上下文是清楚的并且除非明确地陈述,否则,如本申请中使用的术语“存储装置”、“存储子系统”或“存储单元”还包括诸如盘驱动单元之类的存储系统。一些配置中的处理系统可以包括声音输出装置、以及网络接口装置。

[0239] 在一些实施例中,非暂时的计算机可读介质被配置具有(例如被编码具有)指令(例如,逻辑),该指令在被处理系统(诸如包括至少一个处理器元件和存储子系统的数字信号处理装置或者子系统)的一个或更多个处理器执行时使得实现如在本申请中描述的方法。一些实施例具有逻辑本身的形式。非暂时的计算机可读介质是具体地不是暂时传播的信号或暂时的载波或其它一些暂时的传输介质的任何计算机可读介质。术语“非暂时的计算机可读介质”因此覆盖任何有形的计算机可读的存储介质。非暂时的计算机可读介质包括任何有形的计算机可读的存储介质,并且可以采取包括非易失性存储介质和易失性存储介质在内的许多形式。非易失性存储介质包括,例如静态 RAM、光盘、磁盘以及磁光盘。易失性存储介质包括动态内存(诸如处理系统中的主内存),以及处理系统中的硬件寄存器。在如上所述的典型的处理系统中,存储子系统是计算机可读的存储介质,其被配置具有(例如被编码具有)指令,例如,逻辑,例如,软件,其在由一个或更多个处理器执行时使得实现在

本申请中描述的方法步骤中的一个或多个。软件可以存在于硬盘中,或还可以在其由计算机系统执行期间全部或至少部分地存在于内存(例如, RAM)内和 / 或处理器寄存器内。因此,内存以及处理器寄存器也构成其上可以被编码有指令以便使得在运行时实现方法步骤的非暂时的计算机可读介质。

[0240] 虽然计算机可读介质在示例实施例中所示出为单个介质,但是术语“介质”应该被视为包括存储一个或多个组的指令的单个介质或者多个介质(例如,若干内存、集中式或者分布式数据库、和 / 或关联的闪存和服务器的)。

[0241] 此外,非暂时的计算机可读介质,例如,计算机可读的存储介质可以形成计算机程序产品或者被包括在计算机程序产品内。

[0242] 在可替代实施例中,一个或多个处理器操作作为独立的装置或可以在网络连接的部署中连接(例如,网络连接)到一个或多个其它处理器,其中一个或多个处理器可以在服务器-客户端网络环境中以服务器的身份或客户端机器的身份操作,或者在对等或分布式网络环境中操作作为对等机器。除非在本申请中明确地排除,否则术语“处理系统”包括所有这样的可能性。一个或多个处理器可以形成个人计算机(PC)、媒体回放装置、耳机装置、免提通信装置、平板 PC、机顶盒(STB)、个人数字助理(PDA)、游戏机、蜂窝式电话、Web 用具、网络路由器、开关或桥、或者能够执行指定要由该机器采取的动作的一组指令(顺序的或者其它)的任何机器。

[0243] 注意,虽然一些图仅仅示出了单个处理器和单个存储子系统,例如,存储包括指令的逻辑的单个内存,但是本领域技术人员将理解上面描述的组件中的许多组件被包括,但为了不模糊有创造力的方面不明确地示出或描述。例如,虽然仅仅示出单个机器,但是术语“机器”应该还被视为包括独立地或联合地执行一组(或多组)指令以便执行在本申请中讨论的方法论中的任何一个或多个的机器的任何集合。

[0244] 因此,如本领域技术人员将明白的,本发明的实施例可以被具体实现为方法、设备(诸如特殊用途设备)、设备(诸如数据处理系统)、逻辑(例如,具体实现在非暂时的计算机可读介质中)、或者编码有指令的非暂时的计算机可读介质(例如,配置作为计算机程序产品的计算机可读存储介质)。计算机可读介质被配置具有在由一个或多个处理器执行时使得实现方法步骤的一组指令。因此,本发明的方面可以采取方法、完全硬件实施例、完全软件实施例或者结合软件和硬件方面的实施例的形式。此外,本发明可以采取程序逻辑(例如,计算机可读的存储介质上的计算机程序)或者配置具有计算机可读的程序代码的计算机可读存储介质(例如,计算机程序产品)的形式。

[0245] 还将理解,本发明的实施例不限于任何特定的实现方式或者编程技术并且本发明可以使用用于实现在本申请中描述的功能的任何适当的技术来实现。此外,实施例不限于任何特定的编程语言或者操作系统。

[0246] 在整个本说明书期间提及“一个实施例”或者“实施例”意味着与该实施例结合描述的特定的特征、结构或者特性被包括在本发明的至少一个实施例内。因此,在整个本说明书中的各个地方中的短语“在一个实施例中”或者“在实施例中”的出现不一定都指的是同一个实施例,但是可能。此外,如本领域技术人员根据本公开会清楚的,可以在一个或多个实施例中以任何合适的方式结合特定的特征、结构或者特性。

[0247] 类似地,应当明白,在本发明的示例实施例的上述描述中,出于使本公开简单化并

且帮助理解各个有独创性的方面中的一个或更多个的目的,本发明的各个特征有时被集合在一起在单个实施例、图或者其说明中。然而,公开的此方法不被解释为反映所要求保护的发明需要比每个权利要求中明确叙述的特征更多的特征的意图。相反,如下面的权利要求反映的,有独创性的方面依赖于小于单个在前公开的实施例的所有特征。因此,具体实施方式之后的权利要求由此被明确地并入本具体实施方式中,每个权利要求自身保持为单独的本发明的实施例。

[0248] 此外,虽然在本申请中描述的一些实施例包括包含在其它实施例内的一些但不包括其它特征,但是不同的实施例的特征的组合意图在本发明范围内,并且形成不同的实施例,如本领域技术人员会理解的。例如,在下面权利要求中,能以任何组合来使用任何要求保护的实施例。

[0249] 此外,在本申请中描述一些实施例作为能由计算机系统的处理器或者由实现该功能的其它部件实现的方法或者方法的要素的组合。因此,具有用于实现这种方法或者方法的要素的必需的指令的处理器形成用于实现方法或者方法的要素的部件。此外,在本申请中描述的设备实施例的元件是用于实现由出于实现发明的目的的元件执行的功能的部件的示例。

[0250] 在本申请中提供的说明中,陈述了许多的具体细节。然而,应当理解本发明的实施例可以在没有这些具体细节的情况下被实践。在其它例子中,为了不模糊对该说明的理解,没有详细示出公知的方法、结构和技术。

[0251] 如在此使用的,除非另有规定,用于描述共同的对象的顺序形容词“第一”、“第二”、“第三”等的使用仅仅指示类似的对象的不同实例正在被提及,并不意图暗示如此描述的对象必须时间地、空间地、分等级地或者以任何其它方式具有给出的序列。

[0252] 虽然在一个实施例中,短时间傅里叶变换(STFT)被用来获得频带,但是本发明不限于 STFT。诸如 STFT 之类的变换通常被称为循环的(circulant)变换。循环的变换的最一般形式可以由缓冲、加窗、扭曲(twist)(实数值到复数值变换)以及 DFT(例如 FFT)代表。DFT 之后的复数的扭曲可以被用来调节频域表示以便匹配具体的变换定义。本发明可以通过这类变换(包括修改 DFT(MDFT)、短时间傅里叶变换(STFT)以及具有更长窗口和缠绕(wrapping)、共轭正交镜像滤波器(CQMF))中的任意一个实现。其它标准变换(诸如修改离散余弦变换(MDCT)和修改离散正弦变换(MDST))也可以被使用,具有频域格的额外的复数的扭曲,其不改变底层的变换的处理能力或者频率分辨率并且因此可以被保留直到处理链的结束,并且如果需要的话被应用在重映射中。

[0253] 在本申请中引用的所有美国专利、美国专利申请、以及指定美国的国际(PCT)专利申请通过参考被并入于此,除了那些不容许通过参考并入的权限(jurisdiction)之外,在该情况下本申请人保留通过修改将所有这样的材料或者其任何部分插入说明书中而这种插入不被认为是新东西(new matter)的权利。在专利实施细则或者法规不容许通过参考并入本身通过参考并入信息的材料的情况下,在本申请中的材料的通过参考的并入在这种通过参考并入的材料中排除通过参考并入的任何信息,除非这种信息明确地通过参考并入于此。

[0254] 本说明书中的其它任何现有技术的讨论决不应该被认为是在发明时这种现有技术是普遍已知的、公众已知的或者形成本领域中的常识的一部分的认可。

[0255] 在下面的权利要求以及本申请中的说明中,术语“包括”、“由...组成”或者“包含”中的任何一个都是开放术语,其意指至少包括随后的元件/特征,但不排除其它。因此,在权利要求中使用术语“包括”不应该被解释为限制其后列出的部件或者元件或者步骤。例如,表述“装置包括A和B”的范围不应该被限于仅由元件A和B组成的装置。如本文中使用的术语“包含”或者“其包含”或者“包含……的”中的任何一个也是开放术语,其也意指至少包含术语随后的元件/特征,但不排除其它。因此,“包含”和“包括”同义并且意指“包括”。

[0256] 类似地,要注意,在权利要求中使用术语“耦接”不应该被解释为限制为仅仅直接连接。可以使用术语“耦接”和“连接”及其派生词。应当理解这些术语不意图作为彼此同义词,但是可以是。因此,表述“装置A耦接到装置B”的范围不应该限于其中装置A的输入或输出直接连接到装置B的输出或输入的装置或者系统。这意指在装置A和装置B之间存在可以是两者之间包含其它装置或者部件的路径的路径。此外,“耦接到”不暗示方向。因此,表述“装置A耦接到装置B”可以和表述“装置B耦接到装置A”同义。“耦接”可以意指两个或更多个元件直接物理接触或者电接触,或者两个或更多个元件不彼此直接接触但仍然合作或者彼此相互作用。

[0257] 另外,“一(a)”或“一个(an)”的使用被用来描述在本申请中实施例的元件和组件。这仅仅为了方便起见进行并且给出本发明的一般含义。这个描述应该被阅读为包括一个或至少一个,并且单数也包括多个,除非很明显它意味着“一个”。

[0258] 因此,虽然已经描述了被认为是本发明的优选实施例的实施例,但是本领域技术人员将认识到在不脱离本发明的精神的情况下可以对其进行其它和进一步修改,并且意图要求保护如落入本发明的范围内的所有这样的改变和修改。例如,上面给出的任何公式仅仅代表可以使用的过程。可以在框图中添加或者从框图中删除功能,并且可以在功能块之间交换操作。可以在本发明范围内对描述的方法添加或者删除步骤。

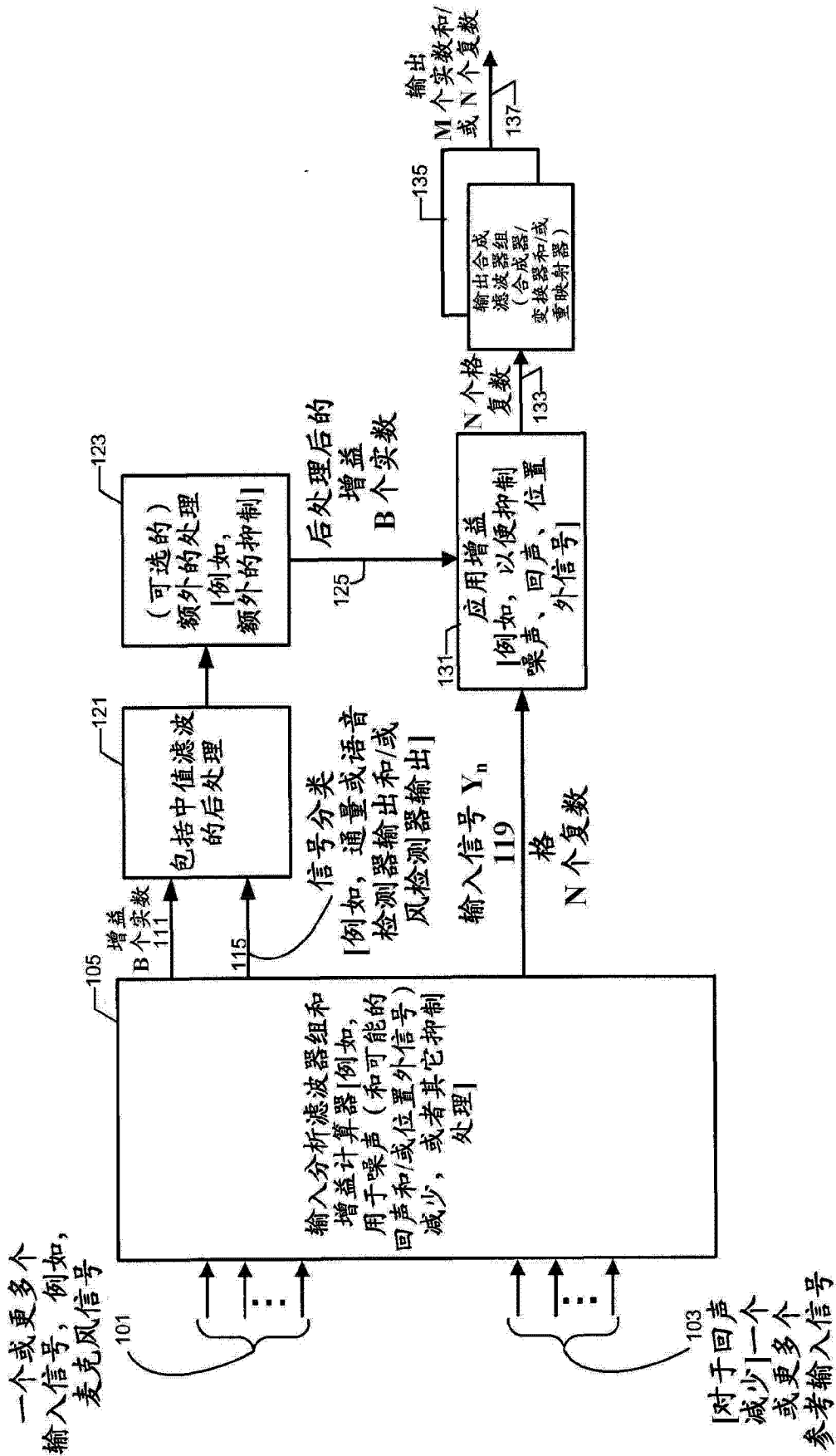


图 1

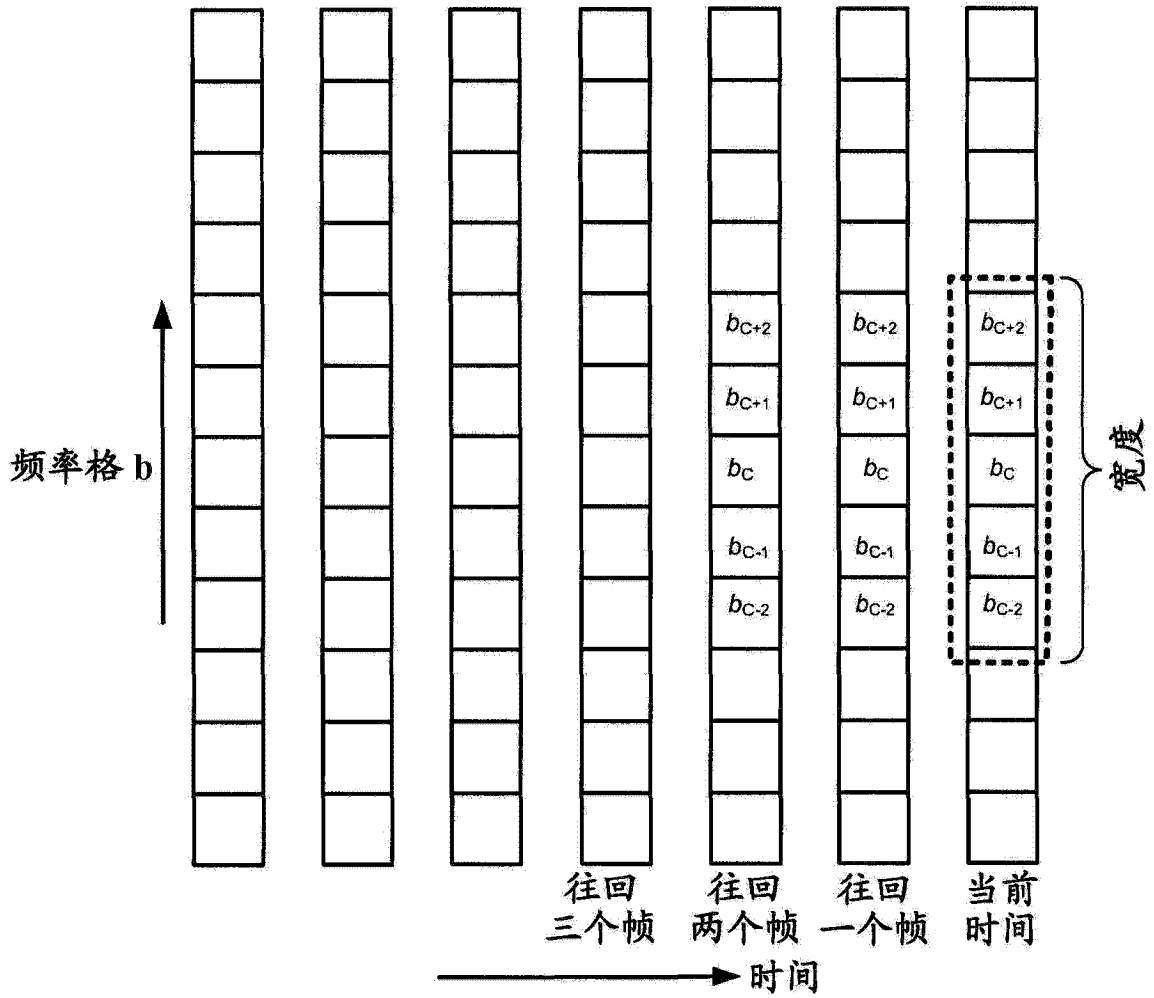


图 2

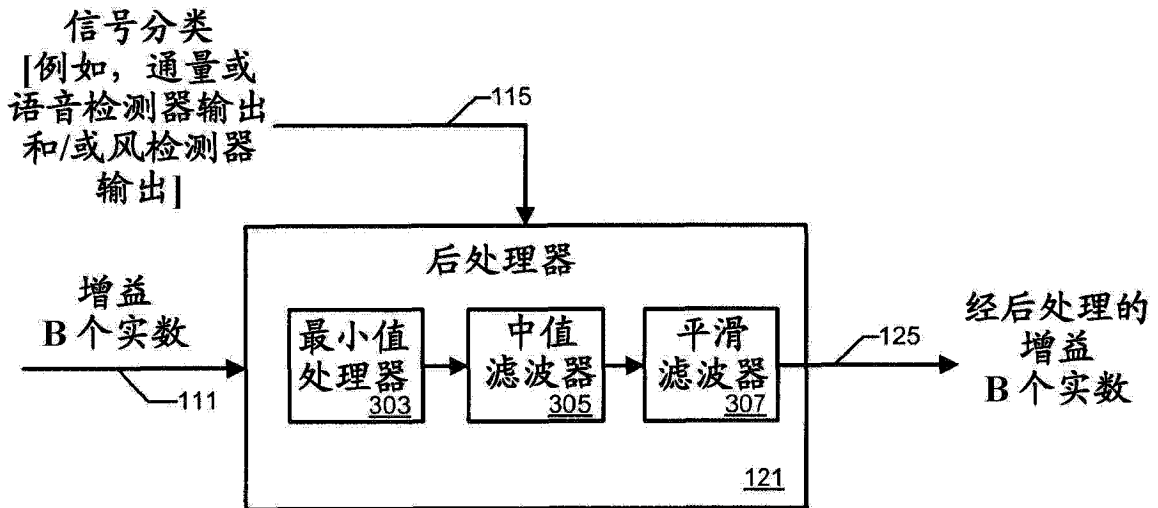


图 3A

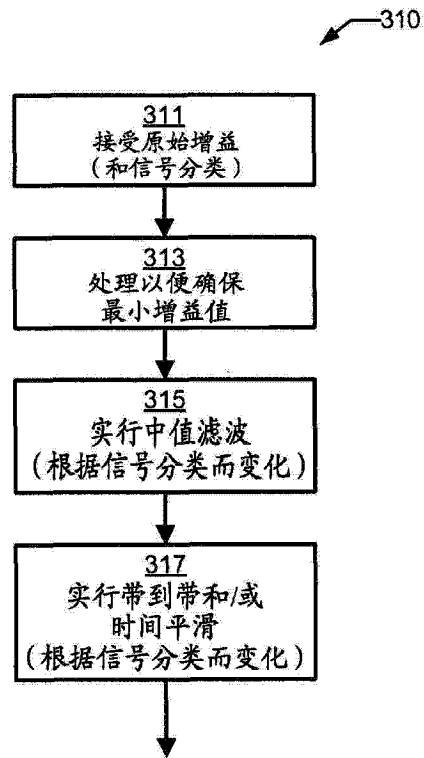


图 3B

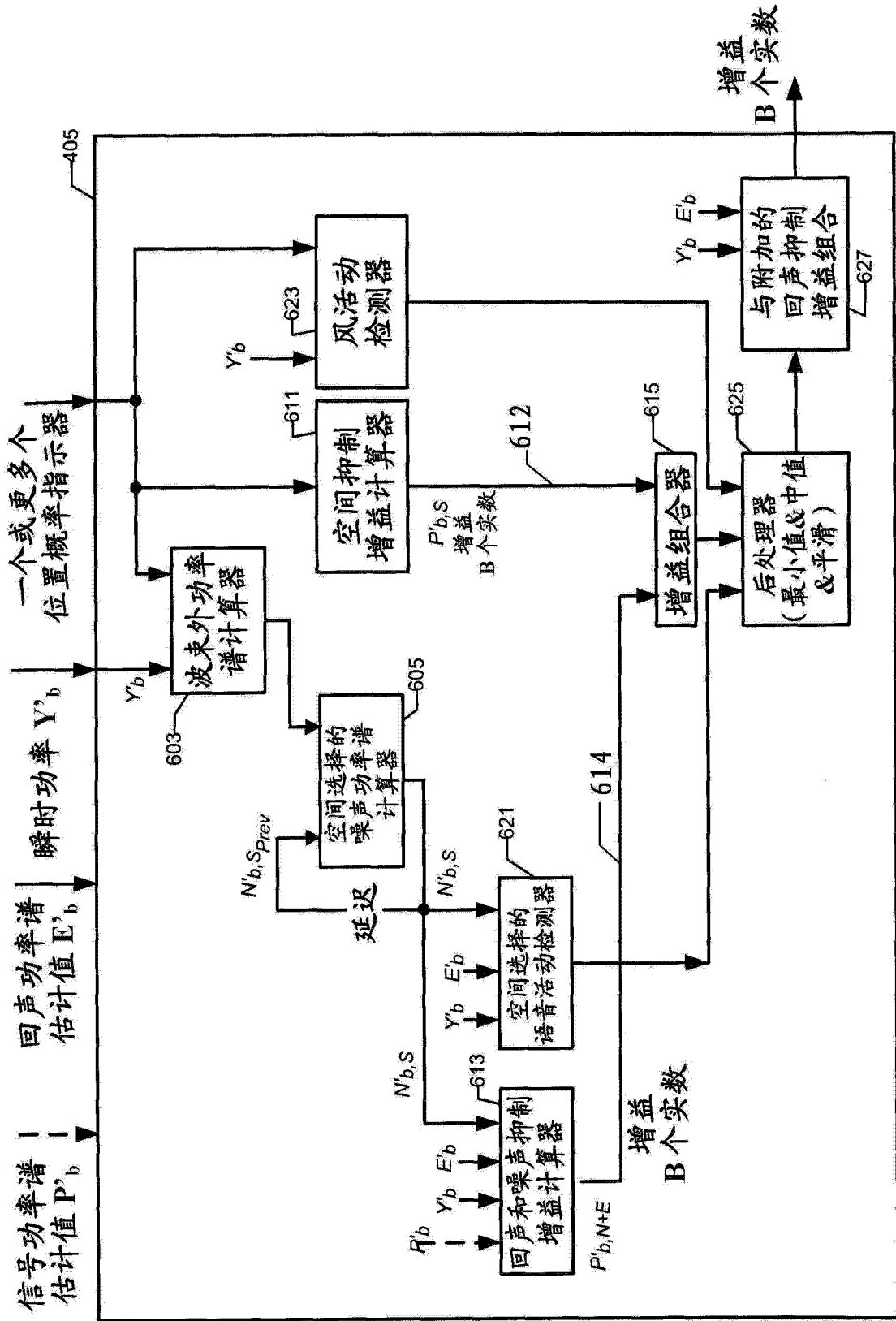


图 6

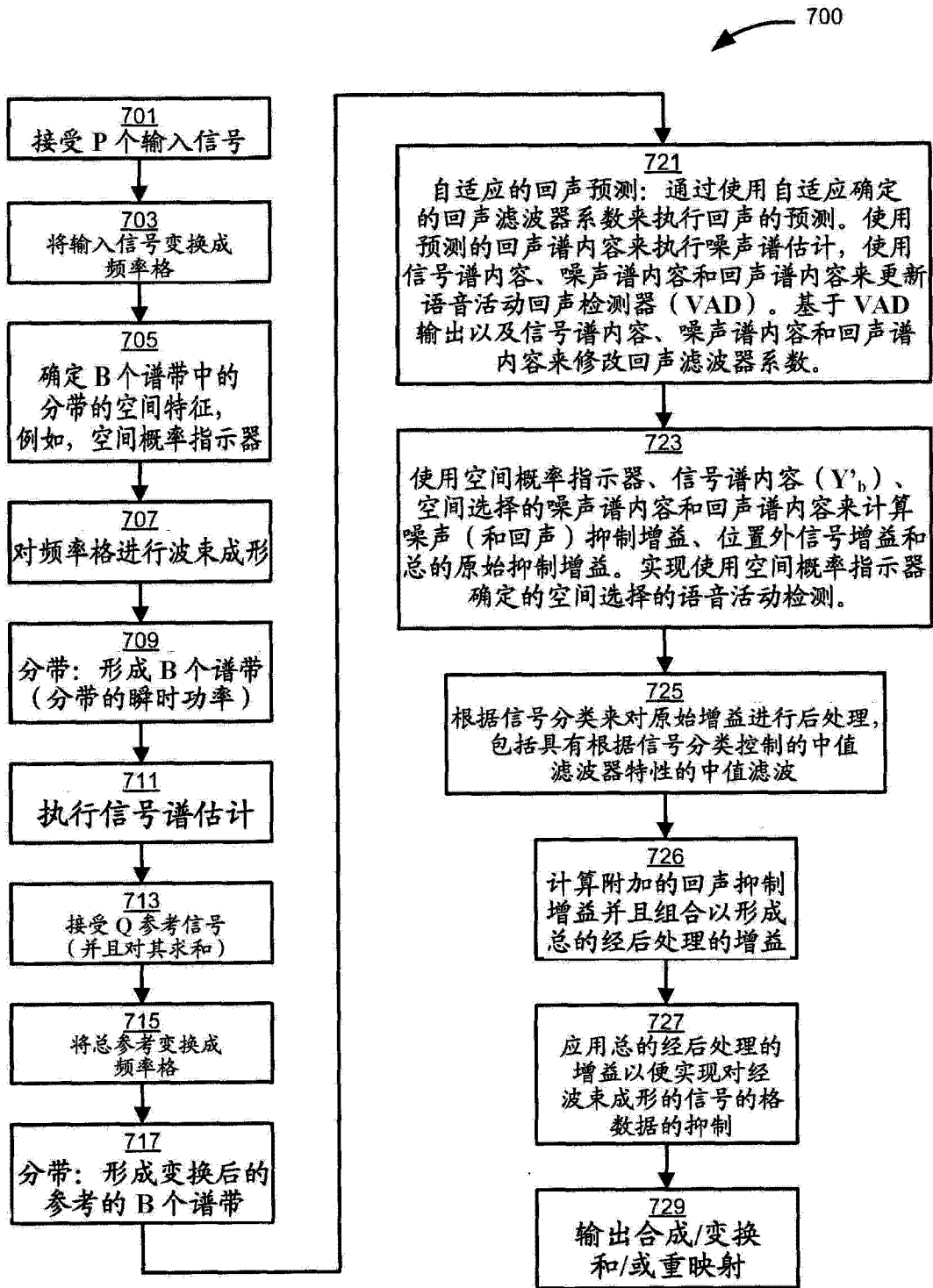


图 7

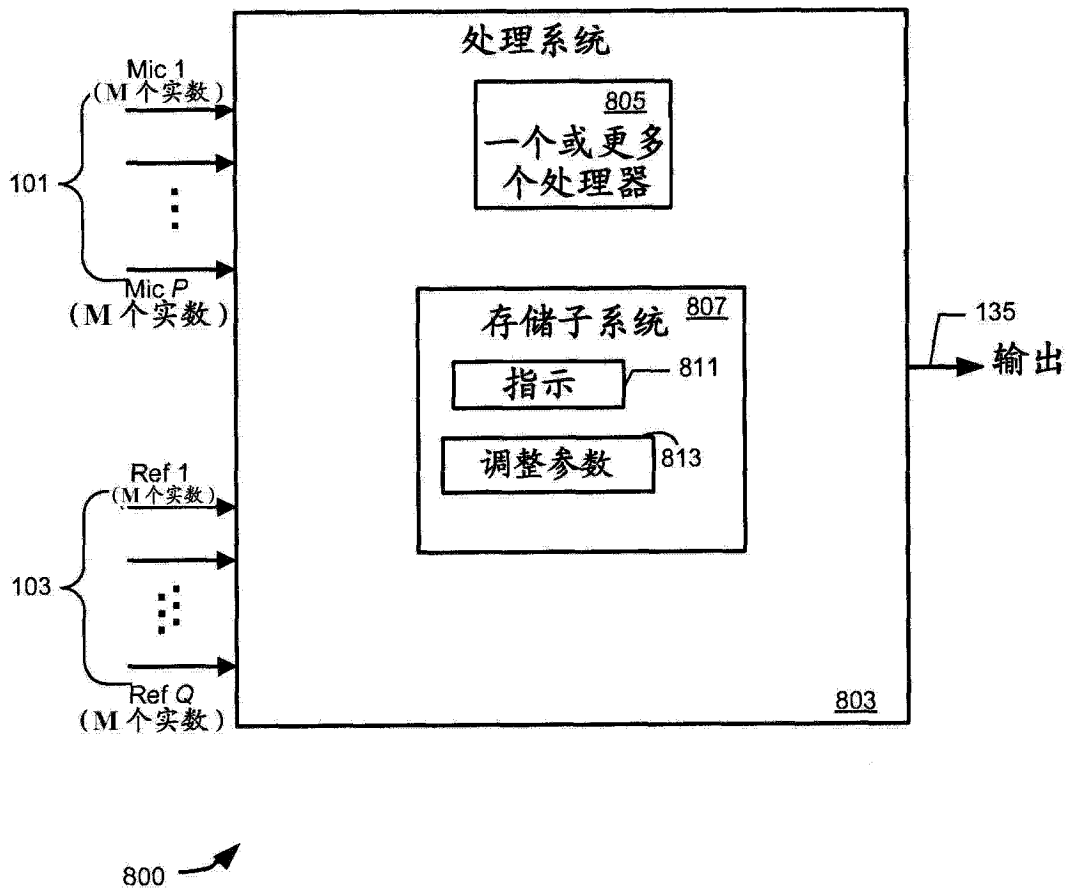


图 8