



[12] 发明专利申请公开说明书

[21] 申请号 01818918.0

[43] 公开日 2004 年 3 月 31 日

[11] 公开号 CN 1486566A

[22] 申请日 2001.9.14 [21] 申请号 01818918.0

[30] 优先权

[32] 2000. 9. 15 [33] US [31] 09/663,748

[86] 国际申请 PCT/US01/28825 2001.9.14

[87] 国际公布 WO02/23755 英 2002.3.21

[85] 进入国家阶段日期 2003.5.15

[71] 申请人 英特尔公司

地址 美国加利福尼亚州

[72] 发明人 M·E·戴舍尔

[74] 专利代理机构 中国专利代理(香港)有限公司

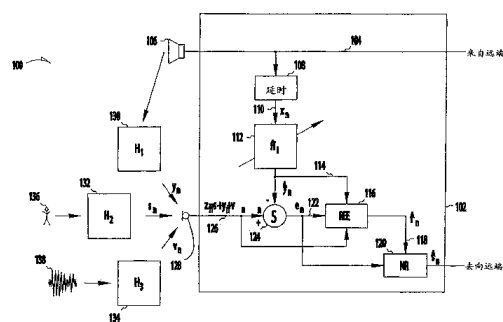
代理人 程天正 王 勇

权利要求书 5 页 说明书 10 页 附图 4 页

[54] 发明名称 用于回声抵消的剩余回声估计

[57] 摘要

回声抵消系统中剩余回声的离散傅立叶变换(DFT)的实部和虚部分量是通过使用没有作在近端话筒输入信号的分开分量之间的统计独立性的假设得出的变换函数被分开地估计的。剩余回声估值器通过使用不带有回声的话筒输入信号的近似,计算分量的实部和虚部。取决于近端讲话人和/或远端讲话人是否正在讲话,使用不同的近似。



1. 声波回声抵消系统, 包括:

自适应滤波器, 用于产生回声估值和包括剩余回声的误差项, 以及

5 剩余回声估值器, 用于独立地计算剩余回声的离散傅立叶变换 (DFT) 的实部和虚部的估值。

2. 权利要求 1 的波声回声抵消系统, 其中剩余回声估值器包括:

第一输入节点, 用于接收复合信号, 其中包括近端语音、噪声、和远端语音的回声;

10 第二输入节点, 用于接收回声估值; 以及

第三输入节点, 用于接收在回声估值与复合信号之间的差值。

3. 权利要求 2 的声波回声抵消系统, 其中剩余回声估值器被配置成当近端语音为活动的和远端语音处于静默时通过使用复合信号作为近端语音与噪声的估值来计算剩余回声的 DFT 的实部和虚部。

15 4. 权利要求 2 的波声回声抵消系统, 其中剩余回声估值器被配置成当近端语音和远端语音两者都是活动的时通过使用复合信号与先前的输出样本的平均值作为近端语音与噪声的估值来计算剩余回声的 DFT 的实部和虚部。

20 5. 权利要求 2 的声波回声抵消系统, 其中剩余回声估值器被配置成当近端语音是静默和远端语音是活动的时通过使用基本上为零的数值作为近端语音与噪声的估值来计算剩余回声的 DFT 的实部和虚部。

6. 权利要求 1 的声波回声抵消系统, 还包括一个响应于剩余回声的 DFT 的实部和虚部以及在回声估值与复合信号之间差值的噪声减小电路。

25 7. 扬声器电话, 包括:

一个被耦合到扬声器的输出端口, 用于把远端语音发送到声音环境;

一个被耦合到话筒的输入端口, 用于从声音环境接收包括近端语音、噪声、和远端语音的回声的复合信号;

30 自适应滤波器, 用于产生远端语音的回声的估值; 以及

剩余回声估值器, 用于独立地估计在远端语音的回声与远端语音的回声的估值之间的差值的 DFT 的实部和虚部。

8. 权利要求 7 的扬声器电话, 其中剩余回声估值器包括:

矢量项计算器, 用于计算由于缺少近端语音、噪声、和远端语音的回声在统计上是独立的假设而造成的矢量项的估值。

5 9. 权利要求 8 的扬声器电话, 其中矢量项计算器被配置成: 根据究竟是远端语音还是近端语音具有大于门限值的能量, 从而不同地估计近端语音与噪声的和值。

10 10. 权利要求 9 的扬声器电话, 其中矢量项计算器被配置成: 当近端语音活动性检测器指示近端语音具有大于门限值的能量以及双向讲话检测器指示远端语音具有低于门限值的能量时, 估计该和值作为复合信号。

11. 权利要求 8 的扬声器电话, 还包括:

噪声减小电路, 响应于剩余回声的实部和虚部以产生近端语音的估值; 以及

15 用来估计近端语音与噪声的和值的电路, 该电路响应于双向讲话指示器, 该电路还被配置成用来估计该和值作为复合信号与近端语音的先前的估值的平均值。

12. 权利要求 8 的扬声器电话, 其中扬声器电话在计算机中实施。

13. 估计声波回声抵消系统中的剩余回声的方法, 该方法包括分开地估计剩余回声的 DFT 的实部和虚部。

20 14. 权利要求 13 的方法, 其中分开地估计包括:

根据近端复合信号的 DFT 的实部、自适应滤波器输出信号的 DFT 的实部、由近端复合信号与自适应滤波器输出信号之间的差值产生的误差信号的 DFT 的实部、和近端语音信号的 DFT 的实部与近端噪声信号的 DFT 的实部的和值的近似值, 从而估计剩余回声的 DFT 的实部。

25 15. 权利要求 14 的方法, 其中剩余回声的 DFT 的实部的平方按照基本上类似于下式的公式而被估计:

其中:

是剩余回声的 DFT 的实部;

是误差信号的 DFT 的实部;

30 是近端复合信号的 DFT 的实部;

是自适应滤波器输出信号的 DFT 的实部; 以及

表示近端语音信号的 DFT 的实部与近端噪声信号的 DFT 的实部的

和值的近似值。

16. 权利要求 15 的方法，还包括：

当近端讲话人处于静默时，使用近端复合信号的 DFT 的实部来近似近端语音信号的 DFT 的实部与近端噪声信号的 DFT 的实部的和值。

5 17. 权利要求 16 的方法，其中剩余回声的 DFT 的实部的平方按照基本上类似于下式的公式而被估计：

18. 权利要求 15 的方法，还包括：

10 只在近端讲话人处于静默时，才使用基本上为零的数值来近似近端语音信号的 DFT 的实部与近端噪声信号的 DFT 的实部的和值。

19. 权利要求 18 的方法，其中剩余回声的 DFT 的实部的平方按照基本上类似于下式的公式而被估计：

15

20. 权利要求 15 的方法，还包括：

当存在双向讲话事件时，使用近端复合信号的 DFT 的实部与来自回声抵消系统的以前输出的 DFT 的和值来近似近端语音信号的 DFT 的实部与近端噪声信号的 DFT 的实部的和值。

20 21. 权利要求 20 的方法，其中剩余回声的 DFT 的实部的平方按照基本上类似于下式的公式而被估计：

其中代表来自回声抵消系统的以前输出的 DFT 的实部。

25 22. 权利要求 13 的方法，其中分开地估计包括：

通过使用包括由于缺少在话筒处接收的近端复合信号中的信号分量在统计上是独立的假设而造成的矢量项的公式来估计剩余回声的 DFT 的虚部。

23. 权利要求 22 的方法，其中矢量项包括回声估值的 DFT 的虚部和近端语音信号的 DFT 的虚部与近端噪声信号的 DFT 的虚部的和值的乘积，以及其中估计剩余回声的 DFT 的虚部的步骤包括：

30 用一个不同的乘积来代替矢量项，以作为近端语音信号是否包含

足够的能量以及远端语音信号是否包含足够的能量的函数。

24. 权利要求 22 的方法，其中剩余回声的 DFT 的虚部的平方按照基本上类似于下式的公式而被估计：

其中：

- 5 是剩余回声的 DFT 的虚部；
- 是误差信号的 DFT 的虚部；
- 是近端复合信号的 DFT 的虚部；
- 是自适应滤波器输出信号的 DFT 的虚部；以及
- 表示近端语音信号的 DFT 的虚部与近端噪声信号的 DFT 的虚部的
- 10 和值的近似值。

25. 权利要求 24 的方法，还包括：

当近端讲话人处于静默时，使用近端复合信号的 DFT 的虚部来近似近端语音信号的 DFT 的虚部与近端噪声信号的 DFT 的虚部的和值。

- 26. 权利要求 25 的方法，其中剩余回声的 DFT 的虚部的平方按照
- 15 基本上类似于下式的公式而被估计：

27. 权利要求 24 的方法，还包括：

- 只在近端讲话人处于静默时，才使用基本上为零的数值来近似近
- 20 端语音信号的 DFT 的虚部与近端噪声信号的 DFT 的虚部的和值。

28. 权利要求 27 的方法，其中剩余回声的 DFT 的虚部的平方按照基本上类似于下式的公式而被估计：

- 25 29. 权利要求 22 的方法，还包括：

当存在双向讲话事件时，使用近端复合信号的 DFT 的虚部与来自声波回声抵消系统的以前输出的 DFT 的和值来近似近端语音信号的 DFT 的虚部与近端噪声信号的 DFT 的虚部的和值。

- 30 30. 权利要求 20 的方法，其中剩余回声的 DFT 的虚部的平方按照基本上类似于下式的公式而被估计：

其中表示来自回声抵消系统的以前输出的 DFT 的虚部。

用于回声抵消的剩余回声估计

技术领域

- 5 本发明总的涉及回声抵消系统，更具体地，涉及回声抵消系统中用于减小剩余回声的系统。

背景技术

- 10 在“扬声器电话”应用中，“近区讲话人”是使用扬声器电话的人，以及“远区讲话人”是在电话线的远端的人。远区讲话人的语音被广播（通过扬声器播放）到扬声器电话和近区讲话人所在的室内（或其他的声音封闭空间）。回声是通过远区讲话人的语音传播到室内并随后在话筒处被接收而产生的。声波回声抵消器（AEC）被使用来抵消在话筒处接收的回声。声波回声抵消器典型地是自适应滤波器，它模拟室内的各个回声路径。存在声波回声抵消器基本上不能抵消回声的极端的
- 15 情形，诸如，拉长的双重讲话或快速的近区讲话人的运动。在这些情形下，远区讲话人可能听到成为回声的他/她自己的声音的短的突发。

- 剩余回声抑制技术试图去除在声波回声抵消后剩余的回声，由此阻止远区讲话人听见这样的突发，而同时仍旧允许近区讲话人的话音不被打扰地传送通过。剩余回声抑制技术一直是研究的课题，已发表了许多描述有关工作的论文。例子包括：S. Gustafsson 等的“Combined Acoustic Echo Control and Noise Reduction for Hands-Free Telephony（用于免提电话的組合的声波回声控制和噪声减小）”，
- 25 Signal Processing, Vol. 64, pp. 21-32, 1998, 此后称为“Gustafsson”；以及 V. Turbin 等的“Using Psychoacoustic Criteria in 声波 Acoustic Echo Cancellation Algorithms（在回声抵消算法中使用心理声音准则）”，Proc. IWAENC'97, London, pp. 53-56, Sept. 1997, 此后称为“Turbin”。尽管正在进行的研究工作，当前的剩余回声抑制技术在去除剩余回声方面并不是完全有效的。
- 30

附图简述

- 图 1 显示带有剩余回声估计的声波回声抵消系统；
图 2 显示矢量项计算器的第一部分；
图 3 显示矢量项计算器的第二部分；以及
5 图 4 显示在扬声器电话中的语音活动性检测器。

实施例描述

在以下的实施例详细说明中，参考附图并通过说明，显示了其中
可以实施本发明的特定的实施例。在附图上，相同的数字描述在几个
10 图上基本上相同的部件。这些实施例以足够的细节被描述，使得本领域技术人员能够实施本发明。也可以利用其他的实施例，以及可以作出结构的、逻辑的和电学的改变，而不背离本发明的范围。而且，应该理解，本发明的各种实施例虽然不同，但并不一定互相排斥。例如，
在一个实施例中描述的特定的特性，结构，或特征可被包括在其他实
15 施例内。所以，以下的详细说明不是以限制性的意义作出的，本发明的范围只由附属的权利要求书规定，连同由这样的权利要求被授权的等价物的全部范围。

本发明的方法和设备提供在声波回声抵消系统中估计剩余回声的
机制。自适应滤波器产生回声的估值，以及从话筒输入信号中减去回
20 声的估值。当回声和回声的估值不完全一致时，不想要的剩余回声仍保留在信号中。本发明的方法和设备提供了一种剩余回声估值器，它在没有在话筒输入信号的各个分量之间作出统计学的独立性假设的情况下分开地估计剩余回声的离散傅立叶变换 (DFT) 的实部和虚部。剩余回声估值器使用不带有回声的话筒输入信号的近似值来计算实部和
25 虚部。根据是近区讲话人还是远区讲话人在讲话，使用不同的近似值。

图 1 显示具有回声抵消和剩余回声估值的扬声器电话。扬声器电
话 102 包括自适应滤波器 112，剩余回声估值器 (REE) 116，加法器
124，噪声减小 (NR) 电路 120，扬声器 106，和话筒 128。扬声器电
话 102 可包括许多其他电路，诸如处理器，放大器等等。许多这样的
30 其他电路在图 1 上被省略，以便清楚地显示本发明的新颖部分。扬声器电话 102 包括扬声器 106，它播放来自远区讲话人的语音。远区讲话人语音在节点 104 处从远端被接收，以及通过扬声器 106 被播放。在

节点 104 处的语音数据在通过自适应滤波器 112 处理之前还要被延时电路 108 延时。

在某些实施例中，在节点 104 处的语音数据是数字数据。在这些实施例中，数字-模拟 (D/A) 变换器把数字数据变换成模拟波形，供
5 扬声器 106 播放。在这些实施例中，延时电路 108 可以是数字延时单元，诸如寄存器存储器或移位寄存器。在其他实施例中，在节点 104 处的数据是模拟数据。在这些实施例中，延时电路 108 可以是模拟延时单元，诸如集中的或分布的模拟延时线。延时电路 108 可以是固定延时或可变延时。通常，当回声路径是较长时，延时电路 108 中的延
10 时是较长的。例如，当扬声器电话 102 被使用在汽车中时，回声路径是短的，以及延时电路 108 的延时是短的。同样例如，当扬声器电话 102 被使用在大的会议室时，回声路径更长，以及延时电路 108 的延时也相应地更长。

在运行时，扬声器 106 把远端语音播放到声音环境中。来自声音
15 封闭空间的回声信号 y_n 在话筒 128 处被接收。声音封闭空间对回声的影响由方块 130 代表，它包含转移函数 H_1 。也就是说，由扬声器播放的语音受到转移函数 H_1 的作用，然后在话筒 128 处被接收。除了回声 y_n 以外，话筒 128 也接收近端语音信号 s_n 和近端噪声信号 v_n 。近端语音信号 s_n 代表来自近区讲话人 136 的语音，由转移函数 H_2 建模，如方
20 块 132 所示，以及噪声信号 v_n 代表噪声 138，由转移函数 H_3 建模，如方块 134 所示。所以，话筒 128 接收 y_n 、 s_n 和 v_n 的总和。这在节点 126 处显示为 z_n 。

至少在某种程度上， y_n 和 v_n 代表 z_n 中的不希望的分量。例如， y_n 是由扬声器 106 与话筒 128 的耦合造成的。当 y_n 的大的分量没有被抵
25 消，以及成为发回到远端的信号的一部分时，远区讲话人会听到他/她的经过一段延时后的语音，这会在打电话期间引起混乱。另外，同样例如，如果 v_n 过分大，则远区讲话人听到大量噪声，以及这会打断电话交谈。与此相反，近端语音信号通常是 z_n 中的想要的分量。当近端语音忠实地传输到远端时，可以进行有效的通信。

30 自适应滤波器 112 把回声的转移函数近似为 $\hat{H}$ ，以及从延时的远端语音信号 x_n 产生回声信号 y_n 的估值 $\hat{y}_n$ 。加法器 124 从 z_n 中减去 $\hat{y}_n$ ，在节点 122 处产生 e_n 。自适应滤波器 112 可以是适合于自适应回声抵消系统

使用的任何滤波器。例子包括(但不限于): S. Gay, S. Tavathia, "The fast affine projection algorithm(快速拟似投射算法)", Proc. IEEE ICASSP, Detroit, USA, 1995, pp. 3023-3026; 和 G. Glentis, K. Berberidis, "A Unified view: Efficient least squares adaptive algorithms for FIR transversal filtering(统一观点: 用于 FIR 横向滤波的有效的最小平方自适应算法)", IEEE Signal Processing Magazine, vol. 16, pp. 13-41, July 1999.

z_n , y_n 和 e_n 被输入到剩余回声估值器 116, 产生剩余回声的估值。剩余回声估值和 e_n 被输入到噪声减小(NR)电路 120 以减小在 e_n 中的噪声。理想地, $\hat{y}_n$ 与 y_n 完全一致, 以及 e_n 只包含近端语音信号 s_n 以及或许某些背景噪声 v_n 。然而, 实际上这是不可能的, e_n 总是包含某些剩余回声量 $r_n = y_n - \hat{y}_n$, 即, $e_n = s_n + v_n + r_n$ 。当远区讲话人处于静默时($x_n = 0$), 加到噪声减小算法上的输入只是语音加背景噪声。当远区讲话人讲话时, "噪声"(或不希望的分量)包括背景噪声 v_n 加回声 y_n 。噪声减小电路 120 使用噪声谱估值器来跟踪近区静止背景噪声。准静止剩余回声通过剩余回声估值器(REE)被分开地估值。然后通过使用复合的"噪声"估值实现 NR。

在某些实施例中, 噪声减小电路 120 留下小量的近区静止背景噪声来掩蔽剩余回声。这可以通过使用在 Tubin 中讨论的算法(见以上的背景部分)来完成。

任何的各种类别的语音谱估值器可用来修正噪声功率谱估值, 以便计入剩余回声。适当的噪声减小算法的例子和电路可以在以下论文中找到: S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction(通过使用频谱相减抑制语音中的噪声声音)", IEEE Trans. Acoust. Speech Signal Process. ASSP-27(April 1979) 113-120; 和 Y. Ephraim, D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator(使用均方误差短时间频谱幅度估值器的语音增强)", IEEE Trans. Acoust. Speech Signal Process. 32(6) (December 1984) 1109-1121.

剩余回声估值

Gustafsson 的以前的著作（见以上的背景部分）假设：先验的剩余回声统计可以通过一个加到自适应滤波器输出端的实数值确定性变换函数 F_n 而被估值。

变换函数是：

- 5 其中大写字母代表由小写字母代表的相应变量的离散傅立叶变换（DFT）分量。例如， $E_n(k)$ 代表 e_n 的 DFT 分量。

10 Gustafsson 从上述的他们的著作中省略关于 $E\{|Z_n(k)|^2\}$ ， $E\{|E_n(k)|^2\}$ 和的估值的细节，但它们典型地使用样本值的“泄漏平均”来估值。使用在真实值与估计的剩余回波功率谱之间的欧几里得距离以量度估值器的成功程度而得到的实验结果表明：瞬时值给出非常好的结果。还是在 Gustafsson 中，虽然没有讨论子频带的数目和宽度， F_n 被划分成子频带以及每个 $F_n(k)$ 被子频带中的平均值代替。在太小的情况下， $F_n(k)$ 没有被包括在平均值中。实验结果表示，当子频带尺寸是一个 DFT 频点（bin）时达到低的估值误差。

- 15 以上的公式（1）和（2）作了两个假设： F_n 是实数值，以及是 s_n ， v_n 和 y_n 是统计上独立的。对于剩余回声估值在使用公式（1）和（2）的实验结果时产生的估值剩余回声功率谱的信噪比最好为约 2.0dB。在记住这两个假设的情况下，现在讨论公式（2）的推导。然后去掉两个假设，就可给出结果。

- 20 像 SNR 那样的测量值是

通过规定实数值转移函数 F_n ，将剩余回声谱分量与回声 $Y_n(k)$ 相联系，这样，

- 25 假设变换函数 F_n 是实数值，把（4）代入到（3），则产生这导致

在（6）中使用（4）的关系式，则产生

- 30 假设 s_n ， v_n ， y_n 是在统计上独立的，以及具有零中值，则以下的关系式保持正确。

从（8a）减去（8b）和从（6）和（7）代入，产生

对 $F_n(k)$ 求解二次方程，产生两个解， $F_n(k)=1$ 和更有意义的解：

- 5 对于 $s_n=0$ 的实验结果表明，在确定性变换中的误差是来源于推导中的以下两个假设： F 是实数值；以及 s_n, v_n 和 y_n 是在统计上独立的。实验结果表明，更精确的剩余回声估值是连同着一个在没有假设统计独立性的情况下得出的非实数值变换函数 F 而达到的。

去除 F 是实数值的假设

- 10 令 $F_n(k)=F_{r,n}(k)+jF_{i,n}(k)$ ，以及把这个定义通过 (4) 和 (5) 代入 (6) 和 (7)，并最后代入 (9)，导致二次方程的系列解。当 $F_{i,n}(k)$ 被强迫为零时，实现 (10) 的解。当 $F_{r,n}(k)$ 被强迫为零时，解为：

- 15 实验结果表明，这个选择不一定比已提出的更好。然而，重新组织该问题，导致非模糊的变换。可以把分开的变换函数用于 $R_n(k)$ 的实部和虚部以代替 (4)。

这导致对于实部和虚部的分开的解。方程 (10) 由以下形式的两个解代替：

- 20 这个方法导致实验结果的重大的改进。剩余回声估值改进到 45dB 的 SNR (没有近区讲话人的无噪声情形)。下面，在 (16) 中显示的数字的具有良好性能的公式也被使用来得到这个结果。

去除统计独立的假设

- 25 去除统计独立的假设使公式 (10) 变为下式：

当与 (13) 相组合时，这成为以下形式的两个方程

- 30 矢量项 (cross-term) 取决于两个未知项，即语音和噪声谱分量。在计算变换函数时的非确定性被分离为矢量项估值。对于包括和 $0.5[Z_n(k)+$ 的矢量项的近似，研究了几个选择。近端的失真作为在与 s_n 之间的 Itakura-Saito 失真被测量。远端抑制则使用回声反射损耗增强 (ERLE) 来测量。已经发现，把矢量项包括在内可以适当地改进在双向讲话期间的剩余回声估值。然而，近端不是没有失真的，而且

某些剩余回声仍旧是可听见的。结果（下面讨论）显示于表 1。

避免数字敏感性

公式 (7) 对于 $F_n(k)$ 的数值接近于 1 时是敏感的。以上作为替换
5 的公式也遇到这个问题。在一个实施例中，当或很小 (<10.0) 时，通过使

和设置 $R_n(k)$ 为零，这个情形可避免。这个改变大大改善了实验结果。
使用这种修正，在 (1) 中计算的剩余回声估值从约 2dB 改进到约
10 11.5dB SNR（没有近区讲话人的无噪声情形）。

表 1 对于各种矢量项的带有和不带有噪声的模拟的近端、远端
和双向讲话情形的样本结果

	仅仅 ABC	仅仅 NR	非矢量项	矢量项=	
无噪声					
ISD-DT	9.795	9.207	4.425	1.001	0.956
ISD-近端	0.000	0.073	0.381	0.089	0.154
ERLE-远端	46.4dB	44.0dB	-23.8dB	33.2dB	35.1dB
有噪声					
ISD-DT	14.523	9.182	4.321	1.142	1.081
ISD-近端	1.219	0.124	0.456	0.156	0.220
ERLE-远端	53.6dB	46.9dB	-22.7dB	-1.9dB	43.0dB

表 1 显示对于带有和不带有近端噪声的模拟的实验结果。在近端
15 单个讲话期间的语音失真作为在和 s_n 之间的 Itakura-Saito 失真 ISD
被测量。这些结果显示在标以“ISD-近端”的各行中。可以得到纯净
信号 s_n ，因为使用了室内模拟器来产生话筒输入 z_n 。Itakura-Saito
失真也用来测量双向讲话性能。这些结果显示在标以“ISD-DT”的各行中。
平均的 ERLE 用来测量在远端单个讲话期间的性能。这些结果被
20 显示在标以“ERLE-远端”的各行中。不是用话筒输入来对 ERLE 进行
归一化，而是用近端信号 s_n 来对 ERLE 实施归一化。不论在具有噪声和
没有噪声的情形下，在远端单个讲话期间，当矢量项被设置为 0 时，

5 剩余回声被很好地抑制 (-23dB ERLE)。当近端在工作时, 矢量项显得很重要。在这些情形下, 对矢量项使用估测可以得到更好的性能。在双向讲话期间, 用代替矢量项可提供非常良好的结果 (ISD 中 10-14 倍的改进)。在近端单个讲话期间, 当存在背景噪声时, 使用代替矢量项的使用给出给出非常好的结果 (ISD 中 8 倍的改进)。在这两种情形下, 剩余回声抑制是轻微的, 远端的失真是可听见的。

10 在以上表格形式中显示的结果, 现在以对于每个近端单独讲话、远端单独讲话、和双向讲话情形的分开的方程来表示。对于近端单个讲话, 剩余回声的实部和虚部由系统通过使用以下的两个方程进行估值。

对于远端单独讲话, 剩余回声的实部和虚部由系统通过使用以下的两个方程进行估值。

以及对于双向讲话, 剩余回声的实部和虚部由系统通过使用以下的两个方程进行估值。

15 如上所示的、能够计算矢量项的剩余回声估值器显示在其余的图上以及参照这些图进行描述。

20 图 2 显示剩余回声估值器的第一部分的方框图。电路 200 产生, e_n 和 z_n 的频谱分量的实部。信号, e_n 和 z_n 分别被显示在节点 114, 122 和 126 上, 它们被分别输入到窗函数 202, 204 和 206。下标“n”是方块的下标。可以使用任何适当的窗函数, 包括矩形、Hamming、或 Hann 窗。在某些实施例中, 窗函数 202, 204 和 206 被省略。

25 在某些实施例中, AEC 和 NR 是在时域中执行的。在这些实施例中, 快速傅立叶变换 (FFT) 方块 208, 210 和 212 被使用来产生信号, e_n 和 z_n 的频域表示, 以便在剩余回声估值器中使用。在其他的实施例中, AEC 和 NR 在频域中执行, 窗方块 202, 204 和 206 以及 FFT 方块 208, 210 和 212 不存在。

30 FFT 方块 208 在节点 220 产生, 以及把同一个结果提供到在节点 222 产生的平方函数 214。FFT 方块 210 产生以及平方函数 216 在节点 224 产生。FFT 方块 212 在节点 232 产生, 以及把同一个结果提供给用于在节点 232 产生的平方函数 218。加法器 226 分别接收在节点 222, 224 和 230 处的, , 以及在节点 228 产生。

电路 200 产生, e_n 和 z_n 频谱分量的实部。完整的剩余回声估值器包

括相应于电路 200 的另一个电路，它产生同一个频谱分量的虚部。

图 3 显示剩余回声估值器的第二部分的方框图。电路 300 在节点 232, 220, 228 和 222 接收信号，这些节点相应于图 2 上相同的号码的节点。电路 300 也在节点 302 接收信号，它相应于来自噪声减小电路（诸如噪声减小电路 120（图 1））的以前的输出频谱分量。加法器 304 和乘法器 303 与 305 产生，它被用作为当出现双向讲话事件时的矢量项。乘法器 308 产生，它只在有近端语音时才被使用。当存在不同的语音活动时，乘法器 306 和 312 实施在矢量项之间的切换。

在某些实施例中（软判决实施例），在节点 310 和 314 产生的矢量项以变化的幅度被产生，以及由加法器 316 相加在一起。结果，在节点 318 产生的分子包括两种类型的矢量项的分量。在软判决实施例中，乘法器 306 和 312 产生可变的输出值，它们是频谱分量 $-$ 和 $-$ 的函数，以及输出被加法器 316 相加在一起。

在其他的实施例中（硬判决实施例）， $-$ 和 $-$ 取数值“1”和“0”，以及乘法器 306 和 312 用作为开关，它们或者让输入通过，或者不让输入通过。例如，当 $-$ 是“1”时，乘法器 306 用作为开关，以及把传送到节点 310。在这种情形下，乘法器 306 用作为闭合的开关以及乘法器 312 用作为打开的开关。另外，当 $-$ 是“1”时，乘法器 312 把传送到节点 314。在这种情形下，乘法器 316 用作为闭合的开关以及乘法器 306 用作为打开的开关。下面参照图 4 讨论 $-$ 和 $-$ 的产生。

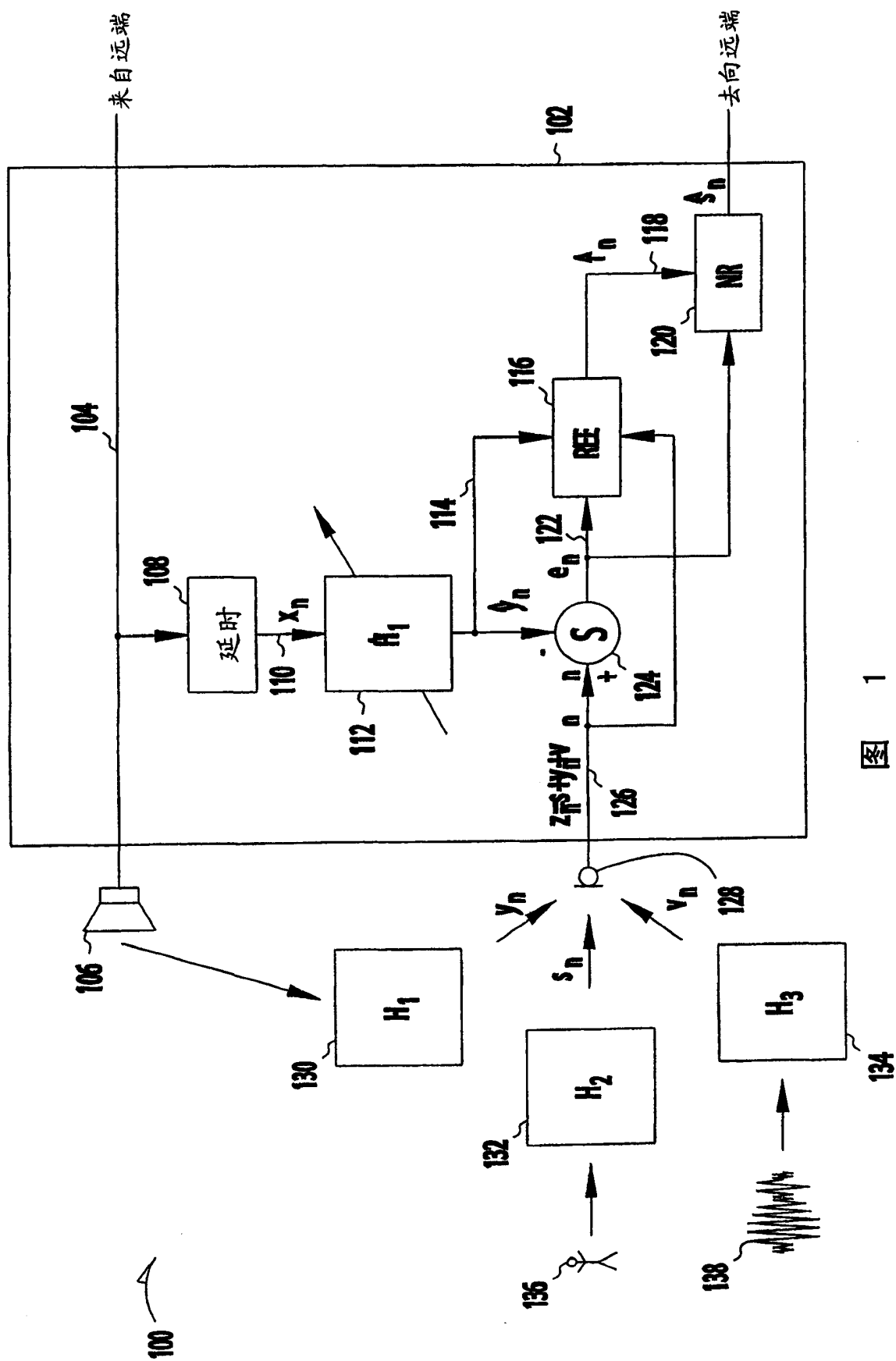
除法器 320 把节点 318 处的分子除以节点 222 处的，在节点 322 处产生。节点 322 处的代表剩余回声的谱分量的实部的平方的估值。

加法器 304 与 316，和乘法器 306, 308 与 312 形成矢量项计算器，它计算以上表 1 所示的各种矢量项。本领域技术人员将会看到，其他实施例可被使用于矢量项计算器而不背离本发明的范围。例如，电路 300 的全部或一部分可以用软件实施。软件可以在通用计算机，数字信号处理器（DSP），或其他处理器上执行。另外，电路 300 的全部或一部分可以专用硬件实施，诸如专用集成电路（ASIC）。

图 4 显示在扬声器电话中的语音活动性检测器的方框图。双向讲话检测器 402 在节点 104 接收远端音频，以及在节点 126 接收由话筒 128 获取的近端音频。在硬判定实施例中，当双向讲话在进行时，双向讲话检测器 402 把 $-$ 设置为等于 1，否则，把 $-$ 设置为等于 0。语音活动

- 性检测器 404 在节点 126 接收近端音频，以及确定在话筒处是否存在语音信号。与门 406 接收语音活动性检测器 404 的输出和在节点 408 上 β 的逻辑反。与门 406 在节点 410 产生变量 β 。在硬判定实施例中，如果没有双向讲话但有近端语音活动，则 β 设置为 1。否则 β 设置为零。
- 5 在软判决实施例中， β 和 β 的范围取 0 与 1 之间的实数值（包括 0 和 1 在内）。例如，在一个实施例中， β 取 0 与 1 之间的实数值，以及 β 由乘法器而不是由与门 406 生成。在这样的实施例中，乘法器按 $\beta = (1 - \beta)$ 来产生 β （语音活动性检测器输出）。

- 将会看到，以上的说明的意图是说明性的，而不是限制性的。在
- 10 阅读和了解以上的说明后，许多其他实施例对于本领域技术人员是显而易见的。所以，本发明的范围是参照附属的权利要求书被确定的，连同由这样的权利要求书被授权的等价物的全部范围。



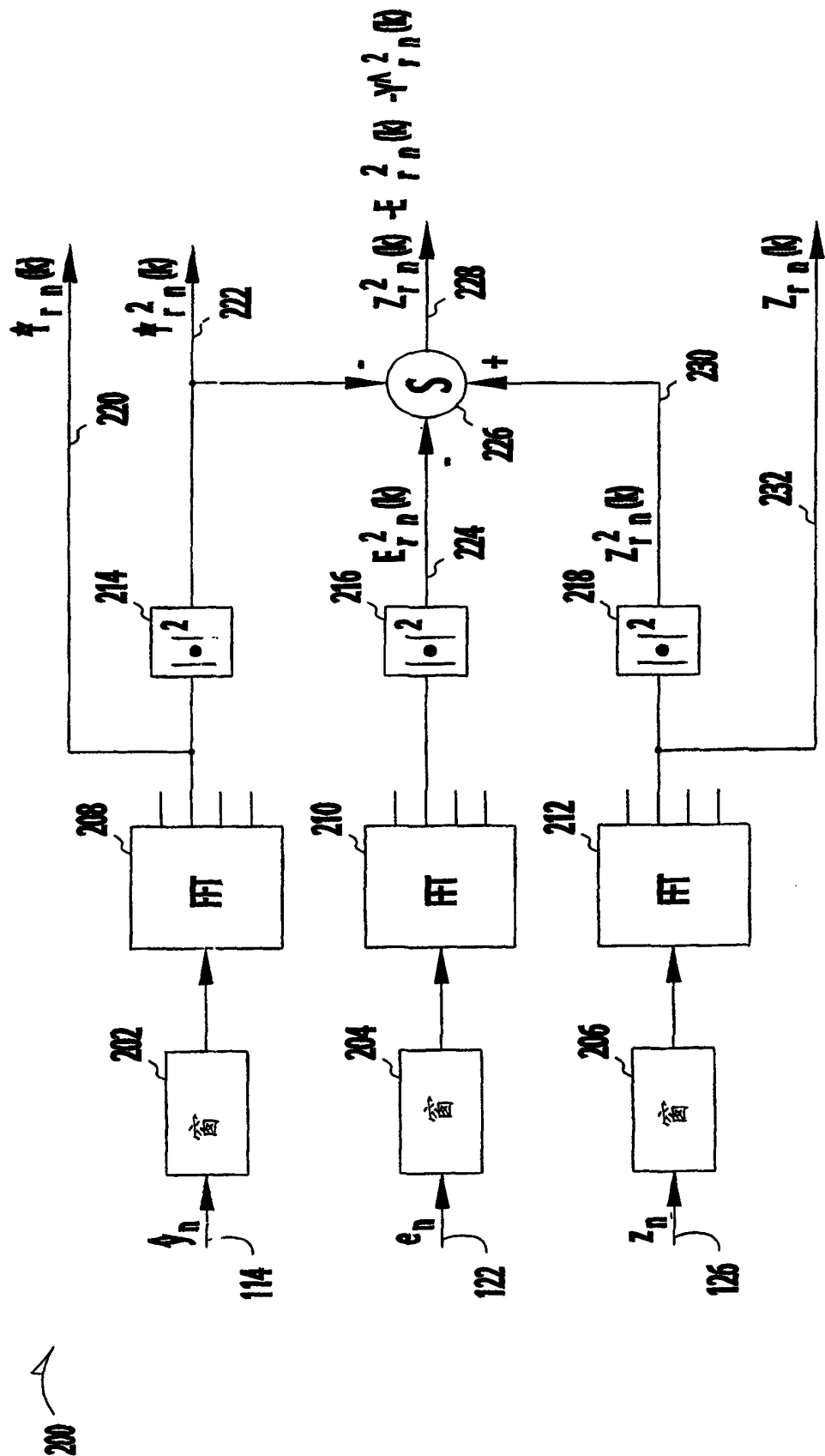


图 2

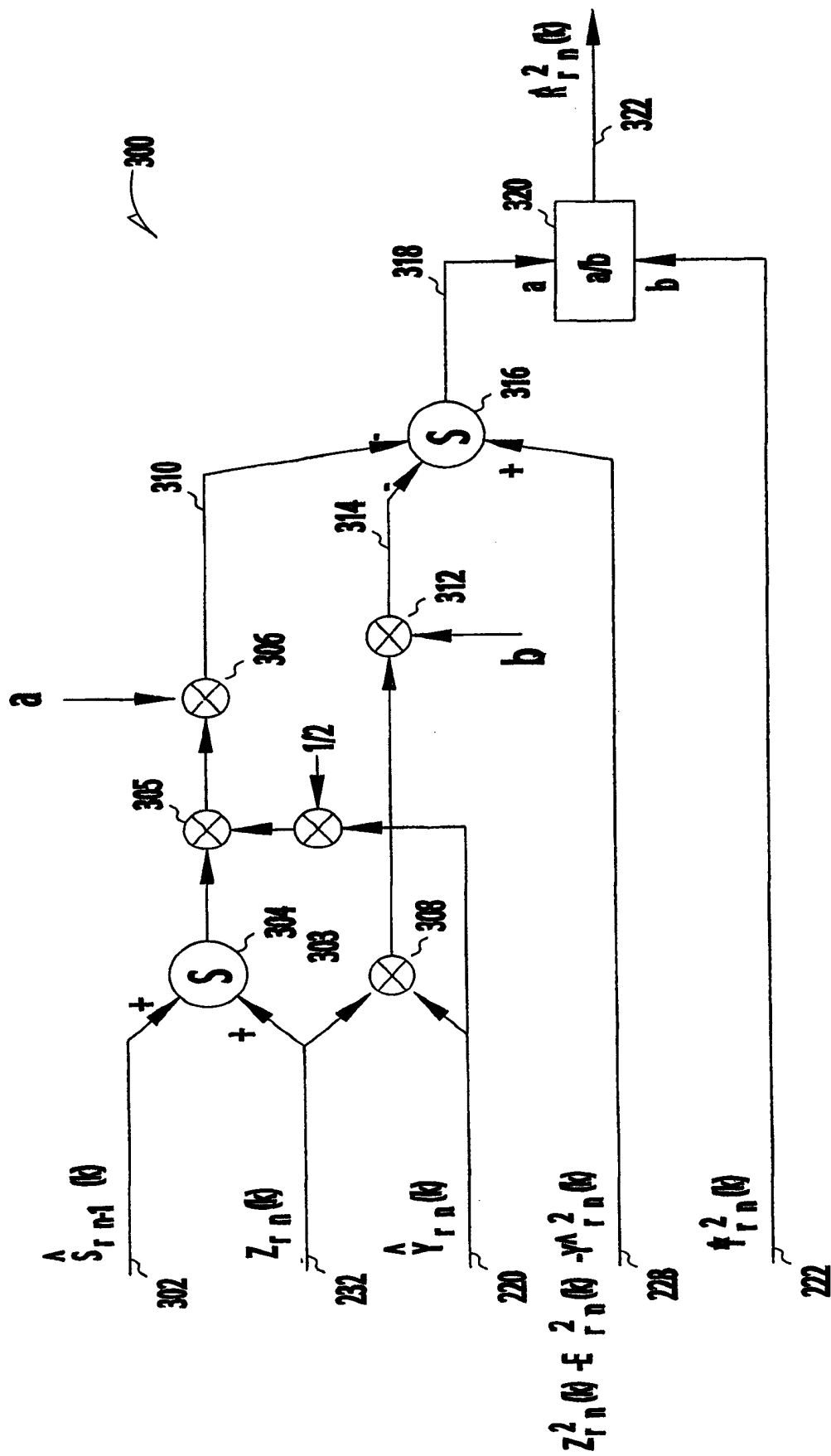


图 3

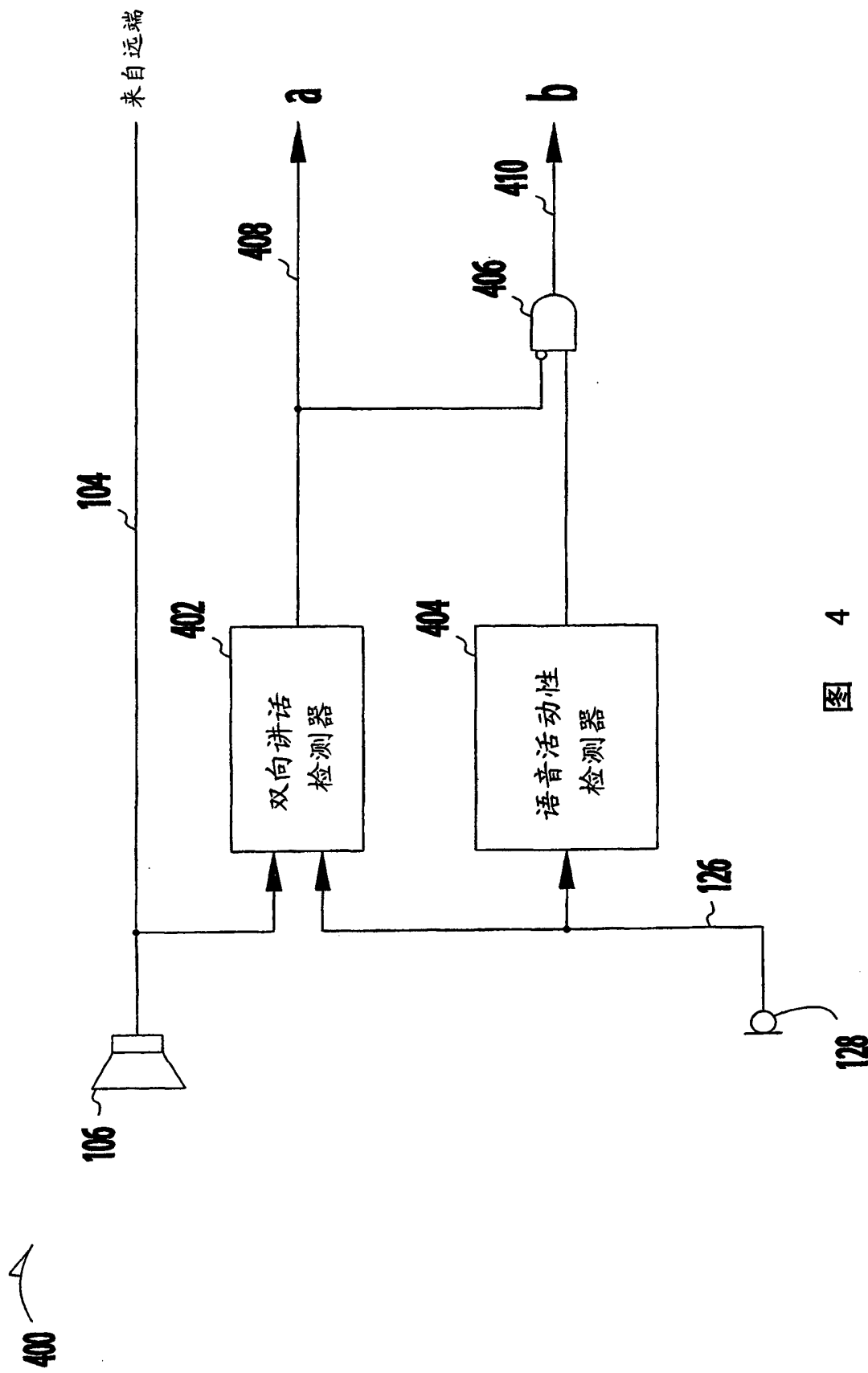


图 4