



(12) 发明专利申请

(10) 申请公布号 CN 103329566 A

(43) 申请公布日 2013. 09. 25

(21) 申请号 201080071117. 9

(51) Int. Cl.

(22) 申请日 2010. 12. 20

H04R 3/00 (2006. 01)

(85) PCT申请进入国家阶段日

H04R 3/02 (2006. 01)

2013. 07. 19

H04R 27/00 (2006. 01)

(86) PCT申请的申请数据

PCT/EP2010/070282 2010. 12. 20

(87) PCT申请的公布数据

W02011/027005 EN 2011. 03. 10

(71) 申请人 峰力公司

地址 瑞士施泰法

(72) 发明人 H·米尔德

(74) 专利代理机构 永新专利商标代理有限公司

72002

代理人 陈松涛 王英

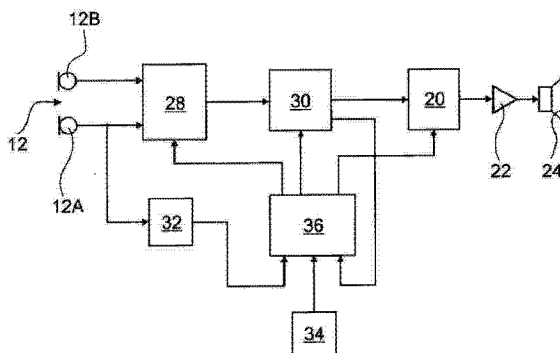
权利要求书2页 说明书4页 附图2页

(54) 发明名称

用于房间中的语音增强的方法和系统

(57) 摘要

本发明涉及一种用于房间(10)中的语音增强的系统,包括:包括至少两个间隔开的、用于捕获来自发声者的话音的音频信号的麦克风(12A, 12B)的麦克风设备(12);声学波束形成单元(28),用于处理捕获的音频信号,以便将多个不同的指向性图案中的一个提供给所述麦克风设备;反馈消除单元(30),用于对经处理的音频信号应用反馈消除算法,并且用于提供反馈状态信号,该反馈状态信号指示所述系统的声学反馈环路(12、17、20、22、24、28、30、40、46)接近反馈的程度;用于放大经处理的音频信号的装置(20、22);待设置于房间中的扬声器设备(24),其用于根据经放大的音频信号产生声音;以及用于根据所述反馈状态信号来选择提供给所述麦克风设备的所述指向性图案的装置(36)。



1. 一种用于房间(10)中的语音增强的系统,包括:

麦克风设备(12),其包括至少两个间隔开的、用于捕获来自发声者的话音的音频信号的麦克风(12A、12B),

声学波束形成单元(28),其用于处理所捕获的音频信号,以便将多个不同的指向性图案中的一个指向性图案提供给所述麦克风设备;

反馈消除单元(30),其用于对经处理的音频信号应用反馈消除算法,并且用于提供反馈状态信号,该反馈状态信号指示所述系统的声学反馈环路(12、17、20、22、24、28、30、40、46)接近反馈的程度;

用于放大经处理的音频信号的装置(20、22);

将被设置于所述房间中的扬声器设备(24),其用于根据经放大的音频信号产生声音;
以及

用于根据所述反馈状态信号来选择提供给所述麦克风设备的所述指向性图案的装置(36)。

2. 根据权利要求1所述的系统,其中,所述麦克风设备(12)是佩带式麦克风设备。

3. 根据权利要求1或2所述的系统,其中,所述麦克风设备(12)的所述麦克风(12A、12B)是全向型的。

4. 根据前述权利要求中的任意一项所述的系统,其中,所述波束形成单元(28)是延迟-相加型的。

5. 根据前述权利要求中的任意一项所述的系统,其中,所述麦克风设备(12)形成传送单元(16)的一部分或者连接到所述传送单元(16),所述传送单元(16)用于通过无线链路(17)将音频信号传送给接收单元(18),所述接收单元(18)形成所述扬声器设备(24)的一部分或者连接到所述扬声器设备(24)。

6. 根据权利要求5所述的系统,其中,所述波束形成器单元(28)形成所述传送单元(16)的一部分,并且所述反馈消除单元(30)形成所述接收单元(18)的一部分,其中所述接收单元包括用于通过所述无线链路(17)将所述反馈状态信号传送给所述传送单元的装置(44、46)。

7. 一种房间中的语音增强的方法,包括:

通过麦克风设备(12)从发声者的话音捕获音频信号,所述麦克风设备(12)包括至少两个间隔开的麦克风(12A、12B);

对所捕获的音频信号应用声学波束形成处理,以便将多个不同的指向性图案中的一个指向性图案提供给所述麦克风设备;

对经处理的音频信号应用反馈消除算法;

通过扬声器设备(24)来根据经处理的音频信号产生声音;

提供反馈状态信号,该反馈状态信号指示所述系统的声学反馈环路(12、17、20、22、24、28、30、40、46)接近反馈的程度;以及

根据所述反馈状态信号来选择提供给所述麦克风设备的所述指向性图案。

8. 根据前述权利要求中的任意一项所述的系统,其中,对将提供给所述麦克风设备(12)的所述指向性图案的选择采用了迟滞。

9. 根据前述权利要求中的任意一项所述的系统,其中,所述反馈状态信号具有指示所

述系统极为接近于所述反馈条件的第一值和指示所述系统没有极为接近于所述反馈条件的第二值,其中,当所述反馈状态信号具有所述第一值时选择第一指向性图案,当所述反馈状态信号具有所述第二值时选择第二指向性图案。

10. 根据前述权利要求中的任意一项所述的系统,其中,通过估计所述系统的所述反馈环路(12、17、20、22、24、28、30、40、46)的增益来提供所述反馈状态信号。

11. 根据权利要求9或10所述的系统,其中,当所述系统的所述反馈环路(12、17、20、22、24、28、30、40、46)的所估计的增益等于或大于预定的增益阈值时,所述反馈状态信号具有所述第一值,并且其中,当所述系统的所述反馈环路的所估计的增益低于所述增益阈值时,所述反馈状态具有所述第二值。

12. 根据权利要求9或11所述的系统,其中,对于所述第二指向性图案而言的对在水平面上冲击到所述麦克风设备(12)上的声音的灵敏度与对在垂直方向上冲击到所述麦克风设备上的声音的灵敏度之比低于对于所述第一指向性图案而言的对在水平面上冲击到所述麦克风设备(12)上的声音的灵敏度与对在垂直方向上冲击到所述麦克风设备上的声音的灵敏度之比。

13. 根据权利要求12所述的系统,其中,所述第一指向性图案是心形图案。

14. 根据权利要求13所述的系统,其中,所述心形图案的最高灵敏度的方向被取向为基本上朝向所述发声者(14)的嘴。

15. 根据权利要求12至14中的任意一项所述的系统,其中,所述第二指向性图案是双向图案。

16. 根据权利要求15所述的系统,其中,所述双向图案的最高灵敏度的方向被取向为基本上朝向所述发声者(14)的嘴。

17. 根据前述权利要求中的任意一项所述的系统,其中,布置所述麦克风设备(12),使得连接所述麦克风设备的所述两个麦克风(12A、12B)的假想线的取向基本上垂直。

18. 根据前述权利要求中的任意一项所述的系统,其中,根据所述反馈状态信号来从多个反馈消除算法中选择反馈消除算法。

19. 根据权利要求18所述的系统,其中,当所述系统的所述反馈环路(12、17、20、22、24、28、30、40、46)的所估计的增益低于阈值时,选择时域反馈消除算法,并且当所述系统的所述反馈环路的所估计的增益等于或大于所述阈值时,选择频域反馈消除算法。

20. 根据权利要求19所述的系统,其中,所述频域反馈消除算法对所述音频信号应用维纳滤波器。

21. 根据权利要求19或20所述的系统,其中,所述频域反馈消除算法估计所述系统的所述反馈环路(12、17、20、22、24、28、30、40、46)的传递函数,并且对所述音频信号应用与逆估计传递函数相对应的滤波器,以便消除由反馈引起的信号部分。

22. 根据权利要求18至21中的任意一项所述的系统,其中,对所述反馈消除算法的选择采用了迟滞。

23. 根据前述权利要求中的任意一项所述的系统,其中,所述扬声器设备(24)设置在所述房间的墙壁上或者所述扬声器设备(24)被设置成接近于所述房间的墙壁。

用于房间中的语音增强的方法和系统

[0001] 本发明涉及一种用于房间中的语音增强的系统,该系统包括:麦克风设备,该麦克风设备包括至少两个间隔开的、用于从发声者的话音(voice)捕获音频信号的麦克风;声学波束形成器单元,用于以将指向性图案(directional pattern)提供给麦克风设备的方式来处理所捕获的音频信号;反馈消除单元,用于对经处理的音频信号应用反馈消除算法;用于对经处理的音频信号进行放大的装置;以及位于房间中的扬声器设备,用于根据放大的音频信号产生声音。

[0002] 例如在 WO 2010 / 00878 A2 中描述了这样的系统。

[0003] US 2004/0170284 A1 涉及语音增强系统,其中,麦克风设置有麦克风波束形成器,并且多个扬声器设置有自适应扬声器波束形成器,其中后者能创建波束图案,该波束图案能够利用麦克风在发声者的方向上创建空值(null)以防止反馈噪声。

[0004] US4489442 涉及语音增强系统,该系统包括多个麦克风阵列,每个阵列均包括单向的前侧麦克风和单向的后侧麦克风,两个麦克风均具有心形灵敏度图案且被布置在阵列的两端。麦克风也可以具有诸如双向的或者全向的之类的其它灵敏度图案。该系统用作话音激活检测器,其中接收语音的麦克风阵列被激活,而另一个麦克风阵列被禁用。

[0005] US 2005/0254640A1 涉及语音增强系统,该系统包括多个指向性麦克风和包括回音消除单元的信号处理块。

[0006] 在如下文献中描述了包括声学波束形成的助听器,例如,US 5473701、EP 1005783B1、US 6522756、EP 1391138A2、EP 1320281 A1 和 WO 00/68703A2。

[0007] 反馈噪声是语音增强系统中的主要问题,尤其是在使用佩带式麦克风时。反馈限制了可以被应用的增益和/或限制了佩带式麦克风的用户的移动性(该佩带式麦克风可以是无线的);另外,反馈可能引起大声的令人生厌的啸声。公知的是,通过在反馈环路中应用反馈消除算法以及通过利用指向性麦克风,可以在某种程度上减少反馈问题。

[0008] 本发明的目的是提供房间中的语音增强系统,其中,将在实现高信噪比的同时减少反馈。本发明的目的也在于提供相应的速度增强方法。

[0009] 根据本发明,分别通过权利要求 1 所限定的系统以及权利要求 22 所限定的方法实现了这些目的。

[0010] 本发明的益处在于:通过根据反馈状态信号(其由反馈消除单元提供且指示系统接近声学反馈条件的程度)来选择由波束形成单元提供给麦克风设备的指向性图案,可以在较低的增益条件下(例如,当系统充分远离反馈时)最优化信噪比,例如,通过选择被优化以捕获来自发声者的嘴的语音的指向性图案,同时可以例如通过选择在扬声器设备方向上具有低灵敏度的指向性图案,而在较高增益条件下(即当系统接近反馈时)降低系统对反馈的灵敏度。

[0011] 该系统尤其适用于佩带式麦克风设备,因为佩带式麦克风特别容易反馈。

[0012] 在低增益条件下选择的指向性图案可以是心形图案,而在接近反馈的高增益条件下选择的指向性图案可以是双向图案。心形图案的最高灵敏度面向上、朝向发声者的嘴且最低灵敏度面向下,其好处在于向左和向右的头部移动并不会使由麦克风设备获得的声音

的等级过多地劣化。双向图案的好处在于,与心形图案相比在水平面上具有降低了的灵敏度;这对位于扬声器设备近区的麦克风设备尤其有用,在该近区大部分声能在水平方向上进行传送,如果扬声器设备是位于在说话者的嘴的最高处的垂直位置上的线性阵列就会出现这个情况。然而,双向图案与心形图案相比,在头部向左或者向右移动的情况下,对于由麦克风设备获得的声音等级的变化更加敏感;因此,双向图案不应用在当系统充分远离反馈时的低增益条件下。

[0013] 优选地,通过无线链路来传送由麦克风设备捕获的音频信号以允许发声者自由移动。

[0014] 以下,将参考附图介绍本发明的优选实施例,其中:

[0015] 图 1 是根据本发明的语音增强系统的示意性框图;

[0016] 图 2 是根据本发明的语音增强系统的示例的更具体的框图;

[0017] 图 3 是根据本发明的无线语音增强系统的示例的框图。

[0018] 图 1 是用于房间 10 中的语音的增强的系统的示意图。系统包括用于从发声者 14 的话音捕获音频信号的麦克风设备 12。麦克风设备 12 包括至少两个间隔开的声学传感器/麦克风 12A、12B (见图 2),用于实现声学灵敏度的指向性图案。将音频信号提供给单元 16,该单元 16 可以提供音频信号的预放大,且在无线麦克风设备的情况下包括用于建立诸如模拟 FM 链路或者优选的数字链路之类的无线音频链路 17 的发送器或收发器。音频信号通过线路或者在无线麦克风设备的情况下经由音频信号接收器 18 提供到用于处理音频信号的音频信号处理单元 20,尤其是为了对音频信号应用频谱滤波和增益控制(替代地,这样的音频信号处理或至少其部分可以在单元 16 中进行)。将经处理的音频信号提供给工作在恒定增益或自适应增益(优选地取决于环境噪声等级)下的功率放大器 22,以便将放大的音频信号提供给扬声器设备 24,从而根据经处理的音频信号来产生由听者 26 感知的放大的声音。

[0019] 图 2 示出这样的系统的更具体的示例,其中麦克风设备 12 包括两个间隔开的麦克风 12A 和 12B,所述麦克风 12A 和 12B 捕获被提供给波束形成器单元 28 的音频信号,该波束形成器单元 28 用于处理音频信号,以便将特定的指向性图案提供给麦克风设备 12。根据本发明,波束形成器单元 28 适于提供至少两个不同的指向性图案,其中,当前所应用的指向性图案根据系统的反馈状态来选择,如以下将更详细地解释的那样。将由波束形成器单元 28 处理了的音频信号提供给反馈消除单元 30,该反馈消除单元 30 用于对音频信号应用反馈消除算法,以便减少反馈噪声。反馈消除单元 30 也提供指示系统接近声学反馈条件的程度的反馈状态信号。将由反馈消除单元 30 处理了的音频信号提供给对音频信号应用频谱滤波和增益控制的音频信号处理单元 20。将由音频信号处理单元 20 处理了的音频信号提供给功率放大器 22,并且从功率放大器 22 提供给扬声器设备 24。

[0020] 根据本发明,由反馈消除单元 30 所提供的反馈状态信号来控制对由波束形成器单元 28 所提供的指向性图案的选择。反馈状态信号也可以用来在反馈消除单元 30 中根据系统的当前现行反馈状态来选择特定的反馈消除算法。

[0021] 典型地,系统也包括用于分析由麦克风设备 12 所捕获的音频信号的音频信号分析器单元 32。这样的分析器单元 32 可以包括用于确定麦克风设备 12 的用户是否当前正在发声的话音激活检测器(VAD),以及用于估计环境话音等级的环境噪声等级估计器。分析器

单元 32 的输出信号可以用来控制音频信号处理单元 20 中的音频信号处理,例如通过根据由分析器单元 32 所提供的信息来调节增益和 / 或频谱滤波。典型地,系统也包括用于使系统的用户能够提供对系统的单独调节(例如提供对所期望的增益的调节)的用户接口 34。

[0022] 典型地,系统包括用于控制系统的操作的控制器 36。特别地,控制器 36 可以接收用户接口 34、分析器单元 32 的输出信号以及来自反馈消除单元 30 的反馈状态信号,以便据此来控制波束形成器单元 28、反馈消除单元 30 以及音频信号处理单元 20 的操作。

[0023] 优选地,麦克风设备是佩带式麦克风,并且麦克风 12A、12B 优选地是全方向型麦克风。典型地,麦克风设备 12 将以如下方式布置:连接两个麦克风 12A、12B 的假想线大致垂直地取向,即麦克风设备 12 被相应地固定在发声者的衣服上。反馈状态信号可以由反馈消除单元 30 通过估计反馈环路(在图 2 的示例中,所述反馈路由麦克风设备 12、诸如单元 28、30、20 之类的处理音频信号的电子部件、功率放大器 22 以及扬声器设备 24 来形成)的增益来提供。

[0024] 例如,反馈状态信号可以具有当反馈环路的估计增益等于或大于预定总增益阈值时的第一值,和当反馈环路的估计增益低于所述总增益阈值时的第二值,其中该第一值指示系统接近于反馈,该第二值指示系统没有极为接近反馈(当反馈环路的增益为 1 时(“拉森条件”)达到反馈)。反馈环路的增益取决于使用系统的特定的环境,例如手动增益调节,使用系统的房间内的声学条件,麦克风设备 12 和扬声器设备 24 的定位和取向等。当反馈状态信号具有第一值时针对音频信号处理波束形成器单元 28 选择第一指向性图案,并且当反馈状态信号具有第二值时选择第二指向性图案。

[0025] 典型地,扬声器设备 24 由设置于房间的墙壁上或者接近于房间的墙壁设置的扬声器阵列形成。优选地,对于第二指向性图案而言比对于第一指向性图案而言的对在水平面上冲击到麦克风设备 12 上的声音的灵敏度与对在垂直方向上冲击到麦克风设备 12 上的声音的灵敏度之比更低,以便在系统接近于反馈时减少从扬声器设备 24 获得的声音。

[0026] 优选地,第一指向性图案是心形图案,其中最高灵敏度的方向的取向基本上朝向发声者 14 的嘴。第二指向性图案优选为双向图案(也称为“图形 8 图案”),其中最高灵敏度的方向的取向基本上朝向发声者的嘴。因此,当系统接近于反馈时所选择的第二指向性图案对于由扬声器设备 24 产生的声音(该声音典型地具有在水平面上有高贡献的指向性图案)具有降低了的灵敏度,由此降低反馈环路中的总增益,因为当以双向图案操作麦克风设备 12 时,麦克风设备 12 从扬声器设备 24 获得更少的声音,由此提高相对于反馈的稳定性。

[0027] 应当被理解的是,取决于系统的特定设计,可以使用其它的指向性图案。

[0028] 如以上已经提及的,麦克风设备 12 的相应的指向性图案由波束形成器单元 28 通过相应地处理来自麦克风 12A、12B 的音频信号输入来创建。例如,心形图案可以由波束形成器单元 28 的简单的延迟 - 相加设计(即,在两个麦克风信号被结合之前延迟两个信号中的一个)来创建。双向图案例如可以通过简单地两个麦克风 12A、12B 的信号相减来创建(即,在不应用延迟的情况下,通过在将一个信号乘以 -1 之后再相加)。例如,更多的用于波束形成的先进技术可以涉及空间频率或虚拟麦克风的概念,例如参见 <http://www.eurasip.org/Proceedings/Eusipco/Eusipco2000/sessions/ThuAm/P01/cr1355.pdf> 上, M. Arcienega、A. Drygajlo 和 J. Maisano 的“Robust phase shift estimation in noise for microphone arrays with virtual sensors”。例如,在如下文献中可以找到关于

声学波束形成概念的概述:在 <http://www.idiap.ch/~mccowan/arrays/tutorial.pdf> 上, I. McCowan 的“Microphone Arrays: A Tutorial”(也可以参见 I. McCowan 的博士论文“Robust Speech Recognition using Microphone Arrays”, Queensland University of Technology, Australia 2001) 以及 M. Brandstein 和 D. Ward (Eds.) 的“Microphone Arrays”, SPRINGER, 2001。

[0029] 优选地, 对由波束形成器单元 28 所提供的指向性图案的选择采用了一些迟滞(hysteresis), 即, 波束形成器单元 28 从一个指向性图案切换到另一个指向性图案时的反馈状态信号的值取决于进行切换的方向, 即, 取决于系统是否从第一图案切换到第二图案(即当发现反馈环路增加时)或者波束形成单元 28 是否从第二图案切换到第一图案(即当发现反馈环路中的增益减小时), 阈值可以是不同的。

[0030] 如以上已经提及的, 反馈消除单元 30 可以根据反馈环路中的估计增益应用不同的反馈消除算法。例如, 当反馈环路中的增益低于特定阈值时可以选择时域反馈消除算法, 当反馈环路中的增益等于或大于阈值时可以选择频域反馈消除算法。时域反馈消除的优点在于没有由于频域中的信号处理引起的音频信号的延迟。然而, 由于频域反馈消除通常更高效, 因此在反馈环路中的高增益的情况下优选采用频域反馈消除。典型地, 选择一切换将采用某种迟滞, 例如相对于反馈环路中的所估计增益的 3 分贝。频域反馈消除算法可以对音频信号应用维纳滤波器。特别地, 频域反馈消除算法可以估计反馈环路的传递函数并且对音频信号应用与逆估计(inverse estimated)传递函数相对应的滤波器, 以消除由反馈引起的信号部分。当然, 也可以使用本领域所公知的其它的反馈消除算法。

[0031] 图 3 中示出无线语音增强系统的框图, 其中麦克风设备 12 连接到传送单元 16, 该传送单元 16 包括: 波束形成器单元 28, 音频信号分析器单元 32, 用户接口 34 以及控制器 36, 其中这些元件具有与图 2 示出的系统同样的功能。另外, 传送单元 16 包括增益模型单元 38, 数字收发器 40 以及天线 42。将由波束形成器单元 28 处理了的音频信号提供给增益模型单元 38 以对音频信号应用适当的增益模型(典型地, 在相对于中间输入等级的较低和较高输入等级减小增益的增益模型)。将增益模型单元 38 的输出提供给数字收发器, 以便经由数字链路 17 将音频信号发送给接收单元 18。另外, 将分析器单元 32 的输出以及关于从用户接口 34 接收的命令/数据的控制器 36 的输出也提供给收发器 40, 以便将相应的数据/命令传送给接收单元 18。

[0032] 接收单元 18 包括天线 44 和收发器 46, 用以接收从传送单元 16 经由链路 17 所传送的音频信号及其它数据和命令。将所接收的音频信号提供给与图 2 的系统中的反馈消除单元 30 相对应的反馈消除单元 30。将由反馈消除单元 30 所提供的反馈状态信号提供给收发器 46, 以便将反馈状态信号传送给传输单元 16 的收发器 40, 该反馈状态信号被从收发器 40 提供给控制器 36, 以便使用它来控制波束形成器单元 28。将反馈消除单元 30 的输出提供给音频信号处理单元 20, 从音频信号处理单元 20 将音频信号经由功率放大器 22 提供给扬声器设备 24, 其中这些元件对应于图 2 的系统中的相应元件。将从分析器单元 32 和控制器 36 产生的、经由链路 17 所接收的数据和命令从收发器 46 提供给反馈消除单元 30 和音频信号处理单元 20。

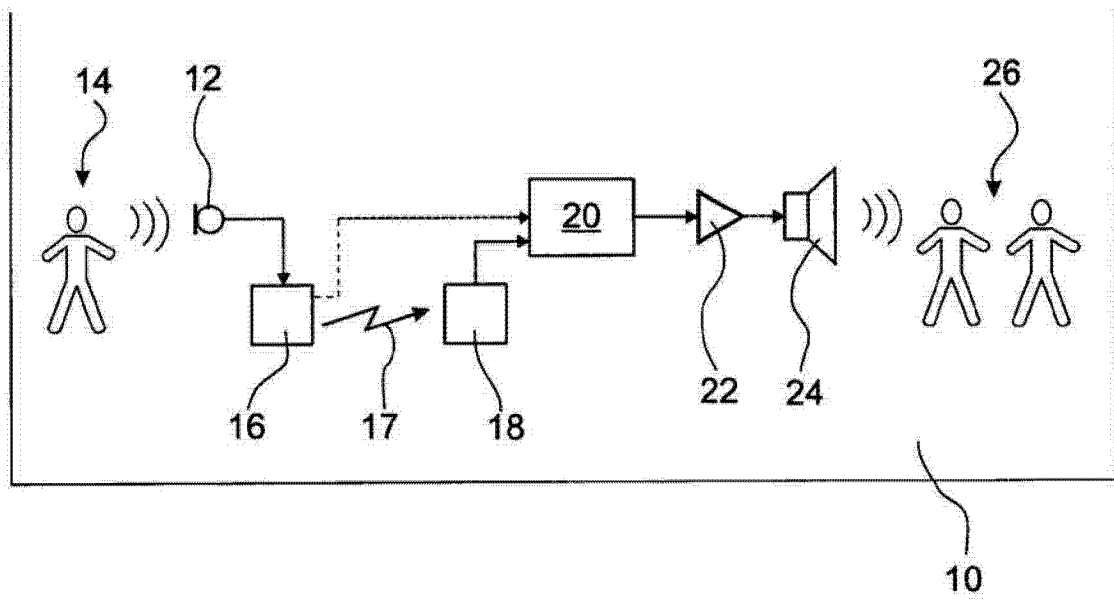


图 1

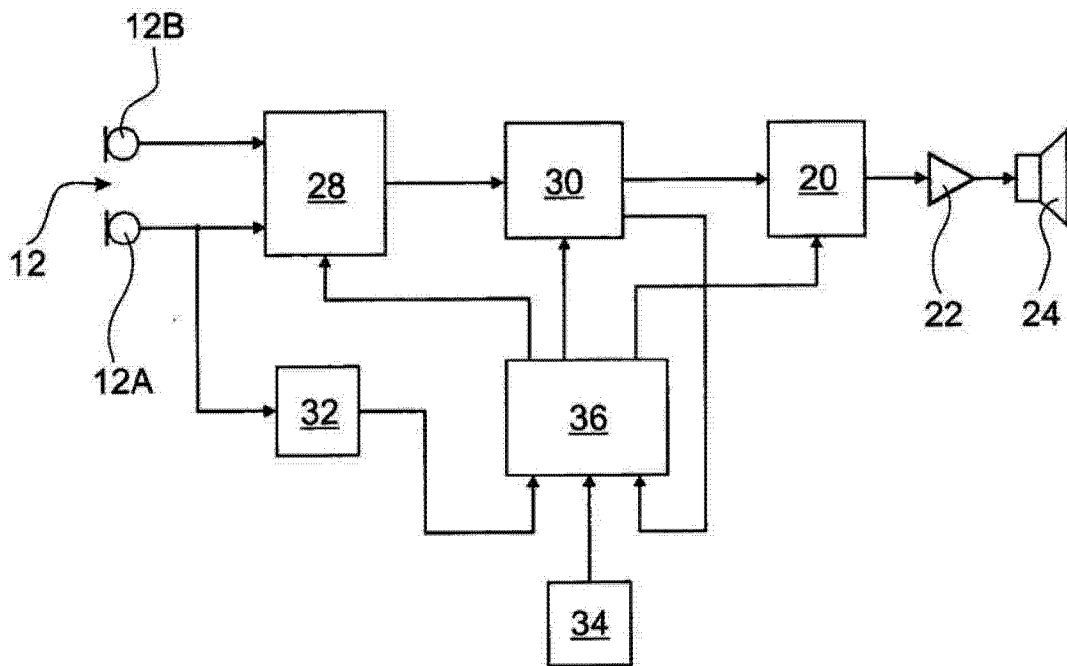


图 2

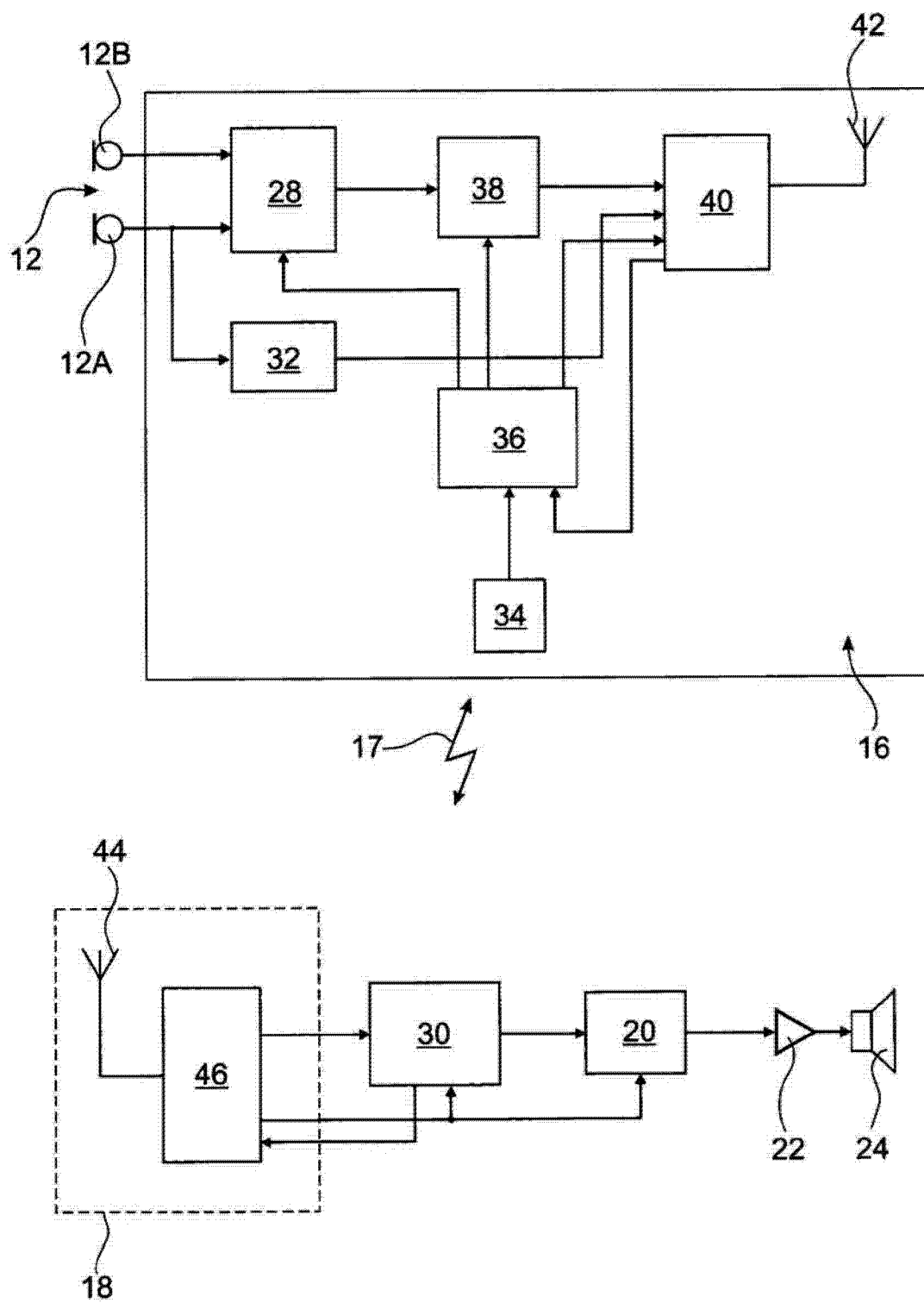


图 3