



(12) 发明专利

(10) 授权公告号 CN 109151663 B

(45) 授权公告日 2021.07.06

(21) 申请号 201810626638.4

(22) 申请日 2018.06.15

(65) 同一申请的已公布的文献号
申请公布号 CN 109151663 A

(43) 申请公布日 2019.01.04

(30) 优先权数据
17176486.3 2017.06.16 EP

(73) 专利权人 恩智浦有限公司
地址 荷兰埃因霍温高科技园区60邮编:
5656AG

(72) 发明人 安·埃尔维雷·F·斯普里耶
沃特·约斯·蒂瑞

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021
代理人 倪斌

(51) Int. Cl.
H04R 3/00 (2006.01)

(56) 对比文件

US 2006239471 A1,2006.10.26

US 2006095256 A1,2006.05.04

US 2004234079 A1,2004.11.25

EP 3089483 A1,2016.11.02

AU 2017202000 A1,2017.04.13

WO 2014145761 A3,2014.12.24

CN 103222192 A,2013.07.24

CN 101467380 A,2009.06.24

CN 1567433 A,2005.01.19

CN 101222786 A,2008.07.16

S. S. Patel, S. Gupta and D.

Ghodgaonkar.Ku-band low noise multistage
amplifier MIC design having performance
comparable to MMIC.《2014 International
Conference on Advances in Computing,
Communications and Informatics》.2014,

蔡渤.基于阵列麦克风的语音采集增强系统
设计.《襄阳职业技术学院学报》.2016,

审查员 刘畅

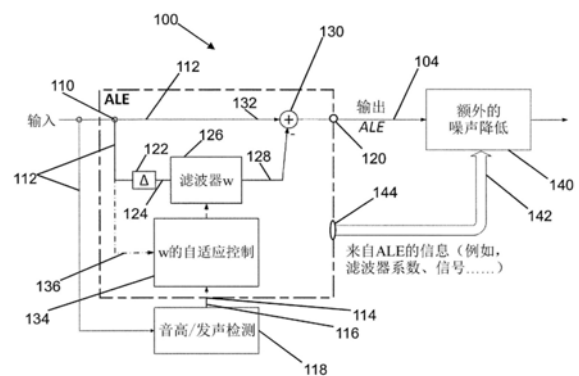
权利要求书2页 说明书13页 附图4页

(54) 发明名称

信号处理器和信号处理系统

(57) 摘要

一种信号处理器,包括:输入端,被配置成接收输入信号;发声端,被配置成接收表示所述输入信号的有声语音分量的发声信号;输出端;延迟块,被配置成接收所述输入信号并提供滤波器输入信号作为所述输入信号的延迟表示;滤波器块,被配置成:接收所述滤波器输入信号;并且通过过滤所述滤波器输入信号来提供噪声估计信号;组合器块,被配置成:接收表示所述输入信号的组合器输入信号;接收所述噪声估计信号;并且组合所述组合器输入信号与所述噪声估计信号以提供输出信号到所述输出端;以及滤波器控制块,被配置成:接收所述发声信号;接收表示所述输入信号的信令;并且根据所述发声信号和所述输入信号来设置所述滤波器块的滤波器系数。



CN 109151663 B

1. 一种信号处理器,其特征在于,包括:
 - 输入端,所述输入端被配置成接收输入信号;
 - 发声端,所述发声端被配置成接收表示所述输入信号的有声语音分量的发声信号;
 - 输出端;
 - 延迟块,所述延迟块被配置成接收所述输入信号并提供滤波器输入信号作为所述输入信号的延迟表示;
 - 滤波器块,所述滤波器块被配置成:
 - 接收所述滤波器输入信号;并且
 - 通过过滤所述滤波器输入信号来提供噪声估计信号;
 - 组合器块,所述组合器块被配置成:
 - 接收表示所述输入信号的组合器输入信号;
 - 接收所述噪声估计信号;并且
 - 组合所述组合器输入信号与所述噪声估计信号以提供输出信号到所述输出端;以及
 - 滤波器控制块,所述滤波器控制块被配置成:
 - 接收所述发声信号;
 - 接收表示所述输入信号的信令;并且
 - 根据所述发声信号和所述输入信号来设置所述滤波器块的滤波器系数;其中所述信号处理器另外被配置成提供额外的输出信号到额外的输出端,其中所述额外的输出信号表示所述滤波器系数和/或所述噪声估计信号。
2. 根据权利要求1所述的信号处理器,其特征在于,所述滤波器控制块被配置成:
 - 接收表示所述输出信号和/或滤波器输入信号的信令;并且
 - 根据所述输出信号和/或所述输入信号的延迟表示来设置所述滤波器块的所述滤波器系数。
3. 根据权利要求1或权利要求2所述的信号处理器,其特征在于,所述输入信号和所述输出信号是与离散频点有关的频域信号,并且所述滤波器系数具有复数值。
4. 根据权利要求1所述的信号处理器,其特征在于,所述发声信号表示以下各项中的一项或多项:
 - 所述输入信号的声音分量的音高的基本频率;
 - 所述输入信号的声音分量的谐波频率;以及
 - 所述输入信号包括有声语音分量的概率和/或所述有声语音分量的强度。
5. 根据权利要求1所述的信号处理器,其特征在于,所述滤波器控制块被配置成基于之前的滤波器系数、步长参数、所述输入信号、以及所述输出信号和较早的所述输入信号的延迟表示中的一者或两者来设置所述滤波器系数。
6. 根据权利要求5所述的信号处理器,其特征在于,所述滤波器控制块被配置成根据以下各项中的一项或多项来设置所述步长参数:
 - 所述输入信号的声音分量的音高的基本频率;
 - 所述输入信号的声音分量的谐波频率;
 - 表示所述输入信号的功率的输入功率;
 - 表示所述输出信号的功率的输出功率;以及

所述输入信号包括有声语音分量的概率和/或所述有声语音分量的强度。

7. 根据权利要求1所述的信号处理器,其特征在於,进一步包括:

噪声估计块,所述噪声估计块被配置成基于所述输入信号和所述输出信号来提供背景噪声估计信号;

先验信噪比估计块和/或后验信噪比估计块,所述先验信噪比估计块和/或所述后验信噪比估计块被配置成基于所述输入信号、所述输出信号和所述背景噪声估计信号来提供先验信噪比估计信号和/或后验信噪比估计信号;以及

增益块,所述增益块被配置成基于以下各项来提供增强的输出信号:(i)所述输入信号;以及(ii)所述先验信噪比估计信号和/或所述后验信噪比估计信号。

8. 根据权利要求1所述的信号处理器,其特征在於,所述输入信号是时域信号并且所述发声信号表示以下各项中的一项或多项:

所述输入信号包括有声语音分量的概率;以及

所述输入信号中所述有声语音分量的强度。

9. 一种信号处理系统,其特征在於,包括多个根据权利要求1到8中任一项所述的信号处理器,其中每个信号处理器被配置成接收是频域点信号的输入信号,并且每个频域点信号与不同的频点有关。

信号处理器和信号处理系统

技术领域

[0001] 本公开涉及信号处理器,并且具体地说(但并非必需的),涉及被配置成处理包含语音分量和噪声分量两者的信号的信号处理器。

背景技术

[0002] 背景噪声可严重地降低麦克风捕获的语音信号的质量和可懂度。因此,一些语音处理应用(例如,声音呼叫、人机交互、助听处理)结合了噪声降低处理以增强捕获的语音。单通道噪声降低法可通过实值增益函数来修改麦克风信号的幅度频谱。为了设计增益函数,可能依赖对背景噪声统计数据的估计。常见的假设可以是,噪声的振幅频谱随时间推移是平稳的。因此,单通道噪声降低法仅可抑制比较长期平稳的噪声分量。另外,由于单通道法仅应用了实值增益函数,因此并未利用相位信息。

[0003] 许多日常噪声包含确定性的周期性噪声分量。一些例子是交通噪声中的喇叭型声音以及餐厅噪声中的盘碟清洗。这些声音无法通过单通道噪声降低方案得到充分地抑制,尤其是当噪声在时长上相对较短(例如,少于几秒)时。

发明内容

[0004] 根据本公开的第一方面,提供了一种信号处理器,包括:

[0005] 输入端,所述输入端被配置成接收输入信号;

[0006] 发声端,所述发声端被配置成接收表示所述输入信号的有声语音分量的发声信号;

[0007] 输出端;

[0008] 延迟块,所述延迟块被配置成接收所述输入信号并提供滤波器输入信号作为所述输入信号的延迟表示;

[0009] 滤波器块,所述滤波器块被配置成:

[0010] 接收所述滤波器输入信号;并且

[0011] 通过过滤所述滤波器输入信号来提供噪声估计信号;

[0012] 组合器块,所述组合器块被配置成:

[0013] 接收表示所述输入信号的组合器输入信号;

[0014] 接收所述噪声估计信号;并且

[0015] 组合所述组合器输入信号与所述噪声估计信号以提供输出信号到所述输出端;以及

[0016] 滤波器控制块,所述滤波器控制块被配置成:

[0017] 接收所述发声信号;

[0018] 接收表示所述输入信号的信令;并且

[0019] 根据所述发声信号和所述输入信号来设置所述滤波器块的滤波器系数。

[0020] 在一个或多个实施例中,所述滤波器控制块可被配置成:接收表示所述输出信号

和/或延迟输入信号的信令;并且根据所述输出信号和/或所述延迟输入信号来设置所述滤波器块的所述滤波器系数。

[0021] 在一个或多个实施例中,所述输入信号和所述输出信号可以是与离散频点有关的频域信号。所述滤波器系数可具有复数值。

[0022] 在一个或多个实施例中,所述发声信号可表示以下各项中的一项或多项:所述输入信号的声音分量的音高的基本频率;所述输入信号的声音分量的谐波频率;以及所述输入信号包括有声语音分量的概率和/或所述有声语音分量的强度。

[0023] 在一个或多个实施例中,所述滤波器控制块可被配置成基于之前的滤波器系数、步长参数、所述输入信号、以及所述输出信号和所述延迟的较早输入的信号中的一者或两者来设置所述滤波器系数。

[0024] 在一个或多个实施例中,所述滤波器控制块可被配置成根据以下各项中的一项或多项来设置所述步长参数:所述输入信号的声音分量的音高的基本频率;所述输入信号的声音分量的谐波频率;表示所述输入信号的功率的输入功率;表示所述输出信号的功率的输出功率;以及所述输入信号包括有声语音分量的概率和/或所述有声语音分量的强度。

[0025] 在一个或多个实施例中,所述滤波器控制块可被配置成:根据所述发声信号来确定漏泄因数;并且通过将滤波器系数乘以所述漏泄因数来设置所述滤波器系数。

[0026] 在一个或多个实施例中,所述滤波器控制块可被配置成根据所述输入信号包括声音信号的概率的递减函数来设置所述漏泄因数。

[0027] 在一个或多个实施例中,所述滤波器控制块可被配置成基于以下各项来确定所述概率:所述输入信号的音高谐波与所述输入信号的频率之间的距离;或者所述输入信号的倒频谱峰值的高度。

[0028] 在一个或多个实施例中,本公开的信号处理器可另外包括混合块,所述混合块被配置成基于所述输入信号与所述输出信号的线性组合来提供混合输出信号。

[0029] 在一个或多个实施例中,本公开的信号处理器可另外包括:噪声估计块,所述噪声估计块被配置成基于所述输入信号和所述输出信号来提供背景噪声估计信号;先验信噪比估计块和/或后验信噪比估计块,所述先验信噪比估计块和/或所述后验信噪比估计块被配置成基于所述输入信号、所述输出信号和所述背景噪声估计信号来提供先验信噪比估计信号和/或后验信噪比估计信号;以及增益块,所述增益块被配置成基于以下各项来提供增强的输出信号:(i)所述输入信号;以及(ii)所述先验信噪比估计信号和/或所述后验信噪比估计信号。

[0030] 在一个或多个实施例中,本公开的信号处理器可另外被配置成提供额外的输出信号到额外的输出端,其中所述额外的输出信号可表示所述滤波器系数和/或所述噪声估计信号。

[0031] 在一个或多个实施例中,所述输入信号可以是时域信号并且所述发声信号可表示以下各项中的一项或多项:所述输入信号包括有声语音分量的概率;以及所述输入信号中所述有声语音分量的强度。

[0032] 在一个或多个实施例中,可提供一种系统,包括多个本公开的信号处理器,其中每个信号处理器可被配置成接收是频域点信号的输入信号,并且每个频域点信号可与不同的频点有关。

[0033] 在一个或多个实施例中,可提供一种计算机程序,所述计算机程序在计算机上运行时使所述计算机对本公开的任何信号处理器或所述系统进行配置。

[0034] 在一个或多个实施例中,可提供一种集成电路或一种电子装置,包括本公开的任何信号处理器或所述系统。

[0035] 虽然本公开可采用各种修改和替代形式,但是本公开的细节已通过举例的方式示出在附图中并且将更加详细地进行描述。然而,应理解的是,除了所描述的特定实施例之外,其它实施例也是可能的。落入所附权利要求书的精神和范围内的所有修改、等效物和替代性实施例也被涵盖。

[0036] 以上讨论并不旨在表示当前或未来权利要求组的范围内的每个示例实施例或每种实施方式。随后的附图和具体实施方式也例证了各个示例实施例。在结合附图考虑以下具体实施方式时,可以更加完整地理解各个示例实施例。

附图说明

[0037] 现在将参考附图仅通过举例的方式对一个或多个实施例进行描述,在附图中:

[0038] 图1a示出了对滤波器系数进行自适应控制的信号处理器的示例实施例;

[0039] 图1b示出了与图1a的信号处理器类似但具有额外的特征的信号处理器的示例实施例;

[0040] 图2示出了包含与图1a和图1b的那些信号处理器类似的多个信号处理器的系统的示例实施例,每个信号处理器被配置成处理与不同频点有关的信号;

[0041] 图3示出了与图2的系统类似的被配置成提供混合输出信号的系统的示例实施例;并且

[0042] 图4示出了被设计成将自适应增益函数应用于输入信号以提供增强的输出信号的系统的示例实施例。

具体实施方式

[0043] 图1a示出了可被称为语音驱动的自适应线谱增强器(ALE)的信号处理器100的框图。输入信号112由信号处理器100进行处理以生成输出信号104。信号处理器100的功能是从输入信号112除去周期性噪声分量以提供输出信号104,其中噪声分量得到抑制、但并无对输入信号112的语音分量的无用抑制。有利的是,信号处理器100可使用表示输入信号112的声音分量的发声信号116来执行发声驱动的自适应控制。在一些例子中,发声信号116可表示输入信号112的有声语音分量。随后,术语“声音分量”和“有声语音分量”可被视为是同义的。

[0044] 发声驱动的自适应控制可应用于时域信号处理器和频域信号处理器两者。对于时域中的信号处理,发声信号116可表示输入信号112的声音分量的音高的强度/振幅(或其高次谐波)或者发声信号116可表示发声的概率或强度。因此,发声的概率或强度是指输入信号112包含声音或语音信号的概率或者是指所述声音或语音信号的强度或振幅。这仅可提供为用二进制值来表示存在的语音或不存在的语音的发声指示器。对于频域中的信号处理,发声信号116还可表示输入信号112的声音分量的音高的频率。在此类例子中,声音分量的音高可以音高信号的形式提供,所述音高信号是发声信号116的例子。有利地是,音高驱

动的频域信号处理器可提供高于时域处理器的频率选择性以及因此将语音谐波与噪声分离的增强的能力。频域信号处理器因此可以为输出信号提供显著降低的噪声。

[0045] 因此,输入信号112和输出信号104可以是时域信号(在时域自适应线谱增强器的情况下)或频域信号,如表示频域中的一个或多个点/带的信号(在表示音频信号所需的各个频点/带上进行操作的子带或频域线谱增强器的情况下)。

[0046] 信号处理器100具有被配置成接收输入信号112的输入端110。信号处理器100具有被配置成接收发声信号116的发声端114。在此例子中,发声信号116由不同于信号处理器100的音高检测块118提供,但是在其它例子中,音高检测块118可与信号处理器100整合。下文关于图2另外详细地描述了音高检测块118。信号处理器100还具有用于提供输出信号104的输出端120。

[0047] 信号处理器100具有延迟块122,延迟块122可接收输入信号112 并提供滤波器输入信号124作为输入信号112的延迟表示。在一些例子中,延迟块122可被实施为线性相位滤波器。信号处理器100具有滤波器块126,滤波器块126可接收滤波器输入信号124并通过过滤滤波器输入信号124来提供噪声估计信号128。在信号处理器100被设计成处理频域信号时,滤波器系数可有利地具有复数值,从而使得可以操纵滤波器输入信号124的振幅和相位两者。

[0048] 为了避免或减少输入信号112中语音谐波的自适应或抑制,由控制块134执行的滤波器块126的自适应由音高信号116来控制(并且可选地通过发声检测进行,如下文另外描述的)。滤波器块126的发声驱动控制可以减缓由信号处理器100提供的对输入信号112的语音谐波的自适应(例如,通过操纵步长,如下文另外讨论的)并且因此有利地避免或至少减少语音衰减。

[0049] 信号处理器100具有组合器块130,组合器块130被配置成接收表示输入信号112的组合器输入信号132。在此例子中,组合器输入信号 132与输入信号112相同,但是将了解的是,在其它例子中,可执行额外的信号处理步骤从输入信号112提供组合器输入信号132。组合器块 130还被配置成接收噪声估计信号128并且将组合器输入信号132与噪声估计信号128组合以便将输出信号104提供到输出端120。在此例子中,输出信号104然后被提供到可选的额外的噪声降低块140(额外的噪声降低块140可提供额外的噪声降低,如例如,频谱噪声降低)。

[0050] 在此例子中,组合器块130被配置成从组合器输入信号132(表示输入信号112)中减去延迟输入信号的过滤版本(即,噪声估计信号128) 并且因此可除去输入信号112的与延迟版本有关的部分。

[0051] 信号处理器100具有滤波器控制块134,滤波器控制块134接收:(i)发声信号116;以及(ii)表示输入信号112的信令136。表示输入信号112的信令136可以是输入信号112。可替代地,可对输入信号112 执行某种额外的信号处理以提供表示信号136。滤波器控制块134可根据发声信号116和输入信号112来设置滤波器块126的滤波器系数,如将在下文更加详细地讨论的。

[0052] 在此例子中,信号处理器100可提供额外的输出信号142到额外的输出端144,额外的输出信号142进而被提供到额外的噪声降低块140。以此方式,额外的噪声降低块140可使用滤波器系数和/或噪声估计信号 128,滤波器系数和/或噪声估计信号128中的任一者或

两者均可由额外的输出信号142来表示。这可以使能改善额外的噪声降低块140的功能,从而允许更高效的噪声抑制。

[0053] 更一般地说,本公开的信号处理器(未示出)可具有额外的输出端,所述额外的输出端被配置成提供由滤波器块或滤波器控制块生成的作为额外的输出信号的任何信号,所述信号可有利地被任何额外的噪声降低块用于改善噪声降低性能。

[0054] 图1b示出了与图1a的信号处理器类似但具有一些额外的特征和功能的信号处理器100的框图。信号处理器100的与图1a所示的那些特征类似的特征被赋予了相同的附图标号并且可能不一定在这里另外进行讨论。

[0055] 信号处理器100具有滤波器控制块134,滤波器控制块134被配置成接收表示输出信号104的信令138和表示滤波器输入信号124的信令125。在一些例子中,表示输出信号104的信令138可以是输出信号104,并且类似地,表示滤波器输入信号124的信令125可以是滤波器输入信号。可替代地,可对输出信号104或滤波器输入信号124执行某种额外的信号处理以提供表示信号125、138。滤波器控制块134可根据输出信号104和/或滤波器输入信号124来设置滤波器块126的滤波器系数,如将在下文更加详细地讨论的。

[0056] 将了解,在其它例子(未示出)中,滤波器控制块可被配置成接收表示输入信号的信令或表示输出信号的信令。滤波器输入信号是延迟输入信号的例子,因为所述滤波器输入信号是输入信号的延迟表示。在其它例子中,滤波器控制块反而可被配置成接收延迟输入信号,所述延迟输入信号是与滤波器输入信号不同的输入信号的延迟表示,因为例如延迟输入信号具有与滤波器输入信号不同的相对于输入信号的延迟。滤波器控制块可基于延迟输入信号来设置滤波器系数。

[0057] 例如,在滤波器控制块134被配置成接收输入信号和延迟输入信号125两者时,滤波器控制块134可使用基于矩阵的处理(如通过使用最小二乘优化)来确定滤波器系数。在此情况下,可基于输入信号112和延迟输入信号125来计算滤波器系数并且不需要输出信号104。可使用对(延迟输入信号125的)自相关矩阵的估计以及延迟输入信号125与输入信号112之间的互相关矢量来计算滤波器重量。发声信号116可被滤波器控制块134用于控制自相关矩阵和互相关矢量的更新速度。

[0058] 图2示出了使用加权重叠相加框架的包括具有音高驱动的自适应控制的频域自适应线谱增强器的实施方式的系统200。将了解,根据本公开的其它系统不限于使用重叠相加框架;本公开的系统可结合重叠保留框架使用(例如,以基于重叠保留的(经分区的块)频域实施方式)。

[0059] 每个传入的输入信号212(可具有用于在不同的较早或较晚输入的信号之间进行区分的帧索引 n)被加窗并由快速傅里叶变换[FFT]块250借助于时间到频率变换(例如,使用 N 点FFT)转换成频域。这产生了频域信号 $X(k, n)$, $k=0, \dots, N-1$,其中 k 指代频率索引并且 n 指代帧索引。由于输入信号是实值信号,因此仅需要处理 $M=N/2+1$ 频点(其它频点可作为点1到点 $N/2-1$ 的复数共轭而被发现)。需要处理的每个频域信号 $X(k, n)$ 由不同的信号处理器260进行处理。在图2中,仅示出了两个信号处理器:第一信号处理器260a和第二信号处理器260b,但将了解的是,本公开的系统可具有任何数目的多个信号处理器。第二信号处理器260b的特征被赋予了与第一信号处理器260a的相应特征类似的附图标号并且可能不一定在这里另外进行描述。

[0060] 每个频率分量 k 的频域信号 $X(k, n)$ 在被由 L_k 个滤波器抽头组成的滤波器 w_k 过滤之前延迟(Δ_k)。因此,属于与第一离散频点有关的频域信号的第一输入信号262a被提供到第一延迟块264a,第一延迟块264a进而将第一滤波器输入信号265a提供到第一滤波器块264a。由于系统200中所使用的滤波器是复数值的,因此振幅和相位信息均用于减少周期性噪声分量。延迟 Δ_k 可被称为解相关参数,所述解相关参数提供了语音保存与结构化噪声抑制之间的折衷。延迟 Δ_k 不一定需要对于所有频点而言都相同。延迟越大,信号处理器260将会越不适应语音的短期相关,而结构化噪声也可能越少地被抑制。

[0061] 各个滤波器块266a、266b提供用 $Y(k, n)$ 指代的噪声估计信号,所述噪声估计信号包括在第 k 个频点对输入信号中的周期性噪声分量的估计。滤波器控制块234设置各个滤波器块266a、266b的滤波器系数,如上文中关于图1a和图1b所描述的。有利的是,滤波器控制块234可基于从音高检测块274接收的音高信号216来设置各个滤波器块266a、266b的不同滤波器系数。因此,各个信号处理器260a、260b可被配置成使用针对在被处理的特定输入信号262a、262b适当地设置的滤波器系数。

[0062] 音高检测块274接收:(i)表示来自时间到频率块250的输入信号212的时间到频率信令276;以及(ii)表示来自额外的频谱处理块272的输出信号269a、269b的频谱信令278。在其它例子(未示出)中,音高检测块274可接收输入信号212和输出信号269a、269b并且通过在时域中进行处理来检测音高。可以本领域技术人员所熟知的任何方式(如在倒频谱域中)来估计音高频率,如下文另外讨论的。

[0063] 每个信号处理器206a、206b包括用于从输入信号262a、262b中减去估计的周期性噪声分量 $Y(k, n)$ 以提供增强的频谱 $E(k, n)$, $k=0, \dots, M-1$ 的组合器268a、268b,所述频谱是输出信号269a、269b的例子。频率到时间块270将增强的频率分量 $E(k, n)$, $k=0, \dots, M-1$ 转换回时域(例如,通过重叠相加或重叠保留)。分别由时间到频率块250和频率到时间块270执行的时间到频率转换和/或频率到时间转换可与任何其它频谱处理算法(例如,最先进的单通道噪声降低)共享。

[0064] 在此例子中,在各个信号处理器260a、260b与频率到时间块270之间提供了可选的额外的频谱处理块272,以便在执行频率到时间转换之前提供对输出信号269a、269b的额外的处理。

[0065] 若干个不同的优化标准(例如,最小均方误差)和所得的更新等式(例如,基于最小二乘的方法、基于归一化最小均方[NLMS]的方法或基于递归最小二乘[RLS]的方法)可被滤波器控制块234用于更新各个频点的滤波器系数。与上文关于图1b所描述的滤波器控制块类似的滤波器控制块234接收输入信号262a、262b和输出信号269a、269b两者以计算滤波器块266a、266b的滤波器系数。为了清楚起见,图2中未示出将输入信号262a、262b和输出信号269a、269b提供到滤波器控制块234。

[0066] 下文呈现的是示例等式,所述示例等式用于更新基于NLMS的自适应的滤波器系数,从而使均方误差最小化。

[0067] 对于每个输入信号262a、262b,滤波器系数可被滤波器控制块234用以下更新递归来更新,从而结合频率相关步长参数 $\mu(k, n)$:

$$[0068] \quad w_k(n+1) = w_k(n) + \mu(k, n) E^*(k, n) x_k(n)$$

$$[0069] \quad w_k(n+1) = (1 - \lambda(k, n)) w_k(n+1)。$$

[0070] 这些等式中使用了以下定义：

$$[0071] \quad x_k(n) = [X(k, n - \Delta_k), \dots, X(k, n - \Delta_k - L_k + 1)]^T,$$

$$[0072] \quad w_k(n) = [W(k, n), \dots, W(k, n - L_k + 1)]^T,$$

$$[0073] \quad E(k, n) = X(k, n) - w_k^H(n) x_k(n).$$

[0074] 为了避免大的滤波器系数并因此限制信号处理器260a、260b对输出信号269a、269bE(k, n)的影响,此例子中使用漏泄因数 $0 < \lambda(k, n) < 1$ 来实施所谓的漏泄NLMS法。

[0075] 在一些基于NLMS的自适应中,步长 $\mu(k, n)$ 可取决于分别属于输入信号 $x_k(n)$ 262和误差信号E(k, n) 269的功率 $P_X(k, n)$ 和 $P_E(k, n)$ 中的一者或两者。在一些例子中,还可能基于对音高频点的估计 $k_{\text{音高}}$ 来适应步长 $\mu(k, n)$,所述估计可通过音高检测块274来计算,如上文所讨论的。

[0076] 以此方式适应步长的优点是,可能可以减缓滤波器系数以与语音谐波相对应的频率的自适应并且从而避免输入信号的期望语音分量的不利衰减。下文示出了可实现这一点的示例步长:

$$[0077] \quad \mu(k, n) = \left(1 - \text{概率}(\text{点}(k, n) = \text{语音谐波})\right) \frac{\mu_c(k)}{P_X(k, n) + \alpha(k)P_E(k, n) + \delta}.$$

[0078] 这里, δ 是最小常数以避免除以零, $\alpha(k)$ 控制误差功率 $P_E(k, n)$ 对步长的贡献并且 $\mu_c(k)$ 是被选择用于处理第k个频点的常数(即,独立于帧大小n)步长因素。

[0079] 第k个点包含语音信令的概率 $\text{概率}(\text{点}(k, n) = \text{语音谐波})$ 越高,滤波器系数对第k个点的自适应减少得越多。

[0080] 除了或代替音高驱动的步长,音高驱动的漏泄机制可用于朝零减小滤波器系数以用于处理语音谐波,例如:

$$[0081] \quad w_k(n+1) = (1 - \lambda(k, n, k_{\text{音高}})) w_k(n),$$

[0082] 其中较高的漏泄因数 λ 可用于语音谐波。

[0083] 时间-频率点(k, n)包含语音谐波的概率可基于如通过音高检测块274确定的对音高频率的估计 $k_{\text{音高}}$ 推导出来。可由音高检测块274执行的估计方法的示例是为了在倒频谱域中通过计算输入信号的倒频谱峰值的索引 $q_{\text{音高}}(n)$ 来确定音高频率在可能的语音音高范围内(如在大约50Hz与500Hz之间):

$$[0084] \quad k_{\text{音高}}(n) = \frac{N}{q_{\text{音高}}(n)},$$

[0085] 其中N是时间到频率分解的FFT大小。代替基于输入信号来推导出音高估计信号,音高估计还可从预先增强的输入频谱推导出来(例如,在将最先进的单通道噪声降低应用于最初的音频输入信号之后)。

[0086] 对 $\text{概率}(\text{点}(k, n) = \text{语音谐波})$ 的估计可例如使用以下表达式发现:

$$\text{概率}(\text{点}(k, n) = \text{语音谐波})$$

$$[0087] \quad \approx \text{概率}(\text{帧 } n = \text{有声}) \cdot f \left(\min_{i=1, \dots, P_n} \text{距离}(k, i * k_{\text{音高}}(n)) \right).$$

[0088] 这里, $\text{概率}(\text{帧 } n = \text{有声})$ 测得第n个帧是有声语音帧的概率并且

最小距离 $\left(k, i * k_{\text{音高}}(n)\right)$ 测得第k个频点到最近的音高谐波的距离。 P_n 等于当前帧中的

音高谐波数。映射函数f将距离映射到概率：第k个频点到最近的音高谐波的距离越大，音高谐波存在于第k个频点中的概率越低。下文示出了可能的二进制映射的例子：

$$f\left(\text{最小距离}\left(k, i * k_{\text{音高}}(n)\right)\right)_{i=1, \dots, P_n}$$

[0089]

$$= \begin{cases} 1 & \text{如果 } k \in \left[i * k_{\text{音高}}(n) - \text{偏移}(k), i * k_{\text{音高}}(n) + \text{偏移}(k) \right] \\ 0 & \text{如果 } k \notin \left[i * k_{\text{音高}}(n) - \text{偏移}(k), i * k_{\text{音高}}(n) + \text{偏移}(k) \right] \end{cases}$$

[0090] 其中(可选地频率相关的)偏移偏移(k)考虑了实际语音谐波频率与估计的语音谐波频率之间的小偏差。以此方式,如果k不大于 $i * k_{\text{音高}}$ 或不小于 $i * k_{\text{音高}}$ 超过偏移值,则函数等于1,否则,函数等于零。

[0091] 在可选的例子中,概率概率(点(k,n)=语音谐波)可通过结合当前帧有声的概率概率(帧n=有声)来改进,从而将来自其它频点的信息结合到对第k个频点的概率的计算中。

[0092] 发声概率可例如从倒频谱域中输入信号262a、262b的倒频谱峰值的高度推导出来。在一些例子中,输入信号262a、262b的所有分量可用于确定发声概率,即,可以使用时域输入信号、或频域输入信号的所有频点。漏泄因数 $\lambda(k,n)$ 可根据输入信号262a、262b包括声音信号的的概率的递减函数进行设置。

[0093] 上文中音高驱动的步长控制可减少语音谐波的自适应,而语音谐波之间的噪声的自适应仍可实现。因此,有利的是,对周期性噪声抑制与谐波语音保存之间的折衷的需求减小。

[0094] 如上文关于图1a、图1b和图2所讨论的,来自自适应线谱增强器的输出信号可用作次级或额外的频谱噪声抑制处理器的经改善的输入信号。在此类情况下,经改善的频谱噪声抑制方法可通过使用来自线谱增强器的信息(如滤波器系数的值或周期性噪声估计)来获得。

[0095] 图3示出了与图2的系统类似的系统300,在系统300中,类似的特征被赋予了类似的附图标号并且因此可能不一定在下文另外进行讨论。

[0096] 各个信号处理器360a、360b被耦接到输入乘法器380a、380b、输出乘法器382a、382b以及混合块384a、384b。输入乘法器380a、380b将输入信号362a、362b乘以乘法因数 α 以生成相乘输入信令386a、386b。输出乘法器382a、382b将输出信号369a、369b乘以乘法因数 $1-\alpha$ 以生成相乘输出信令388a、388b。各个混合块384a、384b从对应的输入乘法器380a、380b接收相乘输入信令386a、386b(表示输入信号362a、362b)。各个混合块384a、384b还从对应的输出乘法器382a、382b接收相乘输出信令388a、388b(表示输出信号369a、369b)。各个混合块384a、384b通过使对应的相乘输出信令388a、388b加上对应的相乘输入信令386a、386b来提供混合输出信号390a、390b。各个混合块384a、384b因此可基于对应的相乘输入信令386a、386b与对应的相乘输出信令388a、388b的线性组合来提供混合输出信号390a、390b。

[0097] 额外的频谱处理块372可通过处理最初的输入信号 $X(k, n)$ 362或各个信号处理器360a、360b的输出信号 $E(k, n)$ 369a、369b或者处理两者的组合(即, $\alpha X(k, n) + (1-\alpha)E(k, n)$, $\alpha \in [0, 1]$) 来执行经改善的频谱噪声抑制。在此类情况下, 可通过适当配置的混合块来提供因数 α 和 $1-\alpha$ 进行的乘法。

[0098] 图4示出了被配置成执行频谱噪声抑制法的系统400, 所述频谱噪声抑制法包括将实值频谱增益函数 $G(k, n)$ 应用于输入信号402 $X(k, n)$ 。增益函数的计算可以是基于对背景噪声的估计 $\hat{N}(k, n)$ 450以及可选地对后验信噪比(SNR)和先验SNR中的一者或两者的估计, 所述后验 SNR和所述先验SNR可分别用 $\gamma(k, n)$ 和 $\varepsilon(k, n)$ 来指代。

[0099] 图4示出了与上文关于图1a、图1b和图2所描述的信号处理器类似的信号处理器410, 信号处理器410被配置成处理输入信号402, 输入信号402在此例子中为频域信号, 所述处理可涉及最初时域音频输入信号的整个频率范围。

[0100] 信号处理器410被配置成将输出信号 $E(k, n)$ 404和噪声估计信号 $Y(k, n)$ 406提供到噪声估计块412。噪声估计块412也被配置成接收输入信号 $X(k, n)$ 402并且基于输入信号 $X(k, n)$ 402、输出信号 $E(k, n)$ 404 和可选地噪声估计信号 $Y(k, n)$ 406来提供背景噪声估计信号 $\hat{N}(k, n)$ 450。

[0101] 所述系统具有被配置成接收输入信号 $X(k, n)$ 402、输出信号 $E(k, n)$ 404和适应的背景噪声估计信号414的SNR估计块420。如下文将讨论的, 适应的背景噪声估计信号414在此例子中是以下各项的乘积: (i) 背景噪声估计信号 $\hat{N}(k, n)$ 450; 以及(ii) 过减因数信号 $\zeta(k, n)$ 456。SNR估计块420然后可基于输入信号 $X(k, n)$ 402、输出信号 $E(k, n)$ 404 和适应的背景噪声估计信号414来提供SNR信令422。SNR信令422 在此例子中表示先验SNR估计和后验SNR估计。在其它例子中, 本公开的系统可提供表示仅先验SNR估计或仅后验SNR估计的SNR信令。

[0102] 所述系统具有被配置成接收输入信号 $X(k, n)$ 402和SNR信令422 的增益块430, 所述接收在此例子中包括接收先验信噪比估计信号和后验信噪比估计信号。增益块430被配置成基于输入信号 $X(k, n)$ 402和SNR信令422来提供增强的输出信号 $X_{\text{增强}}(k, n)$ 432。

[0103] 先验信噪比和后验信噪比可使用判决引导法进行估计, 如通过以下等式例证的:

$$[0104] \quad \varepsilon(k, n) = \beta \left| \frac{G(k, n-1)X(k, n-1)}{\hat{N}(k, n-1)} \right|^2 + (1-\beta) \text{最大}(\gamma(k, n) - 1, 0)$$

$$[0105] \quad \gamma(k, n) = \left| \frac{X(k, n)}{\hat{N}(k, n)} \right|^2$$

[0106] 输入信号402 $X(k, n)$ 、噪声估计信号406 $Y(k, n)$ 和输出信号404 $E(k, n)$ 可用于生成表示周期性背景噪声分量的背景噪声估计信号442 $\hat{N}_{\text{周期性}}(k, n)$ 。这些信号还可用于改善由SNR块420执行的先验SNR计算。

[0107] 在图4所示的系统400中, 增益块430将增益函数应用于输入信号 402 $X(k, n)$ 以提供增强的输出信号 $X_{\text{增强}}(k, n)$ 432。然而, 在其它例子中, 代替将增益函数应用于输入信号402 $X(k, n)$, 增益块430可将增益函数应用于输出信号404 $E(k, n)$ 或输入信号402 $X(k, n)$ 与输出信号404 $E(k, n)$ 两者的组合, 如上文关于图3所描述的。

[0108] 在此例子中,噪声估计块412包括下文所描述的若干个子块。

[0109] 第一子块是周期性噪声估计块440,周期性噪声估计块440被配置成接收输入信号 $X(k,n)$ 402、输出信号 $E(k,n)$ 404和噪声估计信号 $Y(k,n)$ 406并且基于上述接收到的信号来提供周期性噪声估计信号442 $\hat{N}_{\text{周期性}}(k,n)$ 。

[0110] 第二子块是最先进的噪声估计块444,最先进的噪声估计块444被配置成接收输入信号 $X(k,n)$ 402并且提供最先进的噪声估计信号446。在此例子中,基于输入信号 $X(k,n)$ 402的功率或幅度频谱来确定最先进的噪声估计信号446,所述功率或幅度频谱可借助于最小跟踪来提供。最先进的噪声估计信号446表示存在于输入信号 $X(k,n)$ 402中的仅长期平稳的噪声分量。

[0111] 可基于 $Y(k,n)$ 的幅度频谱或根据以下等式通过从 $E(k,n)$ 中频谱减去 $X(k,n)$ 来估计周期性噪声估计信号442 $\hat{N}_{\text{周期性}}(k,n)$ 的幅度频谱,所述幅度频谱可用 $|\hat{N}_{\text{周期性}}(k,n)|$ 来指代:

$$[0112] \quad |\hat{N}_{\text{周期性}}(k,n)| = \text{最小} \left(1, \text{最大} (1 - |E(k,n)|/|X(k,n)|, 0) \right) |X(k,n)|。$$

[0113] 最先进的噪声估计信号446和周期性噪声估计信号 $\hat{N}_{\text{周期性}}(k,n)$ 442均被提供到最大块448。最大块448被配置成通过采取周期性噪声估计信号 $\hat{N}_{\text{周期性}}(k,n)$ 442和最先进的噪声估计信号446中较大的信号来组合这两者,以将表示较大信号的背景噪声估计信号 $\hat{N}(k,n)$ 450提供到组合器块452。

[0114] 噪声估计块412还具有过减因数块454,过减因数块454被配置成接收输入信号 $X(k,n)$ 402、输出信号 $E(k,n)$ 404和噪声估计信号 $Y(k,n)$ 406并且基于上述接收到的信号来提供过减因数信号 $\zeta(k,n)$ 456。

[0115] 在此例子中,组合器块452将背景噪声估计信号 $\hat{N}(k,n)$ 450乘以过减因数信号456 $\zeta(k,n)$ 以提供适应的背景噪声估计信号414。在检测到周期性噪声时,过减因数信号456 $\zeta(k,n)$ 被确定成使得过减因数信号 456提供较高的过减因数信号456 $\zeta(k,n)$ 以及因此增强的噪声抑制。例如,过减因数信号456 $\zeta(k,n)$ 可根据以下表达式来确定:

$$[0116] \quad \zeta(k,n) \sim \text{最小} (1, \text{最大} (1 - |E(k,n)|/|X(k,n)|, 0))$$

[0117] 在一些例子中,代替输入信号402 $X(k,n)$,输出信号404 $E(k,n)$ 可被SNR估计块420用于计算先验信噪比,这可以提供语音噪声与周期性噪声之间的经改善的辨别。

[0118] 在不使用音高驱动的自适应线谱增强器的一些系统中,自适应线谱增强器可用于生成背景噪声估计但不进行任何实际的噪声抑制。一种这样的方法利用了级联的两个时域线谱增强器。自适应线谱增强器关注通过设置适当的延迟分别除去周期性噪声或谐波语音:通过使用大延迟,主要取消了周期性噪声,而通过使用较短延迟,主要关注的是除去语音谐波。如果并无音高信息用于设置时域线谱增强器的步长控制,则与本公开的信号处理器相比,性能可能降低。例如,在使用大延迟时,较多持久性语音谐波可能衰减,而在使用短延迟时,一些周期性噪声分量也可能衰减。在此类情况下,语音谐波的保存与周期性噪声

估计和抑制之间仍可存在折衷。

[0119] 在本公开的信号处理器中,可能基于语音信息(即,音高估计)来重新计算每个短期输入信号(所述短期输入信号在时长上可以是大约10 ms)期间的步长。与其它频点相比,可更加缓慢地适应与估计的音高相对应的频点。因此,信号的语音分量可得到保护,包括在长期周期性噪声的存在下。另外,由于自适应仅在与音高谐波相对应的频点上减少,因此,仍可有效地抑制短期周期性噪声。在其它例子中,可能基于噪声的周期性而不基于有声语音的存在来控制步长。在结构化的周期性噪声存在时,此类方法可以仅更新频域信号处理器。可基于相对较长的时间段来估计周期性并且可针对每个例如3秒时长的连续块来重新计算步长。

[0120] 在本公开的信号处理器中,可使用复数值处理并且因此可利用相位信息。代替延迟到ALE的输入,延迟期望的信号。音高可用于自适应地设置线谱增强器的延迟。这可以保持重量在有声语音期间是高的且并不是为了防止ALE适应有声语音。在其它例子中,噪声抑制可能主要针对随机噪声抑制而非周期性噪声抑制。此类线谱增强器可对频谱幅度进行操作。然而,仅实值增益函数典型地用于此类方法中并且因此并未利用相位信息。

[0121] 本公开的信号处理器可包括适应周期性噪声分量但不适应语音谐波的自适应线谱增强器。因此,信号处理器的输出可由除去了(或至少抑制了)周期性噪声分量的麦克风信号组成。在其它例子中,自适应线谱增强器的目的可以通过使用等于音高周期的延迟来适应音高谐波。此类自适应线谱增强器的输出可由抑制了音高谐波的麦克风信号组成。

[0122] 在本公开的信号处理器中,可能可以根据音高来控制线谱增强器的自适应,从而使得可能可以避免/降低语音谐波的自适应并且从而提供经改善的语音信号。在其它例子中,线谱增强器的自适应不由音高控制:仅延迟可基于音高频率进行设置。

[0123] 本公开的信号处理器可包括线谱增强器,所述线谱增强器提供可用于生成对周期性噪声分量的估计(不一定是完整的背景噪声)的信号。周期性噪声估计可用于噪声抑制(即,在不考虑发声的情况下)。另外,线谱增强器的输出可用作计算先验信噪比时经改善的语音估计,如上文关于图4所讨论的。在其它例子中,在频谱减法中,线谱增强器的输出(在所述线谱增强器中除去了音高谐波)在有声语音分段期间可用于估计背景噪声。

[0124] 根据本公开的自适应线谱增强器的音高驱动的自适应提供了优势。自适应线谱增强器的音高驱动的(频率选择性的)自适应控制使能抑制周期性噪声分量,同时保存了谐波语音分量。另外,在自适应线谱增强器的频谱增益函数的设计中使用来自所述自适应线谱增强器的信息的基于ALE的频谱噪声降低法也可提供优越的性能。与其它方法相比,基于ALE的频谱噪声降低法提供了对周期性噪声分量的经改善的抑制。

[0125] 本公开的信号处理器可用于任何单或多通道语音增强方法中以用于抑制结构化的周期性噪声分量。例如,可能的应用包括用于声音呼叫的语音增强、用于自动语音识别的语音增强前端、以及助听信号处理。

[0126] 本公开的信号处理器可提供嘈杂且混响的环境中声音呼叫的经改善的语音质量和可懂度,包括针对移动和智能家居语音用户界面应用两者。可提供此类信号处理器以用于通过噪声降低、回声消除和消混响来改善移动和智能家居应用(例如,智能电视)的人机交互。

[0127] 本公开的信号处理器的重要特征是自适应线谱增强器的音高驱动的自适应。音高

驱动的自适应控制可使能抑制周期性噪声分量,同时可保存谐波语音分量。在时域线谱增强器的情况下,可基于估计的音高或发声的强度或振幅来控制自适应。对等的频域方法利用对音高频率的估计和其谐波来减缓或停止线谱增强器对语音谐波的自适应,同时保持了对不包含语音谐波的噪声频点的自适应。可使用本领域技术人员熟知的最先进的技术(例如,在时域、倒频谱域或频谱域中)来估计音高。音高估计的准确度对于使所述方法起作用来说并不重要。在有声语音期间,对连续帧的音高估计将常常重叠,而在噪声期间,估计的音高频率将跨时间变化较多。因此,将自然而然地避免对语音谐波的自适应。因此,有声/无声分类对于使所述方法起作用来说并不重要。然而,此类技术可用于进一步改进自适应。

[0128] 音高驱动的自适应线谱增强器的输出可用作到任何最先进的噪声降低方法的经改善的输入。此外,本公开示出了自适应线谱增强器信号是如何在对周期性噪声分量进行经改善的抑制的情况下可用于操纵经修改的噪声降低系统。

[0129] 自适应线谱增强器(ALE)可通过利用当前麦克风输入与其延迟版本之间的相关来抑制确定性的周期性噪声分量。由于ALE利用幅度和相位信息两者,因此与受限于实值增益处理的系统相比,可以实现对确定性的周期性噪声分量的更高抑制。然而,有声语音分量在性质上也是周期性的。额外的控制机制因此可用于在使周期性噪声衰减的同时保存目标语音。

[0130] 本公开的信号处理器通过使用音高驱动的自适应控制在没有折衷的情况下提供了结构化的周期性噪声抑制和目标语音保存两者。音高驱动的自适应减缓了线谱增强器对语音谐波的自适应。原则上,所述概念可结合时域以及子带和频域线谱增强器使用。

[0131] 与时域线谱增强器相比,频域实施方式考虑了频率选择性自适应以及因此语音谐波的保存与周期性噪声分量的抑制之间的更好的折衷。

[0132] 通过估计音高频率和其谐波而进行的频率选择性自适应可减缓对与语音谐波相对应的频率的自适应,同时保持对语音谐波之间的噪声分量的快速自适应。

[0133] 可通过利用有声/无声检测结合音高来改进频率选择性自适应控制。然而,有声/无声检测对于使所述方法起作用来说并不是必需的。在有声语音期间,连续音高估计有望跨时间变化较慢,而在噪声期间,音高估计将会变化得较快。因此,主要是对有声语音的自适应将会减缓,对噪声的自适应将不会减缓,即使是在进行了一些错误的音高检测时。最先进的音高估计器因此对于使所述方法起作用来说足够准确。

[0134] 线谱增强器的输出可用作到另一个最先进的噪声降低系统的经改善的输入。此外,线谱增强器的信号可用于经修改的噪声降低系统的设计,从而产生与其它系统相比对周期性噪声分量更好的抑制。

[0135] 以上附图中的指令和/或流程图步骤可以任何顺序执行,除非明确规定了特定顺序。而且,本领域技术人员将意识到,虽然已经讨论了一个示例指令集/方法,但是本说明书中的材料可以各种方式组合以产生其它例子并且将在本详细说明提供的上下文内进行理解。

[0136] 在一些示例实施例中,上述指令集/方法步骤被实现为被具体化为可执行指令集的功能和软件指令,所述功能和软件指令在用所述可执行指令进行编程且由其控制的计算机或机器上实现。这种指令被加载以供在处理器(如一个或多个CPU)上执行。术语“处理器”包括微处理器、微控制器、处理器模块或子系统(包括一个或多个微处理器或微控制器)、或

其它控制或计算装置。处理器可以指单个组件或多个组件。

[0137] 在其它例子中,本文中所展示的指令集/方法以及与其相关联的数据和指令存储在对应的存储装置上,所述存储装置被实现为一个或多个非暂态机器或计算机可读或计算机可用存储介质。(多个)这种计算机可读或计算机可用存储介质被视为物品(或制品)的一部分。物品或制品可以指任何经制造的单个组件或多个组件。本文中所限定的(多个)非暂态机器或计算机可用介质排除信号,但是(多个)这种介质可以能够接收和处理来自信号和/或其它暂态介质的信息。

[0138] 本说明书中讨论的材料的示例实施例可全部或部分地通过网络、计算机或基于数据的装置和/或服务来实现。这些可包括云、互联网、内联网、移动装置、台式计算机、处理器、查找表、微控制器、消费者设备、基础设施或其它使能装置和服务。如本文中且在权利要求书中所使用的,提供了以下非排他性定义。

[0139] 在一个例子中,本文所讨论的一个或多个指令或步骤是自动化的。术语“自动化”或“自动地”(及其类似变化)意指在不需要人工干预、观察、努力和/或决策的情况下使用计算机和/或机械/电气装置对设备、系统和/或过程进行的受控操作。

[0140] 将了解的是,被称为被耦接的任何组件可以直接或间接地联接或连接。在间接耦接的情况下,额外的组件可位于被称为被耦接的两个组件之间。

[0141] 在本说明书中,已经就所选择的一组细节呈现了示例实施例。然而,本领域的普通技术人员将理解,可以实践包括不同的所选择的一组这些细节的许多其它示例实施例。以下权利要求书旨在涵盖所有可能的示例实施例。

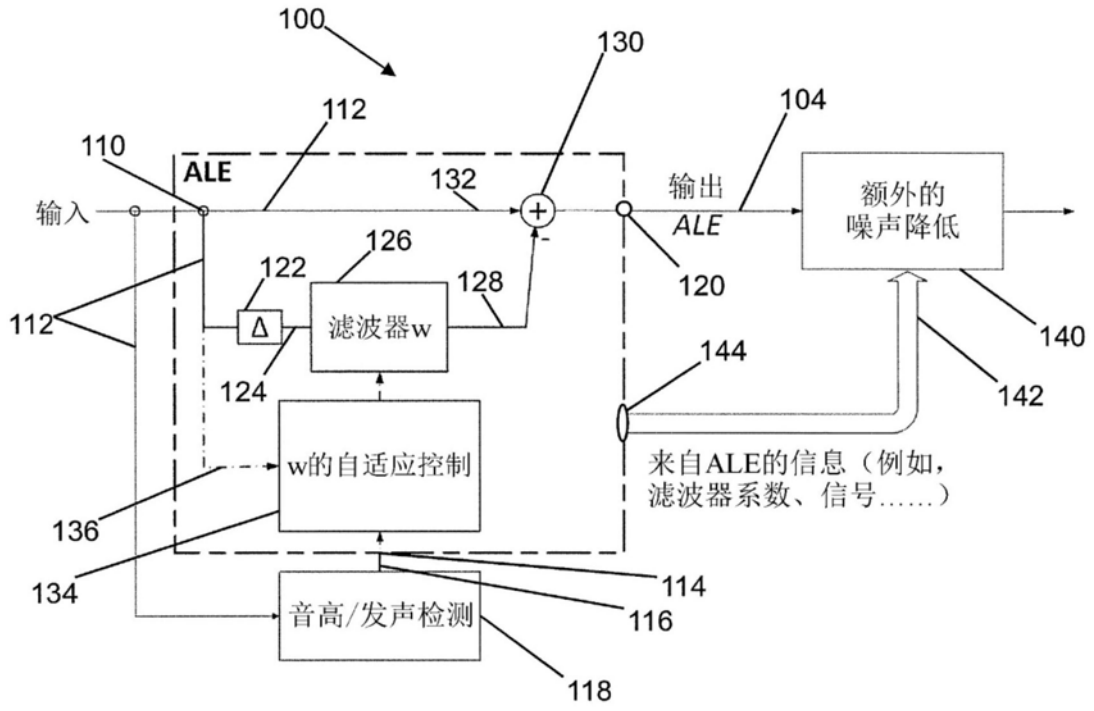


图1a

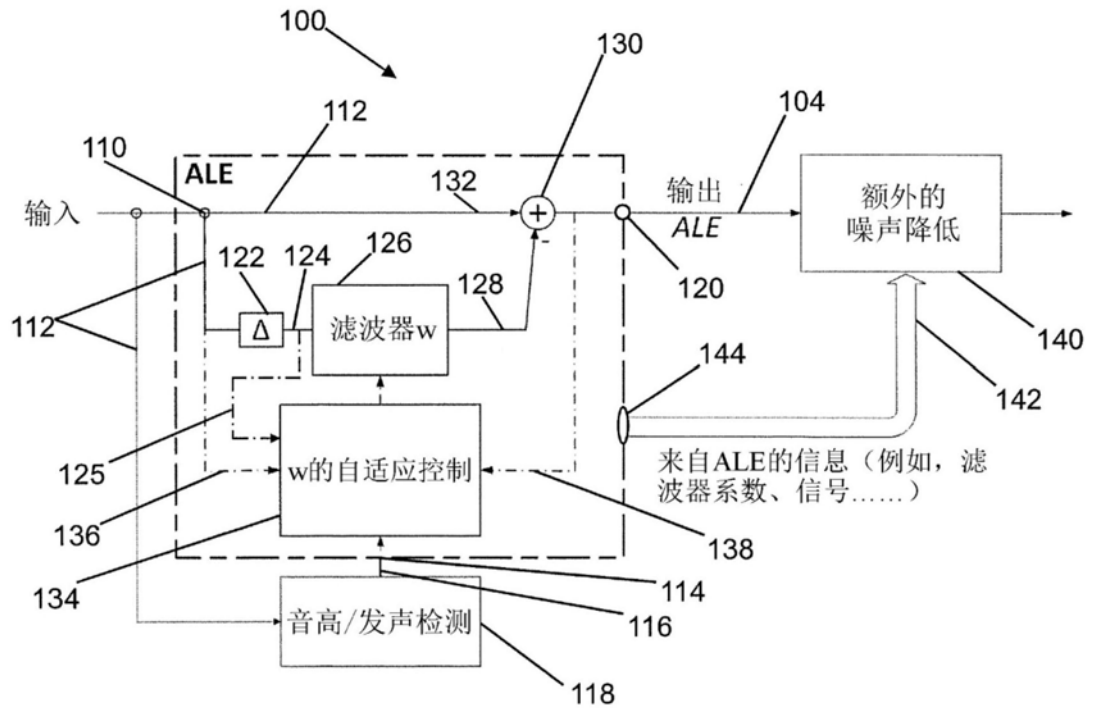


图1b

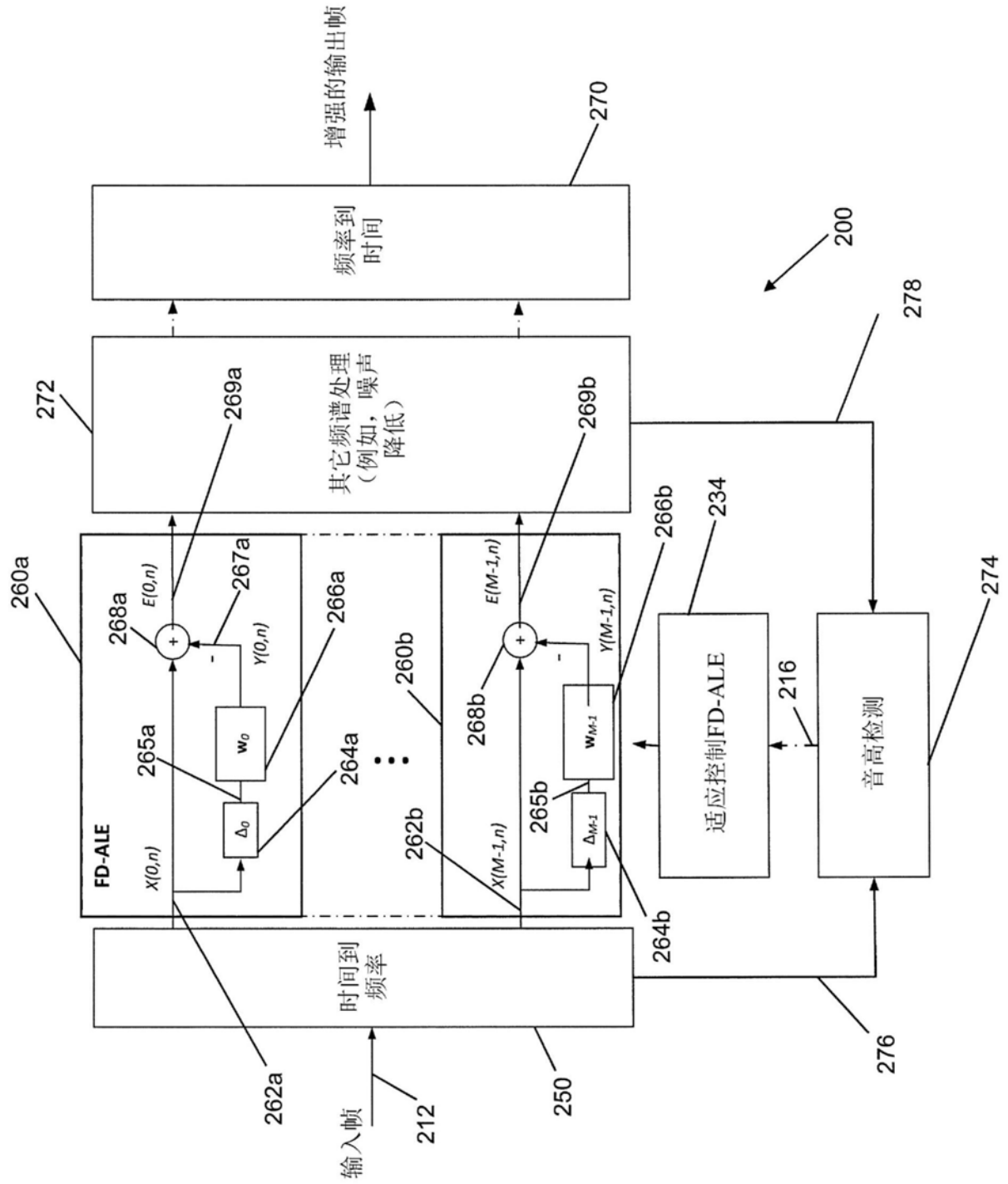


图2

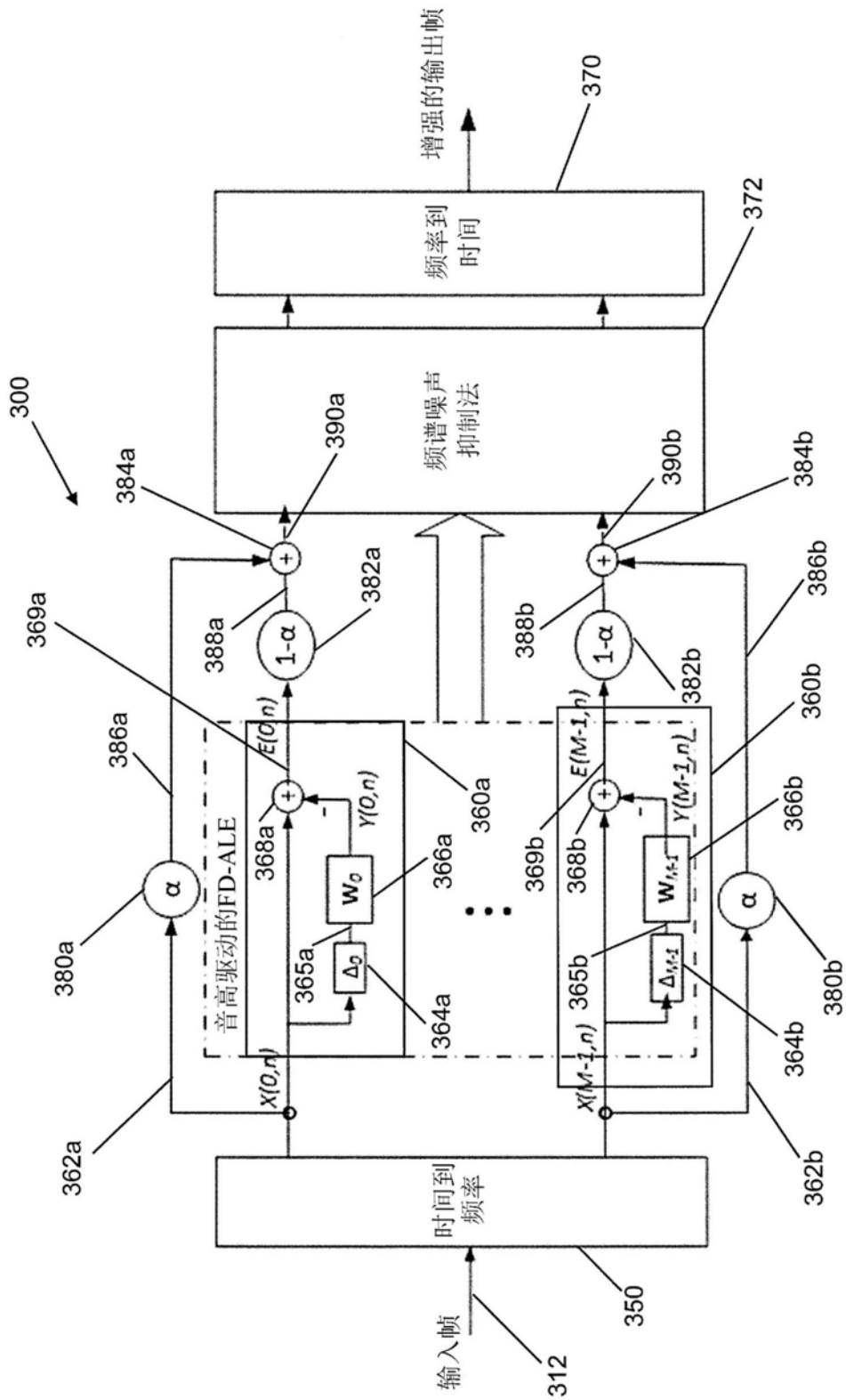


图3

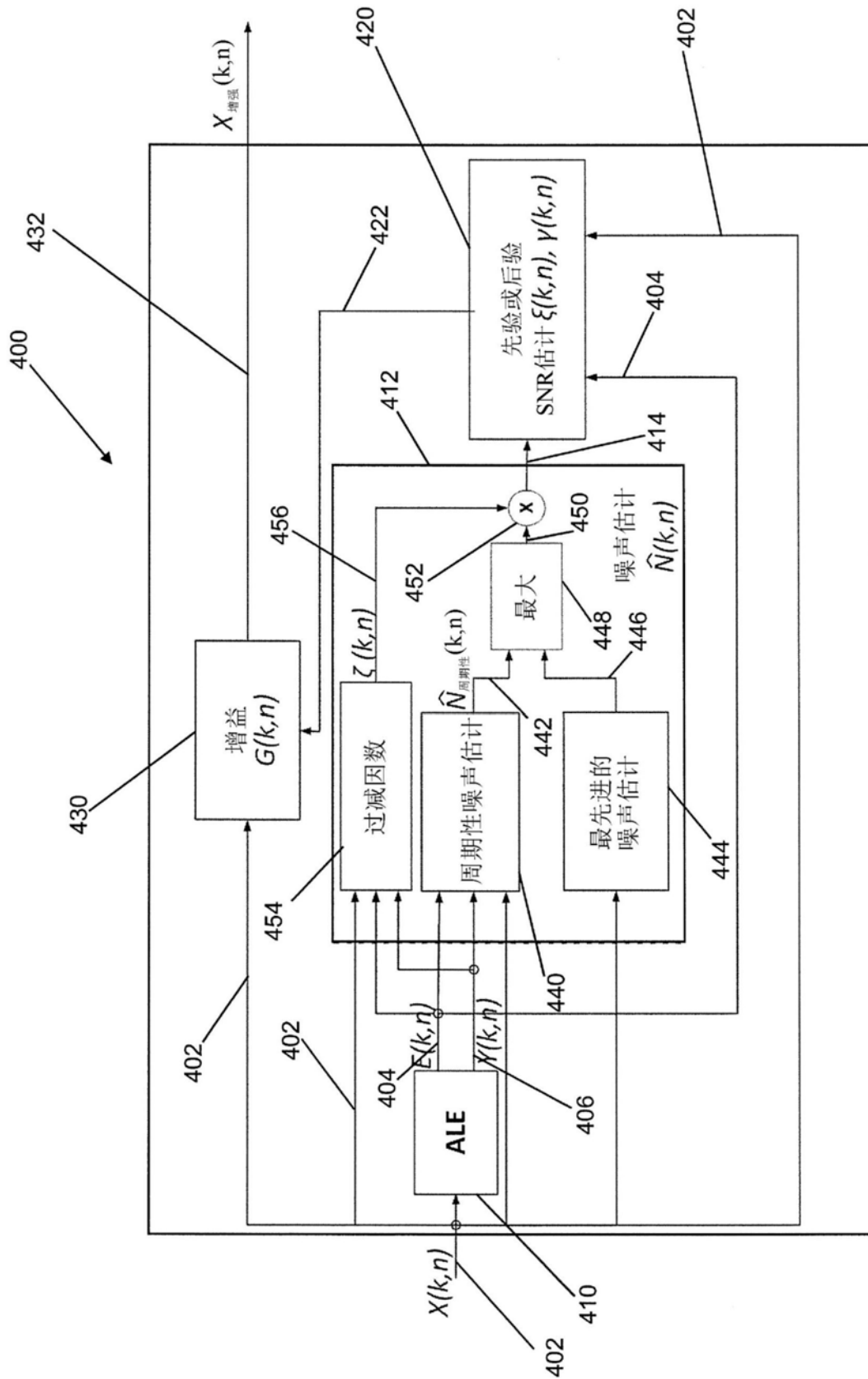


图4