



(12)发明专利

(10)授权公告号 CN 105165026 B

(45)授权公告日 2019.08.13

(21)申请号 201380073406.6

(22)申请日 2013.11.25

(65)同一申请的已公布的文献号
申请公布号 CN 105165026 A

(43)申请公布日 2015.12.16

(30)优先权数据
13169163.6 2013.05.24 EP
61/740,866 2012.12.21 US

(85)PCT国际申请进入国家阶段日
2015.08.19

(86)PCT国际申请的申请数据
PCT/EP2013/074650 2013.11.25

(87)PCT国际申请的公布数据
W02014/095250 EN 2014.06.26

(73)专利权人 弗劳恩霍夫应用研究促进协会
地址 德国慕尼黑

(72)发明人 埃马努埃尔·哈贝茨
奥利弗·蒂尔加特
塞巴斯蒂安·布劳恩
马亚·塔塞斯卡

(74)专利代理机构 北京康信知识产权代理有限
责任公司 11240

代理人 梁丽超 陈鹏

(51)Int.Cl.
H04R 3/00(2006.01)
G10K 11/34(2006.01)

(56)对比文件
WO 2005004532 A1,2005.01.13,
US 2012027219 A1,2012.02.02,
CN 1565144 A,2005.01.12,
CN 101288334 A,2008.10.15,

审查员 黄懈

权利要求书3页 说明书19页 附图14页

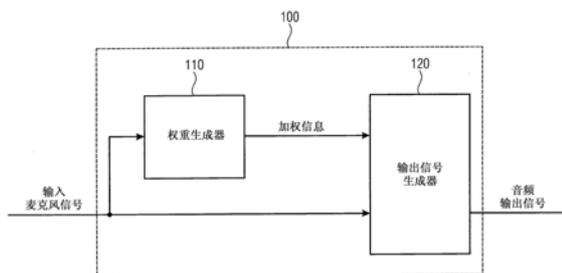
(54)发明名称

使用多个瞬时到达方向估计的知情空间滤波的滤波器及方法

(57)摘要

提供了一种滤波器(100),用于基于两个或更多个输入麦克风信号生成音频输出信号,音频输出信号包括多个音频输出信号样本。在时频域中表示音频输出信号以及两个或更多个输入麦克风信号,其中,多个音频输出信号样本中的每个被分配给多个时频段((k,n))中的时频段((k,n))。滤波器(100)包括权重生成器(110),其被适配成为多个时频段((k,n))中的每个,接收一个或多个声源的一个或多个声音分量的到达方向信息或一个或多个声源的位置信息,并且被适配为根据所述时频段((k,n))的一个或多个声源的一个或多个声音分量的到达方向信息或者根据所述时频段((k,n))的一个或多个声源的位置信息,为所述多个时频段((k,n))中的每个,生成加权信息。而且,滤波器包括输出信号生成器(120),其用于通过根据所述时频段((k,n))的加

权信息并且根据两个或更多个输入麦克风信号中的每个的被分配给所述时频段((k,n))的音频输入样本,为多个时频段((k,n))中的每个生成分配给所述时频段((k,n))的多个音频输出信号样本中的一个,来生成音频输出信号。



CN 105165026 B

1. 一种滤波器 (100), 用于基于两个或更多个输入麦克风信号生成音频输出信号, 所述音频输出信号包括多个音频输出信号样本, 其中, 在时频域中表示所述音频输出信号以及所述两个或更多个输入麦克风信号, 其中, 所述多个音频输出信号样本中的每个被分配给多个时频段中的一个时频段, 并且其中, 所述滤波器 (100) 包括:

权重生成器 (110), 被适配成: 对于所述多个时频段中的每个, 接收一个或多个声源的一个或多个声音分量的到达方向信息或所述一个或多个声源的位置信息, 并且所述权重生成器 (110) 被适配为根据所述时频段的所述一个或多个声源的所述一个或多个声音分量的所述到达方向信息或者根据所述时频段的所述一个或多个声源的所述位置信息, 为所述多个时频段中的每个生成加权信息; 其中, 所述权重生成器 (110) 被适配为根据第一噪声信息并根据第二噪声信息为所述多个时频段中的每个生成所述加权信息, 所述第一噪声信息指示关于所述两个或更多个输入麦克风信号的第一噪声分量的第一相干矩阵的信息, 所述第二噪声信息指示关于所述两个或更多个输入麦克风信号的第二噪声分量的第二相干矩阵的信息; 以及

输出信号生成器 (120), 用于通过根据所述时频段的所述加权信息并且根据所述两个或更多个输入麦克风信号中的每个的被分配给所述时频段的音频输入样本, 为所述多个时频段中的每个生成分配给所述时频段的所述多个音频输出信号样本中的一个, 来生成所述音频输出信号。

2. 根据权利要求1所述的滤波器 (100), 其中, 所述权重生成器 (110) 被配置为通过利用统计信息来生成所述第一噪声信息, 并且其中, 所述权重生成器 (110) 被配置为在不利用所述统计信息的情况下生成所述第二噪声信息, 其中, 预定义所述统计信息。

3. 根据权利要求1所述的滤波器 (100), 其中, 所述权重生成器 (110) 被适配为根据以下公式为所述多个时频段中的每个生成所述加权信息:

$$\mathbf{w}_{nd} = \Phi_u^{-1} \mathbf{A} [\mathbf{A}^H \Phi_u^{-1} \mathbf{A}]^{-1} \mathbf{g},$$

其中, $\Phi_u = \Phi_d + \Phi_n$,

其中, Φ_d 是所述两个或更多个输入麦克风信号的所述第一噪声分量的第一功率谱密度矩阵,

其中, Φ_n 是所述两个或更多个输入麦克风信号的所述第二噪声分量的第二功率谱密度矩阵,

其中, \mathbf{A} 表示所述到达方向信息,

其中, \mathbf{w}_{nd} 是表示所述加权信息的向量,

其中, \mathbf{g} 表示增益,

其中, $G[k | \varphi_1(k, n)]$ 是根据所述到达方向信息的第一实值或复值的预定义的方向性函数, 并且

其中, $G[k | \varphi_l(k, n)]$ 是根据所述到达方向信息的其他实值或复值的预定义的方向性函数,

其中, k 表示频率, 并且其中, n 表示时间。

4. 根据权利要求1所述的滤波器 (100), 其中, 所述权重生成器 (110) 被配置为根据所述

两个或更多个输入麦克风信号的所述第一噪声分量的至少一些之间的一个或多个相干性，来确定所述第一噪声信息，其中，预定义所述一个或多个相干性。

5. 根据权利要求1所述的滤波器(100)，其中，所述权重生成器(110)被配置为根据相干矩阵 $\Gamma_d(k)$ 来确定所述第一噪声信息，所述相干矩阵 $\Gamma_d(k)$ 指示由于所述两个或更多个输入麦克风信号的所述第一噪声分量引起的相干性，其中，预定义所述相干矩阵 $\Gamma_d(k)$ ，其中， k 表示频率。

6. 根据权利要求5所述的滤波器(100)，其中，所述权重生成器(110)被配置为根据以下公式来确定所述第一噪声信息：

$$\Phi_d(k, n) = \phi_d(k, n) \Gamma_d(k),$$

其中， $\Gamma_d(k)$ 是所述相干矩阵，其中，预定义所述相干矩阵，

其中， $\Phi_d(k, n)$ 是所述第一噪声信息，并且

其中， $\phi_d(k, n)$ 是所述两个或更多个输入麦克风信号的所述第一噪声分量的预期功率，其中， k 表示频率，并且其中， n 表示时间。

7. 根据权利要求1所述的滤波器(100)，其中，所述权重生成器(110)被配置为根据所述第二噪声信息并且根据所述到达方向信息，来确定所述第一噪声信息。

8. 根据权利要求1所述的滤波器(100)，

其中，所述权重生成器(110)被配置为生成所述加权信息作为第一加权信息 w_Ψ ，并且

其中，所述权重生成器(110)被配置为通过确定第二加权信息来生成所述第一加权信息，

其中，所述权重生成器(110)被配置为通过应用以下公式来生成所述第一加权信息 w_Ψ

$$w_\Psi = \arg \min_w w^H w$$

以便满足以下公式：

$$w^H a[k | \varphi_l(k, n)] = 0,$$

其中， $\varphi_l(k, n)$ 表示所述到达方向信息，

其中， $a[k | \varphi_l(k, n)]$ 表示传播向量，并且

其中， w 表示所述第二加权信息，

其中， k 表示频率，并且其中， n 表示时间。

9. 根据权利要求8所述的滤波器(100)，其中，所述权重生成器(110)被配置为根据所述第二加权信息并且根据所述两个或更多个输入麦克风信号来生成扩散噪声信息或扩散分量的功率，以确定所述第一加权信息。

10. 根据权利要求1所述的滤波器(100)，其中，所述权重生成器(110)被配置为通过应用参数维纳滤波器来确定所述加权信息，其中，所述参数维纳滤波器取决于关于所述两个或更多个输入麦克风信号的信号分量的统计信息，并且其中，所述参数维纳滤波器取决于关于所述两个或更多个输入麦克风信号的噪声分量的统计信息。

11. 根据权利要求1所述的滤波器(100)，其中，所述权重生成器(110)被配置为根据指示一个或多个平面波的到达方向的所述到达方向信息，来确定所述加权信息。

12. 根据权利要求1所述的滤波器(100)，

其中,所述权重生成器(110)包括用于提供预定义的传递函数的传递函数选择模块(104),并且

其中,所述权重生成器(110)被配置为根据所述到达方向信息并且根据所述预定义的传递函数,来生成所述加权信息。

13. 根据权利要求12所述的滤波器(100),其中,所述传递函数选择模块(104)被配置为提供所述预定义的传递函数,以便所述预定义的传递函数根据所述到达方向信息指示任意拾取图案,以便所述预定义的传递函数根据所述到达方向信息指示扬声器增益,或者以便所述预定义的传递函数根据所述到达方向信息指示头相关传递函数。

14. 一种用于基于两个或更多个输入麦克风信号生成音频输出信号的方法,所述音频输出信号包括多个音频输出信号样本,其中,在时频域中表示所述音频输出信号以及所述两个或更多个输入麦克风信号,其中,所述多个音频输出信号样本中的每个被分配给多个时频段中的一个时频段,并且其中,所述方法包括:

为所述多个时频段中的每个,接收一个或多个声源的一个或多个声音分量的到达方向信息或所述一个或多个声源的位置信息,

根据所述时频段的所述一个或多个声源的所述一个或多个声音分量的所述到达方向信息或者根据所述时频段的所述一个或多个声源的所述位置信息,为所述多个时频段中的每个生成加权信息,其中,根据第一噪声信息并根据第二噪声信息为所述多个时频段中的每个生成所述加权信息,所述第一噪声信息指示关于所述两个或更多个输入麦克风信号的第一噪声分量的第一相干矩阵的信息,所述第二噪声信息指示关于所述两个或更多个输入麦克风信号的第二噪声分量的第二相干矩阵的信息;并且

通过根据所述时频段的所述加权信息并且根据所述两个或更多个输入麦克风信号中的每个的被分配给所述时频段的音频输入样本,为所述多个时频段中的每个生成分配给所述时频段的所述多个音频输出信号样本中的一个,来生成所述音频输出信号。

15. 一种计算机可读存储介质,存储有计算机可执行代码,用于当在计算机或信号处理器上执行时实施根据权利要求14所述的方法。

使用多个瞬时到达方向估计的知情空间滤波的滤波器及方法

技术领域

[0001] 本发明涉及音频信号处理,并且尤其涉及使用多个瞬时到达方向估计的知情空间滤波的滤波器及方法。

背景技术

[0002] 在现代通信系统中通常发现在噪声和混响条件下提取声源。在过去的四十年,提出了大量空间滤波技术,以完成这个任务。在观察的信号符合信号模型时并且在计算滤波器所需要的信息精确时,现有空间滤波器是最佳的。然而,实际上,通常违背信号模型,并且估计所需要的信息是一项重大的挑战。

[0003] 现有空间滤波器可广义地分成线性空间滤波器(例如,参照[1,2,3,4])和参数空间滤波器(parametric spatial filter)(例如,参照[5,6,7,8])。通常,线性空间滤波器需要估计一个或多个传播向量或者期望的一个或多个源的二阶统计量(SOS)加上干扰的SOS。一些空间滤波器被设计为提取混响或去混响的单个源信号(例如,参照[9,10,11,12,13,14,15,16]),而其他空间滤波器被设计为提取两个或更多个混响源信号的和(例如,参照[17,18])。上述方法需要一个或多个期望源的方向的先验知识或仅仅期望源单独地或者同时激活的周期。

[0004] 这些方法的缺点在于,不能足够快速地适应新情况,例如,在期望源激活时变激活的源移动或竞争扬声器。参数空间滤波器通常基于比较简单的信号模型,例如,在时频域内接收的信号由单个平面波加上扩散声构成,并且根据模型参数的瞬时估计,计算参数空间滤波器。参数空间滤波器的优点是高度灵活的方向响应、扩散声音和干涉源的较强抑制、以及快速适合于新情况的能力。然而,如在[19]中所示,实际上可以容易地违背基本单平面波信号模型,这显著降低参数空间滤波器的性能。应注意的是,现有技术参数空间滤波器使用所有可用的麦克风信号,来估计模型参数,而仅单个麦克风信号和实值增益用于计算最终的输出信号。用于使多个可用麦克风信号相结合以便找出增强的输出信号的扩展不直接(straight forward)。

[0005] 因此,如果提供用于获得对声源的期望空间响应的改进概念,则被高度支持。

发明内容

[0006] 因此,本发明的目标在于,提供用于提取声源的改进概念。由根据权利要求1所述的滤波器、根据权利要求17所述的方法以及由根据权利要求18所述的计算机程序,解决本发明的目标。

[0007] 提供了一种滤波器,该滤波器用于基于两个或更多个输入麦克风信号生成音频输出信号(其包括多个音频输出信号样本)。在时频域中表示音频输出信号和两个或更多个输入麦克风信号,其中,多个音频输出信号样本中的每个被分配给多个时频段(time-frequency bin)中的一个时频段。

[0008] 该滤波器包括权重生成器,权重生成器被适配成为多个时频段中的每个,接收一

个或多个声源的一个或多个声音分量的到达方向信息或一个或多个声源的位置信息,并且被适配为根据所述时频段的一个或多个声源的一个或多个声音分量的到达方向信息或者根据所述时频段的一个或多个声源的位置信息,为多个时频段中的每个生成加权信息。

[0009] 而且,该滤波器包括输出信号生成器,输出信号生成器用于通过为多个时频段中的每个,生成多个音频输出信号样本中的分配给所述时频段的一个音频输出信号样本,来根据所述时频段的加权信息并且根据分配给所述时频段的音频输入样本,生成两个或更多个输入麦克风信号中的每个的音频输出信号。

[0010] 实施方式提供了一种空间滤波器,用于获得至多L个同时激活的声源的期望的响应。通过最小化经受L个线性约束的在滤波器的输出处的扩散加噪声功率,获得所提供的空间滤波器。与现有技术概念相反,L个约束基于瞬时窄带到达方向估计。此外,提供用于扩散噪声比/扩散功率的新型估计器,其呈现足够高的时间和光谱分辨率,以实现去混响和噪声降低。

[0011] 根据一些实施方式,提供了用于获得在每个时间-频率瞬间同时激活的至多L个声源的期望的任意空间响应的概念。为此,关于声音场景的瞬时参数信息 (IPI) 结合到空间滤波器的设计内,产生“知情空间滤波器 (informed spatial filter)”。

[0012] 在一些实施方式中,例如,这种知情空间滤波器基于复杂权重 (complex weight) 使所有可用的麦克风信号组合,以提供增强的输出信号。

[0013] 根据实施方式,例如,知情空间滤波器可实现为线性约束最小方差 (LCMV) 空间滤波器或者参数多通道维纳滤波器。

[0014] 在一些实施方式中,例如,通过最小化经受L个线性约束的扩散加自噪声功率,获得所提供的知情空间滤波器。

[0015] 在一些实施方式中,相对于现有技术,L个约束基于瞬时到达方向 (DOA) 估计,并且所产生的对L个DOA的响应与特定的期望方向性对应。

[0016] 而且,提供用于所需要的信号和噪声统计 (例如,扩散噪声比 (DNR)) 的新型估计器,其呈现足够高的时间和光谱分辨率,例如,以减少混响和噪声。

[0017] 而且,提供了一种用于基于两个或更多个输入麦克风信号生成音频输出信号的方法,所述音频输出信号包括多个音频输出信号样本。在时频域中表示音频输出信号和两个或更多个输入麦克风信号,其中,多个音频输出信号样本中的每个被分配给多个时频段中的一个时频段。该方法包括:

[0018] -为多个时频段 $((k, n))$ 中的每个,接收一个或多个声源的一个或多个声音分量的到达方向信息或一个或多个声源的位置信息,

[0019] -根据所述时频段的一个或多个声源的一个或多个声音分量的到达方向信息或者根据所述时频段的一个或多个声源的位置信息,为多个时频段中的每个生成加权信息,并且

[0020] -通过为多个时频段 $((k, n))$ 中的每个,生成分配给所述时频段 $((k, n))$ 的多个音频输出信号样本中的一个,来根据所述时频段 $((k, n))$ 的加权信息并且根据分配给所述时频段 $((k, n))$ 的音频输入样本,生成两个或更多个输入麦克风信号中的每个的音频输出信号。

[0021] 而且,提供了一种计算机程序,用于在计算机或信号处理器上执行时,实施上述方

法。

附图说明

[0022] 在下文中,参照附图,更详细地描述本发明的实施方式,其中:

[0023] 图1a示出了根据实施方式的滤波器;

[0024] 图1b示出了根据实施方式的滤波器的一个可能的应用场景;

[0025] 图2示出了根据实施方式的滤波器以及多个麦克风;

[0026] 图3示出了根据实施方式的权重生成器;

[0027] 图4示出了根据实施方式的两个实例响应的幅度;

[0028] 图5示出了根据实施线性约束最小方差方法的另一个实施方式的权重生成器;

[0029] 图6示出了根据实施参数多通道维纳滤波器方法的进一步实施方式的权重生成器;

[0030] 图7示出了根据时间和频率的真实的和估计的扩散噪声比;

[0031] 图8示出了比较空间滤波器的方向性指数和白噪声增益;

[0032] 图9示出了估计的到达方向以及所产生的增益;以及

[0033] 图10示出了立体扬声器再现的情况的实例。

具体实施方式

[0034] 图1a示出了滤波器100,其用于基于两个或更多个输入麦克风信号生成音频输出信号(其包括多个音频输出信号样本)。在时频域中表示音频输出信号以及两个或更多个输入麦克风信号,其中,多个音频输出信号样本中的每个被分配给多个时频段(k, n)中的一个时频段(k, n)。

[0035] 滤波器100包括权重生成器110,其被适配成为多个时频段(k, n)中的每个,接收一个或多个声源的一个或多个声音分量的到达方向信息或一个或多个声源的位置信息,并且被适配为根据所述时频段(k, n)的一个或多个声源的一个或多个声音分量的到达方向信息或者根据所述时频段(k, n)的一个或多个声源的位置信息,为多个时频段(k, n)中的每个,生成加权信息。

[0036] 而且,滤波器包括输出信号生成器120,其用于通过根据所述时频段(k, n)的加权信息并且根据分配给所述时频段(k, n)的音频输入样本,为多个时频段(k, n)中的每个生成分配给所述时频段(k, n)的所述多个音频输出信号样本中的一个,来生成两个或更多个输入麦克风信号中的每个的音频输出信号。

[0037] 例如,两个或更多个输入麦克风信号中的每个包括多个音频输入样本,其中,每个音频输入样本被分配给一个时频段(k, n),并且音频信号生成器120可被适配为根据所述时频段(k, n)的加权信息并且根据两个或更多个输入麦克风信号中的每个的音频输入样本中的一个,即,根据两个或更多个输入麦克风信号中的每个的音频输入样本中的被分配给所述时频段(k, n)的一个音频输入样本,生成多个音频输出信号样本中的被分配给所述时频段(k, n)的一个音频输出信号样本。

[0038] 对于每个时频段(k, n)的要生成的每个音频输出信号样本,权重生成器110重新生成单独的加权信息。然后,音频信号生成器120基于为该时频段生成的加权信息,生成考虑

的时频段(k,n)的音频输出信号样本。换言之,通过权重生成器110为要生成音频输出信号样本的每个时频段计算新加权信息。

[0039] 在生成加权信息时,权重生成器110被适配为考虑一个或多个声源的信息。

[0040] 例如,权重生成器110可考虑第一声源的位置。在实施方式中,权重生成器还可考虑第二声源的位置。

[0041] 例如,或者,第一声源可发射具有第一声音分量的第一声波。具有第一声音分量的第一声波到达麦克风,并且权重生成器110可考虑第一声音分量的/声波的到达方向。借此,权重生成器110将关于第一声源的信息考虑在内。而且,第二声源可发射具有第二声音分量的第二声波。具有第二声音分量的第二声波到达麦克风,并且权重生成器110可考虑第二声音分量的/第二声波的到达方向。借此,权重生成器110还将关于第二声源的信息考虑在内。

[0042] 图1b示出了根据实施方式的滤波器100的可能的应用场景。具有第一声音分量的第一声波由第一扬声器121(第一声源)发射并且到达第一麦克风111。考虑在第一麦克风111处的第一声音分量的到达方向(=第一声波的到达方向)。而且,具有第二声音分量的第二声波由第二扬声器122(第二声源)发射并且到达第一麦克风111。权重生成器110还能够考虑在第一麦克风111处的第二声音分量的到达方向,以确定加权信息。而且,权重生成器还可考虑在其他麦克风(例如,麦克风112)处的声音分量的到达方向(=声波的到达方向),以确定加权信息。

[0043] 应注意的是,声源可以(例如)是在环境中物理上存在的物理声源,例如,扬声器、乐器或个人说话。

[0044] 然而,应注意的是,镜像源(mirror image source)也是声源。例如,由扬声器122发射的声波可被壁125反射,并且然后,声波似乎从与实际上发射声波的扬声器的位置不同的位置123发射。这种镜像源123也被视为声源。权重生成器110可被适配为根据与镜像源相关的到达方向信息或者根据关于一个、两个或更多个镜像源的位置信息来生成加权信息。

[0045] 图2示出了根据实施方式的滤波器100以及多个麦克风111、112、113、...、11n。在图2的实施方式中,滤波器100进一步包括滤波器组101。而且,在图2的实施方式中,权重生成器110包括信息计算模块102、权重计算模块103以及传递函数选择模块104。

[0046] 在时频域中进行处理,分别地,k表示频率指数,并且n表示时间指数。将M个时域麦克风信号 $x_{1...M}(t)$ 从麦克风111、112、113、...、11n中输入设备(滤波器100)中,通过滤波器组101将这些时域麦克风信号转换成时频域。由以下向量给出转换的麦克风信号:

[0047] $x(k,n) = [X_1(k,n) X_2(k,n) \dots X_M(k,n)]^T$ 。

[0048] 滤波器100输出期望信号 $Y(k,n)$ (音频输出信号)。音频输出信号(期望信号) $Y(k,n)$ 可例如表示用于单声道再现的增强信号、用于双耳声音再现的耳机信号或者用于具有任意的扬声器设置的空间声音再现的扬声器信号。

[0049] 由输出信号生成器120生成期望信号 $Y(k,n)$,例如,通过例如利用以下公式基于瞬时复杂权重 $w(k,n) = [W_1(k,n) W_2(k,n) \dots W_M(k,n)]^T$ 进行M个麦克风信号 $x(k,n)$ 的线性组合:

[0050] $Y(k,n) = w^H(k,n) x(k,n)$ (1)

[0051] 通过权重计算模块103确定权重 $w(k,n)$ 。对于每个k和每个n,重新确定确定 $w(k,n)$ 。换言之,对于每个时频段(k,n),进行权重 $w(k,n)$ 的确定。更具体而言,例如,基于瞬时参数信息(IPI) $\mathcal{J}(k,n)$ 并且基于相应的期望传递函数 $G(k,n)$,计算权重 $w(k,n)$ 。

[0052] 信息计算模块102被配置为从麦克风信号 $x(k, n)$ 计算IPI $\mathcal{J}(k, n)$ 。IPI描述在给定的时频瞬间 (k, n) 的信号以及包含在麦克风信号 $x(k, n)$ 内的噪声分量的特定特征。

[0053] 图3示出了根据实施方式的权重生成器110。权重生成器110包括信息计算模块102、权重计算模块103以及传递函数选择模块104。

[0054] 如在图3中的实例中所示，IPI主要包括例如由DOA估计模块201计算的一个或方向性声音分量(例如，平面波)的瞬时到达方向(DOA)。

[0055] 如下面所解释的，DOA信息可通过空间频率(例如，通过 $\mu[k | \varphi(k, n)]$)、通过相移(例如，通过 $a[k | \varphi(k, n)]$)、通过麦克风之间的时间延迟、通过传播向量(例如，由 $a[k | \varphi(k, n)]$)、通过双耳间强度差(ILD)或者由两耳时差(ITD)表示为角度(例如，通过[方位角 $\varphi(k, n)$ 、仰角 $\theta(k, n)$])。

[0056] 而且，IPI $\mathcal{J}(k, n)$ 可例如包括其他信息，例如，信号或噪声分量的二阶统计量(SOS)。

[0057] 在实施方式中，权重生成器110被适配为根据关于两个或更多个输入麦克风信号的信号或噪声分量的统计信息，为多个时频段 (k, n) 中的每个生成加权信息。例如，这种统计信息是在此处提及的二阶统计量。例如，统计信息可以是噪声分量的功率、信号扩散(signal-to-diffuse)信息、信号噪声(signal-to-noise)信息、扩散噪声(diffuse-to-noise)信息、信号分量的功率、扩散分量的功率、或者两个或更多个输入麦克风信号的信号分量或噪声分量的功率谱密度矩阵。

[0058] 二阶统计量可以由统计计算模块205计算。该二阶统计量信息可例如包括固定噪声分量(例如，自噪声)的功率、非固定噪声分量(例如，扩散噪声)的功率、信号扩散比(SDR)、信噪比(SNR)或扩散噪声比(DNR)。该信息允许根据特定的优化标准计算最佳权重 $w(k, n)$ 。

[0059] “固定噪声分量”/“缓慢变化的噪声分量”例如是具有随着时间不变化或缓慢变化的统计特性的噪声分量。

[0060] “非固定噪声分量”例如是具有随着时间快速变化的统计特性的噪声分量。

[0061] 在实施方式中，权重生成器110被适配为根据表示关于两个或更多个输入麦克风信号的第一噪声分量的信息的第一噪声信息并且根据表示关于两个或更多个输入麦克风信号的第二噪声分量的信息的第二噪声信息，为多个时频段 (k, n) 中的每个生成加权信息。

[0062] 例如，第一噪声分量可以是非固定噪声分量，并且第一噪声信息可以是关于非固定噪声分量的信息。

[0063] 例如，第二噪声分量可以是固定噪声分量/缓慢变化的噪声分量，并且第二噪声信息可以是关于固定的/缓慢变化的噪声分量的信息。

[0064] 在实施方式中，权重生成器110被配置为通过利用例如预定义的统计信息(例如，由非固定噪声分量引起的关于两个或更多个输入麦克风信号之间的空间相干性的信息)生成第一噪声信息(例如，关于非固定的/非缓慢变化的噪声分量的信息)，并且其中，权重生成器110被配置为在不利用统计信息的情况下生成第二噪声信息(例如，关于固定的/缓慢变化的噪声分量的信息)。

[0065] 关于快速变化的噪声分量,输入麦克风信号不能单独提供足够的信息来确定关于这种噪声分量的信息。例如,此外,需要统计信息来确定关于快速变化的噪声分量的信息。

[0066] 然而,关于不变化或不快速变化的噪声分量,不需要统计信息来确定关于这些噪声分量的信息。相反,评估麦克风信号是足够的。

[0067] 应注意的是,可利用估计的DOA信息计算统计信息,如在图3中所示。应进一步注意的是,还可在外部提供IPI。例如,声音(分别是声源的位置)的DOA可由摄影机以及脸部识别算法确定,假设人类谈话者形成声音场景。

[0068] 传递函数选择模块104被配置为提供传递函数 $G(k, n)$ 。图2和图3的(可能复杂的)传递函数 $G(k, n)$ 描述了系统的期望响应,给定(例如,当前参数)IPI $\mathcal{J}(k, n)$ 。例如, $G(k, n)$ 可以描述用于在单声道再现中的信号增强的期望空间麦克风的任意拾取图案(pick-up pattern)、用于扬声器再现的DOA相关的扬声器增益、或者双耳再现的头相关传递函数(HRTF)。

[0069] 应注意的是,通常,记录的声音场景的统计信息随着时间和频率快速地变化。结果,IPI $\mathcal{J}(k, n)$ 以及相应的最佳加权 $w(k, n)$ 仅仅对于特定的时频指数有效,因此,给每个 k 和 n 重新计算。因此,系统可立即适配于当前的记录情况。

[0070] 应进一步注意的是, M 个输入麦克风可形成单个麦克风阵列,或者可分布为在不同的位置形成多个阵列。而且,IPI $\mathcal{J}(k, n)$ 可包括位置信息,而非DOA信息,例如,声源在三维空间中的位置。借此,可定义空间滤波器,这些滤波器不仅根据需要对特定的方向滤波,而且对记录场景的三维空间区域进行滤波。

[0071] 在声源的位置信息是可用的时,对DOA提供的所有解释同样适用。例如,位置信息可由DOA(角度)和距离表示。在使用这种位置表示时,可从位置信息中立即获得DOA。或者,位置信息可例如由 x, y, z 坐标描述。然后,基于声源的位置信息并且基于记录各自的输入麦克风信号的麦克风的位置,可容易计算DOA。

[0072] 在下文中,描述进一步实施方式。

[0073] 一些实施方式允许通过去混响和噪声降低进行在空间上的选择性声音记录。在该背景下,提供实施方式,应用空间滤波,用于在源提取、去混响以及噪声降低方面的信号增强。这种实施方式的目的在于,计算与具有任意拾取图案的方向麦克风的输出对应的信号 $Y(k, n)$ 。这意味着方向性声音(例如,单个平面波)根据其DOA根据需要衰减或保存,同时抑制扩散声音或麦克风自噪声。根据实施方式,所提供的空间滤波器尤其结合使现有技术空间滤波器的优点,提供在具有高DNR的情况下的高方向性指数(DI)以及高白噪声增益(WNG)。根据一些实施方式,空间滤波器可仅仅受到线性约束,这允许快速计算权重。例如,图2和图3的传递函数 $G(k, n)$ 可例如表示方向性麦克风的期望拾取图案。

[0074] 在下文中,提供该问题的公式表示。然后,提供权重计算模块103和IPI计算模块102的实施方式,以利用去混响和噪声降低来进行在空间上的选择性声音记录。而且,描述了相应TF选择模块104的实施方式。

[0075] 首先,提供该问题的公式表示。考虑位于 $d_1 \dots d_M$ 处的 M 个全方向性麦克风的阵列。对于每个 (k, n) ,假设声场由在各向同性以及在空间上均匀的扩散声场中传播的 $L < M$ 个平面波(方向声音)构成。麦克风信号 $x(k, n)$ 可写为:

$$[0076] \quad \mathbf{x}(k, n) = \sum_{l=1}^L \mathbf{x}_l(k, n) + \mathbf{x}_d(k, n) + \mathbf{x}_n(k, n), \quad (2)$$

[0077] 其中, $\mathbf{x}_l(k, n) = [X_l(k, n, d_1) \dots X_l(k, n, d_M)]^T$ 包括与第1个平面波的声压成比例的麦克风信号, $\mathbf{x}_d(k, n)$ 是所测量的非固定噪声(例如, 扩散噪声), 并且 $\mathbf{x}_n(k, n)$ 是固定噪声/缓慢变化的噪声(例如, 麦克风自噪声)。

[0078] 假设在公式(2)中的三个分量彼此不相关, 则麦克风信号的功率谱密度(PSD)矩阵可由以下公式描述:

$$[0079] \quad \begin{aligned} \Phi(k, n) &= E \{ \mathbf{x}(k, n) \mathbf{x}^H(k, n) \} \\ &= \sum_{l=1}^L \Phi_l(k, n) + \Phi_d(k, n) + \Phi_n(k, n) \end{aligned}, \quad (3)$$

[0080] 其中, $\Phi_d(k, n) = \phi_d(k, n) \Gamma_d(k)$ (4)

[0081] 其中, $\Phi_n(k, n)$ 是固定噪声/缓慢变化的噪声的PSD矩阵, 并且 $\phi_d(k, n)$ 是非固定噪声的预期功率, 该功率可随着时间和频率快速地变化。由 $\gamma_{ij}(k)$ 表示的相干矩阵 $\Gamma_d(k)$ 的第 ij 个元素是由非固定噪声引起的在麦克风 i 和 j 之间的相干性。例如, 对于球面各向同性扩散场, $\gamma_{ij}(k) = \text{sinc}(\kappa r_{ij})$ [20], 其中, 波数 κ 和 $r_{ij} = ||d_j - d_i||$ 。相干矩阵 $\Gamma_d(k)$ 的第 ij 个元素是由固定噪声/缓慢变化的噪声引起的在麦克风 i 和 j 之间的相干性。对于麦克风自噪声, $\Phi_n(k, n) = \phi_n(k, n) \mathbf{I}$, 其中, \mathbf{I} 是单位矩阵并且 $\phi_n(k, n)$ 是自噪声的期望功率。

[0082] 在(2)中的方向性声音 $\mathbf{x}_l(k, n)$ 可写为:

$$[0083] \quad \mathbf{x}_l(k, n) = \mathbf{a}[k | \varphi_l(k, n)] X_l(k, n, \mathbf{d}_1), \quad (5)$$

[0084] 其中, $\varphi_l(k, n)$ 是第1个平面波的DOA的方位角 ($\varphi=0$ 表示阵列垂射) 并且 $\mathbf{a}[k | \varphi_l(k, n)] = [a_1[k | \varphi_l(k, n)] \dots a_M[k | \varphi_l(k, n)]]^T$ 是传播向量。 $\mathbf{a}[k | \varphi_l(k, n)]$ 的第 i 个元素

$$[0085] \quad a_i[k | \varphi_l(k, n)] = \exp\{j\kappa r_i \sin \varphi_l(k, n)\}, \quad (6)$$

[0086] 描述了从第一到第 i 个麦克风的第1个平面波的相移。应注意的是, $r_i = ||d_i - d_1||$ 等于第一与第 i 个麦克风之间的距离。

[0087] 角度 $\angle a_i[k | \varphi_l(k, n)] = \mu_i[k | \varphi_l(k, n)]$ 通常称为空间频率。第1个波的DOA可由 $\varphi_l(k, n)$ 、 $a_i[k | \varphi_l(k, n)]$ 、 $\mathbf{a}[k | \varphi_l(k, n)]$ 或者由 $\mu_i[k | \varphi_l(k, n)]$ 表示。

[0088] 如上所述, 实施方式的目的, 对麦克风信号 $\mathbf{x}(k, n)$ 进行滤波, 使得从特定空间区域到达的方向性声音根据需要衰减或放大, 同时抑制固定和非固定噪声。因此, 期望的信号可表示为:

$$[0089] \quad Y(k, n) = \sum_{l=1}^L G[k | \varphi_l(k, n)] X_l(k, n, \mathbf{d}_1), \quad (7)$$

[0090] 其中, $G[k | \varphi(k, n)]$ 是实值或复值任意的 (例如, 预定义的) 方向性函数, 该函数可以与频率相关。

[0091] 图4涉及根据实施方式的具有两个任意的方向性函数和源位置的场景。尤其地, 图4显示了两个实例方向性 $G_1[k | \varphi(k, n)]$ 和 $G_2[k | \varphi(k, n)]$ 的幅度。在使用 $G_1[k | \varphi(k, n)]$ (参照在图4中的实线) 时, 从 $\varphi < 45^\circ$ 中达到的方向声音衰减21dB, 而其他方向的方向声音不衰减。原则上, 任意方向性可被设计为偶函数, 例如, $G_2[k | \varphi(k, n)]$ (参照在图4中的虚线)。而且, $G[k | \varphi(k, n)]$ 可以设计为随着时间变化, 例如, 一旦定位, 就提取移动的或出现的声源。

[0092] 通过麦克风信号的线性组合, 获得信号 $Y(k, n)$ 的估计, 例如, 通过

$$[0093] \quad \hat{Y}(k, n) = \mathbf{w}^H(k, n) \mathbf{x}(k, n) \quad (8)$$

[0094] 其中, $\mathbf{w}(k, n)$ 是长度 M 的复杂权重向量。在下面获得相应的最佳权重向量 $\mathbf{w}(k, n)$ 。在下文中, 为了简单起见, 省略权重 $\mathbf{w}(k, n)$ 对 k 和 n 的依赖性。

[0095] 现在, 描述在图2和图3中的权重计算模块103的两个实施方式。

[0096] 从 (5) 和 (7) 中, 遵循 $\mathbf{w}(k, n)$ 应满足线性约束:

$$[0097] \quad \mathbf{w}^H(k, n) \mathbf{a}[k | \varphi_l(k, n)] = G[k | \varphi_l(k, n)], \quad l \in \{1, 2, \dots, L\}. \quad (9)$$

[0098] 而且, 在滤波器的输出处的非固定和固定/缓慢变化的噪声功率应尽可能减小。

[0099] 图5描述了应用空间滤波的本发明的实施方式。尤其地, 图5示出了根据另一个实施方式的权重生成器110。再次, 权重生成器110包括信息计算模块102、权重计算模块103以及传递函数选择模块104。

[0100] 更尤其地, 图5示出了线性约束的最小方差 (LCMV) 方法。在该实施方式中 (见图5), 根据包括 L 个平面波的DOA的 $IPI \ I(k, n)$ 以及固定和非固定的噪声的统计信息, 计算权重 $\mathbf{w}(k, n)$ 。后面的信息可包括DNR、两个噪声分量的单独功率 $\Phi_n(k, n)$ 以及 $\Phi_d(k, n)$ 、或者两个噪声分量的PSD矩阵 Φ_n 以及 Φ_d 。

[0101] 例如, Φ_d 可被视为关于两个噪声分量的第一噪声分量的第一噪声信息, 并且 Φ_n 可被视为关于两个噪声分量的第二噪声分量的第二噪声信息。

[0102] 例如, 权重生成器110可被配置为根据在一个或多个麦克风输入信号的至少一些第一噪声分量之间的一个或多个相干性, 来确定第一噪声信息 Φ_d 。例如, 权重生成器110可被配置为根据表示由两个或更多个输入麦克风信号的第一噪声分量引起的相干性的相干矩阵 $\Gamma_d(k)$, 来确定第一噪声信息, 例如, 通过应用公式 $\Phi_d(k, n) = \Phi_d(k, n) \Gamma_d(k)$ 。

[0103] 通过在滤波器的输出处最小化自噪声 (固定噪声/缓慢变化的噪声) 和扩散声音功率 (非固定噪声) 的总和, 找出用于解决在 (8) 中的问题的权重 $\mathbf{w}(k, n)$, 即,

$$[0104] \quad \mathbf{w}_{nd} = \arg \min_{\mathbf{w}} \mathbf{w}^H \underbrace{[\Phi_d(k, n) + \Phi_n(k, n)]}_{\Phi_u(k, n)} \mathbf{w} \quad (10)$$

$$[0105] \quad = \arg \min_{\mathbf{w}} \mathbf{w}^H \underbrace{[\phi_d(k, n) \Gamma_d(k) + \Phi_n(k, n)]}_{\Phi_u(k, n)} \mathbf{w} \quad (11)$$

[0106] 使用(4)并且假设 $\Phi_n(k, n) = \phi_n(k, n) \mathbf{I}$, 优化问题可表示为:

$$[0107] \quad \mathbf{w}_{nd} = \arg \min_{\mathbf{w}} \mathbf{w}^H \underbrace{[\Psi(k, n) \Gamma_d(k) + \mathbf{I}]}_{C(k, n)} \mathbf{w} \quad (12)$$

$$[0108] \quad \text{其中, } \Psi(k, n) = \frac{\phi_d(k, n)}{\phi_n(k, n)} \quad (13)$$

[0109] 是在麦克风处的时变输入DNR。给定约束条件(9), 对于(10)和(12)的解决方法是[21]。

$$[0110] \quad \mathbf{w}_{nd} = \Phi_u^{-1} \mathbf{A} [\mathbf{A}^H \Phi_u^{-1} \mathbf{A}]^{-1} \mathbf{g} \quad (14)$$

$$[0111] \quad = \mathbf{C}^{-1} \mathbf{A} [\mathbf{A}^H \mathbf{C}^{-1} \mathbf{A}]^{-1} \mathbf{g}, \quad (15)$$

[0112] 其中, $\mathbf{A}(k, n) = [\mathbf{a}[k | \varphi_1(k, n)] \dots \mathbf{a}[k | \varphi_l(k, n)]]$ 包括根据传播向量的L个平面波的DOA信息。相应的期望增益由以下公式给出:

$$[0113] \quad \mathbf{g}(k, n) = [G[k | \varphi_1(k, n)] \dots G[k | \varphi_l(k, n)]]^T. \quad (16)$$

[0114] 下面描述 $\Psi(k, n)$ 和其他需要的IPI的估计的实施方式。

[0115] 其他实施方式基于参数多通道维纳滤波器。在这种实施方式中, 如图6所示, IPI进一步包括关于信号统计的信息, 例如, 包括L个平面波(方向声音)的功率的信号PSD矩阵 $\Phi_s(k, n)$ 。而且, 可选控制参数 $\lambda_{1 \dots L}(k, n)$ 被视为控制L个平面波中的每个的信号失真的程度。

[0116] 图6示出了实施利用参数多通道维纳滤波器的权重生成器110的空间滤波器的应用的实施方式。在此, 权重生成器110包括信息计算模块102、权重计算模块103以及传递函数选择模块104。

[0117] 通过参数多通道维纳滤波器方法, 计算权重 $\mathbf{w}(k, n)$ 。维纳滤波器最小化在输出处的残余信号的功率, 即,

$$[0118] \quad \mathbf{w}_{nd} = \arg \min_{\mathbf{w}} \underbrace{\mathbf{E} \{ |\hat{Y}(k, n) - Y(k, n)|^2 \}}_{C(k, n)} \quad (17)$$

[0119] 被最小化的价值函数(cost function) $C(k, n)$ 可写为:

$$[0120] \quad C(k, n) = \mathbf{E} \{ |\hat{Y}(k, n) - Y(k, n)|^2 \} \quad (18)$$

$$[0121] \quad = [\mathbf{g} - \mathbf{A}^H(k, n) \mathbf{w}]^H \Phi_s(k, n) [\mathbf{g} - \mathbf{A}^H(k, n) \mathbf{w}] + \mathbf{w}^H \Phi_u(k, n) \mathbf{w}, \quad (19)$$

[0122] 其中, $\Phi_s(k, n) = \mathbf{E} \{ x_s(k, n) x_s(k, n)^H \}$ 包括方向声音PSD, 并且 $x_s(k, n) = [X_1(k, n,$

$d_1) X_2(k, n, d_1) \cdots X_L(k, n, d_1)]$ 包括与在参考麦克风处的 L 个平面波的声压成比例的信号。要注意的是, $\Phi_s(k, n)$ 是对角矩阵, 其中, 对角元素 $\text{diag}\{\Phi_s(k, n)\} = [\Phi_1(k, n) \dots \Phi_L(k, n)]^T$ 是到达的平面波的功率。为了控制引入的信号失真, 可以包括对角矩阵 $\Lambda(k, n)$, 其包括时间和频率相关的控制参数 $\text{diag}\{\Lambda\} = [\lambda_1(k, n) \lambda_2(k, n) \dots \lambda_L(k, n)]^T$, 即,

$$[0123] \quad C_{PW}(k, n) = [g-A^H(k, n)w]^H \Lambda(k, n) \Phi_s(k, n) [g-A^H(k, n)w]$$

$$[0124] \quad w^H \Phi_u(k, n) w. \quad (20)$$

[0125] 考虑 $C_{PW}(k, n)$ 的在 (17) 中的最小化问题的解决方案是:

$$[0126] \quad w = [A^H \Lambda(k, n) \Phi_s(k, n) A + \Phi_u]^{-1} A \Lambda(k, n) \Phi_s(k, n) g. \quad (21)$$

[0127] 这等同于

$$[0128] \quad \mathbf{w} = \Phi_u^{-1} \mathbf{A} [\Lambda^{-1} \Phi_s^{-1} + \mathbf{A}^H \Phi_u^{-1} \mathbf{A}]^{-1} \mathbf{g} \quad (22)$$

[0129] 应注意的是, 对于, $\Lambda^{-1} = 0$, 获得在 (14) 中的 LCMV 解决方案。对于 $\Lambda^{-1} = \mathbf{I}$, 获得多通道维纳滤波器。对于其他值 $\lambda_{1 \dots L}(k, n)$, 可分别控制相应源信号的失真程度以及残余噪声抑制的程度。因此, 通常根据可用的参数信息, 限定 $\lambda_l^{-1}(k, n)$, 即,

$$[0130] \quad \lambda_l^{-1}(k, n) = f(\mathcal{J}(k, n)), \quad (23)$$

[0131] 其中, $f(\cdot)$ 是任意用户定义的函数。例如, 可根据以下公式选择 $\lambda_{1 \dots L}(k, n)$:

$$[0132] \quad \lambda_l^{-1}(k, n) = \frac{1}{1 + \frac{\phi_l(k, n)}{\phi_u(k, n)}}, \quad (24)$$

[0133] 其中, $\phi_1(k, n)$ 是第 1 个信号 (第 1 个平面波) 的功率, 并且 $\phi_u(k, n) = \phi_n(k, n) + \phi_d(k, n)$ 是不期望信号 (固定噪声/缓慢变化的噪声加上非固定的噪声) 的功率。借此, 参数维纳滤波器取决于关于两个或更多个输入麦克风信号的信号分量的统计信息, 因此, 参数维纳滤波器进一步取决于关于两个或更多个输入麦克风信号的噪声分量的统计信息。

[0134] 如果与噪声相比, 源 1 强烈, 则获得接近 0 的 $\lambda_l^{-1}(k, n)$, 意味着获得 LCMV 解决方案 (没有源信号的失真)。如果与源功率相比, 噪声强烈, 则获得接近 1 的 $\lambda_l^{-1}(k, n)$, 意味着获得参数维纳滤波器 (强烈抑制噪声)。

[0135] 下面描述 $\Phi_s(k, n)$ 和 $\Phi_u(k, n)$ 的估计。

[0136] 在下文中, 描述瞬时参数估计模块 102 的实施方式。

[0137] 在能够计算权重之前, 需要估计不同的 IPI。可通过众所周知的窄带 DOA 估计器 (例如, ESPRIT [22] 或根 MUSIC [23]) 或其他现有技术估计器获得在模块 201 中计算的 L 个平面波。例如, 这些算法可为到达阵列的一个或多个波提供 (例如) 方位角 $\varphi(\mathbf{k}, n)$ 、空间频率 $\mu[\mathbf{k} | \varphi(\mathbf{k}, n)]$ 、相移 $\mathbf{a}[\mathbf{k} | \varphi(\mathbf{k}, n)]$ 、或传播向量 $\mathbf{a}[\mathbf{k} | \varphi(\mathbf{k}, n)]$ 。将不进一步讨论 DOA 估计, 因为 DOA 估计本身在本领域中是熟知的。

[0138] 在下文中, 描述扩散噪声比 (DNR) 估计。尤其地, 描述输入 DNR $\Psi(k, n)$ 的估计, 即, 在图 5 中的模块 202 的实现。DNR 估计利用在模块 201 中获得的 DOA 信息。为了估计 $\Psi(k, n)$, 可使用额外的空间滤波器, 该滤波器去除 L 个平面波, 以便仅仅捕捉扩散声音。例如, 通过最大

化阵列的WNG,找出这个空间滤波器的权重,即,

$$[0139] \quad \mathbf{w}_\Psi = \arg \min_{\mathbf{w}} \mathbf{w}^H \mathbf{w} \quad (25)$$

[0140] 服从

$$[0141] \quad \mathbf{w}^H \mathbf{a}[k | \varphi_l(k, n)] = 0, \quad l \in \{1, 2, \dots, L\}, \quad (26)$$

$$[0142] \quad \mathbf{w}^H \mathbf{a}[k | \varphi_0(k, n)] = 1. \quad (27)$$

[0143] 约束条件(27)确保非零加权 w_Ψ 。传播向量 $\mathbf{a}[k | \varphi_0(\mathbf{k}, n)]$ 与特定的方向 $\varphi_0(\mathbf{k}, n)$ 对应,方向与L个平面波的DOA $\varphi_l(\mathbf{k}, n)$ 不同。在下文中,对于 $\varphi_0(\mathbf{k}, n)$,选择与所有 $\varphi_l(\mathbf{k}, n)$ 相距最大距离的方向,即,

$$[0144] \quad \varphi_0(k, n) = \arg \max_{\varphi} (\min_l |\varphi - \varphi_l(k, n)|), \quad (28)$$

[0145] 其中, $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ 。给定权重 w_Ψ ,由以下公式给出额外空间滤波器的输出功率:

$$[0146] \quad \mathbf{w}_\Psi^H \Phi(k, n) \mathbf{w}_\Psi = \phi_d(k, n) \mathbf{w}_\Psi^H \Gamma_d(k) \mathbf{w}_\Psi + \phi_n(k, n) \mathbf{w}_\Psi^H \mathbf{w}_\Psi. \quad (29)$$

[0147] 现在,可通过(13)和(29)计算输入DNR,即,

$$[0148] \quad \Psi(k, n) = \frac{\mathbf{w}_\Psi^H \Phi(k, n) \mathbf{w}_\Psi - \phi_n(k, n) \mathbf{w}_\Psi^H \mathbf{w}_\Psi}{\phi_n(k, n) \mathbf{w}_\Psi^H \Gamma_d(k) \mathbf{w}_\Psi}. \quad (30)$$

[0149] 假设功率随着时间是恒定的或者缓慢地变化,在静音(silence)时,例如,可以估计所需要的麦克风自噪声 $\phi_n(k, n)$ 的预期功率。要注意的是,由于所选择的优化标准(45),所以所提出的DNR估计器实际上不必提供最低估计方差,但是提供公正的结果。

[0150] 在下文中,讨论非固定PSD $\phi_d(k, n)$ 的估计,即,在图5中的模块(202)的另一个实现。可以使用以下公式,估计非固定噪声的功率(PSD):

$$[0151] \quad \phi_d(k, n) = \frac{\mathbf{w}_\Psi^H [\Phi(k, n) - \Phi_n(k, n)] \mathbf{w}_\Psi}{\mathbf{w}_\Psi^H \Gamma_d(k) \mathbf{w}_\Psi}, \quad (31)$$

[0152] 其中,在前一个段落中,定义了 w_Ψ 。应注意的是,在静音时(即,在没有信号和非固定噪声时),可估计固定/缓慢变化的噪声PSD矩阵 $\Phi_n(k, n)$,即,

$$[0153] \quad \Phi_n(k, n) = E \{ \mathbf{x}(k, n) \mathbf{x}^H(k, n) \}, \quad (32)$$

[0154] 其中,通过对静音帧(silent frame) n 求平均来逼近期望值。可以通过现有技术方法检测静音帧。

[0155] 在下文中,讨论不期望信号PSD矩阵(见模块203)的估计。

[0156] 可通过以下公式:

$$[0157] \quad \Phi_u(k, n) = \phi_n(k, n) (\Psi(k, n) \Gamma_d(k) + \Gamma_n(k)), \quad (33)$$

[0158] 或者更通常通过以下公式:

$$[0159] \quad \Phi_u(k, n) = \phi_d(k, n) \Gamma_d(k) + \Phi_n(k, n), \quad (34),$$

[0160] 获得不期望信号(固定/缓慢变化的噪声加上非固定的噪声) $\Phi_u(k, n)$ 的PSD矩阵。

[0161] 其中, $\Gamma_d(k)$ 和 $\Gamma_n(k)$ 可用作先验信息(见上文)。如上所解释的,可计算DNR $\Psi(k, n)$ 、固定/缓慢变化的噪声功率 $\phi_n(k, n)$ 以及其他需要的量。因此, $\Phi_u(k, n)$ 估计利用由模块 201获得的DOA信息。

[0162] 在下文中,描述信号PSD矩阵(见模块204)的估计。

[0163] 可通过以下公式,计算用于计算 $\Phi_s(k, n)$ 所需要的到达平面波的功率 $\phi_{1\dots L}(k, n)$:

$$[0164] \quad \begin{bmatrix} \phi_1(k, n) \\ \vdots \\ \phi_L(k, n) \end{bmatrix} = \begin{bmatrix} \mathbf{w}_1(k, n) \\ \vdots \\ \mathbf{w}_L(k, n) \end{bmatrix}^H [\Phi_x(k, n) - \Phi_u(k, n)] \begin{bmatrix} \mathbf{w}_1(k, n) \\ \vdots \\ \mathbf{w}_L(k, n) \end{bmatrix}, \quad (35)$$

[0165] 其中,权重 w_l 抑制所有到达的平面波,除了第 l 个波,即,

$$[0166] \quad \mathbf{w}_l(k, n)^H \mathbf{a}_{l'}(k, n) = \begin{cases} 1 & \text{如果 } l = l' \\ 0 & \text{否则} \end{cases} \quad (36)$$

[0167] 例如,

$$[0168] \quad \mathbf{w}_l = \arg \min_{\mathbf{w}} \mathbf{w}^H \mathbf{w} \quad (37)$$

[0169] 经受(36)。 $\Phi_s(k, n)$ 估计利用在模块(201)中获得的DOA信息。如在前一段落中所解释的,可计算不期望信号 $\Phi_u(k, n)$ 的所需要的PSD矩阵。

[0170] 现在,描述根据实施方式的传递函数选择模块104。

[0171] 在本申请中,可以根据DOA信息 $\varphi_l(k, n)$ 找出相应的平面波 l 的增益 $G[k | \varphi_l(k, n)]$ 。不同DOA $\varphi(k, n)$ 的传递函数 $G[k | \varphi(k, n)]$ 可用于该系统,例如,作为用户定义的先验信息。还可基于图像的分析,例如,使用所检测的脸部的位置,计算增益。在图4中描述两个实例。这些传递函数与方向麦克风的期望的拾取图案对应。可提供传递函数 $G[k | \varphi(k, n)]$, 例如,作为查找表,即,对于估计的 $\varphi_l(k, n)$, 从查找表中选择相应的增益 $G[k | \varphi_l(k, n)]$ 。要注意的是,还可根据空间频率 $\mu[k | \varphi(k, n)]$, 而非方位角 $\varphi_l(k, n)$, 即,根据 $G(k, \mu)$, 而非 $G[k | \varphi(k, n)]$, 定义传递函数。还可基于源位置信息而非DOA信息,计算增益

[0172] 现在,提供实验结果。以下模拟结果展示了上述实施方式的可适用性。比较所提出的系统和现有技术系统,下面将进行解释。然后,讨论实验设置并且提供结果。

[0173] 首先,考虑现有空间滤波器。

[0174] 虽然在静音期间,可估计PSD $\Phi_n(k, n)$, 但是通常假设 $\Phi_d(k, n)$ 未知并且难以察觉。因此,考虑可在无这种知识的情况下计算的两个现有空间滤波器。

[0175] 第一空间滤波器已知为延时求和波束成形器,并且最小化在滤波器的输出处的自噪声功率[即,最大化WNG][1]。然后,通过以下公式,获得使经受(9)的在(7)与(8)之间的均方误差(MSE)最小化的优化权重向量:

$$[0176] \quad \mathbf{w}_n = \arg \min_{\mathbf{w}} \underbrace{\mathbf{w}^H \Phi_n(k, n) \mathbf{w}}_{\mathbf{w}^H \mathbf{w}} \quad (38)$$

[0177] 存在针对 (38) 的封闭式解决方案[1], 该解决方案允许快速计算 \mathbf{w}_n 。应注意的是, 该滤波器不必提供最大的DI。

[0178] 第二空间滤波器已知为稳健的超方向性 (SD) 波束成形器, 并且通过在WNG上的下限, 最小化在滤波器的输出处的扩散声音功率[即, 最大化DI] [24]。在WNG上的下限增大对在传播向量内的误差的稳健性, 并且限制自噪声的放大[24]。然后, 通过以下公式, 获得最小化经受 (9) 的在 (7) 与 (8) 之间的MSE并且满足在WNG上的下限的优化加权向量:

$$[0179] \quad \mathbf{w}_d = \arg \min_{\mathbf{w}} \underbrace{\mathbf{w}^H \Phi_d(k, n) \mathbf{w}}_{\mathbf{w}^H \Gamma_d(k, n) \mathbf{w}} \quad (39)$$

[0180] 并且经受二次约束 $\mathbf{w}^H \mathbf{w} < \beta$ 。参数 β^{-1} 定义最小WNG并且确定滤波器的可实现的DI。实际上, 通常难以在低SNR情况下的充分WNG与在高SNR情况下的足够高的DI之间找出最佳权衡。而且, 由于二次约束, 所以解决 (39), 造成非凸优化问题, 这需要时间来解决。由于时变约束 (9) 造成需要给每个 k 和 n 重新计算复杂加权向量, 所以这尤其是个问题。

[0181] 现在, 考虑实验设置。假设在 (2) 中的模型中的 $L=2$ 个平面波以及具有麦克风间的间距3cm的 $M=4$ 个麦克风的均匀线性阵列 (ULA), 使用源图像方法[25、26], 模拟鞋盒空间 ($7.0 \times 5.4 \times 2.4 \text{m}^3$ 、 $RT_{60} \approx 380 \text{ms}$), 在 $\varphi_A = 86^\circ$ 以及 $\varphi_B = 11^\circ$ 时分别具有两个声源 (距离1.75m, 比较图4)。信号包括0.6s静音, 其后跟随模糊词 (double talk)。将白高斯噪声加到麦克风信号中, 产生26dB的分段信噪比 (SSNR)。声音以16kHz取样并且使用512点STFT转换成时频域, 具有50%重叠。

[0182] 考虑图4的方向性函数 $G_1(\varphi)$, 即, 应在不失真的情况下提取源A, 同时源B的功率衰减21dB。考虑上面这两个空间滤波器以及所提供的空间滤波器。对于稳健的SD波束成形器 (39), 最小的WNG设为-12dB。对于提供的空间滤波器 (12), 如上面所解释的, 估计DNR $\Psi(k, n)$ 。在开始时, 从静音信号部分计算自噪声功率 $\Phi_n(k, n)$ 。在 (3) 中的预期由超过 $\tau=50 \text{ms}$ 的递归平均时间逼近。

[0183] 在下文中, 考虑非时变方向约束。

[0184] 为了这个模拟, 假设关于这两个源位置 φ_A 和 φ_B 的先验知识。在所有处理步骤中, 使用 $\varphi_1(\mathbf{k}, n) = \varphi_A$ 和 $\varphi_2(\mathbf{k}, n) = \varphi_B$ 。因此, 在 (9) 和 (26) 中的方向约束不随着时间变化。

[0185] 图7示出了真实的和估计的DNR $\Psi(k, n)$ 。这两个标记的区域分别表示信号的静音和活动部分。尤其地, 图7描述了根据时间和频率的真实的和估计的DNR $\Psi(k, n)$ 。由于混响环境, 所以在语音活动期间, 获得较高的DNR。由于结合的时间平均化处理, 所以在图7 (b) 中的估计的DNR处理有限的时间分辨率。然而, $\Psi(k, n)$ 估计足够精确, 如以下结果所示。

[0186] 图8 (a) 描述了 \mathbf{w}_n 和 \mathbf{w}_d (均与信号无关) 的以及所提出的空间滤波器 \mathbf{w}_{nd} (与信号相关) 的平均DI。对于所提出的空间滤波器, 显示了信号的静音部分以及在语音活动期间 [在图7 (b) 中标记这两个信号部分] 的DI。在静音期间, 所提出的空间滤波器 (虚线 \mathbf{w}_{nd}) 提供与 \mathbf{w}_n 相同的低DI。在语音活动期间 (实线 \mathbf{w}_{nd}), 所获得的DI与稳健的SD波束成形器 (\mathbf{w}_d) 一样高。图

8 (b) 显示了相应的WNG。在静音期间,所提出的空间滤波器(虚线 w_{nd})实现了高WNG,而在信号活动期间,WNG较低。

[0187] 图8:比较空间滤波器的DI和WNG。对于 w_d ,最小WNG设为-12dB,以使空间滤波器稳健,抵抗麦克风自噪声。

[0188] 通常,图8示出了所提出的空间滤波器使这两个现有空间滤波器的优点相结合:在静音部分期间,提供最大WNG,导致最小自噪声放大,即,高稳健性。

[0189] 在通常掩蔽自噪声的信号活动和高混响期间,提供高DI(以低WNG为代价),导致扩散声音的最佳降低。在这种情况下,甚至相当小的WNG可容忍。

[0190] 要注意的是,对于更高的频率($f > 5\text{kHz}$),由于在(39)和(12)中的相关矩阵 $\Gamma_d(k)$ 几乎等于单位矩阵,所以所有空间滤波器几乎相同地运行。

[0191] 在下文中,考虑瞬时方向约束。

[0192] 对于该模拟,假设没有关于 φ_A 和 φ_B 的可用先验信息。通过ESPRIT估计DOAs $\varphi_1(k, n)$ 和 $\varphi_2(k, n)$ 。因此,约束条件(9)随着时间变化。仅仅对于稳健的SD波束成形器(w_d),使用与固定观测方向 $\varphi_A = 86^\circ$ 对应的单个和非时变约束条件(9)。这个波束成形器用作参考。

[0193] 图9描述了估计的DOA $\varphi_1(k, n)$ 和所产生的增益 $G[k | \varphi_1(k, n)]$ 。尤其地,图9示出了估计的DOA $\varphi_1(k, n)$ 和所产生的增益 $\|G[k | \varphi_1(k, n)]\|^2$ 。如果DOA在图4中位于空间窗口内部(实线),则到达的平面波不衰减。否则,波的功率衰减21dB。

[0194] 表1示出了所有空间滤波器(*未处理的)的性能。在括号内的值表示非时变方向约束条件,不在括号内的值表示瞬时方向约束条件。在计算SIR、SRR以及SSNR之前,信号进行A加权。

| | SIR [dB] | SRR [dB] | SSNR [dB] | PESQ |
|--------------|----------------|---------------|----------------|------------------|
| * | 11 (11) | -7 (-7) | 26 (26) | 1.5 (1.5) |
| [0195] w_n | 21 (32) | -2 (-3) | 33 (31) | 2.0 (1.7) |
| w_d | 26 (35) | 0 (-1) | 22 (24) | 2.1 (2.0) |
| w_{nd} | 25 (35) | 1 (-1) | 28 (26) | 2.1 (2.0) |

[0196] 表1

[0197] 尤其地,表1总结了空间滤波器在滤波器的输出处的在信号干扰比(SIR)、信号混响比(SRR)以及SSNR方面的总体性能。在SIR和SRR(源分离、去混响)方面,所提出的方法(w_{nd})和稳健的SD波束成形器(w_d)提供最高性能。然而,所提出的 w_{nd} 的SSNR比 w_d 的SSNR高6dB,这表示可以清楚听见的优点。使用 w_n ,获得在SSNR方面的最佳性能。在PESQ方面, w_{nd} 和 w_d 胜过 w_n 。使用瞬时方向约束条件,而非时变约束条件(在括号内的值)主要降低可实现的SIR,但是在改变源位置的情况下,提供快速适应。应注意的是,每个时间帧的所有需要的复杂加权的计算时间对于 w_d 大于80s(CVX工具箱[27,28])并且对于所提出的方法小于0.08s(MATLAB R2012b、MacBook Pro 2008)。

[0198] 在下文中,描述空间声音再现的实施方式。实施方式的目的在于,(例如)通过麦克风阵列捕捉声音场景,并且通过任意的声音再现系统(例如,5.1扬声器设置、耳机再现)再

现空间声音,以便再创建原始空间印象。假设声音再现系统包括N个通道,即,计算N个输出信号 $Y(k, n)$ 。

[0199] 首先,提供问题公式表示。考虑信号模型(见以上公式(2)),并且公式表示相似的问题。固定/缓慢变化的噪声与不期望的麦克风自噪声对应,而非固定的噪声与期望的扩散声音对应。由于再现记录场景的原始空间印象至关重要,所以在本申请中期望扩散声音。

[0200] 在下文中,应在没有失真的情况下从相应的DOA $\varphi_l(k, n)$ 实现再现方向声音 $X_l(k, n, d_1)$ 。而且,应利用来自所有方向的正确能量再现扩散声音,同时抑制麦克风自噪声。因此,在(7)中的期望的信号 $Y(k, n)$ 现在表示为:

$$[0201] \quad Y_i(k, n) = \sum_{l=1}^L G_l[k | \varphi_l(k, n)] X_l(k, n, \mathbf{d}_1) + G_d(k, n) X_{d,i}(k, n, \mathbf{d}), \quad (40)$$

[0202] 其中, $Y_i(k, n)$ 是声音再现系统的第*i*个信道的信号($i = \{1, \dots, N\}$), $X_{d,i}(k, n, \mathbf{d})$ 是在从扬声器*i*中再现的任意点(例如,在第一麦克风 d_1)处测量的扩散声音,并且 $G_d(k, n)$ 是扩散声音的增益函数,以在再现期间确保扩散声音的正确功率(通常 $G_d(k, n) = 1/\sqrt{N}$)。理想地,信号 $X_{d,i}(k, n)$ 具有正确的扩散声音功率,并且在信道*i*上相互不相关,即,

$$[0203] \quad E \{ X_{d,i}(k, n) X_{d,j}^*(k, n) \} = \begin{cases} \phi_d(k, n) & \text{如果 } i = j \\ 0 & \text{否则} \end{cases} \quad (41)$$

[0204] 方向声音分量的传递函数 $G_l[k | \varphi_l(k, n)]$ 与和DOA相关的扬声器增益函数对应。在图10中描述用于立体扬声器再现的情况的实例。如果波1从 $\varphi_1(k, n) = 30^\circ$ 中到达,那么 $G_1 = 1$ 并且 $G_2 = 0$ 。这表示仅从再现系统的信道*i* = 1(左信道)中再现这个方向声音。对于 $\varphi_1(k, n) = 0^\circ$,具有 $G_1 = G_2 = \sqrt{0.5}$,即,通过这两个扬声器的相等功率,再现方向声音。或者,如果期望双耳再现,那么 $G_l[k | \varphi_l(k, n)]$ 可以与HRTF对应。

[0205] 如上所述,基于复杂权重 $w(k, n)$,通过麦克风信号的线性组合,估计信号 $Y_i(k, n)$,即,

$$[0206] \quad \hat{Y}_i(k, n) = \mathbf{w}_i^H(k, n) \mathbf{x}(k, n), \quad (42)$$

[0207] 经受特定的约束条件。在接下来的小部分中解释加权 $w(k, n)$ 的约束条件和计算。

[0208] 在下文中,在本背景下,考虑根据相应的实施方式的权重计算模块103,提供图2的权重计算模块103的两个实施方式。从公式(5)和公式(40)中得出以下结论: $w_i(k, n)$ 应满足线性约束:

$$[0209] \quad \mathbf{w}_i^H(k, n) \mathbf{a}[k | \varphi_l(k, n)] = G_l[k | \varphi_l(k, n)], \quad l \in \{1, 2, \dots, L\}, i \in \{1, 2, \dots, N\}. \quad (43)$$

[0210] 而且,应保持扩散声音功率。因此, $w_i(k, n)$ 可以满足二次约束:

$$[0211] \quad \mathbf{w}_i^H \mathbf{\Gamma}_d(k, n) \mathbf{w}_i = |G_d(k, n)|^2, \quad \forall i. \quad (44)$$

[0212] 而且,在滤波器的输出处的自噪声功率应最小化。因此,最佳权重可计算为:

$$[0213] \quad \mathbf{w}_i = \arg \min_{\mathbf{w}} \mathbf{w}^H \mathbf{w} \quad (45)$$

[0214] 经受公式(43)和公式(44)。这导致(例如)通过熟知的数值方法[29]可以解决的凸优化问题。

[0215] 对于瞬时参数估计模块102,根据相应的实施方式,通过熟知的窄带DOA估计器(例如,ESPRIT[22]或根MUSIC[23])或其他现有技术估计器,可以获得L个平面波的DOA $\varphi_1(\mathbf{k}, n)$ 。

[0216] 现在,考虑根据相应的实施方式的传递函数选择模块104。在本申请中,根据DOA信息 $\varphi_1(\mathbf{k}, n)$ 为相应的方向声音1找出信道i的增益 $G_i[\mathbf{k} | \varphi_1(\mathbf{k}, n)]$ 。用于不同的DOA $\varphi(\mathbf{k}, n)$ 和信道i的传递函数 $G_i[\mathbf{k} | \varphi(\mathbf{k}, n)]$ 可用于系统,例如,作为用户定义的先验信息。还可基于图像的分析,例如,使用所检测的脸部的位置,计算增益。

[0217] 通常提供传递函数 $G_i[\mathbf{k} | \varphi(\mathbf{k}, n)]$,作为查找表,即,为估计的 $\varphi_1(\mathbf{k}, n)$,从查找表中选择相应的增益 $G_i[\mathbf{k} | \varphi_1(\mathbf{k}, n)]$ 。要注意的是,还可以根据空间频率 $\mu[\mathbf{k} | \varphi(\mathbf{k}, n)]$,而非方位角 $\varphi_1(\mathbf{k}, n)$,即,根据 $G_i(k, \mu)$,而非 $G_i[\mathbf{k} | \varphi(\mathbf{k}, n)]$,定义传递函数。要注意的是,传递函数还可与能够进行双耳声再现的HRTF对应。在这种情况下, $G_i[\mathbf{k} | \varphi(\mathbf{k}, n)]$ 通常是复值。要注意的是,还可以根据源位置信息,而非DOA信息,计算增益或传递函数。

[0218] 在图10中描述了立体扬声器再现的一个实例。尤其地,图10示出了立体再现的增益函数。

[0219] 虽然在设备的背景下描述了一些方面,但是显然,这些方面还表示相应方法的描述,其中,方框或装置与方法步骤或方法步骤的特征对应。相似地,在方法步骤的背景下描述的方面还表示相应设备的相应方框或项目或特征的描述。

[0220] 本发明的分解信号可存储在数字储存介质上或者可在传输介质上传输,例如,无线传输介质或有线传输介质,诸如,互联网。

[0221] 根据某些实施要求,本发明的实施方式可以在硬件中或者在软件中实施。可使用数字储存介质(例如,软盘、DVD、CD、ROM、PROM、EPROM、EEPROM或FLASH执行实施),在所述数字储存介质上存储了电子可读的控制信号,其与可编程计算机系统协作(或者能够协作),以便执行相应的方法。

[0222] 根据本发明的一些实施方式包括非瞬时性数据载体,其具有电子可读控制信号,这些信号能够与可编程计算机系统协和,以便执行在本文中描述的一种方法。

[0223] 通常,本发明的实施方式可实施为具有程序代码的计算机程序产品,当计算机程序产品在计算机上运行时,该程序代码可操作为执行一种方法。例如,该程序代码可存储在机器可读载体上。

[0224] 其他实施方式包括存储在机器可读载体上的用于执行在本文中描述的一种方法的计算机程序。

[0225] 换言之,因此,本发明的方法的实施方式是具有程序代码的计算机程序产品,当计算机程序在计算机上运行时,所述程序代码用于执行在本文中描述的一种方法。

[0226] 因此,本发明的方法的进一步实施方式是数据载体(或数字存储介质或计算机可读介质),其包括存记录在其上的用于执行在本文中描述的一种方法的计算机程序

[0227] 因此,本发明的方法的进一步实施方式是数据流或信号序列,其表示用于执行在本文中描述的一种方法的计算机程序。例如,数据流或信号序列可被配置为通过数据通信连接(例如,通过互联网)传输。

[0228] 进一步实施方式包括处理器件,例如,计算机或可编程逻辑装置,其被配置为或者适配为执行在本文中描述的一种方法。

[0229] 进一步实施方式包括具有安装在其上的计算机程序的计算机,所述计算机程序用于执行在本文中描述的一种方法。

[0230] 在一些实施方式中,可编程逻辑装置(例如,现场可编程门阵列)可用于执行在本文中描述的方法的一些或所有功能。在一些实施方式中,现场可编程门阵列可与微处理器协作,以便执行在本文中描述的一种方法。通常,这些方法优选地由任何硬件设备执行。

[0231] 上述实施方式仅仅说明本发明的原理。要理解的是,在本文中描述的设置和细节的修改和变化对于本领域的技术人员将是显而易见的。因此,其目的在于,仅受到待审专利权利要求的范围的限制,而不受到通过在本文中的实施方式的描述和解释呈现的具体细节的限制。

[0232] 参考文献

[0233] [1] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Berlin, Germany: Springer-Verlag, 2008.

[0234] [2] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming for hearing aid applications," in *Handbook on Array Processing and Sensor Networks*, S. Haykin and K. Ray Liu, Eds. Wiley, 2008, ch. 9.

[0235] [3] S. Gannot and I. Cohen, "Adaptive beamforming and postfiltering," in *Springer Handbook of Speech Processing*, J. Benesty, M. M. Sondhi, and Y. Huang, Eds. Springer-Verlag, 2008, ch. 47.

[0236] [4] J. Benesty, J. Chen, and E. A. P. Habets, *Speech Enhancement in the STFT Domain*, ser. SpringerBriefs in Electrical and Computer Engineering. Springer-Verlag, 2011.

[0237] [5] I. Tashev, M. Seltzer, and A. Acero, "Microphone array for headset with spatial noise suppressor," in *Proc. Ninth International Workshop on Acoustic, Echo and Noise Control (IWAENC)*, Eindhoven, The Netherlands, 2005.

[0238] [6] M. Kallinger, G. Del Galdo, F. Kuech, D. Mahne, and R. Schultz-Amling, "Spatial filtering using directional audio coding parameters," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2009, pp. 217-220.

[0239] [7] M. Kallinger, G. D. Galdo, F. Kuech, and O. Thiergart, "Dereverberation in the spatial audio coding domain," in *Audio Engineering Society Convention 130*, London UK, May 2011.

[0240] [8] G. Del Galdo, O. Thiergart, T. Weller, and E. A. P. Habets, "Generating

virtual microphone signals using geometrical information gathered by distributed arrays,” in Proc. Hands-Free Speech Communication and Microphone Arrays (HSCMA), Edinburgh, United Kingdom, May 2011.

[0241] [9] S. Nordholm, I. Claesson, and B. Bengtsson, “Adaptive array noise suppression of handsfree speaker input in cars,” *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 514-518, Nov. 1993.

[0242] [10] O. Hoshuyama, A. Sugiyama, and A. Hirano, “A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters,” *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2677-2684, Oct. 1999.

[0243] [11] S. Gannot, D. Burshtein, and E. Weinstein, “Signal enhancement using beamforming and nonstationarity with applications to speech,” *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1614-1626, Aug. 2001.

[0244] [12] W. Herbordt and W. Kellermann, “Adaptive beamforming for audio signal acquisition,” in *Adaptive Signal Processing: Applications to real-world problems*, ser. Signals and Communication Technology, J. Benesty and Y. Huang, Eds. Berlin, Germany: Springer-Verlag, 2003, ch. 6, pp. 155-194.

[0245] [13] R. Talmon, I. Cohen, and S. Gannot, “Convolutional transfer function generalized sidelobe canceler,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 7, pp. 1420-1434, Sep. 2009.

[0246] [14] A. Krueger, E. Warsitz, and R. Haeb-Umbach, “Speech enhancement with a GSC-like structure employing eigenvector-based transfer function ratios estimation,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 1, pp. 206-219, Jan. 2011.

[0247] [15] E. A. P. Habets and J. Benesty, “Joint dereverberation and noise reduction using a two-stage beamforming approach,” in Proc. Hands-Free Speech Communication and Microphone Arrays (HSCMA), 2011, pp. 191-195.

[0248] [16] M. Taseska and E. A. P. Habets, “MMSE-based blind source extraction in diffuse noise fields using a complex coherence-based a priori SAP estimator,” in Proc. Intl. Workshop Acoust. Signal Enhancement (IWAENC), Sep. 2012.

[0249] [17] G. Reuven, S. Gannot, and I. Cohen, “Dual source transfer-function generalized sidelobe canceller,” *IEEE Trans. Speech Audio Process.*, vol. 16, no. 4, pp. 711-727, May 2008.

[0250] [18] S. Markovich, S. Gannot, and I. Cohen, “Multichannel eigenspace beamforming in a reverberant noisy environment with multiple interfering speech signals,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 6, pp. 1071-1086, Aug. 2009.

[0251] [19] O. Thiergart and E. A. P. Habets, “Sound field model violations in parametric spatial sound processing,” in Proc. Intl. Workshop Acoust. Signal Enhancement (IWAENC), Sep. 2012.

- [0252] [20] R.K.Cook, R.V.Waterhouse, R.D.Berendt, S.Edelman, and M.C.Thompson Jr., "Measurement of correlation coefficients in reverberant sound fields," The Journal of the Acoustical Society of America, vol.27, no.6, pp.1072-1077, 1955.
- [0253] [21] O.L.Frost, III, "An algorithm for linearly constrained adaptive array processing," Proc.IEEE, vol.60, no.8, pp.926-935, Aug.1972.
- [0254] [22] R.Roy and T.Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," Acoustics, Speech and Signal Processing, IEEE Transactions on, vol.37, no.7, pp.984-995, July 1989.
- [0255] [23] B.Rao and K.Hari, "Performance analysis of root-music*," in Signals, Systems and Computers, 1988. Twenty-Second Asilomar Conference on, vol.2, 1988, pp.578-582.
- [0256] [24] H.Cox, R.M.Zeskind, and M.M.Owen, "Robust adaptive beamforming," IEEE Trans.Acoust., Speech, Signal Process., vol.35, no.10, pp.1365-1376, Oct.1987.
- [0257] [25] J.B.Allen and D.A.Berkley, "Image method for efficiently simulating small-room acoustics," J.Acoust.Soc.Am., vol.65, no.4, pp.943-950, Apr.1979.
- [0258] [26] E.A.P.Habets. (2008, May) Room impulse response (RIR) generator. [Online]. Available: <http://home.tiscali.nl/ehabets/rirgenerator.html>; 还见: http://web.archive.org/web/20120730003147/http://home.tiscali.nl/ehabets/rir_generator.html
- [0259] [27] I.CVX Research, "CVX: Matlab software for disciplined convex programming, version 2.0beta," <http://cvxr.com/cvx>, September 2012.
- [0260] [28] M.Grant and S.Boyd, "Graph implementations for nonsmooth convex programs," in Recent Advances in Learning and Control, ser.Lecture Notes in Control and Information Sciences, V.Blondel, S.Boyd, and H.Kimura, Eds.Springer-Verlag Limited, 2008, pp.95-110.
- [0261] [29] H.L.Van Trees, Detection, Estimation, and Modulation Theory: Part IV: Optimum Array Processing. John Wiley&Sons, April 2002, vol.1.

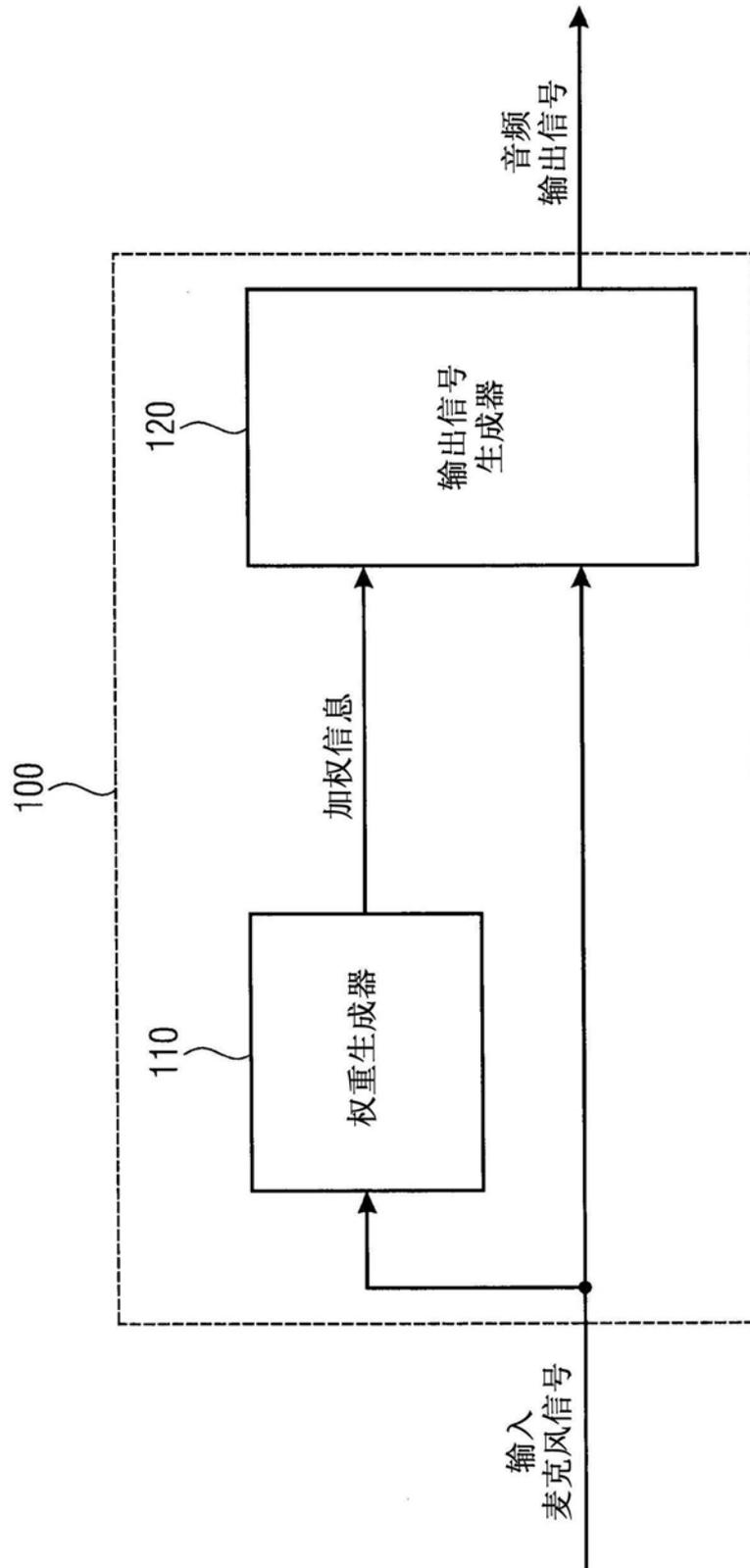


图1A

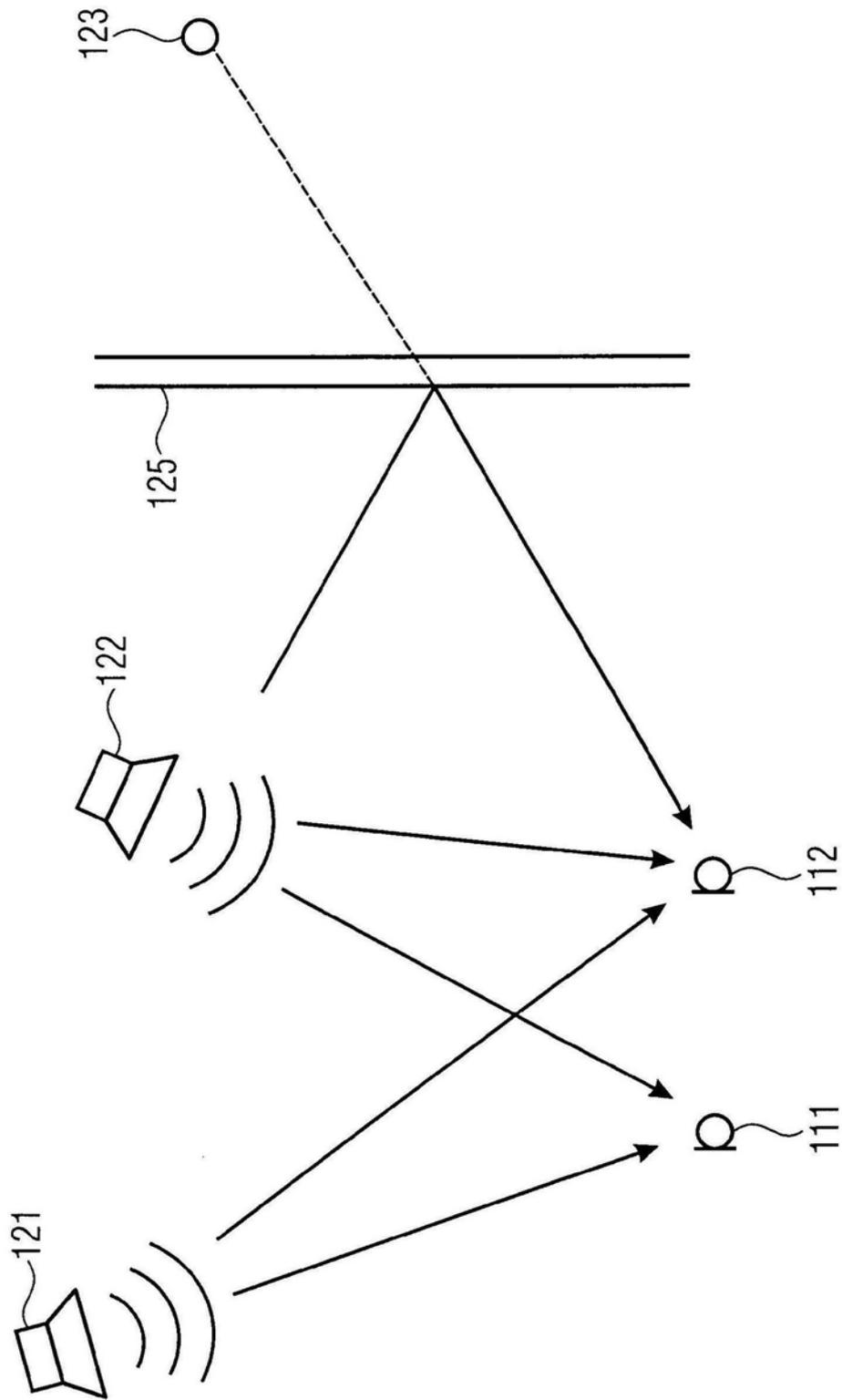


图1B

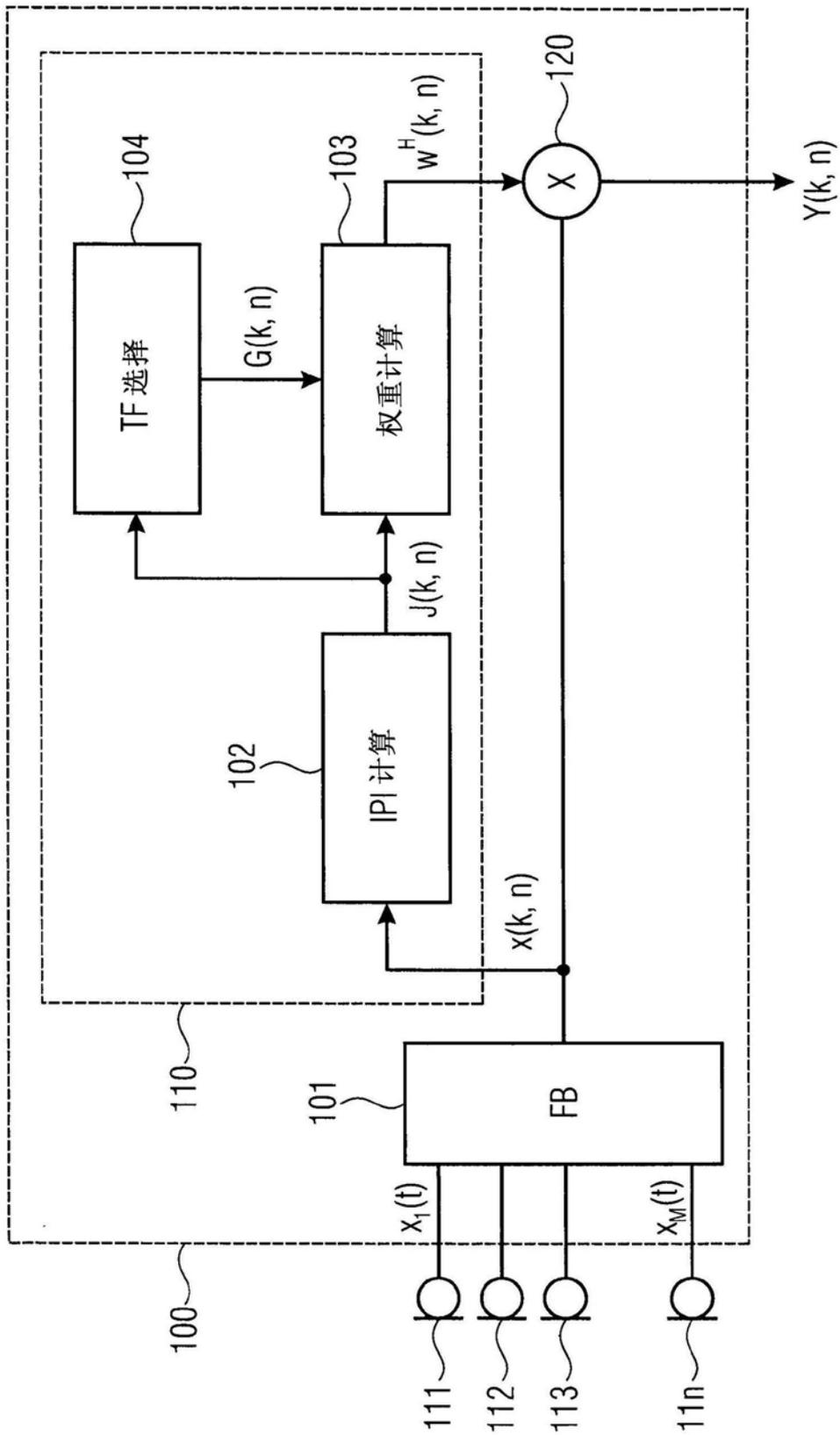


图2

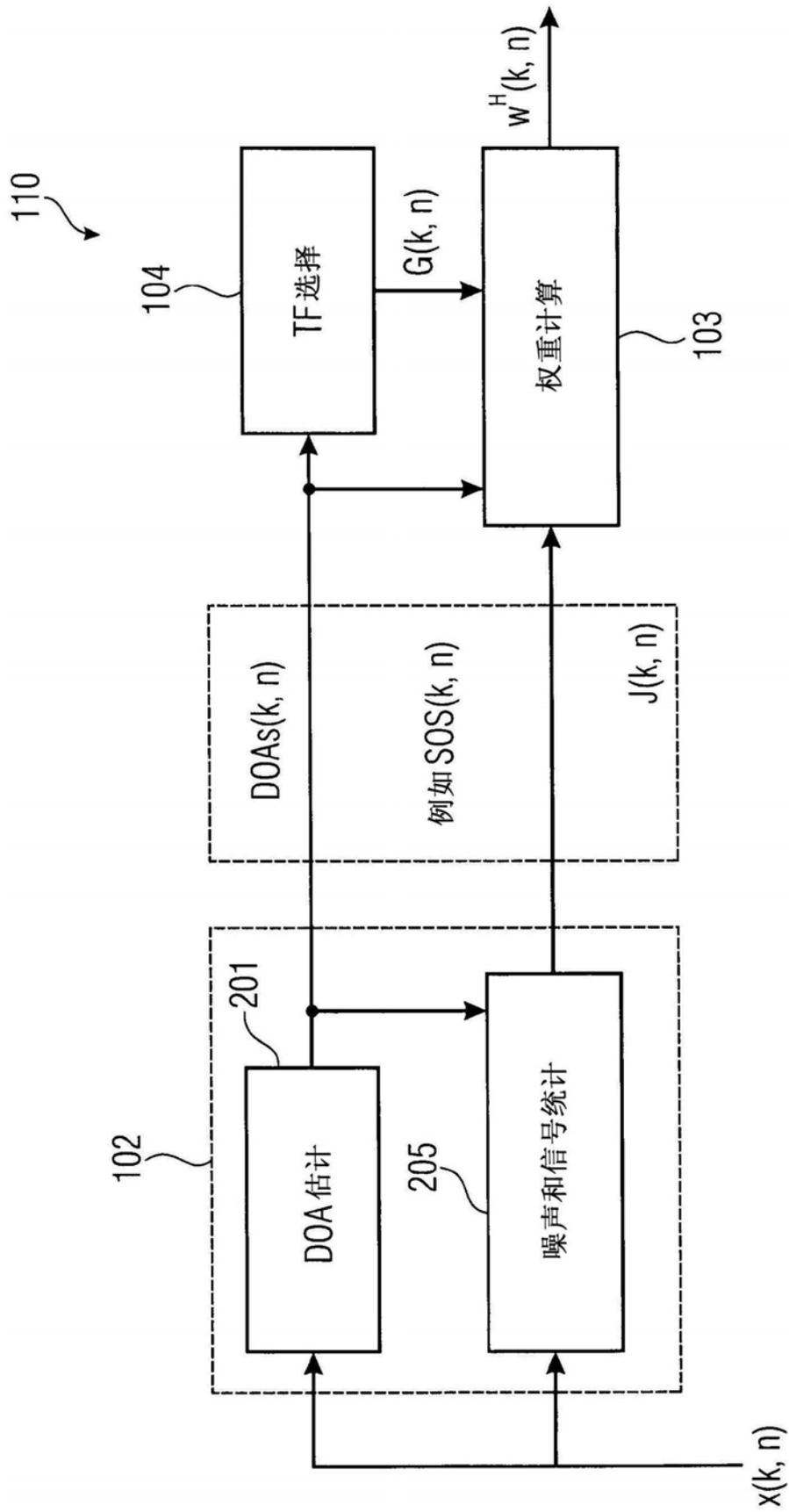


图3

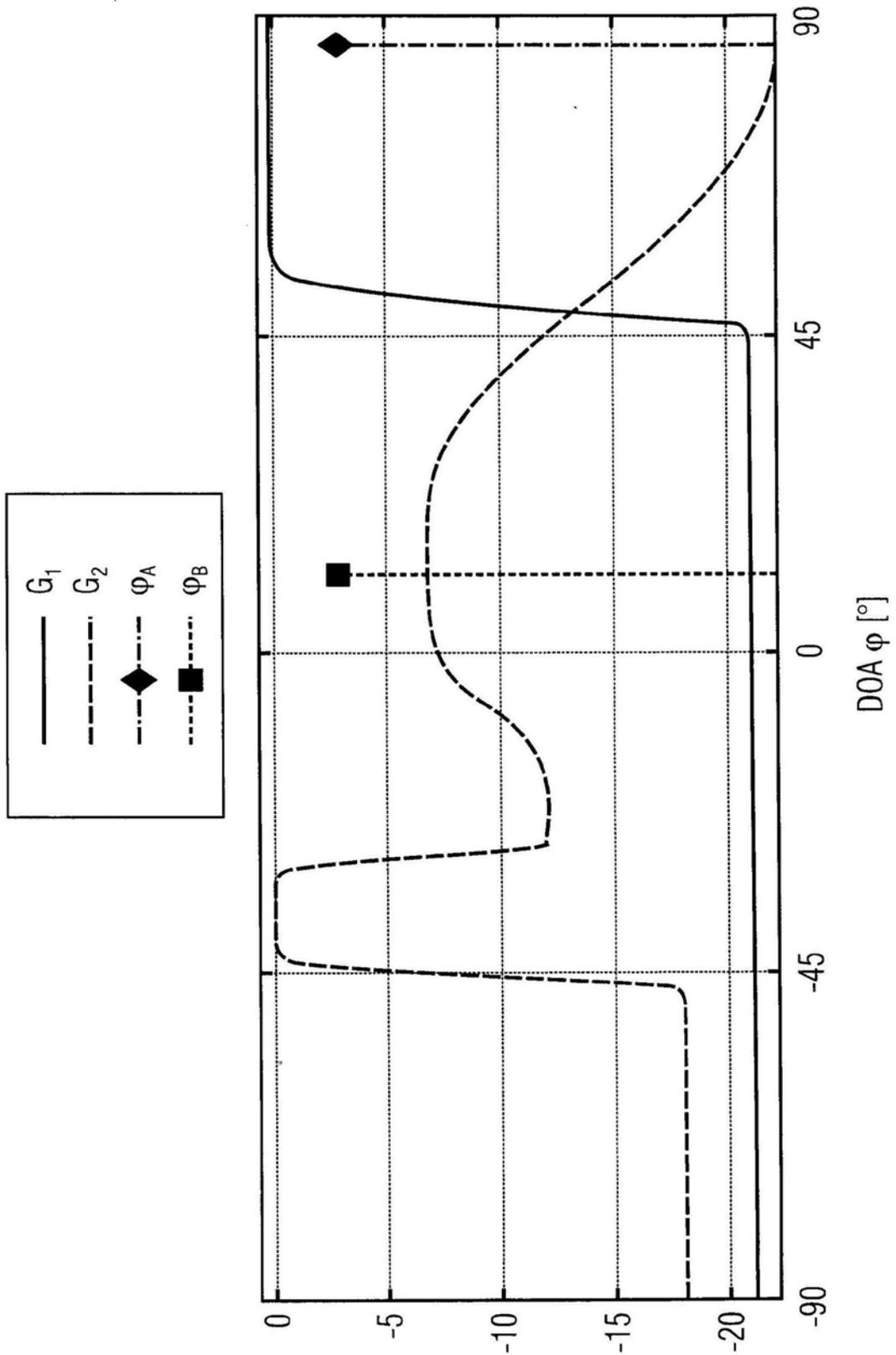


图4

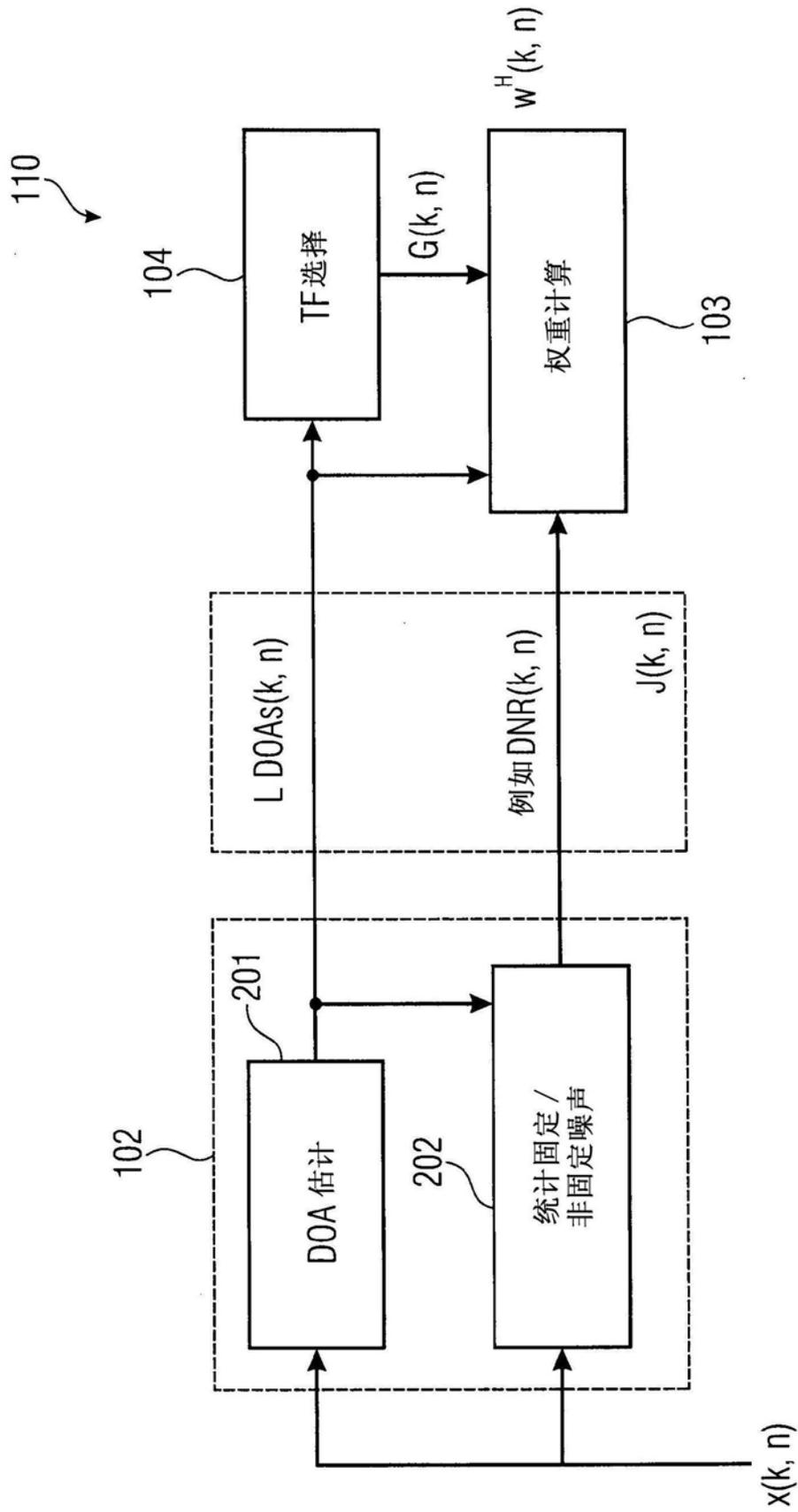


图5

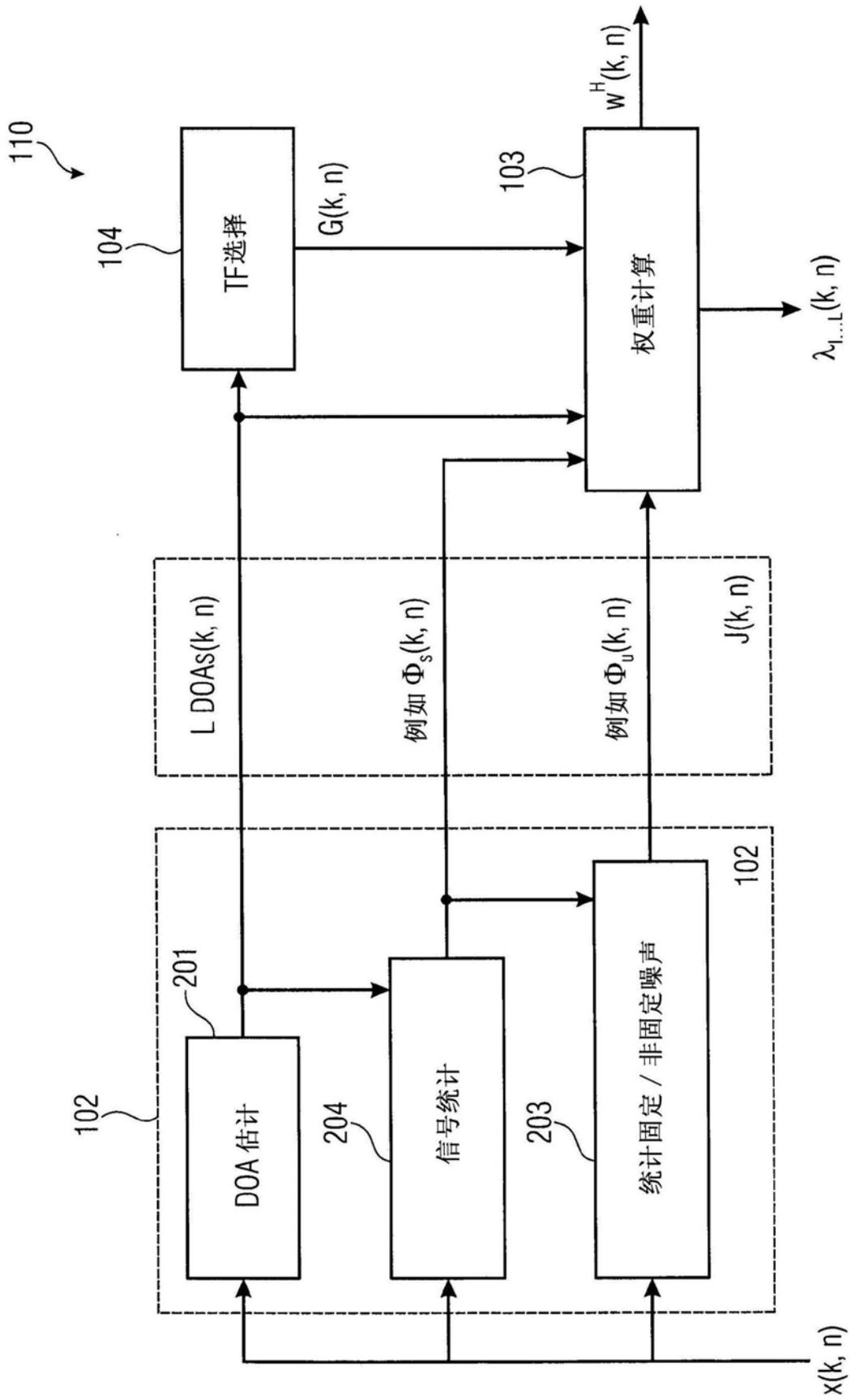


图6

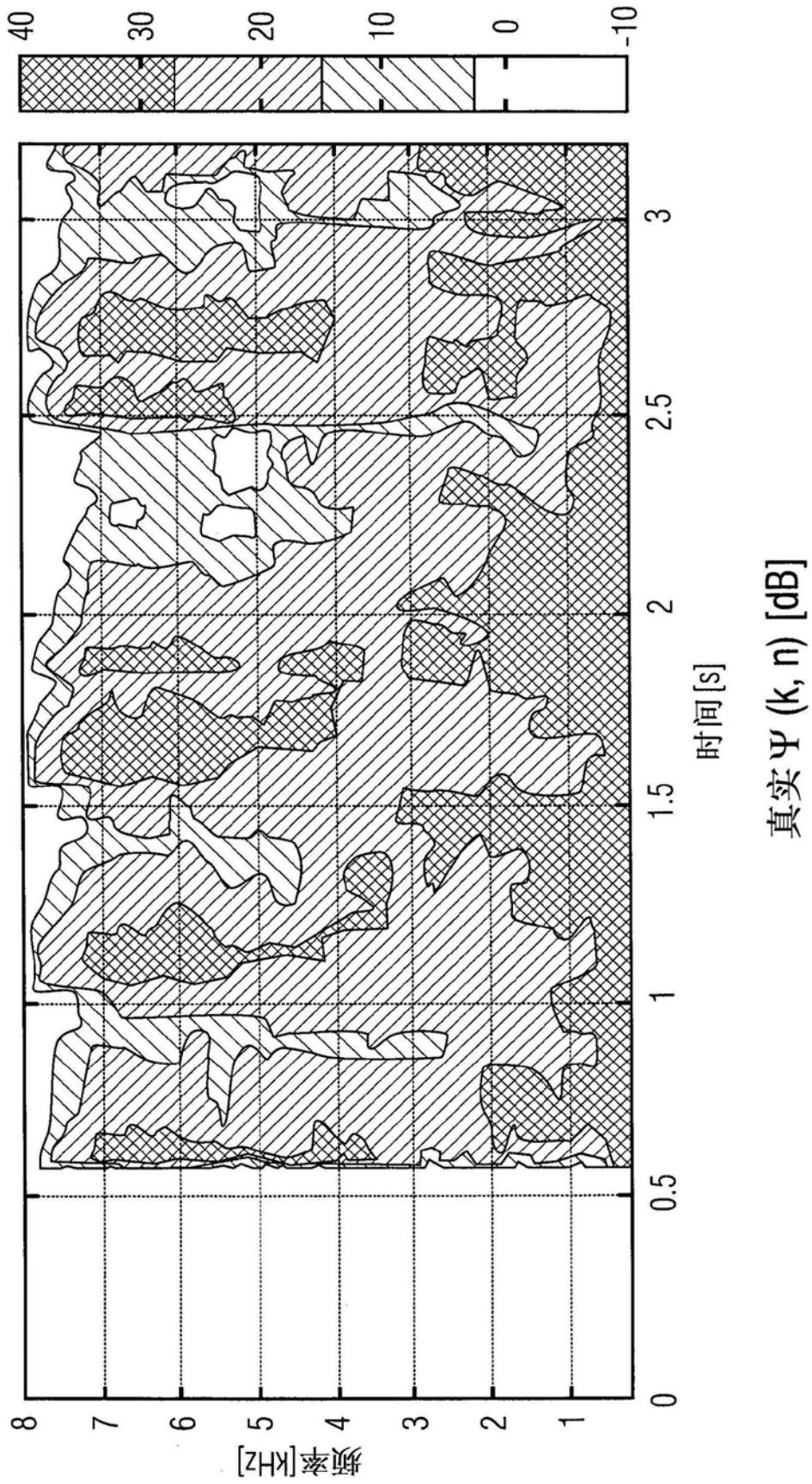


图7A

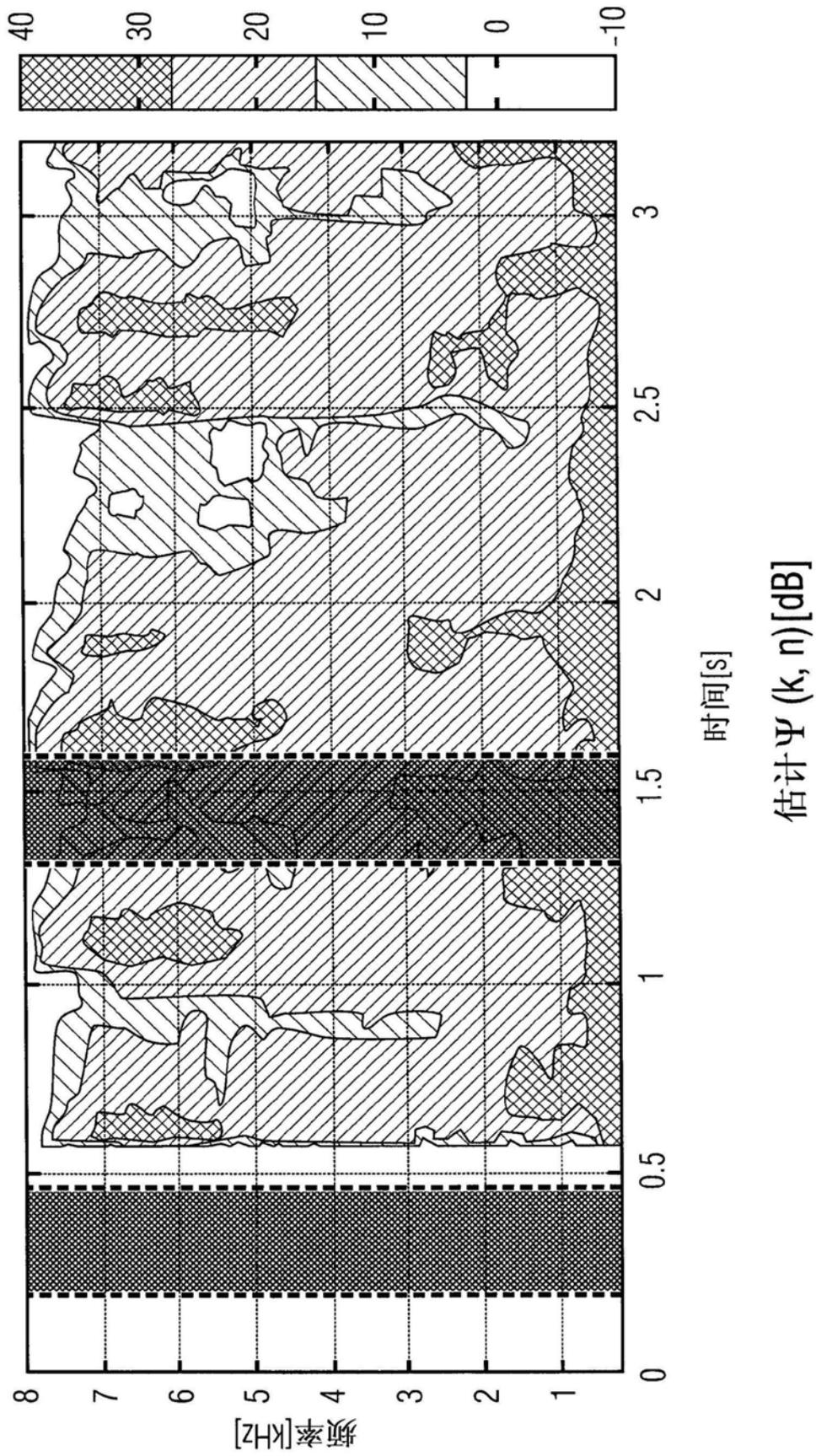


图7B

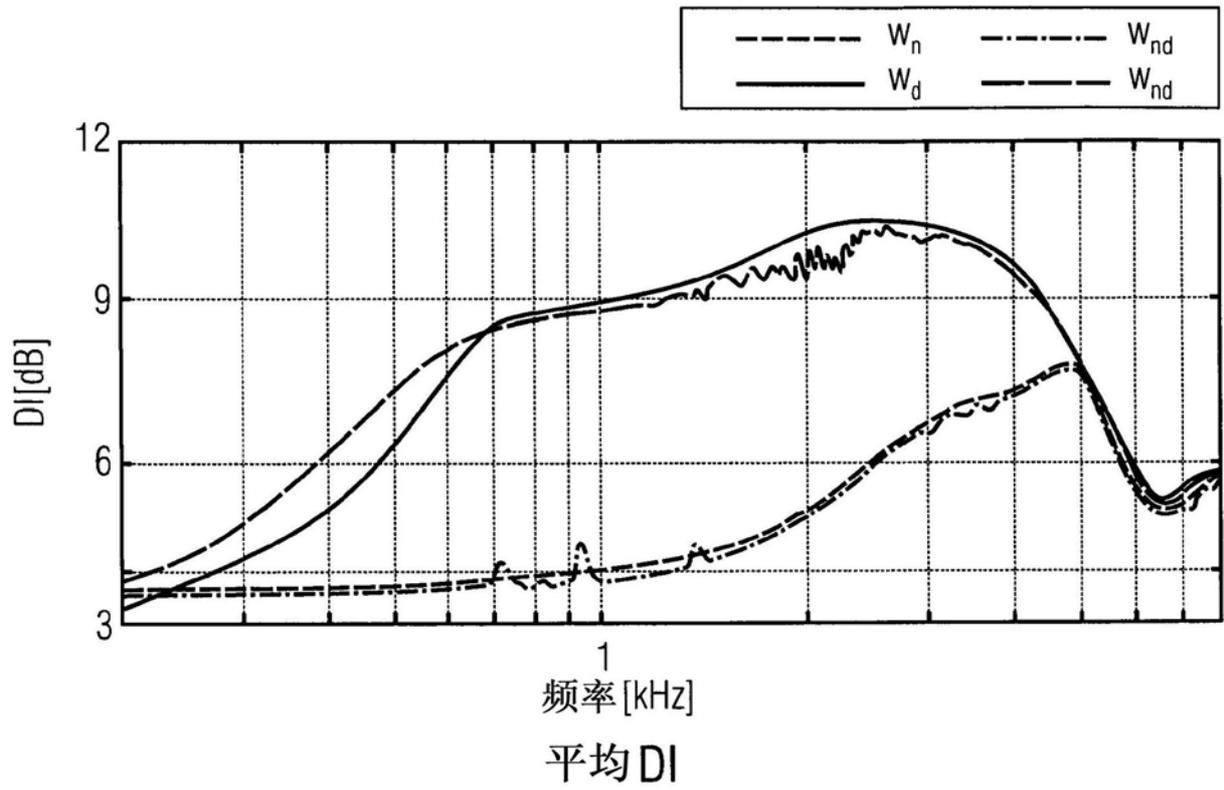


图8A

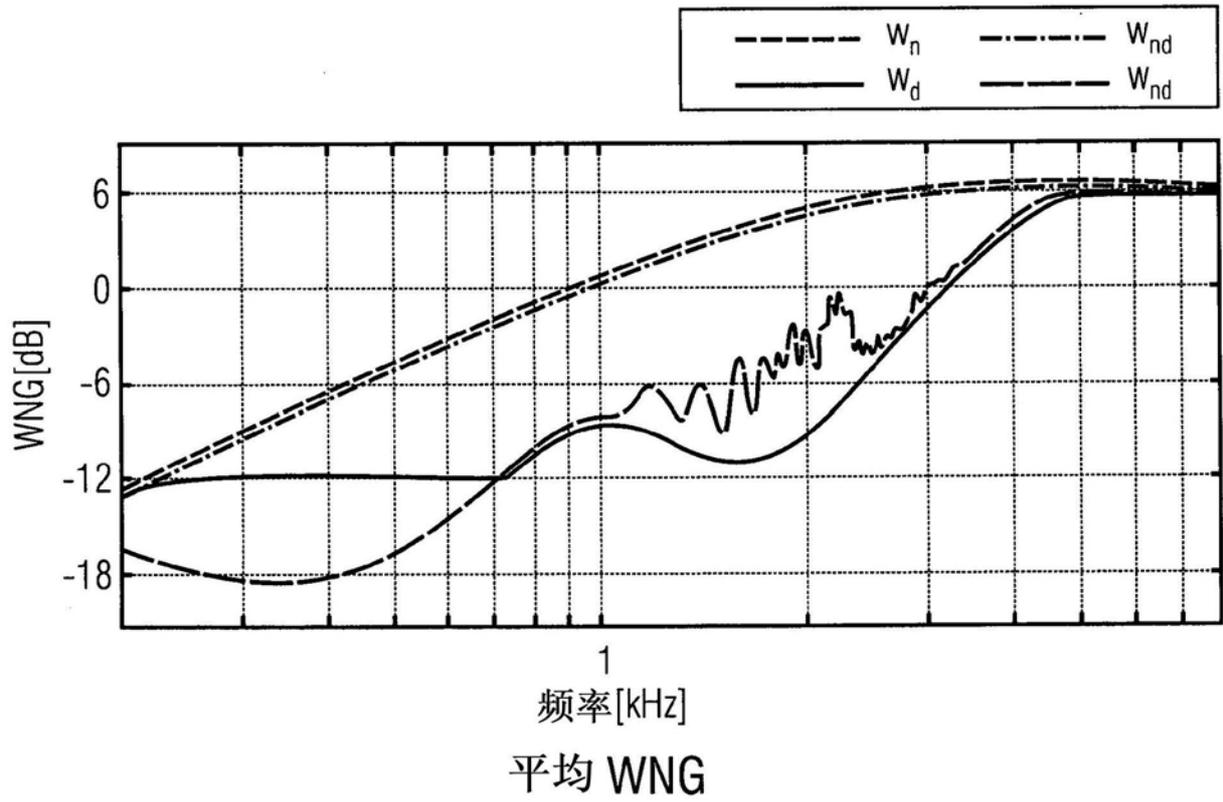


图8B

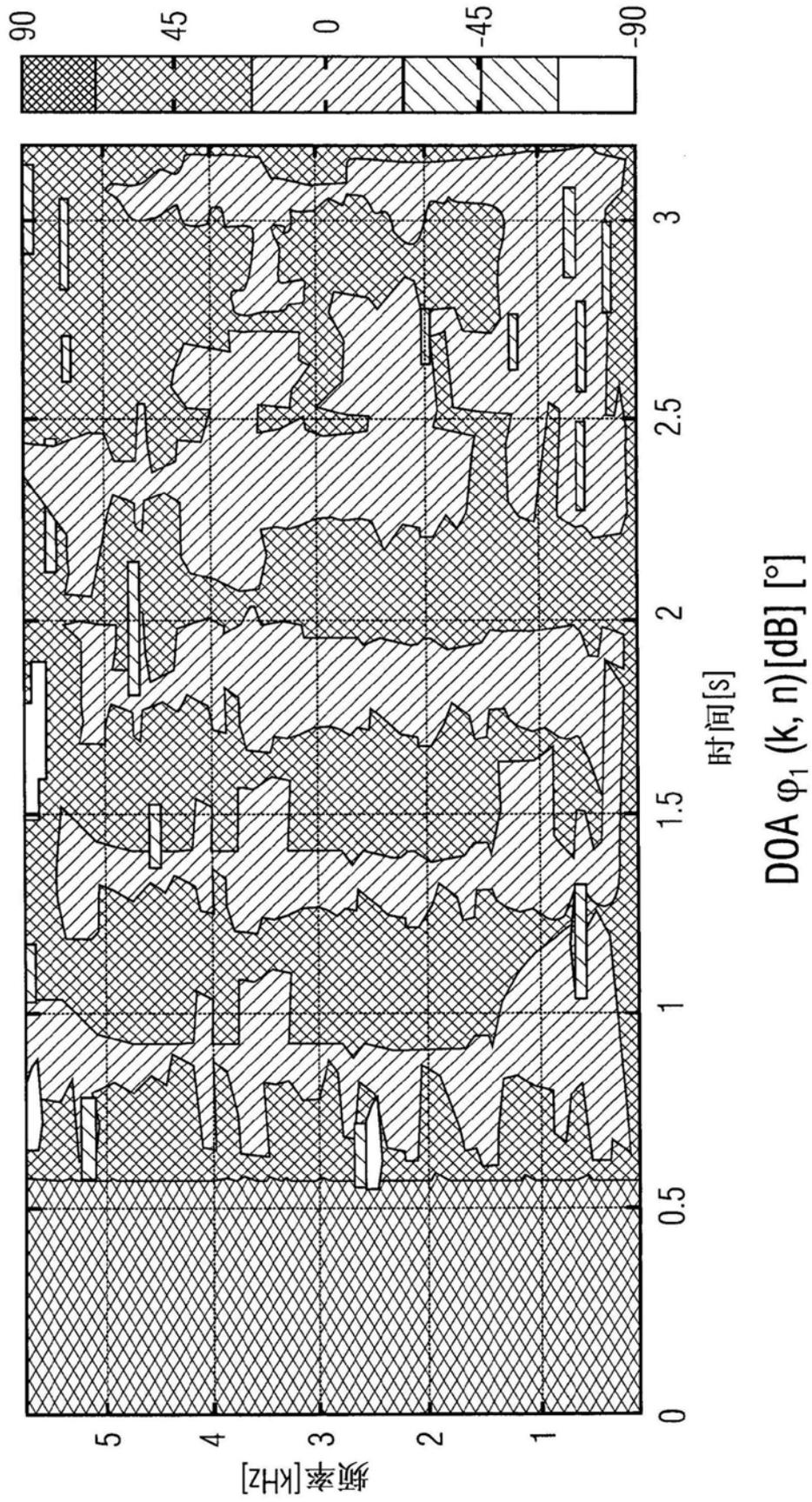


图9A

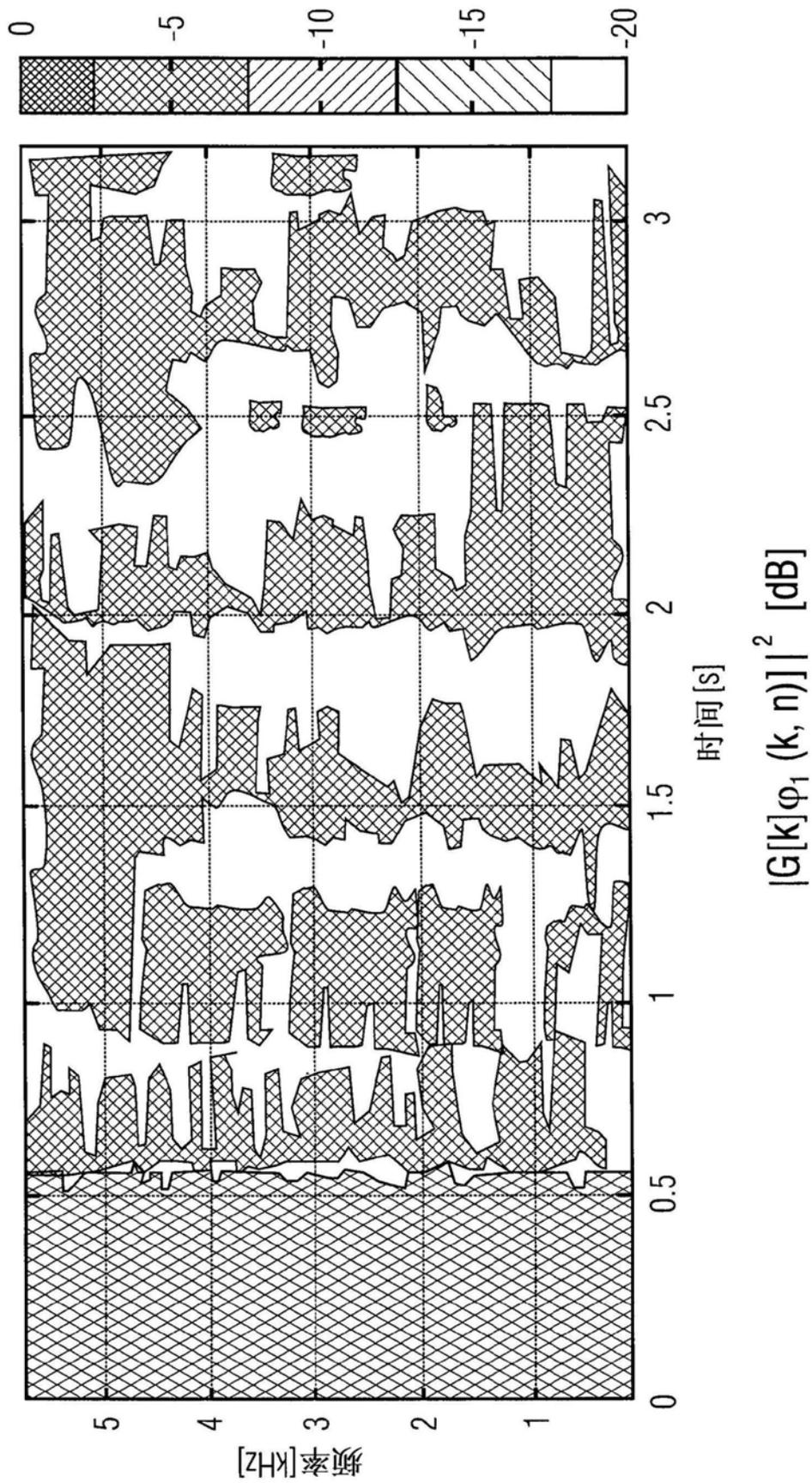


图9B

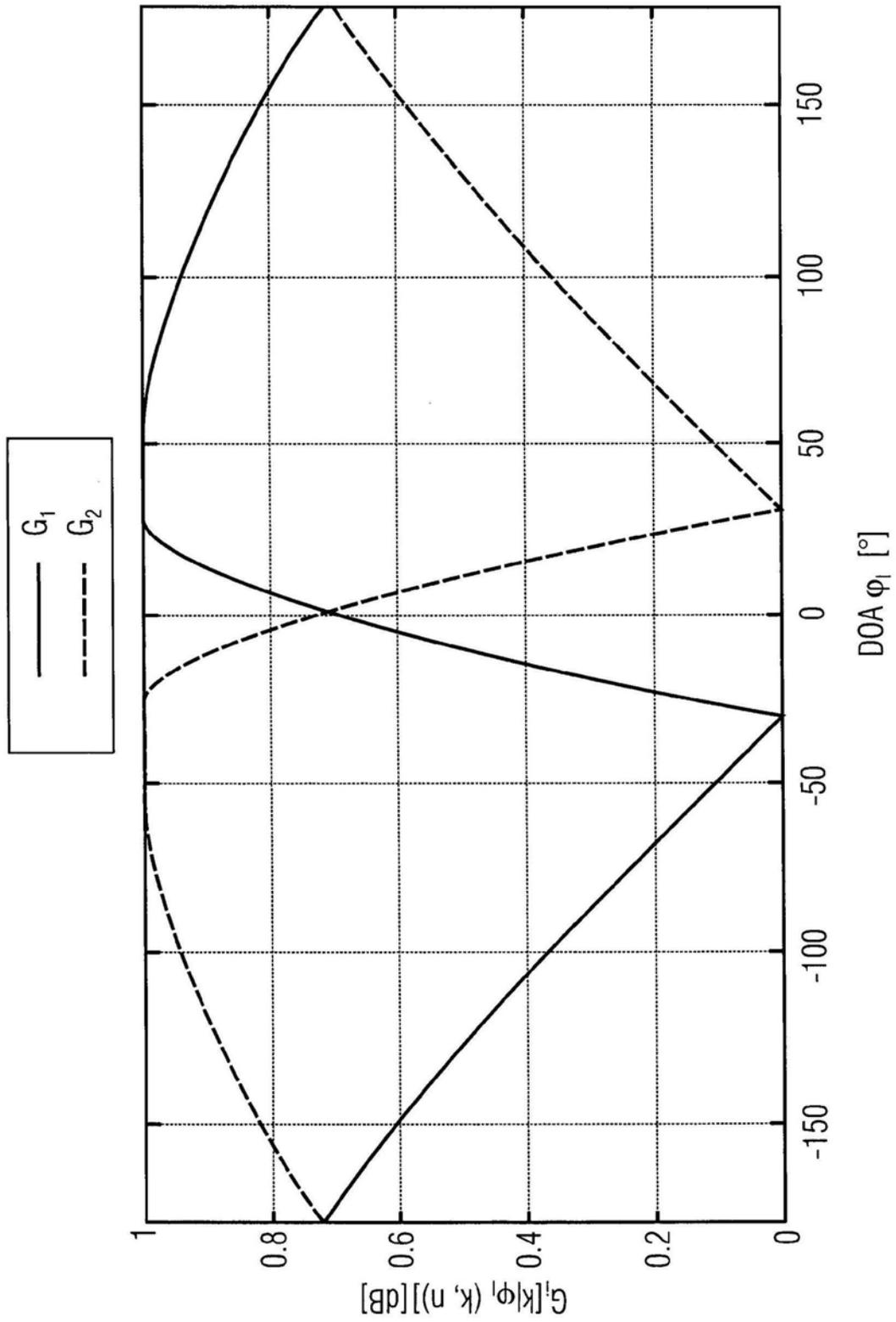


图10