(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 106664501 A
(43)申请公布日 2017.05.10

(21)申请号 201580036158.7

(22)申请日 2015.04.23

(30)优先权数据
14167053.9 2014.05.05 EP
14183855.7 2014.09.05 EP

(85)PCT国际申请进入国家阶段日
2016.12.30

(86)PCT国际申请的申请数据
PCT/EP2015/058859 2015.04.23

(87)PCT国际申请的公布数据
WO2015/169618 EN 2015.11.12

(71)申请人 弗劳恩霍夫应用研究促进协会
地址 德国慕尼黑

(72)发明人 伊曼纽尔·哈比兹
奥利弗·迪尔加特
科纳德·科瓦奇克

(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 闫睥

(51)Int.Cl.
H04S 7/00(2006.01)

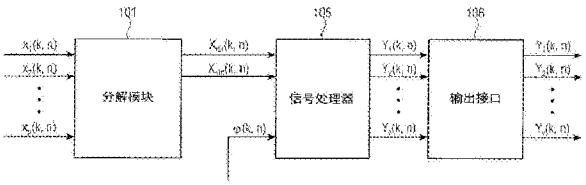
权利要求书4页 说明书34页 附图14页

(54)发明名称

基于所通知的空间滤波的一致声学场景再现的系统、装置和方法

(57)摘要

提供了一种用于生成一个或更多个音频输出信号的系统。该系统包括分解模块(101)、信号处理器(105)和输出接口(106)。分解模块(101)被配置为接收两个或更多个音频输入信号，其中分解模块(101)被配置为生成包括两个或更多个音频输入信号的直达信号分量在内的直达分量信号，并且其中分解模块(101)被配置为生成包括所述两个或更多个音频输入信号的扩散信号分量在内的扩散分量信号。信号处理器(105)被配置为接收直达分量信号、扩散分量信号和方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向。此外，信号处理器(105)被配置为根据扩散分量信号生成一个或更多个经处理的扩散信号。对于一个或更多个音频输出信号的每个音频输出信号，信号处理器(105)被配置为根据到达方向确定直达增益，并且信号处理器(105)被配置为将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，且所述信号处理器(105)被配置为将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个扩散信号进行组合以生成所述音频输出信号。输出接口(106)被配置为输出所述一个或更多个音频输出信号。



CN 106664501 A

1.一种用于生成一个或更多个音频输出信号的系统，包括：

分解模块(101)；

信号处理器(105)；以及

输出接口(106)，

其中分解模块(101)被配置为接收两个或更多个音频输入信号，其中分解模块(101)被配置为生成包括两个或更多个音频输入信号的直达信号分量在内的直达分量信号，并且其中分解模块(101)被配置为生成包括所述两个或更多个音频输入信号的扩散信号分量在内的扩散分量信号，

其中信号处理器(105)被配置为接收直达分量信号、扩散分量信号和方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向，

其中信号处理器(105)被配置为根据扩散分量信号生成一个或更多个经处理的扩散信号，

其中，对于所述一个或更多个音频输出信号中的每个音频输出信号，信号处理器(105)被配置为根据到达方向确定直达增益，并且信号处理器(105)被配置为将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，并且信号处理器(105)被配置为将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个进行组合以生成所述音频输出信号，以及

其中输出接口(106)被配置为输出所述一个或更多个音频输出信号。

2.根据权利要求1所述的系统，

其中信号处理器(105)被配置为确定两个或更多个音频输出信号，

其中，对于所述两个或更多个音频输出信号的每个音频输出信号，将平移增益函数分配给所述音频输出信号，

其中，所述两个或更多个音频输出信号中的每一个的平移增益函数包括多个平移函数自变量值，其中平移函数返回值被分配给所述平移函数自变量值中的每一个，其中，当所述平移增益函数接收到所述平移函数自变量值中的一个值时，所述平移增益函数被配置为返回被分配给所述平移函数自变量值中的所述一个值的平移函数返回值，以及

其中，信号处理器(105)被配置为根据分配给所述音频输出信号的平移增益函数的平移函数自变量值中的取决于方向的自变量值来确定所述两个或更多个音频输出信号中的每一个，其中所述取决于方向的自变量值取决于到达方向。

3.根据权利要求2所述的系统，

其中所述两个或更多个音频输出信号中的每一个的平移增益函数具有作为平移函数自变量值之一的一个或更多个全局最大值，其中对于每个平移增益函数的一个或更多个全局最大值中的每一个，不存在使得所述平移增益函数返回比所述全局最大值使所述平移增益函数返回的增益函数返回值更大的平移函数返回值的其他平移函数自变量值，以及

其中对于所述两个或更多个音频输出信号中的第一音频输出信号和第二音频输出信号的每对，第一音频输出信号的平移增益函数的一个或更多个全局最大值中的至少一个不同于第二音频输出信号的平移增益函数的一个或更多个全局最大值中的任一个。

4.根据权利要求2或3所述的系统，

其中信号处理器(105)被配置为根据窗增益函数来生成所述一个或更多个音频输出信

号的每个音频输出信号，

其中窗增益函数被配置为在接收到窗函数自变量值时返回窗函数返回值，

其中，如果窗函数自变量值大于下窗阈值并且小于上窗阈值，则窗增益函数被配置为返回比在窗函数自变量值小于下阈值或大于上阈值的情况下由所述窗增益函数返回的任何窗函数返回值大的窗函数返回值。

5.根据权利要求2至4中任一项所述的系统，其中信号处理器(105)被配置为进一步接收指示观看方向相对于到达方向的角位移的取向信息，并且其中平移增益函数和窗增益函数中的至少一个取决于所述取向信息，或者

其中增益函数计算模块(104)被配置为进一步接收缩放信息，其中所述缩放信息指示相机的打开角度，并且平移增益函数和窗增益函数中的至少一个取决于所述缩放信息，或者

其中增益函数计算模块(104)被配置为进一步接收校准参数，并且平移增益函数和窗增益函数中的至少一个取决于所述校准参数。

6.根据前述权利要求中的任一项所述的系统，

其中信号处理器(105)被配置为接收距离信息，

其中信号处理器(105)被配置为根据距离信息来生成所述一个或更多个音频输出信号中的每个音频输出信号。

7.根据权利要求6所述的系统，

其中信号处理器(105)被配置为接收取决于原始到达方向的原始角度值并且被配置为接收距离信息，所述原始到达方向是所述两个或更多个音频输入信号的直达信号分量的到达方向，

其中信号处理器(105)被配置为根据原始角度值并根据距离信息来计算修改的角度值，以及

其中信号处理器(105)被配置为根据修改的角度值来生成所述一个或更多个音频输出信号中的每个音频输出信号。

8.根据权利要求6或7所述的系统，其中信号处理器(105)被配置为通过进行低通滤波、或通过添加延迟的直达声音、或通过进行直达声音衰减、或通过进行时间平滑、或者通过进行到达方向扩展、或通过进行去相关来生成所述一个或更多个音频输出信号。

9.根据前述权利要求中的任一项所述的系统，

其中信号处理器(105)被配置为生成两个或更多个音频输出声道，

其中信号处理器(105)被配置为对扩散分量信号应用扩散增益以获得中间扩散信号，以及

其中信号处理器(105)被配置为通过执行去相关，从所述中间扩散信号中生成一个或更多个去相关信号，

其中所述一个或更多个去相关信号形成所述一个或更多个经处理的扩散信号，或所述中间扩散信号和所述一个或更多个去相关信号形成所述一个或更多个经处理的扩散信号。

10.根据前述权利要求中的任一项所述的系统，

其中直达分量信号和一个或更多个另外的直达分量信号形成两个或更多个直达分量信号的组，其中分解模块(101)被配置为生成所述一个或更多个另外的直达分量信号，所述

一个或更多个另外的直达分量信号包括所述两个或更多个音频输入信号的另外的直达信号分量，

其中到达方向和一个或更多个另外的到达方向形成两个或更多个到达方向的组，其中所述两个或更多个到达方向的组中的每个到达方向被分配给所述两个或更多个直达分量信号的组中的恰好一个直达分量信号，其中所述两个或更多个直达分量信号的直达分量信号的数量与所述两个到达方向的到达方向的数量相等，

其中信号处理器(105)被配置为接收所述两个或更多个直达分量信号的组以及所述两个或更多个到达方向的组，以及

其中，对于所述一个或更多个音频输出信号中的每个音频输出信号，

信号处理器(105)被配置为针对所述两个或更多个直达分量信号的组的每个直达分量信号，取决于所述直达分量信号的到达方向确定直达增益，

信号处理器(105)被配置为通过针对所述两个或更多个直达分量信号的组的每个直达分量信号将所述直达分量信号的直达增益应用于所述直达分量信号，来生成两个或更多个经处理的直达信号的组，以及

信号处理器(105)被配置为将所述一个或更多个经处理的扩散信号中的一个与所述两个或更多个经处理的信号的组中的每个经处理的信号进行组合，以生成所述音频输出信号。

11.根据权利要求10所述的系统，其中所述两个或更多个直达分量信号的组中的直达分量信号的数量加1小于由接收接口(101)接收的音频输入信号的数量。

12.一种包括根据权利要求1至11中任一项所述的系统在内的助听器或助听设备。

13.一种用于生成一个或更多个音频输出信号的装置，包括：

信号处理器(105)；以及

输出接口(106)，

其中，信号处理器(105)被配置为接收包括所述两个或更多个原始音频信号的直达信号分量在内的直达分量信号，其中信号处理器(105)被配置为接收包括所述两个或更多个原始音频信号的扩散信号分量在内的扩散分量信号，并且其中信号处理器(105)被配置为接收方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向，

其中信号处理器(105)被配置为根据扩散分量信号生成一个或更多个经处理的扩散信号，

其中，对于所述一个或更多个音频输出信号中的每个音频输出信号，信号处理器(105)被配置为根据到达方向确定直达增益，并且信号处理器(105)被配置为将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，并且信号处理器(105)被配置为将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个进行组合以生成所述音频输出信号，以及

其中输出接口(106)被配置为输出所述一个或更多个音频输出信号。

14.一种用于生成一个或更多个音频输出信号的方法，包括：

接收两个或更多个音频输入信号，

生成包括所述两个或更多个音频输入信号的直达信号分量在内的直达分量信号，

生成包括所述两个或更多个音频输入信号的扩散信号分量在内的扩散分量信号，

接收取决于所述两个或更多个音频输入信号的直达信号分量的到达方向的方向信息，

根据扩散分量信号生成一个或更多个经处理的扩散信号，

对于一个或更多个音频输出信号中的每个音频输出信号，根据到达方向确定直达增益，将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，以及将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个进行组合以生成所述音频输出信号，以及

输出所述一个或更多个音频输出信号。

15.一种用于生成一个或更多个音频输出信号的方法，包括：

接收包括所述两个或更多个原始音频信号的直达信号分量在内的直达分量信号，

接收包括所述两个或更多个原始音频信号的扩散信号分量在内的扩散分量信号，

接收方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向，

根据扩散分量信号生成一个或更多个经处理的扩散信号，

对于一个或更多个音频输出信号中的每个音频输出信号，根据到达方向确定直达增益，将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，以及将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个进行组合以生成所述音频输出信号，以及

输出所述一个或更多个音频输出信号。

16.一种计算机程序，用于在计算机或信号处理器上执行时实施根据权利要求14或15所述的方法。

基于所通知的空间滤波的一致声学场景再现的系统、装置和方法

技术领域

[0001] 本发明涉及音频信号处理，具体地，涉及用于基于所通知的空间滤波的一致声学场景再现的系统、装置和方法。

背景技术

[0002] 在空间声音再现中，利用多个麦克风捕获记录位置(近端侧)处的声音，然后使用多个扬声器或耳机在再现侧(远端侧)再现。在许多应用中，期望再现所记录的声音，使得在远端侧重建的空间图像与在近端侧的原始空间图像一致。这意味着例如声源的声音从源存在于原始记录场景中的方向再现。备选地，当例如视频对所记录的音频进行补充时，期望再现声音，使得重建的声学图像与视频图像一致。这意味着例如声源的声音从源在视频中可见的方向再现。另外，视频相机可以配备有视觉缩放功能，或者在远端侧的用户可以对视频应用数字缩放，从而改变视觉图像。在这种情况下，再现的空间声音的声学图像将相应地改变。在许多情况下，远端侧确定应与再现声音一致的空间图像在远端侧或在回放期间(例如当涉及视频图像时)被确定。因此，在近端侧的空间声音必须被记录、处理和传输，使得在远端侧，我们仍然可以控制重建的声学图像。

[0003] 在许多现代应用中需要再现与期望的空间图像一致的所记录的声学场景的可能性。例如，诸如数字相机或移动电话之类的现代消费者设备通常配备有视频相机和多个麦克风。这使得视频能够与空间声音(例如立体声)一起被记录。当与视频一起再现记录的音频时，期望视觉和声学图像是一致的。当用户用相机放大时，期望在声学上重新创建视觉缩放效果，使得在观看视频时视觉和声学图像是对齐的。例如，当用户放大人物时，随着人物看起来更靠近相机，该人物的声音的混响应越来越小。此外，人的语音应当从与人在视觉图像中出现的方向相同的方向再现。在下文中声学地模拟相机的视觉缩放被称为声学缩放，并且表示一致的音频-视频再现的一个示例。可能涉及声学缩放的一致的音频-视频再现在电视会议中也是有用的，其中近端侧的空间声音在远端侧与视觉图像一起再现。此外，期望以声学方式再现视觉缩放效果，使得视觉和声学图像对齐。

[0004] 声学缩放的第一种实现在[1]中提出，其中，通过增加二阶定向麦克风的方向性来获得缩放效果，二阶定向麦克风的信号是基于线性麦克风阵列的信号生成的。这种方法在[2]中被扩展到立体声缩放。在[3]中提出了最近的用于单声道或立体声缩放的方法，其包括改变声源水平，使得来自正面方向的源被保留，而来自其他方向的源和扩散声音被衰减。[1]、[2]中提出的方法导致直达与混响比(DRR)的增加，并且[3]中的方法额外地允许抑制不期望的源。上述方法假设声源位于相机的正面，但不旨在捕获与视频图像一致的声学图像。

[0005] 用于灵活的空间声音记录和再现的公知方法由定向音频编码(DirAC)表示[4]。在DirAC中，根据音频信号和参数辅助信息(即，声音的到达方向(DOA)和扩散性)来描述近端侧的空间声音。参数描述使得能够利用任意扬声器设置再现原始空间图像。这意味着在远

端侧的重建空间图像与在近端侧在记录期间的空间图像一致。然而，如果例如视频对记录的音频进行补充，则再现的空间声音不一定与视频图像对齐。此外，当视觉图像改变时，例如当相机的观看方向和缩放改变时，不能调整重建的声学图像。这意味着DirAC不提供将重建的声学图像调整为任意期望的空间图像的可能性。

[0006] 在[5]中，基于DirAC实现了声学缩放。DirAC表示实现声学缩放的合理基础，因为它基于简单而强大的信号模型，该模型假设时域-频域中的声场由单个平面波加扩散声音组成。基础模型参数(例如DOA和扩散)被用来分离直达声音和扩散声音，并产生声学缩放效果。空间声音的参数描述使得能够将声音场景有效地传输到远端侧，同时仍然向用户提供对缩放效果和空间声音再现的完全控制。即使DirAC使用多个麦克风来估计模型参数，也仅应用单声道滤波器来提取直达声音和扩散声音，从而限制了再现声音的质量。此外，假设声音场景中的所有源位于圆上，并且参考与视觉缩放不一致的音频-视觉相机的改变位置来执行空间声音再现。实际上，缩放改变了相机的视角，而到视觉对象的距离和它们在图像中的相对位置保持不变，这与移动相机相反。

[0007] 相关的方法是所谓的虚拟麦克风(VM)技术[6]、[7]，其考虑与DirAC相同的信号模型，但允许在声音场景中的任意位置合成不存在的(虚拟)麦克风的信号。将VM朝向声源移动类似于相机到新位置的移动。使用多声道滤波器来实现VM以提高声音质量，但需要若干分布式麦克风阵列来估计模型参数。

[0008] 然而，提供用于音频信号处理的进一步改进的构思是非常有利的。

发明内容

[0009] 因此，本发明的目的是提供用于音频信号处理的改进的构思。通过根据权利要求1所述的系统、根据权利要求13所述的装置、根据权利要求14所述的方法、根据权利要求15所述的方法以及根据权利要求16所述的计算机程序来实现本发明的目的。

[0010] 提供了一种用于生成一个或更多个音频输出信号的系统。所述系统包括分解模块、信号处理器和输出接口。分解模块被配置为接收两个或更多个音频输入信号，其中分解模块被配置为生成包括所述两个或更多个音频输入信号的直达信号分量在内的直达分量信号，并且其中分解模块被配置为生成包括所述两个或更多个音频输入信号的扩散信号分量在内的扩散分量信号。信号处理器被配置为接收直达分量信号、扩散分量信号和方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向。此外，信号处理器被配置为根据扩散分量信号生成一个或更多个经处理的扩散信号。对于一个或更多个音频输出信号的每个音频输出信号，信号处理器被配置为根据到达方向确定直达增益，并且信号处理器被配置为将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，且所述信号处理器被配置为将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个扩散信号进行组合以生成所述音频输出信号。输出接口被配置为输出所述一个或更多个音频输出信号。

[0011] 根据实施例，提供用于实现空间声音记录和再现的构思，使得重建的声学图像可以例如与期望的空间图像一致，所述期望的空间图像例如由用户在远端侧确定或者由视频图像确定。提出的方法在近端侧使用麦克风阵列，这允许我们将捕获的声音分解为直达声音分量和扩散声音分量。然后将提取的声音分量发送到远端侧。一致的空间声音再现可以

例如通过所提取的直达声音和扩散声音的加权和来实现，其中权重取决于应当与再现的声音一致的期望的空间图像，例如，权重取决于视频相机的观看方向和缩放因子，所述视频相机可以例如补充音频记录。提供了采用所通知的多声道滤波器来提取直达声音和扩散声音的构思。

[0012] 根据实施例，信号处理器可以例如被配置为确定两个或更多个音频输出信号，其中对于所述两个或更多个音频输出信号的每个音频输出信号，可以例如将平移增益函数分配给所述音频输出信号，其中所述两个或更多个音频输出信号中的每一个信号的所述平移增益函数包括多个平移函数自变量值，其中，平移函数返回值可以例如被分配给所述平移函数自变量值中的每一个值，其中，当所述平移增益函数接收所述平移函数自变量值中的一个值时，所述平移增益函数可以例如被配置为返回被分配给所述平移函数自变量值中的所述一个值的平移函数返回值，并且其中，信号处理器例如被配置为根据分配给所述音频输出信号的平移增益函数的平移函数自变量值中的取决于方向的自变量值，来确定所述两个或更多个音频输出信号中的每一个信号，其中所述取决于方向的自变量值取决于到达方向。

[0013] 在实施例中，所述两个或更多个音频输出信号中的每一个信号的平移增益函数具有作为平移函数自变量值之一的一个或更多个全局最大值，其中对于每个平移增益函数的一个或更多个全局最大值中的每一个最大值，不存在使得所述平移增益函数返回比所述全局最大值使所述平移增益函数返回的增益函数返回值更大的平移函数返回值的其他平移函数自变量值，并且其中对于所述两个或更多个音频输出信号的第一音频输出信号和第二音频输出信号的每对，第一音频输出信号的平移增益函数的一个或更多个全局最大值中的至少一个最大值可以例如不同于第二音频输出信号的平移增益函数的一个或更多个全局最大值中的任一个最大值。

[0014] 根据实施例，信号处理器可以例如被配置为根据窗增益函数来生成所述一个或更多个音频输出信号的每个音频输出信号，其中窗增益函数可以例如被配置为在接收到窗函数自变量值时返回窗函数返回值，其中，如果窗函数自变量值可以例如大于下窗阈值并小于上窗阈值，窗增益函数可以例如被配置为返回比在窗函数自变量值可以例如小于下阈值或大于上阈值的情况下由窗增益函数返回的任何窗函数返回值大的窗函数返回值。

[0015] 在实施例中，信号处理器可以例如被配置为进一步接收指示相对于到达方向的观看方向的角位移的取向信息，并且其中，平移增益函数和窗增益函数中的至少一个取决于所述取向信息；或者其中增益函数计算模块可以例如被配置为进一步接收缩放信息，其中所述缩放信息指示相机的打开角度，并且其中平移增益函数和窗增益函数中的至少一个取决于所述缩放信息；或者其中增益函数计算模块可以例如被配置为进一步接收校准参数，并且其中，平移增益函数和窗增益函数中的至少一个取决于所述校准参数。

[0016] 根据实施例，信号处理器可以例如被配置为接收距离信息，其中信号处理器可以例如被配置为根据所述距离信息生成所述一个或更多个音频输出信号中的每个音频输出信号。

[0017] 根据实施例，信号处理器可以例如被配置为接收取决于原始到达方向的原始角度值，原始到达方向是所述两个或更多音频输入信号的直达信号分量的到达方向，并且信号处理器可以例如被配置为接收距离信息，其中信号处理器可以例如被配置为根据原始角度

值并根据距离信息计算修改的角度值，并且其中信号处理器可以例如被配置为根据修改的角度值来生成所述一个或更多个音频输出信号中的每个音频输出信号。

[0018] 根据实施例，信号处理器可以例如被配置为通过进行低通滤波、或通过添加延迟的直达声音、或通过进行直达声音衰减、或通过进行时间平滑、或者通过进行到达方向扩展、或通过进行去相关来生成所述一个或更多个音频输出信号。

[0019] 在实施例中，信号处理器可以例如被配置为生成两个或更多个音频输出声道，其中信号处理器可以例如被配置为对扩散分量信号应用扩散增益以获得中间扩散信号，并且其中信号处理器可以例如被配置为通过执行去相关从中间扩散信号生成一个或更多个去相关信号，其中所述一个或更多个去相关信号形成所述一个或更多个经处理的扩散信号，或其中所述中间扩散信号和所述一个或更多个去相关信号形成所述一个或更多个经处理的扩散信号。

[0020] 根据实施例，直达分量信号和一个或更多个另外的直达分量信号形成两个或更多个直达分量信号的组，其中分解模块可以例如被配置为生成包括所述两个或更多个音频输入信号的另外的直达信号分量在内的所述一个或更多个另外的直达分量信号，其中所述到达方向和一个或更多个另外的到达方向形成两个或更多个到达方向的组，其中所述两个或更多个到达方向的组中的每个到达方向例如可以被分配给所述两个或更多个直达分量信号的组中的恰好一个直达分量信号，其中所述两个或更多个直达分量信号的直达分量信号数量和所述两个到达方向的到达方向数量可以例如相等，其中信号处理器可以例如被配置为接收所述两个或更多个直达分量信号的组、以及所述两个或更多个到达方向的组，并且其中对于所述一个或更多个音频输出信号中的每个音频输出信号，信号处理器可以例如被配置为针对所述两个或更多个直达分量信号的组中的每个直达分量信号，根据所述直达分量信号的到达方向确定直达增益，并且信号处理器可以例如被配置为通过针对所述两个或更多个直达分量信号的组中的每个直达分量信号，对所述直达分量信号应用所述直达分量信号的直达增益，来生成两个或更多个经处理的直达信号的组，并且信号处理器可以例如被配置为对所述一个或更多个经处理的扩散信号与所述一个或更多个经处理的信号的组中的每个经处理的信号进行组合，来生成所述音频输出信号。

[0021] 在实施例中，所述两个或更多个直达分量信号的组中的直达分量信号的数量加1可以例如小于由接收接口接收的音频输入信号的数量。

[0022] 此外，可以例如提供包括如上所述的系统的助听器或助听设备。

[0023] 此外，提供了一种用于生成一个或更多个音频输出信号的装置。该装置包括信号处理器和输出接口。信号处理器被配置为接收包括两个或更多个原始音频信号的直达信号分量在内的直达分量信号，其中信号处理器被配置为接收包括所述两个或更多个原始音频信号的扩散信号分量在内的扩散分量信号，并且其中信号处理器被配置为接收方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向。此外，信号处理器被配置为根据扩散分量信号生成一个或更多个经处理的扩散信号。对于一个或更多个音频输出信号的每个音频输出信号，信号处理器被配置为根据到达方向确定直达增益，并且信号处理器被配置为将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，且所述信号处理器被配置为将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个扩散信号进行组合以生成所述音频输出信号。输出接口被配置为输出

所述一个或更多个音频输出信号。

[0024] 此外，提供了一种用于生成一个或更多个音频输出信号的方法。所述方法包括：

[0025] -接收两个或更多个音频输入信号。

[0026] -生成包括所述两个或更多个音频输入信号的直达信号分量在内的直达分量信号。

[0027] -生成包括所述两个或更多个音频输入信号的扩散信号分量在内的扩散分量信号。

[0028] -接收取决于所述两个或更多个音频输入信号的直达信号分量的到达方向的方向信息。

[0029] -根据扩散分量信号生成一个或更多个经处理的扩散信号。

[0030] -对于一个或更多个音频输出信号的每个音频输出信号，根据到达方向确定直达增益，将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，以及将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个扩散信号进行组合以生成所述音频输出信号。以及：

[0031] -输出所述一个或更多个音频输出信号。

[0032] 此外，提供了一种用于生成一个或更多个音频输出信号的方法。所述方法包括：

[0033] -接收包括所述两个或更多个原始音频信号的直达信号分量在内的直达分量信号。

[0034] -接收包括所述两个或更多个原始音频信号的扩散信号分量在内的扩散分量信号。

[0035] -接收方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向。

[0036] -根据扩散分量信号生成一个或更多个经处理的扩散信号。

[0037] -对于一个或更多个音频输出信号的每个音频输出信号，根据到达方向确定直达增益，将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，以及将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个扩散信号进行组合以生成所述音频输出信号。以及：

[0038] -输出所述一个或更多个音频输出信号。

[0039] 此外，提供了计算机程序，其中每个计算机程序被配置为当在计算机或信号处理器上执行时实现上述方法之一，使得上述方法中的每一个由计算机程序之一来实现。

[0040] 此外，提供了一种用于生成一个或更多个音频输出信号的系统。所述系统包括分解模块、信号处理器和输出接口。分解模块被配置为接收两个或更多个音频输入信号，其中分解模块被配置为生成包括所述两个或更多个音频输入信号的直达信号分量在内的直达分量信号，并且其中分解模块被配置为生成包括所述两个或更多个音频输入信号的扩散信号分量在内的扩散分量信号。信号处理器被配置为接收直达分量信号、扩散分量信号和方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向。此外，信号处理器被配置为根据扩散分量信号生成一个或更多个经处理的扩散信号。对于一个或更多个音频输出信号的每个音频输出信号，信号处理器被配置为根据到达方向确定直达增益，并且信号处理器被配置为将所述直达增益应用于所述直达分量信号以获得经

处理的直达信号，且所述信号处理器被配置为将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个扩散信号进行组合以生成所述音频输出信号。输出接口被配置为输出所述一个或更多个音频输出信号。信号处理器包括用于计算一个或更多个增益函数的增益函数计算模块，其中所述一个或更多个增益函数中的每个增益函数包括多个增益函数自变量值，其中增益函数返回值被分配给每个所述增益函数自变量值，其中，当所述增益函数接收到所述增益函数自变量值中的一个值时，其中所述增益函数被配置为返回分配给所述增益函数自变量值中的所述一个值的增益函数返回值。此外，信号处理器还包括信号修改器，用于根据到达方向从所述一个或更多个增益函数的增益函数中的增益函数自变量值中选择取决于方向的自变量值，以用于从所述增益函数获得分配给所述取决于方向的自变量值的增益函数返回值，并且用于根据从所述增益函数获得的所述增益函数返回值来确定所述一个或更多个音频输出信号中的至少一个信号的增益值。

[0041] 根据实施例，增益函数计算模块可以例如被配置为针对所述一个或更多个增益函数的每个增益函数生成查找表，其中查找表包括多个条目，其中查找表的每个条目包括增益函数自变量值之一和被分配给所述增益函数自变量值的增益函数返回值，其中增益函数计算模块可以例如被配置为将每个增益函数的查找表存储在持久性或非持久性存储器中，并且其中信号修改器可以例如被配置为通过从存储在存储器中的所述一个或更多个查找表之一中读取所述增益函数返回值，来获得被分配给所述取决于方向的自变量值的增益函数返回值。

[0042] 在实施例中，信号处理器可以例如被配置为确定两个或更多个音频输出信号，其中增益函数计算模块可以例如被配置为计算两个或更多个增益函数，其中对于所述两个或更多个音频输出信号中的每个音频输出信号，增益函数计算模块可以例如被配置为计算被分配给所述音频输出信号的平移增益函数作为所述两个或更多个增益函数之一，其中信号修改器可以例如被配置为根据所述平移增益函数生成所述音频输出信号。

[0043] 根据实施例，所述两个或更多个音频输出信号中的每一个信号的平移增益函数可以例如具有作为所述平移增益函数的增益函数自变量值之一的一个或更多个全局最大值，其中对于所述平移增益函数的一个或更多个全局最大值中的每一个最大值，不存在使得所述平移增益函数返回比所述全局最大值使所述平移增益函数返回的增益函数返回值更大的增益函数返回值的其他增益函数自变量值，并且其中对于所述两个或更多个音频输出信号的第一音频输出信号和第二音频输出信号的每对，第一音频输出信号的平移增益函数的一个或更多个全局最大值中的至少一个最大值可以例如不同于第二音频输出信号的平移增益函数的一个或更多个全局最大值中的任一个最大值。

[0044] 根据实施例，对于所述两个或更多个音频输出信号中的每个音频输出信号，增益函数计算模块可以例如被配置为计算被分配给所述音频输出信号的窗增益函数作为所述两个或更多个增益函数之一，其中所述信号修改器可以例如被配置为根据所述窗增益函数生成所述音频输出信号，并且其中如果所述窗增益函数的自变量值大于下窗阈值并且小于上窗阈值，则窗增益函数被配置为返回比在窗函数自变量值小于下阈值或大于上阈值的情况下由所述窗增益函数返回的任何增益函数返回值大的增益函数返回值。

[0045] 在实施例中，所述两个或更多个音频输出信号中的每一个信号的窗增益函数具有作为所述窗增益函数的增益函数自变量值之一的一个或更多个全局最大值，其中对于所述

窗增益函数的一个或更多个全局最大值中的每一个最大值，不存在使得所述窗增益函数返回比所述全局最大值使所述平移增益函数返回的增益函数返回值更大的增益函数返回值的其他增益函数自变量值，并且其中对于所述两个或更多个音频输出信号的第一音频输出信号和第二音频输出信号的每对，第一音频输出信号的窗增益函数的一个或更多个全局最大值中的至少一个最大值可以例如等于第二音频输出信号的窗增益函数的一个或更多个全局最大值中的一个最大值。

[0046] 根据实施例，增益函数计算模块可以例如被配置为进一步接收指示观看方向相对于到达方向的角位移的取向信息，并且其中增益函数计算模块可以例如被配置为根据所述取向信息生成每个音频输出信号的平移增益函数。

[0047] 在实施例中，增益函数计算模块可以例如被配置为根据取向信息生成每个音频输出信号的窗增益函数。

[0048] 根据实施例，增益函数计算模块可以例如被配置为进一步接收缩放信息，其中缩放信息指示相机的打开角度，并且其中增益函数计算模块可以例如被配置为根据缩放信息生成每个音频输出信号的平移增益函数。

[0049] 在实施例中，增益函数计算模块可以例如被配置为根据缩放信息生成每个音频输出信号的窗增益函数。

[0050] 根据实施例，增益函数计算模块可以例如被配置为进一步接收用于对齐视觉图像和声学图像的校准参数，并且其中增益函数计算模块可以例如被配置为根据校准参数生成每个音频输出信号的平移增益函数。

[0051] 在实施例中，增益函数计算模块可以例如被配置为根据校准参数生成每个音频输出信号的窗增益函数。

[0052] 根据前述任一权利要求所述的系统，增益函数计算模块可以例如被配置为接收关于视觉图像的信息，并且增益函数计算模块可以例如被配置为根据关于视觉图像的信息生成模糊函数返回复数增益以实现声源的感知扩展。

[0053] 此外，提供了一种用于生成一个或更多个音频输出信号的装置。该装置包括信号处理器和输出接口。信号处理器被配置为接收包括两个或更多个原始音频信号的直达信号分量在内的直达分量信号，其中信号处理器被配置为接收包括所述两个或更多个原始音频信号的扩散信号分量在内的扩散分量信号，并且其中信号处理器被配置为接收方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向。此外，信号处理器被配置为根据扩散分量信号生成一个或更多个经处理的扩散信号。对于一个或更多个音频输出信号的每个音频输出信号，信号处理器被配置为根据到达方向确定直达增益，并且信号处理器被配置为将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，且所述信号处理器被配置为将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个扩散信号进行组合以生成所述音频输出信号。输出接口被配置为输出所述一个或更多个音频输出信号。信号处理器包括用于计算一个或更多个增益函数的增益函数计算模块，其中所述一个或更多个增益函数中的每个增益函数包括多个增益函数自变量值，其中增益函数返回值被分配给每个所述增益函数自变量值，其中，当所述增益函数接收到所述增益函数自变量值中的一个值时，其中所述增益函数被配置为返回分配给所述增益函数自变量值中的所述一个值的增益函数返回值。此外，信号处理器还包括信号修改器，

用于根据到达方向从所述一个或更多个增益函数的增益函数中的增益函数自变量值中选择取决于方向的自变量值，以用于从所述增益函数获得分配给所述取决于方向的自变量值的增益函数返回值，并且用于根据从所述增益函数获得的所述增益函数返回值来确定所述一个或更多个音频输出信号中的至少一个信号的增益值。

[0054] 此外，提供了一种用于生成一个或更多个音频输出信号的方法。所述方法包括：

[0055] -接收两个或更多个音频输入信号。

[0056] -生成包括所述两个或更多个音频输入信号的直达信号分量在内的直达分量信号。

[0057] -生成包括所述两个或更多个音频输入信号的扩散信号分量在内的扩散分量信号。

[0058] -接收取决于所述两个或更多个音频输入信号的直达信号分量的到达方向的方向信息。

[0059] -根据扩散分量信号生成一个或更多个经处理的扩散信号。

[0060] -对于一个或更多个音频输出信号的每个音频输出信号，根据到达方向确定直达增益，将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，以及将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个扩散信号进行组合以生成所述音频输出信号。以及：

[0061] -输出所述一个或更多个音频输出信号。

[0062] 生成所述一个或更多个音频输出信号包括：计算一个或更多个增益函数，其中所述一个或更多个增益函数中的每个增益函数包括多个增益函数自变量值，其中增益函数返回值被分配给每个所述增益函数自变量值，其中，当所述增益函数接收到所述增益函数自变量值中的一个值时，其中所述增益函数被配置为返回分配给所述增益函数自变量值中的所述一个值的增益函数返回值。此外，生成所述一个或更多个音频输出信号包括：根据到达方向从所述一个或更多个增益函数的增益函数中的增益函数自变量值中选择取决于方向的自变量值，以用于从所述增益函数获得分配给所述取决于方向的自变量值的增益函数返回值，并且用于根据从所述增益函数获得的所述增益函数返回值来确定所述一个或更多个音频输出信号中的至少一个信号的增益值。

[0063] 此外，提供了一种用于生成一个或更多个音频输出信号的方法。所述方法包括：

[0064] -接收包括所述两个或更多个原始音频信号的直达信号分量在内的直达分量信号。

[0065] -接收包括所述两个或更多个原始音频信号的扩散信号分量在内的扩散分量信号。

[0066] -接收方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向。

[0067] -根据扩散分量信号生成一个或更多个经处理的扩散信号。

[0068] -对于一个或更多个音频输出信号的每个音频输出信号，根据到达方向确定直达增益，将所述直达增益应用于所述直达分量信号以获得经处理的直达信号，以及将所述经处理的直达信号与所述一个或更多个经处理的扩散信号中的一个扩散信号进行组合以生成所述音频输出信号。以及：

[0069] -输出所述一个或更多个音频输出信号。

[0070] 生成所述一个或更多个音频输出信号包括：计算一个或更多个增益函数，其中所述一个或更多个增益函数中的每个增益函数包括多个增益函数自变量值，其中增益函数返回值被分配给每个所述增益函数自变量值，其中，当所述增益函数接收到所述增益函数自变量值中的一个值时，其中所述增益函数被配置为返回分配给所述增益函数自变量值中的所述一个值的增益函数返回值。此外，生成所述一个或更多个音频输出信号包括：根据到达方向从所述一个或更多个增益函数的增益函数中的增益函数自变量值中选择取决于方向的自变量值，以用于从所述增益函数获得分配给所述取决于方向的自变量值的增益函数返回值，并且用于根据从所述增益函数获得的所述增益函数返回值来确定所述一个或更多个音频输出信号中的至少一个信号的增益值。

[0071] 此外，提供了计算机程序，其中每个计算机程序被配置为当在计算机或信号处理器上执行时实现上述方法之一，使得上述方法中的每一个由计算机程序之一来实现。

附图说明

[0072] 参考附图更详细地描述本发明的实施例，其中：

[0073] 图1a示出了根据实施例的系统，

[0074] 图1b示出了根据实施例的装置，

[0075] 图1c示出了根据另一实施例的系统，

[0076] 图1d示出了根据另一实施例的装置，

[0077] 图2示出了根据另一实施例的系统，

[0078] 图3示出了根据实施例的用于直达/扩散分解和用于对系统的估计的参数的模块，

[0079] 图4示出了根据实施例的具有声学缩放的声学场景再现的第一几何形状，其中声源位于焦平面上，

[0080] 图5示出了用于一致的场景再现和声学缩放的平移函数，

[0081] 图6示出了根据实施例的另外的用于一致的场景再现和声学缩放的平移函数，

[0082] 图7示出了根据实施例的用于各种情况的示例窗增益函数，

[0083] 图8示出了根据实施例的扩散增益函数，

[0084] 图9示出了根据实施例的具有声学缩放的声学场景再现的第二几何形状，其中声源不位于焦平面上，

[0085] 图10示出了用于解释直达声音模糊的函数，以及

[0086] 图11示出了根据实施例的助听器。

具体实施方式

[0087] 图1a示出了一种用于生成一个或更多个音频输出信号的系统。该系统包括分解模块101、信号处理器105和输出接口106。

[0088] 分解模块101被配置为生成直达分量信号$X_{dir}(k,n)$，其包括两个或更多音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$的直达信号分量。此外，分解模块101被配置为生成扩散分量信号$X_{diff}(k,n)$，其包括两个或更多音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$的扩散信号分量。

[0089] 信号处理器105被配置为接收直达分量信号$X_{dir}(k,n)$、扩散分量信号$X_{diff}(k,n)$和方向信息，所述方向信息取决于两个或更多个音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$的直达信号分量的到达方向。

[0090] 此外，信号处理器105被配置为根据扩散分量信号$X_{diff}(k,n)$生成一个或更多个经处理的扩散信号$Y_{diff,1}(k,n)$，$Y_{diff,2}(k,n)$，...，$Yd_{iff,v}(k,n)$。

[0091] 对于一个或更多个音频输出信号$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$的每个音频输出信号$Y_i(k,n)$，信号处理器105被配置为根据到达方向确定直达增益$G_i(k,n)$，信号处理器105被配置为将所述直达增益$G_i(k,n)$应用于直达分量信号$X_{dir}(k,n)$以获得经处理的直达信号$Y_{dir,i}(k,n)$，并且信号处理器105被配置为将所述经处理的直达信号$Y_{dir,i}(k,n)$与一个或更多个经处理的扩散信号$Y_{diff,1}(k,n)$，$Y_{diff,2}(k,n)$，...，$Y_{diff,v}(k,n)$中的一个$Y_{diff,i}(k,n)$组合，以生成音频输出信号$Y_i(k,n)$。

[0092] 输出接口106被配置为输出一个或更多个音频输出信号$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$。

[0093] 如概述的，方向信息取决于两个或更多个音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$的直达信号分量的到达方向$\varphi(k,n)$。例如，两个或更多个音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$的直达信号分量的到达方向例如本身可以是方向信息。或者，例如，方向信息可以例如是两个或更多个音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$的直达信号分量的传播方向。当到达方向从接收麦克风阵列指向声源时，传播方向从声源指向接收麦克风阵列。因此，传播方向精确地指向到达方向的相反方向，并且因此取决于到达方向。

[0094] 为了生成一个或更多个音频输出信号$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$的一个$Y_i(k,n)$，信号处理器105：

[0095] -根据到达方向确定直达增益$G_i(k,n)$，

[0096] -将所述直达增益应用于直达分量信号$X_{dir}(k,n)$以获得经处理的直达信号$Y_{dir,i}(k,n)$，以及

[0097] -将所述经处理的直达信号$Y_{dir,i}(k,n)$和所述一个或更多个经处理的扩散信号$Y_{diff,1}(k,n)$，$Y_{diff,2}(k,n)$，...，$Yd_{iff,v}(k,n)$的一个$Y_{diff,i}(k,n)$组合以生成所述音频输出信号$Y_i(k,n)$。

[0098] 针对应被生成的$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$的一个或更多个音频输出信号$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$中的每个执行所述操作。信号处理器可以例如被配置为生成一个、两个、三个或更多个音频输出信号$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$。

[0099] 关于一个或更多个经处理的扩散信号$Y_{diff,1}(k,n)$，$Y_{diff,2}(k,n)$，...，$Y_{diff,v}(k,n)$，根据实施例，信号处理器105可以例如被配置为通过将扩散增益$Q(k,n)$应用于扩散分量信号$X_{diff}(k,n)$，来生成一个或更多个经处理的扩散信号$Y_{diff,1}(k,n)$，$Y_{diff,2}(k,n)$，...，$Y_{diff,v}(k,n)$。

[0100] 分解模块101被配置为可以例如通过将一个或更多个音频输入信号分解成直达分量信号和分解成扩散分量信号，生成包括两个或更多个音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$的直达信号分量在内的直达分量信号$X_{dir}(k,n)$、以及包括两个或更多个音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$的扩散信号分量在内的扩散分量信号$X_{diff}(k,n)$。

[0101] 在具体实施例中，信号处理器105可以例如被配置为生成两个或更多个音频输出

信号$Y_1(k,n)$,$Y_2(k,n)$,...,$Y_v(k,n)$。信号处理器105可以例如被配置为将扩散增益$Q(k,n)$应用于扩散分量信号$X_{diff}(k,n)$以获得中间扩散信号。此外,信号处理器105可以例如被配置为通过执行去相关来从中间扩散信号生成一个或更多个去相关信号,其中一个或更多个去相关信号形成一个或更多个经处理的扩散信号$Y_{diff,1}(k,n)$,$Y_{diff,2}(k,n)$,...,$Y_{diff,v}(k,n)$,或其中中间扩散信号和一个或更多个去相关信号形成一个或更多个经处理的扩散信号$Y_{diff,1}(k,n)$,$Y_{diff,2}(k,n)$,...,$Y_{diff,v}(k,n)$。

[0102] 例如,经处理的扩散信号$Y_{diff,1}(k,n)$,$Y_{diff,2}(k,n)$,...,$Y_{diff,v}(k,n)$的数量和音频输出信号$Y_1(k,n)$,$Y_2(k,n)$,...,$Y_v(k,n)$的数量可以例如相等。

[0103] 从中间扩散信号生成一个或更多个去相关信号可以例如通过对中间扩散信号应用延迟、或者例如通过使中间扩散信号与噪声突发进行卷积、或者例如通过使中间扩散信号与脉冲响应进行卷积等来执行。可以例如备选地或附加地应用任何其他现有技术的去相关技术。

[0104] 为了获得v个音频输出信号$Y_1(k,n)$,$Y_2(k,n)$,...,$Y_v(k,n)$,可以例如对v个直达增益$G_1(k,n)$,$G_2(k,n)$,...,$G_v(k,n)$进行v次确定、以及对一个或更多个直达分量信号$X_{dir}(k,n)$应用v次相应增益,来获得v个音频输出信号$Y_1(k,n)$,$Y_2(k,n)$,...,$Y_v(k,n)$。

[0105] 例如,可以仅需要单个扩散分量信号$X_{diff}(k,n)$、单个扩散增益$Q(k,n)$的一次确定和对扩散分量信号$X_{diff}(k,n)$应用一次扩散增益$Q(k,n)$,来获得v个音频输出信号$Y_1(k,n)$,$Y_2(k,n)$,...,$Y_v(k,n)$。为了实现去相关,可以仅在已经将扩散增益应用于扩散分量信号之后应用去相关技术。

[0106] 根据图1a的实施例,然后将相同的经处理的扩散信号$Y_{diff}(k,n)$与经处理的直达信号的相应的一个信号($Y_{dir,i}(k,n)$)组合,以获得相应的一个音频输出信号($Y_i(k,n)$)。

[0107] 图1a的实施例考虑了两个或更多音频输入信号$x_1(k,n)$,$x_2(k,n)$,...$x_p(k,n)$的直达信号分量的到达方向。因此,通过根据到达方向灵活调整直达分量信号$X_{dir}(k,n)$和扩散分量信号$X_{diff}(k,n)$,可以生成音频输出信号$Y_1(k,n)$,$Y_2(k,n)$,...,$Y_v(k,n)$。实现了高级适配可能性。

[0108] 根据实施例,例如可以针对时频域的每个时间频率仓(k,n)来确定音频输出信号$Y_1(k,n)$,$Y_2(k,n)$,...,$Y_v(k,n)$。

[0109] 根据实施例,分解模块101可以例如被配置为接收两个或更多个音频输入信号$x_1(k,n)$,$x_2(k,n)$,...$x_p(k,n)$。在另一实施例中,分解模块101可以例如被配置为接收三个或更多个音频输入信号$x_1(k,n)$,$x_2(k,n)$,...$x_p(k,n)$。分解模块101可以例如被配置为将两个或更多个(或者三个或更多个)音频输入信号$x_1(k,n)$,$x_2(k,n)$,...$x_p(k,n)$分解为不是多声道信号的扩散分量信号$X_{diff}(k,n)$、以及一个或更多个直达分量信号$X_{dir}(k,n)$。音频信号不是多声道信号意味着音频信号本身不包括多于一个音频声道。因此,多个音频输入信号的音频信息在两个分量信号($X_{dir}(k,n)$,$X_{diff}(k,n)$)(以及可能的附加辅助信息)内传输,这可实现高效传输。

[0110] 信号处理器105可以例如被配置为通过以下操作来生成两个或更多个音频输出信号$Y_1(k,n)$,$Y_2(k,n)$,...,$Y_v(k,n)$的每个音频输出信号$Y_i(k,n)$:将直达增益$G_i(k,n)$应用于所述音频输出信号$Y_i(k,n)$,将所述直达增益$G_i(k,n)$应用于一个或更多个直达分量信号$X_{dir}(k,n)$以获得针对所述音频输出信号$Y_i(k,n)$的经处理的直达信号$Y_{dir,i}(k,n)$,以及将用于

所述音频输出信号$Y_i(k,n)$的所述经处理的直达信号$Y_{dir,i}(k,n)$与经处理的扩散信号$Y_{diff}(k,n)$组合以生成所述音频输出信号$Y_i(k,n)$。输出接口106被配置为输出两个或更多个音频输出信号$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$。通过仅确定单个经处理的扩散信号$Y_{diff}(k,n)$来生成两个或更多个音频输出信号$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$是尤其有益的。

[0111] 图1b示出了根据实施例的用于生成一个或更多个音频输出信号$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$的装置。该装置实现了图1a的系统中的所谓的“远端”侧。

[0112] 图1b的装置包括信号处理器105和输出接口106。

[0113] 信号处理器105被配置为接收直达分量信号$X_{dir}(k,n)$，其包括两个或更多个原始音频信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$(例如，图1a的音频输入信号)的直达信号分量。此外，信号处理器105被配置为接收扩散分量信号$X_{diff}(k,n)$，其包括两个或更多原始音频信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$的扩散信号分量。此外，信号处理器105被配置为接收方向信息，所述方向信息取决于所述两个或更多个音频输入信号的直达信号分量的到达方向。

[0114] 信号处理器105被配置为根据扩散分量信号$X_{diff}(k,n)$生成一个或更多个经处理的扩散信号$Y_{diff,1}(k,n)$，$Y_{diff,2}(k,n)$，...，$Y_{diff,v}(k,n)$。

[0115] 对于一个或更多个音频输出信号$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$的每个音频输出信号$Y_i(k,n)$，信号处理器105被配置为根据根据到达方向确定直达增益$G_i(k,n)$，信号处理器105被配置为将所述直达增益$G_i(k,n)$应用于直达分量信号$X_{dir}(k,n)$以获得经处理的直达信号$Y_{dir,i}(k,n)$，并且信号处理器105被配置为将所述经处理的直达信号$Y_{dir,i}(k,n)$与一个或更多个经处理的扩散信号$Y_{diff,1}(k,n)$，$Y_{diff,2}(k,n)$，...，$Y_{diff,v}(k,n)$中的一个$Y_{diff,i}(k,n)$组合，以生成所述音频输出信号$Y_i(k,n)$。

[0116] 输出接口106被配置为输出所述一个或更多个音频输出信号$Y_1(k,n)$，$Y_2(k,n)$，...，$Y_v(k,n)$。

[0117] 下面参考系统描述的信号处理器105的所有配置也可以在根据图1b的装置中实现。这具体涉及下文描述的信号修改器103和增益函数计算模块104的各种配置。这同样适用于下述构思的各种应用示例。

[0118] 图1c示出了根据另一实施例的系统。在图1c中，图1a的信号处理器105还包括用于计算一个或更多个增益函数的增益函数计算模块104，其中所述一个或更多个增益函数中的每个增益函数包括多个增益函数自变量值，其中增益函数返回值被分配给每个所述增益函数自变量值，其中，当所述增益函数接收到所述增益函数自变量值中的一个值时，其中所述增益函数被配置为返回分配给所述增益函数自变量值中的所述一个值的增益函数返回值。

[0119] 此外，信号处理器105还包括信号修改器103，用于根据到达方向从所述一个或更多个增益函数的增益函数的增益函数自变量值中选择取决于方向的自变量值，以用于从所述增益函数获得分配给所述取决于方向的自变量值的增益函数返回值，并且用于根据从所述增益函数获得的所述增益函数返回值来确定所述一个或更多个音频输出信号中的至少一个信号的增益值。

[0120] 图1d示出了根据另一实施例的系统。在图1d中，图1b的信号处理器105还包括用于计算一个或更多个增益函数的增益函数计算模块104，其中所述一个或更多个增益函数中的每个增益函数包括多个增益函数自变量值，其中增益函数返回值被分配给每个所述增益

函数自变量值，其中，当所述增益函数接收到所述增益函数自变量值中的一个值时，其中所述增益函数被配置为返回分配给所述增益函数自变量值中的所述一个值的增益函数返回值。

[0121] 此外，信号处理器105还包括信号修改器103，用于根据到达方向从所述一个或更多个增益函数的增益函数的增益函数自变量值中选择取决于方向的自变量值，以用于从所述增益函数获得分配给所述取决于方向的自变量值的增益函数返回值，并且用于根据从所述增益函数获得的所述增益函数返回值来确定所述一个或更多个音频输出信号中的至少一个信号的增益值。

[0122] 实施例提供了记录和再现空间声音，使得声学图像与期望的空间图像一致，该期望的空间图像例如由补充远端侧的音频的视频确定。一些实施例基于利用位于混响近端侧的麦克风阵列的记录。实施例提供例如与相机的视觉缩放一致的声学缩放。例如，当放大时，从扬声器将位于缩放的视觉图像中的方向再现扬声器的直达声音，使得视觉图像和声学图像对齐。如果在放大之后扬声器位于视觉图像之外(或者在期望的空间区域之外)，则这些扬声器的直达声音可以被衰减，因为这些扬声器不再可见，或者例如来自这些扬声器的直达声音不是所期望的。此外，例如，当放大以模拟视觉相机的较小打开角度时，可以增加直达与混响比。

[0123] 实施例基于以下构思：通过在近端侧应用两个近期的多声道滤波器，将记录的麦克风信号分离为声源的直达声音和扩散声音(例如，混响声音)。这些多声道滤波器可以例如基于声场的参数信息，例如直达声音的DOA。在一些实施例中，分离的直达声音和扩散声音可以例如与参数信息一起被发送到远端侧。

[0124] 例如，在远端侧，可以例如将特定权重应用于提取的直达声音和扩散声音，这样可调整再现的声学图像，使得得到的音频输出信号与期望的空间图像一致。这些权重例如模拟声学缩放效果并且例如取决于直达声音的到达方向(DOA)以及例如取决于相机的缩放因子和/或观看方向。然后，可以例如通过对加权的直达声音和扩散声音求和来获得最终的音频输出信号。

[0125] 所提供的构思实现了在上述具有消费者设备的视频记录场景中或在电话会议场景中的高效使用：例如，在视频记录场景中，其可以例如足以存储或发送所提取的直达声音和扩散声音(而不是所有麦克风信号)，同时仍然能够控制所重建的空间图像。

[0126] 这意味着，如果例如在后处理步骤(数字缩放)中应用视觉缩放，则声学图像仍然可以被相应地修改，而不需要存储和访问原始麦克风信号。在电话会议场景中，所提出的构思也可以被有效地使用，因为直达和扩散声音提取可以在近端侧执行，同时仍然能够在远端侧控制空间声音再现(例如，改变扬声器设置)并且将声学图像和视觉图像对齐。因此，只需要发送很少的音频信号和估计的DOA作为辅助信息，同时远端侧的计算复杂度低。

[0127] 图2示出了根据实施例的系统。近端侧包括模块101和102。远端侧包括模块105和106。模块105本身包括模块103和104。当参考近端侧和远端侧时，应当理解，在一些实施例中，第一装置可以实现近端侧(例如，包括模块101和102)，并且第二装置可以实现远端侧(例如，包括模块103和104)，而在其他实施例中，单个装置实现近端侧以及远端侧，其中这样的单个装置例如包括模块101、102、103和104。

[0128] 特别地，图2示出了根据实施例的系统，其包括分解模块101、参数估计模块102、信

号处理器105和输出接口106。在图2中，信号处理器105包括增益函数计算模块104和信号修改器103。信号处理器105和输出接口106可以例如实现如图1b所示的装置。

[0129] 在图2中，参数估计模块102可以例如被配置为接收两个或更多个音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$。此外，参数估计模块102可以例如被配置为根据两个或更多个音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$估计所述两个或更多音频输入信号的直达信号分量的到达方向。信号处理器105可以例如被配置为从参数估计模块102接收包括两个或更多个音频输入信号的直达信号分量的到达方向在内的到达方向信息。

[0130] 图2的系统的输入包括在时频域(频率索引k，时间索引n)中的M个麦克风信号$X_{1...M}(k,n)$。例如，可以假设由麦克风捕获的声场存在于在各向同性扩散场中传播的平面波的每个(k,n)。平面波对声源(例如，扬声器)的直达声音进行建模，而扩散声音对混响进行建模。

[0131] 根据这种模型，第m个麦克风信号可以写为

[0132] $$X_m(k,n)=X_{dir,m}(k,n)+X_{diff,m}(k,n)+X_{n,m}(k,n), \quad (1)$$

[0133] 其中$X_{dir,m}(k,n)$是测量的直达声音(平面波)，$X_{diff,m}(k,n)$是测量的扩散声音，$X_{n,m}(k,n)$是噪声分量(例如，麦克风自噪声)。

[0134] 在图2中的分解模块101中(直达/扩散分解)，从麦克风信号中提取直达声音$X_{dir}(k,n)$和扩散声音$X_{diff}(k,n)$。为此目的，例如，可以采用如下所述的所通知的多声道滤波器。对于直达/扩散分解，例如可以采用关于声场的特定参数信息，例如直达声音的$\mathrm{DOA}\varphi(k,n)$。该参数信息可以例如在参数估计模块102中从麦克风信号中估计。除了直达声音的$\mathrm{DOA}\varphi(k,n)$之外，在一些实施例中，例如可以估计距离信息$r(k,n)$。该距离信息可以例如描述麦克风阵列和发射平面波的声源之间的距离。对于参数估计，例如可以采用距离估计器和/或现有技术的DOA估计器。例如，可以在下面描述相应的估计器。

[0135] 提取的直达声音$X_{dir}(k,n)$、提取的扩散声音$X_{diff}(k,n)$和直达声音的估计的参数信息例如$\mathrm{DOA}\ \varphi(k,n)$和/或距离$r(k,n)$随后可以例如被存储，被发送到远端侧，或者立即被用于生成具有期望的空间图像的空间声音，例如以创建声学缩放效果。

[0136] 使用提取的直达声音$X_{dir}(k,n)$、提取的扩散声音$X_{diff}(k,n)$和估计的参数信息$\varphi(k,n)$和/或$r(k,n)$，在信号修改器103中生成期望的声学图像，例如声学缩放效果。

[0137] 信号修改器103可以例如计算在时频域中的一个或更多个输出信号$Y_i(k,n)$，其重建声学图像，使得它与期望的空间图像一致。例如，输出信号$Y_i(k,n)$模拟声学缩放效果。这些信号可以最终被变换回时域并且例如通过扬声器或耳机被回放。第i个输出信号$Y_i(k,n)$被计算为提取的直达声音$X_{dir}(k,n)$和扩散声音$X_{diff}(k,n)$的加权和，例如，

[0138]
$$Y_i(k,n)=G_i(k,n)X_{dir}(k,n)+f_i\{\underbrace{QX_{diff}(k,n)}_{Y_{diff}(k,n)}\} \quad (2a)$$

$$=Y_{dir,i}(k,n)+Y_{diff,i}(k,n). \quad (2b)$$

[0139] 在公式(2a)和(2b)中，权重$G_i(k,n)$和Q是用于创建期望声学图像(例如声学缩放效果)的参数。例如，当放大时，可以减小参数Q，使得再现的扩散声音被衰减。

[0140] 此外，利用权重$G_i(k,n)$，可以控制从哪个方向再现直达声音，使得视觉图像和声学图像对齐。此外，可以将声学模糊效果与直达声音对齐。

[0141] 在一些实施例中，可以例如在增益选择单元201和202中确定权重$G_i(k,n)$和Q。这些单元可以例如根据估计的参数信息$\varphi(k,n)$和$r(k,n)$，从由g_i和q表示的两个增益函数中选择适当的权重$G_i(k,n)$和Q。在数学上表达为，

[0142] $$G_i(k,n) = g_i(\varphi, r), \quad (3a)$$

[0143] $$Q(k,n) = q(r). \quad (3b)$$

[0144] 在一些实施例中，增益函数g_i和q可以取决于应用，并且例如可以在增益函数计算模块104中生成。增益函数描述了对于给定参数信息、$\varphi(k,n)$和/或$r(k,n)$应在(2a)中使用哪些权重$G_i(k,n)$和Q，使得获得期望的一致空间图像。

[0145] 例如，当用可视相机放大时，调整增益函数，使得从源在视频中可见的方向再现声音。下面进一步描述权重$G_i(k,n)$和Q以及基本增益函数g_i和q。应当注意，权重$G_i(k,n)$和Q以及基本增益函数g_i和q可以例如是复数值的。计算增益函数需要诸如缩放因子、视觉图像的宽度、期望的观看方向和扬声器设置之类的信息。

[0146] 在其他实施例中，在信号修改器103内直接计算的权重$G_i(k,n)$和Q，而不是首先在模块104中计算增益函数，然后在增益选择单元201和202中从计算的增益函数中选择权重$G_i(k,n)$和Q。

[0147] 根据实施例，例如可以针对每个时间-频率对多于一个的平面波进行具体处理。例如，来自两个不同方向的相同频带中的两个或更多个平面波可以例如由相同时间点的麦克风阵列记录。这两个平面波可以各自具有不同的到达方向。在这种情况下，可以例如单独考虑两个或更多个平面波的直达信号分量及其到达方向。

[0148] 根据实施例，直达分量信号$X_{dir1}(k,n)$和一个或更多个另外的直达分量信号$X_{dir2}(k,n),\ldots,X_{dir\ q}(k,n)$可以例如形成两个或更多个直达分量分量信号$X_{dir1}(k,n),X_{dir2}(k,n),\ldots,X_{dir\ q}(k,n)$的组，其中分解模块101可以例如被配置为生成一个或更多个另外的直达分量信号$X_{dir2}(k,n),\ldots,X_{dir\ q}(k,n)$，所述直达分量信号包括两个或更多个音频输入信号$x_1(k,n),x_2(k,n),\ldots x_p(k,n)$的另外的直达信号分量。

[0149] 到达方向和一个或更多个另外的到达方向形成两个或更多个到达方向的组，其中两个或更多个到达方向的组中的每个方向被分配给所述两个或更多个直达分量信号$X_{dir1}(k,n),X_{dir2}(k,n),\ldots,X_{dir\ q,m}(k,n)$的组中的恰好一个直达分量信号$X_{dir\ j}(k,n)$，其中所述两个或更多个直达分量信号的直达分量信号数量与所述两个到达方向的到达方向数量相等。

[0150] 信号处理器105可以例如被配置为接收两个或更多个直达分量信号$X_{dir1}(k,n),X_{dir2}(k,n),\ldots,X_{dir\ q}(k,n)$的组、以及两个或更多个到达方向的组。

[0151] 对于一个或更多个音频输出信号$Y_1(k,n),Y_2(k,n),\ldots,Y_v(k,n)$的每个音频输出信号$Y_i(k,n)$，

[0152] -信号处理器105可以例如被配置为针对两个或更多个直达分量信号$X_{dir1}(k,n),X_{dir2}(k,n),\ldots,X_{dir\ q}(k,n)$的组中的每个直达分量信号$X_{dir\ j}(k,n)$，根据所述直达分量信号$X_{dir\ j}(k,n)$的到达方向确定直达增益$G_{j,i}(k,n)$，

[0153] -信号处理器105可以例如被配置为通过针对所述两个或更多个直达分量信号$X_{dir1}(k,n),X_{dir2}(k,n),\ldots,X_{dir\ q}(k,n)$的组中的每个直达分量信号$X_{dir\ j}(k,n)$，将所述直达

分量信号$X_{dir\ j}(k,n)$的直达增益$G_{j,i}(k,n)$应用于所述直达分量信号$X_{dir\ j}(k,n)$，来生成两个或更多个经处理的直达信号$Y_{dir1,i}(k,n)$，$Y_{dir2,i}(k,n)$，...，$Y_{dirq,i}(k,n)$的组。并且：

[0154] -信号处理器105可以例如被配置为将一个或更多个经处理的扩散信号$Y_{diff,1}(k,n)$，$Y_{diff,2}(k,n)$，...，$Y_{diff,v}(k,n)$中的一个$Y_{diff,i}(k,n)$与两个或更多个经处理的信号$Y_{dir1,i}(k,n)$，$Y_{dir2,i}(k,n)$，...，$Y_{dir\ q,i}(k,n)$的组中的每个经处理的信号$Y_{dirj,i}(k,n)$进行组合，来生成所述音频输出信号$Y_i(k,n)$。

[0155] 因此，如果分别考虑两个或更多个平面波，则公式(1)的模型变为：

[0156] $X_m(k,n)=X_{dir1,m}(k,n)+X_{dir2,m}(k,n)+...+X_{dirq,m}(k,n)+X_{diff,m}(k,n)+X_{n,m}(k,n)$

[0157] 并且可以例如根据下式与公式(2a)和(2b)相类似地计算权重：

[0158] $Y_i(k,n)=G_{1,i}(k,n)X_{dir1}(k,n)+G_{2,i}(k,n)X_{dir2}(k,n)+...+G_{q,i}(k,n)X_{dir\ q}(k,n)+QX_{diff,m}(k,n)$

[0159] $=Y_{dir1,i}(k,n)+Y_{dir2,i}(k,n)+...+Y_{dirq,i}(k,n)+Y_{diff,i}(k,n)$

[0160] 仅一些直达分量信号、扩散分量信号和辅助信息从近端侧发送到远端侧也是足够的。在实施例中，两个或更多个直达分量信号$X_{dir1}(k,n)$，$X_{dir2}(k,n)$，...，$X_{dir\ q}(k,n)$的组中的直达分量信号的数量加1小于由接收接口101接收到的音频输入信号$x_1(k,n)$，$x_2(k,n)$，...$x_p(k,n)$的数量。(使用指数：$q+1<p$)“加1”表示所需的扩散分量信号$X_{diff}(k,n)$。

[0161] 当在下文中提供关于单个平面波、关于单个到达方向和关于单个直达分量信号的解释时，应当理解，所解释的构思同样适用于多于一个平面波、多于一个到达方向和多于一个直达分量信号。

[0162] 在下文中，描述了直达和扩散声音提取。提供了实现直达/扩散分解的图2的分解模块101的实际实现。

[0163] 在实施例中，为了实现一致的空间声音再现，对在[8]和[9]中描述的两个最近提出的所通知的线性约束最小方差(LCMV)滤波器的输出进行组合，这在假设与在DirAC(直达音频编码)中相似的声场模型的情况下，实现利用期望的任意响应对直达声音和扩散声音的精确多声道提取。现在在下文描述根据实施例组合这些滤波器的具体方式：

[0164] 首先，描述根据实施例的直达声音提取。

[0165] 使用最近提出的在[8]中描述的所通知的空间滤波器来提取直达声音。在下文中简要回顾该滤波器，然后将其制定为使得其可用于根据图2的实施例。

[0166] (2b)和图2中的第i个扬声器声道的估计的期望直达信号$\hat{Y}_{dir,i}(k,n)$通过将线性多声道滤波器应用于麦克风信号来计算，例如，

[0167] $$\hat{Y}_{\mathrm{dir},i}(k,n)=\mathbf{w}_{\mathrm{dir},i}^{\mathrm{H}}(k,n)\mathbf{x}(k,n), \tag{4}$$

[0168] 其中，向量$\mathbf{x}(k,n)=[X_1(k,n),...,X_M(k,n)]^T$包括M个麦克风信号，并且$w_{dir,i}$是复数值的权重向量。这里，滤波器权重最小化麦克风所包括的噪声和扩散声音并且同时以期望增益$G_i(k,n)$捕获直达声音声音。在数学上表示，权重可以例如被计算为

[0169] $$\mathbf{w}_{\mathrm{dir},i}(k,n)=\arg\min_{\mathbf{w}}\mathbf{w}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{u}}(k,n)\mathbf{w} \tag{5}$$

[0170] 受到线性约束

[0171] $$\mathbf{w}^{\mathrm{H}}\mathbf{a}(k,\varphi)=G_i(k,n). \tag{6}$$

[0172] 这里，$\mathbf{a}(k,\varphi)$是所谓的阵列传播向量。该向量的第m个元素是第m个麦克风和阵列的参考麦克风之间的直达声音的相对传递函数(不失一般性，在下面的描述中使用位置d_1处的第一麦克风)。该向量取决于直达声音的DOA $\varphi(k,n)$。

[0173] 例如，在[8]中定义了阵列传播向量。在文献[8]的公式(6)中，根据下式定义阵列传播向量

[0174] $$\mathbf{a}(k,\varphi_l)=[a_1(k,\varphi_l)\ \ldots\ a_M(k,\varphi_l)]^{\mathrm{T}}$$

[0175] 其中φ_l是第l个平面波的到达方向的方位角。因此，阵列传播向量取决于到达方向。如果仅存在或考虑一个平面波，则可以省略索引l。

[0176] 根据[8]的公式(6)，阵列传播向量a的第i个元素a_i描述了从第一个到第i个麦克风的第l个平面波的相移是根据下式定义的

[0177] $$a_i(k,\varphi_l)=\exp\{\jmath\kappa r_i\sin\varphi_l(k,n)\}$$

[0178] 例如，r_i等于第一个和第i个麦克风之间的距离，κ表示平面波的波数，并且$\jmath$是虚数。

[0179] 关于阵列传播向量a及其元素a_i的更多信息可以在[8]中找到，其通过引用明确地并入本文。

[0180] (5)中的M×M矩阵$\Phi_u(k,n)$是噪声和扩散声音的功率谱密度(PSD)矩阵，其可以如[8]中所解释的那样来确定。(5)的解出下式给出

[0181] $$\mathbf{w}_{\mathrm{dir},i}(k,n)=\mathbf{h}_{\mathrm{dir}}(k,n)\,G_i^*(k,n), \tag{7}$$

[0182] 其中

[0183] $$\mathbf{h}_{\mathrm{dir}}(k,n)=\mathbf{\Phi}_{\mathrm{u}}^{-1}(k,n)\mathbf{a}(k,\varphi)\left[\mathbf{a}^{\mathrm{H}}(k,\varphi)\mathbf{\Phi}_{\mathrm{u}}^{-1}\mathbf{a}(k,\varphi)\right]^{-1}. \tag{8}$$

[0184] 计算滤波器需要阵列传播向量$\mathbf{a}(k,\varphi)$，其可以在直达声音的DOA$\varphi(k,n)$被估计之后被确定[8]。如上所述，阵列传播向量以及滤波器取决于DOA。可以如下所述对DOA进行估计。

[0185] 在[8]中提出的例如使用(4)和(7)的直达声音提取的所通知的空间滤波器不能直接用于图2的实施例中。实际上，该计算需要麦克风信号x(k,n)以及直达声音增益$G_i(k,n)$。从图2中可以看出，麦克风信号x(k,n)仅在近端侧可用，而直达声音增益$G_i(k,n)$仅在远端侧可用。

[0186] 为了在本发明的实施例中使用所通知的空间滤波器，提供了修改，其中我们将(7)代入(4)，导致

[0187] $$\widehat{Y}_{\mathrm{dir},i}(k,n)=G_i(k,n)\widehat{X}_{\mathrm{dir}}(k,n\ ), \tag{9}$$

[0188] 其中

[0189] $$\widehat{X}_{\mathrm{dir}}(k,n)=\mathbf{h}_{\mathrm{dir}}^{\mathrm{H}}(k,n)\mathbf{x}(k,n). \tag{10}$$

[0190] 该修改的滤波器$h_{dir}(k,n)$独立于权重$G_i(k,n)$。因此，可以在近端侧应用滤波器以

获得直达声音$\hat{X}_{dir}(k,n)$，然后可以将该直达声音与估计的DOA(和距离)一起作为辅助信息发送到远端侧，以提供对直达声音的再现的完全控制。可以在位置d_1处相对于参考麦克风确定直达声音$\hat{X}_{dir}(k,n)$。因此，也可以将直达声音分量与$\hat{X}_{dir}(k,n,\mathbf{d}_1)$相关联，因此：

[0191] $$\widehat{X}_{\mathrm{dir}}(k,n,\mathbf{d}_1)=\mathbf{h}_{\mathrm{dir}}^{\mathrm{H}}(k,n)\mathbf{x}(k,n). \qquad (10a)$$

[0192] 所以根据实施例，分解模块101可以例如被配置为通过根据下式对两个或更多个音频输入信号应用滤波器来生成直达分量信号：

[0193] $$\widehat{X}_{\mathrm{dir}}(k,n)=\mathbf{h}_{\mathrm{dir}}^{\mathrm{H}}(k,n)\mathbf{x}(k,n)\ ,$$

[0194] 其中，k表示频率，并且其中n表示时间，其中$\hat{X}_{dir}(k,n)$表示直达分量信号，其中x(k,n)表示两个或更多个音频输入信号，其中$h_{dir}(k,n)$表示滤波器，并且

[0195] $$\mathbf{h}_{\mathrm{dir}}(k,n)=\mathbf{\Phi}_{\mathrm{u}}^{-1}(k,n)\mathbf{a}(k,\varphi)\left[\mathbf{a}^{\mathrm{H}}(k,\varphi)\mathbf{\Phi}_{\mathrm{u}}^{-1}\mathbf{a}(k,\varphi)\right]^{-1}.$$

[0196] 其中$\Phi_u(k,n)$表示所述两个或更多个音频输入信号的噪声和扩散声音的功率谱密度矩阵，其中$\mathbf{a}(k,\varphi)$表示阵列传播向量，并且其中φ表示所述两个或更多个音频输入信号的直达信号分量的到达方向的方位角。

[0197] 图3示出根据实施例的参数估计模块102和实现直达/扩散分解的分解模块101。

[0198] 图3示出的实施例实现了直达声音提取模块203的直达声音提取和扩散声音提取模块204的扩散声音提取。

[0199] 在直达声音提取模块203中通过将滤波器权重应用于如(10)中给出的麦克风信号来执行直达声音提取。在直达权重计算单元301中计算直达滤波器权重，其可以例如用(8)来实现。然后，例如方程式(9)的增益$G_i(k,n)$被应用在远端侧，如图2所示。

[0200] 在下文中，描述了扩散声音提取。扩散声音提取可以例如由图3的扩散声音提取模块204来实现。在例如下文描述的图3的扩散权重计算单元302中计算扩散滤波器权重。

[0201] 在实施例中，扩散声音可以例如使用最近在[9]中提出的空间滤波器来提取。(2a)和图2中的扩散声音$X_{diff}(k,n)$可以例如通过将第二空间滤波器应用于麦克风信号来估计，例如，

[0202] $$\widehat{X}_{\mathrm{diff}}(k,n)=\mathbf{h}_{\mathrm{diff}}^{\mathrm{H}}(k,n)\mathbf{x}(k,n). \qquad (11)$$

[0203] 为了找到用于扩散声音$h_{diff}(k,n)$的最佳滤波器，我们考虑最近提出的[9]中的滤波器，它可以提取具有期望的任意响应的扩散声，同时最小化滤波器输出处的噪声。对于空间白噪声，滤波器由下式给出

[0204] $$\mathbf{h}_{\mathrm{diff}}(k,n)=\arg\min_{\mathbf{h}}\mathbf{h}^{\mathrm{H}}\mathbf{h} \qquad (12)$$

[0205] 符合$\mathbf{h}^{\mathrm{H}}\mathbf{a}(k,\varphi)=0$以及$h^H\gamma_1(k)=1$。第一线性约束确保直达声音被抑制，而第二约束确保平均来说以所需增益Q捕获扩散声音，参见文献[9]。注意，$\gamma_1(k)$是在[9]中定义的扩散声音相干向量。(12)的解由下式给出

[0206] $$\mathbf{h}_{\mathrm{diff}}(k,n)=\mathbf{\Lambda}\gamma_{\mathrm{diff}}(k)\left[\gamma_{\mathrm{diff}}^{\mathrm{H}}(k)\mathbf{\Lambda}\gamma_{\mathrm{diff}}(k)\right]^{-1}, \qquad (13)$$

[0207] 其中

[0208] $$\mathbf{\Lambda}(k,\varphi)=\mathbf{I}-\mathbf{a}(k,\varphi)\left[\mathbf{a}^{\mathrm{H}}(k,\varphi)\mathbf{a}(k,\varphi)\right]^{-1}\mathbf{a}^{\mathrm{H}}(k,\varphi) \quad (14)$$

[0209] 其中，I是大小为M×M的单位矩阵。滤波器$h_{diff}(k,n)$不取决于权重$G_i(k,n)$和Q，因此，可以在近端侧计算并应用该滤波器来获得$\hat{X}_{diff}(k,n)$。为此，仅需要将单个音频信号发送到远端侧，即$\hat{X}_{diff}(k,n)$，同时仍然能够完全控制扩散声音的空间声音再现。

[0210] 图3还示出了根据实施例的扩散声音提取。在扩散声音提取模块204中通过将滤波器权重应用于如公式(11)中给出的麦克风信号来执行扩散声音提取。在扩散权重计算单元302中计算滤波器权重，其可以例如通过使用公式(13)来实现。

[0211] 在下文中，描述了参数估计。参数估计可以例如由参数估计模块102进行，其中可以例如估计关于所记录的声音场景的参数信息。该参数信息用于计算分解模块101中的两个空间滤波器以及用于在信号修改器103中对一致的空间音频再现进行增益选择。

[0212] 首先，描述了DOA信息的确定/估计。

[0213] 在下文中描述了实施例，其中参数估计模块(102)包括用于直达声音(例如用于源自声源位置并到达麦克风阵列的平面波)的DOA估计器。在不失一般性的情况下，假设对于每个时间和频率存在单个平面波。其他实施例考虑存在多个平面波的情况，并且将这里描述的单个平面波构思扩展到多个平面波是显而易见的。因此，本发明还涵盖具有多个平面波的实施例。

[0214] 可以使用现有技术的窄带DOA估计器之一(例如ESPRIT[10]或根MUSIC[11])，从麦克风信号估计窄带DOA。针对到达麦克风阵列的一个或更多个波，除方位角$\varphi(k,n)$以外，DOA信息也可以被提供为空间频率$\mu[k\mid\varphi(k,n)]$、相移或传播向量$\mathbf{a}[k\mid\varphi(k,n)]$的形式。应当注意，DOA信息也可以在外部提供。例如，平面波的DOA可以与假设人类说话者形成声学场景的面部识别算法一起由视频相机来确定。

[0215] 最后，应当注意，DOA信息也可以在3D(三维)中估计。在这种情况下，在参数估计模块102中估计方位角$\varphi(k,n)$和仰角$\vartheta(k,n)$，并且平面波的DOA在这种情况下被提供为例如(φ,ϑ)。

[0216] 因此，当在下文中提及DOA的方位角时，应当理解，所有解释也可应用于DOA的仰角、DOA的方位角或从DOA的方位角导出的角度、DOA的仰角或从DOA的仰角导出的角度、或者从DOA的方位角和仰角导出的角度。更一般地，下文提供的所有解释同样适用于取决于DOA的任何角度。

[0217] 现在，描述距离信息确定/估计。

[0218] 一些实施例涉及基于DOA和距离的顶部声学缩放。在这样的实施例中，参数估计模块102可以例如包括两个子模块，例如上述DOA估计器子模块和距离估计子模块，该距离估计子模块估计从记录位置到声源$r(k,n)$的距离。在这样的实施例中，例如可以假定到达记录麦克风阵列的每个平面波源自声源并沿着直线传播到该阵列(其也被称为直达传播路径)。

[0219] 存在几种使用麦克风信号进行距离估计的现有技术方法。例如，到源的距离可以

通过计算麦克风信号之间的功率比来找到，如[12]中所述。备选地，可以基于估计的信号与扩散比(SDR)来计算到声学环境(例如，房间)中的源r(k,n)的距离[13]。然后可以将SDR估计与房间的混响时间(已知的或使用现有技术方法估计的混响时间)组合以计算距离。对于高SDR，与扩散声音相比，直达声音能量高，这表示到源的距离小。当SDR值为低时，与房间混响相比，直达声音功率弱，这表示到源的距离大。

[0220] 在其他实施例中，取代通过在参数估计模块102中采用距离计算模块来计算/估计距离，可以例如从视觉系统接收外部距离信息。例如，可以采用能够提供距离信息(例如，飞行时间(ToFu)、立体视觉和结构光)的在视觉中使用的现有技术。例如，在ToF相机中，可以根据由相机发出的、行进到源并返回到相机传感器的光信号的测量的飞行时间来计算到源的距离。例如，计算机立体视觉使用两个有利点，从这两个点中捕获视觉图像以计算到源的距离。

[0221] 或者，例如，可以采用结构化光相机，其中已知的像素图案被投影在视觉场景上。投影之后的变形分析使得视觉系统能够估计到源的距离。应当注意，对于一致的音频场景再现，需要针对每个时间-频率仓的距离信息r(k,n)。如果距离信息由视觉系统在外部提供，则到与DOA$\varphi(k,\ n)$相对应的源r(k,n)的距离可以例如被选为来自视觉系统的与该特定方向$\varphi(k,\ n)$相对应的距离值。

[0222] 在下文中，考虑一致的声学场景再现。首先，考虑基于DOA的声学场景再现。

[0223] 可以进行声学场景再现，使得其与记录的声场景一致。或者，可以进行声学场景再现，使得其与视觉图像一致。可以提供对应的视觉信息以实现与视觉图像的一致性。

[0224] 例如，可以通过调整(2a)中的权重$G_i(k,n)$和Q来实现一致性。根据实施例，信号修改器103可以例如存在于近端侧，或者如图2所示，可以在远端侧例如接收直达声音$\hat{X}_{dir}(k,n)$和扩散声音$\hat{X}_{diff}(k,n)$作为输入，同时接收DOA估计$\varphi(k,\ n)$作为辅助信息。基于所接收的信息，可以例如根据公式(2a)生成用于可用的再现系统的输出信号$Y_i(k,n)$。

[0225] 在一些实施例中，在增益选择单元201和202中，分别从由增益函数计算模块104提供的两个增益函数$g_i(\varphi(k,\ n))$和$q(k,n)$中选择参数$G_i(k,n)$和Q。

[0226] 根据实施例，例如可以仅基于DOA信息来选择$G_i(k,n)$，并且Q可以例如具有常数值。然而，在其他实施例中，其他权重$G_i(k,n)$可以例如基于进一步的信息来确定，并且权重Q可以例如以多种方式来确定。

[0227] 首先，考虑实现与记录的声学场景的一致性的实施。之后，考虑实现与图像信息/与视觉图像的一致性的实施例。

[0228] 在下文中，描述了权重$G_i(k,n)$和Q的计算，用于再现与所记录的声学场景一致的声学场景，例如，使得位于再现系统的最佳点的收听者将声源感知为从所记录的声学场景中的声源的DOA到达，具有与所记录的场景中相同的功率，并且再现对周围的扩散声音的相同感知。

[0229] 对于已知的扬声器设置，例如可以通过由增益选择单元201从由增益函数计算模块104针对估计的DOA $\varphi(k,\ n)$所提供的固定查找表中选择直达声音增益$G_i(k,n)$(“直达增益选择”)，来实现对来自方向$\varphi(k,n)$的声源的再现，其可以写为

[0230] $$G_i(k,n) = g_i(\varphi(k,n)), \qquad (15)$$

[0231] 其中$g_i(\varphi) = p_i(\varphi)$是为第i个扬声器的所有DOA返回平移增益的函数。平移增益函数$p_i(\varphi)$取决于扬声器设置和平移方案。

[0232] 图5(a)中示出了用于立体声再现中的左、右扬声器的由向量基幅度平移(VBAP)[14]定义的平移增益函数$p_i(\varphi)$的示例。

[0233] 在图5(a)中,示出了用于立体声设置的VBAP平移增益函数$p_{b,i}$的示例,图5(b)中示出了用于一致再现的平移增益。

[0234] 例如,如果直达声音从$\varphi(k,n) = 30^\circ$到达,则右扬声器增益为$G_r(k,n) = g_r(30^\circ) = p_r(30^\circ) = 1$,左扬声器增益为$G_l(k,n) = g_l(30^\circ) = p_l(30^\circ) = 0$。对于从$\varphi(k,n) = 0^\circ$到达的直达声音,最终的立体声扬声器增益是$G_r(k,n) = G_l(k,n) = \sqrt{0.5}$。

[0235] 在实施例中,在双耳声音再现的情况下,平移增益函数(例如,$p_i(\varphi)$)可以是例如头相关传递函数(HRTF)。

[0236] 例如,如果HRTF$g_i(\varphi) = p_i(\varphi)$返回复数值,则在增益选择单元201中选择的直达声音增益$G_i(k,n)$可以例如是复数值的。

[0237] 如果将生成三个或更多个音频输出信号,则可以例如采用对应的现有技术的平移概念来将输入信号平移到该三个或更多个音频输出信号。例如,可以采用用于三个或更多个音频输出信号的VBAP。

[0238] 在一致的声学场景再现中,扩散声音的功率应与所记录的场景保持相同。因此,对于具有例如等间隔扬声器的扬声器系统,扩散声音增益具有常数值:

[0239] $$Q = q_i = \frac{1}{\sqrt{I}}, \qquad (16)$$

[0240] 其中I是输出扬声器声道的数量。这意味着增益函数计算模块104根据可用于再现的扬声器的数量为第i个扬声器(或耳机声道)提供单个输出值,并且该值被用作所有频率上的扩散增益Q。通过对在(2b)中获得的$Y_{diff}(k,n)$进行去相关来获得第i个扬声器声道的最终扩散声音$Y_{diff,i}(k,n)$。

[0241] 因此,可以通过以下操作来实现与所记录的声学场景一致的声学场景再现:例如根据例如到达方向确定每个音频输出信号的增益,将多个确定的增益$G_i(k,n)$应用于直达声音信号$\hat{X}_{dir}(k,n)$以确定多个直达输出信号分量$\hat{Y}_{dir,i}(k,n)$,将确定的增益Q应用于扩散声音信号$\hat{X}_{diff}(k,n)$以获得扩散输出信号分量$\hat{Y}_{diff}(k,n)$,以及将所述多个直达输出信号分量$\hat{Y}_{dir,i}(k,n)$中的每一个与扩散输出信号分量$\hat{Y}_{diff}(k,n)$进行组合以获得一个或更多个音频输出信号$Y_i(k,n)$。

[0242] 现在,描述根据实施例的实现与视觉场景的一致性的音频输出信号生成。具体地,描述了根据实施例的用于再现与视觉场景一致的声学场景的权重$G_i(k,n)$和Q的计算。其目的在于重建声像,其中来自源的直达声音从源在视频/图像中可见的方向被再现。

[0243] 可以考虑如图4所示的几何形状,其中1对应于视觉相机的观看方向。不失一般性

地，我们可以在坐标系的y轴上定义l。

[0244] 在所描绘的(x，y)坐标系中，直达声音的DOA的方位角由$\varphi(k,\ n)$给出，并且源在x轴上的位置由$x_g(k,n)$给出。这里，假设所有声源位于与x轴相距相同的距离g处，例如，源位置位于左虚线上，其在光学中被称为焦平面。应当注意，该假设仅用于确保视觉和声音图像对齐，并且对于所呈现的处理不需要实际距离值g。

[0245] 在再现侧(远端侧)，显示器位于b，并且显示器上的源的位置由$x_b(k,n)$给出。此外，x_d是显示器尺寸(或者，在一些实施例中，例如，x_d表示显示器尺寸的一半)，φ_d是相应的最大视角，S是声音再现系统的最佳点，$\varphi_b(k,\ n)$是直达声音应被再现为使得视觉图像和声音图像对齐的角度。$\varphi_b(k,\ n)$取决于$x_b(k,n)$以及最佳点S与位于b处的显示器之间的距离。此外，$x_b(k,n)$取决于几个参数，例如源与相机的距离g、图像传感器尺寸和显示器尺寸x_d。不幸的是，这些参数中的至少一些在实践中经常是未知的，使得对于给定的DOA$\varphi_g(k,\ n)$不能确定$x_b(k,n)$和$\varphi_b(k,\ n)$。然而，假设光学系统是线性的，根据公式(17)：

[0246] $$\tan\varphi_b(k,n)=c\ \tan\varphi(k,n), \qquad (17)$$

[0247] 其中c是补偿上述未知参数的未知常数。应当注意，仅当所有源位置具有与x轴相同的距离g时，c才是常数。

[0248] 在下文中，假设c为校准参数，其应当在校准阶段期间被调整，直到视觉图像和声音图像一致。为了执行校准，声源应当被定位在焦平面上，并且找到c的值以使得视觉图像和声音图像被对齐。一旦校准，c的值保持不变，并且直达声音应该被再现的角度由下式给出

[0249] $$\varphi_b(k,n)=\tan^{-1}[c\ \tan\varphi(k,n)]. \qquad (18)$$

[0250] 为了确保声学场景和视觉场景两者一致，将原始平移函数$p_i(\varphi)$修改为一致(修改的)平移函数$p_{b,i}(\varphi)$。现在根据下式来选择直达声音增益$G_i(k,n)$

[0251] $$G_i(k,n)=g_i(\varphi(k,n)), \qquad (19)$$

[0252] $$g_i(\varphi)=p_{b,i}(\varphi), \qquad (20)$$

[0253] 其中$p_{b,i}(\varphi)$是一致的平移函数，其在所有可能的源DOA中返回用于第i个扬声器的平移增益。对于c的固定值，在增益函数计算模块104中从原始(例如，VBAP)平移增益表将这样的一致平移函数计算为

[0254] $$p_{b,i}(\varphi)=p_i(\tan^{-1}[c\ \tan\varphi]). \qquad (21)$$

[0255] 因此，在实施例中，信号处理器105可以例如被配置为针对一个或更多个音频输出信号的每个音频输出信号进行确定，使得直达增益$G_i(k,n)$根据下式来定义

[0256] $G_i(k,\ n)=p_i(\tan^{-1}[c\ \tan(\varphi(k,\ n))])$。

[0257] 其中，i表示所述音频输出信号的索引，其中k表示频率，并且其中n表示时间，其中$G_i(k,n)$表示直达增益，其中$\varphi(k,\ n)$表示取决于到达方向的角度(例如，到达方向的方位

角)，其中c表示常数值，并且其中p_i表示平移函数。

[0258] 在实施例中，在增益选择单元201中基于来自由增益函数计算模块104提供的固定查找表的估计的DOA $\varphi(k, n)$来选择直达声音增益，其在使用(19)时(在校准阶段之后)被计算一次。

[0259] 因此，根据实施例，信号处理器105可以例如被配置为针对一个或更多个音频输出信号的每个音频输出信号，取决于到达方向从查找表获得用于所述音频输出信号的直达增益。

[0260] 在实施例中，信号处理器105计算用于直达增益函数$g_i(k,n)$的查找表。例如，对于DOA的方位角值φ的每个可能的全度数，例如1°、2°、3°、...，可以预先计算和存储直达增益$G_i(k,n)$。然后，当接收到到达方向的当前方位角值φ时，信号处理器105从查找表读取用于当前方位角值φ的直达增益$G_i(k,n)$。(当前方位角值φ可以例如是查找表自变量值；并且直达增益$G_i(k,n)$可以例如是查找表返回值)。取代DOA的方位角φ，在其他实施例中，可以针对取决于到达方向的任意角度计算查找表。其优点在于，不必总是针对每个时间点或者针对每个时间-频率仓计算增益值，而是相反地，计算查找表一次，然后针对接收角φ，从查找表读取直达增益$G_i(k,n)$。

[0261] 因此，根据实施例，信号处理器105可以例如被配置为计算查找表，其中查找表包括多个条目，其中每个条目包括查找表自变量值和被分配给所述自变量值的查找表返回值。信号处理器105可以例如被配置为通过取决于到达方向来选择查找表的查找表自变量值之一，从查找表获得查找表返回值之一。此外，信号处理器105可以例如被配置为根据从查找表获得的查找表返回值中的一个来确定一个或更多个音频输出信号中的至少一个信号的增益值。

[0262] 信号处理器105可以例如被配置为通过取决于另一个到达方向选择查找表自变量值中的另一个自变量值，从(相同)查找表获得查找表返回值中的另一个返回值，以确定增益值。例如，信号处理器可以例如在稍后的时间点接收取决于所述另一个到达方向的另一个方向信息。

[0263] 图5(a)和5(b)中示出了VBAP平移和一致的平移增益函数的示例。

[0264] 应当注意，取代重新计算平移增益表，可以备选地计算用于显示器的DOA $\varphi_b(k, n)$，并将其应用于原始平移函数中作为$\varphi_i(\varphi_b(k, n))$。这是真的，因为以下关系成立：

[0265] $$p_{b,i}(\varphi(k,n)) = p_i(\varphi_b(k,n)). \quad (22)$$

[0266] 然而，这将要求增益函数计算模块104还接收估计的DOA$\varphi(k, n)$作为输入，并且然后将针对每个时间索引n执行例如根据公式(18)进行的DOA重新计算。

[0267] 关于扩散声音再现，当以与没有视觉的情况下所解释的方式相同的方式进行处理时，例如当扩散声音的功率保持与记录场景中的扩散功率相同，并且扬声器信号是$Y_{diff}(k,n)$的不相关版本时，一致地重建声学图像和视觉图像。对于等间隔的扬声器，扩散声音增益具有例如由公式(16)给出的常数值。结果，增益函数计算模块104为第i个扬声器(或耳机声道)提供在所有频率上用作扩散增益Q的单个输出值。通过对由公式(2b)给出的$Y_{diff}(k,n)$进行去相关来获得第i个扬声器声道的最终扩散声音$Y_{diff,i}(k,n)$。

[0268] 现在，考虑提供基于DOA的声学缩放的实施例。在这样的实施例中，可以考虑与视觉缩放一致的用于声学缩放的处理。通过调整例如在公式(2a)中采用的权重$G_i(k,n)$和Q来实现这种一致的视听缩放，如图2的信号修改器103所示。

[0269] 在实施例中，例如，可以在增益选择单元201中从直达增益函数$g_i(k,n)$中选择直达增益$G_i(k,n)$，其中，所述直达增益函数是在增益函数计算模块104中基于参数估计模块102中估计的DOA来计算的。在增益选择单元202中从在增益函数计算模块104中计算的扩散增益函数$q(\beta)$中选择扩散增益Q。在其他实施例中，直达增益$G_i(k,n)$和扩散增益Q由信号修改器103计算，而不需要首先计算相应的增益函数然后选择增益。

[0270] 应当注意，与上述实施例相反，基于缩放因子β确定扩散增益函数$q(\beta)$。在实施例中，不使用距离信息，因此，在这样的实施例中，不在参数估计模块102中估计距离信息。

[0271] 为了在(2a)中导出缩放参数$G_i(k,n)$和Q，考虑图4中的几何图形。图中所示的参数类似于在上述实施例中参考图4所描述的参数。

[0272] 类似于上述实施例，假设所有声源位于焦平面上，所述焦平面以距离g与x轴平行。应当注意，一些自动聚焦系统能够提供g，例如到焦平面的距离。这允许假设图像中的所有源都是锐利的。在再现(远端)侧，显示器上的DOA$\varphi_b(k,n)$和位置$x_b(k,n)$取决于许多参数，例如源与相机的距离g、图像传感器尺寸、显示器尺寸x_d和相机的缩放因子(例如，相机的打开角度)β。假设光学系统是线性的，根据公式(23)：

[0273] $$\tan\varphi_b(k,n) = \beta c \tan\varphi(k,n), \qquad (23)$$

[0274] 其中c是补偿未知光学参数的校准参数，$\beta \geq 1$是用户控制的缩放因子。应当注意，在视觉相机中，以因子β放大等于将$x_b(k,n)$乘以β。此外，仅当所有源位置与x轴具有相同的距离g时，c才是常数。在这种情况下，c可以被认为是校准参数，其被调整一次使得视觉图形和声音图像对齐。从直达增益函数$g_i(\varphi)$中选择直达声音增益$G_i(k,n)$，如下

[0275] $$G_i(k,n) = g_i(\varphi(k,n)), \qquad (24)$$

[0276] $$g_i(\varphi) = p_{b,i}(\varphi)\, w_b(\varphi), \qquad (25)$$

[0277] 其中$p_{b,i}(\varphi)$表示平移增益函数，$w_b(\varphi)$是用于一致的视听缩放的窗增益函数。在增益函数计算模块104中从原始(例如，VBAP)平移增益函数$p_i(\varphi)$计算用于一致的视听缩放的平移增益函数，如下

[0278] $$p_{b,i}(\varphi) = p_i(\tan^{-1}[\beta c \tan\varphi]). \qquad (26)$$

[0279] 因此，例如在增益选择单元201中选择的直达声音增益$G_i(k,n)$基于来自在增益函数计算模块104中计算的查找平移表的估计的DOA$\varphi(k,n)$来确定，如果β不改变，则所述估计的DOA$\varphi(k,n)$是固定的。应当注意，在一些实施例中，每次修改缩放因子β时，需要通过使用例如公式(26)来重新计算$p_{b,i}(\varphi)$。

[0280] 图6(参照图6(a)和图6(b))中示出了β＝1和β＝3的示例立体声平移增益函数。特别地，图6(a)示出了β＝1的示例平移增益函数$p_{b,i}$；图6(b)示出了在β＝3的缩放之后的平移增益；以及图6(c)示出了在具有角位移的β＝3的缩放之后的平移增益。

[0281] 在该示例中可以看出，当直达声音从$\varphi(k, n) = 10°$到达时，对于大的β值，左扬声器的平移增益增加，而右扬声器的平移函数，且β＝3返回比β＝1小的值。当缩放因子β增加时，这种平移有效地将感知的源位置更多地向外部方向移动。

[0282] 根据实施例，信号处理器105可以例如被配置为确定两个或更多个音频输出信号。对于两个或更多个音频输出信号的每个音频输出信号，将平移增益函数分配给所述音频输出信号。

[0283] 两个或更多个音频输出信号中的每一个的平移增益函数包括多个平移函数自变量值，其中平移函数返回值被分配给所述平移函数自变量值中的每一个，其中，当所述平移函数接收到所述平移函数自变量值之一时，所述平移函数被配置为返回被分配给所述平移函数自变量值中的所述一个值的平移函数返回值。

[0284] 信号处理器105被配置为根据分配给所述音频输出信号的平移增益函数的平移函数自变量值的取决于方向的自变量值来确定两个或更多个音频输出信号中的每一个，其中所述取决于方向的自变量值取决于到达方向。

[0285] 根据实施例，两个或更多个音频输出信号中的每一个的平移增益函数具有作为平移函数自变量值之一的一个或更多个全局最大值，其中对于每个平移增益函数的一个或更多个全局最大值中的每一个，不存在使得所述平移增益函数返回比所述全局最大值使所述平移增益函数返回的增益函数返回值更大的平移函数返回值的其他平移函数自变量值。

[0286] 对于两个或更多个音频输出信号的第一音频输出信号和第二音频输出信号的每对，第一音频输出信号的平移增益函数的一个或更多个全局最大值中的至少一个不同于第二音频输出信号的平移增益函数的一个或更多个全局最大值中的任一个。

[0287] 简言之，实现平移函数使得不同的平移函数的全局最大值(的至少一个)不同。

[0288] 例如，在图6(a)中，$p_{b,l}(\varphi)$的局部最大值在-45°至-28°的范围内，并且$p_{b,r}(\varphi)$的局部最大值在+28°至+45°的范围内，因此全局最大值不同。

[0289] 例如，在图6(b)中，$p_{b,l}(\varphi)$的局部最大值在-45°至-8°的范围内，并且$p_{b,r}(\varphi)$的局部最大值在+8°至+45°的范围内，因此全局最大值也不同。

[0290] 例如，在图6(c)中，$p_{b,l}(\varphi)$的局部最大值在-45°至+2°的范围内，并且$p_{b,r}(\varphi)$的局部最大值在+18°至+45°的范围内，因此全局最大值也不同。

[0291] 平移增益函数可以例如被实现为查找表。

[0292] 在这样的实施例中，信号处理器105可以例如被配置为计算用于至少一个音频输出信号的平移增益函数的平移查找表。

[0293] 所述至少一个音频输出信号的每个音频输出信号的平移查找表可以例如包括多个条目，其中每个条目包括所述音频输出信号的平移增益函数的平移函数自变量值，并且所述平移函数返回值被分配给所述平移函数自变量值，其中信号处理器105被配置为通过根据到达方向来从平移查找表选择取决于方向的自变量值，来从所述平移查找表获得平移函数返回值之一，并且其中信号处理器105被配置为根据从所述平移查找表获得的所述平移函数返回值之一来确定所述音频输出信号的增益值。

[0294] 在下文中，描述了采用直达声音窗的实施例。根据这样的实施例，根据下式来计算用于一致的缩放的直达声窗$w_b(\varphi)$

[0295] $$w_b(\varphi) = w(\tan^{-1}[\beta c \tan \varphi]), \quad (27)$$

[0296] 其中$w_b(\varphi)$是用于声学缩放的窗增益函数，其中如果源被映射到缩放因子β的视觉图像之外的位置，则所述窗增益函数衰减直达声音。

[0297] 例如，可以针对β＝1设置窗函数$w(\varphi)$，使得在视觉图像之外的源的直达声音减小到期望的水平，并且可以例如通过采用公式(27)在每次缩放参数改变时都对其进行重新计算。应当注意，对于所有扬声器声道，$w_b(\varphi)$是相同的。图7(a-b)中示出了β＝1和β＝3的示例窗函数，其中对于增加的β值，窗宽度减小。

[0298] 图7中示出了一致的窗增益函数的示例。特别地，图7(a)示出了没有缩放(缩放因子β＝1)的窗增益函数w_b，图7(b)示出了缩放之后(缩放因子β＝3)的窗增益函数，图7(c)示出了在具有角位移的缩放之后(缩放因子β＝3)的窗增益函数。例如，角位移可以实现窗向观察方向的旋转。

[0299] 例如，在图7(a)、7(b)和7(c)中，如果DOAφ位于窗内，则窗增益函数返回增益1，如果DOAφ位于窗外，则窗增益函数返回增益0.18，并且如果φ位于窗的边界处，则窗增益函数返回0.18和1之间的增益。

[0300] 根据实施例，信号处理器105被配置为根据窗增益函数来生成一个或更多个音频输出信号的每个音频输出信号。窗增益函数被配置为在接收到窗函数自变量值时返回窗函数返回值。

[0301] 如果窗函数自变量值大于下窗阈值并且小于上窗阈值，则窗增益函数被配置为返回比在窗函数自变量值小于下阈值或大于上阈值的情况下由所述窗增益函数返回的任何窗函数返回值大的窗函数返回值。

[0302] 例如，在公式(27)中

[0303] $$w_b(\varphi) = w(\tan^{-1}[\beta c \tan \varphi]),$$

[0304] 到达方向的方位角φ是窗增益函数$w_b(\varphi)$的窗函数自变量值。窗增益函数$w_b(\varphi)$取决于缩放信息，这里为缩放因子β。

[0305] 为了解释窗增益函数的定义，可以参考图7(a)。

[0306] 如果DOA的方位角φ大于-20°(下阈值)且小于+20°(上阈值)，则窗增益函数返回的所有值都大于0.6。否则，如果DOA的方位角φ小于-20°(下阈值)或大于+20°(上阈值)，则窗增益函数返回的所有值都小于0.6。

[0307] 在实施例中，信号处理器105被配置为接收缩放信息。此外，信号处理器105被配置为根据窗增益函数生成一个或更多个音频输出信号的每个音频输出信号，其中窗增益函数取决于缩放信息。

[0308] 在其他值被认为是下/上阈值，或者其他值被认为是返回值的情况下，这可以通过图7(b)和图7(c)的(修改的)窗增益函数看出。参考图7(a)、7(b)和7(c)，可以看出，窗增益函数取决于缩放信息：缩放因子β。

[0309] 窗增益函数可以例如被实现为查找表。在这样的实施例中，信号处理器105被配置为计算窗查找表，其中窗查找表包括多个条目，其中每个条目包括窗增益函数的窗函数自变量值和窗增益函数的被分配给所述窗函数自变量值的窗函数返回值。信号处理器105被

配置为通过取决于到达方向选择窗查找表的窗函数自变量值之一，从窗查找表获得窗函数返回值之一。此外，信号处理器105被配置为根据从窗查找表获得的窗函数返回值中的所述一个值来确定一个或更多个音频输出信号中的至少一个信号的增益值。

[0310] 除了缩放概念之外，窗和平移函数可以移动位移角度θ。该角度可以对应于相机观看方向1的旋转或者通过类比于相机中的数字缩放在视觉图像内移动。在前一种情况下，针对显示器上的角度重新计算相机旋转角度，例如，类似于公式(23)。在后一种情况下，θ可以是用于一致的声学缩放的窗和平移函数(例如$w_b(\varphi)$和$p_{b,i}(\varphi)$)的直接偏移。在图5(c)和图6(c)中描绘了对两个函数进行位移的示意性示例。

[0311] 应注意的是，取代重新计算平移增益和窗函数，可以例如根据公式(23)计算显示器的DOA$\varphi_b(k, n)$，并且将其分别应用于原始平移和窗函数作为$p_i(\varphi)$和$w(\varphi_b)$。这种处理是等效的，因为以下关系成立：

[0312] $$p_{b,i}(\varphi(k,n)) = p_i(\varphi_b(k,n)), \quad (28)$$

[0313] $$w_b(\varphi(k,n)) = w(\varphi_b(k,n)). \quad (29)$$

[0314] 然而，这将要求增益函数计算模块104接收估计的DOA$\varphi(k, n)$作为输入，并且在每个连续时间帧中执行例如根据公式(18)的DOA重新计算，而不管β是否改变。

[0315] 对于扩散声音，例如在增益函数计算模块104中计算扩散增益函数q(β)仅需要知道可用于再现的扬声器I的数量。因此，其可以独立于视觉相机或显示器的参数来设置。

[0316] 例如，对于等间隔的扬声器，在增益选择单元202中基于缩放参数β选择公式(2a)中的实值扩散声音增益$Q \in [0, 1/\sqrt{I}]$。使用扩散增益的目的是根据缩放因子衰减扩散声音，例如，缩放增加了再现信号的DRR。这通过针对较大的β而降低Q来实现。事实上，放大意味着相机的打开角度变小，例如，自然声学对应将是捕获较少扩散声音的更直达的麦克风。

[0317] 为了模拟这种效果，实施例可以例如采用图8所示的增益函数。图8示出了扩散增益函数q(β)的示例。

[0318] 在其他实施例中，增益函数被不同地定义。通过对例如根据公式(2b)的$Y_{diff}(k,n)$进行去相关来获得第i个扬声器声道的最终扩散声音$Y_{diff,i}(k,n)$。

[0319] 在下文中，考虑基于DOA和距离的声学缩放。

[0320] 根据一些实施例，信号处理器105可以例如被配置为接收距离信息，其中信号处理器105可以例如被配置为根据所述距离信息生成一个或更多个音频输出信号中的每个音频输出信号。

[0321] 一些实施例采用基于估计的DOA$\varphi(k, n)$和距离值r(k,n)的一致的声学缩放的处理。这些实施例的构思还可以应用于在不进行缩放的情况下将所记录的声学场景与视频对齐，其中源不位于与之前在可用的距离信息r(k,n)中假设的距离相同的距离，这使得我们能够创建针对在视觉图像中不出现尖锐的声源(例如针对不位于相机的焦平面上的源)创建声学模糊效果。

[0322] 为了利用对位于不同距离处的源进行模糊来促进一致的声音再现(例如声学缩放)，可以在公式(2a)中基于两个估计的参数(即$\varphi(k,n)$和r(k,n))并根据缩放因子β来调整

增益$G_i(k,n)$和Q，如在图2的信号修改器103中所示。如果不涉及缩放，则β可以被设置为β＝1。

[0323] 例如，可以如上所述在参数估计模块102中估计参数$\varphi(k,\ n)$和r(k,n)。在该实施例中，基于来自一个或更多个直达增益函数gi，j(k,n)（其可以例如在增益函数计算模块104中计算）的DOA和距离信息来确定直达增益$G_i(k,n)$（例如通过在增益选择单元201中选择）。与如针对上述实施例所描述的相类似，可以例如在增益选择单元202中从扩散增益函数q(β)中选择扩散增益Q，例如，基于缩放因子β在增益函数计算模块104中计算。

[0324] 在其他实施例中，直达增益$G_i(k,n)$和扩散增益Q由信号修改器103计算，而不需要首先计算相应的增益函数然后选择增益。

[0325] 为了解释不同距离处的声源的声学再现和声学缩放，参考图9。图9中表示的参数与上文描述的那些类似。

[0326] 在图9中，声源位于与x轴相距距离R(k,n)的位置P′。距离r可以是例如是(k,n)特定的（时间-频率特定的：r(k,n)）表示源位置和焦平面（通过g的左垂直线）之间的距离。应当注意，一些自动聚焦系统能够提供g，例如到焦平面的距离。

[0327] 来自麦克风阵列的视点的直达声音的DOA由$\varphi'(k,\ n)$表示。与其他实施例不同，不假设所有源位于距相机镜头相同的距离g处。因此，例如，位置P′可以具有相对于x轴的任意距离R(k,n)。

[0328] 如果源不位于焦平面上，则视频中的源将显得模糊。此外，实施例基于如下发现：如果源位于虚线910上的任何位置，则它将出现在视频中的相同位置$x_b(k,n)$。然而，实施例基于如下的发现：如果源沿着虚线910移动，则直达声音的估计的DOA$\varphi'(k,\ n)$将改变。换句话说，基于实施例采用的发现，如果源平行于y轴移动，则估计的DOA$\varphi'(k,\ n)$将在x_b（进而应该再现声音的DOA$\varphi_b(k,\ n)$）保持相同。因此，如果如在先前实施例中所描述的那样将估计的DOA$\varphi'(k,\ n)$发送到远端侧并且用于声音再现，则如果源改变其距离R(k,n)，声学图像和视觉图像不再对齐。

[0329] 为了补偿该效应并实现一致的声音再现，例如在参数估计模块102中进行的DOA估计好像源位于位置P处的焦平面上那样对直达声音的DOA进行估计。该位置表示P′在焦平面上的投影。相应的DOA由图9中的$\varphi(k,\ n)$表示，并且在远端侧用于一致的声音再现，与前述实施例相类似。如果r和g是已知的，则可以基于几何考虑从估计的（原始）DOA$\varphi'(k,\ n)$计算（修改的）DOA$\varphi(k,\ n)$。

[0330] 例如，在图9中，信号处理器105可以例如根据下式从$\varphi'(k,\ n)$ r和g计算$\varphi(k,\ n)$：

[0331] $$\varphi = \arctan\left(\frac{\tan\varphi' \cdot (r+g)}{g}\right)。$$

[0332] 因此，根据实施例，信号处理器105可以例如被配置为接收到达方向的原始方位角$\varphi'(k,\ n)$，所述到达方向是两个或更多个音频输入信号的直达信号分量的到达方向，并且信号处理器被配置为还接收距离信息，并且可以例如被配置为还接收距离信息r。信号处理器105可以例如被配置为根据原始到达方向的方位角$\varphi'(k,\ n)$并根据到达方向的距离信息r

和g来计算到达方向的修改的方位角$\varphi(k, n)$。信号处理器105可以例如被配置为根据修改的到达方向的方位角$\varphi(k, n)$生成一个或更多个音频输出信号中的每个音频输出信号。

[0333] 可以如上所述估计所需的距离信息(焦平面的距离g可以从透镜系统或者自动聚焦信息获得)。应当注意,例如,在本实施例中,源和焦平面之间的距离r(k,n)与(映射的)DOA $\varphi(k, n)$一起被发送到远端侧。

[0334] 此外,通过类比于视觉缩放,位于距焦平面大距离r处的源在图像中不显得锐利。这种效应在光学中是公知的,称为所谓的场深(DOF),其定义了源距离在视觉图像中看起来锐利的可接受的范围。

[0335] 作为距离r的函数的DOF曲线的示例在图10(a)中示出。

[0336] 图10示出了用于场深的示例图(图10(a))、用于低通滤波器的截止频率的示例图(图10(b))和用于重复直达声音的以ms为单位的时延的示例图(图10(c))。

[0337] 在图10(a)中,距离焦平面小距离处的源仍然是锐利的,而较远距离(距离相机更近或更远)的源显得模糊。因此,根据实施例,相应的声源被模糊,使得它们的视觉图像和声学图像是一致的。

[0338] 为了导出实现声学模糊和一致的空间声音再现的(2a)中的增益$G_i(k,n)$和Q,考虑位于$P(\varphi, r)$处的源将出现在显示器上的角度。模糊的源将显示在

[0339] $$\tan\varphi_b(k,n) = \beta c \tan\varphi(k,n), \qquad (30)$$

[0340] 其中c是校准参数,$\beta \geq 1$是用户控制的缩放因子,$\varphi(k, n)$是例如在参数估计模块102中估计的(映射的)DOA。如前所述,这种实施例中的直达增益$G_i(k,n)$可以例如根据多个直达增益函数$g_{i,j}$来计算。特别地,例如可以使用两个增益函数$g_{i,1}(\varphi(k, n))$和$g_{i,2}(r(k,n))$,其中第一增益函数取决于DOA $\varphi(k, n)$,并且其中第二增益函数取决于距离r(k,n)。直达增益$G_i(k,n)$可以计算为:

[0341] $$G_i(k,n) = g_{i,1}(\varphi(k,n))\, g_{i,2}(r(k,n)) \qquad (31)$$

[0342] $$g_{i,1}(\varphi) = p_{b,i}(\varphi)\, w_b(\varphi) \qquad (32)$$

[0343] $$g_{i,2}(r) = b(r), \qquad (33)$$

[0344] 其中$p_{b,i}(\varphi)$表示平移增益函数(以确保声音从右方向再现),其中$w_b(\varphi)$是窗增益函数(以确保直达声音在源在视频中不可见的情况下被衰减),并且其中b(r)是模糊函数(在源不位于焦平面上的情况下对源进行声学模糊化)。

[0345] 应当注意,所有增益函数可以被定义为取决于频率(为了简洁在此省略)。还应当注意,在该实施例中,通过选择和乘以来自两个不同增益函数的增益来找到直达增益G_i,如公式(32)所示。

[0346] 两个增益函数$p_{b,i}(\varphi)$和$w_b(\varphi)$被如上所述类似地定义。例如,可以例如在增益函数计算模块104中使用公式(26)和(27)计算它们,并且它们保持固定,除非缩放因子β改变。上文已经提供了对这两个函数的详细描述。模糊函数b(r)返回导致源的模糊(例如,感知扩展)的复数增益,因此总增益函数g_i通常也将返回复数。为了简单起见,在下文中,将模糊表

示为到焦平面的距离的函数b(r)。

[0347] 可以获得模糊效果作为以下模糊效果中的选定的一个或组合:低通滤波、添加延迟的直达声音、直达声音衰减、时间平滑和/或DOA扩展。因此,根据实施例,信号处理器105可以例如被配置为通过进行低通滤波、或通过添加延迟的直达声音、或通过进行直达声音衰减、或通过进行时间平滑、或者通过进行到达方向扩展来生成一个或更多个音频输出信号。

[0348] 低通滤波:在视觉中,可以通过低通滤波获得非锐利的视觉图像,其有效地合并视觉图像中的相邻像素。类似地,可以通过对具有截止频率的直达声音的低通滤波来获得声学模糊效果,其中所述截止频率是基于源到焦平面r的估计距离来选择的。这种情况下,模糊函数b(r,k)针对频率k和距离r返回低通滤波器增益。图10(b)中示出了用于16kHz的采样频率的一阶低通滤波器的截止频率的示例曲线。对于小距离r,截止频率接近奈奎斯特频率,因此几乎没有有效地执行低通滤波。对于较大的距离值,截止频率减小,直到其在3kHz处稳定,此时声学图像被充分模糊。

[0349] 添加延迟的直达声音:为了钝化源的声学图像,我们可以例如通过在某个延迟τ(例如,在1和30ms之间)之后重复衰减直达声音来对直达声音进行去相关。这样的处理可以例如根据公式(34)的复数增益函数来进行:

[0350] $b(r,k)=1+a(r)e^{-j\omega\tau(r)}$ (34)

[0351] 其中a表示重复声音的衰减增益,τ是直达声音被重复之后的延迟。图10(c)中示出示例延迟曲线(以ms为单位)。对于小距离,不重复延迟的信号,并且将a设置为零。对于更大的距离,时间延迟随着距离的增加而增加,这导致声源的感知扩展。

[0352] 直达声衰减:当直达声音以常数因子衰减时,源也可以被感知为模糊的。在这种情况下,$b(r)=\mathrm{const}<1$。如上所述,模糊函数b(r)可以由任何所提到的模糊效应或这些效果的组合构成。此外,可以使用模糊源的备选处理。

[0353] 时间平滑:直达声音随时间的平滑可以例如用于感知地模糊声源。这可以通过随着时间对所提取的直达信号的包络进行平滑来实现。

[0354] DOA扩展:钝化声源的另一种方法在于仅从估计方向再现来自方向范围的源信号。这可以通过对角度进行随机化(例如通过从以估计的φ为中心的高斯分布取随机角度)来实现。增加这种分布的方差从而扩大可能的DOA范围,增加了模糊的感觉。

[0355] 与如上所述相类似地,在一些实施例中,在增益函数计算模块104中计算扩散增益函数q(β)可以仅需要知道可用于再现的扬声器I的数量。因此,在这种实施例中,可以根据应用的需要来设置扩散增益函数q(β)。例如,对于等间隔的扬声器,在增益选择单元202中基于缩放参数β选择公式(2a)中的实值扩散声音增益$Q\in\left[0,1/\sqrt{I}\right]$。使用扩散增益的目的是根据缩放因子衰减扩散声音,例如,缩放增加了再现信号的DRR。这通过针对较大的β而降低Q来实现。事实上,放大意味着相机的打开角度变小,例如,自然声学对应将是捕获较少扩散声音的更直达的麦克风。为了模拟这种效果,我们可以使用例如图8所示的增益函数。显然,增益函数也可以不同地定义。可选地,通过对在公式(2b)中获得的$Y_{diff}(k,n)$进行去相关来获得第i个扬声器声道的最终扩散声音$Y_{diff,i}(k,n)$。

[0356] 现在,考虑实现针对助听器和助听设备的应用的实施例。图11示出了这种助听器应用。

[0357] 一些实施例涉及双耳助听器。在这种情况下，假设每个助听器配备有至少一个麦克风，并且可以在两个助听器之间交换信息。由于一些听力损失，听觉受损的人可能难以对期望的声音进行聚焦(例如，集中于来自特定点或方向的声音)。为了帮助听力受损人士的大脑处理由助听器再现的声音，使声学图像与助听器用户的焦点或方向一致。可以想到，焦点或方向是预定义的，用户定义的或由脑机接口定义的。这样的实施例确保期望的声音(假定从焦点或聚焦方向到达)和不期望的声音在空间上分离。

[0358] 在这样的实施例中，可以以不同的方式估计直达声音的方向。根据实施例，基于使用两个助听器(参见[15]和[16])确定的耳间电平差(ILD)和/或耳间时间差(ITD)来确定方向。

[0359] 根据其他实施例，使用配备有至少两个麦克风的助听器独立地估计左侧和右侧的直达声音的方向(参见[17])。基于左右助听器处的声压级或左右助听器处的空间相干性，可以确定(fuss)估计的方向。由于头部遮蔽效应，可以对不同的频带(例如，在高频处的ILD和在低频处的ITD)采用不同的估计器。

[0360] 在一些实施例中，直达声音信号和扩散声音信号可以例如使用上述通知的空间滤波技术来估计。在这种情况下，可以(例如，通过改变参考麦克风)单独地估计在左、右助听器处接收的直达和扩散声音，或者可以以与在先前实施例中获得不同扬声器或耳机信号相类似的方式，分别使用用于左、右助听器输出的增益函数来生成左、右输出信号。

[0361] 为了在空间上分离期望的声音和非期望的声音，可以应用在上述实施例中说明的声学缩放。在这种情况下，对焦点或对焦方向决定了缩放因子。

[0362] 因此，根据实施例，可以提供助听器或助听设备，其中助听器或助听设备包括如上所述的系统，其中上述系统的信号处理器105例如根据聚焦方向或聚焦点，针对一个或更多个音频输出信号中的每一个确定直达增益。

[0363] 在实施例中，上述系统的信号处理器105可以例如被配置为接收缩放信息。上述系统的信号处理器105例如可以被配置为根据窗增益函数生成一个或更多个音频输出信号的每个音频输出信号，其中窗增益函数取决于缩放信息。采用与参考图7(a)、7(b)和7(c)解释的相同的构思。

[0364] 如果取决于聚焦方向或聚焦点的窗函数自变量值大于下阈值并且小于上阈值，则窗增益函数被配置为返回比在窗函数自变量值小于下阈值或大于上阈值的情况下由所述窗增益函数返回的任何窗增益大的窗增益。

[0365] 例如，在聚焦方向的情况下，聚焦方向本身可以是窗函数自变量(因此，窗函数自变量取决于聚焦方向)。在聚焦位置的情况下，可以例如从聚焦位置导出窗函数自变量。

[0366] 类似地，本发明可以应用于包括辅助收听设备或诸如Google眼镜之类的设备的其他可穿戴设备。应当注意，一些可穿戴设备还配备有一个或更多个相机或ToF传感器，其可以用于估计物体到佩戴该设备的人的距离。

[0367] 虽然已经在装置的上下文中描述了一些方面，但是将清楚的是，这些方面还表示对相应方法的描述，其中，框或设备对应于方法步骤或方法步骤的特征。类似地，在方法步骤的上下文中描述的方案也表示对相应块或项或者相应装置的特征的描述。

[0368] 创造性的分解信号可以存储在数字存储介质上，或者可以在诸如无线传输介质或有线传输介质(例如，互联网)等的传输介质上传输。

[0369] 取决于某些实现要求，可以在硬件中或在软件中实现本发明的实施例。可以使用其上存储有电子可读控制信号的数字存储介质(例如，软盘、DVD、CD、ROM、PROM、EPROM、EEPROM或闪存)来执行该实现，该电子可读控制信号与可编程计算机系统协作(或者能够与之协作)从而执行相应方法。

[0370] 根据本发明的一些实施例包括具有电子可读控制信号的非暂时性数据载体，该电子可读控制信号能够与可编程计算机系统协作从而执行本文所述的方法之一。

[0371] 通常，本发明的实施例可以实现为具有程序代码的计算机程序产品，程序代码可操作以在计算机程序产品在计算机上运行时执行方法之一。程序代码可以例如存储在机器可读载体上。

[0372] 其他实施例包括存储在机器可读载体上的计算机程序，该计算机程序用于执行本文所述的方法之一。

[0373] 换言之，本发明方法的实施例因此是具有程序代码的计算机程序，该程序代码用于在计算机程序在计算机上运行时执行本文所述的方法之一。

[0374] 因此，本发明方法的另一实施例是其上记录有计算机程序的数据载体(或者数字存储介质或计算机可读介质)，该计算机程序用于执行本文所述的方法之一。

[0375] 因此，本发明方法的另一实施例是表示计算机程序的数据流或信号序列，所述计算机程序用于执行本文所述的方法之一。数据流或信号序列可以例如被配置为经由数据通信连接(例如，经由互联网)传递。

[0376] 另一实施例包括处理装置，例如，计算机或可编程逻辑器件，所述处理装置被配置为或适于执行本文所述的方法之一。

[0377] 另一实施例包括其上安装有计算机程序的计算机，该计算机程序用于执行本文所述的方法之一。

[0378] 在一些实施例中，可编程逻辑器件(例如，现场可编程门阵列)可以用于执行本文所述的方法的功能中的一些或全部。在一些实施例中，现场可编程门阵列可以与微处理器协作以执行本文所述的方法之一。通常，方法优选地由任意硬件装置来执行。

[0379] 上述实施例对于本发明的原理仅是说明性的。应当理解的是：本文所述的布置和细节的修改和变形对于本领域其他技术人员将是显而易见的。因此，旨在仅由所附专利权利要求的范围来限制而不是由借助对本文的实施例的描述和解释所给出的具体细节来限制。

[0380] 参考文献

[0381] Y.Ishigaki,M.Yamamoto,K.Totsuka,and N.Miyaji,“Zoom microphone,”in Audio Engineering Society Convention 67,Paper 1713,October 1980.

[0382] M.Matsumoto,H.Naono,H.Saitoh,K.Fujimura,and Y.Yasuno,“Stereo zoom microphone for consumer video cameras,”Consumer Electronics,IEEE Transactions on,vol.35,no.4,pp.759-766,November 1989.August 13,2014

[0383] T.van Waterschoot,W.J.Tirry,and M.Moonen,“Acoustic zooming by multi microphone sound scene manipulation,”J.Audio Eng.Soc,vol.61,no.7/8,pp.489-507,2013.

[0384] V.Pulkki,“Spatial sound reproduction with directional audio coding,”

J.Audio Eng.Soc,vol.55,no.6,pp.503-516,June 2007.

[0385] R.Schultz-Amling,F.Kuech,O.Thiergart,and M.Kallinger,"Acoustical zooming based on a parametric sound field representation,"in Audio Engineering Society Convention 128,Paper 8120,London UK,May 2010.

[0386] O.Thiergart,G.Del Galdo,M.Taseska,and E.Habets,"Geometry-based spatial sound acquisition using distributed microphone arrays,"Audio,Speech, and Language Processing,IEEE Transactiohs on,vol.21,no.12,pp.2583-2594, December 2013.

[0387] K.Kowalczyk,O.Thiergart,A.Craciun,and E.A.P.Habets,"Sound acquisition in noisy and reverberant environments using virtual microphones,"in Applications of Signal Processing to Audio and Acoustics(WASPAA),2013 IEEE Workshop on,October 2013.

[0388] O.Thiergart and E.A.P.Habets,"An informed LCMV filter based on multiple instantaneous direction-of-arrival estimates,"in Acoustics Speech and Signal Processing(ICASSP),2013 IEEE International Conference on,2013, pp.659-663.

[0389] O.Thiergart and E.A.P.Habets,"Extracting reverberant sound using a linearly constrained minimum variance spatial filter,"Signal Processing Letters,IEEE,vol.21,no.5,pp.630-634,May 2014.

[0390] R.Roy and T.Kailath,"ESPRIT-estimation of signal parameters viarotational invariance techniques,"Acoustics,Speech and Signal Processing, IEEE Transactions on,vol.37,no.7,pp.984-995,July 1989.

[0391] 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.

[0392] H.Teutsch and G.Elko,"An adaptive close-talking microphone array,"in Applications of Signal Processing to Audio and Acoustics,2001 IEEE Workshop on the,2001,pp.163-166.

[0393] O.Thiergart,G.D.Galdo,and E.A.P.Habets,"On the spatial coherence in mixed sound fields and its application to signal-to-diffuse ratio estimation,"The Journal of the Acoustical Society of America,vol.132,no.4, pp.2337-2346,2012.

[0394] V.Pulkki,"Virtual sound source positioning using vector base amplitude panning,"J.Audio Eng.Soc,vol.45,no.6,pp.456-466,1997.

[0395] J.Blauert,Spatial hearing,3rd ed.Hirzel-Verlag,2001.

[0396] T.May,S.van de Par,and A.Kohlrausch,"A probabilistic model for robust localization based on a binaural auditory front-end,"IEEE Trans.Audio,Speech, Lang.Process.,vol.19,no.1,pp.1-13,2011.

[0397] J.Ahonen,V.Sivonen,and V.Pulkki,"Parametric spatial sound processing

applied to bilateral hearing aids,"in AES 45th International Conference, Mar.2012.

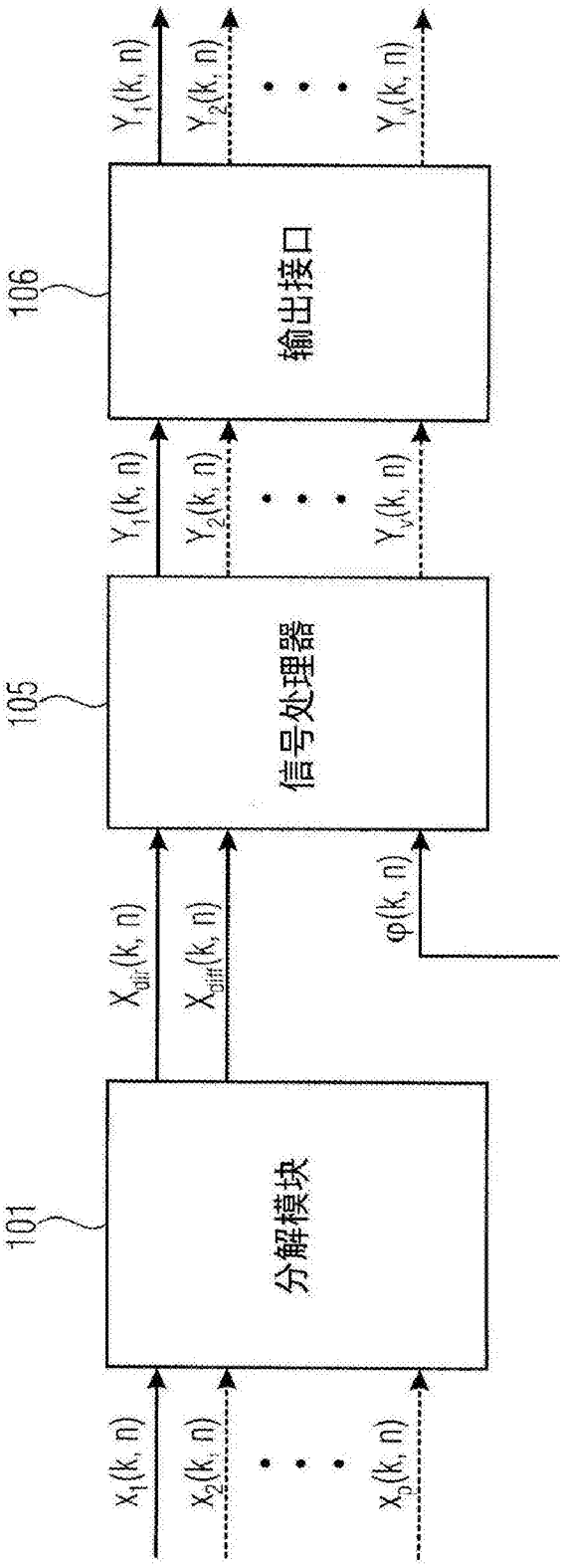

101
分解模块
105
信号处理器
106
输出接口
$X_1(k,n)$
$X_2(k,n)$
$X_5(k,n)$
$X_{dir}(k,n)$
$X_{diff}(k,n)$
$\varphi(k,n)$
$Y_1(k,n)$
$Y_2(k,n)$
$Y_v(k,n)$

图1A

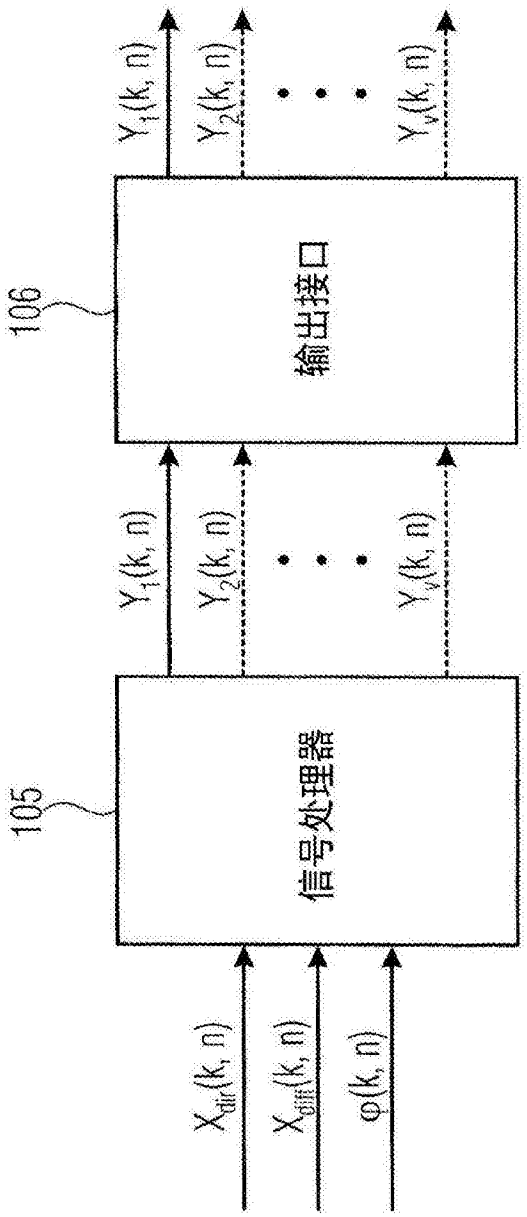

105
信号处理器
106
输出接口
$X_{dir}(k,n)$
$X_{diff}(k,n)$
$\varphi(k,n)$
$Y_1(k,n)$
$Y_2(k,n)$
$Y_v(k,n)$

图1B

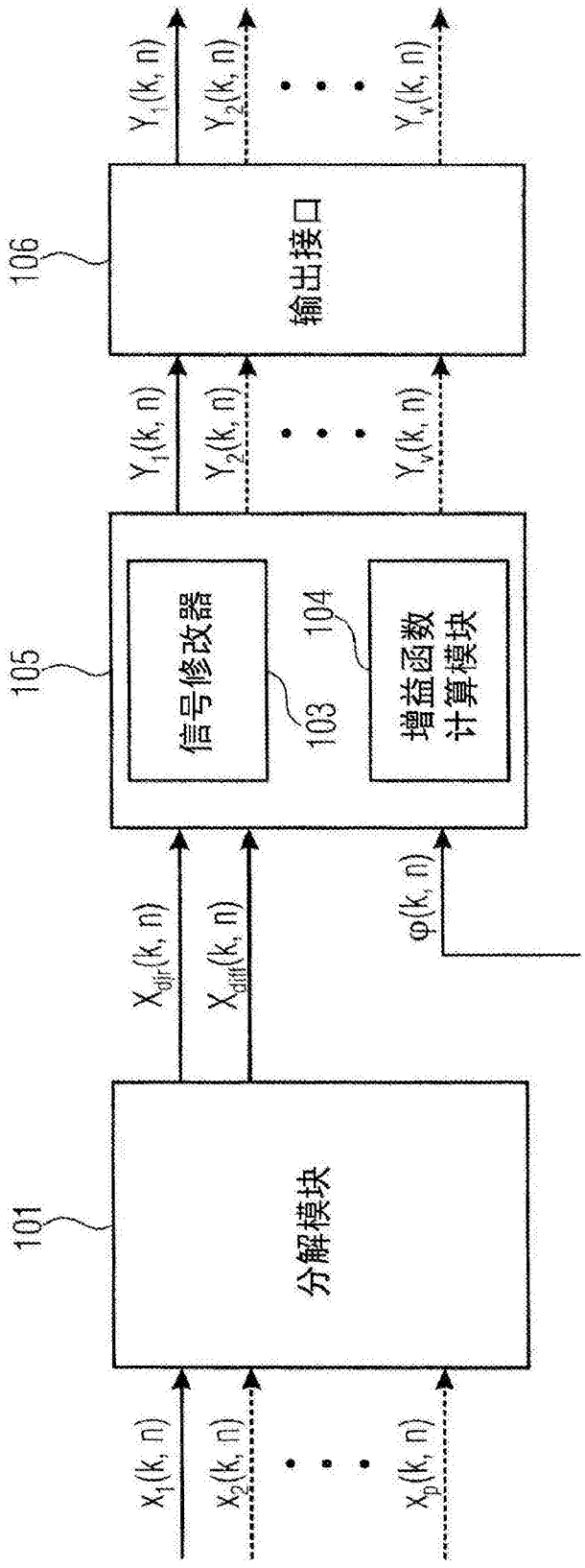

101
分解模块
105
信号修改器
103
104
增益函数
计算模块
106
输出接口
$X_1(k, n)$
$X_2(k, n)$
$X_p(k, n)$
$X_{dir}(k, n)$
$X_{diff}(k, n)$
$\varphi(k, n)$
$Y_1(k, n)$
$Y_2(k, n)$
$Y_v(k, n)$

图1C

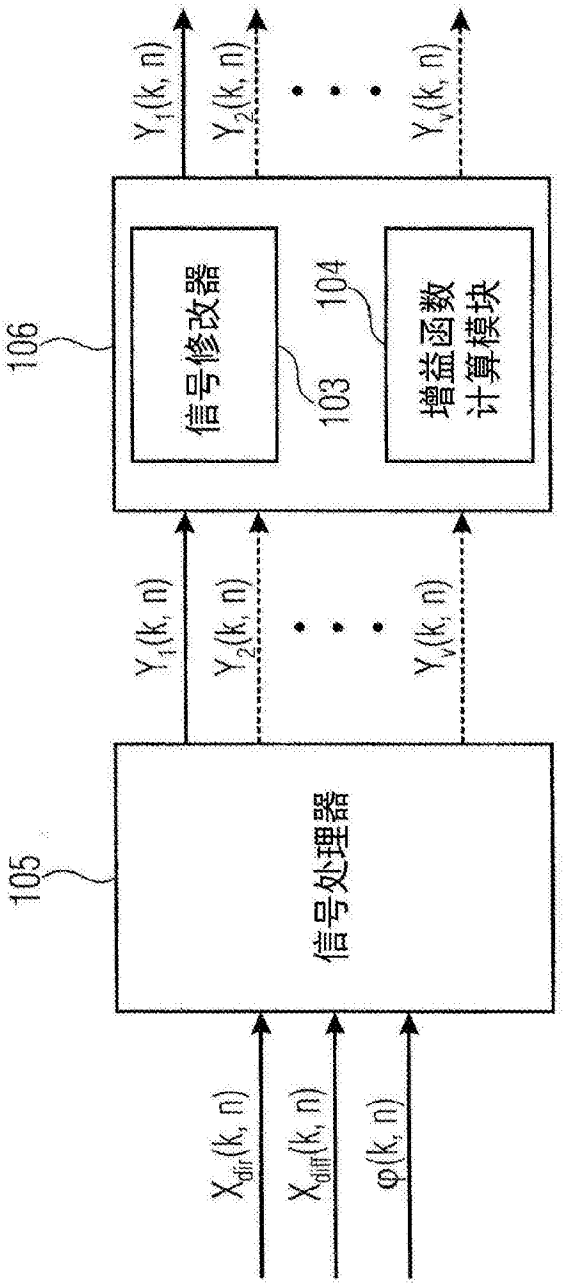

105
信号处理器
106
信号修改器
103
104
增益函数
计算模块
$X_{dir}(k, n)$
$X_{diff}(k, n)$
$\varphi(k, n)$
$Y_1(k, n)$
$Y_2(k, n)$
$Y_v(k, n)$

图1D

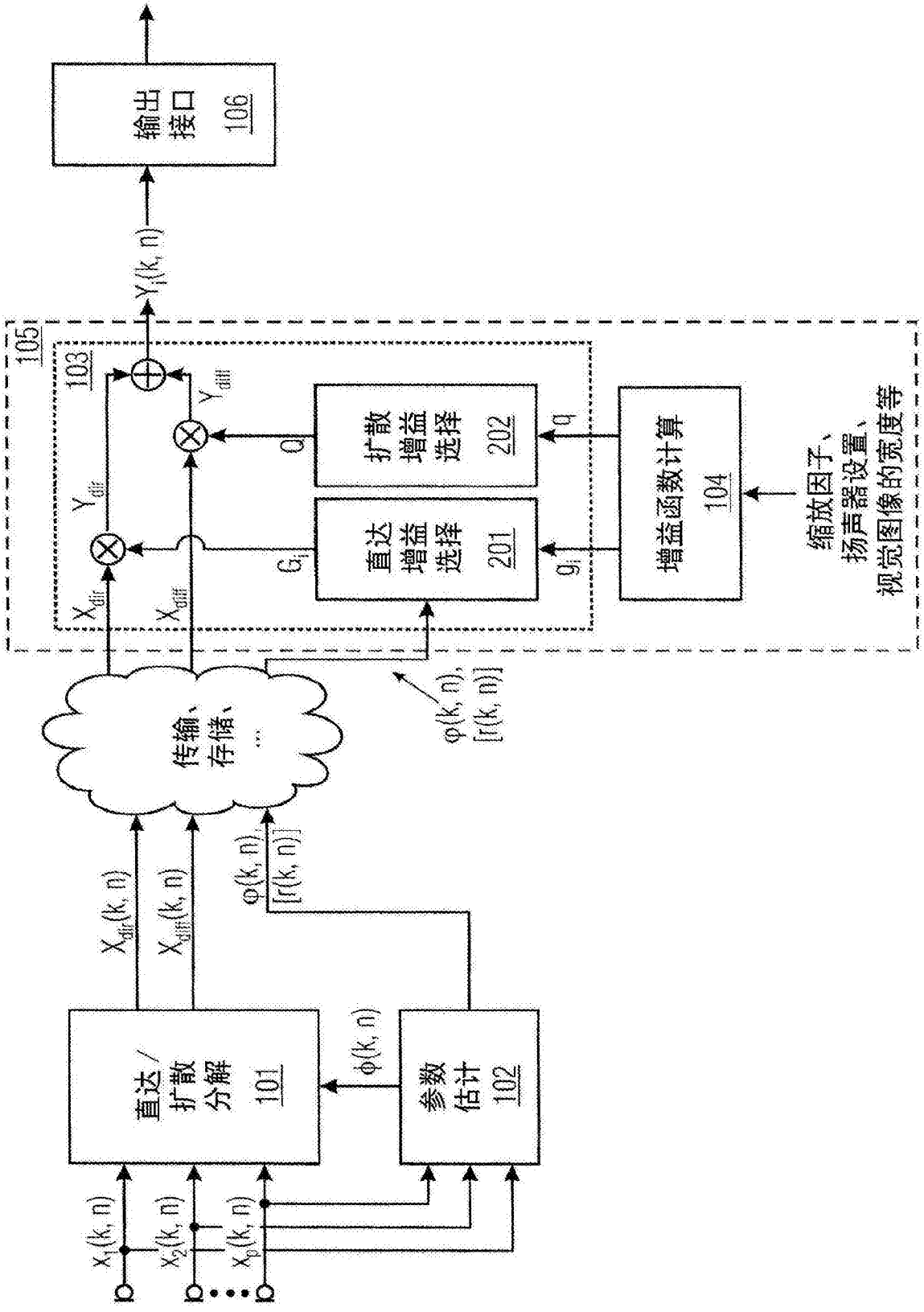

输出接口 106
$Y_1(k, n)$
105
103
Y_{diff}
Y_{dir}
X_{dir}
X_{diff}
Q
q
G_1
g_1
扩散增益选择 202
直达增益选择 201
增益函数计算 104
缩放因子、扬声器设置、视觉图像的宽度等
传输、存储、...
$\varphi(k, n)$, $[r(k, n)]$
$X_{dir}(k, n)$
$X_{diff}(k, n)$
$\varphi(k, n)$, $[r(k, n)]$
直达/扩散分解 101
$\phi(k, n)$
参数估计 102
$x_1(k, n)$
$x_2(k, n)$
$x_p(k, n)$

图2

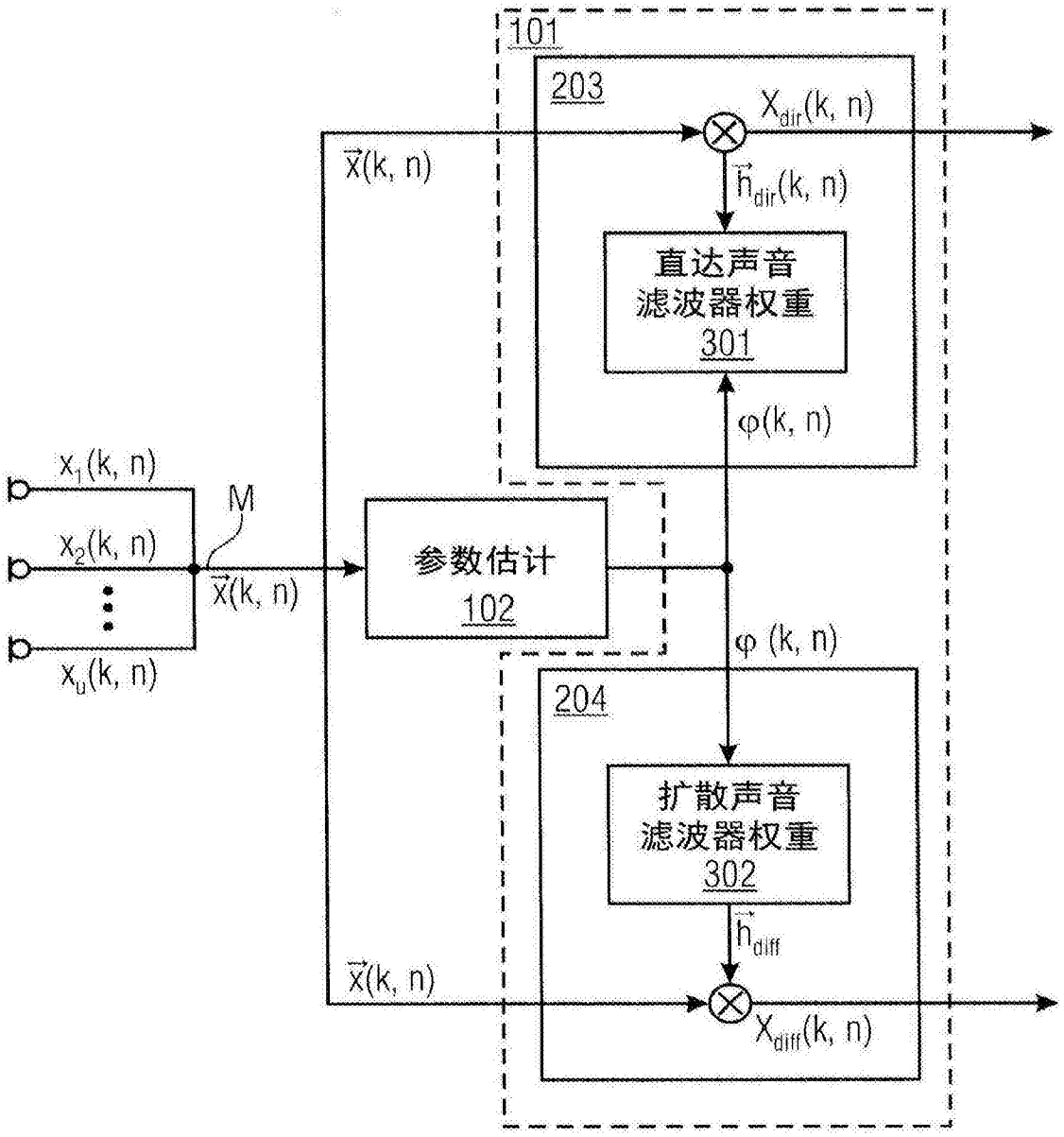

101
203
$X_{dir}(k, n)$
$\vec{x}(k, n)$
$\vec{h}_{dir}(k, n)$
直达声音
滤波器权重
301
$\varphi(k, n)$
$x_1(k, n)$
$x_2(k, n)$
$x_U(k, n)$
M
$\vec{x}(k, n)$
参数估计
102
$\varphi(k, n)$
204
扩散声音
滤波器权重
302
$\vec{h}_{diff}$
$\vec{x}(k, n)$
$X_{diff}(k, n)$

图3

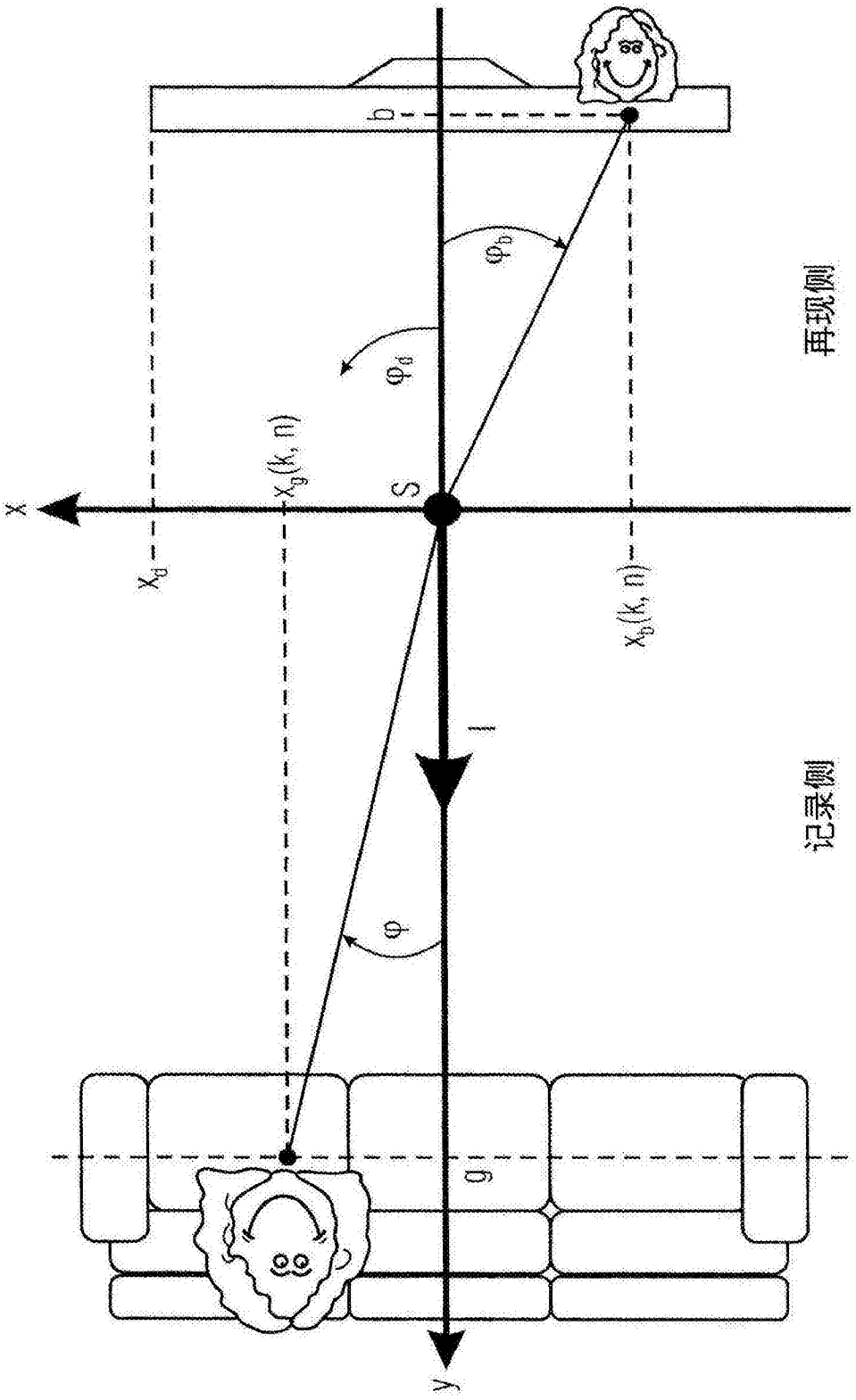


图4

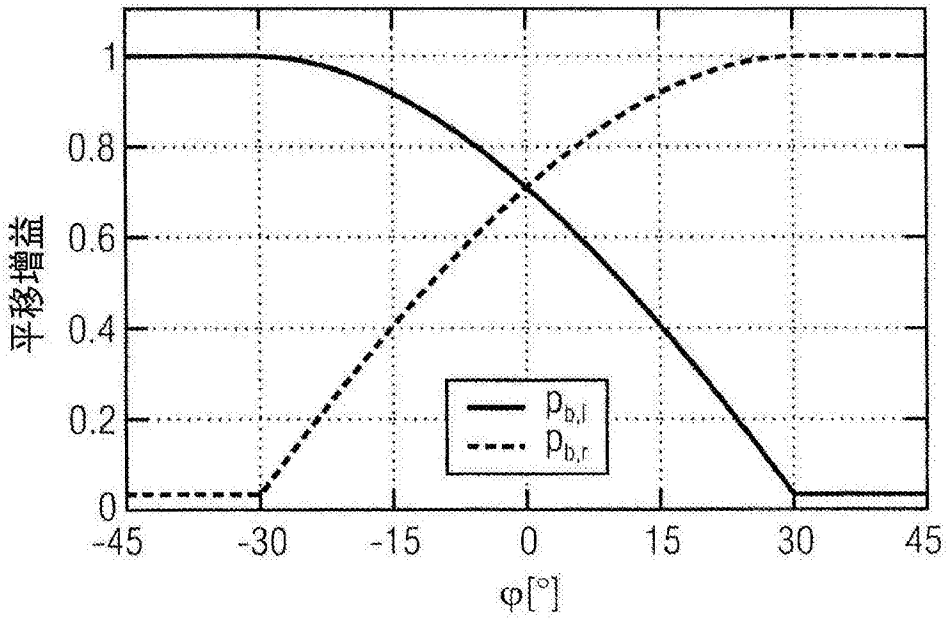

1
0.8
0.6
0.4
0.2
0
平移增益
-45
-30
-15
0
15
30
45
φ[°]
$p_{b,l}$
$p_{b,r}$
VBAP 平移函数

图5A

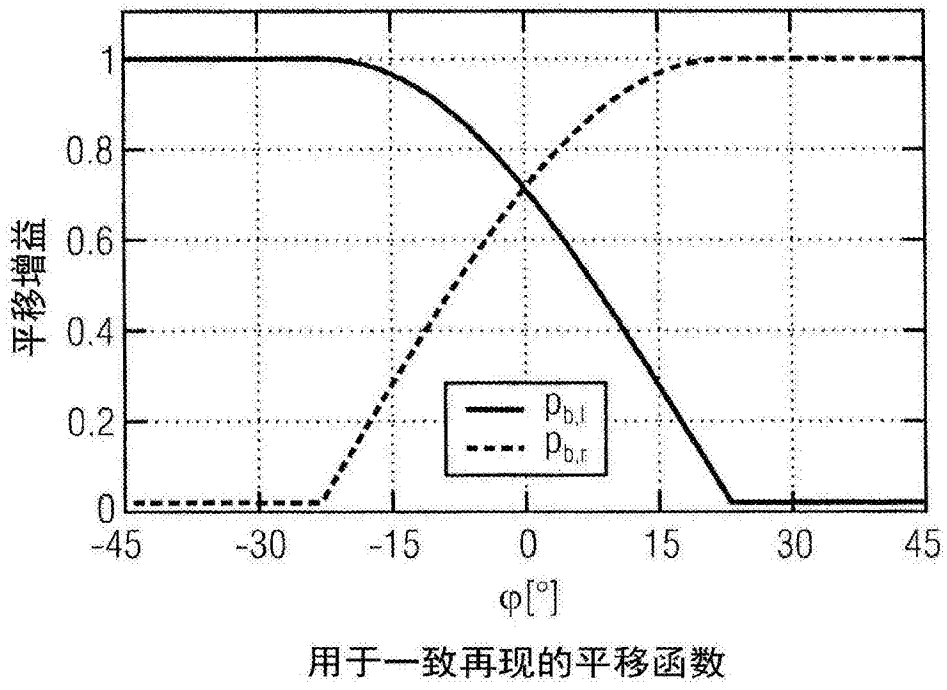

1
0.8
0.6
0.4
0.2
0
平移增益
-45
-30
-15
0
15
30
45
φ[°]
$p_{b,l}$
$p_{b,r}$
用于一致再现的平移函数

图5B

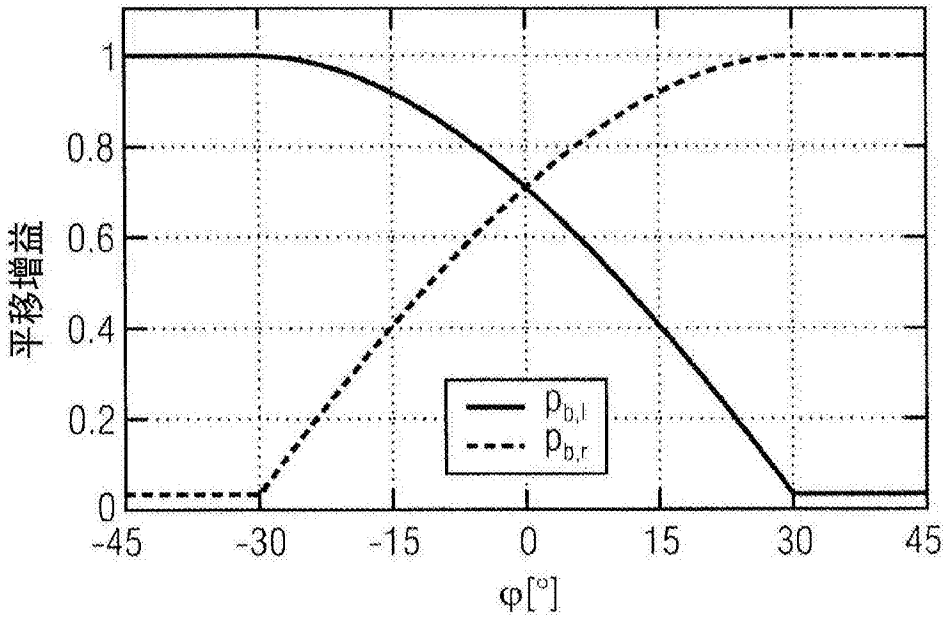

平移增益
1
0.8
0.6
0.4
0.2
0
-45
-30
-15
0
15
30
45
φ[°]
$p_{b,l}$
$p_{b,r}$
VBAP平移函数

图6A

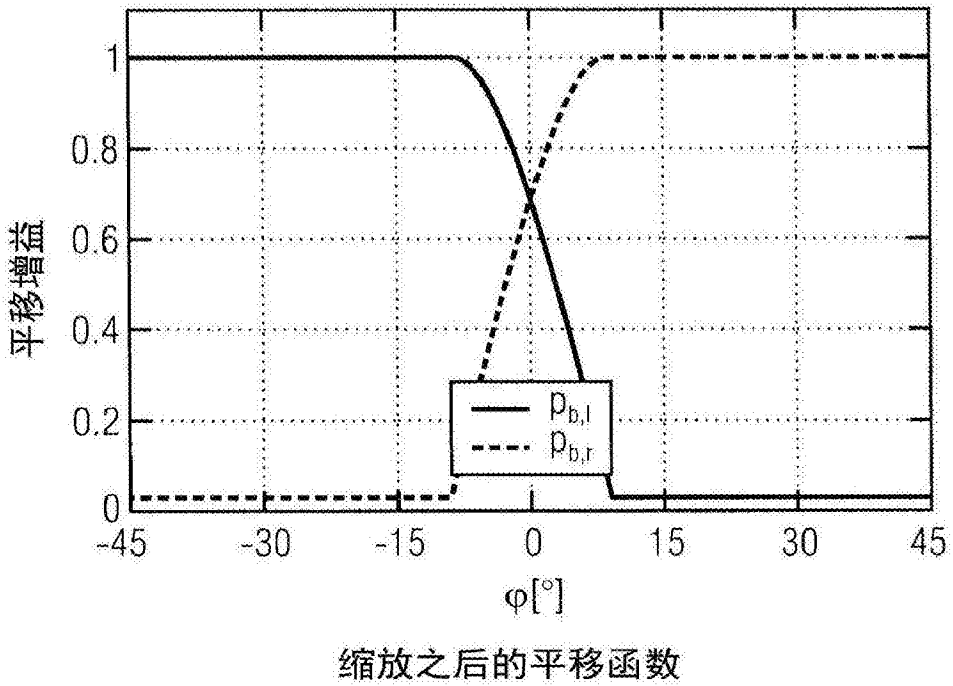

平移增益
1
0.8
0.6
0.4
0.2
0
-45
-30
-15
0
15
30
45
φ[°]
$p_{b,l}$
$p_{b,r}$
缩放之后的平移函数

图6B

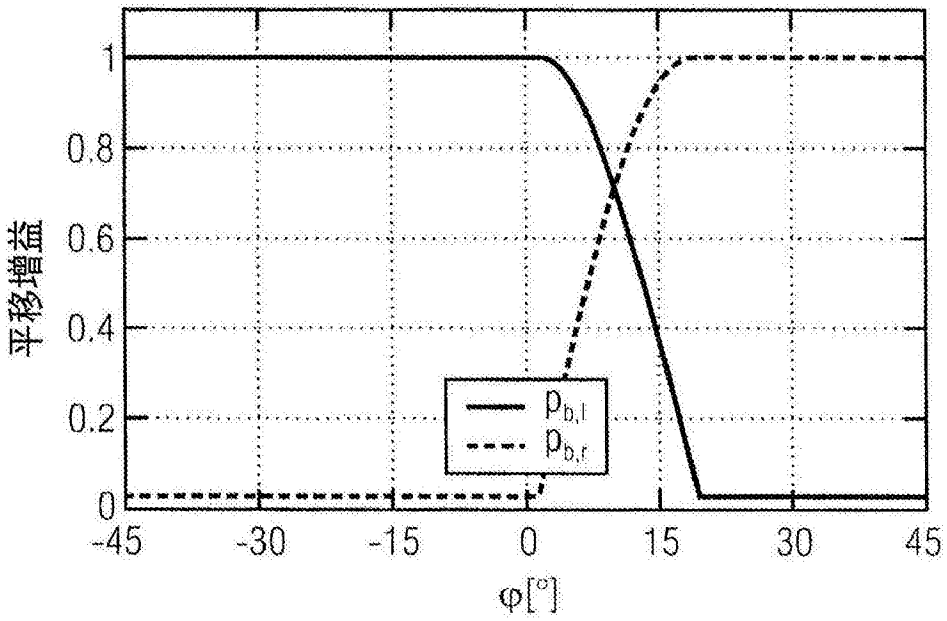

平移增益
1
0.8
0.6
0.4
0.2
0
-45
-30
-15
0
15
30
45
φ[°]
$p_{b,l}$
$p_{b,r}$
具有位移的缩放之后的平移函数

图6C

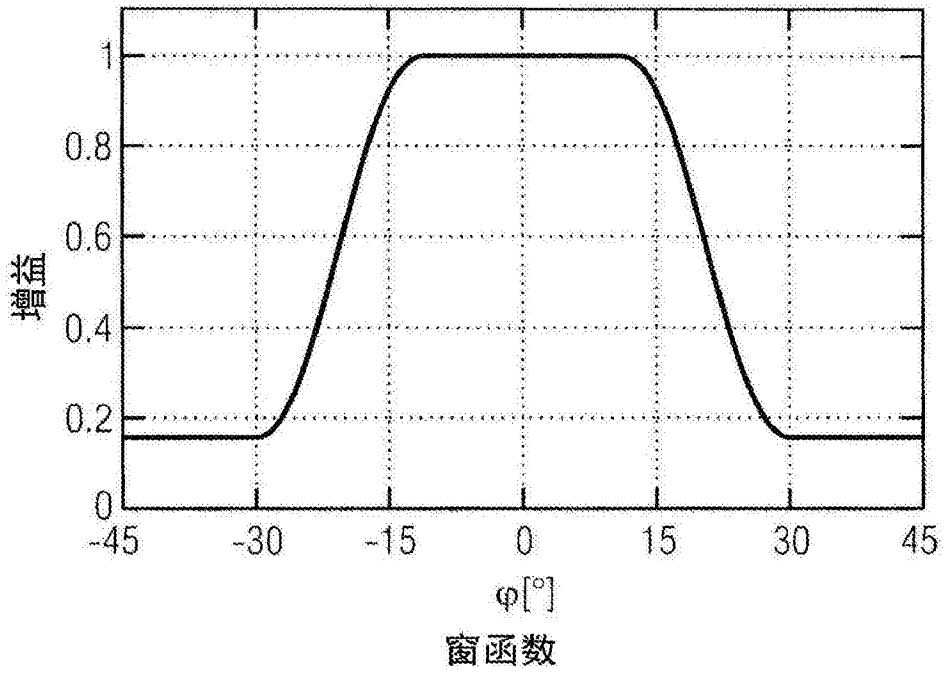

增益
1
0.8
0.6
0.4
0.2
0
-45
-30
-15
0
15
30
45
φ[°]
窗函数

图7A

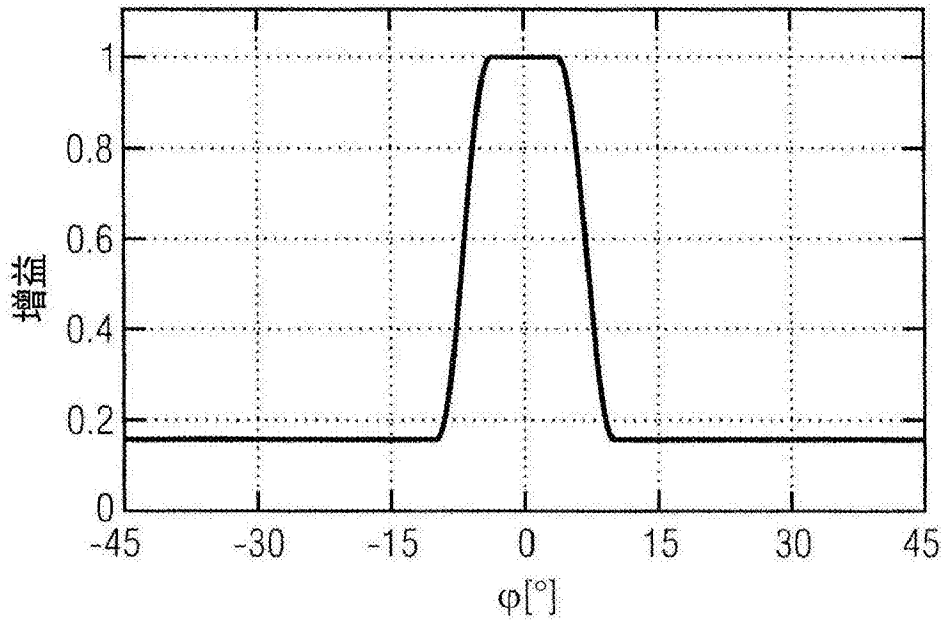

1
0.8
0.6
0.4
0.2
0
幅值
-45
-30
-15
0
15
30
45
φ[°]
缩放之后的窗函数

图7B

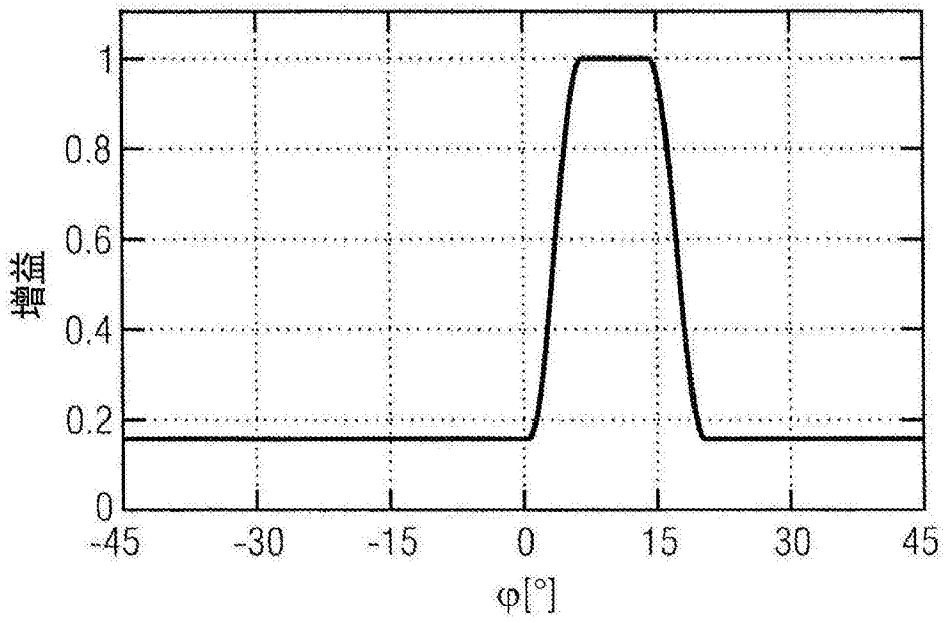

1
0.8
0.6
0.4
0.2
0
幅值
-45
-30
-15
0
15
30
45
φ[°]
具有位移的缩放之后的窗函数

图7C

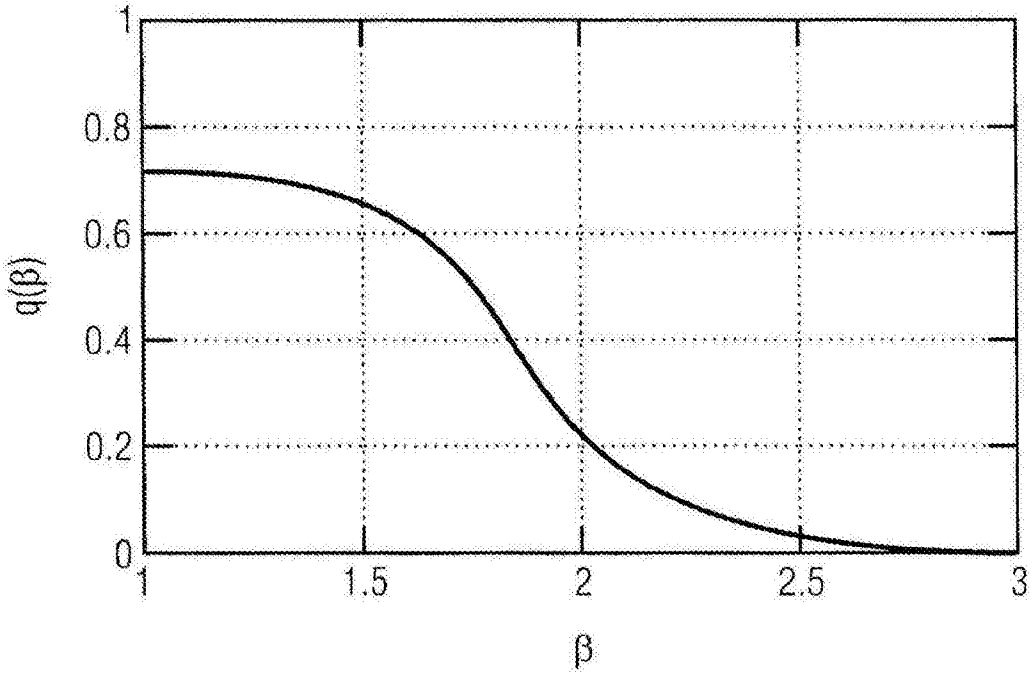


图8

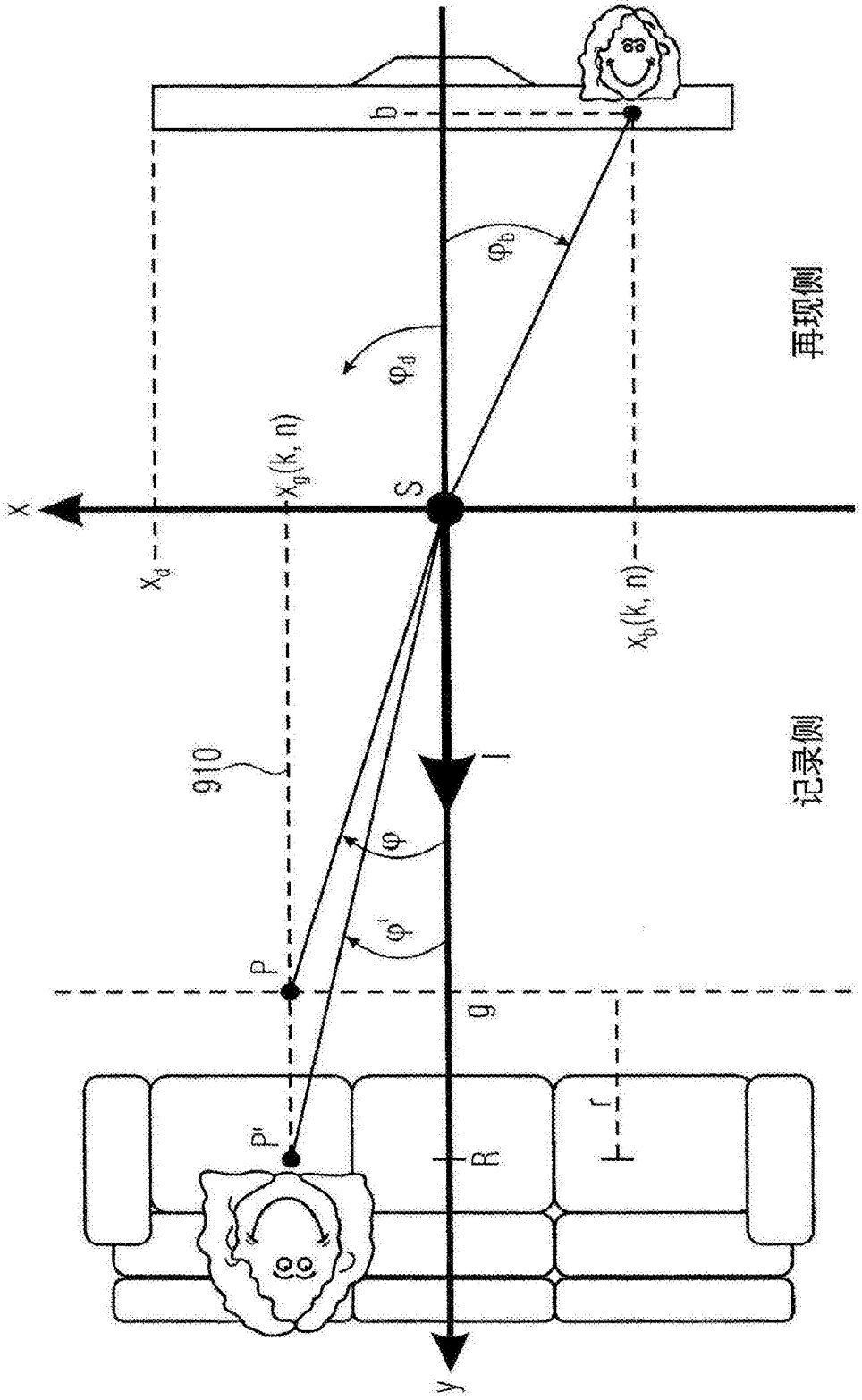

b
φ_b
φ_d
再现侧
$x_g(k, n)$
S
x
x_d
$x_b(k, n)$
910
记录侧
l
φ
φ'
P
g_l
P'
R
r
y

图9

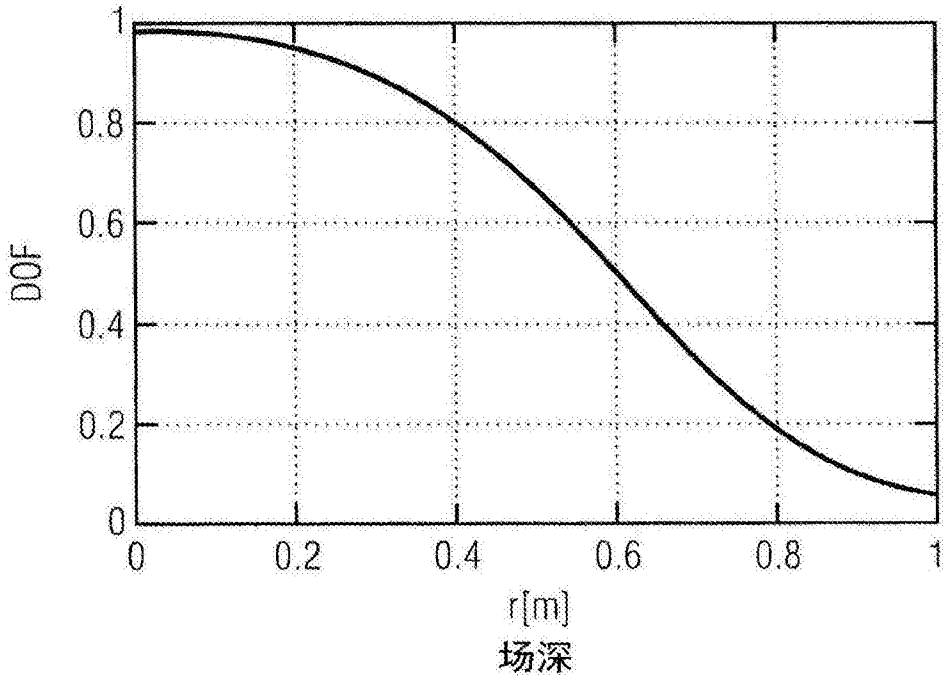

DOF
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
r[m]
场深

图10A

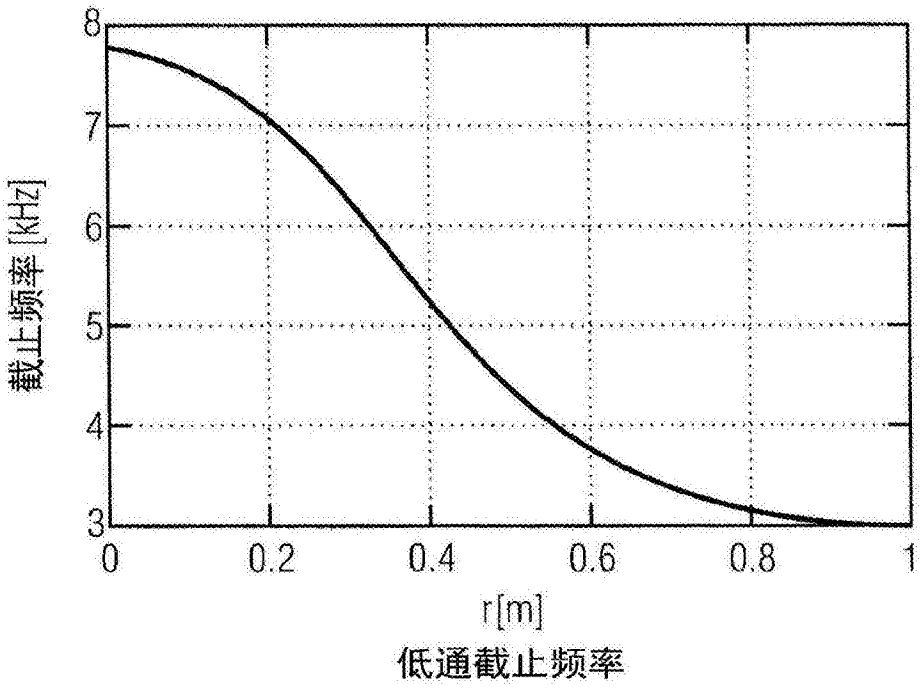

截止频率[kHz]
8
7
6
5
4
3
0
0.2
0.4
0.6
0.8
1
r[m]
低通截止频率

图10B

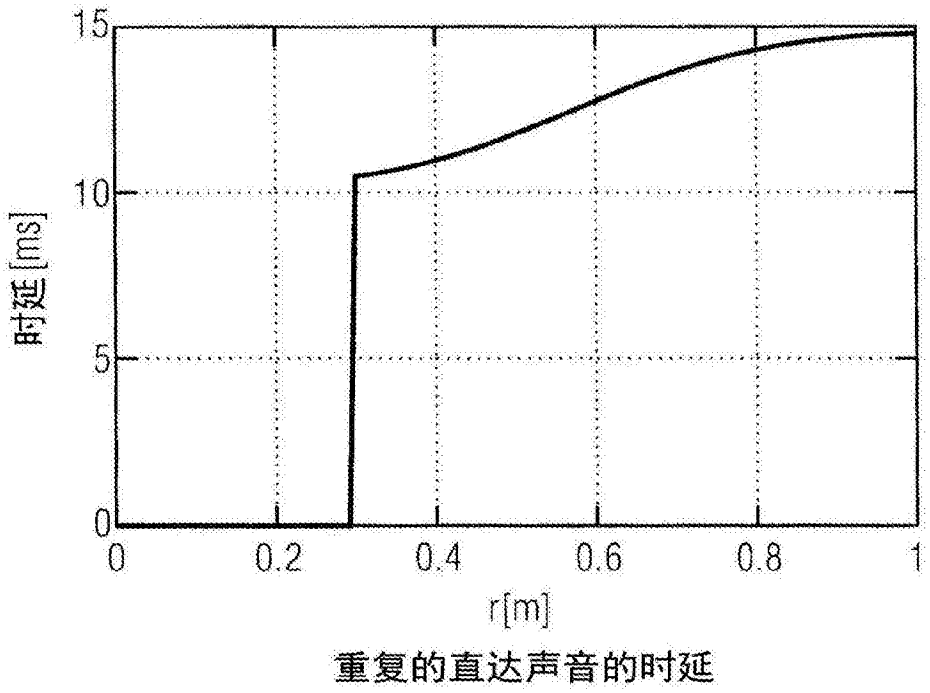


图10C

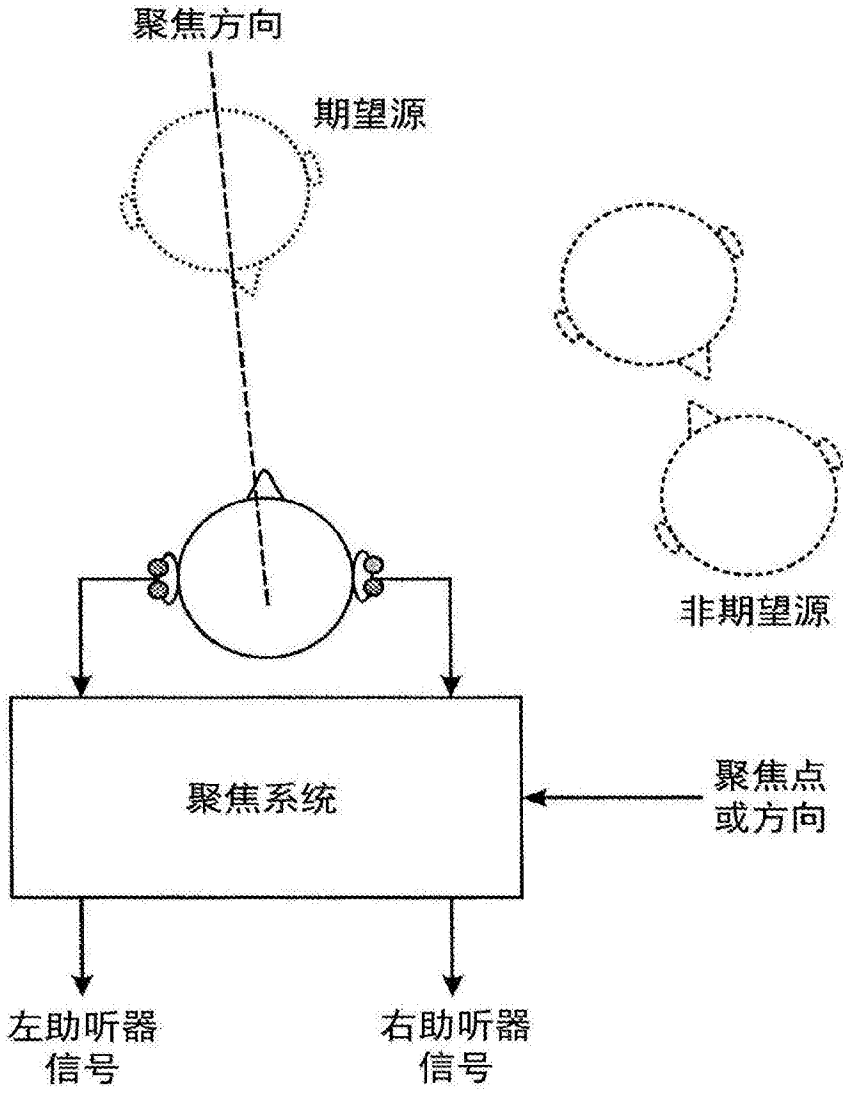


图11