



(12)发明专利

(10)授权公告号 CN 109448743 B

(45)授权公告日 2020.03.10

(21)申请号 201910024898.9

(22)申请日 2013.12.04

(65)同一申请的已公布的文献号

申请公布号 CN 109448743 A

(43)申请公布日 2019.03.08

(30)优先权数据

12306569.0 2012.12.12 EP

(62)分案原申请数据

201380064856.9 2013.12.04

(73)专利权人 杜比国际公司

地址 荷兰阿姆斯特丹

(72)发明人 亚历山大·克鲁格 斯文·科登

约翰内斯·伯姆

(74)专利代理机构 中国国际贸易促进委员会专利商标事务所 11038

代理人 宿小猛

(51)Int.Cl.

G10L 19/008(2013.01)

H04S 3/00(2006.01)

H04S 7/00(2006.01)

(56)对比文件

US 2012155653 A1,2012.06.21,

EP 2268064 A1,2010.12.29,

CN 101606192 A,2009.12.16,

CN 101138274 A,2008.03.05,

EP 2469742 A2,2012.06.27,

审查员 刘红梅

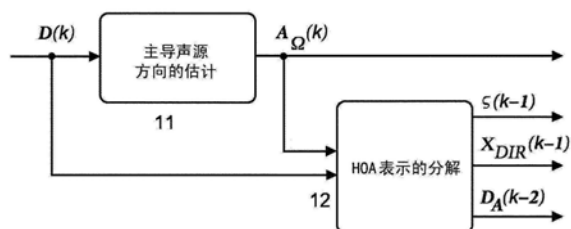
权利要求书2页 说明书14页 附图5页

(54)发明名称

对声场的高阶立体混响表示进行压缩和解压缩的方法和设备

(57)摘要

本公开涉及对声场的高阶立体混响表示进行压缩和解压缩的方法和设备。本发明改善HOA声场表示压缩。针对主导声源的存在来对HOA表示进行分析,并且估计所述主导声源的方向。然后,HOA表示分解为多个主导定向信号和残余分量。该残余分量变换到离散空间域,以便在均匀采样方向获得总的平面波函数,所述均匀采样方向是根据主导定向信号中预测的。最后,预测误差变换回HOA域,并且表示残余环境HOA分量,针对所述残余环境HOA分量来执行阶的降低,随后是主导定向信号和残余分量的感知编码。



1. 一种用于压缩声场的高阶立体混响HOA表示的方法,所述方法包括:

根据HOA系数的当前时间帧,估计主导声源方向;

将所述HOA表示分解为时域中的主导定向信号、和残余HOA分量,其中,所述残余HOA分量变换到离散空间域,以便获得表示所述残余HOA分量的在均匀采样方向上的平面波函数,并且其中根据所述主导定向信号预测所述平面波函数,由此提供描述所述预测的参数;

对降低阶的残余HOA分量进行去相关,以获得对应的残余HOA分量时域信号;

对所述主导定向信号和所述残余HOA分量时域信号进行感知编码,以确定压缩的主导定向信号和压缩的残余分量信号。

2. 根据权利要求1所述的方法,其中,所述分解包括:

根据HOA系数的当前帧的估计的声源方向来计算主导定向信号;

对所述主导定向信号进行时间平滑以确定平滑的主导定向信号;

根据估计的声源方向和平滑的主导定向信号来计算平滑的主导定向信号的HOA表示;

通过均匀网格上的定向信号来表示对应的残余HOA表示;

根据所述平滑的主导定向信号和由定向信号表示的所述残余HOA表示,预测均匀网格上的定向信号,由此计算预测的均匀网格上的定向信号的HOA表示,然后进行时间平滑;

根据平滑的预测的均匀网格上的定向信号、HOA系数的当前帧的二帧延时版本、和平滑的主导定向信号的帧延时版本,计算残余环境声场分量的HOA表示。

3. 一种用于压缩声场的高阶立体混响HOA表示的设备,所述设备包括:

估计器,所述估计器根据HOA系数的当前时间帧,估计主导声源方向;

分解器,所述分解器将所述HOA表示分解为时域中的主导定向信号、和残余HOA分量,其中,所述残余HOA分量变换到离散空间域,以便获得表示所述残余HOA分量的在均匀采样方向上的平面波函数,并且其中根据所述主导定向信号预测所述平面波函数,由此提供描述所述预测的参数;

去相关器,所述去相关器对降低阶的残余HOA分量进行去相关,以获得对应的残余HOA分量时域信号;

编码器,所述编码器对所述主导定向信号和所述残余HOA分量时域信号进行感知编码,以提供压缩的主导定向信号和压缩的残余分量信号。

4. 根据权利要求3所述的设备,其中,所述分解器进一步配置为:

根据HOA系数的当前帧的估计的声源方向来计算主导定向信号;

对所述主导定向信号进行时间平滑,以获得平滑的主导定向信号;

根据估计的声源方向和平滑的主导定向信号来计算平滑的主导定向信号的HOA表示;

通过均匀网格上的定向信号来表示对应的残余HOA表示;

根据所述平滑的主导定向信号和由定向信号表示的所述残余HOA表示,预测均匀网格上的定向信号,由此计算预测的均匀网格上的定向信号的HOA表示,然后进行时间平滑;

根据经平滑的预测的均匀网格上的定向信号、HOA系数的当前帧的二帧延时版本、和平滑的主导定向信号的帧延时版本,计算残余环境声场分量的HOA表示。

5. 一种用于解压缩被压缩的高阶立体混响HOA表示的方法,所述方法包括:

对压缩的主导定向信号和压缩的残余分量信号进行感知解码,从而提供解压缩的主导定向信号和表示空间域中的残余HOA分量的解压缩的时域信号;

对所述解压缩的时域信号进行重新相关,以获得对应的降低阶的残余HOA分量;
通过将所述对应的降低阶的残余HOA分量增大到原始的阶来提供解压缩的残余HOA分量;

基于至少一个参数确定预测的定向信号;

基于所述解压缩的主导定向信号、所述预测的定向信号和所述解压缩的残余HOA分量,确定HOA声场表示。

6. 一种用于解压缩高阶立体混响HOA表示的设备,所述设备包括:

解码器,所述解码器对压缩的主导定向信号和压缩的残余分量信号进行感知解码,从而提供解压缩的主导定向信号和表示空间域中的残余HOA分量的解压缩的时域信号;

重新相关器,所述重新相关器对所述解压缩的时域信号进行重新相关,以获得对应的降低阶的残余HOA分量;

处理器,所述处理器被配置为通过将所述对应的降低阶的残余HOA分量增大到原始阶来提供解压缩的残余HOA分量,所述处理器进一步配置为基于至少一个参数确定预测的定向信号;

其中,所述处理器进一步配置为基于所述解压缩的主导定向信号、所述预测的定向信号和所述解压缩的残余HOA分量,确定HOA声场表示。

对声场的高阶立体混响表示进行压缩和解压缩的方法和设备

[0001] 本申请是申请号为201380064856.9、申请日为2013年12月4日、发明名称为“对声场的高阶立体混响表示进行压缩和解压缩的方法和设备”的发明专利申请的分案申请。

技术领域

[0002] 本发明涉及对声场的高阶立体混响表示进行压缩和解压缩的方法和设备。

背景技术

[0003] 高阶立体混响(表示为HOA)提供了表示三维立体声的一种方式。其它的技术是波场合成(WFS)或者像22.2的基于声道的方法。相比于基于声道的方法,HOA表示提供了独立于特定扬声器配置的优点。然而,这种灵活性是以牺牲解码过程为代价的,对于在特定扬声器配置上的HOA表示的回放,需要解码过程。与需要的扬声器数量通常很大的WFS方法相比,HOA也可以被提供给只包括较少扬声器的配置。HOA的其它优点是,在没有针对对耳机的双耳呈现的任何修改的情况下,也可以采用相同的表示。

[0004] HOA是基于按照截短的球面谐波(SH)展开的、复杂谐波平面波振幅的空间密度的表示。每个展开系数是角频率的函数,所述角频率的函数可以通过时域函数来等价表示。因此,不失一般性地,实际上可以假设完整的HOA声场表示由 O 个时域函数组成,其中 O 表示展开系数的数量。在下文中,这些时域函数将会等地称为HOA系数序列。

[0005] HOA表示的空间分辨率随着展开的最大阶 N 的增长而提高。不幸地,展开系数 O 的数量随着阶 N 二次方地增长,具体地是 $O = (N+1)^2$ 。例如,典型的使用阶 $N=4$ 的HOA表示需要 $O=25$ 的HOA(展开)系数。根据上述考虑,给定期望的单声道采样速率 f_s 以及每个样本的比特数量 N_b ,针对HOA表示的传输的总比特率由 $O \cdot f_s \cdot N_b$ 确定。使用每个样本 $N_b=16$ 个比特、以样本速率 $f_s=48\text{kHz}$ 传输阶 $N=4$ 的HOA表示将会导致 19.2Mbits/s 的比特率,这对于许多实际应用(例如流传输)来说非常的高。因此,非常需要HOA表示的压缩。

发明内容

[0006] 处理HOA表示(具有 $N>1$)的压缩的现有方法是很少的。由E.Hellerud, I. Burnett, A. Solvang and U.P.Svensson, "Encoding Higher Order Ambisonics with AAC", 124th AES Convention, Amsterdam, 2008提出的最直接的方法是使用高级音频编码(AAC)执行各个HOA系数序列的直接编码,所述高级音频编码(AAC)是感知编码算法。然而,该方法固有的问题是从未听见的信号的感知编码。重建的回放信号经常通过HOA系数序列的加权和来获得,并且当解压缩的HOA表示在特定的扬声器配置上呈现时,有很大的可能会暴露感知编码噪音。针对感知编码噪音暴露的主要问题是各个HOA系数序列之间的高互相关性。由于各个HOA系数序列中的编码噪音信号经常是相互之间不相关的,因此可能会出现感知编码噪音的有益的叠加,同时无噪音HOA系数序列在叠加处消除。其它的问题是,这些互相关性导致感知编码器效率下降。

[0007] 为了使两种效果的程度降到最低,在EP 2469742 A2中提出了在感知编码之前,将

H0A表示变换为离散空间域中的等价表示。从形式上看,该离散空间域是在一些离散方向处采样的、复杂谐波平面波振幅的空间密度的时域等价物。因此离散空间域由0个传统时域信号来表示,如果扬声器恰好位于与针对空间域变换假设的方向相同的方向,则传统时域信号可以解释为从采样方向冲击的大体平面波,并且传统时域信号将会与扬声器信号相对应。

[0008] 向离散空间域的变换降低了各个空间域信号之间的互相关性,但是没有完全消除这些互相关性。相对高的互相关性的示例是方向在由空间域信号覆盖的相邻方向中间的方向的定向信号。

[0009] 两种方法的主要缺点是:感知编码信号的数量是 $(N+1)^2$,并且针对压缩的H0A表示的数据速率随着立体混响阶N二次方地增长。

[0010] 为了降低感知编码信号的数量,专利申请EP 2665208 A1提出了将H0A表示分解为给定的最大数量的主导定向信号和残余环境分量。要感知编码的信号数量的降低是通过降低残余环境分量的阶来实现的。该方法背后的原理是:在通过较低阶H0A表示使用足够精度来表示残余的同时,保持关于主导定向信号的高空间分辨率。

[0011] 只要满足关于声场的假设,该方法会很好的工作,即,假设声场由少量的主导定向信号(代表使用完整的阶N编码的大体平面波函数)和没有任何方向性的残余环境分量组成。然而,如果在分解之后残余环境分量仍然包含一些主导定向分量,则阶降低会导致在分解之后的呈现处明显可以感知到的错误。违反了假设的H0A表示的典型示例是以低于N的阶编码的大体平面波。这样的阶低于N的大体平面波可以产生于艺术创作,以便使得声源看起来更广泛,并且这样的阶低于N的大体平面波还可以随着通过球形麦克风记录H0A声场表示而出现。在两种示例中,由大量高度相关的空间域信号来表示声场(其解释还可以参见 Spatial resolution of Higher Order Ambisonics)。

[0012] 本发明要解决的问题是消除专利申请EP 2665208 A1中描述的过程引起的缺点,由此也避免了上述其它引用的现有技术的缺点。该问题是由说明书中所公开的方法解决的。说明书中公开了利用这些方法的对应的设备。

[0013] 本发明改进了专利申请EP 2665208 A1中描述的H0A声场表示压缩过程。首先,像EP 2665208 A1中描述的,针对主导声源的存在对H0A表示进行分析,估计所述主导声源的方向。利用主导声源方向的信息,将H0A表示分解为多个表示大体平面波的主导定向信号和残余分量。然而,将该残余H0A分量的阶变换到离散空间域,而不是立即降低该残余H0A分量的阶,以便获得在表示残余H0A分量的均匀采样方向处的大体平面波函数。此后,根据主导定向信号预测这些平面波函数。该操作的原因在于,残余H0A分量的一部分可能与主导定向信号高度相关。

[0014] 所述预测可以是简单的预测,从而只产生少量的辅助信息。在最简单的情况下,预测由适当的缩放和延时组成。最后,预测误差变换回H0A域,并且当作残余环境H0A分量,针对所述残余环境H0A分量执行阶降低。

[0015] 有利的是,从残余H0A分量中减去可预测的信号的效果是减小其总功率并且保持主导定向信号的数量,并且通过这种方式来减少由于阶降低导致的分解误差。

[0016] 在原则上,本发明的压缩方法适用于压缩声场的高阶立体混响(表示为H0A)表示,所述方法包括以下步骤:

- [0017] -根据H0A系数的当前时帧,估计主导声源方向;
- [0018] -基于所述H0A系数并且基于所述主导声源方向,将所述H0A表示分解为时域中的主导定向信号和残余H0A分量,其中所述残余H0A分量变换到离散空间域,以便在表示所述残余H0A分量的均匀采样方向处获得平面波函数,并且其中所述平面波函数是根据所述主导定向信号预测的,由此提供描述所述预测的参数,并且对应的预测误差变换回H0A域;
- [0019] -将所述残余H0A分量的当前阶降低到更低的阶,得到降阶残余H0A分量;
- [0020] -对所述降阶残余H0A分量进行去相关,以获得对应的残余H0A分量时域信号;
- [0021] -对所述主导定向信号和所述残余H0A分量时域信号进行感知编码,从而提供压缩的主导定向信号和压缩的残余分量信号。
- [0022] 原则上,本发明的压缩设备适用于压缩声场的高阶立体混响(表示为H0A)表示,所述设备包括:
- [0023] -适于根据H0A系数的当前时间帧来估计主导声源方向的装置;
- [0024] -适于基于所述H0A系数并且基于所述主导声源方向,将所述H0A表示分解为时域中的主导定向信号和残余H0A分量的装置,其中所述残余H0A分量变换到离散空间域,以便在表示所述残余H0A分量的均匀采样方向处获得平面波函数,并且其中所述平面波函数是根据所述主导定向信号预测的,由此提供描述所述预测的参数,并且对应的预测误差变换回H0A域;
- [0025] -适于将所述残余H0A分量的当前阶降低到更低的阶,得到降阶残余H0A分量的装置;
- [0026] -适于对所述降阶残余H0A分量进行去相关,以获得对应的残余H0A分量时域信号的装置;
- [0027] -适于对所述主导定向信号和所述残余H0A分量时域信号进行感知编码,从而提供解压缩的主导定向信号和解压缩的残余分量信号的装置;
- [0028] 原则上,本发明的解压缩方法适用于解压缩根据上述压缩方法压缩的高阶立体混响表示,所述解压缩方法包括以下步骤:
- [0029] -对所压缩的主导定向信号和所压缩的残余分量信号进行感知解码,从而提供解压缩的主导定向信号和表示空间域中的残余H0A分量的解压缩的时域信号;
- [0030] -对所述解压缩的时域信号进行重新相关,来获得对应的降阶残余H0A分量;
- [0031] -将所述降阶残余H0A分量的阶增大到原始的阶,从而提供对应的解压缩残余H0A分量;
- [0032] -使用所述解压缩主导定向信号、所述原始阶解压缩残余H0A分量、所述估计的主导声源方向和描述所述预测的所述参数来组成对应的H0A系数的解压缩且重新组成的帧。
- [0033] 在原则上,本发明的解压缩设备适于解压缩根据上述压缩方法压缩的高阶立体混响表示,所述解压缩设备包括:
- [0034] -适于对所压缩的主导定向信号和所压缩的残余分量信号进行感知解码,从而提供解压缩的主导定向信号和表示空间域中的残余H0A分量的解压缩的时域信号的装置;
- [0035] -适于对所述解压缩的时域信号进行重新相关,以获得对应的降阶残余H0A分量的装置;
- [0036] -适于将所述降阶残余H0A分量的阶增大到原始的阶,从而提供对应的解压缩的残

余HOA分量的装置；

[0037] 适于通过使用所述解压缩的主导定向信号、所述原始阶解压缩的残余HOA分量、所述估计的主导声源方向和描述所述预测的所述参数，来组成对应的HOA系数的解压缩且重新组成的帧的装置。

[0038] 在对应的从属权利要求中公开了有利的附加实施例。

附图说明

[0039] 参照附图对本发明的示例性实施例进行描述，其中：

[0040] 图1a 压缩步骤1：将HOA信号分解为多个主导定向信号、残余环境HOA分量和辅助信息；

[0041] 图1b 压缩步骤2：阶降低，针对环境HOA分量进行去相关，以及对两个分量进行感知编码；

[0042] 图2a 解压缩步骤1：对时域信号进行感知解码，对表示残余环境HOA分量的信号进行重新相关，以及阶增大；

[0043] 图2b 解压缩步骤2：总HOA表示的组成；

[0044] 图3 HOA分解

[0045] 图4 HOA组成

[0046] 图5 球形坐标系

[0047] 图6 针对不同的N值的归一化函数 $v_N(\Theta)$ 的示例性曲线

具体实施方式

[0048] 压缩过程

[0049] 根据本发明的压缩过程包括分别在图1a和图1b中示出的两个连续的步骤。各个信号的准确定义在HOA分解和重新组成的详细描述部分中描述。使用了针对长度B的HOA系数序列的非重叠输入帧 $D(k)$ 的压缩的逐帧处理，其中 k 表示帧索引。关于方程式(42)中指定的HOA系数序列，帧定义如下：

$$D(k) := [d((kB+1)T_s) \ d((kB+2)T_s) \ \cdots \ d((kB+B)T_s)] \quad (1)$$

[0051] 其中 T_s 表示采样周期。

[0052] 在图1a中，HOA系数序列的帧 $D(k)$ 输入到主导声源方向估计步骤或阶段11，所述主导声源方向估计步骤或阶段针对主导定向信号的存在来分析HOA表示，估计主导定向信号的方向。可以例如通过专利申请EP 2665208 A1中描述的处理过程来执行方向的估计。估计的方向由 $\hat{\mathbf{n}}_{\text{DOM},1}(k) \dots \hat{\mathbf{n}}_{\text{DOM},D}(k)$ 来表示，其中 D 表示方向估计的最大数量。假设估计的方向如下设置在矩阵 $\mathbf{A}_{\hat{\mathbf{n}}}(k)$ 中：

$$\mathbf{A}_{\hat{\mathbf{n}}}(k) := [\hat{\mathbf{n}}_{\text{DOM},1}(k) \ \dots \ \hat{\mathbf{n}}_{\text{DOM},D}(k)] \quad (2)$$

[0054] 隐含地假设通过将方向估计分配给来自先前的帧的方向估计，来对所述方向估计进行适当的整理。因此，假设各个方向估计的时间序列描述主导声源的方向轨迹。具体地，如果第 d 个主导声源不应当运行，则可以通过向 $\hat{\mathbf{n}}_{\text{DOM},d}(k)$ 分配非有效值来对其进行指示。

然后,在分解步骤或阶段12中,利用 $\mathbf{A}_{\hat{\mathbf{d}}}(k)$ 中估计的方向将HOA表示分解为 $\mathcal{D}$ 个最大主导定向信号 $X_{\text{DIR}}(k-1)$ 、描述根据主导定向信号预测的残余HOA分量的空间域信号的预测的一些参数 $\zeta(k-1)$ 、以及表示预测误差的环境HOA分量 $D_A(k-2)$ 。在HOA解压缩部分提供了所述解压缩的详细描述。

[0055] 在图1b中示出了定向信号 $X_{\text{DIR}}(k-1)$ 的感知编码和残余环境HOA分量 $D_A(k-2)$ 的感知编码。定向信号 $X_{\text{DIR}}(k-1)$ 是能够使用任何现有感知压缩技术来分别压缩的传统时域信号。环境HOA域分量 $D_A(k-2)$ 的压缩在两个连续的步骤或阶段中执行。在阶降低的步骤或阶段13中执行立体混响的阶 N_{RED} 的降低,其中例如 $N_{\text{RED}}=1$,得到环境HOA分量 $D_{A,\text{RED}}(k-2)$ 。通过在 $D_A(k-2)$ 中保留 N_{RED} 个HOA系数并且丢弃其它系数来实现这样的阶的降低。在解码器一侧,如下文的解释,针对省略的值,附加对应的零值。

[0056] 应当注意的是,与专利申请EP 2665208 A1中的方法相比,由于总功率以及残余环境HOA分量的方向性的残余量较小,所以降低的阶 N_{RED} 一般来说可以选择为更小的。因此与专利申请EP 2665208 A1相比,所述阶的降低会导致更小的误差。

[0057] 在后面的去相关步骤或阶段14中,对表示阶降低的环境HOA分量 $D_{A,\text{RED}}(k-2)$ 的HOA系数序列进行去相关,以获得时域信号 $W_{A,\text{RED}}(k-2)$,所述时域信号 $W_{A,\text{RED}}(k-2)$ 输入到(一组)并行的感知编码器或按照任何已知感知压缩技术操作的压缩器15。执行去相关以便在解压缩之后呈现HOA表示时,避免暴露感知编码噪声(其解释参见专利申请EP 12305860.4)。通过将 $D_{A,\text{RED}}(k-2)$ 转化为变换为空间域中 O_{RED} 个等价信号可以实现近似的去相关,所述变换通过应用专利申请EP 2469742 A2中描述的球谐变换来实现。

[0058] 备选地,可以使用专利申请EP 12305861.2中提出的自适应球谐变换,其中将采样方向的网格旋转以实现可能的最佳去相关效果。另一个备选去相关技术是专利申请EP 12305860.4中描述的Karhunen-Loève变换(KLT)。应当注意的是,针对最后两种去相关,要提供表示为 $\alpha(k-2)$ 的某种辅助信息以便能够在HOA解压缩阶段对去相关进行恢复。

[0059] 在一个实施例中,联合地执行所有时域信号 $X_{\text{DIR}}(k-1)$ 和 $D_{A,\text{RED}}(k-2)$ 的感知压缩,以便提高编码效率。

[0060] 感知编码的输出是压缩的定向信号 $\tilde{X}_{\text{DIR}}(k-1)$ 和压缩的环境时域信号 $\tilde{W}_{A,\text{RED}}(k-2)$ 。

[0061] 解压缩步骤

[0062] 图2a和图2b中示出了解压缩过程。与压缩类似,所述解压缩过程由两个连续的步骤组成。在图2a中,在感知解码或解压缩步骤或阶段21中执行对定向信号 $\tilde{X}_{\text{DIR}}(k-1)$ 和表示残余环境HOA分量的时域信号 $\tilde{W}_{A,\text{RED}}(k-2)$ 的感知解压缩。在重新相关步骤或阶段22中对得到的感知解压缩时域信号 $\tilde{W}_{A,\text{RED}}(k-2)$ 进行重新相关,以便提供阶 N_{RED} 的残余分量HOA表示 $\hat{D}_{A,\text{RED}}(k-2)$ 。任选地,重新相关可以使用传输的或存储的(取决于所使用的去相关方法)参数 $\alpha(k-2)$,以与针对步骤/阶段14描述的两种备选过程相反的方式来执行。此后,在阶增大步骤或阶段23中,通过阶增大,根据 $\hat{D}_{A,\text{RED}}(k-2)$ 估计阶 N 的适当的

H0A表示 $\hat{\mathbf{D}}_{\mathbf{A}}(k-2)$ 。阶增大通过将对应的‘零’值行附加到 $\hat{\mathbf{D}}_{\mathbf{A},\text{RED}}(k-2)$ 来实现,由此假设关于更高阶的H0A系数具有零值。

[0063] 在图2b中,在组成步骤或阶段24中,根据解压缩的主导定向信号 $\tilde{\mathbf{x}}_{\text{DIR}}(k-1)$ 连同对应的方向 $\mathbf{A}_{\hat{\Omega}}(k)$ 和预测参数 $\zeta(k-1)$,以及根据残余环境H0A分量 $\hat{\mathbf{D}}_{\mathbf{A}}(k-2)$,来重新组成总的H0A表示,得到解压缩且重新组成的H0A系数的帧 $\hat{\mathbf{D}}(k-2)$ 。

[0064] 在联合地执行所有时域信号 $\mathbf{x}_{\text{DIR}}(k-1)$ 和 $\mathbf{w}_{\mathbf{A},\text{RED}}(k-2)$ 的感知压缩以便提高编码效率的情况下,也以对应的方式联合地执行压缩的定向信号 $\tilde{\mathbf{x}}_{\text{DIR}}(k-1)$ 和压缩的时域信号 $\tilde{\mathbf{w}}_{\mathbf{A},\text{RED}}(k-2)$ 的感知解压缩。

[0065] 在H0A重新组织部分中提供对重新组织的详细描述。

[0066] H0A分解

[0067] 图3中给出了示出针对H0A分解执行的操作的框图。该操作被总结如下:首先,计算平滑的主导定向信号 $\mathbf{x}_{\text{DIR}}(k-1)$,并且将其输出用于感知压缩。接着,由0个定向信号 $\tilde{\mathbf{x}}_{\text{GRID,DIR}}(k-1)$ 来表示主导定向信号的H0A表示 $\mathbf{D}_{\text{DIR}}(k-1)$ 与原始H0A表示 $\mathbf{D}(k-1)$ 之间的残余,其中所述0个定向信号可以被认为是均匀分布的方向上的大体平面波。根据主导定向信号 $\mathbf{x}_{\text{DIR}}(k-1)$ 对这些定向信号进行预测,输出了预测参数 $\zeta(k-1)$ 。最后,计算并输出原始H0A表示 $\mathbf{D}(k-2)$ 与主导定向信号的H0A表示 $\mathbf{D}_{\text{DIR}}(k-1)$ 之间的残余 $\mathbf{D}_{\mathbf{A}}(k-2)$ 以及均匀分布的方向上的预测的定向信号的H0A表示 $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2)$ 。

[0068] 在描述细节之前,需要指出的是,在组成期间,连续帧之间的方向变化可以导致所有计算的信号中断。因此,首先计算针对重叠帧的相应信号的瞬时估计,所述瞬时估计的长度为2B。第二,使用适当的窗口函数使连续的重叠帧的结果平滑。然而,每次平滑引入了单个帧的迟滞。

[0069] 计算瞬时主导定向信号

[0070] 步骤或阶段30中针对H0A系数序列的当前帧 $\mathbf{D}(k)$ 根据 $\mathbf{A}_{\hat{\Omega}}(k)$ 中的估计的声源方向计算瞬时主导方向信号的计算是基于以下文献中描述的模式匹配:M.A.Poletti, "Three-Dimensional Surround Sound Systems Based on Spherical Harmonics", J.Audio Eng.Soc, 53 (11), pages 1004-1025, 2005。具体的,对H0A表示得到给定H0A信号的最佳近似的定向信号进行搜索。

[0071] 此外,不失一般性地,假设一向量可以唯一地指定有效主导声源的每个方向估计 $\hat{\Omega}_{\text{DOM,d}}(k)$,所述向量包含依据以下公式的倾角 $\theta_{\text{DOM,d}}(k) \in [0, \pi]$ 和方位角 $\phi_{\text{DOM,d}}(k) \in [0, 2\pi]$ (其示意参见图5):

$$[0072] \quad \hat{\Omega}_{\text{DOM,d}}(k) := \left(\hat{\theta}_{\text{DOM,d}}(k), \hat{\phi}_{\text{DOM,d}}(k) \right)^T \quad (3)$$

[0073] 首先,根据

$$[0074] \quad \mathbf{E}_{\text{ACT}}(k) := [\mathbf{s}_{\text{DOM,dACT,1}}(k) \quad \mathbf{s}_{\text{DOM,dACT,2}}(k) \quad \dots \quad \mathbf{s}_{\text{DOM,dACT,D}_{\text{ACT}}(k)}(k)] \in \mathbb{R}^{O \times D_{\text{ACT}}(k)} \quad (4)$$

[0075] 对基于有效声源的方向估计的模式矩阵进行计算,其

$$[0076] \quad \mathbf{S}_{\text{DOM},d}(k) := \left[S_0^0(\hat{\boldsymbol{\Omega}}_{\text{DOM},d}(k)), S_1^{-1}(\hat{\boldsymbol{\Omega}}_{\text{DOM},d}(k)), S_1^0(\hat{\boldsymbol{\Omega}}_{\text{DOM},d}(k)), \dots, S_N^N(\hat{\boldsymbol{\Omega}}_{\text{DOM},d}(k)) \right]^T \in \mathbb{R}^O \quad (5)$$

[0077] 在方程式(4)中, $D_{\text{ACT}}(k)$ 表示针对第k个帧的有效方向的数量,并且 $d_{\text{ACT},j}(k)$ ($1 \leq j \leq D_{\text{ACT}}(k)$) 指示它们的索引。 $S_n^m(\cdot)$ 表示实值球谐函数,所述实值球谐函数在实值球谐函数的定义部分中定义。

[0078] 第二,计算定义如下的包含第(k-1)个帧和第k个帧的所有主导定向信号的瞬时估计的矩阵 $\tilde{\mathbf{X}}_{\text{DIR}}(k) \in \mathbb{R}^{D \times 2B}$

$$[0079] \quad \tilde{\mathbf{X}}_{\text{DIR}}(k) := [\tilde{\mathbf{x}}_{\text{DIR}}(k, 1) \quad \tilde{\mathbf{x}}_{\text{DIR}}(k, 2) \quad \dots \quad \tilde{\mathbf{x}}_{\text{DIR}}(k, 2B)] \quad (6)$$

[0080] 其中

$$[0081] \quad \tilde{\mathbf{x}}_{\text{DIR}}(k, l) := [\tilde{x}_{\text{DIR},1}(k, l), \tilde{x}_{\text{DIR},2}(k, l), \dots, \tilde{x}_{\text{DIR},D}(k, l)]^T \in \mathbb{R}^D, 1 \leq l \leq 2B \quad (7)$$

[0082] 这通过两个步骤来实现。在第一个步骤中,将对应于无效方向的行中的定向信号样本设置为零,即

$$[0083] \quad \tilde{x}_{\text{DIR},d}(k, l) = 0 \quad \forall 1 \leq l \leq 2B, \quad \text{if } d \notin \mathcal{M}_{\text{ACT}}(k), \quad (8)$$

[0084] 其中 $\mathcal{M}_{\text{ACT}}(k)$ 指示有效方向的集。在第二个步骤中,通过首先将对应于有效方向的定向信号样本排列在根据以下公式的矩阵中,来获得对应于有效方向的定向信号样本:

$$[0085] \quad \tilde{\mathbf{X}}_{\text{DIR,ACT}}(k) := \begin{bmatrix} \tilde{x}_{\text{DIR},d_{\text{ACT},1}(k)}(k, 1) & \dots & \tilde{x}_{\text{DIR},d_{\text{ACT},1}(k)}(k, 2B) \\ \vdots & \ddots & \vdots \\ \tilde{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k)}(k)}(k, 1) & \dots & \tilde{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k)}(k)}(k, 2B) \end{bmatrix} \quad (9)$$

[0086] 然后对该矩阵进行计算,以使误差的欧几里德范数

$$[0087] \quad \boldsymbol{\varepsilon}_{\text{ACT}}(k) \tilde{\mathbf{X}}_{\text{DIR,ACT}}(k) - [\mathbf{D}(k-1) \quad \mathbf{D}(k)] \quad (10)$$

[0088] 最小化。解是由以下方程式给出的:

$$[0089] \quad \tilde{\mathbf{X}}_{\text{DIR,ACT}}(k) = [\boldsymbol{\varepsilon}_{\text{ACT}}^T(k) \boldsymbol{\varepsilon}_{\text{ACT}}(k)]^{-1} \boldsymbol{\varepsilon}_{\text{ACT}}^T(k) [\mathbf{D}(k-1) \quad \mathbf{D}(k)] \quad (11)$$

[0090] 时间平滑

[0091] 针对步骤或阶段31,只针对定向信号 $\tilde{\mathbf{X}}_{\text{DIR}}(k)$ 解释了平滑,因为其它类型的信号的平滑能够以完全相似的方式来完成。通过以下的适当窗函数来对样本被包含在根据方程式(6)的矩阵 $\tilde{\mathbf{X}}_{\text{DIR}}(k)$ 中的定向信号估计 $\tilde{x}_{\text{DIR},d}(k, l)$ ($1 \leq d \leq D$) 进行加窗:

$$[0092] \quad \tilde{x}_{\text{DIR,WIN},d}(k, l) := \tilde{x}_{\text{DIR},d}(k, l) \cdot w(l), \quad 1 \leq l \leq 2B \quad (12)。$$

[0093] 该窗函数必须满足这样的条件:它与其在以下重叠区域中的偏移版本(假设B样本的偏移)之和为‘1’:

$$[0094] \quad w(l) + w(B + l) = 1 \quad \forall 1 \leq l \leq B \quad (13)。$$

[0095] 由以下方程式定义的周期性Hann窗给出了针对这样的窗函数的示例:

$$[0096] \quad w(l) := 0.5 \left[1 - \cos\left(\frac{2\pi(l-1)}{2B}\right) \right] \quad \text{for } 1 \leq l \leq 2B \quad (14)。$$

[0097] 通过根据以下方程式的加窗的瞬时估计的适当叠加来对第(k-1)个帧的平滑的定向信号进行计算:

$$[0098] \quad x_{\text{DIR},d}((k-1)B+l) = \tilde{x}_{\text{DIR},\text{WIN},d}(k-1, B+l) + \tilde{x}_{\text{DIR},\text{WIN},d}(k, l) \quad (15)。$$

[0099] 针对第(k-1)个帧的所有平滑的定向信号的样本被设置在以下的矩阵中:

$$[0100] \quad \mathbf{X}_{\text{DIR}}(k-1) := [\mathbf{x}_{\text{DIR}}((k-1)B+1) \quad \mathbf{x}_{\text{DIR}}((k-1)B+2) \quad \dots \quad \mathbf{x}_{\text{DIR}}((k-1)B+B)] \in \mathbb{R}^{D \times B} \quad (16)$$

[0101] 其中

$$[0102] \quad \mathbf{x}_{\text{DIR}}(l) = [x_{\text{DIR},1}(l), x_{\text{DIR},2}(l), \dots, x_{\text{DIR},D}(l)]^T \in \mathbb{R}^D \quad (17)。$$

[0103] 平滑的主导定向信号 $\mathbf{X}_{\text{DIR},d}(1)$ 应当是连续地输入到感知编码器的连续信号。

[0104] 计算平滑的主导定向信号的HOA表示

[0105] 在步骤或阶段32中,基于连续信号 $\mathbf{X}_{\text{DIR},d}(1)$,根据 $\mathbf{X}_{\text{DIR}}(k-1)$ 和 $\mathbf{A}_{\hat{\mathbf{n}}}(k)$,对平滑的主导定向信号的HOA表示进行计算,以便对将要针对HOA组成所执行的操作相同的操作进行模仿。由于连续帧之间的方向估计的变化会导致中断,再一次对长度为2B的重叠帧的瞬时HOA表示进行计算,并且通过使用适当的窗函数对连续重叠帧的结果进行平滑。因此,通过以下方程式来获得HOA表示 $\mathbf{D}_{\text{DIR}}(k-1)$:

$$[0106] \quad \mathbf{D}_{\text{DIR}}(k-1) = \Xi_{\text{ACT}}(k) \mathbf{X}_{\text{DIR},\text{ACT},\text{WIN1}}(k-1) + \Xi_{\text{ACT}}(k-1) \mathbf{X}_{\text{DIR},\text{ACT},\text{WIN2}}(k-1) \quad (18),$$

[0107] 其中, $\mathbf{X}_{\text{DIR},\text{ACT},\text{WIN1}}(k-1) :=$

$$[0108] \quad \begin{bmatrix} x_{\text{DIR},d_{\text{ACT},1}(k)}((k-1)B+1) \cdot w(1) & \dots & x_{\text{DIR},d_{\text{ACT},1}(k)}(kB) \cdot w(B) \\ x_{\text{DIR},d_{\text{ACT},2}(k)}((k-1)B+1) \cdot w(1) & & x_{\text{DIR},d_{\text{ACT},2}(k)}(kB) \cdot w(B) \\ \vdots & \ddots & \vdots \\ x_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k)}}((k-1)B+1) \cdot w(1) & \dots & x_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k)}}(kB) \cdot w(B) \end{bmatrix} \quad (19)$$

[0109] 并且 $\mathbf{X}_{\text{DIR},\text{ACT},\text{WIN2}}(k-1) :=$

$$[0110] \quad \begin{bmatrix} x_{\text{DIR},d_{\text{ACT},1}(k-1)}((k-1)B+1) \cdot w(B+1) & \dots & x_{\text{DIR},d_{\text{ACT},1}(k-1)}(kB) \cdot w(2B) \\ x_{\text{DIR},d_{\text{ACT},2}(k-1)}((k-1)B+1) \cdot w(B+1) & & x_{\text{DIR},d_{\text{ACT},2}(k-1)}(kB) \cdot w(2B) \\ \vdots & \ddots & \vdots \\ x_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k-1)}}((k-1)B+1) \cdot w(B+1) & \dots & x_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}(k-1)}}(kB) \cdot w(2B) \end{bmatrix} \quad (20)。$$

[0111] 通过均匀网格上的定向信号来表示残余HOA表示

[0112] 在步骤或阶段33中,根据 $\mathbf{D}_{\text{DIR}}(k-1)$ 和 $\mathbf{D}(k-1)$ (即通过帧延时381延时的 $\mathbf{D}(k)^{D(k)}$),对由均匀网格上的定向信号表示的残余HOA表示进行计算。该操作的目的是:获得从一些固定的、几乎均匀分布的方向 $\hat{\mathbf{n}}_{\text{GRID},o}$ ($1 \leq o \leq O$,也被称为网格方向)冲击的定向信号(即大体平面波函数),以表示残余 $[\mathbf{D}(k-2) \quad \mathbf{D}(k-1)] - [\mathbf{D}_{\text{DIR}}(k-2) \quad \mathbf{D}_{\text{DIR}}(k-1)]$

[0113] 首先,关于网格方向,如下计算模式矩阵 Ξ_{GRID} :

$$[0114] \quad \Xi_{\text{GRID}} := [\mathbf{S}_{\text{GRID},1} \quad \mathbf{S}_{\text{GRID},2} \quad \dots \quad \mathbf{S}_{\text{GRID},O}] \in \mathbb{R}^{O \times O} \quad (21)$$

[0115] 其中

$$[0116] \quad \mathbf{S}_{\text{GRID},o} := [S_0^o(\hat{\mathbf{n}}_{\text{GRID},o}), S_1^{-1}(\hat{\mathbf{n}}_{\text{GRID},o}), S_1^0(\hat{\mathbf{n}}_{\text{GRID},o}), \dots, S_N^N(\hat{\mathbf{n}}_{\text{GRID},o})]^T \in \mathbb{R}^O \quad (22)。$$

[0117] 由于在整个压缩过程期间网格方向是固定的,所以模式矩阵 Ξ_{GRID} 只需要计算一

次。

[0118] 如下获得在对应的网格上的定向信号：

$$[0119] \quad \tilde{\mathbf{X}}_{\text{GRID,DIR}}(k-1) = \mathbf{E}_{\text{GRID}}^{-1}([\mathbf{D}(k-2) \quad \mathbf{D}(k-1)] - [\mathbf{D}_{\text{DIR}}(k-2) \quad \mathbf{D}_{\text{DIR}}(k-1)]) \quad (23)。$$

[0120] 根据主导定向信号预测均匀网格上的定向信号

[0121] 在步骤或阶段34中,根据 $\tilde{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$ 和 $\mathbf{X}_{\text{DIR}}(k-1)$, 对均匀网格上的定向信号进行预测。根据定向信号的在网格方向 $\hat{\mathbf{D}}_{\text{GRID},o}$ ($1 \leq o \leq O$) 组成的均匀网格上的定向信号的预测是基于针对平滑目的的两个连续帧,即网格信号 $\tilde{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$ (长度为2B) 的展开的帧是根据平滑的主导定向信号的展开的帧:

$$[0122] \quad \tilde{\mathbf{X}}_{\text{DIR,EXT}}(k-1) := [\mathbf{X}_{\text{DIR}}(k-3) \quad \mathbf{X}_{\text{DIR}}(k-2) \quad \mathbf{X}_{\text{DIR}}(k-1)] \quad (24)$$

[0123] 预测的。

[0124] 首先,包含在 $\tilde{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$ 中的每个网格信号 $\tilde{\mathbf{x}}_{\text{GRID,DIR},o}(k-1,l)$ ($1 \leq o \leq O$) 分配到包含在 $\tilde{\mathbf{X}}_{\text{DIR,EXT}}(k-1)$ 中的主导定向信号 $\tilde{\mathbf{x}}_{\text{DIR,EXT},d}(k-1,l)$ ($1 \leq d \leq D$) 中。所述分配可以基于网格信号与所有的主导定向信号之间的归一化互相关函数的计算。特别是,该主导定向信号被分配到网格信号,这提供归一化互相关函数的最高值。分配的结果可以由将第o个网格信号分配到第 $f_{\mathcal{A},k-1}(o)$ 个主导定向信号的分配函数 $f_{\mathcal{A},k-1}: \{1, \dots, O\} \rightarrow \{1, \dots, D\}$ 来表示。

[0125] 第二,通过分配的主导定向信号 $\tilde{\mathbf{x}}_{\text{DIR,EXT},f_{\mathcal{A},k-1}(o)}(k-1,l)$ 来预测每个网格信号 $\tilde{\mathbf{x}}_{\text{GRID,DIR},o}(k-1,l)$ 。根据分配的主导定向信号 $\tilde{\mathbf{x}}_{\text{DIR,EXT},f_{\mathcal{A},k-1}(o)}(k-1,l)$, 通过延时和缩放,如下对预测的网格信号 $\hat{\mathbf{x}}_{\text{GRID,DIR},o}(k-1,l)$ 进行计算:

$$[0126] \quad \hat{\mathbf{x}}_{\text{GRID,DIR},o}(k-1,l) = K_o(k-1) \cdot \tilde{\mathbf{x}}_{\text{DIR,EXT},f_{\mathcal{A},k-1}(o)}(k-1,l - \Delta_o(k-1)) \quad (25),$$

[0127] 其中, $K_o(k-1)$ 表示缩放因子并且 $\Delta_o(k-1)$ 指示样本延时。选择这些参数来使预测误差最小化。

[0128] 如果预测误差的功率大于网格信号本身的功率,则假设预测已经失败。然后,对应的预测参数可以设置为任何非有效值。

[0129] 应当注意的是,其它类型的预测也是可以的。例如,替代计算全频带缩放因子,针对感知取向频带来确定缩放因子也是可以的。然而,该操作改进预测是以辅助信息量增加为代价的。

[0130] 所有的预测参数可以如下方程式设置在参数矩阵中:

$$[0131] \quad \boldsymbol{\zeta}(k-1) := \begin{bmatrix} f_{\mathcal{A},k-1}(1) & K_1(k-1) & \Delta_1(k-1) \\ f_{\mathcal{A},k-1}(2) & K_2(k-1) & \Delta_2(k-1) \\ \vdots & \vdots & \vdots \\ f_{\mathcal{A},k-1}(O) & K_O(k-1) & \Delta_O(k-1) \end{bmatrix} \quad (26)。$$

[0132] 假设所有的预测信号 $\hat{\mathbf{x}}_{\text{GRID,DIR},o}(k-1,l)$ ($1 \leq o \leq O$) 设置在矩阵 $\hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$

中。

[0133] 计算预测的均匀网格上的定向信号的HOA表示

[0134] 在步骤或阶段35中,根据以下公式,根据 $\hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$ 计算预测的网格信号的HOA表示:

$$[0135] \quad \hat{\mathbf{D}}_{\text{GRID,DIR}}(k-1) = \mathbf{E}_{\text{GRID}} \hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1) \quad (27)。$$

[0136] 计算残余环境声场分量的HOA表示

[0137] 在步骤或阶段37中,通过公式:

$$[0138] \quad \mathbf{D}_A(k-2) = \mathbf{D}(k-2) - \hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2) - \mathbf{D}_{\text{DIR}}(k-2) \quad (28),$$

[0139] 根据 $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k)$ 的时间平滑版本(在步骤/阶段36中) $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2)$ 、根据 $\mathbf{D}(k)$ 的二帧延时版本(延时381和383) $\mathbf{D}(k-2)$ 、和 $\mathbf{D}_{\text{DIR}}(k-1)$ 的帧延时版本(延时382) $\mathbf{D}_{\text{DIR}}(k-2)$,对残余环境声场分量的HOA表示进行计算。

[0140] HOA表示

[0141] 在对图4中的各个步骤或阶段的过程进行详细描述之前,提供摘要。使用预测参数 $\hat{\mathbf{z}}(k-1)$,根据解码的主导定向信号 $\hat{\mathbf{X}}_{\text{DIR}}(k-1)$ 预测关于均匀分布的方向的定向信号 $\hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1)$ 。接着,总的HOA表示 $\hat{\mathbf{D}}(k-2)$ 由主导定向信号的HOA表示 $\hat{\mathbf{D}}_{\text{DIR}}(k-2)$ 、预测的定向信号的HOA表示 $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2)$ 和残余环境HOA分量 $\hat{\mathbf{D}}_A(k-2)$ 组成。

[0142] 计算主导定向信号的HOA表示

[0143] 将 $\mathbf{A}_d(k)$ 和 $\hat{\mathbf{X}}_{\text{DIR}}(k-1)$ 输入到步骤或阶段41中,用来确定主导定向信号的HOA表示。在已经根据方向估计 $\mathbf{A}_d(k)$ 和 $\mathbf{A}_d(k-1)$ 计算了模式矩阵 $\mathbf{E}_{\text{ACT}}(k)$ 和 $\mathbf{E}_{\text{ACT}}(k-1)$ 之后,基于第 k 个和第 $(k-1)$ 个帧的有效声场的方向估计,通过以下方程式来获得主导定向信号的HOA表示:

$$[0144] \quad \hat{\mathbf{D}}_{\text{DIR}}(k-1) = \mathbf{E}_{\text{ACT}}(k) \mathbf{X}_{\text{DIR,ACT,WIN1}}(k-1) + \mathbf{E}_{\text{ACT}}(k-1) \mathbf{X}_{\text{DIR,ACT,WIN2}}(k-1) \quad (29),$$

[0145] 其中, $\mathbf{X}_{\text{DIR,ACT,WIN1}}(k-1) :=$

$$[0146] \quad \begin{bmatrix} \hat{x}_{\text{DIR},d_{\text{ACT},1}(k)}((k-1)B+1) \cdot w(1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},1}(k)}(kB) \cdot w(B) \\ \hat{x}_{\text{DIR},d_{\text{ACT},2}(k)}((k-1)B+1) \cdot w(1) & & \hat{x}_{\text{DIR},d_{\text{ACT},2}(k)}(kB) \cdot w(B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}}(k)}((k-1)B+1) \cdot w(1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}}(k)}(kB) \cdot w(B) \end{bmatrix} \quad (30),$$

[0147] 并且

$$[0148] \quad \mathbf{X}_{\text{DIR,ACT,WIN2}}(k-1) := \begin{bmatrix} \hat{x}_{\text{DIR},d_{\text{ACT},1}(k-1)}((k-1)B+1) \cdot w(B+1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},1}(k-1)}(kB) \cdot w(2B) \\ \hat{x}_{\text{DIR},d_{\text{ACT},2}(k-1)}((k-1)B+1) \cdot w(B+1) & & \hat{x}_{\text{DIR},d_{\text{ACT},2}(k-1)}(kB) \cdot w(2B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}}(k-1)}((k-1)B+1) \cdot w(B+1) & \dots & \hat{x}_{\text{DIR},d_{\text{ACT},D_{\text{ACT}}}(k-1)}(kB) \cdot w(2B) \end{bmatrix} \quad (31)。$$

[0149] 根据主导定向信号预测均匀网格上的定向信号

[0150] 将 $\hat{\mathbf{z}}(k-1)$ 和 $\hat{\mathbf{X}}_{\text{DIR}}(k-1)$ 输入到步骤或阶段43中,用来根据主导定向信号预测均匀网格上的定向信号。预测的均匀网格上的定向信号的展开的帧由根据以下方程式的单元 $\hat{x}_{\text{GRID,DIR,o}}(k-1,l)$ 组成:

$$[0151] \quad \hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1) = \begin{bmatrix} \hat{x}_{\text{GRID,DIR},1}(k-1,1) & \dots & \hat{x}_{\text{GRID,DIR},1}(k-1,2B) \\ \hat{x}_{\text{GRID,DIR},2}(k-1,1) & & \hat{x}_{\text{GRID,DIR},2}(k-1,2B) \\ \vdots & \ddots & \vdots \\ \hat{x}_{\text{GRID,DIR},O}(k-1,1) & \dots & \hat{x}_{\text{GRID,DIR},O}(k-1,2B) \end{bmatrix} \quad (32),$$

[0152] 所述单元 $\hat{x}_{\text{GRID,DIR},o}(k-1,l)$ 是通过以下方程式根据主导定向信号预测的:

$$[0153] \quad \hat{x}_{\text{GRID,DIR},o}(k-1,l) = K_o(k-1) \cdot \hat{x}_{\text{DIR},f_{A,k-1}(o)}((k-1)B+l-\mathbf{A}_o(k-1)) \quad (33)。$$

[0154] 计算预测的均匀网格上的定向信号的HOA表示

[0155] 在计算均匀网格上的预测的定向信号的HOA表示的步骤或阶段44中,通过方程式

$$\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-1) = \mathbf{\Xi}_{\text{GRID}} \hat{\mathbf{X}}_{\text{GRID,DIR}}(k-1) \quad (34)$$

来获得预测的网格定向信号的HOA表示,其中 $\mathbf{\Xi}_{\text{GRID}}$ 表示关于预定义网格方向的模式矩阵(关于定义,参见方程式(21))。

[0156] 组成HOA声场表示

[0157] 在步骤或阶段46中,如以下方程式,根据 $\hat{\mathbf{D}}_{\text{DIR}}(k-2)$ (即由帧延时42延时的 $\hat{\mathbf{D}}_{\text{DIR}}(k-1)$)、(是步骤/阶段45中 $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-1)$ 的时间平滑版本的) $\hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2)$ 和 $\hat{\mathbf{D}}_A(k-2)$ 来最终组成总的HOA生成表示:

$$[0158] \quad \hat{\mathbf{D}}(k-2) = \hat{\mathbf{D}}_{\text{DIR}}(k-2) + \hat{\mathbf{D}}_{\text{GRID,DIR}}(k-2) + \hat{\mathbf{D}}_A(k-2) \quad (35)$$

[0159] 高阶立体混响的基本原理

[0160] 高阶立体混响是基于感兴趣的紧凑区域中的声场的描述,假设所述紧凑区域中没有声源。在这种情况下,在该感兴趣的区域中,在时间 t 和位置 x 的声压 $p(t, x)$ 的时-空特性物理上完全由均匀波方程来确定。下述内容基于图5中示出的球形坐标系。 x 轴指向正面位置, y 轴指向左方,并且 z 轴指向上方。通过半径 $r > 0$ (即到坐标原点的距离)、从极轴 z 测量的倾角 $\theta \in [0, \pi]$ 和在 x - y 平面中从 x 轴逆时针方向测量的方位角 $\phi \in [0, \pi]$ 来表示空间中的位置 $\mathbf{x} = (r, \theta, \phi)^T$ 。(·)^T 表示转置。

[0161] 可以看出(参见E.G.Williams, "Fourier Acoustics", volume 93 of Applied Mathematical Sciences, Academic Press, 1999), 声压关于时间的傅里叶变换(由 $\mathcal{F}_t(\cdot)$ 表示), 即

$$[0162] \quad P(\omega, x) = \mathcal{F}_t(p(t, x)) = \int_{-\infty}^{\infty} p(t, x) e^{-i\omega t} dt \quad (36),$$

[0163] (其中 ω 表示角频率, i 表示虚数单位) 可以如下展开成一系列球形函数

$$[0164] \quad P(\omega = kc_s, r, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n A_n^m(k) j_n(kr) S_n^m(\theta, \phi) \quad (37),$$

[0165] 其中 c_s 表示声音的速度, 并且 k 表示角波数, 所述角波数 k 通过公式 $k = \frac{\omega}{c_s}$ 与 ω 相关, $j_n(\cdot)$ 表示第一类球形贝赛尔函数, 并且 $S_n^m(\theta, \phi)$ 表示阶为 n 、角度为 m (在实值球谐函数部分定义了) 的实值球谐函数。展开系数 $A_n^m(k)$ 只取决于角波数 k 。需要注意的是, 这里已经隐性地假设声压是空间频带受限的。因此, 该系列关于阶索引 n 在上限 N 处是截断的, 所述上限 N 被称为HOA表示的阶。

[0166] 如果声场由不同角频率 ω 的谐波平面波的无穷大量的叠加来表示,并且声场可以由由角度元组 (θ, ϕ) 指定的所有可能的方向到达,则可以看出(参见B.Rafaely, "Plane-wave Decomposition of the Sound Field on a Sphere by Spherical Convolution", J.Acoust.Soc.Am., 4(116), pages 2149-2157, 2004),对应的平面波复振幅函数可以由以下球谐函数展开来表示:

$$[0167] \quad D(\omega = kc_s, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n D_n^m(k) S_n^m(\theta, \phi) \quad (38),$$

[0168] 其中展开系数 $D_n^m(k)$ 通过以下方程式与展开系数 $A_n^m(k)$ 相关:

$$[0169] \quad A_n^m(k) = 4\pi i^n D_n^m(k). \quad (39)$$

[0170] 假设各个系数 $D_n^m(k = \omega/c_s)$ 是角频率 ω 的函数,傅里叶逆变换(由 $\mathcal{F}_t^{-1}(\cdot)$ 表示)的应用给每个阶 n 和角度 m 提供了如下的时域函数:

$$[0171] \quad d_n^m(t) = \mathcal{F}_t^{-1} \left(D_n^m \left(\frac{\omega}{c_s} \right) \right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} D_n^m \left(\frac{\omega}{c_s} \right) e^{i\omega t} d\omega \quad (40)$$

[0172] 所述函数可以收集在如下的单个矢量中:

$$[0173] \quad \mathbf{d}(t) = \begin{bmatrix} d_0^0(t) & d_1^{-1}(t) & d_1^0(t) & d_1^1(t) & d_2^{-2}(t) & d_2^{-1}(t) & d_2^0(t) & d_2^1(t) & d_2^2(t) & \dots \\ d_N^{N-1}(t) & d_N^N(t) & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}^T \quad (41)$$

[0174] 由 $n(n+1)+1+m$ 来给出矢量 $\mathbf{d}(t)$ 中的时域函数 $d_n^m(t)$ 的位置索引。

[0175] 最终的立体混响格式提供使用采样频率 f_s 的 $\mathbf{d}(t)$ 的采样的版本如下:

$$[0176] \quad \{\mathbf{d}(lT_s)\}_{l \in \mathbb{N}} = \{\mathbf{d}(T_s), \mathbf{d}(2T_s), \mathbf{d}(3T_s), \mathbf{d}(4T_s), \dots\} \quad (42),$$

[0177] 其中 $T_s = 1/f_s$ 表示采样周期。 $\mathbf{d}(lT_s)$ 单元被称为立体混响系数。需要注意的是,时域信号 $d_n^m(t)$ 以及因此立体混响系数是实值的。

[0178] 实值的球谐函数的定义

[0179] 实值的球谐函数 $S_n^m(\theta, \phi)$ 由以下方程式给出:

$$[0180] \quad S_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}} P_{n,|m|}(\cos \theta) \text{trg}_m(\phi) \quad (43),$$

$$[0181] \quad \text{其中 } \text{trg}_m(\phi) = \begin{cases} \sqrt{2}\cos(m\phi) & m > 0 \\ 1 & m = 0 \\ -\sqrt{2}\sin(m\phi) & m < 0 \end{cases} \quad (44)。$$

[0182] 使用勒让德多项式 $P_n(x)$, 并且并不像上文提到的E.G.Williams textbook,在不使用Condon-Shortley项的情况下,如以下方程式定义关联的Legendre函数 $P_{n,m}(X)$:

$$[0183] \quad P_{n,m}(x) = (1-x^2)^{m/2} \frac{d^m}{dx^m} P_n(x), m \geq 0 \quad (45)$$

[0184] 高阶立体混响的空间分辨率

[0185] 从方向 $\Omega_0 = (\theta_0, \phi_0)^T$ 到达的平面波函数 $x(t)$ 在HOA中由以下方程式来表示:

$$[0186] \quad d_n^m(t) = x(t) S_n^m(\Omega_0), \quad 0 \leq n \leq N, |m| \leq n \quad (46)。$$

[0187] 平面波振幅 $\mathbf{d}(t, \Omega) := \mathcal{F}_t^{-1}(D(\omega, \Omega))$ 的对应的空间密度由以下公式给出:

$$[0188] \quad d(t, \Omega) = \sum_{n=0}^N \sum_{m=-n}^n d_n^m(t) S_n^m(\Omega) \quad (47)$$

$$[0189] \quad = x(t) \underbrace{\left[\sum_{n=0}^N \sum_{m=-n}^n S_n^m(\Omega_0) S_n^m(\Omega) \right]}_{v_N(\Theta)} \quad (48)$$

[0190] 可以从方程式(48)中看出,它是大体平面波函数 $x(t)$ 和空间分散函数 $v_N(\Theta)$ 的乘积,空间分散函数 $v_N(\Theta)$ 可以被视为仅取决于 Ω 和 Ω_0 之间的、具有如下特性的角度 Θ :

$$[0191] \quad \cos \Theta = \cos \theta \cos \theta_0 + \cos(\phi - \phi_0) \sin \theta \sin \theta_0 \quad (49)。$$

[0192] 如预期的,在无限的阶的限制下,即 $N \rightarrow \infty$,空间分散函数转换为狄拉克delta函数 $\delta(\cdot)$,即

$$[0193] \quad \lim_{N \rightarrow \infty} v_N(\Theta) = \frac{\delta(\Theta)}{2\pi} \quad (50)。$$

[0194] 然而,在有限阶 N 的情况下,来自方向 Ω_0 的大体平面波的贡献涂到相邻方向,模糊程度随着阶的提高而减少的。图6中示出了针对不同的 N 值的归一化函数 $v_N(\Theta)$ 的曲线。应当指出的是,任何平面波振幅的空间密度的时域特性的方向 Ω 是它在其它任何方向上的特性的倍数。特别是,针对一些固定方向 Ω_1 和 Ω_2 ,函数 $d(t, \Omega_1)$ 和 $d(t, \Omega_2)$ 关于时间 t 相互高度关联。

[0195] 离散空间域

[0196] 如果平面波振幅的空间密度在数量为 O 的、在单位球面上几乎均匀分布的空间方向 $\Omega_o (1 \leq o \leq O)$ 上是离散的,则获得 O 个定向信号 $d(t, \Omega_o)$ 。将这些信号集合到如以下方程式的矢量中:

$$[0197] \quad d_{\text{SPAT}}(t) := [d(t, \Omega_1) \dots d(t, \Omega_O)]^T \quad (51)$$

[0198] 通过使用方程式(47)可以证明,可以通过单一矩阵乘法,根据方程式(41)中限定的连续的立体混响表示 $d(t)$ 来计算该矢量,所述单一矩阵乘法的方程式为:

$$[0199] \quad d_{\text{SPAT}}(t) = \Psi^H d(t), \quad (52)$$

[0200] 其中 $(\cdot)^H$ 指示联合置换和共轭,并且 Ψ 表示由以下方程式限定的模式矩阵:

$$[0201] \quad \Psi := [S_1 \dots S_O] \quad (53),$$

$$[0202] \quad \text{其中 } S_o := [S_o^0(\Omega_o) \ S_o^{-1}(\Omega_o) \ S_o^0(\Omega_o) \ S_o^1(\Omega_o) \ \dots \ S_o^{N-1}(\Omega_o) \ S_o^N(\Omega_o)] \quad (54)。$$

[0203] 由于方向 Ω_o 在单位球面上是几乎均匀分布的,所以模式矩阵一般来说是可逆的。因此,通过方程式

$$[0204] \quad d(t) = \Psi^{-H} d_{\text{SPAT}}(t) \quad (55)$$

[0205] 根据定向信号 $d(t, \Omega_o)$ 可以计算连续的立体混响表示。两个方程式构在立体混响表示与空间域之间的变换和逆变换。在该应用中,这些变换被称为球谐变换和球谐逆变换。

$$[0206] \quad \text{因为在单位球面上方向 } \Omega_o \text{ 是几乎均匀分布的, } \Psi^H \approx \Psi^{-1} \quad (56)$$

[0207] 这证明了在方程式(52)中使用 Ψ^{-1} 而不使用 Ψ^H 是可行的。有利地,上述所有的关系对于离散时域也是有效的。

[0208] 在编码侧以及解码侧,本发明的过程可以通过单一处理器或电路来执行,或者通过若干个处理器或电路并行操作和/或在本发明过程的不同部分中操作。

[0209] 本发明能够用于处理可以在家庭环境中的扬声器设备或电影院中的扬声器设备

上呈现或播放的对应的声音信号。

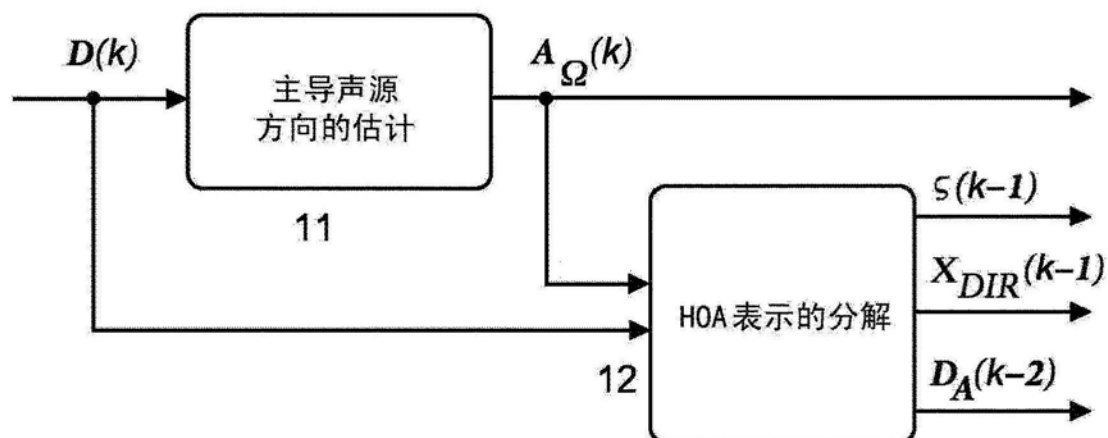


图1a

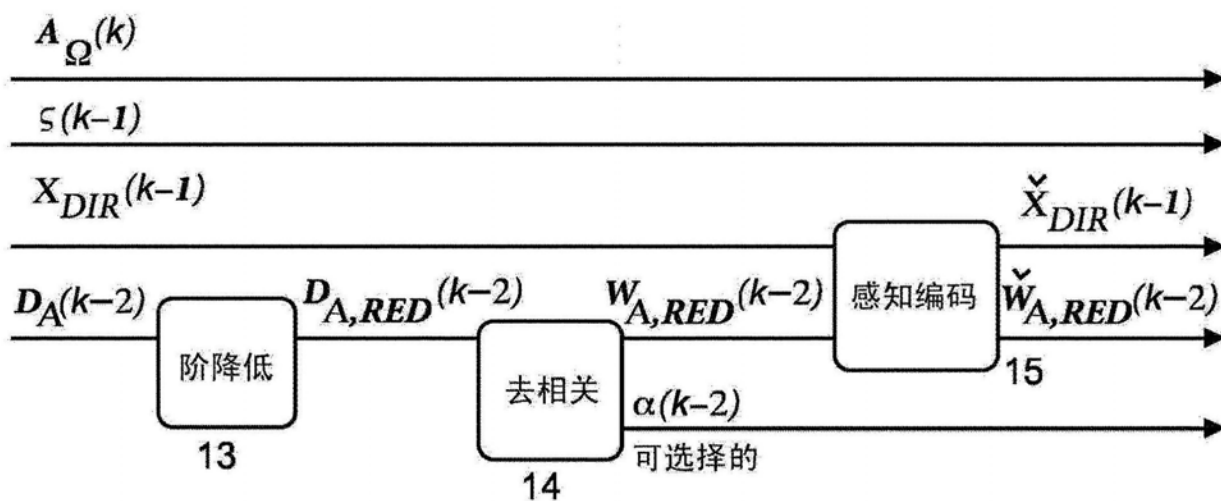


图1b

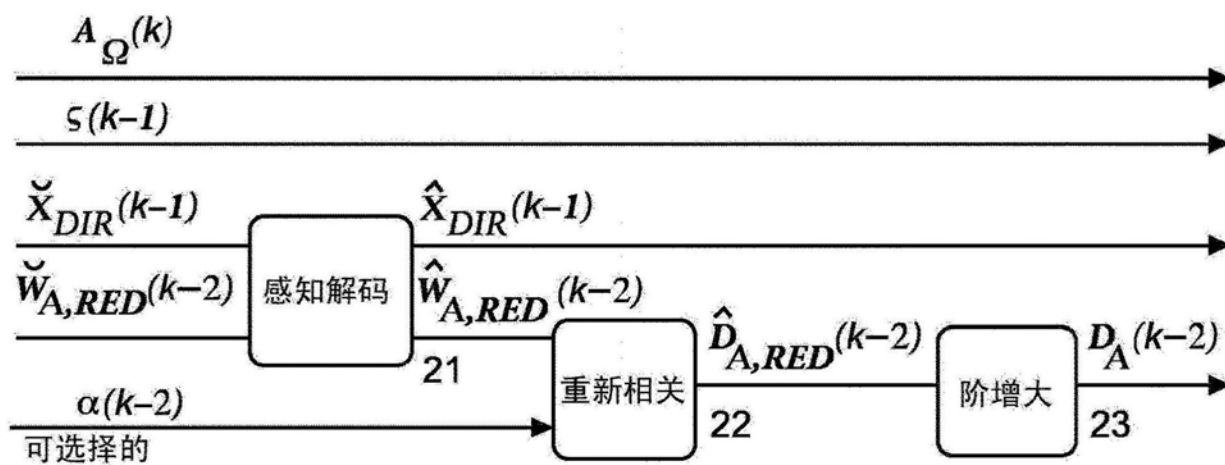


图2a

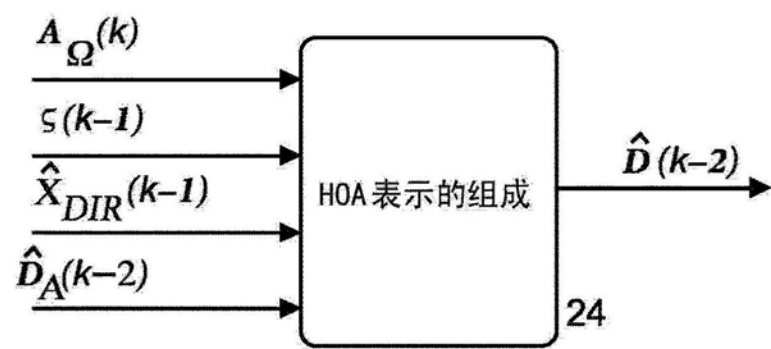


图2b

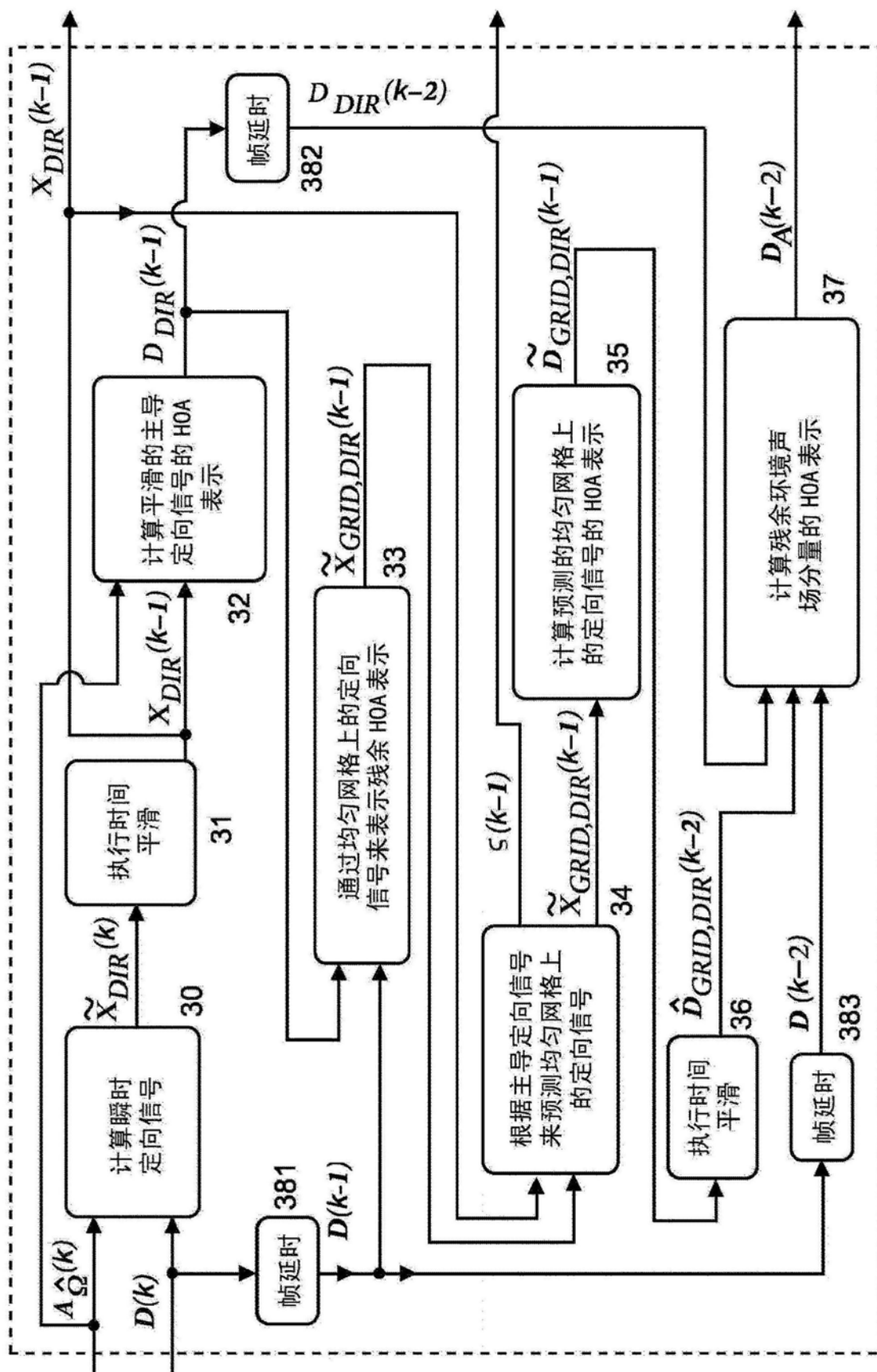


图3

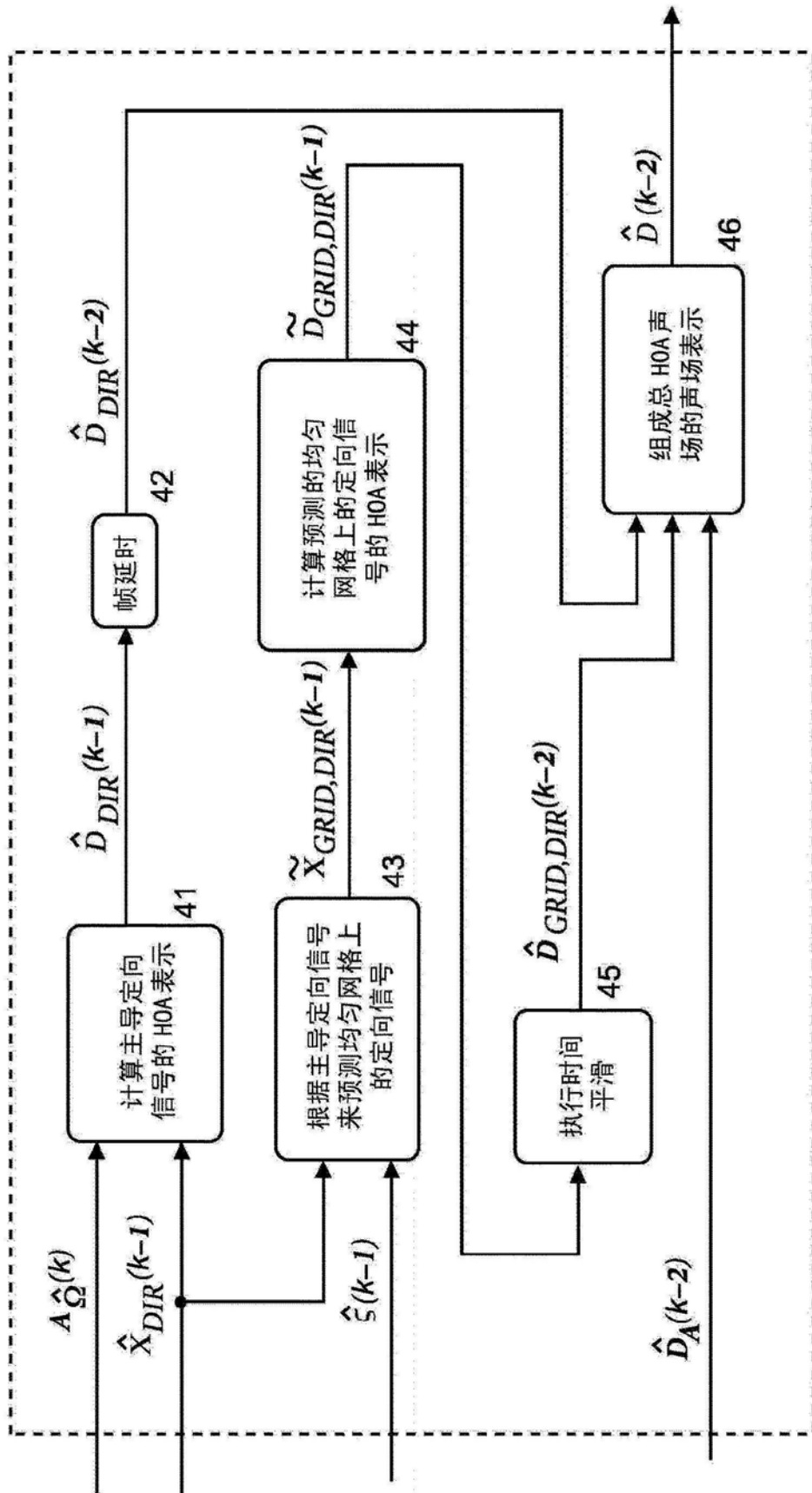


图4

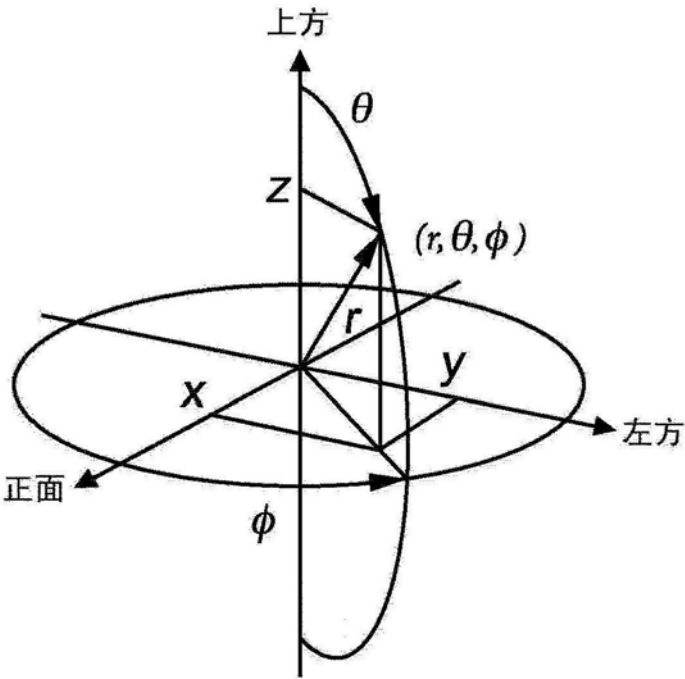


图5

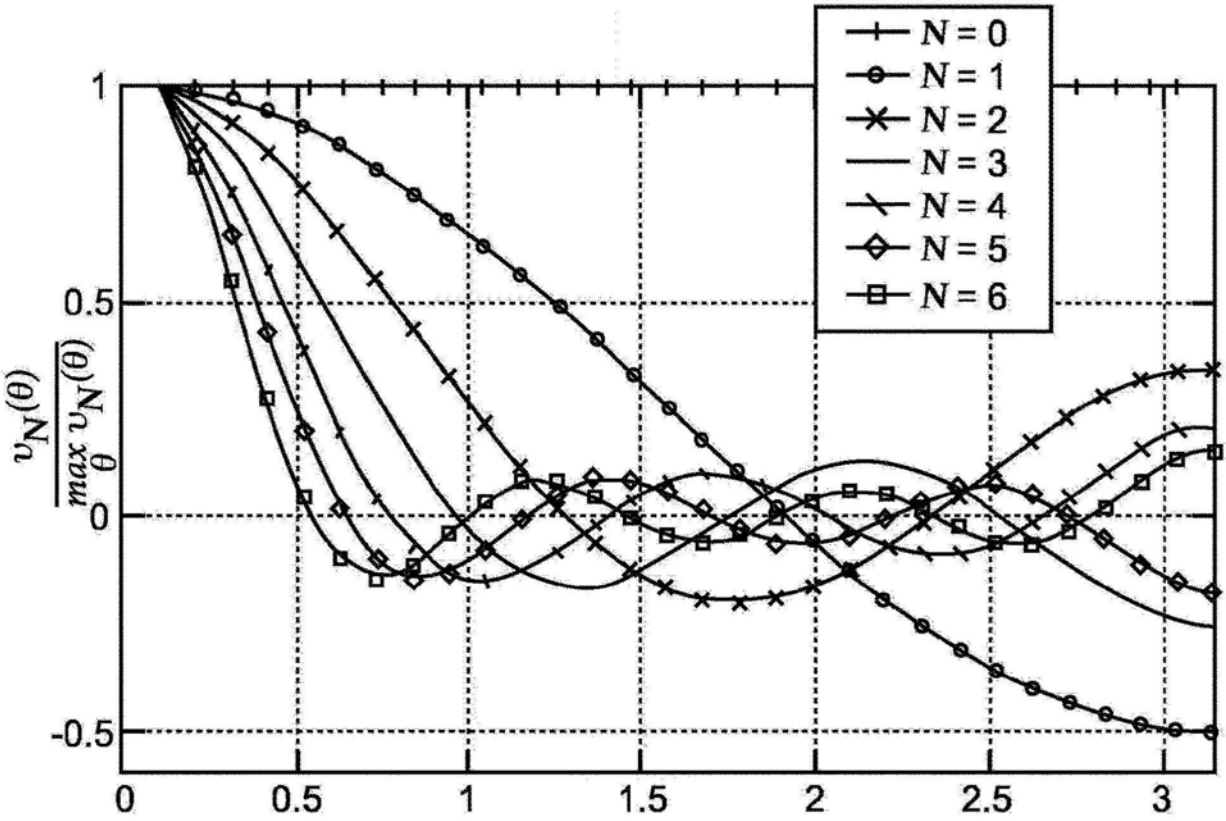


图6