(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 120032652 A
(43)申请公布日 2025.05.23

(21)申请号 202510186602.9
(22)申请日 2015.06.22
(30)优先权数据
14306026.7 2014.06.27 EP
(62)分案原申请数据
201580035127.X 2015.06.22
(71)申请人 杜比国际公司
地址 荷兰
(72)发明人 亚历山大·克鲁格 斯文·科尔东
(74)专利代理机构 中国贸促会专利商标事务所有限公司 11038
专利代理师 张劲松
(51)Int.Cl.
G10L 19/04(2013.01)
G10L 19/24(2013.01)
G10L 19/008(2013.01)

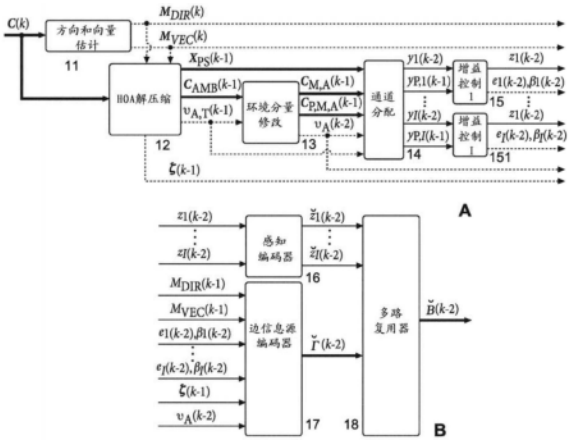
权利要求书2页 说明书15页 附图4页

(54)发明名称

针对HOA数据帧表示的压缩确定表示非差分增益值所需的最小整数比特数的方法

(57)摘要

本发明公开了针对HOA数据帧表示的压缩确定表示非差分增益值所需的最小整数比特数的方法。当对HOA数据帧表示进行压缩时,在每个通道信号被感知地编码(16)之前对其实施增益控制(15,151)。增益值作为边信息以差分的方式被传输。然而,为了开始对这样的流式压缩HOA数据帧表示进行解码,需要绝对增益值,应当以最小数量的比特对该绝对增益值进行编码。为了确定这样的最小整数比特量{βe),在空间域中将HOA数据帧表示(C(k))渲染为位于单位球体上的虚拟扬声器信号,随后对HOA数据帧表示(C(k))进行归一化。然后,将最小整数比特数设置为(AA)。



CN 120032652 A

1.一种用于对声音或声场的压缩的高阶高保真度立体声响复制HOA声音表示进行解码的方法,所述方法包括:

接收包含压缩的HOA表示的比特流,其中,所述比特流包括与所述压缩的HOA表示对应的多个HOA系数,

从所述比特流解复用所述压缩的HOA表示,以及

对所述压缩的HOA表示进行解码以确定解码的HOA表示,其中,所述解码是基于最小整数比特数β_e进行的,并且其中,所述最小整数比特数β_e基于

$\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil + 1) \rceil$确定,

其中,$K_{MAX} = \max_{1 \le N \le N_{MAX}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}\right)$,N是HOA表示的阶数,$N_{MAX}$是感兴趣的HOA表示的最大阶数,$\boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}$是虚拟扬声器的方向,$O=(N+1)^2$是HOA系数序列的数量,并且K是模矩阵的欧几里德范数的平方$||\Psi||_2^2$与O的比值,并且

其中,$\sqrt{K_{MAX}} = 1.5$,并且

其中,所述解码至少包括主要声音合成和环境合成。

2.一种用于对声音或声场的压缩的高阶高保真度立体声响复制HOA声音表示进行解码的设备,所述设备接收包含压缩的HOA表示的比特流,其中,所述比特流包括与所述压缩的HOA表示对应的多个HOA系数,并且所述设备包括:

解复用器,被配置为从所述比特流解复用所述压缩的HOA表示,以及

处理器,被配置为对所述压缩的HOA表示进行解码以确定解码的HOA表示,其中,所述解码是基于最小整数比特数β_e进行的,

其中,所述最小整数比特数β_e基于$\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil + 1) \rceil$确定,

其中,$K_{MAX} = \max_{1 \le N \le N_{MAX}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}\right)$,N是HOA表示的阶数,$N_{MAX}$是感兴趣的HOA表示的最大阶数,$\boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}$是虚拟扬声器的方向,$O=(N+1)^2$是HOA系数序列的数量,并且K是模矩阵的欧几里德范数的平方$||\Psi||_2^2$与O的比值,并且

其中,$\sqrt{K_{MAX}} = 1.5$,并且

其中,所述解码至少包括主要声音合成和环境合成。

3.一种非暂时性计算机可读介质,所述非暂时性计算机可读介质具有存储在其上的使计算机执行根据权利要求1所述的方法的步骤的可执行指令。

4.一种用于对声音或声场的压缩的高阶高保真度立体声响复制HOA声音表示进行解码的方法,所述方法包括:

从比特流解复用压缩的HOA表示,其中,多个HOA系数与所述压缩的HOA表示对应,以及

当独立的访问单元存在于所述比特流中时,基于最小整数比特数β_e对所述压缩的HOA表示进行解码,其中,所述最小整数比特数β_e基于$\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O) \rceil + e_{MAX} + 1) \rceil$确定,

其中，$K_{\mathrm{MAX}} = \max_{1 \le N \le N_{\mathrm{MAX}}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}\right)$，N是HOA表示的阶数，$N_{MAX}$是感兴趣的HOA表示的最大阶数，$\boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}$是虚拟扬声器的方向，$O = (N+1)^2$是HOA系数序列的数量，并且K是模矩阵的欧几里德范数的平方$||\Psi||_2^2$与O的比值，并且$e_{MAX} > 0$用于在通道信号在增益控制之前的采样值幅度小于阈值的情况下增大最小整数比特数β_e，$2^{e_{MAX}}$是通过增益控制施加到通道信号的最大增益。

5.一种用于解码声音或声场的压缩高阶高保真度立体声响复制(HOA)声音表示的方法，所述方法包括：

接收包含所述压缩HOA表示的比特流并且解码所述压缩HOA表示以确定感知解码信号$\hat{\boldsymbol{z}}_i(k),\ i=1,\ldots,I$、相关联的增益校正指数$e_i(k)$以及增益校正异常标记$\beta_i(k)$；

通过对于感知解码信号$\hat{\boldsymbol{z}}_i(k),\ i=1,\ldots,I$、相关联的增益校正指数$e_i(k)$以及增益校正异常标记$\beta_i(k)$执行逆增益控制处理，提供经增益校正的信号帧$\hat{\boldsymbol{y}}_i(k),\ i = 1,\ldots,I$；

在通道重分配期间重新分配经增益校正的信号帧$\hat{\boldsymbol{y}}_i(k),\ i = 1,\ldots,I$，以便重构主要声音信号的帧$\hat{\boldsymbol{X}}_{\mathrm{PS}}(k)$以及环境HOA分量的中间表示的帧$C_{I,AMB}(k)$。

针对HOA数据帧表示的压缩确定表示非差分增益值所需的最小整数比特数的方法

[0001] 本申请是针对申请号为202111089841.0的分案申请(其母案申请的申请号为201580035127.X)提出的分案申请。

技术领域

[0002] 本发明涉及用于针对HOA数据帧表示的压缩而确定表示与所述HOA数据帧中的特定数据帧的通道信号相关联的非差分增益值所需的最小整数比特数的方法。

背景技术

[0003] 表示为HOA的高阶高保真度立体声响复制提供了表示三维声音的一种可能性。其他技术是波场合成(WFS)或如22.2的基于通道的方法。与基于通道的方法相比,HOA表示提供了与特定扬声器设置无关的优点。然而,这种灵活性是以在特定扬声器设置上回放HOA表示所需要的解码过程为代价的。与所需扬声器的数量通常很大的WFS方法相比,HOA也可以被呈现为仅包括几个扬声器的设置。HOA的另一个优点是还可以采用相同的表示而不必对耳机的双耳渲染进行任何修改。

[0004] HOA基于通过截断的球面谐波函数(SH)展开来表示复合谐波平面波幅度的空间密度。每个展开系数是角频率的函数,角频率可以由时域函数等效地表示。因此,在不失一般性的情况下,完整的HOA声场表示实际上可以被假设成由O个时域函数组成,其中,O表示展开系数的数量。这些时域函数在下文中将被等效地称为HOA系数序列或HOA通道。

[0005] HOA表示的空间解析度随着展开的最大阶数N的增长而提高。遗憾的是,展开系数O的数量随着阶数N呈二次方增长,特别地,$O=(N+1)^2$。例如,使用阶数N=4的典型HOA表示需要O=25个HOA(展开)系数。假设期望的单声道采样率为f_S并且每个采样的比特数为N_b,则用于传输HOA表示的总比特率由$O \cdot f_S \cdot N_b$确定。以采用每采样$N_b=16$比特的$f_S=48kHz$采样率传输阶数为N=4的HOA表示,导致19.2MBits/s的比特率,该比特率对于许多实际应用(例如流式传输)来说是非常高的。因此,对HOA表示进行压缩是非常可取的。

[0006] 先前,在EP 2665208 A1、EP 2743922 A1、EP 2800401 A1中提出了HOA声场表示的压缩,参见2014年1月的MPEG-H 3D音频的ISO/IEC JTC1/SC29/WG11、N14264、WD1-HOA文本。这些方法的共同点在于:它们都执行声场分析并且将给定的HOA表示分解成方向分量和残余环境分量。一方面,最终压缩的表示被假设成由若干量化信号组成,这些量化信号由方向信号和基于向量的信号的感知编码以及环境HOA分量的相关系数序列产生。另一方面,最终压缩的表示包括与量化信号相关的附加边信息,根据其压缩版本重构HOA表示需要所述边信息。

[0007] 在被传递至感知编码器之前,要求这些中间时域信号具有在[-1,1]的值范围内的最大幅度,这是为实现当前可用的感知编码器而产生的要求。为了当对HOA表示进行压缩时满足该要求,在感知编码器之前使用平滑地衰减或放大输入信号的增益控制处理单元(参见EP 2824661 A1和上面提到的ISO/IEC JTC1/SC29/WG11 N14264文献)。所产生的信号修

改被假设成是可逆的并且被逐帧地应用,其中特别地,连续帧之间的信号幅度的变化被假设成“2”的幂。为了便于该信号修改在HOA解压缩器中的倒置,相应的归一化边信息包括在总边信息中。该归一化边信息可以由底数为“2”的指数构成,这些指数描述了两个连续帧之间的相对幅度变化。由于连续帧之间的较小幅度变化比较大幅度变化更有可能发生,因此根据上面提到的ISO/IEC JTC1/SC29/WG11 N14264文献利用游程运行长度编码(run length code)对这些指数进行编码。

发明内容

[0008] 例如,在从开始至结束没有任何时间跳跃地对单个文件解压缩的情况下,在HOA解压缩中使用差分编码的幅度变化来重构原始信号幅度是可行的。然而,为了便于随机访问,独立的访问单元必须存在于编码表示(其通常是比特流)中以使得能够与来自在先帧的信息无关地从期望的位置(或至少在其附近)开始解压缩。这种独立的访问单元必须包含由增益控制处理单元引起的从第一帧一直到当前帧的总绝对幅度变化(即,非差分增益值)。假设两个连续帧之间的幅度变化是“2”的幂,则通过底数为“2”的指数来描述总绝对幅度变化就足够了。为了对该指数进行高效编码,在应用增益控制处理单元之前了解信号的可能的最大增益是必要的。然而,该知识高度依赖于对要压缩的HOA表示的值范围的约束规范。遗憾的是,MPEG-H 3D音频文献ISO/IEC JTC1/SC29/WG11 N14264仅提供用于输入HOA表示的格式的描述,而没有对值范围设定任何约束。

[0009] 本发明要解决的问题是提供表示非差分增益值所需的最小整数比特数。

[0010] 本发明确立了输入HOA表示的值范围与信号在HOA压缩器中应用增益控制处理单元之前的可能的最大增益之间的相互关系。

[0011] 基于该相互关系,针对输入HOA表示的值范围的给定规范、针对底数为“2”的指数的有效编码来确定所需比特的量,以在访问单元内描述由增益控制处理单元引起的从第一帧一直到当前帧的修改信号的总绝对幅度变化(即,非差分增益值)。

[0012] 此外,一旦计算用于对指数编码的所需比特量的规则被确定,本发明就使用用于验证给定HOA表示是否满足所需值范围约束的处理,使得给定HOA表示能够被正确压缩。

[0013] 原则上,本发明的方法适合于针对HOA数据帧表示的压缩而确定用于表示所述HOA数据帧中的特定HOA数据帧的通道信号的非差分增益值所需的最小整数比特数β_e,其中,每个帧中的每个通道信号包括一组采样值,并且其中,给所述HOA数据帧中的每个HOA数据帧的每个通道信号分配差分增益值,并且这样的差分增益值引起当前HOA数据帧中的通道信号的采样值的幅度相对于前一HOA数据帧中的通道信号的采样值变化,并且其中,这样的增益调整的通道信号在编码器中被编码,

[0014] 并且其中,所述HOA数据帧表示在空间域中被渲染为O个虚拟扬声器信号$w_j(t)$,其中,所述O个虚拟扬声器的位置位于单位球体上并且与针对β_e的计算而假设的位置不匹配,所述渲染通过矩阵乘法$w(t)=(\Psi)^{-1}\cdot c(t)$来表示,其中,w(t)是包含所有虚拟扬声器信号的向量,Ψ是针对虚拟扬声器位置而计算的模矩阵,并且c(t)是所述HOA数据帧表示的相应HOA系数序列的向量,

[0015] 并且其中,计算最大允许幅度值$\gamma=\min\left(1,\frac{\sqrt{O}\cdot\sqrt{K_{MAX,DES}}}{\|\Psi\|_2}\right)$,并且所述HOA数据帧表

示被归一化,使得$\|\boldsymbol{w}(t)\|_\infty = \max_{1\le j\le O}|w_j(t)| \le 1 \quad \forall t$,

[0016] 所述方法包括以下步骤:

[0017] -通过以下子步骤a)、b)、c)中的一个或更多个从所述归一化HOA数据帧表示中形成所述通道信号:

[0018] a)为了表示所述通道信号中的主要声音信号,将所述HOA系数序列c(t)的向量乘以混合矩阵A,混合矩阵A的欧几里德范数不大于"1",其中,混合矩阵A表示所述归一化HOA数据帧表示的系数序列的线性组合;

[0019] b)为了表示所述通道信号中的环境分量$c_{AMB}(t)$、从所述归一化HOA数据帧表示中减去所述主要声音信号以及选择所述环境分量$c_{AMB}(t)$的系数序列的至少一部分,其中,$||c_{AMB}(t)||_2^2 \le ||c(t)||_2^2$,以及通过计算$\boldsymbol{w}_{MIN}(t) = \boldsymbol{\Psi}_{MIN}^{-1} \cdot \boldsymbol{c}_{AMB,MIN}(t)$对所得到的最小环境分量$c_{AMB,MIN}(t)$进行变换,其中,$\|\boldsymbol{\Psi}_{MIN}^{-1}\|_2 < 1$并且$\Psi_{MIN}$是所述最小环境分量$c_{AMB,MIN}(t)$的模矩阵;

[0020] c)选择所述HOA系数序列c(t)的一部分,其中,所选择的系数序列与对其实施空间变换的环境HOA分量的系数序列相关,并且描述所选择的系数序列的数量的最小阶数N_{MIN}为$B_{MIN} \le 9$;

[0021] -将用于表示所述通道信号的所述非差分增益值所需的所述最小整数比特数β_e设置为$\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX}} \cdot O)\rceil + 1)\rceil$,

[0022] 其中,$K_{MAX,DES} = \max_{1\le N\le N_{MAX,DES}} K(N, \boldsymbol{\Omega}_{DES,1}^{(N)}, \ldots, \boldsymbol{\Omega}_{DES,O}^{(N)})$,N是阶数,$O=(N+1)^2$是HOA系数序列的数量,K是所述模矩阵的欧几里德范数的平方与O之间的比值,并且其中,$N_{MAX,DES}$是感兴趣的阶数,并且$\boldsymbol{\Omega}_{DES,1}^{(N)}, \ldots, \boldsymbol{\Omega}_{DES,1}^{(N)}$是针对每个阶数的虚拟扬声器的方向,其中该方向是为了实现对所述HOA数据帧表示的所述压缩而被假定的,使得通过$\beta_e = \lceil \log_2(\lceil \log_2(\sqrt{K_{MAX,DES}} \cdot O)\rceil + 1)\rceil$来选择$\beta_e$,从而对所述非差分增益值的底数为"2"的指数进行编码,

[0023] 并且其中,对于计算$\gamma = \min\left(1, \frac{\sqrt{O}\cdot\sqrt{K_{MAX,DES}}}{\|\Psi\|_2}\right)$,$||\Psi||_2$是所述模矩阵$\Psi$的欧几里德范数,$K_{MAX} = \max_{1\le N\le N_{MAX}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}\right)$,N是阶数,$N_{MAX}$是感兴趣的最大阶数,$\boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}$是所述虚拟扬声器的方向,$O=(N+1)^2$是HOA系数序列的数量,并且K是所述模矩阵的欧几里德范数的平方$||\Psi||_2^2$与O之间的比值。

附图说明

[0024] 参考附图描述了本发明的示例性实施方式,在附图中示出了:

[0025] 图1 HOA压缩器;

[0026] 图2 HOA解压缩器;

[0027] 图3虚拟方向$\Omega_j^{(N)}$ $(1\le j\le O)$关于HOA阶数$(N=1,\ldots,29)$的缩放值K;

[0028] 图4对于HOA阶数($N_{MIN}=1,...,9$),逆模矩阵Ψ^{-1}关于虚拟方向$\Omega_{MIN,d}$($d=1,...,O_{MIN}$)的欧几里德范数;

[0029] 图5虚拟扬声器在位置$\Omega_j^{(N)}$($1\leq j\leq O$,其中$O=(N+1)^2$)处的信号的最大允许幅度γ_{dB}的确定;

[0030] 图6球坐标系。

具体实施方式

[0031] 即使没有明确描述,也可以在任何组合或子组合中使用以下实施方式。

[0032] 在下文中,介绍HOA压缩和解压缩的原理以提供存在上述问题的更详细背景。该介绍的基础是在MPEG-H 3D音频文档ISO/IEC JTC1/SC29/WG11 N14264(也参见EP 2665208 A1、EP 2800401 A1和EP 2743922 A1)中描述的处理。在N14264中,"方向分量"被扩展至"主要声音分量"。作为方向分量,主要声音分量被假设成部分地由方向信号连同用于根据方向信号预测原始HOA表示的多个部分的一些预测参数一起来表示,方向信号是指具有被假设为从其冲击听者的相应方向的单声道信号。另外,主要声音分量被假设为由"基于向量的信号"表示,基于向量的信号是指具有限定基于向量的信号的方向分布的相应向量的单声道信号。

[0033] HOA压缩

[0034] 图1示出了在EP 2800401 A1中描述的HOA压缩器的总体架构。该HOA压缩器的总体架构具有图1A所示的空间HOA编码部以及图1B所示的感知编码部及源编码部。空间HOA编码器提供由I信号连同描述如何创建其HOA表示的边信息一起构成的第一压缩HOA表示。在对两个编码的表示进行多路复用之前,在感知编码器和边信息源编码器中对I信号进行感知编码,并且对边信息进行源编码。

[0035] 空间HOA编码

[0036] 在第一步骤中,将原始HOA表示的当前第k帧C(k)输入至方向和向量估计处理步骤或阶段11,该当前第k帧C(k)被假设成提供元组集$\mathcal{M}_{\mathrm{DIR}}(k)$和$\mathcal{M}_{\mathrm{VEC}}(k)$。元组集$\mathcal{M}_{\mathrm{DIR}}(k)$由其第一元素表示方向信号的索引并且第二元素表示相应的量化方向的元组构成。元组集$\mathcal{M}_{\mathrm{VEC}}(k)$由其第一元素表示基于向量的信号的索引并且第二元素表示限定信号的方向分布的向量(即,如何计算基于向量的信号的HOA表示)的元组构成。

[0037] 使用两个元组集$\mathcal{M}_{\mathrm{DIR}}(k)$和$\mathcal{M}_{\mathrm{VEC}}(k)$,在HOA分解步骤或阶段12中将初始HOA帧C(k)分解成所有主要声音(即,方向的和基于向量的)信号的帧$X_{PS}(k-1)$和环境HOA分量的帧$C_{AMB}(k-1)$。注意由交叠相加处理引起的一个帧的延迟,以避免堵塞的假象。此外,HOA分解步骤/阶段12被假设成输出描述如何根据方向信号来预测原始HOA表示的多个部分的一些预测参数$\zeta(k-1)$,以丰富主要声音HOA分量。另外,假设提供了包含关于将在HOA分解处理步骤或阶段12中确定的主要声音信号分配给I个可用通道的信息的目标分配向量$v_{A,T}(k-1)$。可以假设要占用受影响的通道,这意味着受影响的通道不能用于在相应的时间帧中传输环境HOA分量的任何系数序列。

[0038] 在环境分量修改处理步骤或阶段13中,根据由目标分配向量$v_{A,T}(k-1)$提供的信息来修改环境HOA分量的帧$C_{AMB}(k-1)$。特别地,(在其他方面中)根据关于哪些通道可用并且还

没有被主要声音信号占据的(包含在目标分配向量$v_{A,T}(k-1)$中的)信息来确定要在给定的I个通道中传输环境HOA分量的哪些系数序列。

[0039] 另外,如果所选择的系数序列的索引在连续帧之间变化,则执行系数序列的淡入淡出。

[0040] 此外,假设环境HOA分量$C_{AMB}(k-2)$的第一O_{MIN}系数序列总是被选择成被感知地编码和传输,其中$O_{MIN}=(N_{MIN}+1)^2$ ($N_{MIN}\leq N$)的阶数通常比原始HOA表示的阶数更小。为了对这些HOA系数序列进行去相关,可以在步骤/阶段13中将它们变换为从一些预定义方向$\Omega_{MIN,d}$ ($d=1,\ldots,O_{MIN}$)冲击的方向信号(即,一般平面波函数)。

[0041] 暂时预测的已修改环境HOA分量$C_{P,M,A}(k-1)$连同修改的环境HOA分量$C_{M,A}(k-1)$一起在步骤/阶段13中被计算,并且被用于增益控制处理步骤/阶段15,…,151以实现合理的预见,其中关于环境HOA分量的修改的信息与在通道分配步骤或阶段14中将所有可能类型的信号分配给可用通道直接相关。关于该分配的最终信息被假设成包含在最终分配向量$v_A(k-2)$中。为了在步骤/阶段13中计算该向量,利用包含在目标分配向量$v_{A,T}(k-1)$中的信息。

[0042] 步骤/阶段14中的通道分配利用由分配向量$v_A(k-2)$提供的信息将包含在帧$X_{PS}(k-2)$中和包含在帧$C_{M,A}(k-2)$中的适当的信号分配给I个可用通道,从而得到信号帧$y_i(k-2)$,$i=1,\ldots,I$。另外,还将包含在帧$X_{PS}(k-1)$和帧$C_{P,AMB}(k-1)$中的适当的信号分配给I个可用通道,从而得到预测的信号帧$y_{P,i}(k-1)$,$i=1,\ldots,I$。

[0043] 信号帧$y_i(k-2)$,$i=1,\ldots,I$中的每一个最终通过增益控制处理步骤/阶段15,…,151来处理,以得到指数$e_i(k-2)$和异常标记$\beta_i(k-2)$,$i=1,\ldots,I$以及信号$z_i(k-2)$,$i=1,\ldots,I$,其中信号增益被平滑地修改以实现适合于感知编码器步骤或阶段16的值范围。步骤/阶段16输出相应的编码信号帧$\breve{\mathbf{z}}_i(k-2),\ i=1,\ldots,I$。预测的信号帧$y_{P,i}(k-1)$,$i=1,\ldots,I$实现了合理的预见以避免连续块之间的较大增益变化。在边信息源编码器步骤或阶段17中,对边信息数据$\mathcal{M}_{\mathrm{DIR}}(k-1)$、$\mathcal{M}_{\mathrm{VEC}}(k-1)$、$e_i(k-2)$、$\beta_i(k-2)$、$\zeta(k-1)$和$v_A(k-2)$进行源编码,以得到经编码的边信息帧$\breve{\boldsymbol{\Gamma}}(k-2)$。在多路复用器18中,对帧(k-2)的编码信号$\breve{\mathbf{z}}_i(k-2)$与该帧的已编码边信息数据$\breve{\boldsymbol{\Gamma}}(k-2)$进行组合,以得到输出帧$\breve{\boldsymbol{B}}(k-2)$。

[0044] 在空间HOA解码器中,增益控制处理步骤/阶段15,…,151中的增益修改被假设成通过使用由指数$e_i(k-2)$和异常标记$\beta_i(k-2)$,$i=1,\ldots,I$构成的增益控制边信息来恢复。

[0045] HOA解压缩

[0046] 图2示出了在EP 2800401 A1中描述的HOA解压缩器的总体架构。该总体架构由HOA压缩器部件的配对部件构成,所述配对部件以相反的顺序布置并且包括图2A所示的感知解码部和源解码部以及图2B所示的空间HOA解码部。

[0047] 在感知解码部和源解码部(代表感知解码器和边信息源解码器)中,解复用步骤或阶段21从比特流接收输入帧$\breve{\boldsymbol{B}}(k)$,并且提供I个信号的感知编码的表示$\breve{\mathbf{z}}_i(k),\ i=1,\ldots,I$以及描述如何创建其HOA表示的已编码边信息数据$\breve{\boldsymbol{\Gamma}}(k)$。在感知解码器步骤或阶段22中对$\breve{\mathbf{z}}_i(k)$信号进行感知解码,以得到解码信号$\hat{\mathbf{z}}_i(k),\ i=1,\ldots,I$。在边信息源解码器步骤或阶段23中对已编码边信息数据$\breve{\boldsymbol{\Gamma}}(k)$进行解码,以得到数据集

$\mathcal{M}_{\mathrm{DIR}}(k+1)$、$\mathcal{M}_{\mathrm{VEC}}(k+1)$、指数$e_i(k)$、异常标记$\beta_i(k)$、预测参数$\zeta(k+1)$和分配向量$v_{AMB,ASSIGN}(k)$。关于$v_A$与$v_{AMB,ASSIGN}$之间的差别,参见上面提到的MPEG文献N14264。

[0048] 空间HOA解码

[0049] 在空间HOA解码部中,感知解码的信号$\hat{\boldsymbol{z}}_i(k),\ i=1,\ldots,I$中的每一个连同其关联的增益校正指数$e_i(k)$以及增益校正异常标记$\beta_i(k)$一起被输入至逆增益控制处理步骤或阶段24、241。第i个逆增益控制处理步骤/阶段提供了经增益校正的信号帧$\hat{\boldsymbol{y}}_i(k)$。

[0050] 全部I个经增益校正的信号帧$\hat{\boldsymbol{y}}_i(k),\ i=1,\ldots,I$连同分配向量$v_{AMB,ASSIGN}(k)$以及元组集$\mathcal{M}_{\mathrm{DIR}}(k+1)$和$\mathcal{M}_{\mathrm{VEC}}(k+1)$一起被馈送至通道重分

[0051] 配步骤或阶段25,参见元组集$\mathcal{M}_{\mathrm{DIR}}(k+1)$和$\mathcal{M}_{\mathrm{VEC}}(k+1)$的上述定义。分配向量$v_{AMB,ASSIGN}(k)$由I个分量构成,所述I个分量针对每个传输通道而指示其是否包含环境HOA分量的系数序列以及它包含哪个系数序列。在通道重分配步骤/阶段25中,经增益校正的信号帧$\hat{\boldsymbol{y}}_i(k)$被重新分配以重构所有主要声音信号(即,所有方向信号和基于向量的信号)的帧$\hat{\boldsymbol{X}}_{\mathrm{PS}}(k)$以及环境HOA分量的中间表示的帧$C_{I,AMB}(k)$。另外,提供了在第k个帧中活跃的环境HOA分量的系数序列的索引的集合$\mathcal{I}_{\mathrm{AMB,ACT}}(k)$,以及在第(k-1)个帧中必须被启用、禁用及保持活跃的环境HOA分量的系数索引的数据集$\mathcal{I}_{\mathrm{E}}(k-1)$、$\mathcal{I}_{\mathrm{D}}(k-1)$和$\mathcal{I}_{\mathrm{U}}(k-1)$。

[0052] 在主要声音合成步骤或阶段26中,利用元组集$\mathcal{M}_{\mathrm{DIR}}(k+1)$、预测参数的集合$\zeta(k+1)$、元组集$\mathcal{M}_{\mathrm{VEC}}(k+1)$以及数据集$\mathcal{I}_{\mathrm{E}}(k-1)$、$\mathcal{I}_{\mathrm{D}}(k-1)$和$\mathcal{I}_{\mathrm{U}}(k-1)$,根据所有主要声音信号的帧$\hat{\boldsymbol{X}}_{\mathrm{PS}}(k)$来计算主要声音分量$\hat{\boldsymbol{C}}_{\mathrm{PS}}(k-1)$的HOA表示。

[0053] 在环境合成步骤或阶段27中,利用在第k个帧中活跃的环境HOA分量的系数序列的索引的集合$\mathcal{I}_{\mathrm{AMB,ACT}}(k)$,根据环境HOA分量的中间表示的帧$C_{I,AMB}(k)$来创建环境HOA分量帧$\hat{\boldsymbol{C}}_{\mathrm{AMB}}(k-1)$。由于与主要声音HOA分量的同步而引入一个帧的延迟。

[0054] 最后,在HOA组成步骤或阶段28中,将环境HOA分量帧$\hat{\boldsymbol{C}}_{\mathrm{AMB}}(k-1)$与主要声音HOA分量的帧$\hat{\boldsymbol{C}}_{\mathrm{PS}}(k-1)$进行叠加,以提供已解码的HOA帧$\hat{\boldsymbol{C}}(k-1)$。

[0055] 此后,空间HOA解码器根据I个信号和边信息来创建重构的HOA表示。

[0056] 在位于编码侧的情况下,环境HOA分量被变换为方向信号,在步骤/阶段27中在解码器侧进行该变换的逆变换。

[0057] 在HOA压缩器内的增益控制处理步骤/阶段15,…,151之前,信号的可能最大增益十分依赖于输入HOA表示的值范围。因此,首先限定输入HOA表示的有意义的值范围,随后在进入增益控制处理步骤/阶段之前对信号的可能最大增益做出结论。

[0058] 输入HOA表示的归一化

[0059] 为了使用本发明的处理,要先执行对(总)输入HOA表示信号的归一化。对于HOA压缩,执行逐帧处理,其中关于在高阶高保真度立体声响复制的章节Basics中的公式(54)中指定的时间连续HOA系数序列的向量c(t),将原始输入HOA表示的第k个帧C(k)限定为

[0060] $$\boldsymbol{C}(k) := \left[\boldsymbol{c}\left((kL+1)T_S\right) \quad \boldsymbol{c}\left((kL+2)T_S\right) \quad \dots \quad \boldsymbol{c}\left((k+1)LT_S\right)\right] \in \mathbb{R}^{O\times L} \qquad (1)$$

[0061] 其中,k表示帧索引,L为(在采样中的)帧长度,O=(N+1)^{2}为HOA系数序列的数量,以及T_S表示采样周期。

[0062] 如EP 2824661 A1中所提及的,从实际角度来看,HOA表示的有意义的归一化不是通过对个别HOA系数序列$c_n^m(t)$的值范围施加约束来实现的,这是因为这些时域函数不是在渲染之后由扬声器实际播放的信号。相反,更方便的是考虑通过将HOA表示渲染为O个虚拟扬声器信号$w_j(t)$,1≤j≤O而获得的"等效空间域表示"。假设相应的虚拟扬声器位置借助于球坐标系来表示,其中假设每个位置位于单位球体上并且半径为"1"。因此,可以通过阶数相关方向$\Omega_j^{(N)}=(\theta_j^{(N)},\varphi_j^{(N)})$,1≤j≤O等效地表达位置,其中$\theta_j^{(N)}$和$\varphi_j^{(N)}$分别表示倾斜度和方位角(也参见图6及其关于球坐标系定义的描述)。例如参见J.Fliege、U.Maier于1999年在多特蒙德大学的专业课范围数学技术报告"A two-stage approach for computing cubature formulae for the sphere",这些方向应当尽可能均匀地分布在单位球体上。在以下网址中可以找到用于特定方向的计算的节点数量:http://www.mathematik.uni-dortmund.de/lsx/research/projects/fliege/node s/nodes.html。这些位置通常取决于"球上的均匀分布"的定义种类,因此是不明确的。

[0063] 通过限定HOA系数序列的值范围来限定虚拟扬声器信号的值范围的优点在于:如常规扬声器信号假定PCM表示的情况那样,虚拟扬声器信号的值范围可以被直观地设置为等于区间[-1,1]。这导致空间上均匀分布的量化误差,使得在与实际聆听相关的域中有利地应用量化。该背景中的一个重要方面是每采样比特数可以被选择成与通常用于常规扬声器信号的比特数(即,16)一样低,与通常需要更高的每采样比特数(例如,24或甚至32)的HOA系数序列的直接量化相比,这提高了效率。

[0064] 为了详细描述空间域中的归一化处理,所有虚拟扬声器信号以向量被概括为w(t):=$[w_1(t)\dots w_0(t)]^T$, (2)

[0065] 其中,$(\bullet)^T$表示转置。用Ψ表示关于虚拟方向$\Omega_j^{(N)}$,1≤j≤O的模矩阵,Ψ被定义为$\boldsymbol{\Psi} := [S_1 \quad \dots \quad S_O] \in \mathbb{R}^{O\times O}$, (3)其中,

[0066] $$\mathbf{S_j} = \left[S_0^0(\boldsymbol{\Omega}_j^{(N)}) \quad S_1^{-1}(\boldsymbol{\Omega}_j^{(N)}) \quad S_1^0(\boldsymbol{\Omega}_j^{(N)}) \quad S_1^1(\boldsymbol{\Omega}_j^{(N)}) \quad \dots \quad S_N^{N-1}(\boldsymbol{\Omega}_j^{(N)}) \quad S_N^N(\boldsymbol{\Omega}_j^{(N)})\right]^T \qquad (4)$$

[0067] ,渲染处理可以用公式表示为矩阵乘积

[0068] $w(t)=(\Psi)^{-1}\bullet c(t)$。 (5)

[0069] 使用这些定义,对虚拟扬声器信号的合理要求是:

[0070] $$\boxed{\|\boldsymbol{w}(lT_S)\|_\infty = \max_{1\le j\le O}\left|w_j(lT_S)\right| \le 1 \quad \forall l}, \qquad (6)$$

[0071] 这意味着每个虚拟扬声器信号的幅度需要落入范围[-1,1]内。时间t的时刻由所述HOA数据帧的采样值的采样索引l和采样周期T_S来表示。

[0072] 扬声器信号的总的幂因此满足条件

[0073] $$\|\boldsymbol{w}(lT_S)\|_2^2 = \sum_{j=1}^{O}\left|w_j(lT_S)\right|^2 \le O \quad \forall l \text{。} \qquad (7)$$

[0074] HOA数据帧表示的渲染和归一化在图1A的输入C(k)的上游执行。

[0075] 增益控制之前的信号值范围结果

[0076] 假设输入HOA表示的归一化是根据输入HOA表示的归一化小节中的描述执行的,下面考虑被输入至HOA压缩器中的增益控制处理单元的信号y_i,i=1,...,I的值范围。这些信号是通过向HOA系数序列或主要声音信号$x_{PS,d}$,d=1,...,D和/或环境HOA分量$c_{AMB,n}$,n=1,...,O的特定系数序列中的一个或更多个分配可用I个通道来创建的,对这些信号中的一部分实施空间变换。因此,在公式(6)中的归一化假设下,有必要分析所提及的这些不同信号类型的可能值范围。由于所有种类的信号是根据原始HOA系数序列而在中间计算出的,因此检查它们可能的值范围。

[0077] 图1A和图2B中未描绘I个通道中包含仅一个或更多个HOA系数序列的情况,即,在这种情况下,不需要HOA分解、环境分量修改块和相应的合成块。

[0078] HOA表示的值范围结果

[0079] 时间连续的HOA表示是通过c(t)=Ψw(t), (8)

[0080] 从虚拟扬声器信号中获得的,公式(8)是公式(5)的逆运算。

[0081] 因此,使用公式(8)和公式(7)来如下限制所有HOA系数序列的总的幂:

[0082] $||c(lT_S)||_2^2 \leq ||\Psi||_2^2 \cdot ||W(lT_S)||_2^2 \leq ||\Psi||_2^2 \cdot O$ (9)

[0083] 在球面谐波函数的N3D归一化的假设下,模矩阵的欧几里德范数的平方可以写成:$||\Psi||_2^2 = K \cdot O$, (10a)

[0084] 其中,

$$K = \frac{\|\Psi\|_2^2}{O} \quad \textbf{(10b)}$$

表示模矩阵的欧几里德范数的平方与HOA系数序列的数量O之间的比值。该比值取决于特定HOA阶数N和特定虚拟扬声器方向$\boldsymbol{\Omega}_j^{(N)}$, $1 \leq j \leq O$,其可以通过向该比值附加相应参数列表来如下表示:

[0085] $$K = K(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}) \quad \textbf{(10c)}$$

[0086] 图3示出了根据上面提到的Fliege等人的文章的虚拟方向$\boldsymbol{\Omega}_j^{(N)}$, $1 \leq j \leq O$关于HOA阶数(N=1,...,29)的K的值。

[0087] 结合所有先前的论证和考虑,提供如下HOA系数序列的幅度的上限:

[0088] $$\boxed{\|\boldsymbol{c}(lT_S)\|_\infty \leq \|\boldsymbol{c}(lT_S)\|_2 \leq \sqrt{K} \cdot O} \quad \textbf{(11)}$$

[0089] 其中,第一个不等式直接从范数定义得出。

[0090] 重要的是注意:公式(6)中的条件意味着公式(11)中的条件,但是相反的情况不成立,即,公式(11)不意味着公式(6)。

[0091] 另一个重要方面是:在虚拟扬声器位置近似均匀分布的假设下,模矩阵Ψ的表示关于虚拟扬声器位置的模向量的列向量几乎彼此正交并且每个都具有欧几里德范数N+1。该特性意味着:除了乘法常数之外,空间变换几乎保持欧几里德范数,即,

[0092] $||c(lT_S)||_2 \approx (N+1)||w(lT_S)||_2$。 (12)

[0093] 真实的范数$||c(lT_S)||_2$与公式(12)中的近似值相差越多,就越违反对模向量的正

交性假设。

[0094] 主要声音信号的值范围结果

[0095] 两种类型的(方向的和基于向量的)主要声音信号的共同点在于:它们对HOA表示的贡献由具有欧几里德范数N+1的单个向量$\boldsymbol{v}_1 \in \mathbb{R}^O$来描述,即,$||v_1||_2$=N+1。 (13)

[0096] 在方向信号的情况下,该向量与关于某个信号源方向$\Omega_{S,1}$的模向量相对应,即,

$$\boldsymbol{v}_1 = \boldsymbol{S}(\boldsymbol{\Omega}_{S,1}) \quad (14)$$

[0097]

$$:= [S_0^0(\boldsymbol{\Omega}_{S,1}) \quad S_1^{-1}(\boldsymbol{\Omega}_{S,1}) \quad S_1^0(\boldsymbol{\Omega}_{S,1}) \quad S_1^1(\boldsymbol{\Omega}_{S,1}) \quad \dots \quad S_N^{N-1}(\boldsymbol{\Omega}_{S,1}) \quad S_N^N(\boldsymbol{\Omega}_{S,1})]^T \quad (15)$$

[0098] 该向量借助于HOA表示将方向束描述为信号源方向$\Omega_{S,1}$。在基于向量的信号的情况下,向量v_1不限于关于任何方向的模向量,因此可以描述基于向量的单声道信号的更一般的方向分布。

[0099] 下面考虑D个主要声音信号$x_d(t)$,d=1,...,D的一般情况,D个主要声音信号可以根据下式被集中在向量x(t)中

[0100] $x(t)=[x_1(t) \ x_2(t)...x_D(t)]^T$ (16)

[0101] 这些信号必须基于以下矩阵来确定:

[0102] $V:=[v_1 \ v_2...v_D]$ (17)

[0103] 该矩阵由表示单声道主要声音信号$x_d(t)$,d=1,...,D的方向分布的所有向量v_d,d=1,...,D构成。

[0104] 为了主要声音信号x(t)的有意义的提取,规定以下约束:

[0105] a)每个主要声音信号是作为原始HOA表示的系数序列的线性组合而获得的,即

[0106] $x(t)=A \cdot c(t)$, (18)

[0107] 其中,$\boldsymbol{A} \in \mathbb{R}^{D \times O}$表示混合矩阵。

[0108] b)混合矩阵A应当被选择成使得其欧几里德范数不超过值"1",即,

[0109] $$\|\boldsymbol{A}\|_2 \overset{!}{\leq} 1 \quad (19)$$

[0110] 并且使得原始HOA表示与主要声音信号的HOA表示之间的残差的欧几里德范数的平方(或幂)不大于原始HOA表示的欧几里德范数的平方(或幂),即

[0111] $$\|\boldsymbol{c}(t) - \boldsymbol{V} \cdot \boldsymbol{x}(t)\|_2^2 \overset{!}{\leq} \|\boldsymbol{c}(t)\|_2^2 \quad (20)$$

[0112] 通过将公式(18)代入公式(20)中,可以看出公式(20)与以下约束相当:

[0113] $$\|\boldsymbol{I} - \boldsymbol{V} \cdot \boldsymbol{A}\|_2 \overset{!}{\leq} 1 \quad (21)$$

[0114] 其中,I表示单位矩阵。

[0115] 使用公式(18)、公式(19)和公式(11)根据公式(18)和公式(19)中的约束以及根据欧几里德矩阵与向量范数的兼容性,通过下式来限定主要声音信号的幅度上限:

$$\|\boldsymbol{x}(lT_S)\|_\infty \le \|\boldsymbol{x}(lT_S)\|_2 \quad (22)$$

[0116] $$\le \|\boldsymbol{A}\|_2 \|\boldsymbol{c}(lT_S)\|_2 \quad (23)$$

$$\le \sqrt{K} \cdot O \ , \quad (24)$$

[0117] 因此,确保主要声音信号保持在与原始HOA系数序列相同的范围内(与公式(11)比较),即,$\boxed{\|\boldsymbol{x}(lT_S)\|_\infty \le \sqrt{K} \cdot O}$。 (25)

[0118] 选择混合矩阵的示例

[0119] 如何确定满足约束(20)的混合矩阵的示例是通过计算主要声音信号使得提取之后的残差的欧几里德范数最小来获得的,即,

[0120] $x(t)=\mathrm{argmin}_{x(t)}||V \cdot x(t)-c(t)||_2$。 (26)

[0121] 公式(26)中的最小化问题的解决方案由下式给出:

[0122] $x(t)=V^{+}c(t)$, (27)

[0123] 其中,$(\cdot)^{+}$表示摩尔-彭罗斯(Moore-Penrose)广义逆。通过将公式(27)与公式(18)进行比较,由此得出,在这种情况下,混合矩阵等于矩阵V的摩尔-彭罗斯广义逆,即$A=V^{+}$。

[0124] 然而,仍然必须选择矩阵V以满足约束(19),即,$\boxed{\|\boldsymbol{V}^{+}\|_2 \overset{!}{\le} 1}$。(28)

[0125] 在仅方向信号的情况下,其中,矩阵V是关于一些源信号方向$\Omega_{S,d,}=1,\ldots,D$的模矩阵,即

[0126] $V=[S(\Omega_{S,1}\ S(\Omega_{S,2})\ldots S(\Omega_{S,D})]$, (29)

[0127] 可以通过选择源信号方向$\Omega S,d,d=1,\ldots,D$使得任何两个相邻方向的距离不是太小来满足约束(28)。

[0128] 环境HOA分量的系数序列的值范围结果

[0129] 环境HOA分量是通过从原始HOA表示中减去主要声音信号的HOA表示来计算的,即,$c_{AMB}(t)=c(t)-V \cdot x(t)$。 (30)

[0130] 如果主要声音信号x(t)的向量是根据标准(20)确定的,则可以得出结论:

$$\|\boldsymbol{c}_{\mathrm{AMB}}(lT_S)\|_\infty \le \|\boldsymbol{c}_{\mathrm{AMB}}(lT_S)\|_2 \quad (31)$$

[0131]

$$\overset{(30)}{=} \|\boldsymbol{c}(lT_S) - \boldsymbol{V} \cdot \boldsymbol{x}(lT_S)\|_2 \quad (32)$$

$$\overset{(20)}{\le} \|\boldsymbol{c}(lT_S)\|_2 \quad (33)$$

[0132]

$$\overset{(11)}{=} \sqrt{K} \cdot O \ 。 \quad (34)$$

[0133] 环境HOA分量的空间变换系数序列的值范围

[0134] EP 2743922 A1和上面提到的MPEG文献N14264中提出的HOA压缩处理的另一方面是:环境HOA分量的第一O_{MIN}系数序列总是被选成被分配给传输通道,其中,$O_{MIN}(N_{MIN}+1)^2$,$N_{MIN} \le N$通常是比原始HOA表示的阶数更小的阶数。为了对这些HOA系数序列去相关,可以将它们变换为从一些预定义方向$\Omega_{MIN,d}$,$d=1,\ldots,O_{MIN}$(类似于输入HOA表示的归一化小节中

描述的概念)冲击的虚拟扬声器信号。

[0135] 用$c_{AMB,MIN}(t)$来定义阶数索引为$n \leq N_{MIN}$的环境HOA分量的所有系数序列的向量并且用Ψ_{MIN}来定义关于虚拟方向$\Omega_{MIN,d}, d=1,\ldots,0_{MIN}$的模矩阵,所有虚拟扬声器信号的向量(被定义为)$w_{MIN}(t)$通过下式获得:

[0136] $$\boldsymbol{w}_{\mathrm{MIN}}(t) = \boldsymbol{\Psi}_{\mathrm{MIN}}^{-1} \cdot \boldsymbol{c}_{\mathrm{AMB,MIN}}(t) \text{。} \quad (35)$$

[0137] 因此,使用欧几里德矩阵与向量范数的兼容性,

$$\|\boldsymbol{w}_{\mathrm{MIN}}(lT_S)\|_\infty \leq \|\boldsymbol{w}_{\mathrm{MIN}}(lT_S)\|_2 \quad (36)$$

[0138]

$$\overset{(35)}{\leq} \|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2 \cdot \left\|\boldsymbol{c}_{\mathrm{AMB,MIN}}(lT_S)\right\|_2 \quad (37)$$

$$\overset{(34)}{\leq} \|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2 \cdot \sqrt{K} \cdot O \text{。} \quad (38)$$

[0139] 在上面提到的MPEG文献N14264中,根据上面提到的Fliege等人的文章来选择虚拟方向$\Omega_{MIN,d}, d=1,\ldots,0_{MIN}$。图4示出了模矩阵$\Psi_{MIN}$的逆矩阵针对阶数($N_{MIN}=1,\ldots,9$)的相应欧几里德范数。可以看出:对于$N_{MIN}=1,\ldots,9$,$\|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2 < 1$。 (39)

[0140] 然而,这通常不适用于$\|\boldsymbol{\Psi}_{\mathrm{MIN}}^{-1}\|_2$的值通常远大于"1"的$N_{MIN}>9$的情况。然而,至少对于$1 \leq N_{MIN} \leq 9$,虚拟扬声器信号的幅度由下式限制:

[0141] $$\boxed{\|\boldsymbol{w}_{\mathrm{MIN}}(lT_S)\|_\infty \overset{(38),\mathrm{Fig.4}}{\leq} \sqrt{K} \cdot O \quad \text{for } 1 \leq N_{MIN} \leq 9} \text{。} \quad (40)$$

[0142] 通过限制输入HOA表示以满足条件(6),其中条件(6)要求根据该HOA表示创建的虚拟扬声器信号的幅度不超过值"1",可以保证在以下条件下,信号在增益控制之前的幅度将不超过值$\sqrt{K} \cdot O$(参见公式(25)、公式(34)和公式(40)):

[0143] a)所有主要声音信号x(t)的向量是根据公式/限制(18)、(19)和(20)来计算的;

[0144] b)如果使用如上述Fliege等人的文章中限定的虚拟扬声器位置,则确定对其实施空间变换的环境HOA分量的第一系数序列的数量0_{MIN}的最小阶数N_{MIN}必须小于"9"。

[0145] 可以进一步得出结论:对于高达感兴趣的最大阶数N_{MAX}的任何阶数N,即$1 \leq N \leq N_{MAX}$,信号在增益控制之前的幅度将不超过值$\sqrt{K_{\mathrm{MAX}}} \cdot O$,其中,

[0146] $$K_{\mathrm{MAX}} = \max_{1 \leq N \leq N_{\mathrm{MAX}}} K\left(N, \boldsymbol{\Omega}_1^{(N)}, \ldots, \boldsymbol{\Omega}_O^{(N)}\right) \text{。} \quad (41a)$$

[0147] 特别地,从图3可以得出结论:如果假设用于初始空间变换的虚拟扬声器方向$\boldsymbol{\Omega}_j^{(N)}, 1 \leq j \leq O$是根据Fliege等人的文章中的分布来选择的,并且如果另外假设感兴趣的最大阶数为$N_{MAX}=29$(例如参见MPEG文献N14264),则信号增益控制之前的幅度将不超过值1.50,这是因为在这种特殊情况下$\sqrt{K_{\mathrm{MAX}}} < 1.5$,即,可以选择$\sqrt{K_{\mathrm{MAX}}} = 1.5$。

[0148] K_{MAX}取决于感兴趣的最大阶数N_{MAX}和虚拟扬声器方向$\boldsymbol{\Omega}_j^{(N)}, 1 \leq j \leq O$,其可以由下

式来表示：

[0149] $$K_{\mathrm{MAX}} = K_{\mathrm{MAX}}\left(\left\{\boldsymbol{\Omega}_1^{(N)},\ldots,\boldsymbol{\Omega}_O^{(N)} \mid 1 \le N \le N_{\mathrm{MAX}}\right\}\right)。 \quad (41b)$$

[0150] 因此，为确保感知编码之前的信号位于区间[-1,1]内而通过增益控制施加的最小增益由$2^{e_{\mathrm{MIN}}}$给出，其中，

[0151] $$e_{\mathrm{MIN}} = -\left\lceil \log_2\left(\sqrt{K_{\mathrm{MAX}}}\cdot O\right)\right\rceil < 0。 \quad (41c)$$

[0152] 在信号在增益控制之前的幅度太小的情况下，在MPEG文献N14264中提出可以用高达$2^{e_{\mathrm{MAX}}}$的因子来平滑地放大它们，其中，$e_{MAX}\geq 0$作为编码HOA表示中的边信息被传输。

[0153] 因此，在访问单元内描述由增益控制处理单元引起的从第一帧一直到当前帧的已修改信号的总绝对幅度变化的底数为“2”的每个指数，可以假设在区间$[e_{MIN},e_{MAX}]$内的任何整数值。因此，编码所需的(最小整数)比特数β_e由下式给出：

[0154] $$\beta_e = \left\lceil \log_2\left(|e_{\mathrm{MIN}}| + e_{\mathrm{MAX}} + 1\right)\right\rceil = \left\lceil \log_2\left(\left\lceil \log_2\left(\sqrt{K_{\mathrm{MAX}}}\cdot O\right)\right\rceil + e_{\mathrm{MAX}} + 1\right)\right\rceil。 \quad (42)$$

[0155] 在信号在增益控制之前的幅度不太小的情况下，公式(42)可以简化为：

[0156] $$\beta_e = \left\lceil \log_2\left(|e_{\mathrm{MIN}}| + 1\right)\right\rceil = \left\lceil \log_2\left(\left\lceil \log_2\left(\sqrt{K_{\mathrm{MAX}}}\cdot O\right)\right\rceil + 1\right)\right\rceil。 \quad (42a)$$

[0157] 可以在增益控制处理步骤/阶段15,…,151的输入处计算该比特数β_e。

[0158] 针对指数使用该比特数β_e确保可以捕获由HOA压缩器增益控制处理单元引起的所有可能的绝对幅度变化，从而允许在压缩表示中的一些预定义入口点处开始解压缩。

[0159] 当在HOA解压缩器中开始对压缩HOA表示进行解压缩时，被分配给一些数据帧的边信息并且除了所接收的数据流$\check{\boldsymbol{B}}$之外从解复用器21接收到的、表示总绝对幅度变化的非差分增益值被用在逆增益控制步骤或阶段24,…,241中，从而以与在增益控制处理步骤/阶段15,…,151中执行的处理相反的方式实施正确的增益控制。

[0160] 另外的实施方式

[0161] 当实现如在章节HOA压缩、空间HOA编码、HOA解压缩和空间HOA解码中描述的特定HOA压缩/解压缩系统时，用于对指数编码的比特数β_e必须依赖于缩放因子$K_{MAX,DES}$根据公式(42)来设定，缩放因子$K_{MAX,DES}$自身取决于要压缩的HOA表示的期望的最大阶数$N_{MAX,DES}$和特定的虚拟扬声器方向$\boldsymbol{\Omega}_{\mathrm{DES},1}^{(N)},\ldots,\boldsymbol{\Omega}_{\mathrm{DES},O}^{(N)},\ 1 \le N \le N_{\mathrm{MAX}}$。

[0162] 例如，当假设$N_{MAX,DES}=29$并且根据Fliege等人的文章来选择虚拟扬声器方向时，合理的选择是$\sqrt{K_{\mathrm{MAX,DES}}} = 1.5$。在该情况下，保证对阶数为$N(1\leq N\leq N_{MAX})$的HOA表示进行正确的压缩，该HOA表示是利用相同的虚拟扬声器方向$\boldsymbol{\Omega}_{\mathrm{DES},1}^{(N)},\ldots,\boldsymbol{\Omega}_{\mathrm{DES},O}^{(N)}$、根据章节输入HOA表示的归一化而被归一化的。然而，在以下HOA表示的情况下不能给出这种保证：该HOA表示还(出于效率原因)由PCM格式的虚拟扬声器信号等效地表示，但是其中虚拟扬声器的方向$\boldsymbol{\Omega}_j^{(N)}, 1 \le j \le O$被选择成与在系统设计阶段假定的虚拟扬声器方向$\boldsymbol{\Omega}_{\mathrm{DES},1}^{(N)},\ldots,\boldsymbol{\Omega}_{\mathrm{DES},O}^{(N)}$不同。

[0163] 由于虚拟扬声器位置的这种不同选择，即使这些虚拟扬声器信号的幅度在区间[-

1,1]内,也不能再保证信号在增益控制之前的幅度将不超过值$\sqrt{K_{\mathrm{MAX,DES}}}\cdot O$。因此,不能保证该HOA表示根据MPEG文献N14264中描述的处理而具有针对压缩的适当的归一化。

[0164] 在这种情况下,具有以下系统是有利的:该系统基于虚拟扬声器位置的知识而提供虚拟扬声器信号的最大允许幅度以确保相应的HOA表示适合于根据在MPEG文献N14264中描述的处理的压缩。在图5中示出了这样的系统。它采用虚拟扬声器位置$\boldsymbol{\Omega}_j^{(N)},\ 1\le j\le O$作为输入,其中,$O=(N+1)^2\ (N\in\mathbb{N}_0)$,并且提供虚拟扬声器信号的最大允许幅度$\gamma_{\mathrm{dB}}$(其以分贝测量)作为输出。在步骤或阶段51中,根据公式(3)来计算关于虚拟扬声器位置的模矩阵Ψ。在随后的步骤或阶段52中,计算模矩阵的欧几里德范数$\|\Psi\|_2$。在第三步骤或阶段53中,将幅度γ计算为"1"与下述值中的最小值:该值为虚拟扬声器位置数量的平方根和$K_{\mathrm{MAX,DES}}$的平方根的乘积与模矩阵的欧几里德范数的商,

[0165] 即

$$\gamma=\min\left(1,\frac{\sqrt{O}\cdot\sqrt{K_{\mathrm{MAX,DES}}}}{\|\boldsymbol{\Psi}\|_2}\right)。\qquad(43)$$

[0166] 以分贝为单位的值通过下式获得:$\gamma_{\mathrm{dB}}=20\log_{10}(\gamma)$。 (44)

[0167] 为了说明:从上面的推导可以看出,如果HOA系数序列的幅度不超过值$\sqrt{K_{\mathrm{MAX,DES}}}\cdot O$,即,如果

[0168] $$\|\boldsymbol{c}(lT_S)\|_\infty\le\sqrt{K_{\mathrm{MAX,DES}}}\cdot O,\qquad(45)$$

[0169] 则在增益控制处理单元之前的所有信号将相应地不超过该值,这是对适当的HOA压缩的要求。

[0170] 从公式(9)发现HOA系数序列的幅度由下式限制

[0171] $||c(lT_S)||_\infty\le||c(lT_S)||_2\le||\Psi||_2\cdot||W(lT_S)||_2$。 (46)

[0172] 因此,如果γ是根据公式(43)设置的并且PCM格式的虚拟扬声器信号满足

[0173] $||w(lT_S)||_\infty\le\gamma$, (47)

[0174] 则从公式(7)得出

$$\|\boldsymbol{w}(lT_S)\|_2\le\gamma\cdot\sqrt{O}\qquad(48)$$

[0175] 并且满足要求(45)。

[0176] 即,公式(6)中的最大幅度值"1"由公式(47)中的最大幅度值γ代替。

[0177] 高阶高保真度立体声响复制的基础

[0178] 高阶高保真度立体声响复制(HOA)基于对感兴趣的密集区域内的声场的描述,其被假设为没有声源。在该情况下,在感兴趣的区域内的时间t和位置x处的声压p(t,x)的时空行为在物理上完全由齐次波动方程确定。在下文中,假设如图6所示的球坐标系。在所使用的坐标系中,x轴指向前面,y轴指向左侧,z轴指向顶部。空间中的位置$x=(r,\theta,\varphi)^T$由半径$r>0$(即,到坐标原点的距离)、从极轴z测量的倾斜角$\theta\in[0,\pi]$以及在x-y平面内从x轴逆时针方向测量的方位角$\varphi\in[0,2\pi[$来表示。此外,$(\cdot)^T$表示转置。

[0179] 然后,从"傅里叶声学"教科书中可以看出,声压关于时间的傅立叶变换由$\mathcal{F}_t(\cdot)$表

示,即,

[0180] $$P(\omega,x)=\mathcal{F}_t\big(p(t,x)\big)=\int_{-\infty}^{\infty}p(t,x)e^{-\mathrm{i}\omega t}dt,\qquad(49)$$

[0181] 其中,ω表示角频率,i表示虚数单位,根据下式可以将上述声压关于时间的傅里叶变换展开成球面谐波函数的级数

[0182] $$P(\omega=kc_s,r,\theta,\phi)=\sum_{n=0}^{N}\sum_{m=-n}^{n}A_n^m(k)j_n(kr)S_n^m(\theta,\phi),\qquad(50)$$

[0183] 其中,c_s表示声速,k表示角波数,其通过$k=\frac{\omega}{c_s}$而与角频率ω相关。此外,$j_n(\bullet)$表示第一类球贝塞尔函数,且$S_n^m(\theta,\phi)$表示阶数为n且度数为m的实值球面谐波函数,在章节实值球面谐波函数的定义中对它们做出了定义。展开系数$A_n^m(k)$仅取决于角波数k。注意,已经隐含地假设声压在空间上是频带受限的。因此,在称为HOA表示的阶数的上限N处关于阶数索引n截断该级数。

[0184] 如果声场是通过从由角度元组(θ,φ)指定的所有可能方向到达的无限个具有不同角频率ω的谐波平面波进行叠加来表示的,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,卷4(116),第2149至2157页,2004年10月),相应的平面波复幅度函数C(ω,θ,φ)可以由以下球面谐波函数展开式来表示

[0185] $$C(\omega=kc_s,\theta,\phi)=\sum_{n=0}^{N}\sum_{m=-n}^{n}C_n^m(k)S_n^m(\theta,\phi),\qquad(51)$$

[0186] 其中,展开系数$C_n^m(k)$通过下式与展开系数$A_n^m(k)$相关:

[0187] $$A_n^m(k)=\mathrm{i}^nC_n^m(k)\qquad(52)$$

[0188] 假设各个系数$C_n^m(k=\omega/c_s)$是角频率ω的函数,则逆傅里叶变换(由$\mathcal{F}^{-1}(\cdot)$表示)的应用针对每个阶数n和度数m提供以下时域函数

[0189] $$c_n^m(t)=\mathcal{F}_t^{-1}\big(C_n^m(\omega/c_s)\big)=\frac{1}{2\pi}\int_{-\infty}^{\infty}C_n^m\left(\frac{\omega}{c_s}\right)e^{\mathrm{i}\omega t}d\omega\qquad(53)$$

[0190] 这些时域函数在这里称为连续时间HOA系数序列,其可以通过下式被集中在单个向量c(t)中

$$\boldsymbol{c}(t)=\qquad(54)$$

[0191] $$\left[c_0^0(t)\quad c_1^{-1}(t)\quad c_1^0(t)\quad c_1^1(t)\quad c_2^{-2}(t)\quad c_2^{-1}(t)\quad c_2^0(t)\quad c_2^1(t)\quad c_2^2(t)\quad \dots\ c_N^{N-1}(t)\quad c_N^N(t)\right]^T$$

[0192] 向量c(t)内的HOA系数序列$c_n^m(t)$的位置索引由n(n+1)+1+m给出。向量c(t)中的总元素数由O=(N+1)^{2}给出。

[0193] 最终的高保真度立体声响复制格式利用采样频率f_S提供c(t)的如下采样版本

[0194] $$\{\boldsymbol{c}(lT_S)\}_{l\in\mathbb{N}}=\{\boldsymbol{c}(T_S),\boldsymbol{c}(2T_S),\boldsymbol{c}(3T_S),\boldsymbol{c}(4T_S),\dots\}\qquad(55)$$

[0195] 其中,$T_S=1/f_S$表示采样周期。元素$c(lT_S)$称为离散时间HOA系数序列,其可以总是实值。该特性也适用于连续时间版本$c_n^m(t)$。实值球面谐波函数的定义

[0196] 实值球面谐波函数$S_n^m(\theta,\phi)$(假设根据以下文献的SN3D归一化:J.Daniel,“Représentation de champs acoustiques,applicationàla transmission etàla reproduction de scènes sonores complexes dans un contexte multimédia”,博士论文,巴黎大学,2001年6月,3.1章)由下式给出

[0197] $$S_n^m(\theta,\phi)=\sqrt{(2n+1)\frac{(n-|m|)!}{(n+|m|)!}}\,P_{n,|m|}(cos\theta)\,\mathrm{trg}_m(\phi) \quad (56)$$

[0198] 其中,

[0199] $$\mathrm{trg}_m(\phi)=\begin{cases}\sqrt{2}cos(m\phi) & m>0\\ 1 & m=0\\ -\sqrt{2}sin(m\phi) & m<0\end{cases}。 \quad (57)$$

[0200] 相关联的勒让德函数$P_{n,m}(x)$被定义为

[0201] $$P_{n,m}(x)=(1-x^2)^{m/2}\frac{d^m}{dx^m}P_n(x),m\geq 0 \quad (58)$$

[0202] 其具有勒让德多项式$P_n(x)$,并且与Academic Press1999年出版的Applied Mathematical Sciences第93卷E.G.Williams的“Fourier Acoustics”中的不同,其没有Condon-Shortley相位项$(-1)^m$。

[0203] 本发明的处理可以由单个处理器或电子电路,或者由并行工作和/或在本发明的处理的不同部分中工作的若干处理器或电子电路执行。

[0204] 用于操作一个或多个处理器的指令可以被存储在一个或更多个存储器中。

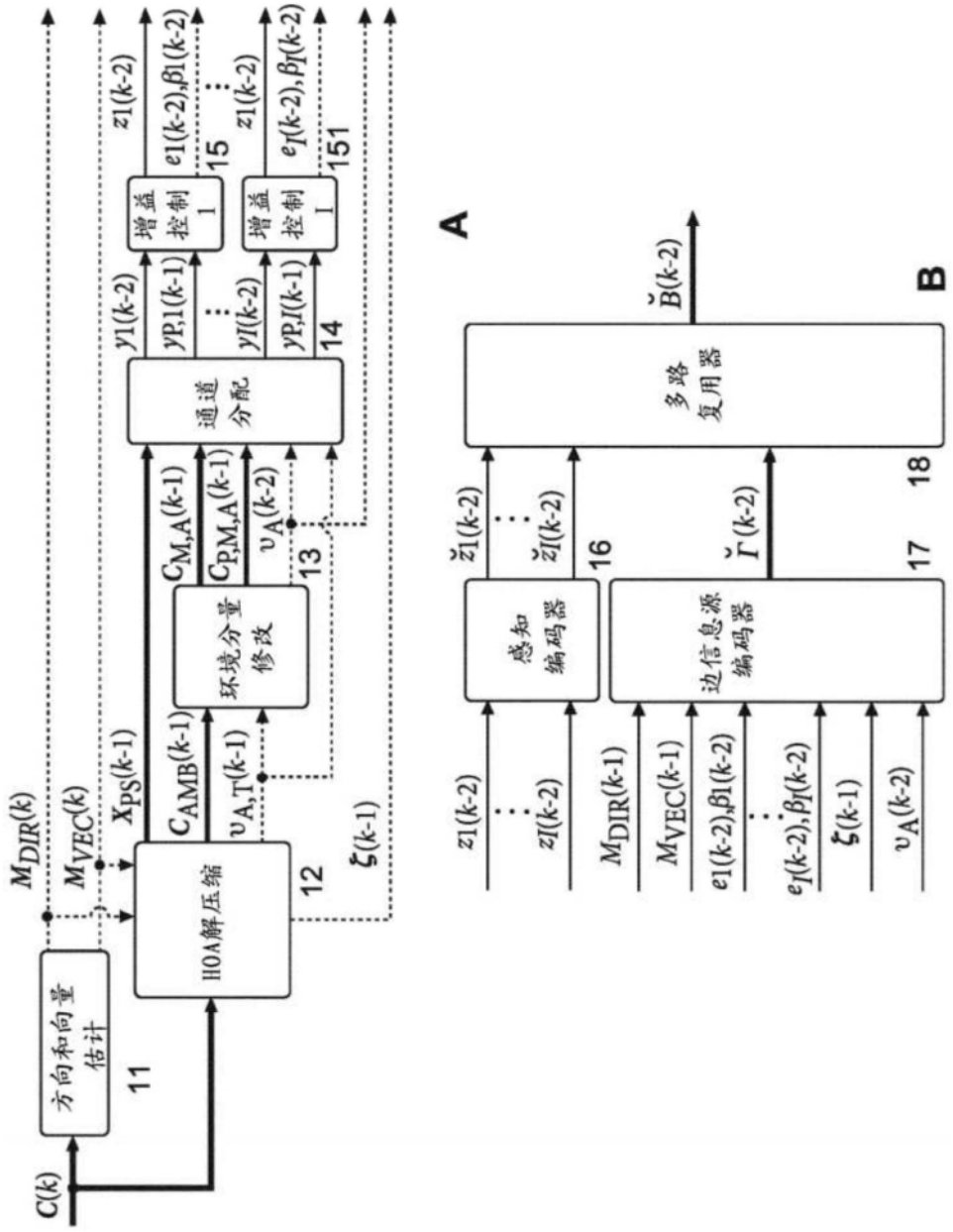

C(k)
方向和向量估计
11
HOA解压缩
12
环境分量修改
13
通道分配
14
增益控制 1
15
增益控制 1
151
感知编码器
16
边信息源编码器
17
多路复用器
18
A
B
$M_{DIR}(k)$
$M_{VEC}(k)$
$x_{PS}(k-1)$
$C_{AMB}(k-1)$
$v_{A,T}(k-1)$
$C_{M,A}(k-1)$
$C_{P,M,A}(k-1)$
$v_A(k-2)$
$\zeta(k-1)$
$y_1(k-2)$
$y_{P,1}(k-1)$
$y_I(k-2)$
$y_{P,I}(k-1)$
$z_1(k-2)$
$e_1(k-2),\beta_1(k-2)$
$e_I(k-2),\beta_I(k-2)$
$\check{z}_1(k-2)$
$\check{z}_I(k-2)$
$\check{B}(k-2)$
$\check{\Gamma}(k-2)$
$M_{DIR}(k-1)$
$M_{VEC}(k-1)$

图1

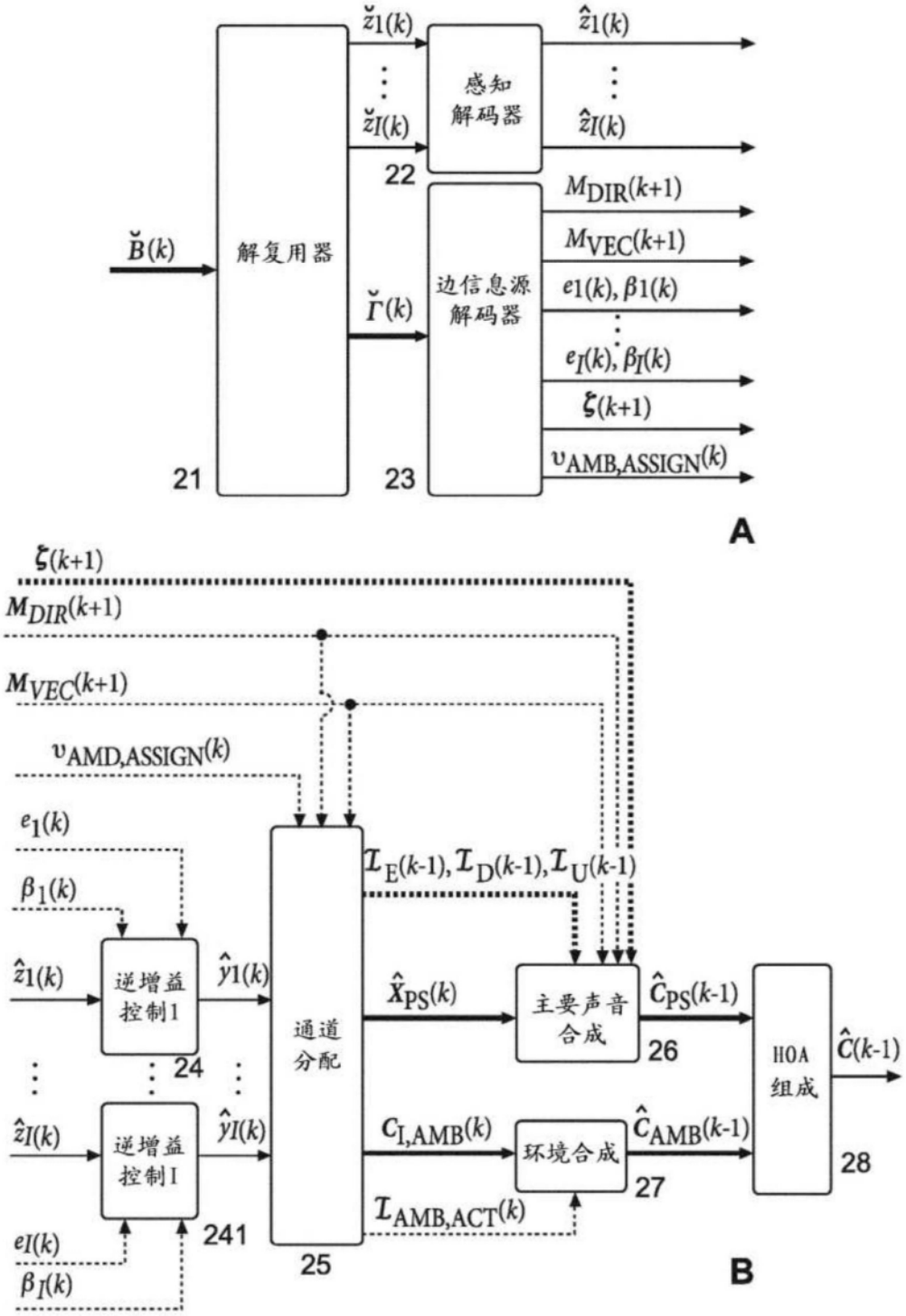


图2

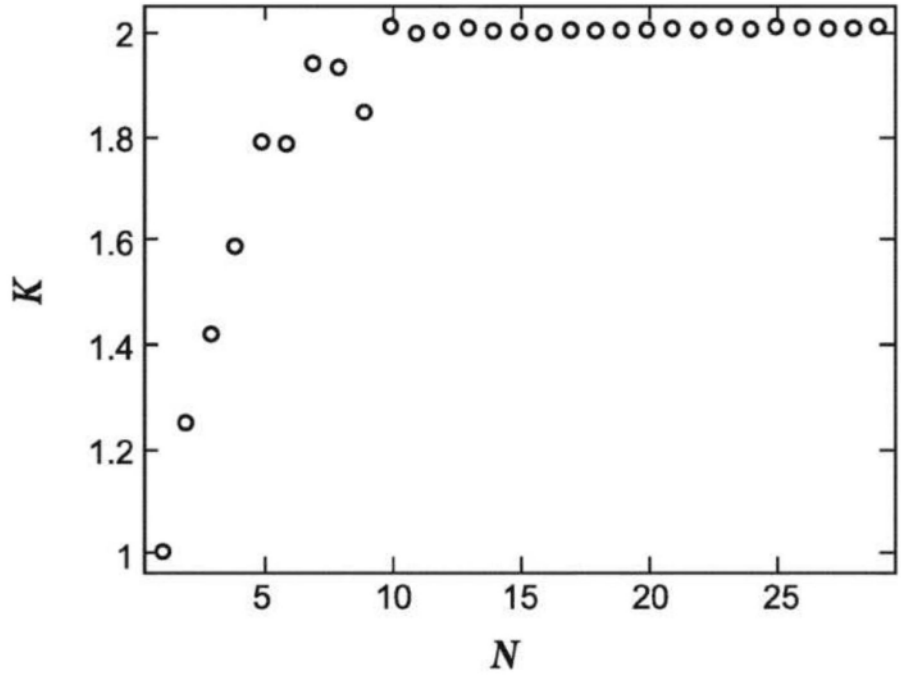


图3

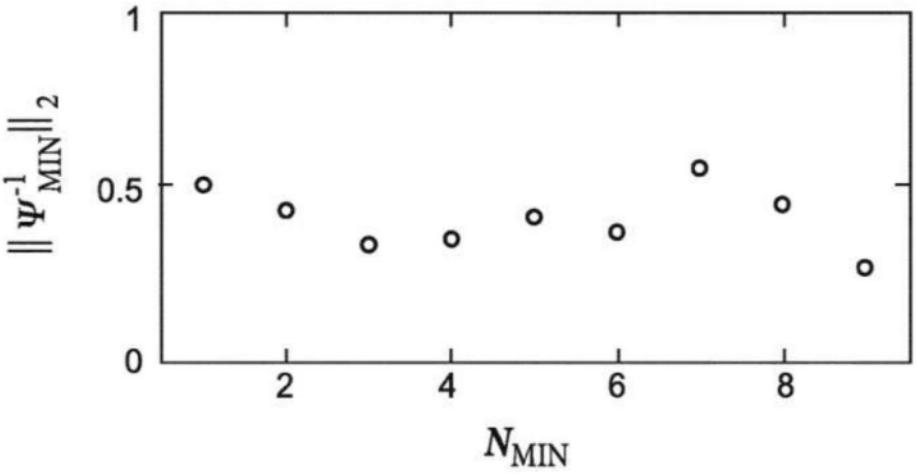


图4

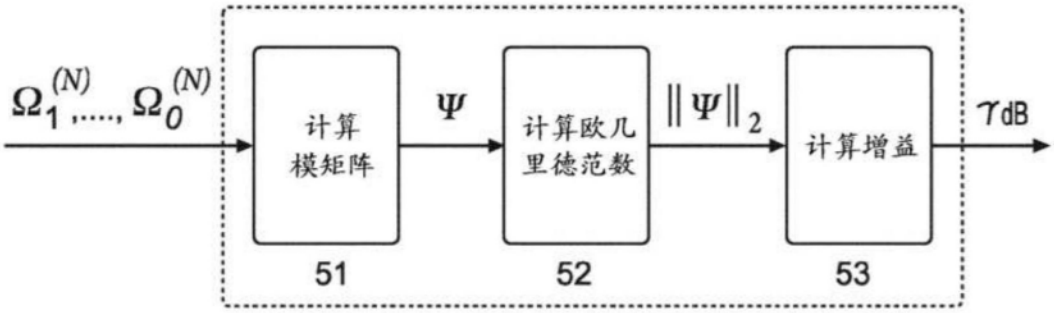

$\Omega_1^{(N)},\ldots,\Omega_0^{(N)}$
计算
模矩阵
51
Ψ
计算欧几
里德范数
52
$\|\Psi\|_2$
计算增益
53
τ_{dB}

图5

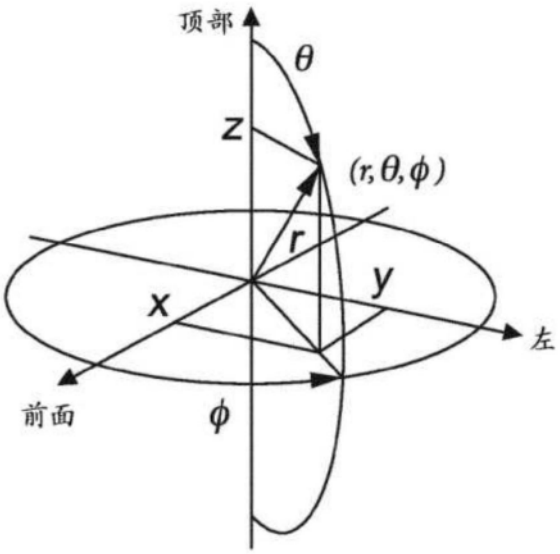

顶部
θ
z
(r,θ,ϕ)
r
y
x
左
前面
ϕ

图6