



(12) 发明专利

(10) 授权公告号 CN 107170458 B

(45) 授权公告日 2021.01.12

(21) 申请号 201710350455.X

(74) 专利代理机构 中国贸促会专利商标事务所

(22) 申请日 2013.05.06

有限公司 11038

(65) 同一申请的已公布的文献号

代理人 宿小猛

申请公布号 CN 107170458 A

(51) Int.CI.

(43) 申请公布日 2017.09.15

G10L 19/008 (2013.01)

(30) 优先权数据

H04S 3/00 (2006.01)

12305537.8 2012.05.14 EP

(56) 对比文件

(62) 分案原申请数据

EP 2450880 A1, 2012.05.09

201380025029.9 2013.05.06

WO 2011117399 A1, 2011.09.29

(73) 专利权人 杜比国际公司

CN 101952888 A, 2011.01.19

地址 荷兰阿姆斯特丹

审查员 涂亮梅

(72) 发明人 A.克鲁格 S.科唐 J.贝姆

权利要求书2页 说明书22页 附图2页

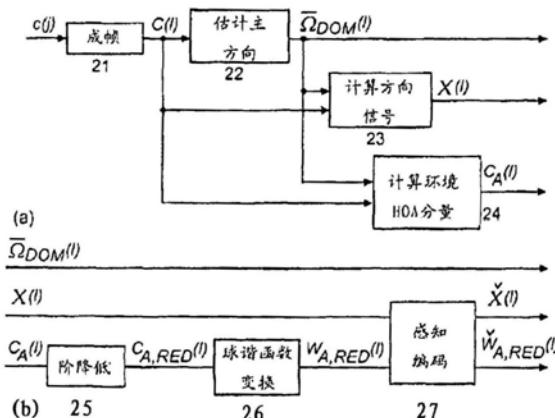
J-M.巴特克

(54) 发明名称

压缩和解压缩高阶高保真度立体声响复制
信号表示的方法及装置

(57) 摘要

本公开涉及压缩和解压缩高阶高保真度立体声响复制信号表示的方法及装置。高阶高保真度立体声响复制(HOA)表示在最佳听音位置附近的完整声场，其不依赖于扩音器结构。高空间分辨率需要大量的HOA系数。在本发明中，估计主声音方向，并且将HOA信号表示分解成时域中的主方向信号和相关的方向信息以及HOA域中的环境分量，继之以通过降低它的阶来压缩环境分量。阶降低后的环境分量被变换到空间域，并且与方向信号一起被感知编码。在接收器侧，编码后的方向信号和阶降低后经编码的环境分量被感知地解压缩，被感知地解压缩的环境信号被变换到降低的阶的HOA域表示，继之以阶扩展。从方向信号、对应的方向信息以及原始阶的环境HOA分量重新组成总的HOA表示。

B
CN 107170458

CN

1. 一种用于解压缩高阶高保真度立体声响复制HOA信号表示的方法,所述方法包括:
接收编码的方向信号和编码的环境信号;
对编码的方向信号和编码的环境信号感知解码以分别产生解码的方向信号和解码的环境信号;
将解码的环境信号从空间域转换到环境信号的HOA域表示;
扩展环境信号的HOA域表示的阶次;以及
从阶次扩展的环境信号的HOA域表示和解码的方向信号重新组成高阶高保真度立体声响复制HOA信号;
其中,该转换包括将逆空间变换应用于解码的环境信号。
2. 根据权利要求1所述的方法,其中,高阶高保真度立体声响复制HOA信号表示具有大于1的阶次。
3. 根据权利要求2所述的方法,其中,解码的环境信号的阶次小于高阶高保真度立体声响复制HOA信号表示的阶次。
4. 根据权利要求1所述的方法,其中,编码的方向信号和编码的环境信号在比特流中被接收到,并且所述比特流被感知解码到多个传输信道中,在该转换和重新组成之前所述多个传输信道中的每个传输信道被重新分配给方向信号或环境信号。
5. 一种用于解压缩高阶高保真度立体声响复制HOA信号表示的设备,所述设备包括:
输入接口,其接收编码的方向信号和编码的环境信号;
音频解码器,其对编码的方向信号和编码的环境信号感知解码以分别产生解码的方向信号和解码的环境信号;
逆变换器,其将解码的环境信号从空间域转换到环境信号的HOA域表示;
阶次扩展器,其扩展环境信号的HOA域表示的阶次;以及
合成器,其从阶次扩展的环境信号的HOA域表示和解码的方向信号重新组成高阶高保真度立体声响复制HOA信号;
其中,所述逆变换器配置为通过将逆空间变换应用于解码的环境信号来进行转换。
6. 根据权利要求5所述的设备,其中,高阶高保真度立体声响复制HOA信号表示具有大于1的阶次。
7. 根据权利要求6所述的设备,其中,解码的环境信号的阶次小于高阶高保真度立体声响复制HOA信号表示的阶次。
8. 根据权利要求5所述的设备,其中,编码的方向信号和编码的环境信号在比特流中被接收到,并且所述比特流被感知解码到多个传输信道中,在该转换和重新组成之前所述多个传输信道中的每个传输信道被重新分配给方向信号或环境信号。
9. 一种用于解压缩高阶高保真度立体声响复制HOA信号表示的方法,所述方法包括:
接收编码的方向信号和编码的环境信号;
对编码的方向信号和编码的环境信号感知解码以分别产生解码的方向信号和解码的环境信号;
将解码的环境信号从空间域转换到环境信号的HOA域表示;
扩展环境信号的HOA域表示的阶次;
从阶次扩展的环境信号的HOA域表示和解码的方向信号重新组成高阶高保真度立体声

响复制HOA信号；以及

平滑化该重新组成的HOA信号，其中，该平滑化是基于窗口函数的。

10. 一种用于解压缩高阶高保真度立体声响复制HOA信号表示的设备，所述设备包括：

输入接口，其接收编码的方向信号和编码的环境信号；

音频解码器，其对编码的方向信号和编码的环境信号感知解码以分别产生解码的方向信号和解码的环境信号；

逆变换器，其将解码的环境信号从空间域转换到环境信号的HOA域表示；

阶次扩展器，其扩展环境信号的HOA域表示的阶次；

合成器，其从阶次扩展的环境信号的HOA域表示和解码的方向信号重新组成高阶高保真度立体声响复制HOA信号；以及

平滑器，其用于平滑化该重新组成的HOA信号，其中，该平滑化是基于窗口函数的。

11. 一种非暂时性计算机可读介质，包含指令，所述指令在由处理器执行时使得执行根据权利要求1-4和9中任一项所述的方法。

12. 一种用于解压缩高阶高保真度立体声响复制HOA信号表示的设备，包括：

一个或多个处理器，以及

一个或多个存储介质，存储有指令，所述指令在被所述一个或多个处理器执行时使得执行根据权利要求1-4和9中任一项所述的方法。

13. 一种用于解压缩高阶高保真度立体声响复制HOA信号表示的装置，包括用于执行根据权利要求1-4和9中任一项所述的方法的部件。

压缩和解压缩高阶高保真度立体声响复制信号表示的方法及装置

[0001] 本申请是申请号为201380025029.9、申请日为2013年5月6日、发明名称为“压缩和解压缩高阶高保真度立体声响复制信号表示的方法及装置”的发明专利申请的分案申请。

技术领域

[0002] 本发明涉及一种压缩和解压缩高阶高保真度立体声响复制(HigherOrder Ambisonics)信号表示的方法及装置,其中以不同的方式处理方向和环境(ambient)分量。

背景技术

[0003] 高阶高保真度立体声响复制(HOA)提供了以下优点:捕获三维空间中的特定位置附近的完整声场,该位置被称为“最佳听音位置(sweet spot)”。与像立体声或环绕声这样的基于信道的技术相反,这种HOA表示不依赖于具体的扩音器结构。但是,这种灵活性以在特定扩音器结构上回放该HOA表示所需的解码处理为代价。

[0004] HOA基于使用截断的球谐函数(SH)展开式的在期望的听者位置附近的位置 x 的单独的角形波数量 k 的气压的复数幅度的描述,其中,在不失一般性的情况下,可以将期望的听者位置假设为球坐标系统的原点。这种表示的空间分辨率随着该展开式的增长的最大阶 N 提高。不幸的是,展开式系数的数量 O 随阶 N 而平方地增长,亦即 $O = (N+1)^2$ 。例如,使用阶 $N=4$ 的典型的HOA表示需要 $O=25$ 个HOA系数。给出期望的采样率 f_s 和每个样本的比特数量 N_b ,传送HOA信号表示的总的比特率按照 $O \cdot f_s \cdot N_b$ 来确定,并且在针对每个样本采用 $N_b=16$ 个比特,采样率为 $f_s=48\text{kHz}$ 的情况下阶 $N=4$ 的HOA信号表示的传输导致 19.2MBits/s 的比特率。因此,压缩HOA信号表示是非常值得做的。

[0005] 关于现存空间音频压缩方法的概述可以在专利申请EP 10306472.1中或者在I.Elfitri、B.Günel、A.M.Kondoz的“Multichannel Audio Coding Based on Analysis by Synthesis”(Proceedings of the IEEE,第99卷,第4期,657-670页,2011年4月)中找到。

[0006] 下面的技术与本发明更相关。

[0007] 可以如V.Pulkki在“Spatial Sound Reproduction with Directional Audio Coding”(Journal of Audio Eng.Society,第55(6)卷,503-516页,2007年)中所述的使用方向音频编码(DirAC)来压缩B格式信号(等效于一阶高保真度立体声响复制表示)。在对电子会议应用提出的一个版本中,将B格式信号编码成单个全向信号以及以单一方向形式的边信息和针对每个频带的扩散参数。然而,作为结果的数据率的显著降低以在再现时得到的较小的信号质量为代价。另外,DirAC受限于一阶高保真度立体声响复制表示的压缩,其受到非常低的空间分辨率的影响。

[0008] 已知的用于压缩具有 $N > 1$ 的HOA表示的方法相当少。其中之一利用感知高级音频编码(AAC)编码解码器对单独的HOA系数序列进行直接编码,参见E.Hellerud、I.Burnett、A.Solvang、U.Peter Svensson的“Encoding Higher Order Ambisonics with AAC”(第124届AES大会,阿姆斯特丹,2008年)。然而,该方法的固有问题永远不会被听到的信号的感

知编码。通常通过HOA系数序列的加权和来获得重构的回放信号。这是为什么当在特定的扩音器结构上呈现解压缩后的HOA表示时未屏蔽感知编码噪声的概率很高的原因。以更技术性的术语,感知编码噪声未屏蔽的主要问题是单独的HOA系数序列之间的高度的互相关性。因为在单独的HOA系数序列中的编码后的噪声信号通常彼此不相关,所以可能出现感知编码噪声的结构重叠,同时与噪声无关的HOA系数序列在重叠处被消去。另一个问题是所提到的互相关性导致感知编码器的效率降低。

[0009] 为了将这些影响的程度最小化,在EP 10306472.1中提出在感知编码之前将HOA表示变换为空间域中的等效表示。空间域信号对应于常规的方向信号,并且如果扩音器被置于与对空间域变换假设的那些方向完全相同的方向上,则将对应于扩音器信号。

[0010] 到空间域的变换降低了单独的空间域信号之间的互相关性。然而,并未彻底消除互相关性。关于相对较高的互相关性的示例是其方向落入空间域信号所覆盖的相邻方向之间的方向信号。

[0011] EP 10306472.1和上述的Hellerud等人的论文的另一个不足是经感知编码的信号的数量是 $(N+1)^2$,其中,N是HOA表示的阶。因此,压缩后的HOA表示的数据率随高保真度立体声复制阶而平方地增长。

[0012] 本发明的压缩处理将HOA声场表示分解为方向分量和环境分量。具体对于计算方向声场分量,在下面描述了一种新的处理,用于估计若干主声音方向。

[0013] 关于基于高保真度立体声复制的方向估计的现存方法,上述的Pulkki的论文描述了一种结合DirAC编码的方法,用于基于B格式声场表示来估计方向。方向根据平均强度矢量获得,其指向声场能量流动的方向。在D.Levin、S.Gannot、E.A.P Habets的“Direction-of-Arrival Estimation using Acoustic Vector Sensors in the Presence of Noise”(IEEE Proc.Of the ICASSP,105-108页,2011年)中提出了一种基于B格式的替代。通过搜索对引入到那个方向的波束形成器输出信号提供最大能量的那个方向,迭代地进行方向估计。

[0014] 然而,对于方向估计,两种方法都受约束于B格式,其受到相对较低的空间分辨率的影响。另一不足之处是该估计被限制于仅仅单个主方向。

[0015] HOA表示提供了改善的空间分辨率,从而允许对若干主方向的改善的估计。现存的基于HOA声场表示对若干方向进行估计的方法相当稀少。在N.Epain、C.Jin、A.van Schaik的“The Application of Compressive Sampling to the Analysis and Synthesis of Spatial Sound Fields”(127th Convention of the AudioEng.Soc.,纽约,2009年)中以及在A.Wabnitz、N.Epain、A.van Schaik、C Jin的“Time Domain Reconstruction of Spatial Sound Fields Using Compressed Sensing”(IEEE Proc.of the ICASSP,465-468页,2011年)中提出了一种基于压缩传感的方法。主要想法是假设声场是空间稀疏的,亦即由仅仅少量的方向信号构成。在球上分配大量的测试方向之后,采用最优化算法以便发现尽可能少的测试方向以及对应的方向信号,使得它们被给出的HOA表示良好地描述。与实际上由给出的HOA表示提供的空间分辨率相比,该方法提供了一种改善的空间分辨率,因为其避开了从给出的HOA表示的有限阶导致的空间离差。然而,该算法的性能高度依赖于是否满足稀疏性假设。具体地,如果声场包括任何的较小的附加环境分量,或者如果HOA表示受到将在从多信道记录计算时出现的噪声的影响,则该方法将失败。

[0016] 另一个更直观的方法是将给出的HOA表示变换成本B.Rafael的“Plane-wave decomposition of the sound field on a sphere by spherical convolution”(J.Acoust.Soc.Am.,第4卷,第116号,2149-2157页,2004年10月)中所述的空间域,然后搜索方向功率中的最大值。该方法的不足之处是环境分量的存在将导致方向功率分布的模糊,并且与不存在任何环境分量相比,将导致方向功率的最大值的移位。

发明内容

[0017] 本发明要解决的问题是提供一种HOA信号的压缩,由此仍然保持HOA信号表示的高空间分辨率。通过在权利要求1和2中所述的方法解决该问题。在权利要求3和4中公开了利用这些方法的装置。

[0018] 本发明解决声场的高阶高保真度立体声响复制HOA表示的压缩。在本申请中,术语“HOA”是指所述高阶高保真度立体声响复制表示以及对应地编码或表示后的音频信号。估计主声音方向,并且将HOA信号表示分解成时域中的若干主方向信号和相关的方向信息以及HOA域中的环境分量,继之以通过降低其阶来压缩环境分量。在该分解之后,将降低了阶的环境HOA分量变换到空间域,并且与方向信号一起进行感知编码。

[0019] 在接收器或解码器侧,感知地解压缩编码后的方向信号和阶降低后经编码的环境分量。将经感知解压缩的环境信号变换为降低了阶的HOA域表示,继之以阶扩展。从方向信号和对应的方向信息以及从原始阶的环境HOA分量重新组成总的HOA表示。

[0020] 有利地,环境声场分量可以通过具有低于原始的阶的HOA表示以足够的准确度来表示,并且主方向信号的提取确保了在压缩和解压缩之后仍然获得高空间分辨率。

[0021] 原则上,本发明的方法适于压缩高阶高保真度立体声响复制HOA信号表示,所述方法包括以下步骤:

[0022] -估计主方向,其中,所述主方向估计取决于能量上的主HOA分量的方向功率分布;

[0023] -将HOA信号表示分解或解码成时域中的若干主方向信号和相关的方向信息以及HOA域中的残差环境分量,其中,所述残差环境分量表示所述HOA信号表示和所述主方向信号的表示之间的差异;

[0024] -通过与所述残差环境分量的原始阶相比降低所述残差环境分量的阶来压缩所述残差环境分量;

[0025] -将降低了阶的所述残差环境HOA分量变换到空间域;

[0026] -对所述主方向信号和所述变换后的残差环境HOA分量进行感知编码。

[0027] 原则上,本发明的方法适于对通过以下步骤进行了压缩的高阶高保真度立体声响复制HOA信号表示进行解压缩:

[0028] -估计主方向,其中,所述主方向估计取决于能量上的主HOA分量的方向功率分布;

[0029] -将HOA信号表示分解或解码成时域中的若干主方向信号和相关的方向信息以及HOA域中的残差环境分量,其中,所述残差环境分量表示所述HOA信号表示和所述主方向信号的表示之间的差异;

[0030] -通过与所述残差环境分量的原始阶相比降低所述残差环境分量的阶来压缩所述残差环境分量;

[0031] -将降低了阶的所述残差环境分量变换到空间域;

- [0032] -对所述主方向信号和所述变换后的残差环境HOA分量进行感知编码；
[0033] 所述方法包括以下步骤：
[0034] -对所述经感知编码的主方向信号和所述经感知编码的变换后的残差环境HOA分量进行感知解码；
[0035] -对经感知解码的变换后的残差环境HOA分量进行逆变换以便取得HOA域表示；
[0036] -对经逆变换的残差环境HOA分量进行阶扩展以便建立原始阶的环境HOA分量；
[0037] -组成所述经感知解码的主方向信号、所述方向信息以及所述经原始阶扩展的环境HOA分量以便取得HOA信号表示。
[0038] 原则上，本发明的装置适于压缩高阶高保真度立体声响复制HOA信号表示，所述装置包括：
[0039] -适于估计主方向的部件，其中，所述主方向估计取决于能量上的主HOA分量的方向功率分布；
[0040] -适于将HOA信号表示分解或解码成时域中的若干主方向信号和相关的方向信息以及HOA域中的残差环境分量的部件，其中，所述残差环境分量表示所述HOA信号表示和所述主方向信号的表示之间的差异；
[0041] -适于通过与所述残差环境分量的原始阶相比降低所述残差环境分量的阶来压缩所述残差环境分量的部件；
[0042] -适于将降低了阶的所述残差环境分量变换到空间域的部件；
[0043] -适于对所述主方向信号和所述变换后的残差环境HOA分量进行感知编码的部件。
[0044] 原则上，本发明的装置适于对通过以下步骤进行了压缩的高阶高保真度立体声响复制HOA信号表示进行解压缩：
[0045] -估计主方向，其中，所述主方向估计取决于能量上的主HOA分量的方向功率分布；
[0046] -将HOA信号表示分解或解码成时域中的若干主方向信号和相关的方向信息以及HOA域中的残差环境分量，其中，所述残差环境分量表示所述HOA信号表示和所述主方向信号的表示之间的差异；
[0047] -通过与所述残差环境分量的原始阶相比降低所述残差环境分量的阶来压缩所述残差环境分量；
[0048] -将降低了阶的所述残差环境分量变换到空间域；
[0049] -对所述主方向信号和所述变换后的残差环境HOA分量进行感知编码；
[0050] 所述装置包括：
[0051] -适于对经感知编码的主方向信号和经感知编码的变换后的残差环境HOA分量进行感知解码的部件；
[0052] -适于对经感知解码的变换后的残差环境HOA分量进行逆变换以便取得HOA域表示的部件；
[0053] -适于对所述经逆变换的残差环境HOA分量进行阶扩展以便建立原始阶的环境HOA分量的部件；
[0054] -适于组成所述经感知解码的主方向信号、所述方向信息以及所述经原始阶扩展的环境HOA分量以便取得HOA信号表示的部件。
[0055] 在相应的从属权利要求中公开了本发明的有利的另外的实施例。

附图说明

- [0056] 参照附图说明本发明的示例性实施例,附图中:
- [0057] 图1是关于不同的高保真度立体声响复制阶N和角 $\Theta \in [0, \pi]$ 的规一化离差函数 $v_N(\Theta)$;
- [0058] 图2是根据本发明的压缩处理的框图;
- [0059] 图3是根据本发明的解压缩处理的框图。

具体实施方式

[0060] 高保真度立体声响复制信号使用球谐函数(SH)展开式描述无源区域内的声场。这种描述的灵活性可以归因于声压的时间和空间行为基本上由波动方程确定这一物理特性。

[0061] 波动方程和球谐函数展开式

[0062] 为了对高保真度立体声响复制进行更详细的描述,下面假设球坐标系统,其中,通过半径 $r > 0$ (亦即,到坐标原点的距离)、从极轴z测量的倾斜角 $\theta \in [0, \pi]$ 以及从x轴在 $x = y$ 平面中测量的方位角 $\phi \in [0, 2\pi]$ [来表示空间 $x = (r, \theta, \phi)^T$ 中的点。在该球坐标系统中,关于连通的无源区域内的声压 $p(t, x)$ (其中,t表示时间)的波动方程由Earl G.Williams的教科书“Fourier Acoustics”(Applied Mathematical Sciences第93卷,Academic Press,1999年)给出:

$$[0063] \frac{1}{r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial p(t, x)}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial p(t, x)}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 p(t, x)}{\partial \phi^2} \right] - \frac{1}{c_s^2} \frac{\partial^2 p(t, x)}{\partial t^2} = 0 \quad (1)$$

[0064] 其中, c_s 指示声音的速度。因此,关于时间的声压的傅里叶变换为

$$[0065] P(\omega, x) := \mathcal{F}_t \{p(t, x)\} \quad (2)$$

$$[0066] := \int_{-\infty}^{\infty} p(t, x) e^{-i\omega t} dt \quad (3)$$

[0067] 其中,i表示虚数单位,根据Williams的教科书可以展开成SH的级数:

$$[0068] P(kc_s, (r, \theta, \phi)^T) = \sum_{n=0}^{\infty} \sum_{m=-n}^n p_n^m(kr) Y_n^m(\theta, \phi) \quad (4)$$

[0069] 应当注意到,该展开式对于连通的无源区域(其对应于序列的收敛的区域)内的所有点x均有效。

[0070] 在等式(4)中,k表示由下式定义的角形波数量:

$$[0071] k := \frac{\omega}{c_s} \quad (5)$$

[0072] 并且 $p_n^m(kr)$ 指示SH展开式系数,其只取决于乘积kr。

[0073] 另外, $Y_n^m(\theta, \phi)$ 是阶n以及次数(degree)m的SH函数:

$$[0074] Y_n^m(\theta, \phi) := \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos \theta) e^{im\phi} \quad (6)$$

[0075] 其中, $P_n^m(\cos \theta)$ 表示相关联的勒让德函数,并且 $(\bullet)!$ 表示阶乘。

[0076] 关于非负次数指数m的相关联的勒让德函数通过勒让德多项式 $P_n(x)$ 定义,如下:

$$[0077] P_n^m(x) := (-1)^m (1-x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n(x), \text{ 其中 } m \geq 0. \quad (7)$$

[0078] 对于负次数指数,亦即 $m < 0$,相关联的勒让德函数定义如下:

[0079] $P_n^m(x) := (-1)^m \frac{(n+m)!}{(n-m)!} P_n^{-m}(x)$, 其中 $m < 0$ 。 (8)

[0080] 继而勒让德多项式 $P_n(x)$ ($n \geq 0$) 可以使用罗德里格公式定义为:

[0081] $P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n$ (9)

[0082] 在现有技术中,例如在M.Poletti的“Unified Description of Ambisonics using Real and Complex Spherical Harmonics”(Proceedings of the Ambisonics Symposium 2009,2009年6月25至27日,格拉茨,奥地利)中,还存在关于SH函数的定义,其通过关于负次数指数m的因子 $(-1)^m$ 从等式(6)得出。

[0083] 替代地,关于时间的声压的傅里叶变换可以使用实数SH函数 $S_n^m(\theta, \phi)$ 表示为

[0084] $P(kc_s, (r, \theta, \phi)^T) = \sum_{n=0}^{\infty} \sum_{m=-n}^n q_n^m(kr) S_n^m(\theta, \phi)$ (10)

[0085] 在文献中,存在关于实数SH函数的多种定义(例如,参见上述的Poletti的论文)。一种在本文档中应用的可行的定义由下式给出:

$$[0086] S_n^m(\theta, \phi) := \begin{cases} \frac{(-1)^m}{\sqrt{2}} [Y_n^m(\theta, \phi) + Y_n^{m*}(\theta, \phi)] & \text{对于 } m > 0 \\ Y_n^m(\theta, \phi) & \text{对于 } m = 0 \\ \frac{(-1)^m}{i\sqrt{2}} [Y_n^m(\theta, \phi) - Y_n^{m*}(\theta, \phi)] & \text{对于 } m < 0 \end{cases} \quad (11)$$

[0087] 其中, $(\cdot)^*$ 表示复数共轭。通过将等式(6)插入到等式(11)中得到一种替代的表示:

[0088] $S_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) \operatorname{trg}_m(\phi)$ (12)

[0089] 其中,

[0090] $\operatorname{trg}_m(\phi) := \begin{cases} (-1)^m \sqrt{2} \cos(m\phi) & \text{对于 } m > 0 \\ 1 & \text{对于 } m = 0 \\ -\sqrt{2} \sin(m\phi) & \text{对于 } m < 0 \end{cases}$ (13)

[0091] 虽然实数SH函数针对每个定义都是实数值的,但是一般地,对于对应的展开式系数 $q_n^m(kr)$,这并不满足。

[0092] 复数SH函数涉及如下实数SH函数:

$$[0093] Y_n^m(\theta, \phi) = \begin{cases} \frac{q_n^m(kr)}{\sqrt{2}} [S_n^m(\theta, \phi) + iS_n^{-m}(\theta, \phi)] & \text{对于 } m > 0 \\ S_n^0(\theta, \phi) & \text{对于 } m = 0 \\ \frac{1}{i\sqrt{2}} [S_n^m(\theta, \phi) + iS_n^{-m}(\theta, \phi)] & \text{对于 } m < 0 \end{cases} \quad (14)$$

[0094] 复数SH函数 $Y_n^m(\theta, \phi)$ 以及具有方向矢量 $\Omega := (\theta, \phi)^T$ 的实数SH函数 $S_n^m(\theta, \phi)$ 形成三维空间中的单位球 \mathcal{S}^2 上的平方可积分复数值函数的正交基,因此满足如下条件:

[0095]
$$\int_{\mathcal{S}^2} Y_n^m(\Omega) Y_{n'}^{m'*}(\Omega) d\Omega = \int_0^{2\pi} \int_0^\pi Y_n^m(\theta, \phi) Y_{n'}^{m'*}(\theta, \phi) \sin\theta d\theta d\phi \\ = \delta_{n-n'} \delta_{m-m'} \quad (15)$$

$$[0096] \int_{\mathcal{S}^2} S_n^m(\Omega) S_{n'}^{m'}(\Omega) d\Omega = \delta_{n-n'} \delta_{m-m'} \quad (16)$$

[0097] 其中, δ 表示克罗内克 δ 函数。使用等式(15) 和等式(11) 中的实数球谐函数的定义可以得出第二结果。

[0098] 内部问题和高保真度立体声响复制系数

[0099] 高保真度立体声响复制的目的是表示坐标原点附近的声场。在不失一般性的情况下, 此处假设感兴趣的这个区域为以坐标原点为中心的半径为 R 的球形体, 其通过集合 $\{x | 0 \leq r \leq R\}$ 指定。关于该表示的关键假设是假定该球形体不包含任何声源。找出该球形体内的声场表示被称为“内部问题”, 参见上述的 Williams 的教科书。

[0100] 可以示出, 关于该内部问题, SH 函数展开式系数 $p_n^m(kr)$ 可以表示为

$$[0101] p_n^m(kr) = a_n^m(k) j_n(kr) \quad (17)$$

[0102] 其中, $j_n(\cdot)$ 表示一阶球贝塞尔函数。根据等式(17), 其满足关于声场的完整信息包含在被称为高保真度立体声响复制系数的系数 $a_n^m(k)$ 中。

[0103] 类似地, 可以对实数 SH 函数展开式 $q_n^m(kr)$ 的系数进行因式分解为

$$[0104] q_n^m(kr) = b_n^m(k) j_n(kr) \quad (18)$$

[0105] 其中, 系数 $b_n^m(k)$ 被称为关于使用实数值的 SH 函数的展开式的高保真度立体声响复制系数。它们还通过下式与 $a_n^m(k)$ 相关:

$$[0106] b_n^m(k) = \begin{cases} \frac{1}{\sqrt{2}} [(-1)^m a_n^m(k) + a_n^{-m}(k)] & \text{对于 } m > 0 \\ a_n^0(k) & \text{对于 } m = 0 \\ \frac{1}{i\sqrt{2}} [a_n^m(k) - (-1)^m a_n^{-m}(k)] & \text{对于 } m < 0 \end{cases} \quad (19)$$

[0107] 平面波分解

[0108] 在中心位于坐标原点的声音无源球形体内的声场可以通过从所有可能方向碰撞到该球形体上的无限数量的角形波数量 k 不同的平面波的重叠表示, 参见上述的 Rafely 的 “Plane-wave decomposition...” 论文。假设来自方向 Ω_0 的具有角形波数量 k 的平面波的复数幅度由 $D(k, \Omega_0)$ 给出, 可以使用等式(11) 和等式(19) 以类似的方式示出关于实数 SH 函数展开式的对应的高保真度立体声响复制系数由下式给出:

$$[0109] b_{n, \text{plane wave}}^m(k; \Omega_0) = 4\pi i^n D(k, \Omega_0) S_n^m(\Omega_0) \quad (20)$$

[0110] 因此, 关于从无限数量的角形波数量为 k 的平面波的重叠得到的声场的高保真度立体声响复制系数从等式(20) 在所有可能的方向 $\Omega_0 \in \mathcal{S}^2$ 的积分得到:

$$[0111] b_n^m(k) = \int_{\mathcal{S}^2} b_{n, \text{plane wave}}^m(k; \Omega_0) d\Omega_0 \quad (21)$$

$$[0112] = 4\pi i^n \int_{\mathcal{S}^2} D(k, \Omega_0) S_n^m(\Omega_0) d\Omega_0 \quad (22)$$

[0112] 函数 $D(k, \Omega)$ 被称为“幅度密度”, 并且假设在单位球 \mathcal{S}^2 上是平方可积分的。可以将其展开成实数 SH 函数的级数, 如下

$$[0113] D(k, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n c_n^m(k) S_n^m(\Omega) \quad (23)$$

[0114] 其中, 展开系数 $c_n^m(k)$ 等于出现在等式(22) 中的积分, 亦即

$$[0115] c_n^m(k) = \int_{\mathcal{S}^2} D(k, \Omega) S_n^m(\Omega) d\Omega \quad (24)$$

[0116] 通过将等式(24)插入到等式(22)中,可以看出高保真度立体声响复制系数 $b_n^m(k)$ 是展开式系数 $c_n^m(k)$ 的缩放后的版本,亦即

$$[0117] b_n^m(k) = 4\pi i^n c_n^m(k) \quad (25)$$

[0118] 在对缩放后的高保真度立体声响复制系数 $c_n^m(k)$ 以及幅度密度函数 $D(k, \Omega)$ 应用关于时间的逆傅里叶变换时,得到对应的时域量

$$[0119] \tilde{c}_n^m(t) := \mathcal{F}_t^{-1} \left\{ c_n^m \left(\frac{\omega}{c_s} \right) \right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} c_n^m \left(\frac{\omega}{c_s} \right) e^{i\omega t} d\omega \quad (26)$$

$$[0120] d(t, \Omega) := \mathcal{F}_t^{-1} \left\{ D \left(\frac{\omega}{c_s}, \Omega \right) \right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} D \left(\frac{\omega}{c_s}, \Omega \right) e^{i\omega t} d\omega \quad (27)$$

[0121] 然后,在时域中,可以将等式(24)用公式表示为

$$[0122] \tilde{c}_n^m(t) = \int_{\mathcal{S}^2} d(t, \Omega) S_n^m(\Omega) d\Omega \quad (28)$$

[0123] 时域方向信号 $d(t, \Omega)$ 可以通过实数SH函数展开式根据下式表示

$$[0124] d(t, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \tilde{c}_n^m(t) S_n^m(\Omega) \quad (29)$$

[0125] 使用SH函数 $S_n^m(\Omega)$ 是实数值的这一事实,其复数共轭可以表示为

$$[0126] d^*(t, \Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \tilde{c}_n^m(t) S_n^m(\Omega) \quad (30)$$

[0127] 假设时域信号 $d(t, \Omega)$ 是实数值的,亦即 $d(t, \Omega) = d^*(t, \Omega)$,根据等式(29)与等式(30)的比较,可以得出系数 $\tilde{c}_n^m(t)$ 在该情况下是实数值的,亦即 $\tilde{c}_n^m(t) = \tilde{c}_n^m(t)$ 。

[0128] 下面,将系数 $\tilde{c}_n^m(t)$ 称为缩放后的时域高保真度立体声响复制系数。

[0129] 下面,还假设声场表示通过将在下面的处理压缩的部分更详细地描述的这些系数给出。

[0130] 注意,通过用于根据本发明的处理的系数 $\tilde{c}_n^m(t)$ 进行的时域HOA表示等效于对应的频域HOA表示 $c_n^m(k)$ 。因此,在对等式进行了较小的相应修改的情况下,可以在频域中等效地实现所述压缩和解压缩。

[0131] 具有有限阶的空间分辨率

[0132] 实践中,仅使用有限数量的阶 $n \leq N$ 的高保真度立体声响复制系数 $c_n^m(k)$ 描述坐标原点附近的声场。相对于真实幅度密度函数 $D(k, \Omega)$,根据下式从截断的SH函数级数计算幅度密度函数引入了一种空间离差

$$[0133] D_N(k, \Omega) := \sum_{n=0}^N \sum_{m=-n}^n c_n^m(k) S_n^m(\Omega) \quad (31)$$

[0134] 参见上述的“Plane-wave decomposition...”论文。这可以通过使用等式(31)对来自方向 Ω_0 的单个平面波计算幅度密度函数来实现:

$$D_N(k, \Omega) = \sum_{n=0}^N \sum_{m=-n}^n \frac{1}{4\pi i^n n!} \cdot b_{n, \text{plane wave}}^m(k; \Omega_0) S_n^m(\Omega) \quad (32)$$

$$= D(k, \Omega_0) \sum_{n=0}^N \sum_{m=-n}^n S_n^m(\Omega_0) S_n^m(\Omega) \quad (33)$$

[0135] $= D(k, \Omega_0) \sum_{n=0}^N \sum_{m=-n}^n Y_n^m(\Omega_0)^* Y_n^m(\Omega) \quad (34)$

$$= D(k, \Omega_0) \sum_{n=0}^N \frac{2n+1}{4\pi} P_n(\cos\Theta) \quad (35)$$

$$= D(k, \Omega_0) \left[\frac{N+1}{4\pi(\cos\Theta-1)} (P_{N+1}(\cos\Theta) - P_N(\cos\Theta)) \right] \quad (36)$$

$$= D(k, \Omega_0) v_N(\Theta) \quad (37)$$

[0136] 其中

$$[0137] v_N(\Theta) := \frac{N+1}{4\pi(\cos\Theta-1)} (P_{N+1}(\cos\Theta) - P_N(\cos\Theta)) \quad (38)$$

[0138] 其中, Θ 表示满足下述属性的指向方向 Ω 和 Ω_0 的两个矢量之间的角

$$[0139] \cos\Theta = \cos\theta \cos\theta_0 + \cos(\phi - \phi_0) \sin\theta \sin\theta_0 \quad (39)$$

[0140] 在等式(34)中, 利用在等式(20)中给出的平面波的高保真度立体声响复制系数, 而在等式(35)和(36)中利用一些数学理论, 参见上述的“Plane-wave decomposition...”论文。可以使用等式(14)示出在等式(33)中的属性。

[0141] 比较等式(37)与真实幅度密度函数

$$[0142] D(k, \Omega) = D(k, \Omega_0) \frac{\delta(\Theta)}{2\pi} \quad (40)$$

[0143] 其中, $\delta(\cdot)$ 表示迪拉克 δ 函数, 从将缩放后的迪拉克 δ 函数替换为离差函数 $v_N(\Theta)$ (其在按照其最大值进行了归一化之后, 针对不同的高保真度立体声响复制阶 N 和角 $\Theta \in [0, \pi]$, 在图1中示出), 空间离差变得显而易见。

[0144] 因为对于 $N \geq 4$, $v_N(\Theta)$ 的第一个零近似地位于 $\frac{\pi}{N}$ (参见上述的“Plane-wave decomposition...”论文), 随着增加高保真度立体声响复制阶 N , 离差效应降低 (并且因此空间分辨率提高)。

[0145] 对于 $N \rightarrow \infty$, 离差函数 $v_N(\Theta)$ 收敛到缩放后的迪拉克 δ 函数。在以下情况下可以看到这一点: 勒让德多项式的完整关系

$$[0146] \sum_{n=0}^{\infty} \frac{2n+1}{2} P_n(x) P_n(x') = \delta(x - x') \quad (41)$$

[0147] 与等式(35)一起使用以将关于 $N \rightarrow \infty$ 的 $v_N(\Theta)$ 的极限表示为

$$\lim_{N \rightarrow \infty} v_N(\Theta) = \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{2n+1}{2} P_n(\cos\Theta) \quad (42)$$

$$= \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{2n+1}{2} P_n(\cos\Theta) P_n(1) \quad (43)$$

[0148] $= \frac{1}{2\pi} \delta(\cos\Theta - 1) \quad (44)$

$$= \frac{1}{2\pi} \delta(\Theta) . \quad (45)$$

[0149] 在通过

$$[0150] \mathbf{S}(\Omega) := (S_0^0(\Omega), S_1^{-1}(\Omega), S_1^0(\Omega), S_1^1(\Omega), S_2^{-2}(\Omega), \dots, S_N^N(\Omega))^T \in \mathbb{R}^o \quad (46)$$

[0151] 定义阶 $n \leq N$ 的实数SH函数的矢量时, 其中, $o = (N+1)^2$, 并且 $(.)^T$ 表示转置, 等式

(37) 与等式(33)比较示出离差函数可以通过两个实数SH矢量的标量乘积表示为

$$[0152] \quad v_N(\Theta) = S^T(\Omega) S(\Omega_0) \quad (47)$$

[0153] 在时域中,可以将离差等效地表示为

$$[0154] \quad d_N(t, \Omega) := \sum_{n=0}^N \sum_{m=-n}^n \tilde{c}_n^m(t) S_n^m(\Omega) \quad (48)$$

$$[0154] \quad = d(t, \Omega_0) v_N(\Theta) \quad (49)$$

[0155] 采样

[0156] 对于一些应用,期望根据在有限数量J的离散方向 Ω_j 上的时域幅度密度函数 $d(t, \Omega)$ 的样本确定缩放后的时域高保真度立体声响复制系数 $\tilde{c}_n^m(t)$ 。然后,根据B.Rafaely的“Analysis and Design of Spherical Microphone Arrays”(IEEE Transactions on Speech and Audio Processing,卷13,第1号,页135-143,2005年1月)通过有限求和近似等式(28)中的积分:

$$[0157] \quad \tilde{c}_n^m(t) \approx \sum_{j=1}^J g_j \cdot d(t, \Omega_j) S_n^m(\Omega_j) \quad (50)$$

[0158] 其中, g_j 表示一些适当选取的采样加权。相对于“Analysis and Design...”论文,近似(50)是指使用实数SH函数的时域表示而不是使用复数SH函数的频域表示。使近似(50)变得精确的必要条件是幅度密度是有限谐函数阶N的,意味着

$$[0159] \quad \tilde{c}_n^m(t) = 0 \quad \text{对于 } n > N. \quad (51)$$

[0160] 如果该条件不满足,则近似(50)受到空间混叠误差的影响,参见B.Rafaely的“Spatial Aliasing in Spherical Microphone Arrays”(IEEE Transactions on Signal Processing,卷55,第3期,第1003-1010页,2007年3月)。

[0161] 第二必要条件需要采样点 Ω_j 和对应的加权满足在“Analysis and Design...”论文中给了的对应条件:

$$[0162] \quad \sum_{j=1}^J g_j S_{n'}^{m'}(\Omega_j) S_n^m(\Omega_j) = \delta_{n-n'} \delta_{m-m'} \quad \text{对于 } m, m' \leq N \quad (52)$$

[0163] 条件(51)和(52)联合起来对于精确采样就足够了。

[0164] 采样条件(52)由一组线性等式组成,可以使用单个矩阵等式简洁地用公式表示为

$$[0165] \quad \Psi G \Psi^H = I \quad (53)$$

[0166] 其中, Ψ 表示由下式定义的模式矩阵

$$[0167] \quad \Psi := [S(\Omega_1) \quad \dots \quad S(\Omega_J)] \in \mathbb{R}^{O \times J} \quad (54)$$

[0168] 并且G表示在其对角线上具有加权的矩阵,亦即

$$[0169] \quad G := \text{diag}(g_1, \dots, g_J) \quad (55)$$

[0170] 从等式(53)可以看出,满足等式(52)的必要条件是采样点的数量J满足 $J \geq 0$ 。将在J个采样点处的时域幅度密度的值聚集到如下矢量中

$$[0171] \quad w(t) := (D(t, \Omega_1), \dots, D(t, \Omega_J)) \quad (56)$$

[0172] 并且通过下式定义缩放后的时域高保真度立体声响复制系数的矢量

$$[0173] \quad c(t) := (\tilde{c}_0^0(t), \tilde{c}_1^{-1}(t), \tilde{c}_1^0(t), \tilde{c}_1^1(t), \tilde{c}_2^{-2}(t), \dots, \tilde{c}_O^0(t))^T \quad (57)$$

[0174] 两个矢量通过SH函数展开式(29)相关。这种关系提供了下面的线性等式系统:

$$[0175] \quad w(t) = \Psi^H c(t) \quad (58)$$

[0176] 使用所引入的矢量记号,从时域幅度密度函数样本的值计算缩放后的时域高保真度立体声响复制系数可以写作:

[0177] $c(t) \approx \Psi G_w(t)$ (59)

[0178] 给出固定的高保真度立体声响复制阶N,经常无法实现通过计算J ≥ 0 数量的采样点 Ω_j 和对应的加权使得满足采样条件等式(52)。然而,如果选取采样点使得良好地近似采样条件,则模式矩阵 Ψ 的秩为0,并且其条件数低。在该情况下,存在模式矩阵 Ψ 的伪逆

[0179] $\Psi^+ := (\Psi \Psi^H)^{-1} \Psi \Psi^+$ (60)

[0180] 并且通过下式给出从时域幅度密度函数样本的矢量到缩放后的时域高保真度立体声响复制系数矢量c(t)的合理近似

[0181] $c(t) \approx \Psi^+ w(t)$ (61)

[0182] 如果J=0并且模式矩阵的秩为0,则其伪逆与其逆一致,因为

[0183] $\Psi^+ = (\Psi \Psi^H)^{-1} \Psi = \Psi^{-H} \Psi^{-1} \Psi = \Psi^{-H}$ (62)

[0184] 如果额外满足采样条件等式(52),则满足

[0185] $\Psi^{-H} = \Psi G$ (63)

[0186] 并且两个近似(59)和(61)是等价的并且是精确的。

[0187] 可以将矢量w(t)解释为空间时域信号的矢量。从HOA域到空间域的变换可以例如通过使用等式(58)进行。这种变换在本申请中被称为“球谐函数变换”(SHT)并且在降低了阶的环境HOA分量变换到空间域时使用。隐含地假设SHT的空间采样点 Ω_j 近似地满足在 $g_j \approx \frac{4\pi}{o}$ ($j = 1, \dots, J$)并且J=0情况下的等式(52)中的采样条件。

[0188] 在这些假设下,SHT矩阵满足 $\Psi^H \approx \frac{4\pi}{o} \Psi^{-1}$ 。在SHT的绝对缩放不重要的情况下,则可以忽略常量 $\frac{4\pi}{o}$ 。

[0189] 压缩

[0190] 本发明涉及对给出的HOA信号表示的压缩。如上所述,将HOA表示分解成时域中的预定义数量的主方向信号以及HOA域中的环境分量,继之以通过降低环境分量的阶来压缩环境分量的HOA表示。该操作利用如下被收听测试支持的假设:环境声场分量可以通过具有低阶的HOA表示以足够的精确度来表示。对主方向信号的提取确保了在压缩和对应的解压缩之后保持高空间分辨率。

[0191] 在分解之后,降低了阶的环境HOA分量被变换到空间域,并且与如在专利申请EP 10306472.1的Exemplary embodiments部分中所述那样与方向信号一起被感知地编码。

[0192] 压缩处理包括在图2中图示的两个相继步骤。在下面的压缩的细节部分描述单独信号的确切定义。

[0193] 在图2a中示出的第一步骤或阶段中,在主方向估计器22中估计主方向,并且进行将高保真度立体声响复制信号C(1)分解成方向分量以及残差或环境分量,其中1表示帧索引。在方向信号计算步骤或阶段23中计算方向分量,由此高保真度立体声响复制表示被转换到由具有对应的方向 $\bar{\Omega}_{DOM}(l)$ 的D个常规方向信号X(l)的集合表示的时域信号。在环境HOA分量计算步骤或阶段24中计算残差的环境分量,并且表示为HOA域系数C_A(1)。

[0194] 在图2b中示出的第二步骤中,对方向信号X(1)和环境HOA分量C_A(1)执行感知编码,如下:

[0195] -可以在感知编码器27中使用任何已知的感知压缩技术单独地压缩常规时域方向信号X(1)。

[0196] -在两个子步骤或阶段中执行环境HOA域分量C_A(1)的压缩。

[0197] 第一子步骤或阶段25执行将原始高保真度立体声响复制阶N降低至N_{RED}, 例如N_{RED}=2, 得到环境HOA分量C_{A,RED}(1)。此处, 利用如下假设: 可以通过具有低阶的HOA足够精确地表示环境声场分量。第二子步骤或阶段26基于在专利申请EP 10306472.1中所述的压缩。通过应用球谐函数变换, 将在子步骤/阶段25计算的环境声场分量的O_{RED}:=(N_{RED}+1)²个HOA信号C_{A,RED}(1)变换成空间域中的O_{RED}个等效信号W_{A,RED}(1), 得到可以输入给一组并行的感知编码解码器27的常规时域信号。可以应用任何已知的感知编码或压缩技术。输出编码后的方向信号 $\bar{X}(l)$ 和阶降低了的编码后的空间域信号 $\bar{W}_{A,RED}(l)$, 并且它们可以被传送或存储。

[0198] 有利地, 可以在感知编码器27中联合地执行对所有时域信号X(1)和W_{A,RED}(1)的感知压缩, 以便通过利用可能剩余的信道间相关性提高总体的编码效率。

[0199] 解压缩

[0200] 在图3中图示了对收到的或重放的信号的解压缩处理。如同压缩处理, 其包括两个相继步骤。

[0201] 在图3a中示出的第一步骤或阶段中, 在感知解码31中执行对编码后的方向信号 $\bar{X}(l)$ 以及阶降低了的编码后的空间域信号 $\bar{W}_{A,RED}(l)$ 的感知解码或者解压缩, 其中, $\bar{X}(l)$ 是表示分量并且 $\bar{W}_{A,RED}(l)$ 表示环境HOA分量。在逆球谐函数变换器32中经由逆球谐函数变换将经感知解码或解压缩的空间域信号 $\widehat{W}_{A,RED}(l)$ 变换成阶为N_{RED}的HOA域表示 $\widehat{C}_{A,RED}(l)$ 。此后, 在阶扩展步骤或阶段33中, 通过阶扩展从 $\widehat{C}_{A,RED}(l)$ 估计阶为N的适当的HOA表示 $\widehat{C}_A(l)$ 。

[0202] 在图3b示出的第二步骤或阶段中, 在HOA信号组装器34中从方向信号 $\widehat{X}(l)$ 和对应的方向信息 $\bar{\Omega}_{DOM}(l)$ 以及从原始阶的环境HOA分量 $\widehat{C}_A(l)$ 重新组成总的HOA表示 $\widehat{C}(l)$ 。

[0203] 可达到的数据率降低

[0204] 本发明所解决的问题是与现有的用于HOA表示的压缩方法相比显著地降低数据率。下面论述与非压缩的HOA表示相比的可达到的压缩率。压缩率得自传送阶为N的非压缩的HOA信号C(1)所需的数据率与传送由D个经感知编码的方向信号和对应的方向 $\bar{\Omega}_{DOM}(l)$ 以及N_{RED}个表示环境HOA分量的经感知编码的空间域信号W_{A,RED}(1)组成的压缩后的信号表示所需的数据率的比较。

[0205] 为了传送非压缩的HOA信号C(1), 需要O·f_S·N_b的数据率。相反, 传送D个经感知编码的方向信号X(1)需要D·f_{b,COD}的数据率, 其中, f_{b,COD}表示经感知编码的信号的比特率。类似地, 传送N_{RED}个经感知编码的空间域信号W_{A,RED}(1)信号需要O_{RED}·f_{b,COD}的比特率。假设基于与采样率f_S相比低得多的速率计算方向 $\bar{\Omega}_{DOM}(l)$, 亦即假设它们对于由占个样本组成的信号帧的持续时间是固定的, 例如对于f_S=48kHz的采样率, B=1200, 并且对于压缩后的HOA信号的总的数据率的计算, 可以忽略对应的数据率份额。

[0206] 因此, 传送压缩后的表示需要大约(D+O_{RED})·f_{b,COD}的数据率。因此, 压缩率r_{COMP}为

$$r_{COMP} \approx \frac{O \cdot f_S \cdot N_b}{(D + O_{RED}) \cdot f_{b,COD}} \quad (64)$$

[0208] 例如,使用降低的HOA阶 $N_{RED}=2$ 并且 $64\frac{\text{kbit}}{\text{s}}$ 的比特率将采用采样率 $f_s=48\text{kHz}$ 并且对于每个样本 $N_b=16$ 比特的阶 $N=4$ 的HOA表示压缩成具有 $D=3$ 个主方向的表示将导致 $r_{COMPR}\approx 25$ 的压缩率。传送压缩后的表示需要大约 $768\frac{\text{kbit}}{\text{s}}$ 的数据率。

[0209] 降低的出现编码噪声未屏蔽的概率

[0210] 如在背景技术中所述,在专利申请EP 10306472.1中所述的空间域信号的感知压缩受到信号之间的剩余的相互相关性的影响,其可能导致未屏蔽感知编码噪声。根据本发明,主方向信号在被感知编码之前,首先从HOA声场表示提取将其提取。这意味着,在组成HOA表示时,在感知解码之后,编码噪声具有与方向信号完全相同的空间方向性。具体地,编码噪声以及方向信号对任何任意方向的影响通过在具有有限阶的空间分辨率部分中解释的空间离差函数确定性地描述。换言之,在任何时刻,表示编码噪声的HOA系数矢量恰好是表示方向信号的HOA系数矢量的倍数。因此,噪声HOA系数的任意加权的和将不会导致对感知编码噪声的任何未屏蔽。

[0211] 另外,正如在EP 10306472.1中所提出那样地处理降低了阶的环境分量,但是因为针对每个定义,环境分量的空间域信号在彼此之间具有相当低的相关性,所以感知噪声未屏蔽的概率很低。

[0212] 改进的方向估计

[0213] 本发明的方向估计取决于能量上的主HOA分量的方向功率分布。从HOA表示的秩降低了的相关性矩阵(其通过对HOA表示的相关性矩阵的特征值分解得到)计算方向功率分布。与在上述的“Plane-wave decomposition...”论文中使用的方向估计相比,提供了更准确这一优点,因为关注于能量上的主HOA分量而不是对方向估计使用完整的HOA表示降低了方向功率分布的空间模糊。

[0214] 与在上述的“The Application of Compressive Sampling to the Analysis and Synthesis of Spatial Sound Fields”和“Time Domain Reconstruction of Spatial Sound Fields Using Compressed Sensing”论文中提出的方向估计相比,提供了更加健壮这一优点。原因是将HOA表示分解成方向分量和环境分量几乎永远不会完美的实现,使得在方向分量中保留少量环境分量。然后,像在这两个论文中那样的压缩采样方法由于它们对环境信号的存在的高度敏感性而无法提供合理的方向估计。

[0215] 有利地,本发明的方向估计不会受到该问题的影响。

[0216] HOA表示分解的替代应用

[0217] 根据在上述的Pulkki的论文“Spatial Sound Reproduction with Directional Audio Coding”中所提出的,所述的将HOA表示分解成带有相关方向信息的若干方向信号以及在HOA域中的环境分量可以用于HOA表示的信号自适应类DirAC呈现。

[0218] 可以不同地呈现每个HOA分量,因为两个分量的物理特征是不同的。例如,可以使用如基于矢量的幅度摇摄(VBAP)这样的信号摇摄技术对扩音器呈现方向信号,参见V.Pulkki的“Virtual Sound Source Positioning Using Vector Base Amplitude Panning”(Journal of Audio Eng.Society,卷45,第6期,第456–466页,1997年)。可以使得已知的标准HOA呈现技术呈现环境HOA分量。

[0219] 这样的呈现不限于阶为“1”的高保真度立体声响复制表示,并且因此可以被视为

到阶N>1的HOA表示的类DirAC呈现的扩展。

[0220] 对来自HOA信号表示的若干方向的估计可以用于任何相关类型的声场分析。

[0221] 下面的部分更详细地描述信号处理步骤。

[0222] 压缩

[0223] 输入格式的定义

[0224] 作为输入,假设在等式(26)中定义的缩放后的时域HOA系数 $\tilde{c}_n^m(t)$ 以速率 $f_s = \frac{1}{T_s}$ 进行采样。将矢量 $c(j)$ 定义为由属于采样时间 $t=jT_s, j \in \mathbb{Z}$ 的所有系数组成,其根据:

[0225] $c(j) := [\tilde{c}_0^0(jT_s), \tilde{c}_1^{-1}(jT_s), \tilde{c}_1^0(jT_s), \tilde{c}_1^1(jT_s), \tilde{c}_2^{-2}(jT_s), \dots, \tilde{c}_N^N(jT_s)]^T \in \mathbb{R}^O \quad (65)$

[0226] 成帧

[0227] 在成帧步骤或阶段21中,对缩放后的HOA系数的进入的矢量 $c(j)$ 进行成帧成为长度为B的非重叠的帧,其根据:

[0228] $C(l) := [c(lB+1) \ c(lB+2) \ \dots \ c(lB+B)] \in \mathbb{R}^{O \times B} \quad (66)$

[0229] 假设 $f_s=48\text{kHz}$ 的采样率,对应于25ms的帧持续时间,适当的帧长度为 $B=1200$ 个样本。

[0230] 主方向的估计

[0231] 对于主方向的估计,计算下面的相关性矩阵

[0232] $B(l) := \frac{1}{LB} \sum_{l'=0}^{L-1} C(l-l')C^T \quad (l-l') \in \mathbb{R}^{O \times O} \quad (67)$

[0233] 在当前帧1和L-1个先前帧上的求和指出方向分析基于具有 $L \cdot B$ 个样本的帧的长重叠组,亦即,对于每个当前帧,考虑邻近帧的内容。这有助于方向分析的稳定性,理由有两个:更长的帧导致更大数量的观测,方向估计由于重叠帧而平滑。

[0234] 假设 $f_s=48\text{kHz}$ 并且 $B=1200$,对应于100ms的总体帧持续时间,L的合理值是4。

[0235] 接下来,根据下式确定相关性矩阵B(l)的特征值分解

[0236] $B(l) = V(l) \Lambda(l) V^T(l) \quad (68)$

[0237] 其中,矩阵V(l)由特征矢量 $v_i(l), 1 \leq i \leq O$ 组成,如下

[0238] $V(l) := [v_1(l) \ v_2(l) \ \dots \ v_O(l)] \in \mathbb{R}^{O \times O} \quad (69)$

[0239] 并且 $\Lambda(l)$ 是具有对应的特征值 $\lambda_i(l), 1 \leq i \leq O$ 的对角矩阵,在其对角线上:

[0240] $\Lambda(l) := \text{diag}(\lambda_1(l), \lambda_2(l), \dots, \lambda_O(l)) \in \mathbb{R}^{O \times O} \quad (70)$

[0241] 假设以非升序编排特征值的索引,亦即,

[0242] $\lambda_1(l) \geq \lambda_2(l) \geq \dots \geq \lambda_O(l) \quad (71)$

[0243] 之后,计算主特征值的索引集合 $\{1, \dots, \tilde{j}(l)\}$ 一种对此进行管理的可行方式是定义所期望的最小宽带方向对环境功率比DAR_{MIN},然后确定 $\tilde{j}(l)$,使得

[0244] $10\log_{10}\left(\frac{\lambda_i(l)}{\lambda_1(l)}\right) \geq -\text{DAR}_{\text{MIN}} \quad \forall i \leq \tilde{j}(l) \quad \text{并且} \quad 10\log_{10}\left(\frac{\lambda_i(l)}{\lambda_1(l)}\right) > -\text{DAR}_{\text{MIN}}$

对于 $i = \tilde{j}(l) + 1 \quad (72)$

[0245] 关于DAR_{MIN}的合理选择是15dB。主特征值的数量进一步地被约束为不大于D,以便集中于不超过D个主方向。这通过将索引集 $\{1, \dots, \tilde{j}(l)\}$ 替换为 $\{1, \dots, J(l)\}$ 来实现,其中

[0246] $\mathcal{I}(l) := \max(\tilde{\mathcal{I}}(l), D)$ (73)

[0247] 接下来,通过下式得到B(1)的 $\mathcal{I}(l)$ 秩近似

[0248] $\mathbf{B}_J(l) := \mathbf{V}_J(l)\mathbf{\Lambda}_J(l)\mathbf{V}_J^T(l)$, 其中 (74)

[0249] $\mathbf{V}_J(l) := [\mathbf{v}_1(l) \quad \mathbf{v}_2(l) \quad \dots \quad \mathbf{v}_{J(l)}(l)] \in \mathbb{R}^{O \times J(l)}$ (75)

[0250] $\mathbf{\Lambda}_J(l) := \text{diag}(\lambda_1(l), \lambda_2(l), \dots, \lambda_{J(l)}(l)) \in \mathbb{R}^{J(l) \times J(l)}$ (76)

[0251] 该矩阵应当包含主方向分量对B(1)的贡献。

[0252] 之后,计算矢量

$\sigma^2(l) := \text{diag}(\mathbf{\Xi}^T \mathbf{B}_J(l) \mathbf{\Xi}) \in \mathbb{R}^Q$ (77)

[0253] $= (\mathbf{S}_1^T \mathbf{B}_J(l) \mathbf{S}_1, \dots, \mathbf{S}_Q^T \mathbf{B}_J(l) \mathbf{S}_Q)^T$ (78)

[0254] 其中, $\mathbf{\Xi}$ 表示关于大量近似相等分布的测试方向 $\Omega_q := (\theta_q, \phi_q)$, $1 \leq q \leq Q$ 的模式矩阵,其中, $\theta_q \in [0, \pi]$ 表示从极轴z测量的倾斜角 $\theta \in [0, \pi]$, 并且 $\phi_q \in [-\pi, \pi]$ 表示从x轴在x=y平面中测量的方位角。

[0255] 通过下式定义模式矩阵 $\mathbf{\Xi}$

[0256] $\mathbf{\Xi} := [\mathbf{S}_1 \quad \mathbf{S}_2 \quad \dots \quad \mathbf{S}_Q] \in \mathbb{R}^{O \times Q}$ (79)

[0257] 其中,对于 $1 \leq q \leq Q$

[0258] $\mathbf{S}_q := [S_0^0(\mathbf{\Omega}_q), S_1^{-1}(\mathbf{\Omega}_q), S_1^0(\mathbf{\Omega}_q), S_1^{-1}(\mathbf{\Omega}_q), S_2^{-2}(\mathbf{\Omega}_q), \dots, S_N^N(\mathbf{\Omega}_q)]^T$ (80)

[0259] $\sigma^2(l)$ 中的 $\sigma_q^2(l)$ 个元素是从方向 Ω_q 入射的对应于主方向信号的平面波的功率的近似。在下面的关于方向搜索算法的解释部分中提供与此相关的理论上的解释。

[0260] 根据 $\sigma^2(l)$, 计算用于方向信号分量的确定的若干 ($\tilde{D}(l)$ 个) 主方向 $\mathbf{\Omega}_{\text{CURRDOM}, \tilde{d}}(l)$, $1 \leq \tilde{d} \leq \tilde{D}(l)$ 。从而约束主方向的数量以满足 $\tilde{D}(l) \leq D$, 以便确保不变的数据率。然而,如果允许可变的数据率,则主方向的数量可以适配于当前的声音场景。

[0261] 计算 $\tilde{D}(l)$ 个主方向的一种可行方式是将第一主方向设置成具有最大功率的那个,亦即, $\mathbf{\Omega}_{\text{CURRDOM}, 1}(l) = \mathbf{\Omega}_{q_1}$, 其中, $q_1 := \text{argmax}_{q \in \mathcal{M}_1} \sigma_q^2(l)$ 并且 $\mathcal{M}_1 := \{1, 2, \dots, Q\}$ 。

假设由主方向信号创建功率最大值,并且考虑使用有限阶N的HOA表示得到方向信号的空间离差的事实(参见,上述的“Plane-wave decomposition...”论文),则可以断定:在 $\Omega_{\text{CURRDOM}, 1}(l)$ 的方向领域中,应当出现属于相同的方向信号的功率分量。因为可以通过函数 $v_N(\Theta_{q, q_1})$ (参见等式(38)) 表示空间信号离差,其中, $\Theta_{q, q_1} := \angle(\mathbf{\Omega}_q, \mathbf{\Omega}_{q_1})$ 表示 Ω_q 和 $\Omega_{\text{CURRDOM}, 1}(l)$ 之间的角,属于方向信号的功率根据 $v_N^2(\Theta_{q, q_1})$ 下降。因此,对于另外的主方向的搜索,排除在具有 $\Theta_{q, 1} \leq \Theta_{\text{MIN}}$ 的 $\mathbf{\Omega}_{q_1}$ 的方向领域中的所有方向 Ω_q ,这是合理的。可以将距离 Θ_{MIN} 选取为 $v_N(x)$ (对于 $N \geq 4$, 其近似地通过 $\frac{\pi}{N}$ 给出) 的第一个零。然后,将第二主方向设置为在剩余的方向 $\mathbf{\Omega}_q \in \mathcal{M}_2$ 上具有最大功率的那个,其中, $\mathcal{M}_2 := \{q \in \mathcal{M}_1 | \Theta_{q, 1} > \Theta_{\text{MIN}}\}$ 。以类似的方式确定剩余的主方向。

[0262] 可以通过以下方式确定主方向的数量 $\tilde{D}(l)$: 考虑分配给单独的主方向 $\mathbf{\Omega}_{q_d}$ 的功率

$\sigma_{q_d}^2(l)$, 并且搜索比值 $\sigma_{q_1}^2(l)/\sigma_{q_d}^2(l)$ 超过所期望的方向对环境率比 DAR_{MIN} 的值的情况。这意味着, $\tilde{D}(l)$ 满足

$$[0263] \quad 10 \log_{10} \left(\frac{\sigma_{q_1}^2(l)}{\sigma_{q_{\tilde{D}(l)}}^2(l)} \right) \leq DAR_{MIN} \wedge \left[10 \log_{10} \left(\frac{\sigma_{q_1}^2(l)}{\sigma_{q_{\tilde{D}(l)+1}}^2(l)} \right) > DAR_{MIN} \vee \tilde{D}(l) = D \right] \quad (81)$$

[0264] 关于计算所有主方向的总体处理可以按照下面执行:

算法 1 给定球上的功率分布的主方向搜索

```

PowerFlag = true
 $\tilde{d} = 1$ 
 $\mathcal{M}_1 = \{1, 2, \dots, Q\}$ 
repeat
     $q_{\tilde{d}} = \operatorname{argmax}_{q \in \mathcal{M}_{\tilde{d}}} \sigma_q^2(l)$ 
    if  $\left[ \tilde{d} > 1 \wedge 10 \log_{10} \left( \frac{\sigma_{q_1}^2(l)}{\sigma_{q_{\tilde{d}}}^2(l)} \right) > DAR_{MIN} \right]$  then
        PowerFlag = false
    else
         $\Omega_{CURRENDOM, \tilde{d}}(l) = \Omega_{q_{\tilde{d}}}$ 
         $\mathcal{M}_{\tilde{d}+1} = \{q \in \mathcal{M}_{\tilde{d}} \mid \angle(\Omega_q, \Omega_{q_{\tilde{d}}}) > \Theta_{MIN}\}$ 
         $\tilde{d} = \tilde{d} + 1$ 
    end if
until  $\left[ \tilde{d} > D \vee PowerFlag = \text{false} \right]$ 
 $\tilde{D}(l) = \tilde{d} - 1$ 

```

[0266] 接下来, 对在当前帧中得到的方向 $\Omega_{CURRENDOM, \tilde{d}}(l)$, $1 \leq \tilde{d} \leq \tilde{D}(l)$ 和先前帧中的方向进行平滑, 得到平滑的方向 $\bar{\Omega}_{DOM, d}(l)$, $1 \leq d \leq D$ 。该操作可以分成两个相继部分:

[0267] (a) 对先前帧中的平滑的方向 $\bar{\Omega}_{DOM, d}(l)$ ($1 \leq d \leq D$) 分配当前的主方向 $\Omega_{CURRENDOM, \tilde{d}}(l)$, $1 \leq \tilde{d} \leq \tilde{D}(l)$ 。确定分配函数 $f_{A, l}: \{1, \dots, \tilde{D}(l)\} \rightarrow \{1, \dots, D\}$, 使得分配的方向之间的角的和

$$[0268] \quad \sum_{\tilde{d}=1}^{\tilde{D}(l)} \angle(\Omega_{CURRENDOM, \tilde{d}}(l), \bar{\Omega}_{DOM, f_{A, l}(\tilde{d})}(l-1)) \quad (82)$$

[0269] 最小化。可以使用著名的匈牙利算法(参见H.W.Kuhn的“The Hungarian method for the assignment problem”, Naval research logistics quarterly 2, 第1-2期, 第83-97页, 1955年)解决这样的分配问题。将当前方向 $\Omega_{CURRENDOM, \tilde{d}}(l)$ 和先前帧 $\bar{\Omega}_{DOM, d}(l-1)$ 中的不活动的方向(关于术语“不活动的方向”的解释, 参见下面)之间的角设置为 $2\Theta_{MIN}$ 。该操作的效果是, 试图将比 $2\Theta_{MIN}$ 更接近于先前活动的方向 $\bar{\Omega}_{DOM, d}(l-1)$ 的当前方向 $\Omega_{CURRENDOM, \tilde{d}}(l)$ 分配给它们。如果距离超过 $2\Theta_{MIN}$, 则假设对应的当前方向属于新的信号, 这意味着其优先分配给先前不活动的方向 $\bar{\Omega}_{DOM, d}(l-1)$ 。注释: 当允许全体压缩算法的更大等待时间时, 相继方向估计的分配可以更健壮的进行。例如, 可以更好地识别突

然的方向改变,而不会将它们与从估计误差得到的离群值混合在一起。

[0270] (b) 使用步骤(a)中的分配计算平滑的方向 $\bar{\Omega}_{\text{DOM},d}(l-1)$, $1 \leq d \leq D$ 。平滑是基于球的几何形状而不是欧几里得几何形状。对于当前的主方向

$\Omega_{\text{CURRDOM},\tilde{d}}(l)$ ($1 \leq \tilde{d} \leq \tilde{D}(l)$) 中的每个,沿着由方向 $\Omega_{\text{CURRDOM},\tilde{d}}(l)$ 知

$\bar{\Omega}_{\text{DOM},d}(l-1)$ 指定的跨越球上的两个点的大圆的劣弧进行平滑。显然,通过用平滑因子 α_Ω 计算经指数加权的移动平均数,独立地平滑方位角和倾斜角。对于倾斜角,这得到下面的平滑操作:

$$\begin{aligned} [0271] \bar{\theta}_{\text{DOM},f_{\mathcal{A}},l(\tilde{d})}(l) &= (1 - \alpha_\Omega) \cdot \bar{\theta}_{\text{DOM},f_{\mathcal{A}},l(\tilde{d})}(l-1) + \alpha_\Omega \cdot \theta_{\text{DOM},\tilde{d}}(l), \\ &1 \leq \tilde{d} \leq \tilde{D}(l) \end{aligned} \quad (83)$$

[0272] 对于方位角,必须修改平滑以在从 $\pi - \varepsilon$ ($\varepsilon > 0$) 到 $-\pi$ 的平移时以及在相反方向的平移时得到正确的平滑。可以对此进行考虑,通过首先将以 2π 为模的差分角计算为

$$[0273] \Delta_{\phi,[0,2\pi],\tilde{d}}(l) := [\phi_{\text{DOM},\tilde{d}}(l) - \bar{\phi}_{\text{DOM},f_{\mathcal{A}},l(\tilde{d})}(l-1)] \bmod 2\pi \quad (84)$$

[0274] 其通过下式被转换到区间 $[-\pi, \pi[$

$$[0275] \Delta_{\phi,[-\pi,\pi],\tilde{d}}(l) := \begin{cases} \Delta_{\phi,[0,2\pi],\tilde{d}}(l) & \text{对于 } \Delta_{\phi,[0,2\pi],\tilde{d}}(l) < \pi \\ \Delta_{\phi,[0,2\pi],\tilde{d}}(l) - 2\pi & \text{对于 } \Delta_{\phi,[0,2\pi],\tilde{d}}(l) \geq \pi \end{cases} \quad (85)$$

[0276] 这个以 2π 为模的平滑后的主方位角被确定为

$$[0277] \bar{\phi}_{\text{DOM},[0,2\pi],\tilde{d}}(l) := [\bar{\phi}_{\text{DOM},\tilde{d}}(l-1) + \alpha_\Omega \cdot \Delta_{\phi,[-\pi,\pi],\tilde{d}}(l)] \bmod 2\pi \quad (86)$$

[0278] 并且通过下式最终被转换成位于区间 $[-\pi, \pi[$ 内

$$[0279] \bar{\phi}_{\text{DOM},\tilde{d}}(l) = \begin{cases} \bar{\phi}_{\text{DOM},[0,2\pi],\tilde{d}}(l) & \text{对于 } \bar{\phi}_{\text{DOM},[0,2\pi],\tilde{d}}(l) < \pi \\ \bar{\phi}_{\text{DOM},[0,2\pi],\tilde{d}}(l) - 2\pi & \text{对于 } \bar{\phi}_{\text{DOM},[0,2\pi],\tilde{d}}(l) \geq \pi \end{cases} \quad (87)$$

[0280] 在 $\tilde{D}(l) < D$ 的情况下,存在未取得分配的当前主方向的先前帧中的方向

$\bar{\Omega}_{\text{DOM},d}(l-1)$ 对应的索引集合被表示为

$$[0281] \mathcal{M}_{\text{NA}}(l) := \{1, \dots, D\} \setminus \{f_{\mathcal{A},l}(\tilde{d}) \mid 1 \leq \tilde{d} \leq D\} \quad (88)$$

[0282] 从上一帧复制相应的方向,亦即,对于 $d \in \mathcal{M}_{\text{NA}}(l)$

$$[0283] \bar{\Omega}_{\text{DOM},d}(l) = \bar{\Omega}_{\text{DOM},d}(l-1) \quad (89)$$

[0284] 对预定数量 (L_{IA}) 的帧未分配的方向被称为是不活动的。

[0285] 之后,计算通过 $\mathcal{M}_{\text{ACT}}(l)$ 表示的活动的方向的索引集合。其基数表示为 $D_{\text{ACT}}(l) := |\mathcal{M}_{\text{ACT}}(l)|$ 。

[0286] 然后,将所有平滑后的方向连接成单个方向矩阵,作为

$$[0287] \bar{\Omega}_{\text{DOM}}(l) := [\bar{\Omega}_{\text{DOM},1}(l) \quad \bar{\Omega}_{\text{DOM},2}(l) \quad \dots \quad \bar{\Omega}_{\text{DOM},D}(l)] \quad (90)$$

[0288] 方向信号的计算

[0289] 方向信号的计算基于模式匹配。具体地,对于那些HOA表示得到给出的HOA信号的最佳近似的方向信号进行搜索。因为相继帧之间的方向的改变会导致方向信号的不连续性,所以可以计算重叠帧的方向信号的估计,继之以使用适当的窗口函数平滑相继的重叠

帧的结果。然而,该平滑引入单个帧的等待时间。

[0290] 下面解释关于方向信号的详细估计:

[0291] 首先,根据下式计算基于平滑后的活动的方向的模式矩阵

$$[\mathbf{E}_{\text{ACT}}(l) := [\mathbf{s}_{\text{DOM},d_{\text{ACT},1}}(l) \quad \mathbf{s}_{\text{DOM},d_{\text{ACT},2}}(l) \quad \dots \quad \mathbf{s}_{\text{DOM},d_{\text{ACT},D_{\text{ACT}}(l)}}(l)] \in \mathbb{R}^{O \times D_{\text{ACT}}(l)} \quad (91)$$

[0293] 其中,

$$[\mathbf{s}_{\text{DOM},d}(l) := [s_0^0(\bar{\mathbf{a}}_{\text{DOM},d}(l)), s_1^{-1}(\bar{\mathbf{a}}_{\text{DOM},d}(l)), s_1^0(\bar{\mathbf{a}}_{\text{DOM},d}(l)), \dots, s_N^0(\bar{\mathbf{a}}_{\text{DOM},d}(l))]^T \in \mathbb{R}^O \quad (92)$$

[0295] 其中, $d_{\text{ACT},j}, 1 \leq j \leq D_{\text{ACT}}(l)$ 表示活动的方向的索引。

[0296] 接下来,计算包含关于第(1-1)个和第1个帧的所有方向信号的非平滑的估计的矩阵 $\mathbf{X}_{\text{INST}}(l)$:

$$[\mathbf{X}_{\text{INST}}(l) := [\mathbf{x}_{\text{INST}}(l, 1) \quad \mathbf{x}_{\text{INST}}(l, 2) \quad \dots \quad \mathbf{x}_{\text{INST}}(l, 2B)] \in \mathbb{R}^{D \times 2B} \quad (93)$$

[0298] 其中,

$$[\mathbf{x}_{\text{INST}}(l, j) = [x_{\text{INST},1}(l, j), x_{\text{INST},2}(l, j), \dots, x_{\text{INST},D}(l, j)]^T \in \mathbb{R}^D, 1 \leq j \leq 2B \quad (94)$$

[0300] 这在两个步骤中完成。在第一步骤中,将对应于不活动的方向的行中的方向信号样本设置成零,亦即

$$[x_{\text{INST},d}(l, j) = 0 \quad \forall 1 \leq j \leq 2B, \quad \text{如果 } d \notin \mathcal{M}_{\text{ACT}}(l) \quad (95)$$

[0302] 在第二步骤中,通过首先根据下式将对应于活动的方向的方向信号样本安排在矩阵中来得到它们

$$[\mathbf{X}_{\text{INST,ACT}}(l) := \begin{bmatrix} x_{\text{INST},d_{\text{ACT},1}}(l, 1) & x_{\text{INST},d_{\text{ACT},1}}(l, 2B) \\ \vdots & \ddots \\ x_{\text{INST},d_{\text{ACT},D_{\text{ACT}}(l)}}(l, 1) & x_{\text{INST},d_{\text{ACT},D_{\text{ACT}}(l)}}(l, 2B). \end{bmatrix} \quad (96)$$

[0304] 然后计算该矩阵,以便将误差的欧几里得范数

[0305] $\mathbf{E}_{\text{ACT}}(l) \mathbf{X}_{\text{INST,ACT}}(l) - [\mathbf{C}(l-1) \mathbf{C}(l)]$ (97) 最小化。其解通过下式给出

$$[\mathbf{X}_{\text{INST,ACT}}(l) = [\mathbf{E}_{\text{ACT}}^T(l) \mathbf{E}_{\text{ACT}}(l)]^{-1} \mathbf{E}_{\text{ACT}}^T(l) [\mathbf{C}(l-1) \quad \mathbf{C}(l)] \quad (98)$$

[0307] 通过适当的窗口函数 $w(j)$ 对方向信号 $x_{\text{INST},d}(l, j)$ ($1 \leq d \leq D$) 的估计进行窗口处理:

$$[\mathbf{x}_{\text{INST,WIN},d}(l, j) := x_{\text{INST},d}(l, j) \cdot w(j), 1 \leq j \leq 2B \quad (99)$$

[0309] 关于窗口函数的示例由周期汉明窗口给出,定义如下

$$[w(j) := \begin{cases} K_w \left[0.54 - 0.46 \cos \left(\frac{2\pi j}{2B+1} \right) \right] & \text{对于 } 1 \leq j \leq 2B \\ 0 & \text{否则} \end{cases} \quad (100)$$

[0311] 其中, K_w 表示被确定为使得移位后的窗口的和等于“1”的缩放因子。根据下式通过进行了窗口处理的非平滑的估计的适当重叠来计算第(1-1)个帧的平滑后的方向信号

$$[x_d((l-1)B+j) = \mathbf{x}_{\text{INST,WIN},d}(l-1, B+j) + \mathbf{x}_{\text{INST,WIN},d}(l, j) \quad (101)$$

[0313] 对第(1-1)个帧的所有平滑后的方向信号的样本布置在矩阵 $\mathbf{X}(l-1)$ 中,如下

$$[\mathbf{X}(l-1) := [\mathbf{x}((l-1)B+1) \quad \mathbf{x}((l-1)B+2) \quad \dots \quad \mathbf{x}((l-1)B+B)] \in \mathbb{R}^{D \times B} \quad (102)$$

[0315] 其中,

$$[\mathbf{x}(j) = [x_1(j), x_2(j), \dots, x_D(j)]^T \in \mathbb{R}^D \quad (103)$$

[0317] 环境HOA分量的计算

[0318] 根据下式通过从总的HOA表示C (l-1) 减去总的方向HOA分量C_{DIR} (l-1) 得到环境HOA分量C_A (l-1)

$$[0319] \mathbf{C}_A(l-1) := \mathbf{C}(l-1) - \mathbf{C}_{DIR}(l-1) \in \mathbb{R}^{O \times B} \quad (104)$$

[0320] 其中,通过下式确定C_{DIR} (l-1)

$$[0321] \mathbf{C}_{DIR}(l-1) := \mathbf{\Xi}_{DOM}(l-1) \begin{bmatrix} x_{INST,WIN,1}(l-1, B+1) & \cdots & x_{INST,WIN,1}(l-1, 2B) \\ \vdots & \ddots & \vdots \\ x_{INST,WIN,D}(l-1, B+1) & \cdots & x_{INST,WIN,D}(l-1, 2B) \end{bmatrix} + \mathbf{\Xi}_{DOM}(l) \begin{bmatrix} x_{INST,WIN,1}(l, 1) & \cdots & x_{INST,WIN,1}(l, B) \\ \vdots & \ddots & \vdots \\ x_{INST,WIN,D}(l, 1) & \cdots & x_{INST,WIN,D}(l, B) \end{bmatrix}, \quad (105)$$

[0322] 其中, $\mathbf{\Xi}_{DOM}(l)$ 表示通过下式定义的基于所有平滑的方向的模式矩阵

$$[0323] \mathbf{\Xi}_{DOM}(l) := [\mathbf{S}_{DOM,1}(l) \ \mathbf{S}_{DOM,2}(l) \ \dots \ \mathbf{S}_{DOM,D}(l)] \in \mathbb{R}^{O \times D} \quad (106)$$

[0324] 因为总的方向HOA分量的计算还基于重叠的相继瞬间总方向HOA分量的空间平滑, 还得到具有单个帧的等待时间的环境HOA分量。

[0325] 环境HOA分量的阶降低

[0326] 通过C_A (l-1) 的分量将其表示为

$$[0327] \mathbf{C}_A(l-1) = \begin{bmatrix} c_{0,A}^0((l-1)B+1) & \cdots & c_{0,A}^0((l-1)B+B) \\ \vdots & \ddots & \vdots \\ c_{N,A}^N((l-1)B+1) & \cdots & c_{N,A}^N((l-1)B+B) \end{bmatrix}, \quad (107)$$

[0328] 通过删去所有n>N_{RED}的HOA系数c_{n,A}^m(j)完成阶降低:

$$[0329] \mathbf{C}_{A,RED}(l-1) := \begin{bmatrix} c_{0,A}^0((l-1)B+1) & \cdots & c_{0,A}^0((l-1)B+B) \\ \vdots & \ddots & \vdots \\ c_{N_{RED},A}^{N_{RED}}((l-1)B+1) & \cdots & c_{N_{RED},A}^{N_{RED}}((l-1)B+B) \end{bmatrix} \in \mathbb{R}^{O_{RED} \times B} \quad (108)$$

[0330] 环境HOA分量的球谐函数变换

[0331] 通过降低了阶的环境HOA分量C_{A,RED} (l) 与模式矩阵的逆的相乘执行球谐函数变换

$$[0332] \mathbf{\Xi}_A := [\mathbf{S}_{A,1} \ \mathbf{S}_{A,2} \ \dots \ \mathbf{S}_{A,O_{RED}}] \in \mathbb{R}^{O_{RED} \times O_{RED}} \quad (109)$$

[0333] 其中,

$$[0334] \mathbf{S}_{A,d} := [S_0^0(\Omega_{A,d}), S_1^{-1}(\Omega_{A,d}), S_1^0(\Omega_{A,d}), \dots, S_{N_{RED}}^{N_{RED}}(\Omega_{A,d})]^T \in \mathbb{R}^{O_{RED}} \quad (110)$$

[0335] 基于O_{RED}是均匀分布的方向 $\Omega_{A,d}$

$$[0336] 1 \leq d \leq O_{RED}: \mathbf{W}_{A,RED}(l) = (\mathbf{\Xi}_A)^{-1} \mathbf{C}_{A,RED}(l) \quad (111)$$

[0337] 解压缩

[0338] 逆球谐函数变换

[0339] 经由逆球谐函数变换通过下式将经感知解压缩的空间域信号 $\widehat{\mathbf{W}}_{A,RED}(l)$ 变换成阶为N_{RED}的HOA域表示 $\widehat{\mathbf{C}}_{A,RED}(l)$

$$[0340] \widehat{\mathbf{C}}_{A,RED}(l) = \mathbf{\Xi}_A \widehat{\mathbf{W}}_{A,RED}(l) \quad (112)$$

[0341] 阶扩展

[0342] 根据下式通过附加零将HOA表示 $\hat{\mathbf{C}}_{A,RED}(l)$ 的高保真度立体声响复制阶扩展成N

$$[0343] \quad \hat{\mathbf{C}}_A(l) := \begin{bmatrix} \hat{\mathbf{C}}_{A,RED}(l) \\ \mathbf{0}_{(O-O_{RED}) \times B} \end{bmatrix} \in \mathbb{R}^{O \times B}, \quad (113)$$

[0344] 其中, $\mathbf{0}_{m \times n}$ 表示具有m行和n列的零矩阵。

[0345] HOA系数组成

[0346] 最终的解压缩后的HOA系数根据下式由方向和环境HOA分量相加组成

$$[0347] \quad \hat{\mathbf{C}}(l-1) := \hat{\mathbf{C}}_A(l-1) + \hat{\mathbf{C}}_{DIR}(l-1) \quad (114)$$

[0348] 在该阶段,再次引入单个帧的等待时间以允许基于空间平滑计算方向HOA分量。由此,避免了在声场的方向分量中由相继帧之间的方向改变导致的可能的不期望的不连续性。

[0349] 为了计算平滑后的方向HOA分量,将包含所有单独方向信号的估计的两个相继帧连接成单个长帧,如下

$$[0350] \quad \hat{\mathbf{X}}_{INST}(l) := [\hat{\mathbf{X}}(l-1) \quad \hat{\mathbf{X}}(l)] \in \mathbb{R}^{D \times 2B} \quad (115)$$

[0351] 在该长帧中包含的每个单独信号选段乘以例如等式(100)的窗口函数。当按照下式通过长帧 $\hat{\mathbf{X}}_{INST}(l)$ 的分量表示该长帧时

$$[0352] \quad \hat{\mathbf{X}}_{INST}(l) = \begin{bmatrix} \hat{x}_{INST,1}(l, 1) & \hat{x}_{INST,1}(l, 2B) \\ \vdots & \ddots \quad \vdots \\ \hat{x}_{INST,D}(l, 1) & \hat{x}_{INST,D}(l, 2B) \end{bmatrix} \quad (116)$$

[0353] 可以将窗口处理操作用公式表示为计算经窗口处理的信息选段 $\hat{x}_{INST,WIN,d}(l, j)$, $1 \leq d \leq D$, 如下

$$[0354] \quad \hat{x}_{INST,WIN,d}(l, j) = \hat{x}_{INST,d}(l, j) \cdot w(j), \quad 1 \leq j \leq 2B, \quad 1 \leq d \leq D \quad (117)$$

[0355] 最后,通过将所有经窗口处理的方向信号选段编码成适当的方向并且以重叠的方式将它们重叠,得到总的方向HOA分量 $C_{DIR}(l-1)$:

$$\begin{aligned} [0356] \quad \hat{\mathbf{C}}_{DIR}(l-1) &= \Xi_{DOM}(l-1) \begin{bmatrix} \hat{x}_{INST,WIN,1}(l-1, B+1) & \hat{x}_{INST,WIN,1}(l-1, 2B) \\ \vdots & \ddots \quad \vdots \\ \hat{x}_{INST,WIN,D}(l-1, B+1) & \hat{x}_{INST,WIN,D}(l-1, 2B) \end{bmatrix} \\ &+ \Xi_{DOM}(l) \begin{bmatrix} \hat{x}_{INST,WIN,1}(l, 1) & \hat{x}_{INST,WIN,1}(l, B) \\ \vdots & \ddots \quad \vdots \\ \hat{x}_{INST,WIN,D}(l, 1) & \hat{x}_{INST,WIN,D}(l, B) \end{bmatrix}. \end{aligned} \quad (118)$$

[0357] 方向搜索算法的解释

[0358] 下面,解释在主方向估计部分中所述的方向搜索处理之后的动机。其基于首先定义的一些假设。

[0359] 假设

[0360] HOA系数矢量 $c(j)$ 通常通过下式与时域幅度密度函数 $d(j, \Omega)$ 相关

$$[0361] \quad \mathbf{c}(j) = \int_{\mathcal{S}^2} d(j, \Omega) \mathbf{S}(\Omega) d\Omega \quad (119)$$

[0362] 假设HOA系数矢量 $c(j)$ 符合以下模型:

$$[0363] \quad c(j) = \sum_{i=1}^I x_i(j) \mathbf{S}(\Omega_{x_i}(l)) + c_A(j) \quad \text{对于 } lB + 1 \leq j \leq (l+1)B \quad (120)$$

[0364] 该模型表明,一方面,HOA系数矢量 $c(j)$ 通过来自第1个帧的方向 $\Omega_{x_i}(l)$ 的 I 个主方向源信号 $x_i(j)$ ($1 \leq i \leq I$) 创建。具体地,假设对于单个帧的持续时间,方向是固定的。假设主源信号的数量 I 明显地小于HOA系数的总数量 0 。另外,假设帧长度 B 明显地大于 0 。另一方面,矢量 $c(j)$ 由残差分量 $c_A(j)$ 组成,可以将其视为表示理想的各向同性环境声场。

[0365] 假设单独的HOA系数矢量分量具有以下性质:

[0366] ●假设主源信号是零平均值,亦即

$$[0367] \quad \sum_{j=lB+1}^{(l+1)B} x_i(j) \approx 0 \quad \forall 1 \leq i \leq I \quad (121)$$

[0368] 并且假设主源信号彼此无关,亦即

$$[0369] \quad \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_i(j) x_{i'}(j) \approx \delta_{i-i'} \bar{\sigma}_{x_i}^2(l) \quad \forall 1 \leq i, i' \leq I \quad (122)$$

[0370] 其中 $\bar{\sigma}_{x_i}^2(l)$ 表示第1个帧的第 i 个信号的平均功率。

[0371] ●假设主源信号与HOA系数矢量的环境分量无关,亦即

$$[0372] \quad \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_i(j) c_A(j) \approx 0 \quad \forall 1 \leq i \leq I \quad (123)$$

[0373] ●假设环境HOA分量矢量是零平均值,并且假设其具有协方差矩阵

$$[0374] \quad \Sigma_A(l) := \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} c_A(j) c_A^T(j) \quad (124)$$

[0375] ●每个帧 l 的方向对环境功率比DAR(l)在此处通过下式定义

$$[0376] \quad \text{DAR}(l) := 10 \log_{10} \left[\frac{\max_{1 \leq i \leq I} \bar{\sigma}_{x_i}^2(l)}{\|\Sigma_A(l)\|^2} \right] \quad (125)$$

[0377] 假设其大于预定义的期望值 DAR_{MIN} ,亦即

$$[0378] \quad \text{DAR}(l) \geq \text{DAR}_{\text{MIN}} \quad (126)$$

[0379] 方向搜索的解释

[0380] 为了进行解释,考虑以下情况:仅基于第1个帧的样本而不考虑 $L-1$ 个先前帧的样本,计算相关性矩阵 $B(l)$ (参见等式(67))。该操作对应于设置 $L=1$ 。因此,相关性矩阵可以表示为

$$[0381] \quad B(l) = \frac{1}{B} C(l) C^T(l) \quad (127)$$

$$= \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} c(j) c^T(j) \quad (128)$$

[0382] 通过将等式(120)中的模型假设替换到等式(128)中,并且通过使用等式(122)和(123)以及等式(124)中的定义,可以将相关性矩阵 $B(l)$ 近似为(129)

$$\begin{aligned}
 \mathbf{B}(l) &= \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} [\sum_{i=1}^l x_i(j) \mathbf{S}(\Omega_{x_i}(l)) + \mathbf{c}_A(j)] [\sum_{i'=1}^l x_{i'}(j) \mathbf{S}(\Omega_{x_{i'}}(l)) + \mathbf{c}_A(j)]^T \\
 &= \sum_{i=1}^l \sum_{i'=1}^l \mathbf{S}(\Omega_{x_i}(l)) \mathbf{S}^T(\Omega_{x_{i'}}(l)) \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_i(j) x_{i'}(j) \\
 [0383] \quad &+ \sum_{i=1}^l \mathbf{S}(\Omega_{x_i}(l)) \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_i(j) \mathbf{c}_A^T(j) + \sum_{i'=1}^l \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} x_{i'}(j) \mathbf{c}_A(j) \mathbf{S}^T(\Omega_{x_{i'}}(l)) \\
 &+ \frac{1}{B} \sum_{j=lB+1}^{(l+1)B} \mathbf{c}_A(j) \mathbf{c}_A^T(j) \tag{130}
 \end{aligned}$$

$$\approx \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) \mathbf{S}(\Omega_{x_i}(l)) \mathbf{S}^T(\Omega_{x_i}(l)) + \Sigma_A(l) . \tag{131}$$

[0384] 根据等式(131)可以看出, $\mathbf{B}(l)$ 近似地由对方向和环境HOA分量有贡献的两个附加分量组成。其 $\mathcal{I}(l)$ 秩近似 $\mathbf{B}_J(l)$ 提供方向HOA分量的近似, 亦即

$$[0385] \quad \mathbf{B}_J(l) \approx \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) \mathbf{S}(\Omega_{x_i}(l)) \mathbf{S}^T(\Omega_{x_i}(l)) \tag{132}$$

[0386] 其根据关于方向对环境功率比的等式(126)得出。

[0387] 然而, 应当强调的是, $\Sigma_A(l)$ 的一部分将不可避免地漏到 $\mathbf{B}_J(l)$ 中, 因为 $\Sigma_A(l)$ 一般具有完整的秩, 因此矩阵的列 $\sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) \mathbf{S}(\Omega_{x_i}(l)) \mathbf{S}^T(\Omega_{x_i}(l))$ 和 $\Sigma_A(l)$ 跨过的子空间彼此不正交。通过等式(132), 用于主方向搜索的等式(77)中的矢量 $\sigma^2(l)$ 可以表示为

$$\sigma^2(l) = \text{diag}(\mathbf{\Xi}^T \mathbf{B}_J(l) \mathbf{\Xi}) \tag{133}$$

$$[0388] \quad = \text{diag} \left(\begin{bmatrix} \mathbf{S}^T & (\Omega_1) \mathbf{B}_J(l) \mathbf{S}(\Omega_1) \\ \vdots & \ddots \\ \mathbf{S}^T & (\Omega_Q) \mathbf{B}_J(l) \mathbf{S}(\Omega_1) \end{bmatrix} \right) \approx \tag{134}$$

$$\begin{aligned}
 [0389] \quad &\text{diag} \left(\begin{bmatrix} \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) v_N^2(\angle(\Omega_1, \Omega_{x_i})) & \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) v_N(\angle(\Omega_1, \Omega_{x_i})) v_N(\angle(\Omega_{x_i}, \Omega_Q)) \\ \vdots & \vdots \\ \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) v_N(\angle(\Omega_Q, \Omega_{x_i})) v_N(\angle(\Omega_{x_i}, \Omega_1)) & \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) v_N^2(\angle(\Omega_Q, \Omega_{x_i})) \end{bmatrix} \right) \\
 &= \left[\sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) v_N^2(\angle(\Omega_1, \Omega_{x_i})) \dots \sum_{i=1}^l \bar{\sigma}_{x_i}^2(l) v_N^2(\angle(\Omega_Q, \Omega_{x_i})) \right]^T . \tag{136}
 \end{aligned}$$

[0390] 在等式(135)中, 使用在等式(47)中示出的球谐函数的以下属性:

$$[0391] \quad S^T(\Omega_q) S(\Omega_{q'}) = v_N(\angle(\Omega_q, \Omega_{q'})) \tag{137}$$

[0392] 等式(136)示出, $\sigma^2(l)$ 的 $\sigma_q^2(l)$ 个分量是来自测试方向 Ω_q ($1 \leq q \leq Q$) 的信号的功率的近似。

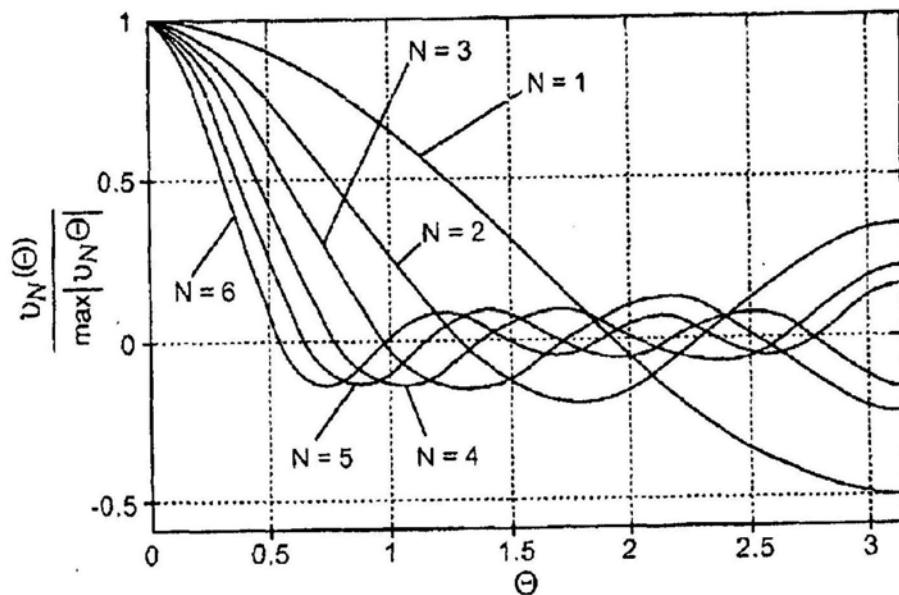


图1

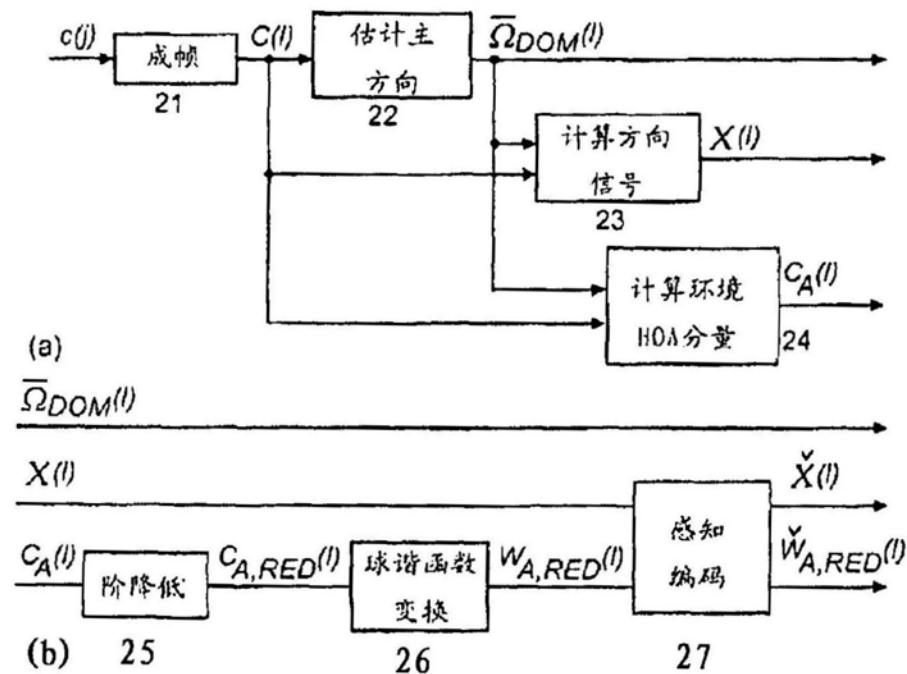


图2

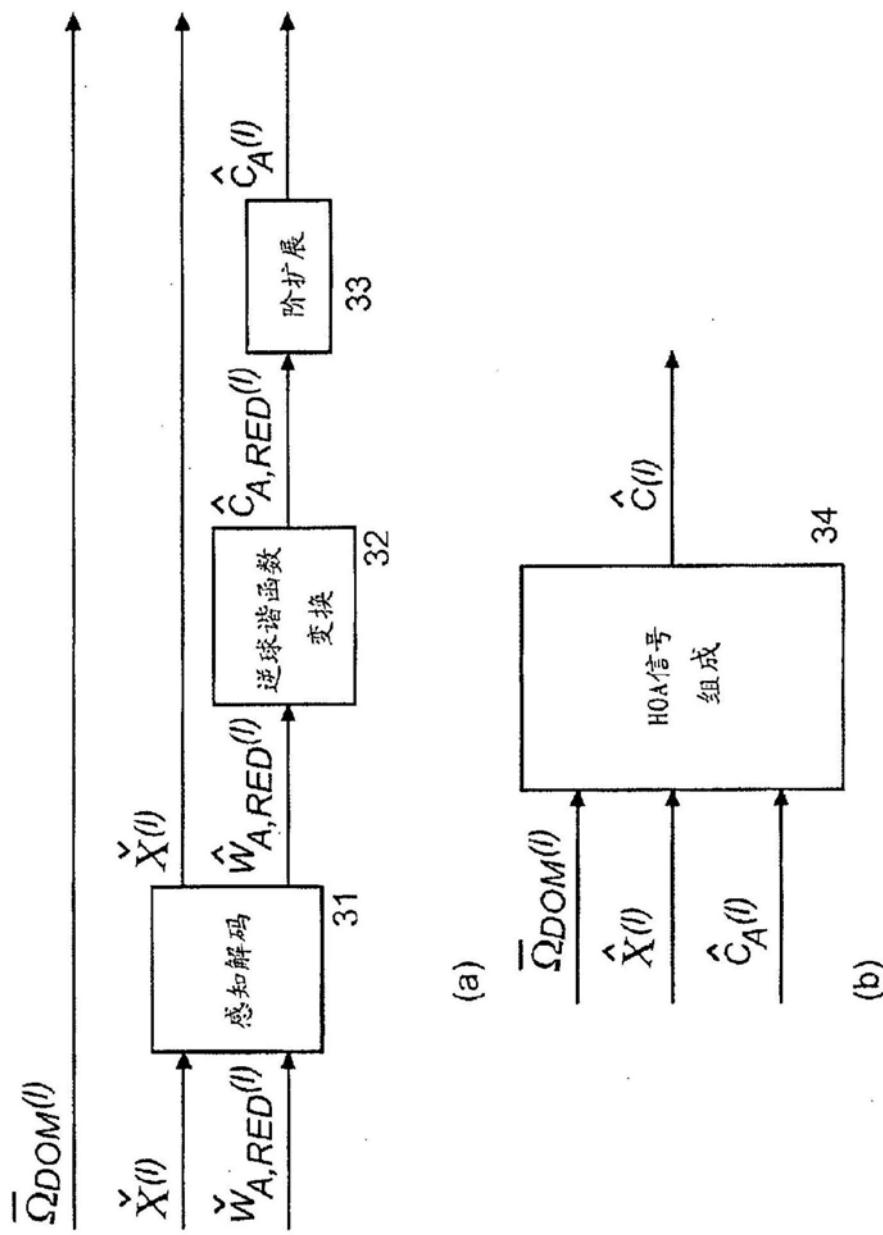


图3