(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 107403625 B
(45)授权公告日 2021.06.04

(21)申请号 201710829618.2
(22)申请日 2013.07.16
(65)同一申请的已公布的文献号
申请公布号 CN 107403625 A
(43)申请公布日 2017.11.28
(30)优先权数据
12305861.2 2012.07.16 EP
(62)分案原申请数据
201380036698.6 2013.07.16
(73)专利权人 杜比国际公司
地址 荷兰阿姆斯特丹
(72)发明人 J.贝姆 S.科唐 A.克鲁格 P.贾克斯
(74)专利代理机构 中国贸促会专利商标事务所有限公司 11038
代理人 张劲松
(51)Int.Cl.
G10L 19/008(2013.01)
H04S 3/02(2006.01)
(56)对比文件
US 2012155653 A1,2012.06.21
CN 107591159 A,2018.01.16
CN 102547549 A,2012.07.04
CN 102318372 A,2012.01.11
WO 2009081406 A3,2009.10.01
CN 1809872 A,2006.07.26
审查员 涂亮梅

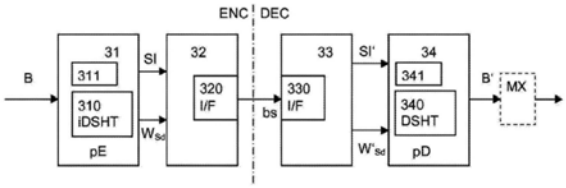
权利要求书1页 说明书17页 附图5页

(54)发明名称

用于对HOA音频信号进行解码的方法、设备和计算机可读介质

(57)摘要

本发明公开了用于对HOA音频信号进行解码的方法、设备和计算机可读介质。一种用于对多信道HOA音频信号进行编码以便降噪的方法，包括以下步骤：使用逆自适应DSHT对信道进行解相关(81)，所述逆自适应DSHT包括旋转运算(330)和逆DSHT(810)，所述旋转运算旋转iDSHT的空间采样网格；对每个经解相关的信道进行感知编码(82)；对旋转信息(SI)进行编码，所述旋转信息包括定义所述旋转运算的参数；以及，传输或存储经感知编码的信道和经编码的旋转信息。



CN 107403625 B

1.一种用于对高阶高保真度立体声响复制HOA音频信号进行解码的方法,所述方法包括:

基于感知解码对HOA音频信号进行解压缩以至少确定与所述HOA音频信号对应的HOA表示,该HOA表示指代经感知解码的信道;

使用自适应DSHT对HOA表示进行空间解码,所述空间解码包括:通过根据接收的信道旋转信息旋转自适应DSHT的默认的球面样本网格来确定旋转的变换,以及基于所述旋转的变换和HOA表示确定旋转的HOA表示,其中,所述旋转的HOA表示基于所述旋转的变换和HOA表示的相乘而确定;以及

对所述旋转的HOA表示进行矩阵化,其中获得映射到扬声器位置的可再现的音频信号。

2.一种用于对高阶高保真度立体声响复制HOA音频信号进行解码的设备,所述设备包括:

解码器,所述解码器被配置为:

基于感知解码对HOA音频信号进行解压缩以至少确定与所述HOA音频信号对应的HOA表示,该HOA表示指代经感知解码的信道;

使用自适应DSHT对HOA表示进行空间解码,所述空间解码包括:通过根据接收的信道旋转信息旋转自适应DSHT的默认的球面样本网格来确定旋转的变换,以及基于所述旋转的变换和HOA表示确定旋转的HOA表示,其中,所述旋转的HOA表示基于所述旋转的变换和HOA表示的相乘而确定;以及

对所述旋转的HOA表示进行矩阵化,其中获得映射到扬声器位置的可再现的音频信号。

3.一种包含指令的非暂时性计算机可读介质,所述指令在被处理器运行时执行如权利要求1所述的方法。

4.一种设备,包括:

一个或多个处理器,

一个或多个非暂时性计算机可读存储介质,所述非暂时性计算机可读存储介质具有存储在其上的指令,所述指令在被所述一个或多个处理器运行时使所述设备执行如权利要求1所述的方法。

用于对HOA音频信号进行解码的方法、设备和计算机可读介质

[0001] 本申请是基于申请号为201380036698.6、申请日为2013年7月16日、发明名称为“用于对多信道HOA音频信号进行编码以便降噪的方法和设备以及用于对多信道HOA音频信号进行解码以便降噪的方法和设备”的专利申请的分案申请。

技术领域

[0002] 本发明涉及用于对多信道高阶高保真度立体声响复制音频信号进行编码以便降噪的方法和设备、以及对多信道高阶高保真度立体声响复制音频信号进行解码以便降噪的方法和设备。

背景技术

[0003] 高阶高保真度立体声响复制(Higher Order Ambisonics,HOA)是多信道声场表示[4],并且HOA信号是多信道音频信号。在特定扬声器装配上回放某些多信道音频信号表示,特别是HOA表示,需要特殊的呈现,这通常包括矩阵化运算。在解码之后,高保真度立体声响复制(Ambisonics)信号被“矩阵化”,亦即,被映射到与例如扬声器的实际空间位置相对应的新的音频信号。通常,在单个信道之间存在高的互相关性。

[0004] 问题是经历到在矩阵化运算之后编码噪声增大。在现有技术下,原因似乎是未知的。当在通过感知编码器进行压缩之前例如通过离散球面谐波变换(Discrete Spherical Harmonics Transform,DSHT)将HOA信号变换到空间域时,也发生该效应。

[0005] 用于高阶高保真度立体声响复制音频信号表示的压缩的通常方法是将独立的感知编码器应用于个体高保真度立体声响复制系数信道[7]。具体地,感知编码器仅考虑对在每个个体单信道信号中发生的噪声掩蔽效应进行编码。然而,这种效应典型地是非线性的。如果将这种单信道矩阵化成新信号,则可能发生噪声去掩蔽(unmasking)。当在用感知编码器进行压缩之前通过离散球面谐波变换将高阶高保真度立体声响复制信号变换到空间域时,也发生该效应[8]。

[0006] 这种多信道音频信号表示的传输或存储通常要求适当的多信道压缩技术。通常,在最终将I个经解码的信号 $\hat{\mathrm{x}}_i(l), i = 1,\ldots,I$ 矩阵化为J个新信号 $\hat{\mathrm{y}}_j(l), j = 1,\ldots,J$ 之前,执行与信道无关的感知解码。术语矩阵化表示以加权的方式添加或混合经解码的信号 $\hat{\hat{\mathrm{x}}}_i(l)$。将所有的信号 $\hat{\mathrm{x}}_i(l), i = 1,\ldots,I$ 以及所有的新信号 $\hat{\mathrm{y}}_j(l), j = 1,\ldots,J$ 布置在根据如下的矢量中:

[0007] $$\hat{\hat{\mathbf{x}}}(l) := \left[\hat{\hat{\mathrm{x}}}_1(l) \ \ldots \ \hat{\hat{\mathrm{x}}}_I(l)\right]^{\mathrm{T}}$$

[0008] $$\hat{\mathbf{y}}(l) := \left[\hat{\mathrm{y}}_1(l) \ \ldots \ \hat{\mathrm{y}}_J(l)\right]^{\mathrm{T}}$$

[0009] 术语“矩阵化”源自以下事实:在数学上通过以下矩阵运算从 $\hat{\hat{\mathrm{x}}}(l)$ 获得 $\hat{\mathrm{y}}(l)$:

[0010] $$\hat{\mathbf{y}}(l) = \boldsymbol{A}\,\hat{\hat{\mathbf{x}}}(l)$$

[0011] 其中,A表示由混合权重(mixing weight)构成的混合矩阵(mixing matrix)。在此同义地使用术语“混合”和“矩阵化”。混合/矩阵化用于呈现任何特定扬声器装配的音频信号的目的。矩阵依赖的特定的个体扬声器装配以及因此用于在运算期间的矩阵化的矩阵在感知编码阶段通常是未知的。

发明内容

[0012] 本发明提供对多信道高阶高保真度立体声响复制音频信号进行编码和/或解码以便获得降噪的改善。具体地,本发明提供对3D音频比率压缩抑制编码噪声解蔽(de-masking)的方式。

[0013] 本发明描述使(不期望的)噪声去掩蔽效应最小化的自适应离散球面谐波变换(aDSHT)的技术。此外,描述如何可以将aDSHT集成在压缩编码器架构中。所描述的技术至少对于HOA信号是特别有利的。本发明的一个优点是减少要传输的边信息(side information)的量。原则上,仅需要传输旋转轴和旋转角。可以通过所传输的信道的数量,间接地用信号通知DSHT采样网格。与需要传输多于一半的相关矩阵的其它方法(例如Karhunen Loève变换(KLT))相比,该边信息的量非常小。

[0014] 根据本发明的一个实施例,用于对多信道HOA音频信号进行编码以便降噪的方法包括以下步骤:使用逆自适应DSHT对信道进行解相关,所述逆自适应DSHT包括旋转运算和逆DSHT(iDSHT),所述旋转运算旋转iDSHT的空间采样网格;对每个经解相关的信道进行感知编码;对旋转信息进行编码,所述旋转信息包括定义所述旋转运算的参数;以及,传输或存储经感知编码的音频信道和经编码的旋转信息。使用逆自适应DSHT对信道进行解相关的步骤在原则上是空间编码步骤。

[0015] 根据本发明的一个实施例,用于对具有降低的噪声的经编码的多信道HOA音频信号进行解码的方法包括以下步骤:接收经编码的多信道HOA音频信号和信道旋转信息;对所接收的数据进行解压缩,其中使用感知解码;使用自适应DSHT(aDSHT)对每个信道进行空间解码,使经感知解码和空间解码的信道相关,其中执行根据所述旋转信息的aDSHT的空间采样网格的旋转;以及,对相关的经感知解码和空间解码的信道进行矩阵化,其中获得映射到扬声器位置的可再现的音频信号。

[0016] 公开一种用于对多信道HOA音频信号进行编码的设备。公开一种用于对多信道HOA音频信号进行解码的设备。

[0017] 一方面,计算机可读介质具有可执行的指令,以使计算机执行包括以上公开的步骤的用于进行编码的方法,或者执行包括以上公开的步骤的用于进行解码的方法。在从属权利要求、下面的描述以及附图中公开本发明的有利的实施例。

附图说明

[0018] 参考附图描述本发明的示例性实施例,附图中:

[0019] 图1示出用于对M个系数的块进行比率压缩的已知的编码器和解码器;

[0020] 图2示出使用传统的DSHT(离散球面谐波变换)和传统的逆DSHT将HOA信号变换到空间域中的已知的编码器和解码器;

[0021] 图3示出使用自适应DSHT和自适应逆DSHT将HOA信号变换到空间域中的编码器和

解码器；

[0022] 图4示出测试信号；

[0023] 图5示出在编码器和解码器构建块中使用的码本的球面采样位置的示例；

[0024] 图6示出信号自适应DSHT构建块(pE和pD)；

[0025] 图7示出本发明的第一实施例；

[0026] 图8示出编码处理和解码处理的流程图；以及

[0027] 图9示出本发明的第二实施例。

具体实施方式

[0028] 图2示出使用逆DSHT将HOA信号变换到空间域中的已知的系统。对信号进行使用iDSHT 21的变换、比率压缩E1/解压缩D1，并且使用DSHT 24重新变换到系数域S24。与此不同，图3示出根据本发明的一个实施例的系统：已知的解决方案的DSHT处理块被替换为分别控制逆自适应DSHT和自适应DSHT的处理块31、34。在比特流bs内传输边信息SI。该系统包括用于对多信道HOA音频信号进行编码的设备的元件以及用于对多信道HOA音频信号进行解码的设备的元件。

[0029] 在一个实施例中，用于对多信道HOA音频信号进行编码以便降噪的设备ENC包括使用逆自适应DSHT(iaDSHT)对信道B进行解相关的解相关器31，所述逆自适应DSHT包括旋转运算单元311和逆DSHT(iDSHT)310。旋转运算单元旋转iDSHT的空间采样网格。解相关器31提供经解相关的信道W_{sd}和包括旋转信息的边信息SI。此外，该设备包括用于对每个经解相关的信道W_{sd}进行感知编码的感知编码器32以及用于对旋转信息进行编码的边信息编码器。旋转信息包括定义所述旋转运算的参数。感知编码器32提供经感知编码的音频信道和经编码的旋转信息，从而降低数据率。最后，用于进行编码的设备包括用于从经感知编码的音频信道和经编码的边信息创建比特流bs以及用于传输或存储比特流bs的接口装置320。

[0030] 用于对具有降低的噪声的多信道HOA音频信号进行解码的设备DEC包括：用于接收经编码的多信道HOA音频信号和信道旋转信息的接口装置330；以及用于对所接收的数据进行解压缩的解压缩模块33，其包括用于对每个信道进行感知解码的感知解码器。解压缩模块33提供所恢复的经感知解码的信道W'_{sd}和所恢复的边信息SI'。此外，用于进行解码的设备包括：使用自适应DSHT(aDSHT)使经感知解码的信道W'_{sd}相关的相关器34，其中执行DSHT和根据所述旋转信息的DSHT的空间采样网格的旋转；以及用于对相关的经感知解码的信道进行矩阵化的混合器MX，其中获得映射到扬声器位置的可再现的音频信号。在相关器34内的DSHT单元340中，至少可以执行aDSHT。在一个实施例中，在网格旋转单元341中完成空间采样网格的旋转，这在原则上重新计算原始的DSHT采样点。在另一实施例中，在DSHT单元340内执行旋转。

[0031] 下面给出定义和描述去掩蔽的数学模型。假设给定的离散时间多信道信号包括I个信道$x_i(m)$，$i=1,\ldots,I$，其中m表示时间样本索引(time sample index)。个体信号可以是实数值或复数值。考虑以时间样本索引$m_{START}+1$开始的M个样本的帧，其中假设个体信号是固定的。根据下式在矩阵$\mathbf{X} \in \mathbb{C}^{I\times M}$内布置对应的样本：

[0032] $X:=[x(m_{START}+1),\ldots,x(m_{START}+M)]$ (1)

[0033] 其中

[0034] $x(l):=[x_1(m),...,x_I(m)]^T$ (2)

[0035] 其中$(\bullet)^T$表示转置。对应的经验相关矩阵由下式给出：

[0036] $\Sigma_X:=XX^H$ (3)

[0037] 其中$(\bullet)^H$表示联合复共轭和转置。

[0038] 现在假设多信道信号帧已被编码，从而在重构时引入编码误差噪声。因此，用$\widehat{\mathbf{X}}$表示的经重构的帧样本的矩阵根据下式由真实样本矩阵X和编码噪声分量E构成：

[0039] $$\widehat{\mathbf{X}} = \mathbf{X} + \mathbf{E} \quad (4)$$

[0040] 其中

[0041] $E:=[e(m_{START}+1),...,e(m_{START}+L)]$ (5)

[0042] 并且

[0043] $e(m):=[e_1(m),...,e_1(m)]^T$ (6)

[0044] 因为假设每个信道已经被独立地编码，所以对于i＝1，...，I，可以假设编码噪声信号$e_i(m)$彼此独立。利用该特性以及噪声信号是零均值的假设，噪声信号的经验相关矩阵由如下对角矩阵给出：

[0045] $$\mathbf{\Sigma_E} = \mathrm{diag}\left(\sigma_{e_1}^2,\ldots,\sigma_{e_I}^2\right) \quad (7)$$

[0046] 这里，$\mathrm{diag}\left(\sigma_{e_1}^2,\ldots,\sigma_{e_I}^2\right)$表示对角矩阵，在其对角线上具有经验噪声信号幂

[0047] $$\sigma_{e_i}^2 = \frac{1}{M}\sum_{m=m_{\mathrm{START}}+1}^{m_{\mathrm{START}}+M}|e_i(m)|^2 \quad (8)$$

[0048] 另外的基本假设是，执行编码使得对于每个信道满足预定义的信噪比(SNR)。在不失一般性的情况下，假设预定义的SNR是对于每个信道相等的，亦即：

[0049] $$\mathrm{SNR}_x = \frac{\sigma_{x_i}^2}{\sigma_{e_i}^2} \quad for\ all\ \ i = 1,\ldots,I \quad (9)$$

[0050] 其中

[0051] $$\sigma_{x_i}^2 := \frac{1}{M}\sum_{m=m_{\mathrm{START}}+1}^{m_{\mathrm{START}}+M}|x_i(m)|^2 \quad (10)$$

[0052] 从现在开始，考虑将经重构的信号矩阵化为J个新信号$y_j(m)$，j＝1，...，J。

[0053] 在不引入任何编码误差的情况下，经矩阵化的信号的样本矩阵可以表示为：

[0054] Y＝AX (11)

[0055] 其中$\mathbf{A} \in \mathbb{C}^{J\times I}$表示混合矩阵，并且其中

[0056] $Y:=[y(m_{START}+1),...,y(m_{START}+M)]$ (12)

[0057] 其中

[0058] $y(m):=[y_1(m),\cdots,y_J(m)]^T$ (13)

[0059] 然而，由于编码噪声，经矩阵化的信号的样本矩阵被给出为：

[0060] $$\widehat{\mathbf{Y}} := \mathbf{Y} + \mathbf{N} \quad (14)$$

[0061] 其中,N是包含经矩阵化的噪声信号的样本的矩阵。其可以表示为:

[0062] $N=AE$ (15)

[0063] $N=[n(m_{START}+1) \ \ldots \ n(m_{START}+M)]$ (16)

[0064] 其中

[0065] $n(m):=[n_1(m) \ \ldots \ n_J(m)]^T$ (17)

[0066] 是在时间样本索引m时的所有的经矩阵化的噪声信号的矢量。

[0067] 使用等式(11),经矩阵化的无噪声信号的经验相关矩阵可以用公式表示为:

[0068] $\Sigma_Y=A\Sigma_X A^H$ (18)

[0069] 因此,作为Σ_Y的对角线上的第j个元素的第j个经矩阵化的无噪声信号的经验幂(empirical power)可以写为:

[0070] $$\sigma_{y_j}^2=\mathbf{a}_j{}^H\boldsymbol{\Sigma}_{\mathbf{X}}\mathbf{a}_j \qquad (19)$$

[0071] 其中a_j是根据下式的A^H的第j列:

[0072] $A^H=[a_1,\ldots,a_J]$ (20)

[0073] 类似地,利用等式(15),经矩阵化的噪声信号的经验相关矩阵可以写为:

[0074] $\Sigma_N=A\Sigma_E A^H$ (21)

[0075] 作为Σ_N的对角线上的第j个元素的第j个经矩阵化的噪声信号的经验幂由下式给出:

[0076] $$\sigma_{n_j}^2=\mathbf{a}_j{}^H\,\boldsymbol{\Sigma}_{\mathbf{E}}\,\mathbf{a}_j \qquad (22)$$

[0077] 因此,对于通过下式定义的经矩阵化的信号的经验SNR,

[0078] $$\mathrm{SNR}_{y_j}:=\frac{\sigma_{y_j}^2}{\sigma_{n_j}^2} \qquad (23)$$

[0079] 可以使用等式(19)和(22)重新用公式表示为:

[0080] $$\mathrm{SNR}_{y_j}=\frac{\mathbf{a}_j{}^H\,\boldsymbol{\Sigma}_{\mathbf{X}}\,\mathbf{a}_j}{\mathbf{a}_j{}^H\,\boldsymbol{\Sigma}_{\mathbf{E}}\,\mathbf{a}_j} \qquad (24)$$

[0081] 通过将Σ_X如下地分解成其对角线分量和非对角线分量:

[0082] $$\boldsymbol{\Sigma}_{\mathbf{X}}=\mathrm{diag}\left(\sigma_{x_1}^2,\ldots,\sigma_{x_I}^2\right)+\boldsymbol{\Sigma}_{\mathbf{X,NG}} \qquad (25)$$

[0083] 以及

[0084] $$\boldsymbol{\Sigma}_{\mathbf{X,NG}}:=\boldsymbol{\Sigma}_{\mathbf{X}}-\mathrm{diag}\left(\sigma_{x_1}^2,\ldots,\sigma_{x_I}^2\right) \qquad (26)$$

[0085] 并且通过利用从假设(7)和(9)以及在所有信道上的SNR常量得到的如下特性:

[0086] $$\mathrm{diag}\left(\sigma_{x_1}^2,\ldots,\sigma_{x_I}^2\right)=SNR_x\cdot\mathrm{diag}\left(\sigma_{e_1}^2,\ldots,\sigma_{e_I}^2\right) \qquad (27)$$

[0087] 最终获得关于经矩阵化的信号的经验SNR的所期望的表达式:

[0088] $$\mathrm{SNR}_{y_j}=\frac{\mathbf{a}_j{}^H\,\mathrm{diag}\left(\sigma_{x_1}^2,,\sigma_{x_I}^2\right)\mathbf{a}_j}{\mathbf{a}_j{}^H\,\boldsymbol{\Sigma}_{\mathbf{E}}\,\mathbf{a}_j}+\frac{\mathbf{a}_j{}^H\,\boldsymbol{\Sigma}_{\mathbf{X,NG}}\,\mathbf{a}_j}{\mathbf{a}_j{}^H\,\boldsymbol{\Sigma}_{\mathbf{E}}\,\mathbf{a}_j} \qquad (28)$$

[0089] $$\mathrm{SNR}_{y_j} = SNR_x \left(1 + \frac{\mathbf{a}_j{}^H \boldsymbol{\Sigma}_{\mathbf{X},\mathrm{NG}}\, \mathbf{a}_j}{\mathbf{a}_j{}^H \,\mathrm{diag}\left(\sigma_{x_1}^2, \ldots, \sigma_{x_I}^2\right)\mathbf{a}_j}\right). \quad (29)$$

[0090] 从该表达式可以看出，从预定义的SNR(SNR_x)，通过乘以取决于信号相关矩阵Σ_X的对角线分量和非对角线分量的项来获得该SNR。具体地，如果信号$x_i(m)$彼此不相关，使得$\Sigma_{X,NG}$变成零矩阵，则经矩阵化的信号的经验SNR等于预定义的SNR，亦即：

[0091] $$\mathrm{SNR}_{y_j} = \mathrm{SNR}_x \quad \text{对于所有的} j = 1, \ldots, J, \ \text{如果} \boldsymbol{\Sigma}_{\mathbf{X},\mathrm{NG}} = \mathbf{0}_{I\times I} \quad (30)$$

[0092] 其中$0_{I\times I}$表示具有I个行和I个列的零矩阵。也就是说，如果信号$x_i(m)$是相关的，则经矩阵化的信号的经验SNR可能偏离预定义的SNR。在最差的情况下，SNR_{y_j}可能比SNR_x低得多。这种现象在此被称为矩阵化时的噪声去掩蔽。

[0093] 下面的部分给出对高阶高保真度立体声响复制(HOA)的简要介绍，并定义要处理的信号(数据率压缩)。

[0094] 高阶高保真度立体声响复制(HOA)基于对在被假设为无声源的所关注的紧密区域内的声场的描述。在该情况下，在时间t时和在所关注的区域内的(以球面坐标的)位置$x=[r,\theta,\phi]^T$处的声压$p(t,x)$的时空行为在物理上完全由齐次波动等式来确定。可以示出，相对于时间的声压的傅立叶变换，亦即，

[0095] $$P(\omega, \boldsymbol{x}) = \mathcal{F}_t\{\, p(t, \boldsymbol{x}) \,\} \quad (31)$$

[0096] 其中ω表示角频率(并且$\mathcal{F}_t\{\}$对应于$\int_{-\infty}^{\infty} p(t,\boldsymbol{x})\, e^{-\omega t} dt$)，

[0097] 可以根据[10]展开为球面谐波级数(SHs)：

[0098] $$P(k\, c_s, \boldsymbol{x}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} A_n^m(k)\ j_n(kr)\ Y_n^m(\theta,\phi) \quad (32)$$

[0099] 在等式(32)中，c_s表示声音的速度，并且$k = \frac{\omega}{c_s}$表示角波数。此外，$j_n(\bullet)$指示第一类的n阶球面贝塞耳(Bessel)函数，$Y_n^m(\cdot)$表示n阶m次球面谐波(SH)。关于声场的完整信息实际上包含在声场系数$A_n^m(k)$内。

[0100] 应当注意到，SHs一般是复数值的函数。然而，通过它们的适当的线性组合，能够获得实数值的函数，并且关于这些函数，能够进行展开。

[0101] 与等式(32)中的压力声场描述相关地，源场(source field)可以被定义为：

[0102] $$D(k\, c_s, \boldsymbol{\Omega}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} B_n^m(k)\ Y_n^m(\boldsymbol{\Omega}) \quad (33)$$

[0103] 其中，源场或幅值密度(amplitude density)[9]$D(k\ c_s, \Omega)$取决于角波数和角方向$\Omega=[\theta,\phi]^T$。源场可以包括远场/近场、离散/连续的源[1]。根据下式[1]，源场系数B_n^m与声场系数A_n^m相关：

[0104] $$A_n^m = \begin{cases} 4\pi i^n B_n^m & \text{对于远场} \\ -i\, k\, h_n^{(2)}(kr_s)\, B_n^m & \text{对于近场}^1 \end{cases} \quad (34)$$

[1] 对于进入的波（与 e^{-ikr} 有关）使用正频率以及第二类球面汉克尔函数 $h_n^{(2)}$。

[0105] 其中 $h_n^{(2)}$ 是第二类的球面汉克尔(Hankel)函数，r_s 是离开原点的源距离。

[0106] 可以在频域或时域中将HOA域中的信号表示为源场或声场系数的逆傅立叶变换。下面的描述将假设使用有限数量的源场系数的时域表示：

[0107] $$b_n^m = i\mathcal{F}_t\{B_n^m\} \quad (35)$$

[0108] 所述有限数量：(33)中的无穷级数在n＝N处被截断。截断对应于空间带宽限制。系数(或HOA信道)的数量由下式给出：

[0109] $O_{3D}=(N+1)^2$ 对于3D (36)

[0110] 或者对于仅仅2D的描述，由 $O_{2D}=2N+1$ 给出。系数 b_n^m 包括用于由扬声器进行稍后再现的一个时间样本m的音频信息。它们可以被存储或传输，并且因此是数据率压缩的主体。系数的单个时间样本m可以由具有 O_{3D} 个元素的矢量b(m)表示：

[0111] $$\boldsymbol{b}(m) := [b_0^0(m), b_1^{-1}(m), b_1^0(m), b_1^1(m), b_2^{-2}(m), \ldots, b_N^N(m)]^T \quad (37)$$

[0112] 并且通过矩阵B表示M个时间样本的块：

[0113] $B:=[b(m_{START}+1), b(m_{START}+2), .., b(m_{START}+M)]$ (38)

[0114] 可以通过圆形谐波的展开来得到声场的二维表示。这可以被看作是使用固定的倾斜 $\theta=\frac{\pi}{2}$、系数的不同加权以及减小到 O_{2D} 个系数(m＝±n)的集合的上述一般描述的特殊情况。因此，所有下面的考虑也适用于2D表示，然后术语球面(sphere)需要替换为术语圆形(circle)。

[0115] 下面描述从HOA系数域到基于信道的空间域的变换，反之亦然。可以对单位球面上的l个离散的空间样本位置 $\Omega_l=[\theta_l, \Phi_l]^T$ 使用时域HOA系数重写等式(33)：

[0116] $$d_{\Omega_l} := \sum_{n=0}^{N}\sum_{m=-n}^{n} b_n^m\, Y_n^m(\boldsymbol{\Omega}_l) \quad (35)$$

[0117] 假设 $L_{sd}=(N+1)^2$ 个球面样本位置 Ω_l，这可以针对HOA数据块B以矢量标记来重写：

[0118] $W=\Psi_i B$ (36)

[0119] 其中，$W:=[w(m_{START}+1), w(m_{START}+2), ..., w(m_{START}+M)]$，并且 $\boldsymbol{w}(m)=\left[d_{\Omega_1}(m), \ldots, d_{\Omega_{L_{sd}}}(m)\right]^T$ 表示 L_{sd} 个多信道信号的单个时间样本，矩阵 $\boldsymbol{\Psi}_i=\left[\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{L_{sd}}\right]^H$，其中矢量 $\boldsymbol{y}_l=[Y_0^0(\boldsymbol{\Omega}_l), Y_1^{-1}(\boldsymbol{\Omega}_l), \ldots, Y_N^N(\boldsymbol{\Omega}_l)]^T$。如果非常规律地选择球面样本位置，则存在矩阵 Ψ_f，其中：

[0120] $\Psi_f\Psi_i=I$, (37)

[0121] 其中，I是 $O_{3D}\times O_{3D}$ 的单位矩阵。然后，到等式(36)的对应变换可以定义为：

[0122] $B=\Psi_f W$ (38)

[0123] 等式(38)将 L_{sd} 个球面信号变换到系数域，并可重写为正向变换(forward

transform)：

[0124] B＝DSHT{W}， (39)

[0125] 其中，DSHT{ }表示离散球面谐波变换。对应的逆变换将O_{3D}系数信号变换到空间域以形成L_{sd}个基于信道的信号，并且等式(36)变成：

[0126] W＝iDSHT{B} (40)

[0127] 这里，离散球面谐波变换的该定义对于关于HOA数据的数据率压缩的考虑是足够的，因为开始于给出的系数B并且仅关注B＝DSHT{iDSHT{B}}的情况。在[2]中给出了离散球面谐波变换的更严格的定义。可以在[3]、[4]、[6]、[5]中回顾DSHT的适当的球面样本位置以及得到这样的位置的过程。在图5中示出采样网格的示例。

[0128] 具体地，图5示出在编码器和解码器构建块pE、pD中使用的码本的球面采样位置的示例，即，在图5a)中对于L_{Sd}＝4，在图5b)中对于L_{Sd}＝9，在图5c)中对于L_{Sd}＝16，并且在图5d)中对于L_{Sd}＝25。

[0129] 下面描述高阶高保真度立体声响复制系数数据的比率压缩和噪声去掩蔽。首先，定义测试信号以强调下面使用的一些特性。

[0130] 位于方向$\boldsymbol{\Omega}_{s_1}$上的单个远场源由M个离散时间样本的矢量$g＝[g(m),...,g(M)]^T$表示，并且可以通过编码由HOA系数的块表示：

[0131] $B_g＝yg^T$ (45)

[0132] 其中，矩阵B_g类似于等式(38)，并且编码矢量

$\boldsymbol{y}=\left[Y_0^{0^*}(\boldsymbol{\Omega}_{s_1}),\ Y_1^{-1^*}(\boldsymbol{\Omega}_{s_1}),\ldots,Y_N^{N^*}(\boldsymbol{\Omega}_{s_1})\right]^T$由在方向$\boldsymbol{\Omega}_{s_1}=\left[\theta_{s_1},\phi_{s_1}\right]^T$上评估的共轭复数球面谐波构成(如果使用实数值的SH，则该共轭无效)。测试信号可以被看作HOA信号的最简单的情况。更复杂的信号由许多这种信号的叠加构成。

[0133] 考虑HOA信道的直接压缩，下面示出为何在HOA系数信道被压缩时出现噪声去掩蔽。实际的HOA数据块B的O_{3D}系数信道的直接压缩和解压缩将引入类似于等式(4)的编码噪声E：

[0134] $$\widehat{\boldsymbol{B}}=\boldsymbol{B}+\mathbf{E} \qquad (46)$$

[0135] 假设如等式(9)中的常量SNR_{B_g}。为了在扬声器上重放该信号，需要呈现该信号。可以将该处理描述为：

[0136] $$\widehat{\boldsymbol{W}}=\boldsymbol{A}\,\widehat{\boldsymbol{B}} \qquad (47)$$

[0137] 其中，解码矩阵$\boldsymbol{A}\in\mathbb{C}^{L\times O_{3D}}$(并且$A^H＝[a_1,...,a_L]$)，并且矩阵$\widehat{\boldsymbol{W}}\in\mathbb{C}^{L\times M}$保持L个扬声器信号的M个时间样本。这类似于(14)。应用所有上述考虑，扬声器信道l的SNR可以描述为(类似于等式(29))：

[0138] $$SNR_{w_l}=SNR_{B_g}\left(1+\frac{\boldsymbol{a}_l{}^H\,\boldsymbol{\Sigma}_{B,\mathrm{NG}}\,\boldsymbol{a}_l}{\boldsymbol{a}_l{}^H\,\mathrm{diag}\left(\sigma_{B_1}^2,\ldots,\sigma_{B_{O_{3D}}}^2\right)\,\boldsymbol{a}_l}\right) \qquad (48)$$

[0139] 其中，$\sigma_{B_o}^2$是第o个对角线元素，并且$\Sigma_{B,NG}$保持：

[0140] $\Sigma_B = B\ B^H$ (49)

[0141] 的非对角线元素。

[0142] 解码矩阵A不应当受到影响(因为其应当能够针对任意的扬声器布局进行解码)，因此矩阵Σ_B需要变成对角线以获得$SNR_{w_l} = SNR_{B_g}$。通过等式(45)和(49)，($B=B_g$)，$\Sigma_B = yg^H gy^H = cyy^H$变成具有常量标量值的非对角线$c=g^T g$。与$SNR_{B_g}$相比，在扬声器信道处的信噪比$SNR_{w_l}$降低。但是由于声源信号g和扬声器布局在编码阶段通常都是未知的，所以系数信道的直接有损压缩可能导致不可控制的去掩蔽效应，特别是对于低数据率。

[0143] 下面描述为何在使用DSHT之后在空间域中压缩HOA系数时出现噪声去掩蔽。

[0144] 在压缩之前使用等式(36)中给出的球面谐波变换将HOA系数数据的当前块B变换到空间域中：

[0145] $W_{Sd} = \Psi_i B$ (50)

[0146] 其中，逆变换矩阵Ψ_i与$L_{Sd} \geq 0_{3D}$个空间样本位置有关，并且空间信号矩阵$\boldsymbol{W}_{SH} \in \mathbb{C}^{L_{Sd} \times M}$。对这些进行压缩和解压缩，并且添加量化噪声(类似于等式(4))：

[0147] $$\widehat{\boldsymbol{W}}_{Sd} = \boldsymbol{W}_{Sd} + \mathbf{E} \qquad (51)$$

[0148] 其中，编码噪声分量E根据等式(5)。再次假设对于所有空间信道均恒定的SNR，即SNR_{Sd}。使用变换矩阵Ψ_f将该信号变换到系数域等式(42)，其具有特性(41)：$\Psi_f \Psi_i = I$。系数的新的块$\widehat{\boldsymbol{B}}$变成：

[0149] $$\widehat{\boldsymbol{B}} = \boldsymbol{\Psi}_f \widehat{\boldsymbol{W}}_{Sd} \qquad (52)$$

[0150] 通过应用解码矩阵$\boldsymbol{A}_D$：$\widehat{\boldsymbol{W}} = \boldsymbol{A}_D \widehat{\boldsymbol{B}}$，将该信号呈现为L个扬声器信号$\widehat{\boldsymbol{W}} \in \mathbb{C}^{L \times M}$。这可以使用(52)和$A = A_D \Psi_f$来重写：

[0151] $$\widehat{\boldsymbol{W}} = \boldsymbol{A}\, \widehat{\boldsymbol{W}}_{Sd} \qquad (53)$$

[0152] 这里，A变成具有$\boldsymbol{A} \in \mathbb{C}^{L \times L_{Sd}}$的混合矩阵。等式(53)应当被看作类似于等式(14)。再次应用所有上述考虑，扬声器信道l的SNR可以被描述为(类似于等式(29))：

[0153] $$SNR_{w_l} = SNR_{s_d} \left(1 + \frac{\boldsymbol{a}_l{}^H \boldsymbol{\Sigma}_{\boldsymbol{W}_{Sd},\mathrm{NG}}\, \boldsymbol{a}_l}{\boldsymbol{a}_l{}^H \operatorname{diag}\left(\sigma^2_{S_{d_1}}, \ldots, \sigma^2_{S_{d_{L_{Sd}}}} \right) \boldsymbol{a}_l} \right) \qquad (54)$$

[0154] 其中，$\sigma^2_{S_{d_l}}$是第l个对角线元素，并且$\boldsymbol{\Sigma}_{\boldsymbol{W}_{Sd},\mathrm{NG}}$保持：

[0155] $$\boldsymbol{\Sigma}_{\boldsymbol{W}_{Sd}} = \boldsymbol{W}_{Sd}\, \boldsymbol{W}_{Sd}^H \qquad (55)$$

[0156] 的非对角线元素。

[0157] 因为决不会影响A_D(由于其应当可以针对任意的扬声器布局来呈现)，并且因此决不会对A有任何影响，所以$\boldsymbol{\Sigma}_{\boldsymbol{W}_{Sd}}$需要变成接近于对角线以保持所期望的SNR：使用来自等式(45)($B=B_g$)的简单测试信号，$\boldsymbol{\Sigma}_{\boldsymbol{W}_{Sd}}$变成：

[0158] $$\boldsymbol{\Sigma}_{W_{Sd}} = c\,\boldsymbol{\Psi}_i\,\boldsymbol{y}\,\boldsymbol{y}^H\,\boldsymbol{\Psi}_i^H \qquad (56)$$

[0159] 其中，$c=g^Tg$恒定。使用固定的球面谐波变换(Ψ_i、Ψ_f固定)，$\boldsymbol{\Sigma}_{W_{Sd}}$可以只在非常稀少的情况下变成对角线，并且变得更差，如上文所述，项$\frac{\boldsymbol{a}_l^H\,\boldsymbol{\Sigma}_{W_{Sd},\mathrm{NG}}\,\boldsymbol{a}_l}{\boldsymbol{a}_l^H\,\mathrm{diag}\left(\sigma^2_{S_{d_1}},\ldots,\sigma^2_{S_{d_{L_{Sd}}}}\right)\boldsymbol{a}_l}$取决于系数信号空间特性。因此，球面域中的HOA系数的低比率有损压缩可能导致SNR的降低和不可控制的去掩蔽效应。

[0160] 本发明的基本思想是通过使用自适应DSHT(aDSHT)来最小化噪声去掩蔽，自适应DSHT由与HOA输入信号的空间特性有关的DSHT的空间采样网格的旋转以及DSHT本身构成。

[0161] 下面描述具有与HOA系数的数量O_{3D}相匹配的许多球面位置L_{Sd}的信号自适应DSHT(aDSHT)，(36)。首先，选择如传统的非自适应DSHT中的默认球面样本网格。对于M个时间样本的块，旋转球面样本网格使得最小化项

[0162] $$\sum_{l=1}^{L_{Sd}}\sum_{j=1}^{L_{Sd}}\left|\Sigma_{W_{Sd}\,l,j}\right| - \sum\left(\sigma^2_{S_{d_1}},\ldots,\sigma^2_{S_{d_{L_{Sd}}}}\right) \qquad (57)$$

[0163] 的对数，其中，$|\Sigma_{W_{Sd}\,l,j}|$是$\boldsymbol{\Sigma}_{W_{Sd}}$(具有矩阵行索引l和列索引j)的元素的绝对值，并且$\sigma^2_{S_{d_l}}$是$\boldsymbol{\Sigma}_{W_{Sd}}$的对角线元素。这等于最小化等式(54)的项$\frac{\boldsymbol{a}_l^H\,\boldsymbol{\Sigma}_{W_{Sd},\mathrm{NG}}\,\boldsymbol{a}_l}{\boldsymbol{a}_l^H\,\mathrm{diag}\left(\sigma^2_{S_{d_1}},\ldots,\sigma^2_{S_{d_{L_{Sd}}}}\right)\boldsymbol{a}_l}$。

[0164] 直观化地，如图4所示，该处理对应于以单个空间样本位置匹配最强的源方向的方式的DSHT的球面采样网格的旋转。使用来自等式(45)($B=B_g$)的简单测试信号，可以示出等式(55)的项W_{Sd}变成矢量$\in \mathbb{C}^{L_{Sd}\times 1}$(其中，除了一个元素之外的所有元素都接近于零)。因此，$\boldsymbol{\Sigma}_{W_{Sd}}$变成接近于对角线，并且可以保持所期望的SNR($SNR_{S_d}$)。

[0165] 图4示出被变换到空间域的测试信号B_g。在图4a)中，使用默认的采样网格，并且在图4b)中，使用aDSHT的旋转的网格。通过对应的样本位置周围的Voronoi单元的颜色/灰度变化示出空间信道的相关的$\boldsymbol{\Sigma}_{W_{Sd}}$的值(以dB)。空间结构的每个单元表示采样点，并且单元的亮度/暗度表示信号强度。如同在图4b)中可以看到的，发现最强的源方向，并且旋转采样网格，使得侧面(side)之一(亦即，单个空间样本位置)匹配最强的源方向。将该侧面图示为白色(对应于强的源方向)，而其它侧面是暗的(对应于低的源方向)。在图4a)中，亦即，在旋转之前，没有侧面匹配最强的源方向，并且若干侧面是更深/更浅的灰色，这意味着在相应的采样点处接收到相当大的(但不是最大的)强度的音频信号。

[0166] 下面描述在压缩编码器和解码器内使用的aDSHT的主要构建块。

[0167] 在图6中示出编码器和解码器处理构建块pE和pD的细节。两个模块拥有作为DSHT的基础的相同的球面采样位置网格的码本。最初，使用系数的数量O_{3D}根据通用码本选择具有$L_{Sd}=O_{3D}$个位置的模块pE中的基础网格。必须将L_{Sd}传输给块pD进行初始化以选择与图3中所指示的相同的基础采样位置网格。通过矩阵$\boldsymbol{\Omega}_{DSHT}=\left[\boldsymbol{\Omega}_1,\ldots,\boldsymbol{\Omega}_{L_{sd}}\right]$描述基础采样网格，其中$\Omega_l=[\theta_l,\Phi_l]^T$定义单位球面上的位置。如上文所述，图5示出基本网格的示例。

[0168] 对旋转发现块(构建块“发现最佳旋转”)320的输入是系数矩阵B。该构建块负责旋转基础采样网格,使得等式(57)的值最小化。该旋转用“轴-角”表示来表示,并且将与该旋转有关的压缩的轴ψ_{rot}和旋转角φ_{rot}输出到该构建块作为边信息SI。可以通过从原点到单位球面上的位置的单位矢量来描述旋转轴ψ_{rot}。在球面坐标中,这可以通过两个角来结合:$\psi_{rot}=[\theta_{axis},\phi_{axis}]^T$,具有不需要传输的一个隐含的相关半径。通过用信号通知重用先前使用的值以创建边信息SI的特殊逃逸模式(escape pattern)对三个角θ_{axis}、ϕ_{axis}、φ_{rot}进行量化和熵编码。

[0169] 构建块“构建Ψ_i”330将旋转轴和角解码为$\widehat{\boldsymbol{\psi}}_{rot}$和$\widehat{\varphi}_{rot}$,并且将该旋转应用于基础采样网格$\mathfrak{D}_{DSHT}$,以得出旋转网格$\widehat{\mathfrak{D}}_{DSHT}=[\widehat{\boldsymbol{\Omega}}_1,\ldots,\widehat{\boldsymbol{\Omega}}_{L_{sd}}]$。其输出从矢量$\boldsymbol{y}_l=[Y_0^0(\widehat{\boldsymbol{\Omega}}_l),\ Y_1^{-1}(\widehat{\boldsymbol{\Omega}}_l),\ldots,Y_N^N(\widehat{\boldsymbol{\Omega}}_l)]^T$得出的iDSHT矩阵$\boldsymbol{\Psi}_i=[\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{L_{sd}}]$。

[0170] 在构建块“iDSHT”310中,通过$W_{Sd}=\Psi_i B$将HOA系数数据的实际块B变换到空间域中。

[0171] 解码处理块pD的构建块“构建Ψ_f”350接收旋转轴和角并将其解码为$\widehat{\boldsymbol{\psi}}_{rot}$和$\widehat{\varphi}_{rot}$,并且将该旋转应用于基础采样网格$\mathfrak{D}_{DSHT}$,以得出旋转网格$\widehat{\mathfrak{D}}_{DSHT}=[\widehat{\boldsymbol{\Omega}}_1,\ldots,\widehat{\boldsymbol{\Omega}}_{L_{sd}}]$。通过用矢量$\boldsymbol{y}_l=[Y_0^0(\widehat{\boldsymbol{\Omega}}_l),\ Y_1^{-1}(\widehat{\boldsymbol{\Omega}}_l),\ldots,Y_N^N(\widehat{\boldsymbol{\Omega}}_l)]^T$得到iDSHT矩阵$\boldsymbol{\Psi}_i=[\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{L_{sd}}]$,并且在解码侧计算DSHT矩阵$\boldsymbol{\Psi}_f=\boldsymbol{\Psi}_i^{-1}$。

[0172] 在解码器处理块34内的构模块“DSHT”340中,将空间域数据的实际块$\widehat{\boldsymbol{W}}_{Sd}$变换回到系数域数据的块:$\widehat{\boldsymbol{B}}=\boldsymbol{\Psi}_f\widehat{\boldsymbol{W}}_{Sd}$。

[0173] 下面描述包括压缩编解码器的总体架构的各种有利的实施例。第一实施例使用单个aDSHT。第二实施例使用谱带中的多个aDSHT。

[0174] 在图7中示出第一(“基本”)实施例。具有O_{3D}个系数信道b(m)的索引m的HOA时间样本首先被存储在缓冲器71中以形成M个样本的块和时间索引μ。在上述的构建块pE 72中使用自适应iDSHT,将B(μ)变换到空间域。将空间信号块$w_{Sd}(\mu)$输入到L_{Sd}个音频压缩单声道(mono)编码器73(如AAC或mp3编码器)或单个AAC多信道编码器(L_{Sd}个信道)。比特流S73包括具有集成的边信息SI的多个编码器比特流帧的复用的帧或集成了边信息SI(优选地作为辅助数据)的单个多信道比特流。

[0175] 在一个实施例中,相应的压缩解码器构建块包括用于将比特流S73分用为L_{Sd}个比特流和边信息SI并且将该比特流馈送给L_{Sd}个单声道解码器的分用器D1,将它们解码为具有M个样本的L_{Sd}个空间音频信道以形成块$\widehat{\boldsymbol{W}}_{Sd}(\mu)$,并且将$\widehat{\boldsymbol{W}}_{Sd}(\mu)$和SI馈送给pD。在不对比特流进行复用的另一实施例中,压缩解码器构建块包括接收器74,接收器74用于接收比特流并且将其解码为L_{Sd}个多信道信号$\widehat{\boldsymbol{W}}_{Sd}(\mu)$,对SI解包,并且将$\widehat{\boldsymbol{W}}_{Sd}(\mu)$和SI馈送给pD。

[0176] 在解码器处理块pD 75中,使用自适应DSHT和SI将$\widehat{\boldsymbol{W}}_{Sd}(\mu)$变换到系数域,以形成HOA信号的块B(μ),其被存储在缓冲器76中以便解帧,以形成系数的时间信号b(m)。

[0177] 在某些条件下，上述的第一实施例可能具有两个缺点：首先，由于空间信号分布的改变，可能存在来自先前块(即，来自块μ至μ+1)的组块伪像(blocking artifact)；其次，可能同时存在多于一个的强信号，并且aDSHT的解相关效应可能相当小。

[0178] 在工作于频域中的第二实施例中解决两个缺点。aDSHT应用于组合多个频带数据的标度因子带数据。通过利用重叠添加(Overlay Add，OLA)处理重叠时频变换(TFT)的块来避免组块伪像。可以通过使用本发明在J个谱带内以传输SI_j的数据率中的增大的开销的成本来实现改善的信号解相关。

[0179] 下面描述图9所示的第二实施例的一些更多的细节：对信号b(m)的每个系数信道进行时频变换(TFT)912。广泛使用的TFT的示例是修正余弦变换(MDCT)。在TFT成帧单元911中，构造50％的重叠数据块(块索引μ)。TFT块变换单元912执行块变换。在谱带化单元913中，组合TFT频带以形成J个新的谱带和有关的信号$\boldsymbol{B}_j(\mu)\in\mathbb{C}^{O_{3D}\times K_j}$，其中$K_J$表示带j中的频率系数的数量。在多个处理模块914中处理这些谱带。对于这些谱带中的每一个，存在一个创建信号$\boldsymbol{W}_{j_{Sd}}(\mu)\in\mathbb{C}^{L_{sd}\times K_j}$和边信息$SI_j$的处理块$pE_j$。谱带可以匹配有损音频压缩方法的谱带(如AAC/mp3标度因子带)，或者具有更粗糙的粒度。在后者的情况下，不利用TFT块915的信道无关的有损音频压缩需要重新布置所述带化。处理块914操作如同将恒定的比特率分配给每个音频信道的频域中的L_{Sd}多信道音频编码器。在比特流包装块916中格式化比特流。

[0180] 解码器接收或存储比特流(至少其若干部分)，将其解包921，并且将音频数据馈送给用于不利用TFT的信道无关的音频解码的多信道音频解码器922，并且将边信息SI_j馈送给多个解码处理块pD_j 923。用于不利用TFT的信道无关的音频解码的音频解码器922对音频信息进行解码，并且格式化J个谱带信号$\widehat{\boldsymbol{W}}_{j_{Sd}}(\mu)$，作为给解码处理块$pD_j$ 923的输入，其中，将这些信号变换到HOA系数域以形成$\widehat{\boldsymbol{B}}_j(\mu)$。在去谱带化块924中，重组J个谱带以匹配TFT的带化。将它们变换到iTFT和OLA块925中的时域，该块使用块重叠的重叠添加(OLA)处理。最后，在TFT解帧块926中，iTFT和OLA模块925的输出被解帧，以创建信号$\widehat{\boldsymbol{b}}(m)$。

[0181] 本发明基于如下发现：由信道之间的互相关性产生SNR增加。感知编码器仅考虑出现在每个个体单信道信号内的编码噪声掩蔽效应。然而，这种效应典型地是非线性的。因此，在将这样的单信道矩阵化为新的信号时，可能发生噪声去掩蔽。这是通常在矩阵化运算之后编码噪声增大的原因。

[0182] 本发明提出通过使不需要的噪声去掩蔽效应最小化的自适应离散球面谐波变换(aDSHT)对信道进行解相关。aDSHT被集成在压缩编码器和解码器架构内。因为其包括针对HOA输入信号的空间特性来调节DSHT的空间采样网格的旋转运算，所以其是自适应的。aDSHT包括自适应旋转和实际的传统DSHT。实际的DSHT是可以如现有技术中描述的那样地构造的矩阵。对该矩阵应用自适应旋转，从而导致信道间相关性的最小化，并且因此导致矩阵化之后的SNR增加的最小化。通过自动搜索运算(而不是分析地)发现旋转轴和角。对旋转轴和角进行编码和传输，以使得能够在解码之后和在矩阵化之前进行重新相关，其中使用逆自适应DSHT(iaDSHT)。

[0183] 在一个实施例中，执行时频变换(TFT)和谱带化，并且将aDSHT/iaDSHT独立地应用

于每个谱带。

[0184] 图8a)示出本发明的一个实施例中的用于对多信道HOA音频信号进行编码以便降噪的方法的流程图。图8b)示出本发明的一个实施例中的用于对多信道HOA音频信号进行解码以便降噪的方法的流程图。

[0185] 在图8a)所示的实施例中,用于对多信道HOA音频信号进行编码以便降噪的方法包括以下步骤:使用逆自适应DSHT对信道进行解相关81,所述逆自适应DSHT包括旋转运算和逆DSHT 812,所述旋转运算旋转811iDSHT的空间采样网格;对每个经解相关的信道进行感知编码82;对(作为边信息SI的)旋转信息进行编码83,所述旋转信息包括定义所述旋转运算的参数;以及,传输或存储84经感知编码的音频信道和经编码的旋转信息。

[0186] 在一个实施例中,逆自适应DSHT包括以下步骤:选择初始的默认球面样本网格;确定最强的源方向;以及,对M个时间样本的块,旋转球面样本网格,使得单个空间样本位置匹配最强的源方向。

[0187] 在一个实施例中,旋转球面样本网格,使得以下项的对数最小化:

[0188] $$\sum_{l=1}^{L_{Sd}}\sum_{j=1}^{L_{Sd}}|\Sigma_{W_{Sd}\,l,j}|-\sum\left(\sigma^2_{S_{d_1}},\ldots,\sigma^2_{S_{d_{L_{Sd}}}}\right)$$

[0189] 其中,$|\Sigma_{W_{Sd}\,l,j}|$是$\boldsymbol{\Sigma}_{\boldsymbol{W}_{Sd}}$(具有矩阵行索引l和列索引j)的元素的绝对值,并且$\sigma^2_{S_{d_l}}$是$\boldsymbol{\Sigma}_{\boldsymbol{W}_{Sd}}$的对角线元素,其中$\boldsymbol{\Sigma}_{\boldsymbol{W}_{Sd}}=\boldsymbol{W}_{Sd}\,\boldsymbol{W}_{Sd}^{H}$并且$W_{Sd}$是音频信道的数量乘以处理样本的块的数量的矩阵,并且W_{Sd}是aDSHT的结果。

[0190] 在图8b)所示的实施例中,一种用于对具有降低的噪声的经编码的多信道HOA音频信号进行解码的方法包括以下步骤:接收85经编码的多信道HOA音频信号和信道旋转信息(在边信息SI内);对接收的数据进行解压缩86,其中使用感知解码;使用自适应DSHT对每个信道进行空间解码87,其中执行DSHT 872和根据所述旋转信息的DSHT的空间采样网格的旋转871,并且其中对经感知解码的信道进行重新相关;以及,对重新相关的经感知解码的信道进行矩阵化88,其中获得映射到扬声器位置的可再现的音频信号。

[0191] 在一个实施例中,自适应DSHT包括以下步骤:选择自适应DSHT的初始的默认球面样本网格;以及,对M个时间样本的块,根据所述旋转信息来旋转球面样本网格。

[0192] 在一个实施例中,旋转信息是具有三个分量的空间矢量$\widehat{\boldsymbol{\psi}}_{rot}$。注意,旋转轴$\psi_{rot}$可以用单位矢量来描述。

[0193] 在一个实施例中,旋转信息是由3个角构成的矢量:θ_{axis}、Φ_{axis}、φ_{rot},其中,θ_{axis}、Φ_{axis}定义关于具有在球面坐标中的一个隐含半径的旋转轴的信息,并且φ_{rot}定义绕该轴的旋转角。

[0194] 在一个实施例中,通过用信号通知(亦即,指示)重用先前的值以便创建边信息(SI)的逃逸模式(亦即,专用比特模式),对角进行量化和熵编码。

[0195] 在一个实施例中,一种用于对多信道HOA音频信号进行编码以便降噪的设备包括:解相关器,用于使用逆自适应DSHT对信道进行解相关,所述逆自适应DSHT包括旋转运算和逆DSHT(iDSHT),其中旋转运算旋转iDSHT的空间采样网格;感知编码器,用于对每个经解相关的信道进行感知编码;边信息编码器,用于对旋转信息进行编码,所述旋转信息包括定义

所述旋转运算的参数;以及接口,用于传输或存储经感知编码的音频信道和经编码的旋转信息。

[0196] 在一个实施例中,一种用于对具有降低的噪声的多信道HOA音频信号进行解码的设备包括:接口装置330,用于接收经编码的多信道HOA音频信号和信道旋转信息;解压缩模块33,用于通过使用用于对每个信道进行感知解码的感知解码器对接收的数据进行解压缩;相关器34,用于对经感知解码的信道进行重新相关,其中执行DSHT和根据所述旋转信息的DSHT的空间采样网格的旋转;以及混合器,用于对相关的经感知解码的信道进行矩阵化,其中获得映射到扬声器位置的可再现的音频信号。原则上,相关器34用作空间解码器。

[0197] 在一个实施例中,一种用于对具有降低的噪声的多信道HOA音频信号进行解码的设备包括:接口装置330,用于接收经编码的多信道HOA音频信号和信道旋转信息;解压缩模块33,用于通过用于对每个信道进行感知解码的感知解码器对接收的数据进行解压缩;相关器34,用于使用aDSHT对经感知解码的信道进行相关,其中执行DSHT和根据所述旋转信息的DSHT的空间采样网格的旋转;以及混合器MX,用于对相关的经感知解码的信道进行矩阵化,其中获得映射到扬声器位置的可再现音频信号。

[0198] 在一个实施例中,用于进行解码的设备中的自适应DSHT包括用于选择自适应DSHT的初始的默认样本网格的装置、用于对M个时间样本的块根据所述旋转信息旋转默认球面样本网格的旋转处理装置、以及用于对旋转的球面样本网格执行DSHT的变换处理装置。

[0199] 在一个实施例中,用于进行解码的设备中的相关器34包括用于使用自适应DSHT同时对每个信道进行空间解码的多个空间解码单元922,还包括用于执行去谱带化的去谱带化单元924、以及用于通过重叠添加处理执行逆时频变换的iTFT和OLA单元925,其中所述去谱带化单元将其输出提供给iTFT和OLA单元。

[0200] 在所有实施例中,术语降低的噪声至少涉及避免编码噪声去掩蔽。

[0201] 对音频信号的感知编码表示适合于对音频的人类感知的编码。应当注意,在对音频信号进行感知编码时,通常不对宽带音频信号样本而是在与人类感知有关的个体频带中执行量化。因此,信号功率与量化噪声之间的比率可以在个体频带之间变化。因此,感知编码通常包括减少冗余和/或无关信息,而空间编码通常涉及信道之间的空间关系。

[0202] 上述的技术可以被看作是对使用Karhunen-Loève变换(KLT)的解相关的替代。本发明的一个优点是极大地减少了边信息量,边信息仅包括三个角。KLT需要块相关矩阵的系数作为边信息,因此需要多得多的数据。此外,在此公开的技术允许对旋转进行调整(或微调),以便减少进行到下一个处理块时的过渡伪像(transition artifact)。这有利于后续的感知编码的压缩质量。

[0203] 表1提供aDSHT与KLT之间的直接比较。尽管存在一些相似性,但是aDSHT提供了超过KLT的显著优点。

[0204]

	sDSHT	KLT
定义	$\boldsymbol{B}$ 是 N 阶 HOA 信号矩阵，$(N+1)^2$ 行（系数），T 列（时间样本）；$\boldsymbol{W}$ 是具有$(N+1)^2$ 行（信道）、T 列（时间样本）的空间矩阵	
编码器，空间变换	逆 aDSHT $\boldsymbol{W}_{Sd} = \boldsymbol{\Psi}_i \boldsymbol{B}$	Karhunen-Loève 变换: $\boldsymbol{W}_k = \boldsymbol{K}\boldsymbol{B}$
变换矩阵	选择具有对编码器和解码器已知的$(N+1)^2$ 个球面样本位置的球面规律采样网格。绕轴$\boldsymbol{\psi}_{rot}$以旋转角φ_{rot}旋转该网格，这已在先前得出（参见下面的备注）。创建该网格的模态矩阵$\boldsymbol{\Psi}_f$（亦即，这些位置的球面谐波）：$\boldsymbol{\Psi}_i = \boldsymbol{\Psi}_f^{-1}$ （或者更一般地，在空间信道的数量变得大于$(N+1)^2$ 时具有$\boldsymbol{\Psi}_f\boldsymbol{\Psi}_i = \boldsymbol{I}$的$\boldsymbol{\Psi}_i = \boldsymbol{\Psi}_f^+$） 该变换矩阵式旋转的球面网格的逆模态矩阵。该旋转是信号驱动的，并且在每个处理块更新。	构建协方差矩阵: $\boldsymbol{C} = \boldsymbol{B}\boldsymbol{B}^H$ 本征值分解: $\boldsymbol{C} = \boldsymbol{K}^H \boldsymbol{\Lambda} \boldsymbol{K}$ 具有 $\boldsymbol{\Lambda}$中的本征值对角线以及$\boldsymbol{K}^H$中布置的相关的本征值矢量，其中$\boldsymbol{K}\boldsymbol{K}^H = \boldsymbol{1}$（如同任何正交变换中那样）。针对每个处理块，从信号 $\boldsymbol{B}$ 得出变换矩阵。
要传输的边信息	例如被编码为如下 3 个值的轴$\boldsymbol{\psi}_{rot}$和旋转角φ_{rot}：θ_{axis}、ϕ_{axis}、φ_{rot}	$\boldsymbol{C}$的多于一半的元素（即，$\frac{(N+1)^4+(N+1)^2}{2}$个值）或 $\boldsymbol{K}$的多于一半的元素（即，$(N+1)^4$个值）

	经有损解压缩的空间信号	空间信号被有损编码（编码噪声 $\boldsymbol{E}_{cod}$）。T 个样本的块被布置为 $\widehat{\boldsymbol{W}}_{Sd}$	空间信号被有损编码（编码噪声 $\widehat{\boldsymbol{E}}_{cod}$）。T 个样本的块被布置为 $\widehat{\boldsymbol{W}}_{k}$
	解码器，逆空间变换	$\widehat{\boldsymbol{B}} = \boldsymbol{\Psi}_f \widehat{\boldsymbol{W}}_{Sd} = \boldsymbol{B} + \boldsymbol{\Psi}_f \boldsymbol{E}_{cod}$	$\widehat{\boldsymbol{B}}_{\boldsymbol{k}} = \boldsymbol{K}\widehat{\boldsymbol{W}}_{k} = \boldsymbol{B} + \boldsymbol{K}\widehat{\boldsymbol{E}}_{cod}$
[0205]	备注	在一个实施例中，旋转网格，使得采样位置匹配 $\boldsymbol{B}$ 内的最强的信号方向。这里可以使用协方差矩阵的分析，如同可以将其用于 KLT 一样。实际上，由于更加简单和更少的计算复杂度，可以使用也允许从块到块平滑地调节/修正旋转的信号跟踪模型，从而避免在有损（感知）编码块中创建组块伪像。	

[0206] 表1 aDSHT对KLT的比较

[0207] 虽然已经示出、描述和指出对本发明的优选的实施例应用的基础的新颖的特征，但是应当理解，本领域的技术人员可以在所描述的设备和方法中，在所公开的装置的形式和细节以及在其操作方面，进行各种省略和替代和变化，而不脱离本发明的精神。显然旨在以基本相同的方式执行基本相同的功能以获得相同的结果的那些元件的所有组合都在本发明的范围内。还充分地预期和设想到从一个所描述的实施例到另一个所描述的实施例的元件的替换。

[0208] 应当理解的是，仅仅通过示例对本发明进行了描述，可以对细节进行修改，而不脱离本发明的范围。

[0209] 在本说明书和(适当之处)权利要求书和附图中公开的每个特征可以独立地或以任何适当的组合来提供。

[0210] 特征可以在适当的情况下被实现为硬件、软件或这二者的组合。连接可以在可应用的情况下被实现为无线连接或者有线的(不必是直接或专用的)连接。

[0211] 在权利要求中出现的标号仅作为示例，而不应当具有对权利要求的范围的限定效果。

[0212] 引用的参考文献

[0213] [1]T.D.Abhayapala。Generalized framework for spherical microphone arrays:Spatial and frequency decomposition。IEEE International Conference on Acoustics,Speech,and Signal Processing(ICASSP)会议，(接受的)第X卷，页，2008年4月，拉斯维加斯，美国。

[0214] [2]James R.Driscoll和Dennis M.Healy Jr.。Computing fourier transforms and convolutions on the 2-sphere。Advances in Applied Mathematics,15:202-250,1994年。

[0215] [3] Jörg Fliege。Integration nodes for the sphere,http://www.personal.soton.ac.uk/jf1w07/nodes/nodes.html

[0216] [4] Jörg Fliege和Ulrike Maier。A two-stage approach for computing cubature formulae for the sphere。技术报告,Fachbereich Mathematik,多特蒙德大学,1999年。

[0217] [5]R.H.Hardin和N.J.A.Sloane。网页:Spherical designs,spherical t-designs。http://www2.research.att.com/-njas/sphdesigns

[0218] [6]R.H.Hardin和N.J.A.Sloane。Mclaren's improved snub cube and other new spherical designs in three dimensions。Discrete and Computational Geometry,15:429-441,1996年。

[0219] [7]Erik Hellerud、lan Burnett、Audun Solvang和U.Peter Svensson.Encoding higher order Ambisonics with AAC。第124届AES会议,阿姆斯特丹,2008年5月。

[0220] [8]Peter Jax、Jan-Mark Batke、Johannes Boehm和Sven Kordon。Perceptual coding of HOA signals in spatial domain。欧洲专利申请EP2469741A1(PD100051)。

[0221] [9]Boaz Rafaely。Plane-wave decomposition of the sound field on a sphere by spherical convolution。J.Acoust.Soc.Am.,4(116):2149-2157,2004年10月。

[0222] [10]Earl G.Williams。Fourier Acoustics,Applied Mathematical Sciences第93卷。Academic Press,1999年。

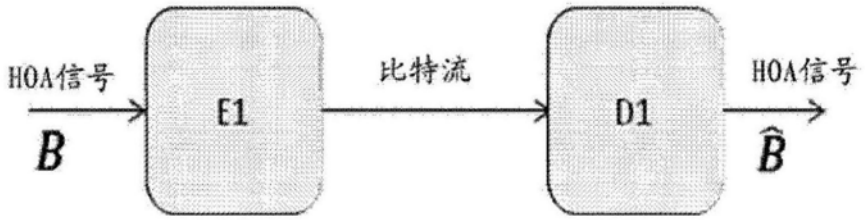

HOA信号
B
E1
比特流
D1
HOA信号
$\hat{B}$

图1

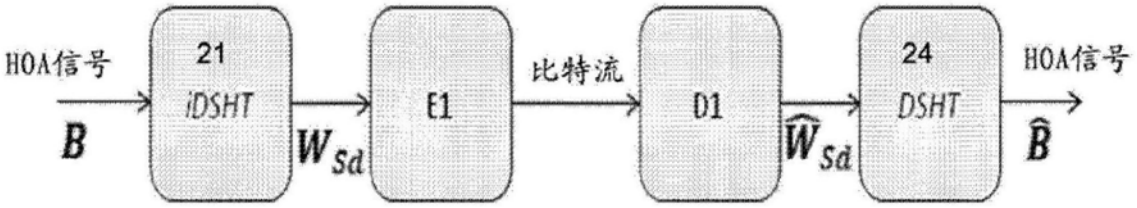

HOA信号
B
21
iDSHT
W_{Sd}
E1
比特流
D1
$\widehat{W}_{Sd}$
24
DSHT
HOA信号
$\hat{B}$

图2

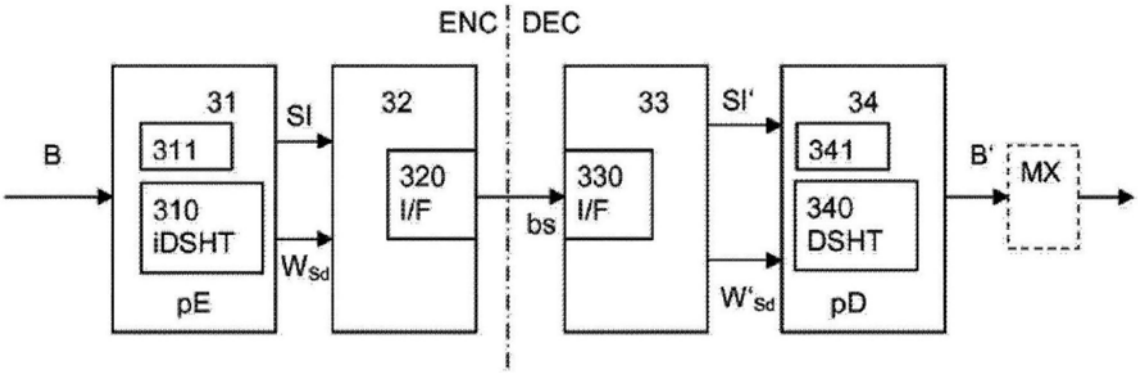
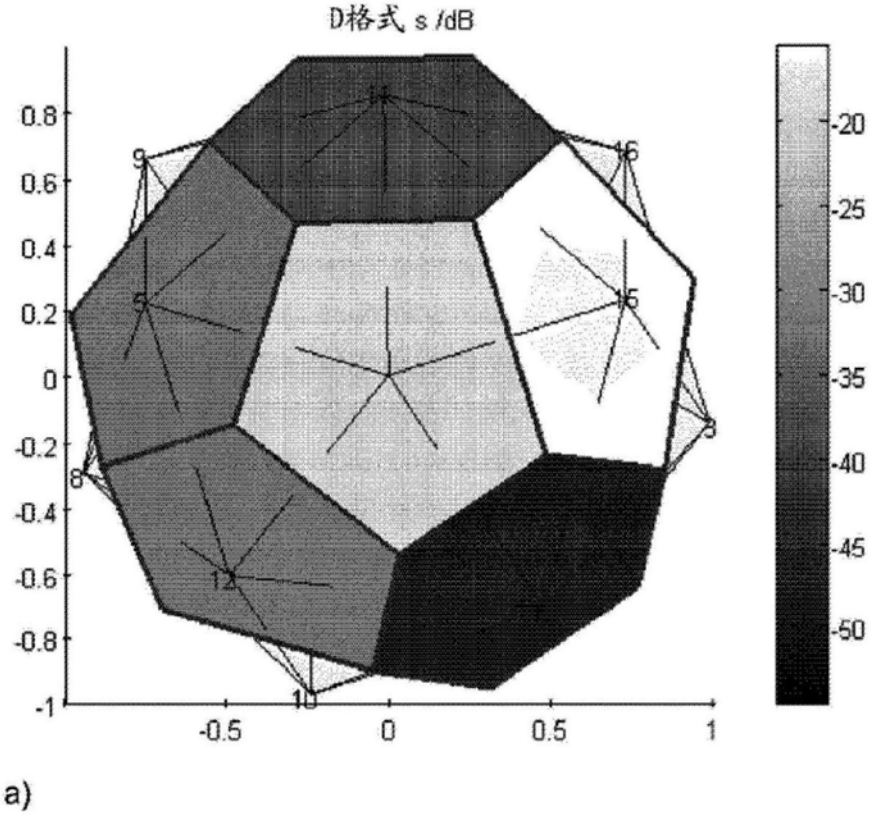
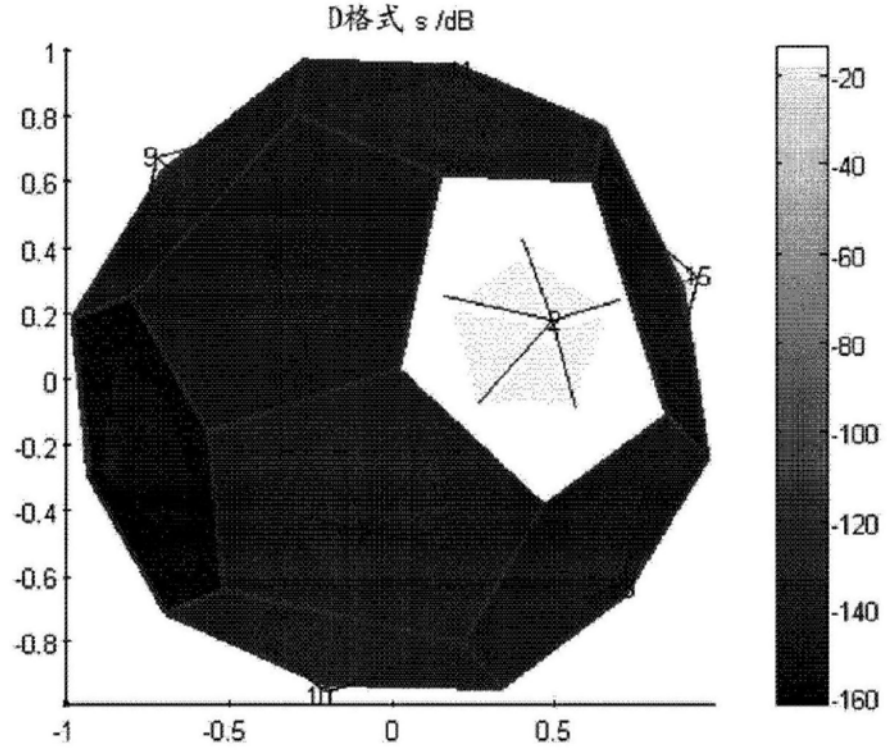

ENC
DEC
B
31
311
310
iDSHT
pE
SI
W_{Sd}
32
320
I/F
bs
33
330
I/F
SI'
W'sd
34
341
340
DSHT
pD
B'
MX

图3


D格式 s /dB
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-0.5
0
0.5
1
-20
-25
-30
-35
-40
-45
-50

a)


D格式 s /dB
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-0.5
0
0.5
-20
-40
-60
-80
-100
-120
-140
-160

b)

图4

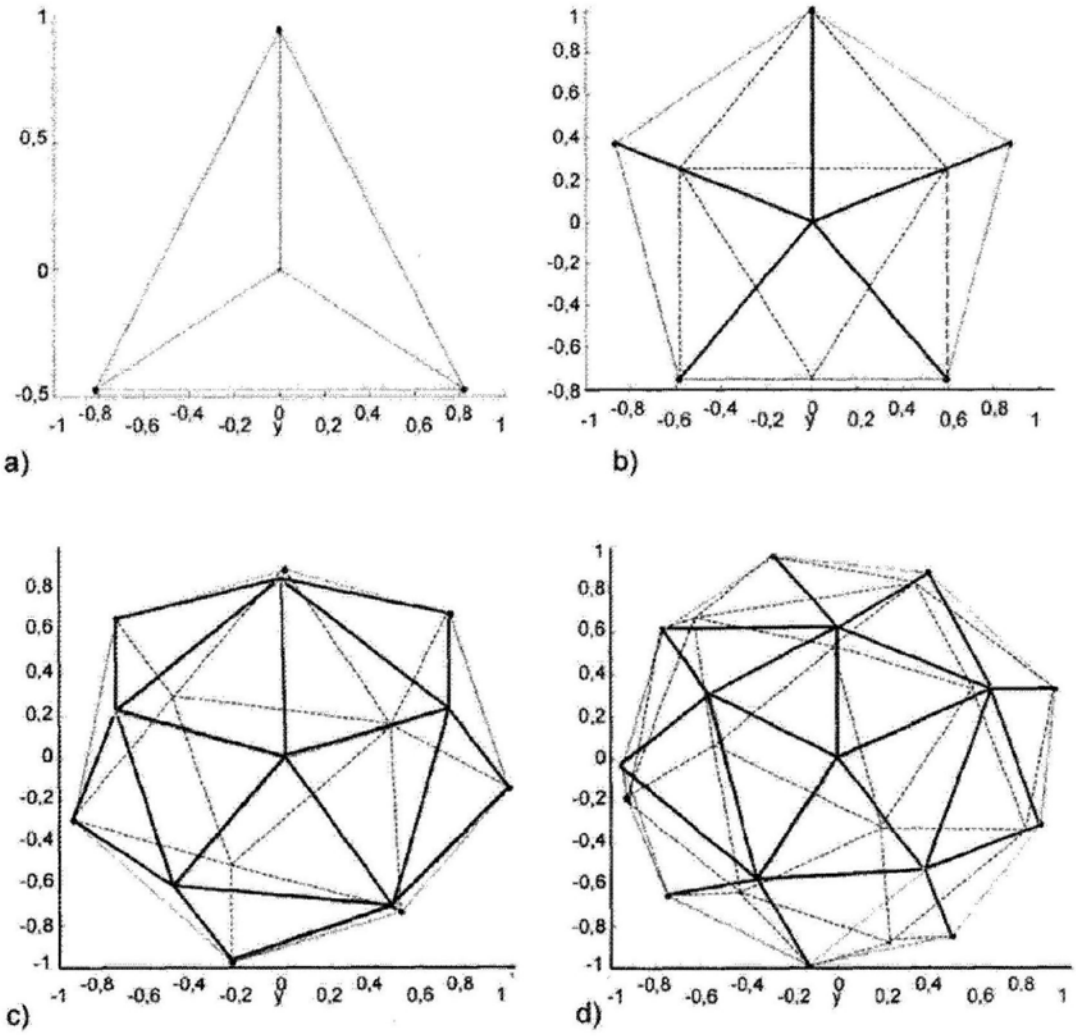

a)
b)
c)
d)
x
y

图5

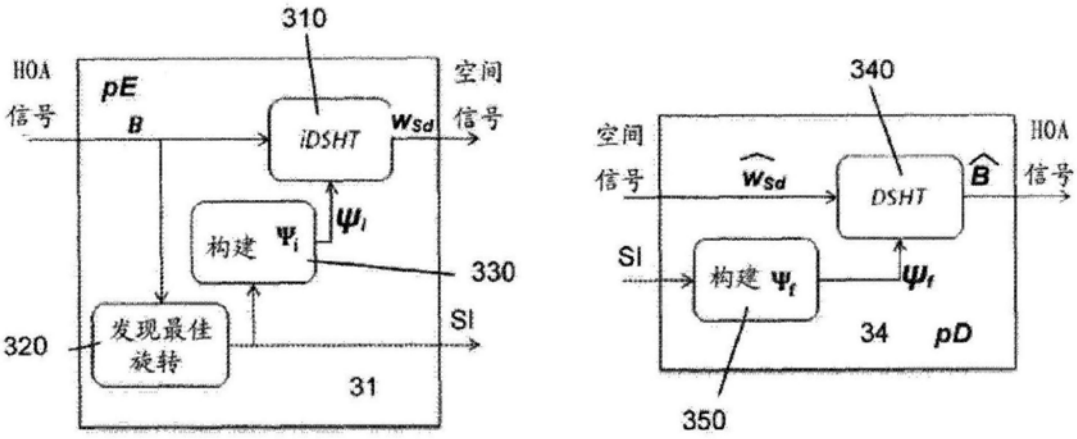

310
HOA
信号
pE
B
iDSHT
wSd
空间
信号
构建 Ψi
ψi
330
320
发现最佳
旋转
SI
31
340
空间
信号
HOA
信号
wSd
DSHT
B
SI
构建 Ψf
ψf
34
pD
350

图6

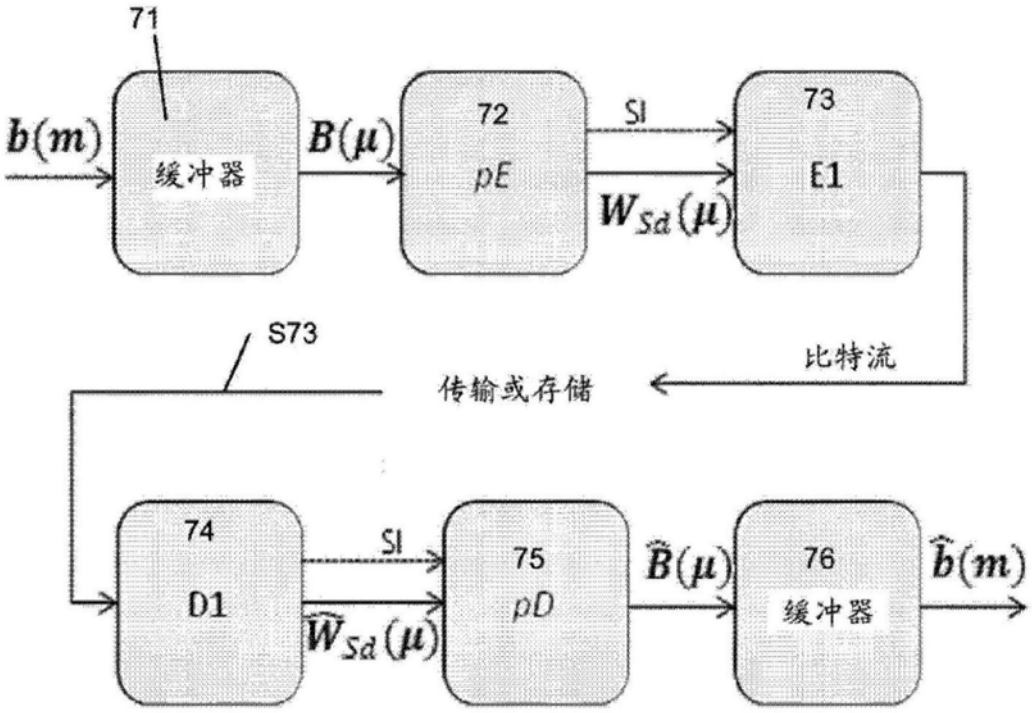

71
b(m)
缓冲器
B(μ)
72
pE
SI
W_Sd(μ)
73
E1
S73
比特流
传输或存储
74
D1
SI
W̃_Sd(μ)
75
pD
B̂(μ)
76
缓冲器
b̂(m)

图7

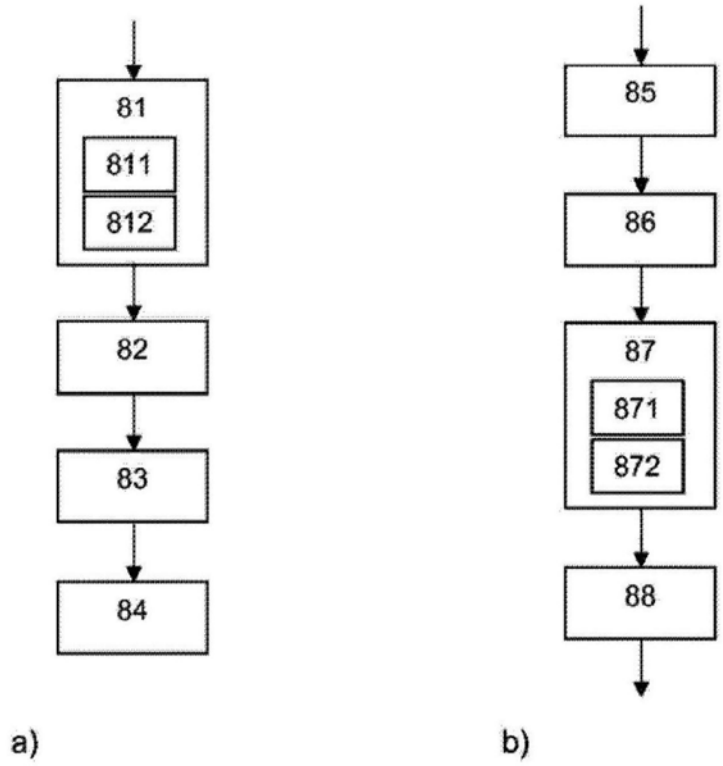

81
811
812
82
83
84
a)
85
86
87
871
872
88
b)

图8

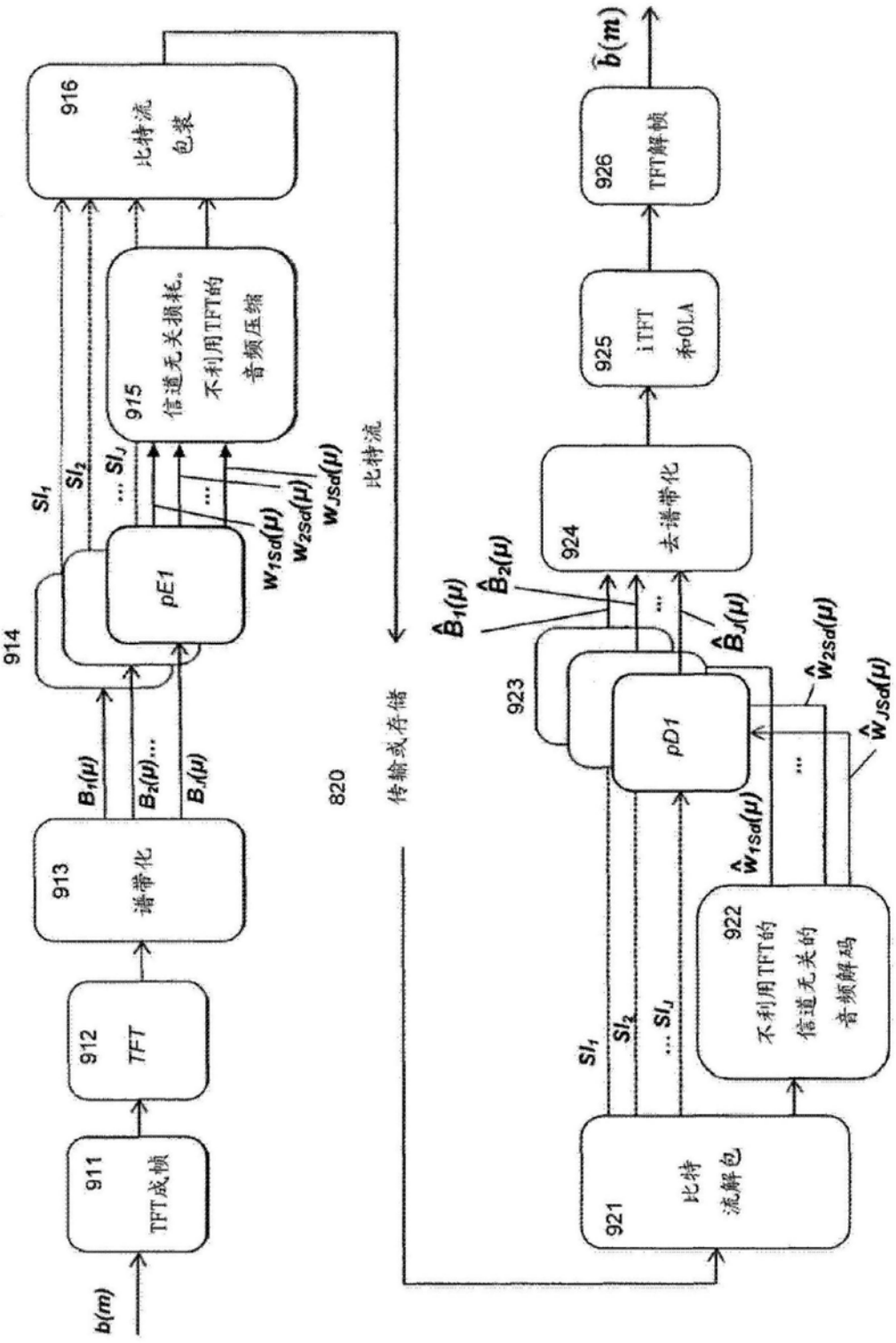

b(m)
911
TFT成帧
912
TFT
913
谱带化
914
pE1
915
信道无关损耗。
不利用TFT的
音频压缩
916
比特流
包装
820
传输或存储
比特流
921
比特
流解包
922
不利用TFT的
信道无关的
音频解码
923
pD1
924
去谱带化
925
iTFT
和OLA
926
TFT解帧
$\hat{b}(m)$

图9