



## (12)发明专利

(10)授权公告号 CN 106463123 B

(45)授权公告日 2020.03.03

(21)申请号 201580014972.9

(22)申请日 2015.03.20

(65)同一申请的已公布的文献号  
申请公布号 CN 106463123 A

(43)申请公布日 2017.02.22

(30)优先权数据  
14305411.2 2014.03.21 EP(85)PCT国际申请进入国家阶段日  
2016.09.20(86)PCT国际申请的申请数据  
PCT/EP2015/055914 2015.03.20(87)PCT国际申请的公布数据  
W02015/140291 EN 2015.09.24

(73)专利权人 杜比国际公司

地址 荷兰阿姆斯特丹

(72)发明人 S·科尔多恩 A·克鲁格  
O·伍埃博尔特(74)专利代理机构 中国国际贸易促进委员会专  
利商标事务所 11038

代理人 宿小猛

(51)Int.Cl.  
G10L 19/008(2006.01)  
G10L 19/24(2006.01)

审查员 薛松

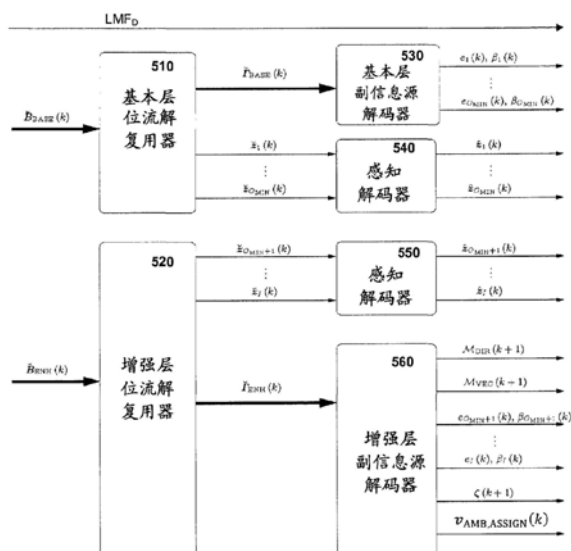
权利要求书4页 说明书18页 附图9页

## (54)发明名称

对压缩的高阶高保真立体声(HOA)表示进行  
解码的方法及装置

## (57)摘要

一种用于压缩HOA信号的方法包括输入时间帧的空间HOA编码以及随后的感知编码和源编码,该HOA信号是具有HOA系数序列的输入时间帧( $C_{(k)}$ )的输入HOA表示。将每个输入时间帧分解(802)为主导声音信号( $X_{PS}(k-1)$ )的帧和环境HOA分量( $C_{AMB}(k-1)$ )的帧。环境HOA分量( $C_{AMB}(k-1)$ )在分层模式下包括在较低位置的输入HOA表示( $c_n(k-1)$ )的第一HOA系数序列以及在其余的较高的位置的第二HOA系数序列( $c_{AMB,n}(k-1)$ )。第二HOA系数序列是输入HOA表示和主导声音信号的HOA表示之间的残差的HOA表示的一部分。



1. 一种对声音或声场的压缩的高阶高保真立体声 (HOA) 表示进行解码的方法, 所述方法包括:

接收包含压缩的HOA表示的位流, 包括压缩的基本层位流和压缩的增强层位流;

确定是否存在与压缩的HOA表示有关的多个层;

基于确定存在多个层从所述位流解码压缩的HOA表示, 以获得解码的HOA表示的序列,

其中, 该解码包括:

对压缩的基本层位流进行解复用, 以获得第一感知编码的输送信号和第一编码的副信息;

对压缩的增强层位流进行解复用, 以获得第二感知编码的输送信号和第二编码的副信息;

对所述第一感知编码的输送信号进行感知解码以获得第一感知解码的输送信号, 并且对所述第二感知编码的输送信号进行感知解码以获得第二感知解码的输送信号;

对所述第一编码的副信息进行解码以获得第一指数和第一异常标志, 并且对第二编码的副信息进行解码以获得第二指数和第二异常标志, 其中关于方向信号的第一元组集合和关于基于矢量的信号的第二元组集合被获得, 所述第一元组集合的每个元组包括方向信号的索引和相应的量化方向, 所述第二元组集合的每个元组包括基于矢量的信号的索引和限定所述基于矢量的信号的方向分布的矢量, 其中, 预测参数和环境分配矢量被获得, 其中所述环境分配矢量包括对于每个传输通道指示它是否包含环境HOA分量的系数序列以及它包含环境HOA分量的哪个系数序列的分量;

根据所述第一指数和所述第一异常标志将所述第一感知解码的输送信号变换为第一增益校正的信号帧, 并且其中, 根据所述第二指数和所述第二异常标志将所述第二感知解码的输送信号变换为第二增益校正的信号帧;

根据所述环境分配矢量以及所述第一元组集合和所述第二元组集合将所述第一增益校正的信号帧和所述第二增益校正的信号帧重新分布给I个通道, 其中, 主导声音信号的帧被重构, 所述主导声音信号包括方向信号和基于矢量的信号, 并且其中, 修正的环境HOA分量被获得;

产生修正的环境HOA分量的在第k帧中起作用的系数序列的第一索引集合、以及修正的环境HOA分量的在第(k-1)帧中必须被启用、被禁用以及保持起作用的系数序列的第二索引集合, 其中所述第一索引集合基于 $0_{\text{MIN}}$ 个通道, 所述第二索引集合基于所述多个层中的至少一个被确定;

从所述主导声音信号合成主导HOA声音分量的HOA表示, 其中, 所述第一和第二元组集合、所述预测参数以及所述第二索引集合被使用;

从所述修正的环境HOA分量合成环境HOA分量; 其中,

如果确定具有至少两个层, 则所述环境HOA分量在其 $0_{\text{MIN}}$ 个最低位置包括所述解压缩的HOA信号的HOA系数序列, 以及在其余的较高位置包括作为所述主导HOA声音分量的HOA表示和所述解压缩的HOA信号之间的残差的HOA表示的一部分的系数序列, 并且

只有最高的 $I - 0_{\text{MIN}}$ 个系数通道通过所述主导HOA声音分量和所述环境HOA分量的相加获得, 所述解压缩的HOA信号的最低的 $0_{\text{MIN}}$ 个系数通道是从所述环境HOA分量复制的。

2. 根据权利要求1所述的方法, 其中, 所述第一索引集合基于 $1 \leq n \leq 0_{\text{MIN}}$ 被确定, 并且所

述第二索引集合基于 $0_{\text{MIN}}+1 \leq n \leq 0$ 被确定,其中,0指示通道的总数,并且 $0_{\text{MIN}}$ 指示1和0之间的数。

3.根据权利要求1所述的方法,其中,对于索引 $n$ 和帧 $k$ ,当 $n$ 是在所述第一索引集合中时,所述第一子集基于对应的环境声音分量 $\hat{c}_{\text{AMB},n}(k-1)$ 被确定,并且当 $n$ 在所述第二索引集合中时,所述第二子集基于对应的主导声音分量 $\hat{c}_{\text{PS},n}(k-1)$ 和对应的环境声音分量 $\hat{c}_{\text{AMB},n}(k-1)$ 的相加而被确定,并且其中,解码的HOA表示至少部分地由下式表示:

$$\tilde{c}_n(k-1) = \begin{cases} \hat{c}_{\text{AMB},n}(k-1) & \text{对于在所述第一索引集合中的 } n \\ \hat{c}_n(k-1) = \hat{c}_{\text{PS},n}(k-1) + \hat{c}_{\text{AMB},n}(k-1), & \text{对于在所述第二索引集合中的 } n \end{cases}$$

4.根据权利要求1所述的方法,其中, $0_{\text{MIN}} = (N_{\text{MIN}}+1)^2$ 且 $N_{\text{MIN}} \leq N$ ,其中 $N$ 是编码的HOA表示的输入帧的阶次。

5.根据权利要求1所述的方法,其中,多个层的指示在所述位流中以信号告知。

6.根据权利要求1所述的方法,其中,所述多个层包括基本层和至少一个增强层。

7.根据权利要求1所述的方法,进一步基于不存在多个层的确定而确定存在单个层,并且基于单个层的确定,对于帧 $k$ 基于对应的主导HOA声音分量( $\hat{c}_{\text{PS}}(k-1)$ )和对应的环境HOA分量( $\hat{c}_{\text{AMB}}(k-1)$ )的相加来确定单层解码的HOA表示。

8.一种对声音或声场的压缩的高阶高保真立体声(HOA)表示进行解码的装置,所述装置包括:

接收器,用于接收包含压缩的HOA表示的位流,包括压缩的基本层位流和压缩的增强层位流;

音频解码器,用于基于存在多个层的确定从所述位流解码压缩的HOA表示,以获得解码的HOA表示的序列,

其中,所述音频解码器被配置为通过以下进行解码:

对压缩的基本层位流进行解复用,以获得第一感知编码的输送信号和第一编码的副信息;

对压缩的增强层位流进行解复用,以获得第二感知编码的输送信号和第二编码的副信息;

对所述第一感知编码的输送信号进行感知解码以获得第一感知解码的输送信号,并且对所述第二感知编码的输送信号进行感知解码以获得第二感知解码的输送信号;

对所述第一编码的副信息进行解码以获得第一指数和第一异常标志,并且对第二编码的副信息进行解码以获得第二指数和第二异常标志,其中关于方向信号的第一元组集合和关于基于矢量的信号的第二元组集合被获得,所述第一元组集合的每个元组包括方向信号的索引和相应的量化方向,所述第二元组集合的每个元组包括基于矢量的信号的索引和限定所述基于矢量的信号的方向分布的矢量,其中,预测参数和环境分配矢量被获得,其中所述环境分配矢量包括对于每个传输通道指示它是否包含环境HOA分量的系数序列以及它包含环境HOA分量的哪个系数序列的分量;

根据所述第一指数和所述第一异常标志将所述第一感知解码的输送信号变换为第一增益校正的信号帧,并且其中,根据所述第二指数和所述第二异常标志将所述第二感知解

码的输送信号变换为第二增益校正的信号帧；

根据所述环境分配矢量以及所述第一元组集合和所述第二元组集合将所述第一增益校正的信号帧和所述第二增益校正的信号帧重新分布给I个通道,其中,主导声音信号的帧被重构,所述主导声音信号包括方向信号和基于矢量的信号,并且其中,修正的环境HOA分量被获得;

产生修正的环境HOA分量的在第k帧中起作用的系数序列的第一索引集合、以及修正的环境HOA分量的在第(k-1)帧中必须被启用、被禁用以及保持起作用的系数序列的第二索引集合,其中所述第一索引集合基于 $0_{\text{MIN}}$ 个通道,所述第二索引集合基于所述多个层中的至少一个被确定;

从所述主导声音信号合成主导HOA声音分量的HOA表示,其中,所述第一和第二元组集合、所述预测参数以及所述第二索引集合被使用;

从所述修正的环境HOA分量合成环境HOA分量;其中,

如果确定具有至少两个层,则所述环境HOA分量在其 $0_{\text{MIN}}$ 个最低位置包括所述解压缩的HOA信号的HOA系数序列,以及在其余的较高位置包括作为所述主导HOA声音分量的HOA表示和所述解压缩的HOA信号之间的残差的HOA表示的一部分的系数序列,并且

只有最高的 $I-0_{\text{MIN}}$ 个系数通道通过所述主导HOA声音分量和所述环境HOA分量的相加获得,所述解压缩的HOA信号的最低的 $0_{\text{MIN}}$ 个系数通道是从所述环境HOA分量复制的。

9. 根据权利要求8所述的装置,其中,所述第一索引集合基于 $1 \leq n \leq 0_{\text{MIN}}$ 被确定,并且所述第二索引集合基于 $0_{\text{MIN}}+1 \leq n \leq 0$ 被确定,其中,0指示通道的总数,并且 $0_{\text{MIN}}$ 指示1和0之间的数。

10. 根据权利要求8所述的装置,其中,对于索引n和帧k,当n是在所述第一索引集合中时,所述第一子集基于对应的环境声音分量 $\hat{\mathbf{c}}_{\text{AMB},n}(k-1)$ 被确定,并且当n在所述第二索引集合中时,所述第二子集基于对应的主导声音分量 $\hat{\mathbf{c}}_{\text{PS},n}(k-1)$ 和对应的环境声音分量 $\hat{\mathbf{c}}_{\text{AMB},n}(k-1)$ 的相加而被确定,并且其中,解码的HOA表示至少部分地由下式表示:

$$\tilde{\mathbf{c}}_n(k-1) = \begin{cases} \hat{\mathbf{c}}_{\text{AMB},n}(k-1) & \text{对于在所述第一索引集合中的 } n \\ \hat{\mathbf{c}}_n(k-1) = \hat{\mathbf{c}}_{\text{PS},n}(k-1) + \hat{\mathbf{c}}_{\text{AMB},n}(k-1), & \text{对于在所述第二索引集合中的 } n \end{cases}$$

11. 根据权利要求8所述的装置,其中, $0_{\text{MIN}} = (N_{\text{MIN}}+1)^2$ 且 $N_{\text{MIN}} \leq N$ ,其中N是编码的HOA表示的输入帧的阶次。

12. 根据权利要求8所述的装置,其中,多个层的指示在所述位流中以信号告知。

13. 根据权利要求8所述的装置,其中,所述多个层包括基本层和至少一个增强层。

14. 根据权利要求8所述的装置,其中,所述音频解码器进一步被配置为基于不存在多个层的确定而确定存在单个层,并且基于单个层的确定,基于对应的主导HOA声音分量 $(\hat{\mathbf{c}}_{\text{PS}}(k-1))$ 和对应的环境HOA声音分量 $(\hat{\mathbf{c}}_{\text{AMB}}(k-1))$ 的相加来确定单层解码的HOA表示。

15. 一种设备,所述设备包括:

一个或多个处理器;和

存储有指令的一个或多个存储介质,所述指令在由所述一个或多个处理器执行时使得

执行根据权利要求1-7中任一项所述的方法。

16. 一种包括用于执行根据权利要求1-7中任一项所述的方法的部件的装置。

17. 一种存储有指令的非暂时性计算机可读存储介质, 所述指令在被一个或多个处理器执行时使得执行根据权利要求1-7中任一项所述的方法。

## 对压缩的高阶高保真立体声 (HOA) 表示进行解码的方法及装置

### 技术领域

[0001] 本发明涉及一种用于压缩高阶高保真立体声 (HOA) 信号的方法、用于解压缩压缩的 HOA 信号的方法、用于压缩 HOA 信号的装置以及用于解压缩压缩的 HOA 信号的装置。

### 背景技术

[0002] 高阶高保真立体声 (HOA) 提供了表示三维声音的可能性。其他已知的技术是波场合成 (WFS) 或基于通道的方法 (比如 22.2)。然而,与基于通道的方法相反,HOA 表示提供了独立于特定扩音器设置的优点。然而,该灵活性是以 HOA 表示在特定扩音器设置上的回放需要解码处理为代价的。与所需扩音器的数量通常非常大的 WFS 方法相比,HOA 也可以被渲染到仅由少量扩音器组成的设置。HOA 的进一步的优点是,相同的表示也可以被没有任何修正地用于对于耳机的双耳渲染。

[0003] HOA 是基于通过截断球谐函数 (SH) 展开得出的复谐波平面波振幅的所谓空间密度的表示。每个展开系数是角频率的函数,其可以等同地用时域函数表示。因此,不失一般性地,整个 HOA 声场表示实际上可以被假定为由 0 个时域函数组成,其中,0 表示展开系数的数量。在下面,这些时域函数将被等同地称为 HOA 系数序列或 HOA 通道。通常,使用球坐标系,在球坐标系中,x 轴指向前向的位置,y 轴指向左边,z 轴指向顶部。空间  $x = (r, \theta, \phi)^T$  中的位置用半径  $r > 0$  (即,到坐标原点的距离)、从极轴 z 测量的倾角  $\theta \in [0, \pi]$ 、以及在 x-y 平面中从 x 轴逆时针测量的方位角  $\phi \in [0, 2\pi]$  表示。此外,  $(\cdot)^T$  表示转置。

[0004] 下面提供 HOA 编码的更详细的描述。

[0005] 用  $\mathcal{F}_t(\cdot)$  表示的声压相对于时间的傅立叶变换 (即,

$P(\omega, x) = \mathcal{F}_t(p(t, x)) = \int_{-\infty}^{\infty} p(t, x) e^{-i\omega t} dt$ , 其中,  $\omega$  表示角频率,  $i$  指示虚数单位) 可以根据

$P(\omega = kc_s, r, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n A_n^m(k) j_n(kr) S_n^m(\theta, \phi)$  被展开为球谐函数的级数。

[0006] 这里,  $c_s$  表示声音的速度,  $k$  表示通过  $k = \frac{\omega}{c_s}$  与角频率  $\omega$  相关的角波数,  $j_n(\cdot)$  表示第一种球贝塞尔函数,  $S_n^m(\theta, \phi)$  表示阶次  $n$  和度数  $m$  的实数值球谐函数。展开系数  $A_n^m(k)$  仅取决于角波数  $k$ 。注意到,已经隐含地假定声压在空间上是带限的。因此,级数关于阶次索引  $n$  在上限  $N$  处被截断,阶次索引  $n$  被称为 HOA 表示的阶次。如果声场用不同角频率  $\omega$  的并且从角度元组  $(\theta, \phi)$  指定的所有可能的方向到达的无限个谐波平面波的叠加来表示,则相应的平面波复振幅函数  $C(\omega, \theta, \phi)$  可以用以下球谐函数展开来表达:

[0007]  $C(\omega = kc_s, \theta, \phi) = \sum_{n=0}^N \sum_{m=-n}^n C_n^m(k) S_n^m(\theta, \phi),$

[0008] 其中,展开系数  $C_n^m(k)$  通过  $A_n^m(k) = i^n C_n^m(k)$  与展开系数  $A_n^m(k)$  相关。

[0009] 假定各个系数  $C_n^m(\omega = kc_s)$  是角频率  $\omega$  的函数,那么逆傅立叶变换 (用  $\mathcal{F}^{-1}(\cdot)$  表示) 的应用对于每个阶次  $n$  和度数  $m$  提供时域函数:



$$[0010] \quad c_n^m(t) = \mathcal{F}_t^{-1}(C_n^m(\omega/c_s)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_n^m\left(\frac{\omega}{c_s}\right) e^{i\omega t} d\omega$$

[0011] 这些时域函数可以通过

$c(t) = [c_0^0(t) \ c_1^{-1}(t) \ c_1^0(t) \ c_1^1(t) \ c_2^{-2}(t) \ c_2^{-1}(t) \ c_2^0(t) \ \dots \ c_N^{N-1}(t) \ c_N^N(t)]^T$  聚集在单个矢量  $c(t)$  中。时域函数  $c_n^m(t)$  在矢量  $c(t)$  内的位置索引由  $n(n+1)+1+m$  给出。矢量  $c(t)$  中的元素的总数由  $0 = (N+1)^2$  给出。函数  $c_n^m(t)$  的离散时间版本被称为高保真立体声系数序列。基于帧的 HOA 表示通过将所有这些序列划分为如下的长度为  $B$ 、索引为  $k$  的帧  $C(k)$  而获得：

$$[0012] \quad C(k) := [c((kB+1)T_s) \ c((kB+2)T_s) \ \dots \ c((kB+B)T_s)],$$

[0013] 其中,  $T_s$  表示采样周期。帧  $C(k)$  本身于是可以如下地表示为其各个行  $c_i(k)$ ,  $i = 1, \dots, 0$  的复合：

$$[0014] \quad C(k) = \begin{bmatrix} c_1(k) \\ c_2(k) \\ \vdots \\ c_0(k) \end{bmatrix}$$

[0015] 其中,  $c_i(k)$  表示高保真立体声系数序列的位置索引为  $i$  的帧。HOA 表示的空间分辨率随着展开的最大阶次  $N$  增长而改进。不幸的是, 展开系数的数量  $0$  随着阶次  $N$  二次方地增长, 具体为  $0 = (N+1)^2$ 。例如, 典型的使用阶次  $N=4$  的 HOA 表示需要  $0=25$  个 HOA (展开) 系数。根据这些考虑, 给定期望的单通道采样速率  $f_s$  和每个采样的位数  $N_b$ , 用于传输 HOA 表示的总位速率由  $0 \cdot f_s \cdot N_b$  确定。因此, 每个采样利用  $N_b=16$  个位、以  $f_s=48\text{kHz}$  的采样速率传输阶次  $N=4$  的 HOA 表示导致  $19.2\text{Mbits/s}$  的位速率, 该位速率对于许多实际的应用 (例如流传输) 是非常高的。因此, HOA 表示的压缩是非常希望的。

[0016] 以前, HOA 声场表示的压缩在欧洲专利申请 EP2743922A、EP2665208A 和 EP2800401A 中被提出过。这些方法的共同之处是, 它们执行声场分析, 并且将给定的 HOA 表示分解为方向分量和残差环境分量。

[0017] 最终的压缩的表示被假定为一方面, 包括若干个由方向信号的感知编码得到的量化信号以及环境 HOA 分量的相关系数序列。另一方面, 它被假定为包括与量化信号相关的附加副信息, 该副信息对于从 HOA 表示的压缩版本重构 HOA 表示是必需的。

[0018] 此外, 类似的方法在 ISO/IEC JTC1/SC29/WG11 N14264 (Working draft 1-HOA text of MPEG-H 3D audio, 2014 年 1 月, San Jose) 中有描述, 其中, 方向分量扩展为所谓的主导声音分量。作为方向分量, 主导声音分量被假定为部分地用方向信号 (即, 具有对应方向的单耳信号, 这些单耳信号被假定为从该方向传到收听者)、连同用于从方向信号预测原始 HOA 表示的各部分的一些预测参数来表示。

[0019] 另外, 主导声音分量被假设为用所谓的基于矢量的信号来表示, 基于矢量的信号意指具有对应矢量的单耳信号, 该矢量限定基于矢量的信号的方向分布。已知的压缩的 HOA 表示由  $I$  个量化的单耳信号和某些附加的副信息构成, 其中, 这些  $I$  个量化的单耳信号中的固定数量  $0_{\text{MIN}}$  个单耳信号表示环境 HOA 分量  $C_{\text{AMB}}(k-2)$  的前  $0_{\text{MIN}}$  个系数序列的空间变换版本。其余的  $I-0_{\text{MIN}}$  个信号的类型在连续的帧之间可以有所变化, 并且可以是方向的、基于矢量的、空的、或者表示环境 HOA 分量  $C_{\text{AMB}}(k-2)$  的附加系数序列。

[0020] 用于压缩具有 HOA 系数序列的输入时间帧 ( $C(k)$ ) 的 HOA 信号表示的已知方法包括对

输入时间帧进行空间HOA编码、随后进行感知编码和源编码。如图1a)所示的空间HOA编码包括在方向和矢量估计模块101中执行HOA信号的方向和矢量估计处理,其中,包括关于方向信号的第一元组集合 $\mathcal{M}_{\text{DIR}}(k)$ 和关于基于矢量的信号的第二元组集合 $\mathcal{M}_{\text{VEC}}(k)$ 的数据被获得。第一元组集合中的每个包括方向信号的索引和相应的量化方向,第二元组集合中的每个包括基于矢量的信号的索引和限定信号的方向分布的矢量。下一个步骤是将HOA系数序列的每个输入时间帧分解103为多个主导声音信号 $X_{\text{PS}}(k-1)$ 的帧和环境 HOA分量 $C_{\text{AMB}}(k-1)$ 的帧,其中,主导声音信号 $X_{\text{PS}}(k-1)$ 包括所述方向声音信号和所述基于矢量的声音信号。分解进一步提供预测参数 $\xi(k-1)$ 和目标分配矢量 $v_{\text{A},\text{T}}(k-1)$ 。预测参数 $\xi(k-1)$ 描述如何从主导声音信号 $X_{\text{PS}}(k-1)$ 内的方向信号预测HOA信号表示的各部分以便丰富主导声音HOA分量,目标分配矢量 $v_{\text{A},\text{T}}(k-1)$ 包含关于如何将主导声音信号分配给给定数量 $I$ 个通道的信息。根据目标分配矢量 $v_{\text{A},\text{T}}(k-1)$ 提供的信息修正104环境 HOA分量 $C_{\text{AMB}}(k-1)$ ,其中,确定环境HOA分量的哪些系数序列将在给定数量 $I$ 个通道中传输,这取决于有多少个通道被主导声音信号占据。修正的环境HOA分量 $C_{\text{M},\text{A}}(k-2)$ 和时间预测的修正的环境HOA分量 $C_{\text{P},\text{M},\text{A}}(k-1)$ 被获得。此外,最终的分配矢量 $v_{\text{A}}(k-2)$ 从目标分配矢量 $v_{\text{A},\text{T}}(k-1)$ 中的信息获得。使用最终的分配矢量 $v_{\text{A}}(k-2)$ 提供的信息将通过分解获得的主导声音信号 $X_{\text{PS}}(k-1)$ 以及所确定的修正的环境HOA分量 $C_{\text{M},\text{A}}(k-2)$ 和时间预测的修正的环境HOA分量 $C_{\text{P},\text{M},\text{A}}(k-1)$ 的系数序列分配到给定数量的通道,其中,输送信号 $y_i(k-2)$ ,  $i=1, \dots, I$ 和预测的输送信号 $y_{\text{P},i}(k-2)$ ,  $i=1, \dots, I$ 被获得。然后,对输送信号 $y_i(k-2)$ 和预测的输送信号 $y_{\text{P},i}(k-2)$ 执行增益控制(或规范化),其中,增益修正的输送信号 $z_i(k-2)$ 、指数 $e_i(k-2)$ 和异常标志 $\beta_i(k-2)$ 被获得。

[0021] 如图1b)所示,感知编码和源编码包括:对增益修正的输送信号 $z_i(k-2)$ 进行感知编码,其中,感知编码的输送信号 $\tilde{z}_i(k-2)$ ,  $i=1, \dots, I$ 被获得;对副信息进行编码,副信息包括所述指数 $e_i(k-2)$ 和异常标志 $\beta_i(k-2)$ 、第一元组集合 $\mathcal{M}_{\text{DIR}}(k)$ 和第二元组集合 $\mathcal{M}_{\text{VEC}}(k)$ 、预测参数 $\xi(k-1)$ 和最终的分配矢量 $v_{\text{A}}(k-2)$ ,并且编码的副信息 $\tilde{\Gamma}(k-2)$ 被获得。最后,感知编码的输送信号 $\tilde{z}_i(k-2)$ 和编码的副信息被复用到位流中。

## 发明内容

[0022] 提议的HOA压缩方法的一个缺点是它提供整体式的(即,不可缩放的)压缩的HOA表示。然而,对于某些应用,比如广播或互联网流传输,可取的是能够将压缩的表示划分为低质量基本层(BL)和高质量增强层(EL)。基本层被假设为提供HOA表示的低质量压缩版本,该版本可以被独立于增强层进行解码。这样的BL对传输误差通常应是高度鲁棒的,并且应被以低数据速率传输以便即使在不良传输状况下也保证解压缩的HOA表示的某个最低质量。EL包含改进解压缩的HOA表示的质量的附加信息。

[0023] 本发明提供一种用于修改现有的HOA压缩方法以便能够提供包括(低质量)基本层和(高质量)增强层的压缩表示的解决方案。此外,本发明提供一种用于修改现有的HOA解压缩方法以便能够对根据本发明压缩的至少包括低质量基本层的压缩表示进行解码的解决方案。

[0024] 一个改进涉及获得自含的(低质量)基本层。根据本发明,被假设为包含环境HOA分



量 $C_{AMB}(k-2)$ 的(不失一般性的)前 $O_{MIN}$ 个系数序列的空间变换版本的 $O_{MIN}$ 个通道被用作基本层。选择前 $O_{MIN}$ 个通道形成基本层的优点是它们的时不变类型。然而,照惯例,各个信号缺少声音场景必需的任何主导声音分量。这从环境HOA分量 $C_{AMB}(k-1)$ 的常规计算也是清楚的,环境HOA分量 $C_{AMB}(k-1)$ 的常规计算是通过根据下式从原始HOA表示 $C(k-1)$ 减去主导声音HOA表示 $C_{PS}(k-1)$ 来进行的:

$$[0025] \quad C_{AMB}(k-1) = C(k-1) - C_{PS}(k-1) \quad (1)$$

[0026] 因此,本发明的一个改进涉及这样的主导声音分量的添加。根据本发明,该问题的解决方案是将低空间分辨率的主导声音分量包括到基本层中。为了这个目的,通过根据本发明的空间HOA编码器中的 HOA分解处理输出的环境HOA分量 $C_{AMB}(k-1)$ 被其修正版本取代。修正的环境HOA分量将原始HOA分量的系数序列包括在被假设总是以空间变换形式传输的前 $O_{MIN}$ 个系数序列中。HOA分解处理的这个改进可以被看作使HOA压缩在分层模式(例如双层模式)下工作的初始操作。该模式提供例如两个位流、或者可以被划分为基本层和增强层的单个位流。使用或者不使用该模式是由总位流的访问单元中的模式指示位(例如,单个位)来信号告知的。

[0027] 在一个实施例中,基本层位流 $\bar{B}_{BASE}(k-2)$ 仅包括感知编码的信号 $\bar{z}_i(k-2)$ ,  $i = 1, \dots, O_{MIN}$ 以及对应的编码的增益控制副信息,该副信息由指数 $e_i(k-2)$ 和异常标志 $\beta_i(k-2)$ ,  $i = 1, \dots, O_{MIN}$ 构成。其余的感知编码的信号 $\bar{z}_i(k-2)$ ,  $i = O_{MIN}+1, \dots, 0$ 和编码的其余的副信息被包括到增强层位流中。在一个实施例中,代替前述总位流 $\bar{B}(k-2)$ ,基本层位流 $\bar{B}_{BASE}(k-2)$ 和增强层位流 $\bar{B}_{ENH}(k-2)$ 然后被联合传输。

[0028] 用于压缩具有高阶高保真立体声(HOA)系数序列的时间帧的 HOA信号表示的方法在权利要求1中公开。用于压缩具有高阶高保真立体声(HOA)系数序列的时间帧的HOA信号表示的装置在权利要求10中公开。

[0029] 用于解压缩具有高阶高保真立体声(HOA)系数序列的时间帧的 HOA信号表示的方法在权利要求8中公开。用于解压缩具有高阶高保真立体声(HOA)系数序列的时间帧的HOA信号表示的装置在权利要求18中公开。

[0030] 具有使计算机执行用于压缩具有高阶高保真立体声(HOA)系数序列的时间帧的HOA信号表示的方法的可执行指令的非暂时性计算机可读存储介质在权利要求20中公开。

[0031] 具有使计算机执行用于解压缩具有高阶高保真立体声(HOA)系数序列的时间帧的HOA信号表示的方法的可执行指令的非暂时性计算机可读存储介质在权利要求21中公开。

[0032] 本发明的有利的实施例在从属权利要求、下面的描述和附图中公开。

## 附图说明

[0033] 参照附图描述本发明的示例性实施例,附图在以下图中示出了:

[0034] 图1HOA压缩器的常规架构的结构;

[0035] 图2HOA解压缩器的常规架构的结构;

[0036] 图3根据本发明的一个实施例的HOA压缩器的空间HOA编码和感知编码部分的架构的结构;

- [0037] 图4根据本发明的一个实施例的HOA压缩器的源编码器部分的架构的结构；
- [0038] 图5根据本发明的一个实施例的HOA解压缩器的感知解码和源解码部分的架构的结构；
- [0039] 图6根据本发明的一个实施例的HOA解压缩器的空间HOA 解码部分的架构的结构；
- [0040] 图7从环境HOA信号到修正的环境HOA信号的帧变换；
- [0041] 图8用于压缩HOA信号的方法的流程图；
- [0042] 图9用于解压缩压缩的HOA信号的方法的流程图；以及
- [0043] 图10根据本发明的一个实施例的HOA解压缩器的空间HOA 解码部分的架构的各部分的细节。

## 具体实施方式

[0044] 为了更易于理解,下面概括了图1和图2中的现有技术的解决方案。

[0045] 图1示出了HOA压缩器的常规架构的结构。在[4]中描述的方法中,方向分量扩展为所谓的主导声音分量。作为方向分量,主导声音分量被假定为部分地用方向信号(指的是具有对应方向的单耳信号,这些单耳信号被假定为从该方向传到收听者)、连同用于从方向信号预测原始HOA表示的各部分的一些预测参数来表示。另外,主导声音分量被假设为用所谓的基于矢量的信号来表示,基于矢量的信号意指具有对应矢量的单耳信号,该矢量限定了基于矢量的信号的方向分布。[4]中提议的HOA压缩器的总体架构在图1中示出。它可以被再分为图1a中描绘的空间HOA编码部分与图1b中描绘的感知和源编码部分。空间HOA编码器提供第一压缩HOA表示,该第一压缩HOA 表示由I个信号连同描述如何创建其HOA表示的副信息构成。在感知和副信息源编码器中,对所提及的I个信号进行感知编码,并且对副信息进行源编码,之后对该两个编码的表示进行复用。

[0046] 照惯例,空间编码如下工作。

[0047] 在第一步中,原始HOA表示的第k帧C(k)输入到方向和矢量估计处理模块,该模块提供元组集合 $\mathcal{M}_{\text{DIR}}(k)$ 和 $\mathcal{M}_{\text{VEC}}(k)$ 。元组集合 $\mathcal{M}_{\text{DIR}}(k)$ 由其第一元素表示方向信号的索引、其第二元素表示相应的量化方向的元组构成。元组集合 $\mathcal{M}_{\text{VEC}}(k)$ 由其第一元素指示基于矢量的信号的索引、其第二元素表示限定信号的方向分布的矢量(即,基于矢量的信号的HOA 表示如何被计算)的元组构成。

[0048] 通过使用这两个元组集合 $\mathcal{M}_{\text{DIR}}(k)$ 和 $\mathcal{M}_{\text{VEC}}(k)$ ,初始HOA帧C(k)在HOA 分解中被分解为所有主导声音信号(即,方向信号和基于矢量的信号)的帧 $X_{\text{PS}}(k-1)$ 、和环境HOA分量的帧 $C_{\text{AMB}}(k-1)$ 。注意到分别有一个帧的延迟,该延迟是由于为了避免成块效应的重叠相加处理而导致的。此外,HOA分解被假定输出一些预测参数 $\zeta(k-1)$ ,这些预测参数描述如何从方向信号预测原始HOA表示的各部分以便丰富主导声音HOA分量。另外,目标分配矢量 $\mathbf{v}_{\text{A,T}}(k-1)$ 被提供,该目标分配矢量 $\mathbf{v}_{\text{A,T}}(k-1)$ 包含关于在HOA分解处理模块中确定的主导声音信号对于I个可用通道的分配的信息。受到影响的通道可以被假定被占据,这意味着它们不可用于在相应时间帧中输送环境HOA分量的任何系数序列。

[0049] 在环境分量修正处理模块中,根据目标分配矢量 $\mathbf{v}_{\text{A,T}}(k-1)$ 提供的信息修正环境HOA分量的帧 $C_{\text{AMB}}(k-1)$ 。特别地,确定环境HOA分量的哪些系数序列将在给定的I个通道中传

输,这尤其取决于关于哪些通道是可用的、但尚未被主导声音信号占据的信息(该信息包含在目标分配矢量 $\mathbf{v}_{A,T}(k-1)$ 中)。另外,如果被选系数序列的索引在连续的帧之间变化,则系数序列的淡入或淡出被执行。

[0050] 此外,假定环境HOA分量 $C_{AMB}(k-2)$ 的前 $0_{MIN}$ 个系数序列总是被选择被进行感知编码并且被传输,其中, $0_{MIN} = (N_{MIN}+1)^2$ ,  $N_{MIN} \leq N$ 通常是比原始HOA表示的阶次小的阶次。为了使这些HOA系数序列去相关,提出了将它们变换为从一些预定义方向 $\Omega_{MIN,d}, d=1, \dots, 0_{MIN}$ 传来的方向信号(即,一般平面波函数)。连同修正的环境HOA分量 $C_{M,A}(k-1)$ 一起,时间预测的修正的环境HOA分量 $C_{P,M,A}(k-1)$ 被计算以供以后用在增益控制处理模块中以便允许有合理的前瞻。

[0051] 关于环境HOA分量的修正的信息与所有可能的类型的信号对于可用通道的分配直接相关。关于分配的最终信息包含在最终的分配矢量 $\mathbf{v}_A(k-2)$ 中。为了计算该矢量,利用目标分配矢量 $\mathbf{v}_{A,T}(k-1)$ 中所包含的信息。

[0052] 通道分配使用分配矢量 $\mathbf{v}_A(k-2)$ 提供的信息将 $X_{PS}(k-2)$ 中所包含的和 $C_{M,A}(k-2)$ 中所包含的适当信号分配给 $I$ 个可用通道,得到信号 $y_i(k-2), i=1, \dots, I$ 。此外, $X_{PS}(k-1)$ 中所包含的和 $C_{P,AMB}(k-1)$ 中所包含的适当信号也被分配给 $I$ 个可用通道,得到预测信号 $y_{P,i}(k-2), i=1, \dots, I$ 。信号 $y_i(k-2), i=1, \dots, I$ 中的每个最后均被增益控制进行处理,其中,信号增益被平滑地修正以实现适合于感知编码器的值范围。预测信号帧 $y_{P,i}(k-2), i=1, \dots, I$ 允许有一种前瞻以便避免连续的块之间的严重的增益变化。增益修正被假定在空间解码器中通过增益控制副信息被恢复,增益控制副信息由指数 $e_i(k-2)$ 和异常标志 $\beta_i(k-2), i=1, \dots, I$ 构成。

[0053] 图2示出了如[4]中提出的HOA解压缩器的常规架构的结构。照惯例,HOA解压缩由HOA压缩器组件的对应物构成,这些对应物显然按相反次序布置。它可以被再分为图2a)中描绘的感知和源解码部分与图2b)中描绘的空间HOA解码部分。

[0054] 在感知和副信息源解码器中,位流首先被解复用为 $I$ 个信号的感知编码的表示以及编码的副信息,该副信息描述如何创建其HOA表示。相继地,执行所述 $I$ 个信号的感知解码和副信息的解码。然后,空间HOA解码器从所述 $I$ 个信号和副信息创建重构的HOA表示。

[0055] 照惯例,空间HOA解码如下工作。

[0056] 在空间HOA解码器中,感知解码的信号 $\hat{\mathbf{z}}_i(k), i \in \{1, \dots, I\}$ 中的每个首先连同相关联的增益校正指数 $e_i(k)$ 和增益校正异常标志 $\beta_i(k)$ 一起输入到逆增益控制处理模块。第 $i$ 逆增益控制处理提供增益校正的信号帧 $\hat{\mathbf{y}}_i(k)$ 。

[0057] 所有的 $I$ 个增益校正的信号帧 $\hat{\mathbf{y}}_i(k), i \in \{1, \dots, I\}$ 连同分配矢量 $\mathbf{v}_{AMB,ASSIGN}(k)$ 以及元组集合 $\mathcal{M}_{DIR}(k+1)$ 和 $\mathcal{M}_{VEC}(k+1)$ 一起被传递到通道再分配。元组集合 $\mathcal{M}_{DIR}(k+1)$ 和 $\mathcal{M}_{VEC}(k+1)$ 在上文被定义(针对空间HOA编码),分配矢量 $\mathbf{v}_{AMB,ASSIGN}(k)$ 由 $I$ 个分量构成,这些分量对于每个传输通道指示它是否包含环境HOA分量的系数序列以及它包含环境HOA分量的哪个系数序列。在通道再分配中,增益校正的信号帧 $\hat{\mathbf{y}}_i(k)$ 被重新分布以重构所有的主导声音信号(即,所有的方向信号和基于矢量的信号)的帧 $\hat{\mathbf{x}}_{PS}(k)$ 以及环境HOA分量的中间表示的帧 $C_{I,AMB}(k)$ 。另外,环境HOA分量的在第 $k$ 帧中起作用的系数序列的索引集合



$\mathcal{I}_{\text{AMB,ACT}}(k)$  以及环境HOA分量的在第  $(k-1)$  帧中必须被启用、被禁用以及保持起作用的系数索引集合  $\mathcal{I}_{\text{E}}(k-1)$ 、 $\mathcal{I}_{\text{D}}(k-1)$  和  $\mathcal{I}_{\text{U}}(k-1)$  被提供。

[0058] 在主导声音合成中,使用元组集合  $\mathcal{M}_{\text{DIR}}(k+1)$  和预测参数集合  $\zeta(k+1)$ 、元组集合  $\mathcal{M}_{\text{VEC}}(k+1)$  和集合  $\mathcal{I}_{\text{E}}(k-1)$ 、 $\mathcal{I}_{\text{D}}(k-1)$  和  $\mathcal{I}_{\text{U}}(k-1)$  从所有主导声音信号的帧  $\hat{\mathbf{X}}_{\text{PS}}(k)$  计算主导声音分量  $\tilde{\mathbf{c}}_{\text{PS}}(k-1)$  的HOA表示。

[0059] 在环境合成中,使用环境HOA分量的在第  $k$  帧中起作用的系数序列的索引集合  $\mathcal{I}_{\text{AMB,ACT}}(k)$  从环境HOA分量的中间表示的帧  $\mathbf{C}_{\text{I,AMB}}(k)$  创建环境HOA分量帧  $\tilde{\mathbf{c}}_{\text{AMB}}(k-1)$ 。注意到有一帧的延迟,该延迟是由于与主导声音HOA分量的同步而引入的。最后,在HOA复合中,环境HOA 分量帧  $\tilde{\mathbf{c}}_{\text{AMB}}(k-1)$  和主导声音HOA分量的帧  $\tilde{\mathbf{c}}_{\text{PS}}(k-1)$  被叠加以提供解码的HOA帧  $\tilde{\mathbf{c}}(k-1)$ 。

[0060] 从以上HOA压缩和解压缩方法的粗略描述已经变得清楚的是,压缩的表示由  $I$  个量化的单耳信号和某些附加的副信息构成。这些  $I$  个量化的单耳信号中的固定数量  $O_{\text{MIN}}$  个单耳信号表示环境HOA分量  $\mathbf{C}_{\text{AMB}}(k-2)$  的前  $O_{\text{MIN}}$  个系数序列的空间变换版本。其余的  $I-O_{\text{MIN}}$  个信号的类型可在连续的帧之间变化,是方向的、基于矢量的、空的、或者表示环境HOA分量  $\mathbf{C}_{\text{AMB}}(k-2)$  的附加系数序列。照其原样来看,压缩的HOA 表示意在于是整体式的。具体地说,一个问题是如何将所描述的表示划分为低质量基本层和增强层。

[0061] 根据公开的本发明,低质量基本层的候选是包含环境HOA分量  $\mathbf{C}_{\text{AMB}}(k-2)$  的前  $O_{\text{MIN}}$  个系数序列的空间变换版本的  $O_{\text{MIN}}$  个通道。使这些 (不失一般性地,前)  $O_{\text{MIN}}$  个通道成为形成低质量基本层的良好选择的是它们的时不变类型。然而,相应的信号缺少声音场景必需的任何主导声音分量。这从环境HOA分量  $\mathbf{C}_{\text{AMB}}(k-1)$  的常规计算也可以看出,环境 HOA分量  $\mathbf{C}_{\text{AMB}}(k-1)$  的常规计算是通过根据下式从原始HOA表示  $\mathbf{C}(k-1)$  减去主导声音HOA表示  $\mathbf{C}_{\text{PS}}(k-1)$  来进行的:

$$[0062] \quad \mathbf{C}_{\text{AMB}}(k-1) = \mathbf{C}(k-1) - \mathbf{C}_{\text{PS}}(k-1) \quad (1)$$

[0063] 该问题的解决方案是将低空间分辨率的主导声音分量包括到基本层中。

[0064] 下面描述提出的对于HOA压缩的修改。

[0065] 图3示出了根据本发明的一个实施例的HOA压缩器的空间HOA 编码和感知编码部分的架构的结构。为了还将低空间分辨率的主导声音分量包括基本层中,通过空间HOA编码器 (参见图1a) 中的HOA 分解处理输出的环境HOA分量  $\mathbf{C}_{\text{AMB}}(k-1)$  被修正版本取代:

$$[0066] \quad \tilde{\mathbf{c}}_{\text{AMB}}(k-1) = \begin{bmatrix} \tilde{\mathbf{c}}_{\text{AMB},1}(k-1) \\ \tilde{\mathbf{c}}_{\text{AMB},2}(k-1) \\ \vdots \\ \tilde{\mathbf{c}}_{\text{AMB},O}(k-1) \end{bmatrix} \quad (2)$$

[0067] 该修正版本的元素由下式给出:

$$[0068] \quad \tilde{\mathbf{c}}_{\text{AMB},n}(k-1) = \begin{cases} \mathbf{c}_n(k-1) & \text{对于 } 1 \leq n \leq O_{\text{MIN}} \\ \mathbf{c}_{\text{AMB},n}(k-1) & \text{对于 } O_{\text{MIN}} + 1 \leq n \leq O \end{cases} \quad (3)$$

[0069] 换句话说,被假设总是被以空间变换形式传输的环境HOA分量的前 $O_{\text{MIN}}$ 个系数序列被原始HOA分量的系数序列取代。空间HOA编码器的其他处理模块可以保持不变。

[0070] 重要的是注意到,HOA分解处理的这个变化可以被看作使HOA 压缩在所谓的“双层”或“两层”模式下工作的初始操作。该模式提供了可以被划分为低质量基本层和增强层的位流。使用或者不使用该模式是由总位流的访问单元中的单个位以信号告知的。

[0071] 提供用于基本层和增强层的位流的位流复用的可能的随后的修正在下面进一步描述的图3和4中示出。

[0072] 基本层位流 $\tilde{\mathbf{B}}_{\text{BASE}}(k-2)$ 仅包括感知编码的信号 $\tilde{\mathbf{z}}_i(k-2)$ ,  $i = 1, \dots, O_{\text{MIN}}$ 以及对应的编码的增益控制副信息,该副信息由指数 $e_i(k-2)$ 和异常标志 $\beta_i(k-2)$ ,  $i = 1, \dots, O_{\text{MIN}}$ 构成。其余的感知编码的信号 $\tilde{\mathbf{z}}_i(k-2)$ ,  $i = O_{\text{MIN}} + 1, \dots, O$ 和编码的其余的副信息被包括到增强层位流中。代替前述总位流 $\tilde{\mathbf{B}}(k-2)$ ,基本层位流 $\tilde{\mathbf{B}}_{\text{BASE}}(k-2)$ 和增强层位流 $\tilde{\mathbf{B}}_{\text{ENH}}(k-2)$ 然后被联合传输。

[0073] 在图3和图4中,示出了用于压缩HOA信号的装置,所述HOA 信号是具有HOA系数序列的输入时间帧(C(k))的输入HOA表示。所述装置包括用于输入时间帧的空间HOA编码以及随后的感知编码的空间HOA编码和感知编码部分(该部分在图3中示出)以及用于源编码的源编码器部分(该部分在图4中示出)。空间HOA编码和感知编码部分包括方向和矢量估计模块301、HOA分解模块303、环境分量修正模块304、通道分配模块305以及多个增益控制模块306。

[0074] 方向和矢量估计模块301适于执行HOA信号的方向和矢量估计处理,其中,包括关于方向信号的第一元组集合 $\mathcal{M}_{\text{DIR}}(k)$ 和关于基于矢量的信号的第二元组集合 $\mathcal{M}_{\text{VEC}}(k)$ 的数据被获得,第一元组集合 $\mathcal{M}_{\text{DIR}}(k)$ 中的每个包括方向信号的索引和相应的量化方向,第二元组集合 $\mathcal{M}_{\text{VEC}}(k)$ 中的每个包括基于矢量的信号的索引和限定信号的方向分布的矢量。

[0075] HOA分解模块303适于将HOA系数序列的每个输入时间帧分解为多个主导声音信号 $X_{\text{PS}}(k-1)$ 的帧和环境HOA分量 $\tilde{\mathbf{C}}_{\text{AMB}}(k-1)$ 的帧,其中,主导声音信号 $X_{\text{PS}}(k-1)$ 包括所述方向声音信号和所述基于矢量的声音信号,并且其中,环境HOA分量 $\tilde{\mathbf{C}}_{\text{AMB}}(k-1)$ 包括表示输入HOA表示和主导声音信号的HOA表示之间的残差的HOA系数序列,并且其中,分解进一步提供预测参数 $\xi(k-1)$ 和目标分配矢量 $\mathbf{v}_{\text{A},\text{T}}(k-1)$ 。预测参数 $\xi(k-1)$ 描述如何从主导声音信号 $X_{\text{PS}}(k-1)$ 内的方向信号预测HOA信号表示的各部分以便丰富主导声音HOA分量,目标分配矢量 $\mathbf{v}_{\text{A},\text{T}}(k-1)$ 包含关于如何将主导声音信号分配给给定数量I个通道的信息。

[0076] 环境分量修正模块304适于根据目标分配矢量 $\mathbf{v}_{\text{A},\text{T}}(k-1)$ 提供的信息修正环境HOA分量 $\mathbf{C}_{\text{AMB}}(k-1)$ ,其中,确定环境HOA分量 $\mathbf{C}_{\text{AMB}}(k-1)$ 的哪些系数序列将在给定数量I个通道中传输,这取决于有多少个通道被主导声音信号占据,并且其中,修正的环境HOA分量 $\mathbf{C}_{\text{M},\text{A}}(k-2)$ 和时间预测的修正的环境HOA分量 $\mathbf{C}_{\text{P},\text{M},\text{A}}(k-1)$ 被获得,并且其中,最终的分配矢量 $\mathbf{v}_{\text{A}}(k-2)$ 从目标分配矢量 $\mathbf{v}_{\text{A},\text{T}}(k-1)$ 中的信息获得。

[0077] 通道分配模块305适于使用最终的分配矢量 $\mathbf{v}_{\text{A}}(k-2)$ 提供的信息将从分解获得的主导声音信号 $X_{\text{PS}}(k-1)$ 、所确定的修正的环境HOA分量 $\mathbf{C}_{\text{M},\text{A}}(k-2)$ 和时间预测的修正的环境



HOA分量 $C_{P,M,A}(k-1)$ 的系数序列分配给给定数量 $I$ 个通道,其中,输送信号 $y_i(k-2)$ ,  $i=1, \dots, I$ 和预测的输送信号 $y_{P,i}(k-2)$ ,  $i=1, \dots, I$ 被获得。

[0078] 所述多个增益控制模块306适于对输送信号 $y_i(k-2)$ 和预测的输送信号 $y_{P,i}(k-2)$ 执行增益控制(805),其中,增益修正的输送信号 $z_i(k-2)$ 、指数 $e_i(k-2)$ 和异常标志 $\beta_i(k-2)$ 被获得。

[0079] 图4示出了根据本发明的一个实施例的HOA压缩器的源编码器部分的架构的结构。如图4所示的源编码器部分包括感知编码器310、具有两个编码器320、330(即,基本层副信息源编码器320和增强层副信息编码器330)的副信息源编码器模块、以及两个复用器340、350(即,基本层位流复用器340和增强层位流复用器350)。副信息源编码器可以在单个副信息源编码器模块中。

[0080] 感知编码器310适于对所述增益修正的输送信号 $z_i(k-2)$ 进行感知编码806,其中,感知编码的输送信号 $\tilde{z}_i(k-2)$ ,  $i=1, \dots, I$ 被获得。

[0081] 副信息源编码器320、330适于对副信息进行编码,副信息包括所述指数 $e_i(k-2)$ 和异常标志 $\beta_i(k-2)$ 、所述第一元组集合 $\mathcal{M}_{DIR}(k)$ 和第二元组集合 $\mathcal{M}_{VEC}(k)$ 、所述预测参数 $\xi(k-1)$ 和所述最终的分配矢量 $v_A(k-2)$ ,其中,编码的副信息 $\tilde{r}(k-2)$ 被获得。

[0082] 复用器340、350适于将感知编码的输送信号 $\tilde{z}_i(k-2)$ 和编码的副信息 $\tilde{r}(k-2)$ 复用为复用的数据流 $\tilde{\mathbf{B}}(k-2)$ ,其中,在分解中获得的环境HOA分量 $\tilde{\mathbf{C}}_{AMB}(k-1)$ 包括输入HOA表示 $c_n(k-1)$ 的在 $0_{MIN}$ 个最低位置(具有最低索引的那些位置)的第一HOA系数序列、以及在其余的较高位置上的第二HOA系数序列 $c_{AMB,n}(k-1)$ 。如下面关于式(4)-(6)说明的,第二HOA系数序列是输入HOA表示和主导声音信号的HOA表示之间的残差的HOA表示的一部分。此外,前 $0_{MIN}$ 个指数 $e_i(k-2)$ ,  $i=1, \dots, 0_{MIN}$ 和异常标志 $\beta_i(k-2)$ ,  $i=1, \dots, 0_{MIN}$ 在基本层副信息源编码器320中被编码,其中,编码的基本层副信息 $\tilde{\mathbf{r}}_{BASE}(k-2)$ 被获得,并且其中, $0_{MIN} = (N_{MIN}+1)^2$ ,  $0 = (N+1)^2$ ,  $N_{MIN} \leq N$ 并且 $0_{MIN} \leq I$ ,  $N_{MIN}$ 是预定义整数值。前 $0_{MIN}$ 个感知编码的输送信号 $\tilde{z}_i(k-2)$ ,  $i=1, \dots, 0_{MIN}$ 和编码的基本层副信息 $\tilde{\mathbf{r}}_{BASE}(k-2)$ 在基本层位流复用器340(其是所述复用器之一)中被复用,其中,基本层位流 $\tilde{\mathbf{B}}_{BASE}(k-2)$ 被获得。基本层副信息源编码器320是副信息源编码器之一,或者它在副信息源编码器模块内。其余的 $I-0_{MIN}$ 个指数 $e_i(k-2)$ ,  $i=0_{MIN}+1, \dots, I$ 和异常标志 $\beta_i(k-2)$ ,  $i=0_{MIN}+1, \dots, I$ 、所述第一元组集合 $\mathcal{M}_{DIR}(k-1)$ 和第二元组集合 $\mathcal{M}_{VEC}(k-1)$ 、所述预测参数 $\xi(k-1)$ 以及所述最终的分配矢量 $v_A(k-2)$ 在增强层副信息编码器330中被编码,其中,编码的增强层副信息 $\tilde{\mathbf{r}}_{ENH}(k-2)$ 被获得。增强层副信息源编码器330是副信息源编码器之一,或者在副信息源编码器模块内。

[0083] 其余的 $I-0_{MIN}$ 个感知编码的输送信号 $\tilde{z}_i(k-2)$ ,  $i=0_{MIN}+1, \dots, I$ 和编码的增强层副信息 $\tilde{\mathbf{r}}_{ENH}(k-2)$ 在增强层位流复用器350(其也是所述复用器之一)中被复用,其中,增强层位流 $\tilde{\mathbf{B}}_{ENH}(k-2)$ 被获得。此外,模式指示 $LMF_E$ 被添加在复用器或指示插入模块中。模式指示 $LMF_E$ 以信号告知分层模式的使用,分层模式用于压缩的信号的正确解压缩。

[0084] 在一个实施例中,用于编码的装置进一步包括适于选择模式的模式选择器,模式由模式指示LMF<sub>E</sub>指示,是分层模式和非分层模式之一。在非分层模式下,环境HOA分量 $\tilde{\mathbf{C}}_{\text{AMB}}(k-1)$ 仅包括表示输入HOA表示和主导声音信号的HOA表示之间的残差的HOA系数序列(即,不包括输入HOA表示的系数序列)。

[0085] 下面描述提出的HOA解压缩的修改。

[0086] 在分层模式下,在HOA解压缩处通过适当地修正HOA复合来考虑HOA压缩中的环境HOA分量 $\mathbf{C}_{\text{AMB}}(k-1)$ 的修正。

[0087] 在HOA解压缩器中,根据图5执行基本层位流和增强层位流的解复用和解码。基本层位流 $\tilde{\mathbf{B}}_{\text{BASE}}(k)$ 被解复用为基本层副信息的编码表示和感知编码的信号。随后,基本层副信息的编码表示和感知编码的信号被解码以一方面提供指数 $e_i(k)$ 和异常标志、另一方面提供感知解码的信号。类似地,增强层位流被解复用并且被解码以提供感知解码的信号和其余的副信息(参见图5)。就该分层模式而言,空间HOA解码部分也必须被修正以考虑空间HOA编码中的环境HOA分量 $\mathbf{C}_{\text{AMB}}(k-1)$ 的修正。修正在HOA复合中实现。

[0088] 具体地说,重构的HOA表示

$$[0089] \quad \tilde{\mathbf{C}}(k-1) = \tilde{\mathbf{C}}_{\text{PS}}(k-1) + \tilde{\mathbf{C}}_{\text{AMB}}(k-1) \quad (4)$$

[0090] 被其修正版本取代:

$$[0091] \quad \tilde{\mathbf{C}}(k-1) = \begin{bmatrix} \tilde{\mathbf{c}}_1(k-1) \\ \tilde{\mathbf{c}}_2(k-1) \\ \vdots \\ \tilde{\mathbf{c}}_O(k-1) \end{bmatrix} \quad (5)$$

[0092] 修正版本的元素由下式给出:

$$[0093] \quad \tilde{\mathbf{c}}_n(k-1) = \begin{cases} \hat{\mathbf{c}}_{\text{AMB},n}(k-1) & \text{对于 } 1 \leq n \leq O_{\text{MIN}} \\ \hat{\mathbf{c}}_n(k-1) & \text{对于 } O_{\text{MIN}} + 1 \leq n \leq O \end{cases} \quad (6)$$

[0094] 这意味着主导声音HOA分量没有被添加到前 $O_{\text{MIN}}$ 个系数序列的环境HOA分量,因为它已经被包括在其中。HOA空间解码器的所有的其他的处理模块保持不变。

[0095] 在下面,简要地考虑仅存在低质量基本层位流 $\tilde{\mathbf{B}}_{\text{BASE}}(k)$ 时的HOA解压缩。

[0096] 该位流首先被解复用并且被解码以提供重构的信号 $\hat{\mathbf{z}}_i(k)$ 和对应的增益控制副信息,该副信息由指数 $e_i(k)$ 和异常标志 $\beta_i(k)$ ,  $i=1, \dots, O_{\text{MIN}}$ 构成。注意到,在不存在增强层时,感知编码的信号 $\hat{\mathbf{z}}_i(k-2)$ ,  $i=O_{\text{MIN}}+1, \dots, O$ 是不可用的。解决这种情况的可能的方式是将信号 $\hat{\mathbf{z}}_i(k)$ ,  $i=O_{\text{MIN}}+1, \dots, O$ 设置为零,这自动地使重构的主导声音分量 $\mathbf{C}_{\text{PS}}(k-1)$ 为零。

[0097] 在下一个步骤中,在空间HOA解码器中,前 $O_{\text{MIN}}$ 个逆增益控制处理模块提供增益校正的信号帧 $\hat{\mathbf{y}}_i(k)$ ,  $i=1, \dots, O_{\text{MIN}}$ ,这些信号帧被用于通过通道再分配来构造环境HOA分量的中间表示的帧 $\mathbf{C}_{\text{I,AMB}}(k)$ 。注意到,环境HOA分量的在第 $k$ 帧中起作用的系数序列的索引集合

$\mathcal{I}_{\text{AMB,ACT}}(k)$  仅包含索引  $1, 2, \dots, 0_{\text{MIN}}$ 。在环境合成中, 前  $0_{\text{MIN}}$  个系数序列的空间变换被恢复以提供环境 HOA 分量帧  $\mathcal{C}_{\text{AMB}}(k-1)$ 。最后, 根据式 (6) 计算重构的 HOA 表示。

[0098] 图5和图6示出了根据本发明的一个实施例的HOA解压缩器的架构的结构。所述装置包括如图5所示的感知解码和源解码部分、如图6所示的空间HOA解码部分、以及适于检测分层模式指示  $\text{LMF}_D$  的模式检测器, 分层模式指示  $\text{LMF}_D$  指示压缩的HOA信号包括压缩的基本层位流  $\tilde{\mathcal{B}}_{\text{BASE}}(k)$  和压缩的增强层位流。

[0099] 图5示出了根据本发明的一个实施例的HOA解压缩器的感知解码和源解码部分的架构的结构。感知解码和源解码部分包括第一解复用器510、第二解复用器520、基本层感知解码器540和增强层感知解码器550、基本层副信息源解码器530和增强层副信息源解码器560。

[0100] 第一解复用器510适于对压缩的基本层位流  $\tilde{\mathcal{B}}_{\text{BASE}}(k)$  进行解复用, 其中, 第一感知编码的输送信号  $\tilde{\mathcal{Z}}_i(k), i = 1, \dots, 0_{\text{MIN}}$  和第一编码的副信息  $\tilde{\mathcal{I}}_{\text{BASE}}(k)$  被获得。第二解复用器520适于对压缩的增强层位流  $\tilde{\mathcal{B}}_{\text{ENH}}(k)$  进行解复用, 其中, 第二感知编码的输送信号  $\tilde{\mathcal{Z}}_i(k), i = 0_{\text{MIN}} + 1, \dots, I$  和第二编码的副信息  $\tilde{\mathcal{I}}_{\text{ENH}}(k)$  被获得。

[0101] 基本层感知解码器540和增强层感知解码器550适于对感知编码的输送信号  $\tilde{\mathcal{Z}}_i(k), i = 1, \dots, I$  进行感知解码904, 其中, 感知解码的输送信号  $\hat{\mathcal{Z}}_i(k)$  被获得, 并且其中, 在基本层感知解码器540中, 基本层的所述第一感知编码的输送信号  $\tilde{\mathcal{Z}}_i(k), i = 1, \dots, 0_{\text{MIN}}$  被解码, 并且第一感知解码的输送信号  $\hat{\mathcal{Z}}_i(k), i = 1, \dots, 0_{\text{MIN}}$  被获得。在增强层感知解码器550中, 增强层的所述第二感知编码的输送信号  $\tilde{\mathcal{Z}}_i(k), i = 0_{\text{MIN}} + 1, \dots, I$  被解码, 并且第二感知解码的输送信号  $\hat{\mathcal{Z}}_i(k), i = 0_{\text{MIN}} + 1, \dots, I$  被获得。

[0102] 基本层副信息源解码器530适于对第一编码的副信息  $\tilde{\mathcal{I}}_{\text{BASE}}(k)$  进行解码905, 其中, 第一指数  $e_i(k), i = 1, \dots, 0_{\text{MIN}}$  和第一异常标志  $\beta_i(k), i = 1, \dots, 0_{\text{MIN}}$  被获得。

[0103] 增强层副信息源解码器560适于对第二编码的副信息  $\tilde{\mathcal{I}}_{\text{ENH}}(k)$  进行解码906, 其中, 第二指数  $e_i(k), i = 0_{\text{MIN}} + 1, \dots, I$  和第二异常标志  $\beta_i(k), i = 0_{\text{MIN}} + 1, \dots, I$  被获得, 并且其中, 进一步的数据被获得。进一步的数据包括关于方向信号的第一元组集合  $\mathcal{M}_{\text{DIR}}(k+1)$  和关于基于矢量的信号的第二元组集合  $\mathcal{M}_{\text{VEC}}(k+1)$ 。第一元组集合  $\mathcal{M}_{\text{DIR}}(k+1)$  的每个元组包括方向信号的索引和相应的量化方向, 第二元组集合  $\mathcal{M}_{\text{VEC}}(k+1)$  的每个元组包括基于矢量的信号的索引和限定基于矢量的信号的方向分布的矢量。此外, 预测参数  $\xi(k+1)$  和环境分配矢量  $\mathcal{V}_{\text{AMB,ASSIGN}}(k)$  被获得, 其中, 环境分配矢量  $\mathcal{V}_{\text{AMB,ASSIGN}}(k)$  包括对于每个传输通道指示它是否包含环境 HOA 分量的系数序列以及它包含环境 HOA 分量的哪个系数序列的分量。

[0104] 图6示出了根据本发明的一个实施例的HOA解压器的空间HOA 解码部分的架构的结构。空间HOA解码部分包括多个逆增益控制单元604、通道再分配模块605、主导声音合成模块606、环境合成模块 607、HOA复合模块608。

[0105] 多个逆增益控制单元604适于执行逆增益控制, 其中, 所述第一感知解码的输送信



号  $\hat{\mathbf{z}}_i(k)$ ,  $i = 1, \dots, O_{\text{MIN}}$  根据第一指数  $e_i(k)$ ,  $i = 1, \dots, O_{\text{MIN}}$  和第一异常标志  $\beta_i(k)$ ,  $i = 1, \dots, O_{\text{MIN}}$  被变换为第一增益校正的信号帧  $\hat{\mathbf{y}}_i(k)$ ,  $i = 1, \dots, O_{\text{MIN}}$ , 并且其中, 第二感知解码的输送信号  $\hat{\mathbf{z}}_i(k)$ ,  $i = O_{\text{MIN}} + 1, \dots, I$  根据第二指数  $e_i(k)$ ,  $i = O_{\text{MIN}} + 1, \dots, I$  和第二异常标志  $\beta_i(k)$ ,  $i = O_{\text{MIN}} + 1, \dots, I$  被变换为第二增益校正的信号帧  $\hat{\mathbf{y}}_i(k)$ ,  $i = O_{\text{MIN}} + 1, \dots, I$ 。

[0106] 通道再分配模块605适于将第一和第二增益校正的信号帧  $\hat{\mathbf{y}}_i(k)$ ,  $i = 1, \dots, I$  重新分布911给I个通道, 其中, 主导声音信号  $\hat{\mathbf{X}}_{PS}(k)$  的帧被重构, 主导声音信号包括方向信号和基于矢量的信号, 并且其中, 修正的环境HOA分量  $\tilde{\mathbf{C}}_{I,AMB}(k)$  被获得, 并且其中, 分配是根据所述环境分配矢量  $\mathbf{V}_{AMB, ASSIGN}(k)$  以及所述第一和第二元组集合  $\mathcal{M}_{\text{DIR}}(k+1)$ 、 $\mathcal{M}_{\text{VEC}}(k+1)$  进行的。

[0107] 此外, 通道再分配模块605适于产生修正的环境HOA分量的在第k帧中起作用的系数序列的第一索引集合  $\mathcal{I}_{AMB, ACT}(k)$ 、以及修正的环境 HOA分量的在第(k-1)帧中必须被启用、被禁用以及保持起作用的系数序列的第二索引集合  $\mathcal{I}_E(k-1)$ 、 $\mathcal{I}_D(k-1)$ 、 $\mathcal{I}_U(k-1)$ 。

[0108] 主导声音合成模块606适于从所述主导声音信号  $\hat{\mathbf{X}}_{PS}(k)$  合成912主导HOA声音分量  $\tilde{\mathbf{C}}_{PS}(k-1)$  的HOA表示, 其中, 第一元组集合  $\mathcal{M}_{\text{DIR}}(k+1)$ 、第二元组集合  $\mathcal{M}_{\text{VEC}}(k+1)$ 、预测参数  $\xi(k+1)$  以及第二索引集合  $\mathcal{I}_E(k-1)$ 、 $\mathcal{I}_D(k-1)$ 、 $\mathcal{I}_U(k-1)$  被使用。

[0109] 环境合成模块607适于从修正的环境HOA分量  $\tilde{\mathbf{C}}_{I,AMB}(k)$  合成913 环境HOA分量  $\tilde{\mathbf{C}}_{AMB}(k-1)$ , 其中, 对前  $O_{\text{MIN}}$  个通道的逆空间变换被执行, 并且其中, 第一索引集合  $\mathcal{I}_{AMB, ACT}(k)$  被使用, 第一索引集合是环境 HOA分量的在第k帧中起作用的系数序列的索引。

[0110] 如果分层模式指示  $\text{LMF}_D$  指示具有至少两个层的分层模式, 则环境HOA分量在其  $O_{\text{MIN}}$  个最低位置 (即, 具有最低索引的那些位置) 包括解压缩的HOA分量  $\tilde{\mathbf{C}}(k-1)$  的HOA系数序列、以及在其余的较高位置包括作为残差的HOA表示的一部分的系数序列。该残差是解压缩的HOA信号  $\tilde{\mathbf{C}}(k-1)$  和主导HOA声音分量  $\tilde{\mathbf{C}}_{PS}(k-1)$  的HOA表示之间的残差。

[0111] 另一方面, 如果分层模式指示  $\text{LMF}_D$  指示单层模式, 则不包括解压缩的HOA信号  $\tilde{\mathbf{C}}(k-1)$  的HOA系数序列, 并且环境HOA分量是解压缩的HOA信号  $\tilde{\mathbf{C}}(k-1)$  和主导声音分量  $\tilde{\mathbf{C}}_{PS}(k-1)$  的HOA表示之间的残差。

[0112] HOA复合模块608适于将主导声音分量的HOA表示与环境HOA分量  $\tilde{\mathbf{C}}_{PS}(k-1)\tilde{\mathbf{C}}_{AMB}(k-1)$  相加, 其中, 主导声音信号的HOA表示的系数和环境HOA分量的对应系数相加, 并且其中, 解压缩的HOA信号  $\tilde{\mathbf{C}}'(k-1)$  被获得, 并且其中,

[0113] 如果分层模式指示  $\text{LMF}_D$  指示具有至少两个层的分层模式, 则只有最高的  $I - O_{\text{MIN}}$  个系数通道通过主导HOA声音分量  $\tilde{\mathbf{C}}_{PS}(k-1)$  和环境 HOA分量  $\tilde{\mathbf{C}}_{AMB}(k-1)$  的相加获得, 解压缩的HOA信号  $\tilde{\mathbf{C}}'(k-1)$  的最低的  $O_{\text{MIN}}$  个系数通道是从环境HOA分量  $\tilde{\mathbf{C}}_{AMB}(k-1)$  复制的。另

一方面,如果分层模式指示LMF<sub>D</sub>指示单层模式,则解压缩的HOA信号 $\tilde{\mathbf{c}}'(k-1)$ 的所有的系数通道通过主导HOA声音分量 $\tilde{\mathbf{c}}_{PS}(k-1)$ 和环境HOA分量 $\tilde{\mathbf{c}}_{AMB}(k-1)$ 的相加获得。

[0114] 图7示出了从环境HOA信号到修正的环境HOA信号的帧变换。

[0115] 图8示出了用于压缩HOA信号的方法的流程图。

[0116] 用于压缩高阶高保真立体声 (HOA) 信号 (该信号是具有HOA 系数序列的输入时间帧C(k)的、阶次为N的输入HOA表示)的方法 800包括输入时间帧的空间HOA编码以及随后的感知编码和源编码。

[0117] 空间HOA编码包括以下步骤:

[0118] 在方向和矢量估计模块301中执行HOA信号的方向和矢量估计处理801,其中,包括关于方向信号的第一元组集合 $\mathcal{M}_{DIR}(k)$ 和关于基于矢量的信号的第二元组集合 $\mathcal{M}_{VEC}(k)$ 的数据被获得,第一元组集合 $\mathcal{M}_{DIR}(k)$ 中的每个包括方向信号的索引和相应的量化方向,第二元组集合中 $\mathcal{M}_{VEC}(k)$ 的每个包括基于矢量的信号的索引和限定信号的方向分布的矢量;

[0119] 在HOA分解模块303中将HOA系数序列的每个输入时间帧分解 802为多个主导声音信号 $X_{PS}(k-1)$ 的帧和环境HOA分量 $C_{AMB}(k-1)$ 的帧,其中,主导声音信号 $X_{PS}(k-1)$ 包括所述方向声音信号和所述基于矢量的声音信号,并且其中,环境HOA分量 $\tilde{\mathbf{c}}_{AMB}(k-1)$ 包括表示输入HOA表示和主导声音信号的HOA表示之间的残差的HOA系数序列,并且其中,分解702进一步提供预测参数 $\xi(k-1)$ 和目标分配矢量 $\mathbf{v}_{A,T}(k-1)$ ,预测参数 $\xi(k-1)$ 描述了如何从主导声音信号 $X_{PS}(k-1)$ 内的方向信号预测HOA信号表示的各部分以便丰富主导声音HOA分量,目标分配矢量 $\mathbf{v}_{A,T}(k-1)$ 包含关于如何将主导声音信号分配给给定数量I个通道的信息;

[0120] 在环境分量修正模块304中根据目标分配矢量 $\mathbf{v}_{A,T}(k-1)$ 提供的信息修正802环境HOA分量 $C_{AMB}(k-1)$ ,其中,确定环境HOA分量 $C_{AMB}(k-1)$ 的哪些系数序列将在给定数量I个通道中传输,这取决于有多少个通道被主导声音信号占据,并且其中,修正的环境HOA分量 $C_{M,A}(k-2)$ 和时间预测的修正的环境HOA分量 $C_{P,M,A}(k-1)$ 被获得,并且其中,最终的分配矢量 $\mathbf{v}_A(k-2)$ 从目标分配矢量 $\mathbf{v}_{A,T}(k-1)$ 中的信息获得;

[0121] 在通道分配模块105中使用最终的分配矢量 $\mathbf{v}_A(k-2)$ 提供的信息将从分解获得的主导声音信号 $X_{PS}(k-1)$ 以及修正的环境HOA分量 $C_{M,A}(k-2)$ 和时间预测的修正的环境HOA分量 $C_{P,M,A}(k-1)$ 的确定的系数序列分配给804给定数量I个通道,其中,输送信号 $y_i(k-2)$ ,  $i=1, \dots, I$ 和预测的输送信号 $y_{P,i}(k-2)$ ,  $i=1, \dots, I$ 被获得;

[0122] 以及在多个增益控制模块306中对输送信号 $y_i(k-2)$ 和预测的输送信号 $y_{P,i}(k-2)$ 执行增益控制805,其中,增益修正的输送信号 $z_i(k-2)$ 、指数 $e_i(k-2)$ 和异常标志 $\beta_i(k-2)$ 被获得。

[0123] 感知编码和源编码包括以下步骤:

[0124] 在感知编码器310中对所述增益修正的输送信号 $z_i(k-2)$ 进行感知编码806,其中,感知编码的输送信号 $\tilde{\mathbf{z}}_i(k-2)$ ,  $i=1, \dots, I$ 被获得;

[0125] 在一个或多个副信息源编码器320、330中对副信息进行编码807,副信息包括所述指数 $e_i(k-2)$ 和异常标志 $\beta_i(k-2)$ 、所述第一元组集合 $\mathcal{M}_{DIR}(k)$ 和第二元组集合 $\mathcal{M}_{VEC}(k)$ 、所述



预测参数 $\xi(k-1)$ 和所述最终的分配矢量 $v_A(k-2)$ ,其中,编码的副信息 $\tilde{r}(k-2)$ 被获得;以及

[0126] 对感知编码的输送信号 $\tilde{z}_i(k-2)$ 和编码的副信息 $\tilde{r}(k-2)$ 进行复用808,其中,复用的数据流 $\tilde{B}(k-2)$ 被获得。

[0127] 在分解步骤802中获得的环境HOA分量 $\tilde{c}_{AMB}(k-1)$ 包括输入HOA表示 $c_n(k-1)$ 的在 $0_{MIN}$ 个最低位置(即,具有最低索引的那些位置)的第一HOA系数序列以及在其余的较高位置的第二HOA系数序列 $c_{AMB,n}(k-1)$ 。第二系数序列是输入HOA表示和主导声音信号的HOA表示之间的残差的HOA表示的一部分。

[0128] 前 $0_{MIN}$ 个指数 $e_i(k-2)$ ,  $i=1, \dots, 0_{MIN}$ 和异常标志 $\beta_i(k-2)$ ,  $i=1, \dots, 0_{MIN}$ 在基本层副信息源编码器320中被编码,其中,编码的基本层副信息 $\tilde{r}_{BASE}(k-2)$ 被获得,并且其中,  $0_{MIN} = (N_{MIN}+1)^2$ ,  $0 = (N+1)^2$ ,  $N_{MIN} \leq N$ 并且 $0_{MIN} \leq I$ ,  $N_{MIN}$ 是预定义整数值。

[0129] 前 $0_{MIN}$ 个感知编码的输送信号 $\tilde{z}_i(k-2)$ ,  $i=1, \dots, 0_{MIN}$ 和编码的基本层副信息 $\tilde{r}_{BASE}(k-2)$ 在基本层位流复用器340中被复用809,其中,基本层位流 $\tilde{B}_{BASE}(k-2)$ 被获得。

[0130] 其余的 $I-0_{MIN}$ 个指数 $e_i(k-2)$ ,  $i=0_{MIN}+1, \dots, I$ 和异常标志 $\beta_i(k-2)$ ,  $i=0_{MIN}+1, \dots, I$ 、所述第一元组集合 $\mathcal{M}_{DIR}(k-1)$ 和第二元组集合 $\mathcal{M}_{VEC}(k-1)$ 、所述预测参数 $\xi(k-1)$ 以及所述最终的分配矢量 $v_A(k-2)$ (在图中也被示为 $v_{AMB, ASSIGN}(k)$ ) 在增强层副信息编码器330中被编码,其中,编码的增强层副信息 $\tilde{r}_{ENH}(k-2)$ 被获得。

[0131] 其余的 $I-0_{MIN}$ 个感知编码的输送信号 $\tilde{z}_i(k-2)$ ,  $i=0_{MIN}+1, \dots, I$ 和编码的增强层副信息 $\tilde{r}_{ENH}(k-2)$ 在增强层位流复用器350中被复用810,其中,增强层位流 $\tilde{B}_{ENH}(k-2)$ 被获得。

[0132] 如上所述,信号告知分层模式的使用的模式指示被添加811。该模式指示由指示插入模块或复用器添加。

[0133] 在一个实施例中,所述方法进一步包括将基本层位流 $\tilde{B}_{BASE}(k-2)$ 、增强层位流 $\tilde{B}_{ENH}(k-2)$ 和模式指示复用到单个位流中的最终步骤。

[0134] 在一个实施例中,所述主导方向估计取决于能量占主导的HOA分量的方向功率分布。

[0135] 在一个实施例中,在修正环境HOA分量中,如果被选HOA系数序列的HOA序列索引在连续的帧之间变化,则系数序列的淡入和淡出被执行。

[0136] 在一个实施例中,在修正环境HOA分量中,环境HOA分量 $c_{AMB}(k-1)$ 的部分去相关被执行。

[0137] 在一个实施例中,第一元组集合 $\mathcal{M}_{DIR}(k)$ 中所包括的量化方向是主导方向。

[0138] 图9示出了用于解压缩压缩的HOA信号的方法的流程图。在本发明的这个实施例中,用于解压缩压缩的HOA信号的方法900包括获得HOA系数序列的输出时间帧 $\tilde{c}(k-1)$ 的感知解码和源解码以及随后的空间HOA解码,并且所述方法包括检测901分层模式指示 $LMF_D$ 的

步骤,分层模式指示LMF<sub>D</sub>指示压缩的高阶高保真立体声(HOA)信号包括压缩的基本层位流 $\tilde{B}_{BASE}(k)$ 和压缩的增强层位流 $\tilde{B}_{ENH}(k)$ 。

[0139] 感知解码和源解码包括以下步骤:

[0140] 对压缩的基本层位流 $\tilde{B}_{BASE}(k)$ 进行解复用902,其中,第一感知编码的输送信号 $\tilde{z}_i(k), i = 1, \dots, O_{MIN}$ 和第一编码的副信息 $\tilde{r}_{BASE}(k)$ 被获得;

[0141] 对压缩的增强层位流 $\tilde{B}_{ENH}(k)$ 进行解复用903,其中,第二感知编码的输送信号 $\tilde{z}_i(k), i = O_{MIN} + 1, \dots, I$ 和第二编码的副信息 $\tilde{r}_{ENH}(k)$ 被获得;

[0142] 对感知编码的输送信号 $\tilde{z}_i(k), i = 1, \dots, I$ 进行感知解码904,其中,感知解码的输送信号 $\hat{z}_i(k)$ 被获得,并且其中,在基本层感知解码器540中,基本层的所述第一感知编码的输送信号 $\tilde{z}_i(k), i = 1, \dots, O_{MIN}$ 被解码,并且第一感知解码的输送信号 $\hat{z}_i(k), i = 1, \dots, O_{MIN}$ 被获得,并且其中,在增强层感知解码器550中,增强层的所述第二感知编码的输送信号 $\tilde{z}_i(k), i = O_{MIN} + 1, \dots, I$ 被解码,并且第二感知解码的输送信号 $\hat{z}_i(k), i = O_{MIN} + 1, \dots, I$ 被获得;

[0143] 在基本层副信息源解码器530中对第一编码的副信息 $\tilde{r}_{BASE}(k)$ 进行解码905,其中,第一指数 $e_i(k), i = 1, \dots, O_{MIN}$ 和第一异常标志 $\beta_i(k), i = 1, \dots, O_{MIN}$ 被获得;以及

[0144] 在增强层副信息源解码器560中对第二编码的副信息 $\tilde{r}_{ENH}(k)$ 进行解码906,其中,第二指数 $e_i(k), i = O_{MIN} + 1, \dots, I$ 和第二异常标志 $\beta_i(k), i = O_{MIN} + 1, \dots, I$ 被获得,并且其中,进一步的数据被获得,进一步的数据包括关于方向信号的第一元组集合 $\mathcal{M}_{DIR}(k+1)$ 和关于基于矢量的信号的第二元组集合 $\mathcal{M}_{VEC}(k+1)$ ,第一元组集合 $\mathcal{M}_{DIR}(k+1)$ 的每个元组包括方向信号的索引和相应的量化方向,第二元组集合 $\mathcal{M}_{VEC}(k+1)$ 的每个元组包括基于矢量的信号的索引和限定基于矢量的信号的方向分布的矢量,并且进一步其中,预测参数 $\xi(k+1)$ 和环境分配矢量 $\mathbf{v}_{AMB, ASSIGN}(k)$ 被获得。环境分配矢量 $\mathbf{v}_{AMB, ASSIGN}(k)$ 包括对于每个传输通道指示它是否包含环境HOA分量的系数序列以及它包含环境HOA分量的哪个系数序列的分量。

[0145] 空间HOA解码包括以下步骤:

[0146] 执行910逆增益控制,其中,所述第一感知解码的输送信号 $\hat{z}_i(k), i = 1, \dots, O_{MIN}$ 根据所述第一指数 $e_i(k), i = 1, \dots, O_{MIN}$ 和所述第一异常标志 $\beta_i(k), i = 1, \dots, O_{MIN}$ 被变换为第一增益校正的信号帧 $\hat{y}_i(k), i = 1, \dots, O_{MIN}$ ,并且其中,所述第二感知解码的输送信号 $\hat{z}_i(k), i = O_{MIN} + 1, \dots, I$ 根据所述第二指数 $e_i(k), i = O_{MIN} + 1, \dots, I$ 和所述第二异常标志 $\beta_i(k), i = O_{MIN} + 1, \dots, I$ 被变换为第二增益校正的信号帧 $\hat{y}_i(k), i = O_{MIN} + 1, \dots, I$ ;

[0147] 在通道再分配模块605中将第一和第二增益校正的信号帧 $\hat{y}_i(k), i = 1, \dots, I$ 重新分布911给I个通道,其中,主导声音信号 $\tilde{\mathbf{X}}_{PS}(k)$ 的帧被重构,主导声音信号包括方向信号和基于矢量的信号,并且其中,修正的环境HOA分量 $\tilde{\mathbf{C}}_{LAMB}(k)$ 被获得,并且其中,分配是根据所述环境分配矢量 $\mathbf{v}_{AMB, ASSIGN}(k)$ 以及所述第一和第二元组集合 $\mathcal{M}_{DIR}(k+1)$ 、 $\mathcal{M}_{VEC}(k+1)$ 进行

的；

[0148] 在通道再分配模块605中产生修正的环境HOA分量的在第k帧中起作用的系数序列的第一索引集合 $\mathcal{I}_{AMB,ACT}(k)$ 、以及修正的环境HOA分量的在第(k-1)帧中必须被启用、被禁用以及保持起作用的系数序列的第二索引集合 $\mathcal{I}_E(k-1)$ 、 $\mathcal{I}_D(k-1)$ 、 $\mathcal{I}_U(k-1)$ ；

[0149] 在主导声音合成模块606中，从所述主导声音信号 $\tilde{X}_{PS}(k)$ 合成912 主导HOA声音分量 $\tilde{C}_{PS}(k-1)$ 的HOA表示，其中，第一元组集合 $\mathcal{M}_{DIR}(k+1)$ 、第二元组集合 $\mathcal{M}_{VEC}(k+1)$ 、预测参数 $\xi(k+1)$ 以及第二索引集合 $\mathcal{I}_E(k-1)$ 、 $\mathcal{I}_D(k-1)$ 、 $\mathcal{I}_U(k-1)$ 被使用；

[0150] 在环境合成模块607中，从修正的环境HOA分量 $\tilde{C}_{I,AMB}(k)$ 合成913 环境HOA分量 $\tilde{C}_{AMB}(k-1)$ ，其中，对前 $O_{MIN}$ 个通道的逆空间变换被执行，并且其中，第一索引集合 $\mathcal{I}_{AMB,ACT}(k)$ 被使用，第一索引集合是环境 HOA分量的在第k帧中起作用的系数序列的索引，其中，环境HOA 分量具有至少两种不同配置之一，这取决于分层模式指示 $LMF_D$ ；以及

[0151] 在HOA复合模块608中将主导HOA声音分量 $\tilde{C}_{PS}(k-1)$ 的HOA 表示与环境HOA分量 $\tilde{C}_{AMB}(k-1)$ 相加914，其中，主导声音信号的HOA 表示的系数和环境HOA分量的对应系数被相加，并且其中，解压缩的HOA信号 $\tilde{C}'(k-1)$ 被获得，并且其中，以下条件适用：

[0152] 如果分层模式指示 $LMF_D$ 指示具有至少两个层的分层模式，则只有最高的 $I-O_{MIN}$ 个系数通道通过主导HOA声音分量 $\tilde{C}_{PS}(k-1)$ 和环境 HOA分量 $\tilde{C}_{AMB}(k-1)$ 的相加获得，解压缩的HOA信号 $\tilde{C}'(k-1)$ 的最低的  $O_{MIN}$ 个系数通道是从环境HOA分量 $\tilde{C}_{AMB}(k-1)$ 复制的。否则，如果分层模式指示 $LMF_D$ 指示单层模式，则解压缩的HOA信号 $\tilde{C}'(k-1)$ 的所有的系数通道通过主导HOA声音分量 $\tilde{C}_{PS}(k-1)$ 和环境HOA分量 $\tilde{C}_{AMB}(k-1)$ 的相加获得。

[0153] 环境HOA分量的取决于分层模式指示 $LMF_D$ 的配置如下：

[0154] 如果分层模式指示 $LMF_D$ 指示具有至少两个层的分层模式，则环境HOA分量在其 $O_{MIN}$ 个最低位置包括解压缩的HOA信号 $\tilde{C}(k-1)$ 的 HOA系数序列以及在其余的较高位置包括如下的系数序列，该系数序列是主导HOA声音分量 $\tilde{C}_{PS}(k-1)$ 的HOA表示和解压缩的HOA信号 $\tilde{C}(k-1)$ 之间的残差的HOA表示的一部分。

[0155] 另一方面，如果分层模式指示 $LMF_D$ 指示单层模式，则环境HOA 分量是主导声音分量 $\tilde{C}_{PS}(k-1)$ 的HOA表示和解压缩的HOA信号 $\tilde{C}(k-1)$ 之间的残差。

[0156] 在一个实施例中，压缩的HOA信号表示在复用的位流中，用于解压缩压缩的HOA信号的方法进一步包括对压缩的HOA信号表示进行解复用的初始步骤，其中，所述压缩的基本层位流 $\tilde{B}_{BASE}(k)$ 、所述压缩的增强层位流 $\tilde{B}_{ENH}(k)$ 以及所述分层模式指示 $LMF_D$ 被获得。

[0157] 图10示出了根据本发明的一个实施例的HOA解压器的空间 HOA解码部分的架构的各部分的细节。

[0158] 有利地，例如，如果没有EL被接收到，或者如果BL质量足够，则可以仅对BL进行解



码。对于这种情况下,EL的信号在解码器处可以被设置为零。于是,在通道再分配模块605中将第一和第二增益校正的信号帧  $\hat{y}_i(k)$ ,  $i=1, \dots, I$  重新分布911给I个通道是非常简单的,因为主导声音信号  $\hat{X}_{PS}(k)$  的帧是空的。修正的环境HOA分量的在第(k-1)帧中必须被启用、被禁用以及保持起作用的系数序列的第二索引集合  $\mathcal{I}_E(k-1)$ 、 $\mathcal{I}_D(k-1)$ 、 $\mathcal{I}_U(k-1)$  被设置为零。在主导声音合成模块606中从主导声音信号  $\hat{X}_{PS}(k)$  合成912主导HOA声音分量  $\tilde{C}_{PS}(k-1)$  的HOA表示因此可以被跳过,并且在环境合成模块607中从修正的环境HOA分量  $\tilde{C}_{LAMB}(k)$  合成913环境HOA分量  $\tilde{C}_{AMB}(k-1)$  对应于常规的HOA合成。

[0159] 用于HOA压缩的原始(即,整体式的、不可缩放的、非分层的)模式对于不需要低质量基本层位流的应用、例如对于基于文件的压缩可能仍然是有用的。对环境HOA分量  $C_{AMB}$  (其是原始HOA表示和方向HOA表示之间的差值)的空间变换的前  $O_{MIN}$  个系数序列、而不是原始HOA分量C的空间变换的系数序列进行感知编码的主要优点是,在前一种情况下,将被感知编码的所有信号之间的互相关性降低。信号  $z_i$ ,  $i=1, \dots, I$  之间的任何互相关性可以引起空间解码处理期间感知编码噪声的建设性叠加,同时无噪声的HOA系数序列在叠加时被取消。这种现象被称为感知噪声揭露。

[0160] 在分层模式下,在信号  $z_i$ ,  $i=1, \dots, O_{MIN}$  中的每个之间、还有在信号  $z_i$ ,  $i=1, \dots, O_{MIN}$  和  $z_i$ ,  $i=O_{MIN}+1, \dots, I$  之间存在高互相关性,因为环境HOA分量  $\tilde{C}_{AMB,n}$ ,  $n=1, \dots, O_{MIN}$  的修正的系数序列包括方向HOA分量的信号(参见式3)。相反,对于原始的非分层模式来说情况并非如此。因此可以推断由分层模式引入的传输鲁棒性可能是以压缩质量为代价的。然而,压缩质量的降低与传输鲁棒性的提高相比低。以上已经表明,提出的分层模式至少在上述情况下是有利的。

[0161] 虽然已经示出、描述并指出了应用于本发明的优选实施例的、本发明的基本的新颖的特征,但是将理解的是,在不脱离本发明的精神的情况下,本领域技术人员可以在所描述的装置和方法中、在所公开的设备的形式和细节上以及在它们的操作中做出各种省略、替换和改变。明确的意图是,以基本上相同的方式执行基本上相同的功能以实现相同结果的那些元件的所有组合都在本发明的范围内。从所描述的一个实施例对于另一个描述的实施例的元件的替换也被充分预期和设想。

[0162] 将理解的是,已经纯粹以举例的方式描述了本发明,并且在不脱离本发明的范围的情况下,可以进行细节的修改。

[0163] 在描述以及(在适当的情况下)权利要求和附图中公开的每个特征可以独立地提供或者按任何适当的组合提供。特征在适当的情况下可以用硬件、软件或两者的组合来实现。连接可以在适用的情况下可以实现为无线连接或有线(不一定是直接的或专用的)连接。

[0164] 在权利要求中出现的标号仅仅是例示说明,不应当限制权利要求的范围。

[0165] 引用的参考文献

[0166] [1]EP12306569.0

[0167] [2]EP12305537.8(作为EP2665208A发布)

[0168] [3]EP133005558.2

[0169] [4]ISO/IEC JTC1/SC29/WG11N14264.Working draft 1-H0A text of MPEG-H 3D audio,2014年1月



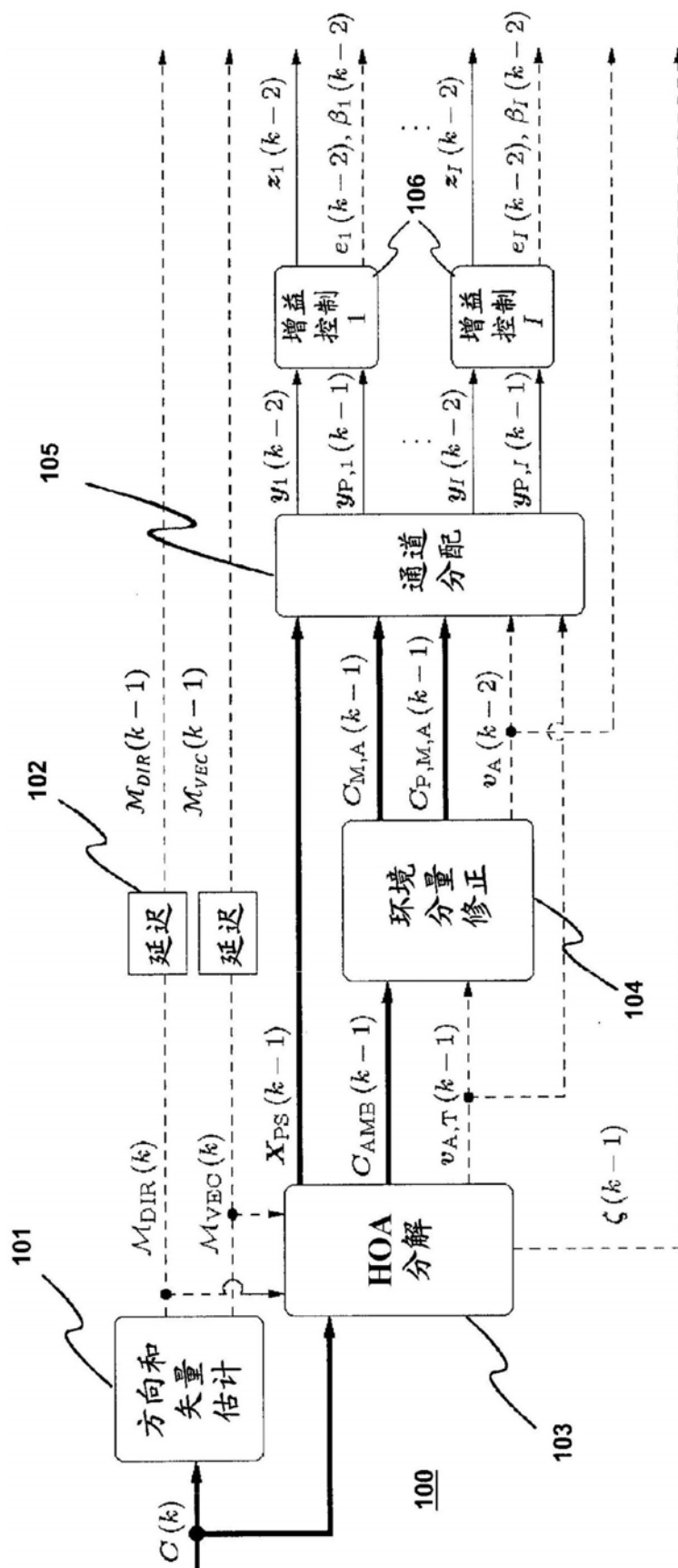


图1a)

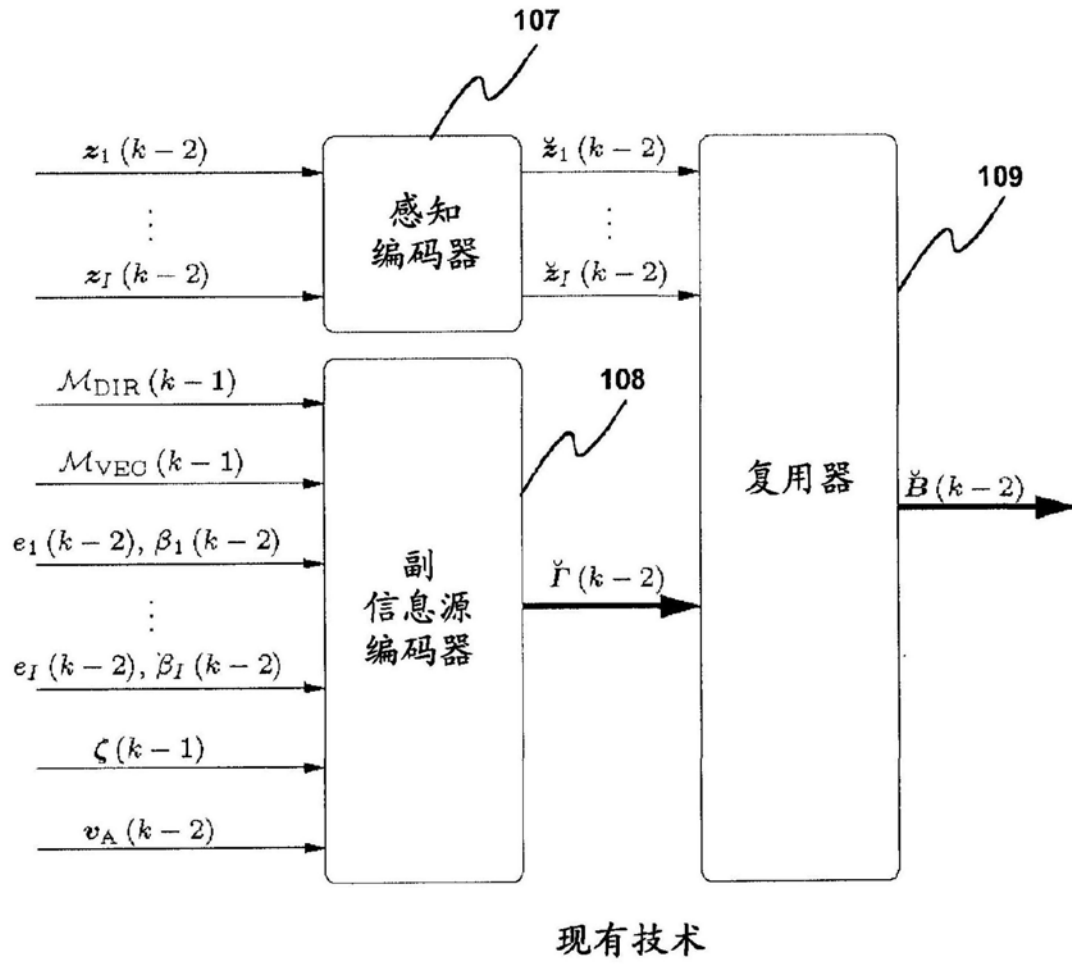
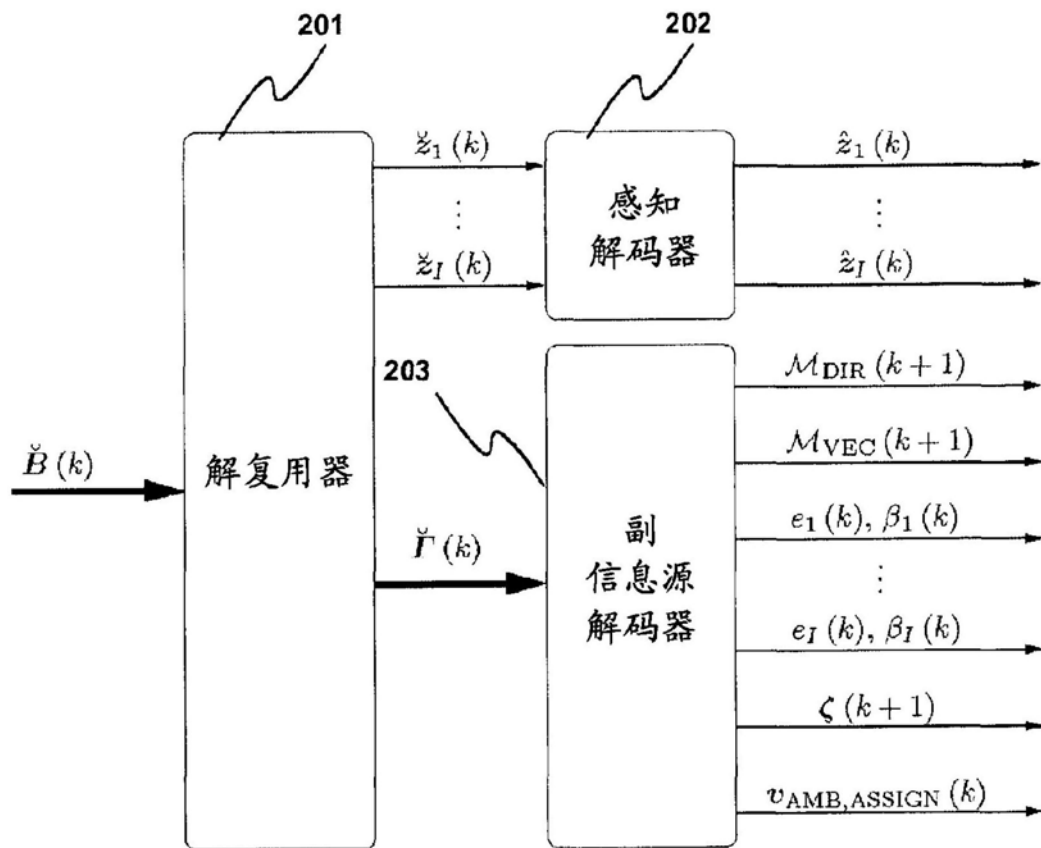
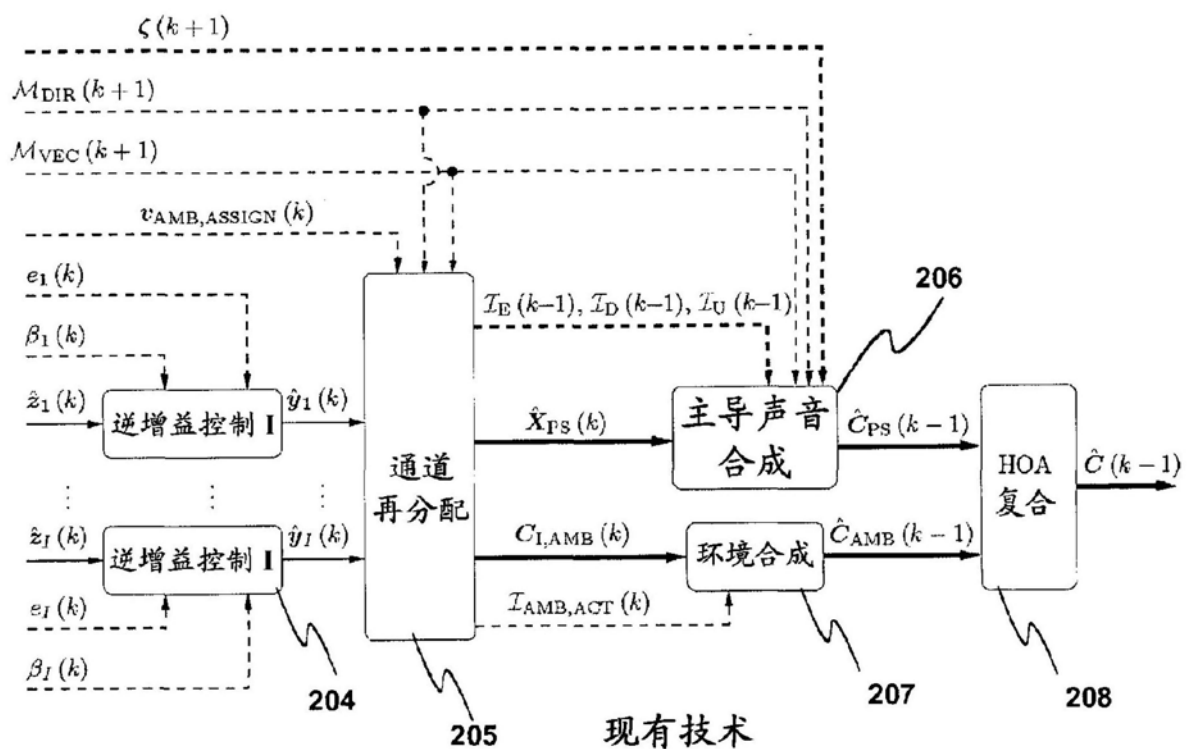


图1b)



a)



b)

图2

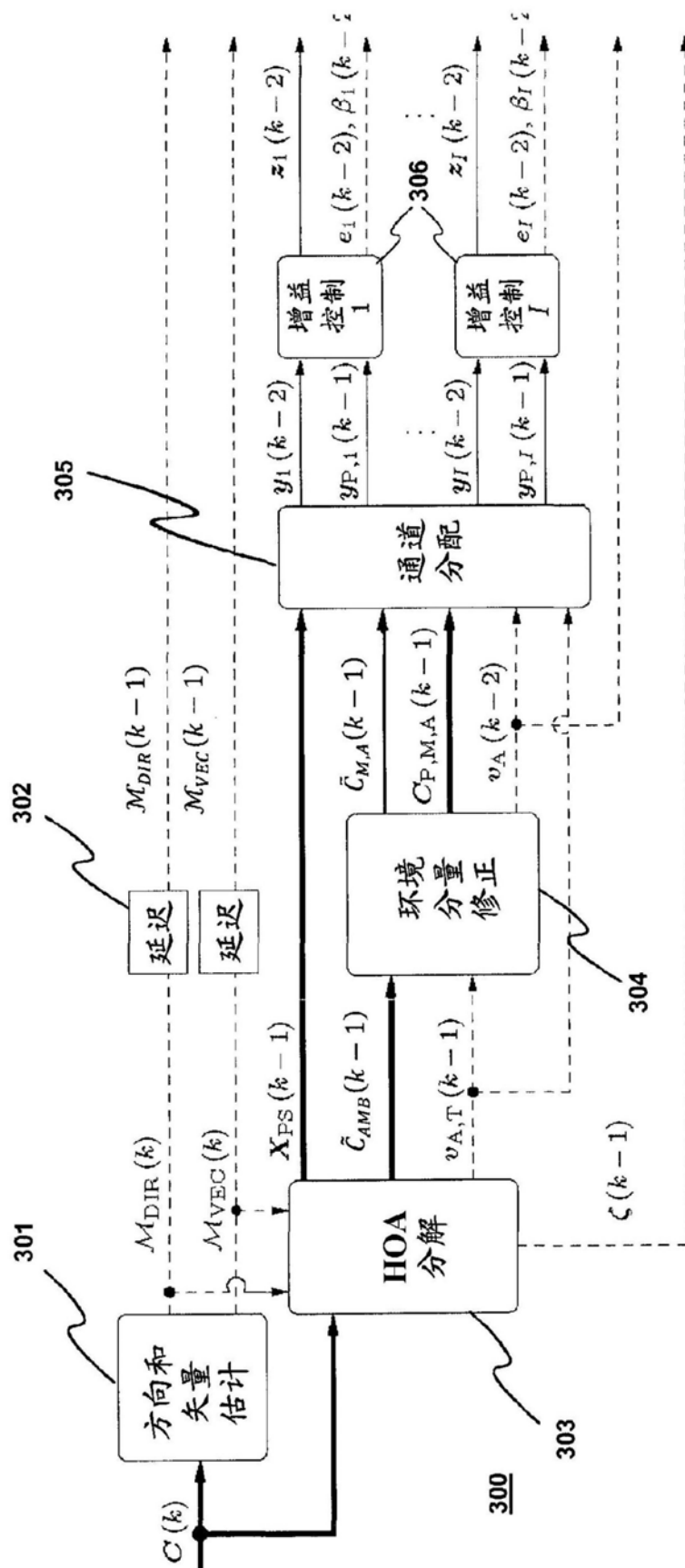


图3

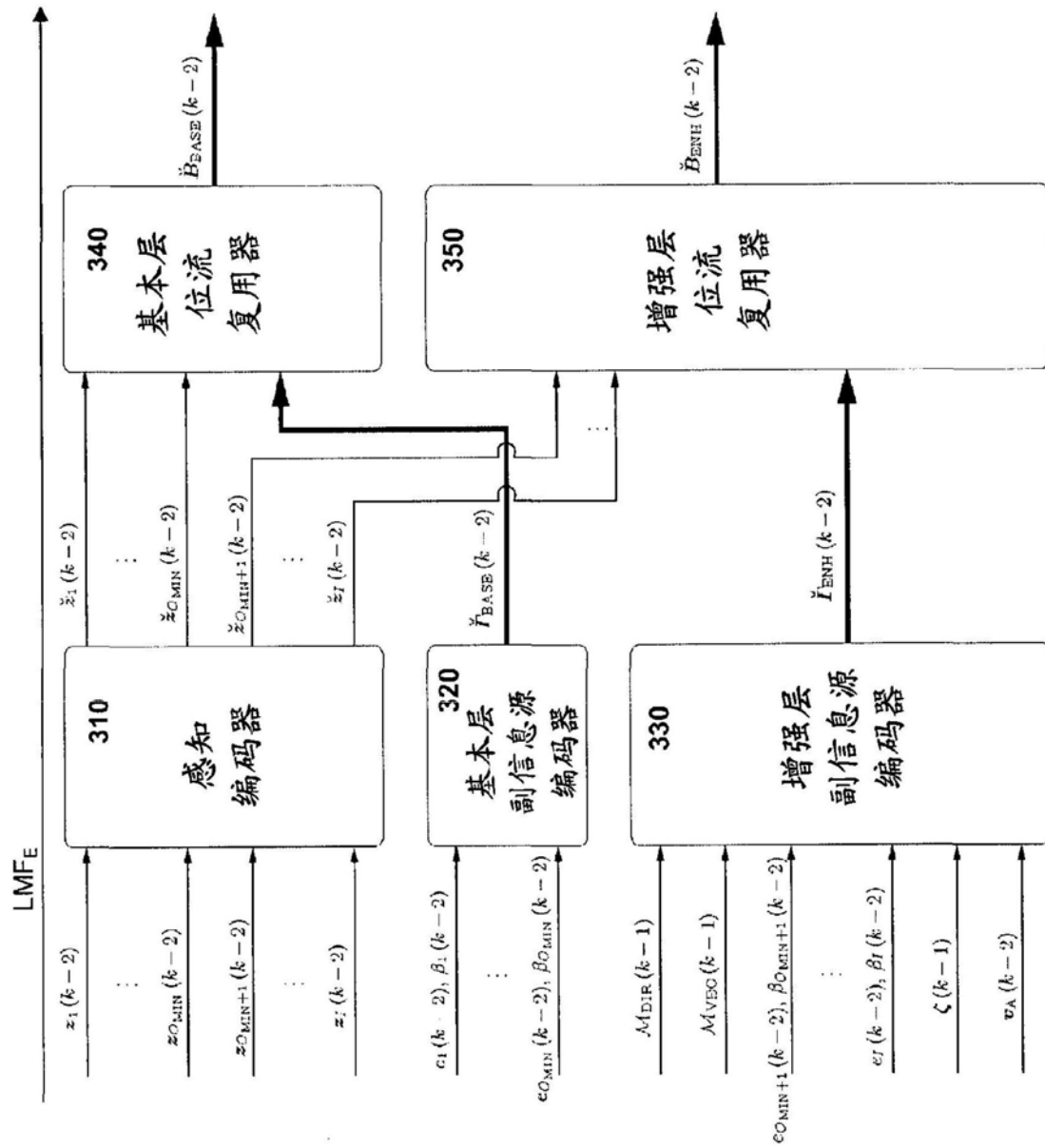


图4



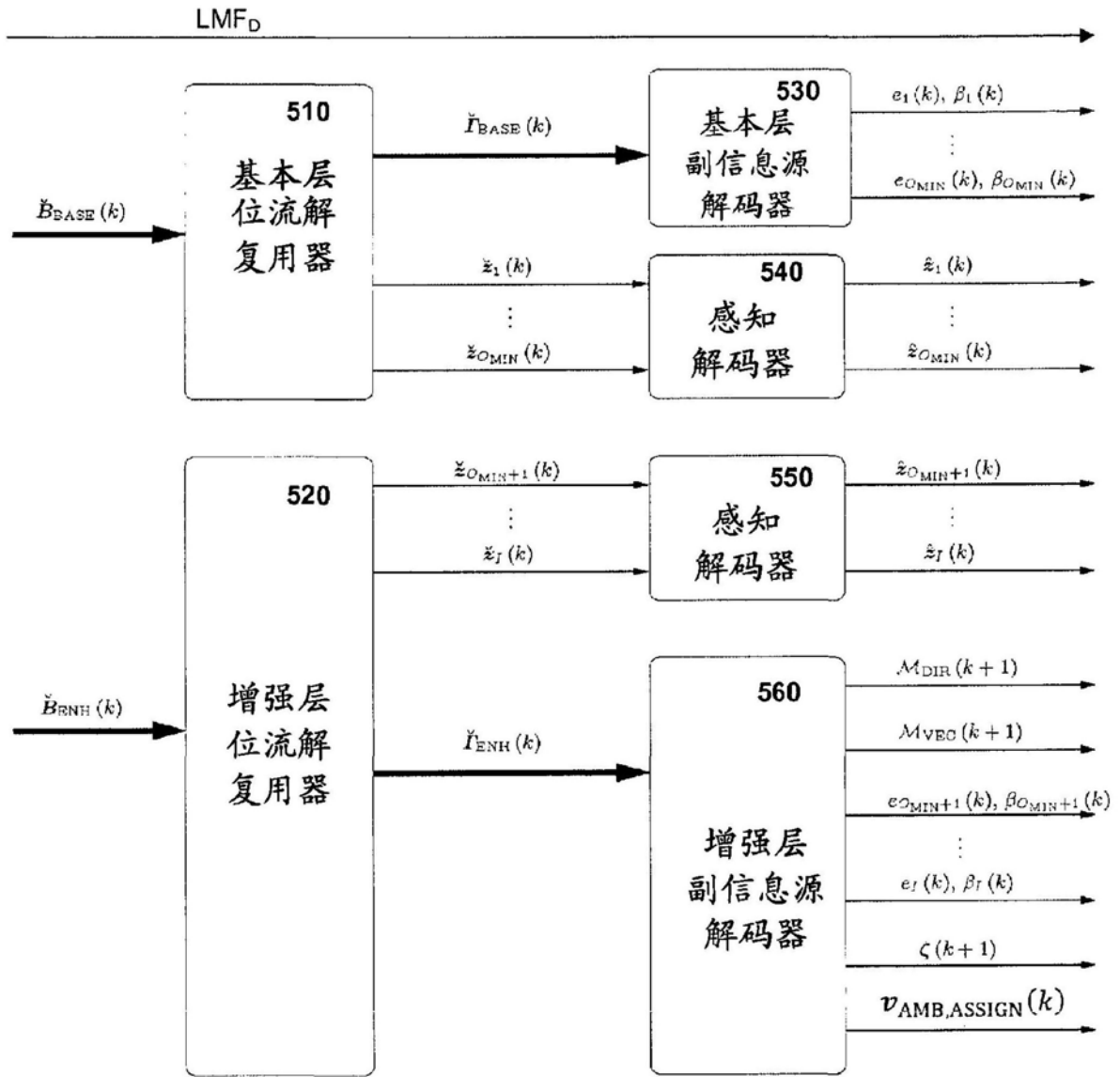


图5

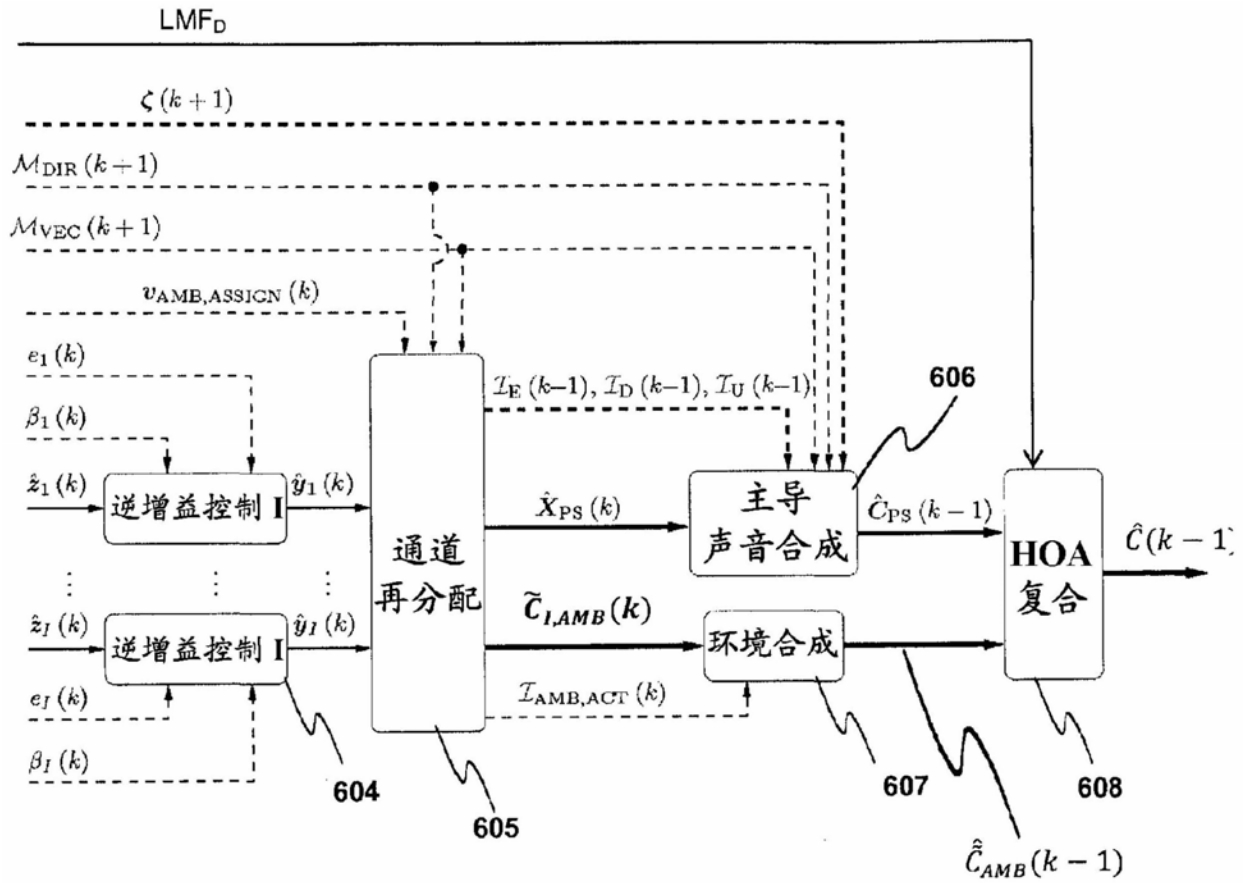


图6

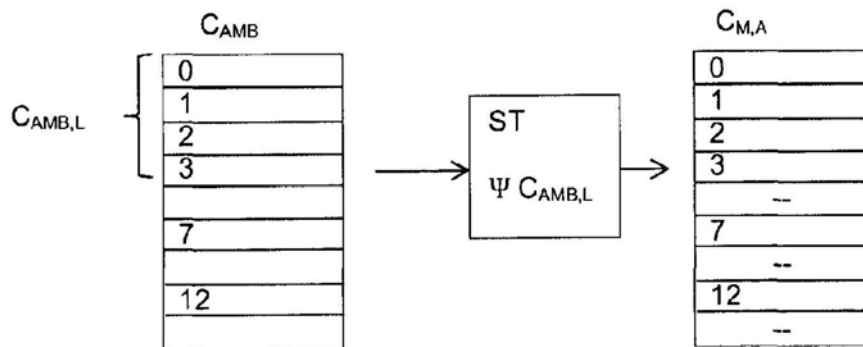


图7

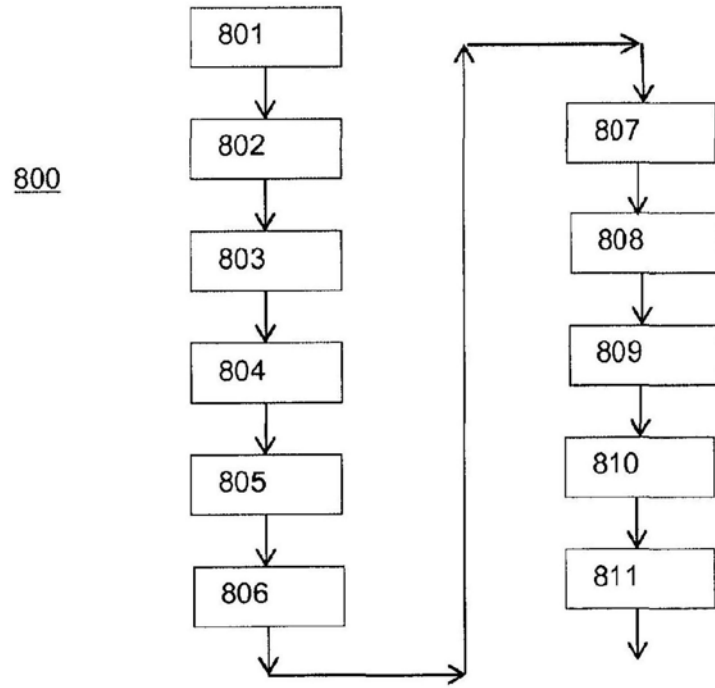


图8

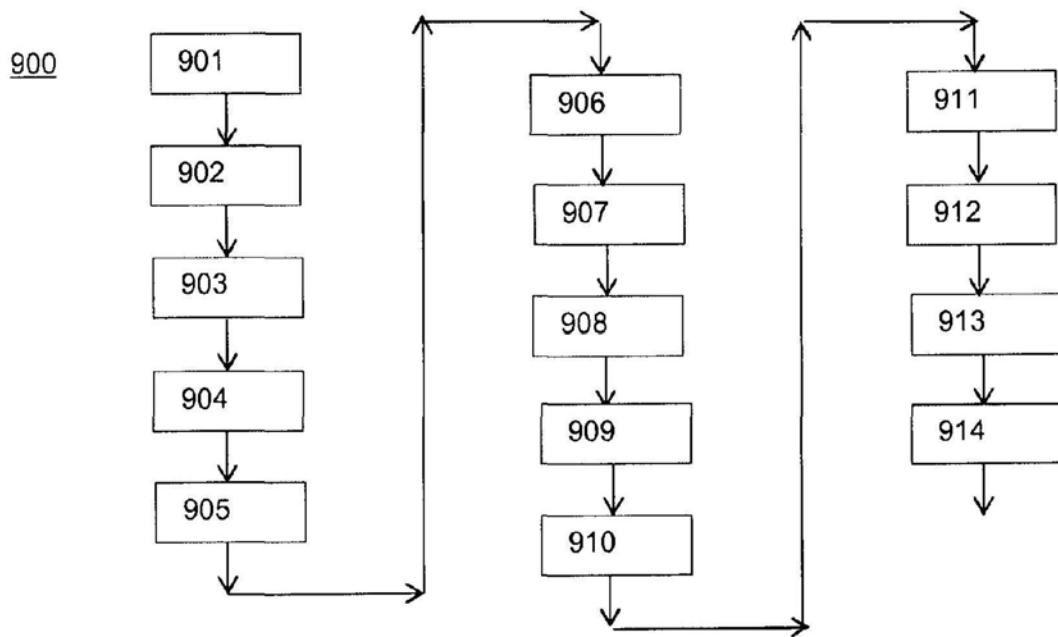


图9

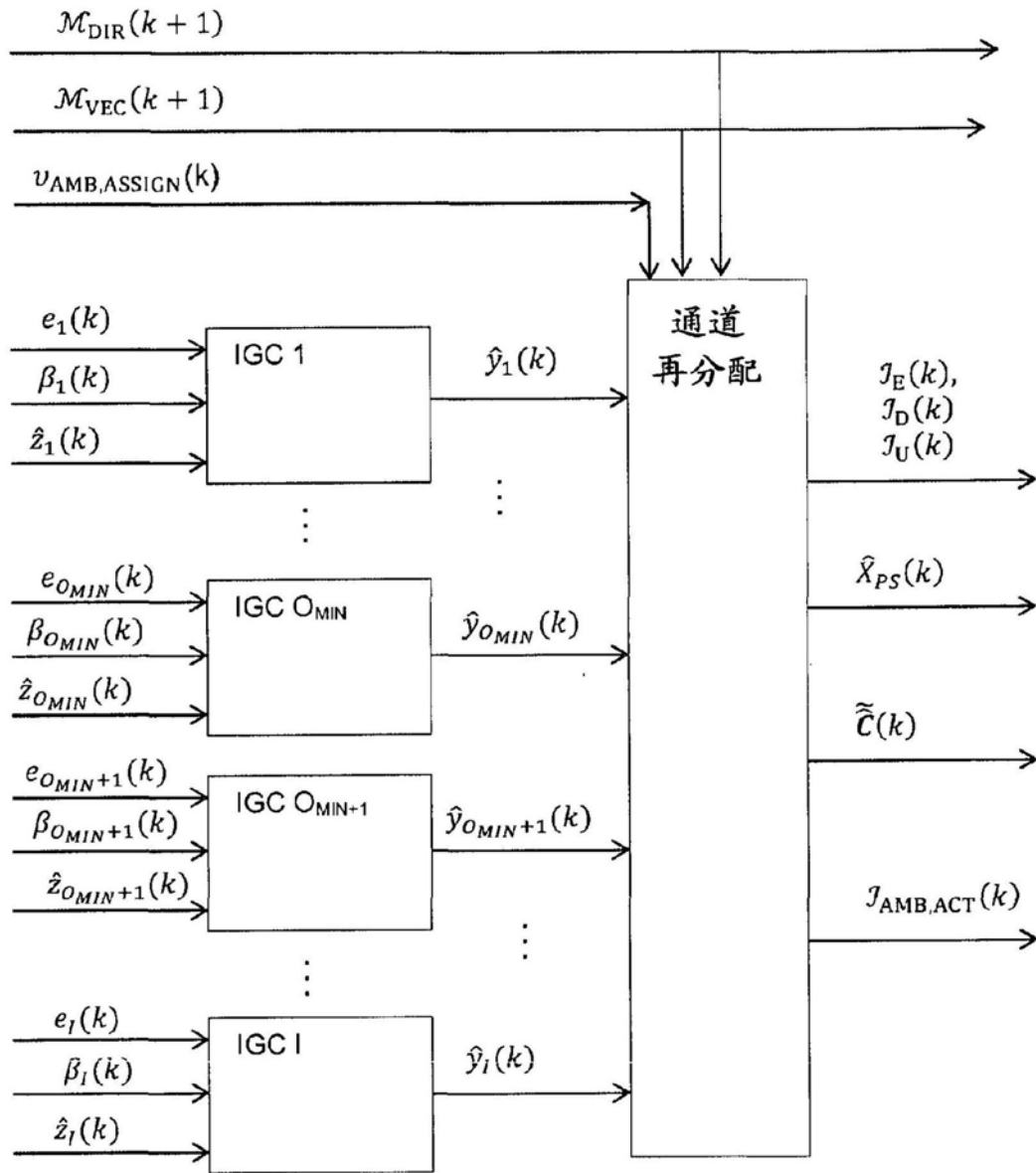


图10