



(12) 发明专利

(10) 授权公告号 CN 106663433 B

(45) 授权公告日 2020.12.29

(21) 申请号 201580033805.9

(72) 发明人 尼尔斯·京特·彼得斯

(22) 申请日 2015.07.02

迪潘让·森

(65) 同一申请的已公布的文献号

马丁·詹姆斯·莫雷尔

申请公布号 CN 106663433 A

(74) 专利代理机构 北京律盟知识产权代理有限公司 11287

(43) 申请公布日 2017.05.10

代理人 宋献涛

(30) 优先权数据

(51) Int.CI.

62/020,348 2014.07.02 US

G10L 19/008 (2013.01)

62/060,512 2014.10.06 US

H04R 5/04 (2006.01)

14/789,961 2015.07.01 US

H04S 3/00 (2006.01)

(85) PCT国际申请进入国家阶段日

(56) 对比文件

2016.12.22

CN 102547549 A, 2012.07.04

(86) PCT国际申请的申请数据

CN 103250207 A, 2013.08.14

PCT/US2015/038943 2015.07.02

EP 2688066 A1, 2014.01.22

(87) PCT国际申请的公布数据

审查员 王立华

W02016/004277 EN 2016.01.07

(73) 专利权人 高通股份有限公司

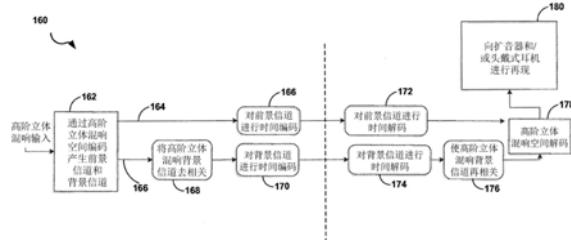
权利要求书3页 说明书28页 附图7页

(54) 发明名称

用于处理音频数据的方法和装置

(57) 摘要

一般来说，本发明描述用于压缩和解码音频数据的技术。用于压缩音频数据的实例装置包含一个或多个处理器，所述一个或多个处理器经配置以将去相关变换应用于环境立体混响系数，并且获得所述环境立体混响系数的去相关表示。所述系数是从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量，其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联。



1.一种用于处理音频数据的方法,其包括:

获得具有至少一左信号和一右信号的环境立体混响系数的经去相关表示,所述环境立体混响系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量,所述环境立体混响系数的所述经去相关表示是已使用基于相位的变换经去相关,其中所述多个高阶立体混响系数中的至少一者与具有阶数一或零的球形基底函数相关联;

将再相关变换应用于所述环境立体混响系数的所述经去相关表示以获得多个相关的环境立体混响系数;和

基于从所述环境立体混响系数的所述经去相关表示获得的所述多个相关的环境立体混响系数而产生扬声器馈送。

2.根据权利要求1所述的方法,其中应用所述再相关变换包括将基于相位的逆变换应用于所述环境立体混响系数。

3.根据权利要求2所述的方法,其中所述基于相位的逆变换已根据全三维归一化经归一化。

4.根据权利要求2所述的方法,其中所述基于相位的逆变换已根据施密特半归一化经归一化。

5.根据权利要求2所述的方法,其中所述环境立体混响系数与具有阶数零或阶数一的球形基底函数相关联,且其中应用所述基于相位的逆变换包括对于所述环境立体混响系数的所述经去相关表示执行所述基于相位的变换的标量乘法。

6.根据权利要求1所述的方法,其进一步包括获得对环境立体混响系数的所述经去相关表示是通过去相关变换经去相关的指示。

7.根据权利要求1所述的方法,其进一步包括获得定义所述声场的前景分量的空间特性的一或多个空间分量,所述空间分量定义于球面谐波域中且通过对于所述多个高阶立体混响系数执行分解而产生,

其中产生所述扬声器馈送包括组合所述相关的环境立体混响系数与基于所述一或多个空间分量获得的一或多个前景信道。

8.一种用于压缩音频数据的方法,其包括:

将基于相位的去相关变换应用于环境立体混响系数以获得所述环境立体混响系数的经去相关表示,所述环境立体混响系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量,其中所述多个高阶立体混响系数中的至少一者与具有阶数一或零的球形基底函数相关联。

9.根据权利要求8所述的方法,其进一步包括根据全三维归一化将所述基于相位的变换进行归一化。

10.根据权利要求8所述的方法,其进一步包括根据施密特半归一化将所述基于相位的变换进行归一化。

11.根据权利要求8所述的方法,其中所述环境立体混响系数与具有阶数零或阶数一的球形基底函数相关联,且其中将所述基于相位的变换应用于所述环境立体混响系数包括对于所述环境立体混响系数的至少一个子集执行所述基于相位的变换的标量乘法。

12.根据权利要求8所述的方法,其进一步包括用信号发送对所述去相关变换已应用于

所述环境立体混响系数的指示。

13. 一种用于处理音频数据的装置，所述装置包括：

存储器，其经配置以存储待处理的所述音频数据的至少一部分；和  
一或多个处理器，其经配置以：

获得具有至少一左信号和一右信号的环境立体混响系数的经去相关表示，所述环境立体混响系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量，所述环境立体混响系数的所述经去相关表示是已使用基于相位的变换经去相关，其中所述多个高阶立体混响系数中的至少一者与具有阶数一或零的球形基底函数相关联；

将再相关变换应用于所述环境立体混响系数的所述经去相关表示以获得多个相关的环境立体混响系数；和

基于所述环境立体混响系数的所述经去相关表示而产生扬声器馈送。

14. 根据权利要求13所述的装置，其中为了产生所述扬声器馈送，所述一或多个处理器经配置以基于所述左信号产生左扬声器馈送且基于所述右信号产生右扬声器馈送，所述左扬声器馈送和所述扬声器馈送用于由立体声再现系统输出。

15. 根据权利要求13所述的装置，其中为了产生所述扬声器馈送，所述一或多个处理器经配置以在不将所述再相关变换应用于所述右信号和所述左信号的情况下，使用所述左信号作为左扬声器馈送且使用所述右信号作为右扬声器馈送。

16. 根据权利要求13所述的装置，其中为了产生所述扬声器馈送，所述一或多个处理器经配置以混合所述左信号与所述右信号以用于由单声道音频系统输出。

17. 根据权利要求13所述的装置，其中为了产生所述扬声器馈送，所述一或多个处理器经配置以组合相关的环境立体混响系数与一或多个前景信道。

18. 根据权利要求13所述的装置，其中所述一或多个处理器进一步经配置以确定没有前景信道可用于与所述相关的环境立体混响系数组合。

19. 根据权利要求13所述的装置，其中所述一或多个处理器进一步经配置以：

确定将经由单声道音频再现系统输出所述声场；和

对包含用于由所述单声道音频再现系统输出的数据的经去相关的环境立体混响系数的至少一个子集进行解码。

20. 根据权利要求13所述的装置，其中所述一或多个处理器进一步经配置以获得对环境立体混响系数的所述经去相关表示是通过去相关变换经去相关的指示。

21. 根据权利要求13所述的装置，其进一步包括扩音器，所述扩音器经配置以输出基于所述环境立体混响系数的所述经去相关表示产生的所述扬声器馈送。

22. 一种用于压缩音频数据的装置，所述装置包括：

存储器，其经配置以存储待压缩的所述音频数据的至少一部分；和  
一或多个处理器，其经配置以：

将基于相位的去相关变换应用于环境立体混响系数以获得所述环境立体混响系数的经去相关表示，所述环境立体混响系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量，其中所述多个高阶立体混响系数中的至少一者与具有阶数一或零的球形基底函数相关联。

23. 根据权利要求22所述的装置，其中所述一或多个处理器进一步经配置以用信号发送所述经去相关的环境立体混响系数以及一或多个前景信道。

24. 根据权利要求22所述的装置，其中为了用信号发送所述经去相关的环境立体混响系数以及一或多个前景信道，所述一或多个处理器经配置以响应于确定目标位速率符合或超过预定阈值而用信号发送所述经去相关的环境立体混响系数以及一或多个前景信道。

25. 根据权利要求22所述的装置，其中所述一或多个处理器进一步经配置以在不用信号发送任何前景信道的情况下用信号发送所述经去相关的环境立体混响系数。

26. 根据权利要求25所述的装置，其中为了在不用信号发送任何前景信道的情况下用信号发送所述经去相关的环境立体混响系数，所述一或多个处理器经配置以响应于确定目标位速率低于预定阈值而在不用信号发送任何前景信道的情况下用信号发送经去相关的环境立体混响系数。

27. 根据权利要求26所述的装置，其中所述一或多个处理器进一步经配置以用信号发送对所述去相关变换已应用于所述环境立体混响系数的指示。

28. 根据权利要求22所述的装置，其进一步包括麦克风，所述麦克风经配置以捕获待压缩的所述音频数据。

## 用于处理音频数据的方法和装置

- [0001] 本申请案主张以下各者的权益：
- [0002] 第62/020,348号美国临时专利申请案，其标题为“减少HOA背景信道之间的相关性 (REDUCING CORRELATION BETWEEN HOA BACKGROUND CHANNELS)”，在2014年7月2日申请；和
- [0003] 第62/060,512号美国临时专利申请案，其标题为“减少HOA背景信道之间的相关性 (REDUCING CORRELATION BETWEEN HOA BACKGROUND CHANNELS)”，在2014年10月6日申请，
- [0004] 其中的每一者的全部内容以引用的方式并入本文中。

### 技术领域

[0005] 本发明涉及音频数据，且更确切地说，涉及高阶立体混响音频数据的译码。

### 背景技术

[0006] 高阶立体混响 (HOA) 信号 (通常由多个球面谐波系数 (SHC) 或其它分层元素表示) 是声场的三维表示。HOA或SHC表示可以独立于用以回放从SHC信号再现的多信道音频信号的局部扬声器几何布置的方式来表示声场。SHC信号还可促进向后兼容性，因为可将SHC信号再现为众所周知的且被广泛采用的多信道格式 (例如，5.1音频信道格式或7.1音频信道格式)。SHC表示因此可实现对声场的更好表示，其也适应向后兼容性。

### 发明内容

[0007] 一般来说，描述用于对高阶立体混响音频数据进行译码的技术。高阶立体混响音频数据可包括对应于具有大于一的阶数的球面谐波基底函数的至少一个高阶立体混响 (HOA) 系数。描述用于减少高阶立体混响 (HOA) 背景信道之间的相关性的技术。

[0008] 在一个方面，一种方法包含：获得具有至少一左信号和一右信号的环境立体混响系数的经去相关表示，所述环境立体混响系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量，其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联；和基于所述环境立体混响系数的所述经去相关表示而产生扬声器馈送。

[0009] 在另一方面，一种方法包含：将去相关变换应用于环境立体混响系数以获得所述环境立体混响系数的经去相关表示，所述环境HOA系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量，其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联。

[0010] 在另一方面，一种用于压缩音频数据的装置包含一或多个处理器，其经配置以：获得具有至少一左信号和一右信号的环境立体混响系数的经去相关表示，所述环境立体混响系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量，其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联；和基于所述环境立体混响系数的所述经去相关表示而产生扬声器馈送。

[0011] 在另一方面，一种用于压缩音频数据的装置包含一或多个处理器，其经配置以：将

去相关变换应用于环境立体混响系数以获得所述环境立体混响系数的经去相关表示,所述环境HOA系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量,其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联。

[0012] 在另一方面,一种用于压缩音频数据的装置包含:用于获得具有至少一左信号和一右信号的环境立体混响系数的经去相关表示的装置,所述环境立体混响系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量,其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联;和用于基于所述环境立体混响系数的所述经去相关表示而产生扬声器馈送的装置。

[0013] 在另一方面,一种用于压缩音频数据的装置包含:用于将去相关变换应用于环境立体混响系数以获得所述环境立体混响系数的经去相关表示的装置,所述环境HOA系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量,其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联;和用于存储所述环境立体混响系数的所述经去相关表示的装置。

[0014] 在另一方面,用指令对计算机可读存储媒体进行编码,所述指令在执行时致使音频压缩装置的一或多个处理器:获得具有至少一左信号和一右信号的环境立体混响系数的经去相关表示,所述环境立体混响系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量,其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联;和基于所述环境立体混响系数的所述经去相关表示而产生扬声器馈送。

[0015] 在另一方面,用指令对计算机可读存储媒体进行编码,所述指令在执行时致使音频压缩装置的一或多个处理器:将去相关变换应用于环境立体混响系数以获得所述环境立体混响系数的经去相关表示,所述环境HOA系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量,其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联。

[0016] 在附图和以下描述中陈述所述技术的一或多个方面的细节。所述技术的其它特征、目标和优点将从所述描述和图式以及权利要求书而显而易见。

## 附图说明

[0017] 图1是说明具有各种阶数和子阶数的球面谐波基底函数的图。

[0018] 图2是说明可执行本发明中描述的技术的各种方面的系统的图。

[0019] 图3是更详细地说明在图2的实例中所展示的可执行本发明中描述的技术的各种方面的音频编码装置的一个实例的框图。

[0020] 图4是更详细地说明图2的音频解码装置的框图。

[0021] 图5是说明音频编码装置执行本发明中描述的基于向量的合成技术的各种方面的示范性操作的流程图。

[0022] 图6A是说明音频解码装置执行本发明中描述的技术的各种方面的示范性操作的流程图。

[0023] 图6B是说明音频编码装置和音频解码装置执行本发明中描述的译码技术的示范

性操作的流程图。

### 具体实施方式

[0024] 环绕声的演进现今已使得许多输出格式可用于娱乐。此等消费型环绕声格式的实例大部分为基于“信道”的，这是因为其以特定几何坐标隐含地指定去往扩音器的馈送。消费型环绕声格式包含普遍的5.1格式(其包含以下六个信道：左前(FL)、右前(FR)、中心或前中心、左后或左环绕、右后或右环绕，以及低频效应(LFE))、发展中的7.1格式、包含高度扬声器的各种格式，例如7.1.4格式和22.2格式(例如，供与超高清电视标准一起使用)。非消费型格式可涵括任何数目个扬声器(成对称和非对称几何布置)，其常常被为“环绕阵列”。此类阵列的一个实例包含定位在截角二十面体的拐角上的坐标处的32个扩音器。

[0025] 去往未来MPEG编码器的输入视情况为以下三个可能格式中的一者：(i)传统的基于信道的音频(如上文所论述)，其意图由处于预先指定的位置处的扩音器播放；(ii)基于对象的音频，其涉及用于单个音频对象的具有含有其位置坐标(以及其他信息)的相关联元数据的离散脉码调制(PCM)数据；以及(iii)基于场景的音频，其涉及使用球面谐波基底函数的系数(也称为“球面谐波系数”或SHC、“高阶立体混响”或HOA以及“HOA系数”)来表示声场。所述未来MPEG编码器更详细地描述于国际标准化组织/国际电工委员会(ISO)/(IEC)JTC1/SC29/WG11/N13411的标题为“要求对于3D音频的提议(Call for Proposals for 3D Audio)”的文献中，所述文献于2013年1月在瑞士日内瓦发布，且可在<http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w13411.zip>获得。

[0026] 在市场中存在各种基于信道的“环绕声”格式。它们的范围(例如)是从5.1家庭影院系统(其在使起居室享有立体声效方面已获得最大成功)到NHK(日本广播协会(Nippon Hoso Kyokai)或日本广播公司(Japan Broadcasting Corporation))所开发的22.2系统。内容创建者(例如，好莱坞工作室)将希望一次产生电影的原声带，而不花费精力来针对每一扬声器配置对其进行再混合。近来，标准开发组织(Standards Developing Organizations)一直在考虑如下方式：提供到标准化位流中的编码，以及后续解码，其可调且不知晓回放位置(涉及再现器)处的扬声器几何布置(和数目)以及声学条件。

[0027] 为向内容创建者提供此类灵活性，可使用分层要素集合来表示声场。所述分层要素集合可指其中元素经排序以使得较低阶元素的基础集合提供模型化声场的完整表示的元素集合。在所述集合经扩展以包含高阶元素时，所述表示变得更详细，从而增加分辨率。

[0028] 分层要素集合的一个实例为球面谐波系数(SHC)集合。以下表达式示范使用SHC对声场的描述或表示：

$$[0029] p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[ 4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t},$$

[0030] 所述表达式展示在时间t处在声场的任何点 $\{r_r, \theta_r, \varphi_r\}$ 处的压力 $p_i$ 可由SHC、 $A_n^m(k)$ 来唯一地表示。此处， $k = \frac{\omega}{c}$ ，c是声音的速度(约343m/s)， $\{r_r, \theta_r, \varphi_r\}$ 是参考点(或观测点)， $j_n(\cdot)$ 是阶n的球面贝塞尔函数，且 $Y_n^m(\theta_r, \varphi_r)$ 是阶数n和子阶数m的球面谐波基底函

数。可认识到,方括号中的术语为信号(即, $S(\omega, r_r, \theta_r, \varphi_r)$ )的频域表示,其可由各种时间-频率变换(例如离散傅里叶变换(DFT)、离散余弦变换(DCT)或小波变换)近似表示。分层集合的其它实例包含小波变换系数的集合和多分辨率基底函数的系数的其它集合。通过截断高阶以使得仅留存零阶和一阶来处理高阶立体混响信号。归因于高阶系数的能量损失,通常会对剩余的信号进行某些能量补偿。

[0031] 本发明的各种方面针对于减少背景信号间的相关性。举例来说,本发明的技术可减少或可能地消除在HOA域中表达的背景信号之间的相关性。减少背景HOA信号之间的相关性的潜在优点是减低噪声去掩蔽。如本文中所使用,表达“噪声去掩蔽”可指将音频对象归属于在空间域中不对应于所述音频对象的位置。除减少与噪声去掩蔽有关的潜在问题之外,本文中所描述的编码技术还可产生表示左音频信号和右音频信号(例如一起形成立体声输出的信号)的输出信号。继而,解码装置可解码左音频信号和右音频信号以获得立体声输出,或可混合左音频信号与右音频信号以获得单声道输出。另外,在经编码位流表示纯水平布局的情境中,解码装置可实施本发明的各种技术以仅解码水平分量去相关HOA背景信号。通过将解码过程限制于水平分量去相关HOA背景信号,解码器可实施所述技术以节省计算资源并且减少带宽消耗。

[0032] 图1是说明从零阶( $n=0$ )到四阶( $n=4$ )的球面谐波基底函数的图。如可见,对于每一阶,存在子阶数m的扩展,出于易于说明的目的,在图1的实例中展示所述子阶数但未明确注释。

[0033] 可通过各种麦克风阵列配置物理地获取(例如,记录)SHC  $A_n^m(k)$ ,或替代地,其可从声场的基于信道或基于对象的描述导出。SHC表示基于场景的音频,其中SHC可输入到音频编码器以获得经编码SHC,所述经编码SHC可促成更有效的传输或存储。举例来说,可使用涉及 $(1+4)^2$ (25,且因此为四阶)系数的四阶表示。

[0034] 如上文所提到,可使用麦克风阵列从麦克风记录导出SHC。可如何从麦克风阵列导出SHC的各种实例描述于波莱蒂·M(Poletti,M)的“基于球面谐波的三维环绕声系统(Three-Dimensional Surround Sound Systems Based on Spherical Harmonics)”(听觉工程学协会会刊(J.Audio Eng.Soc.),第53卷,第11期,2005年11月,第1004到1025页)中。

[0035] 为说明可如何从基于对象的描述导出SHC,考虑以下方程式。可将对应于个别音频对象的声场的系数 $A_n^m(k)$ 表达为:

$$[0036] A_n^m(k) = g(\omega)(-4\pi ik)h_n^{(2)}(kr_s)Y_n^m(\theta_s, \varphi_s),$$

[0037] 其中i是 $\sqrt{-1}$ , $h_n^{(2)}(\cdot)$ 是阶n的球面汉克尔函数(第二种类),且 $\{r_s, \theta_s, \varphi_s\}$ 是对象的位置。已知随频率变化的对象源能量 $g(\omega)$ (例如,使用时间-频率分析技术,例如对PCM流执行快速傅里叶变换)允许将每一PCM对象和对应位置转换成SHC  $A_n^m(k)$ 。此外,可展示(由于以上是线性和正交分解)用于每一对象的 $A_n^m(k)$ 系数是累加的。以此方式,众多PCM对象可由 $A_n^m(k)$ 系数(例如,作为个别对象的系数向量的总和)来表示。基本上,所述系数含有关于声场的信息(随3D坐标而变的压力),且上述情形表示在观测点 $\{r_r, \theta_r, \varphi_r\}$ 附近从个别对象到整个声场的表示的变换。下文在基于对象和基于SHC的音频译码的上下文中描述剩余

各图。

[0038] 图2是说明可执行本发明中描述的技术的各种方面的系统10的图。如图2的实例中所展示,系统10包含内容创建者装置12和内容消费者装置14。尽管在内容创建者装置12和内容消费者装置14的上下文中描述,但可在其中声场的SHCt也可称为HOA系数)或任何其它分层表示经编码以形成表示音频数据的位流的任何上下文中实施所述技术。此外,内容创建者装置12可表示能够实施本发明中所描述的技术的任何形式的计算装置,包含手持机(或蜂窝式电话)、平板计算机、智能电话或台式计算机(提供几个实例)。同样地,内容消费者装置14可表示能够实施本发明中所描述的技术的任何形式的计算装置,包含手持机(或蜂窝式电话)、平板计算机、智能电话、机顶盒或台式计算机(提供几个实例)。

[0039] 内容创建者装置12可由电影工作室或可产生多信道音频内容以供内容消费者装置(例如,内容消费者装置14)的操作者消耗的其它实体来操作。在一些实例中,内容创建者装置12可由将希望压缩HOA系数11的个别用户操作。内容创建者通常产生音频内容与视频内容。内容消费者装置14可由个人操作。内容消费者装置14可包含音频回放系统16,其可指能够再现SHC以供作为多信道音频内容回放的任何形式的音频回放系统。

[0040] 内容创建者装置12包含音频编辑系统18。内容创建者装置12获得各种格式(包含直接作为HOA系数)的实况记录7和音频对象9,内容创建者装置12可使用音频编辑系统18对其进行编辑。麦克风5可捕获实况记录7。内容创建者可在编辑过程期间再现来自音频对象9的HOA系数11,从而收听所再现的扬声器馈送以试图标识需要进一步编辑的声场的各种方面。内容创建者装置12接着可编辑HOA系数11(潜在地通过操纵可以上文所描述的方式从中导出源HOA系数的音频对象9中的不同者而间接地编辑)。内容创建者装置12可采用音频编辑系统18产生HOA系数11。音频编辑系统18表示能够编辑音频数据且输出所述音频数据作为一或多个源球面谐波系数的任何系统。

[0041] 当编辑过程完成时,内容创建者装置12可基于HOA系数11产生位流21。也就是说,内容创建者装置12包含音频编码装置20,所述音频编码装置表示经配置以根据本发明中描述的技术的各种方面编码或以其它方式压缩HOA系数11以产生位流21的装置。音频编码装置20可产生位流21以用于跨越传输信道(其可为有线或无线信道、数据存储装置或其类似者)传输(作为一个实例)。位流21可表示HOA系数11的经编码版本,且可包含主要位流和另一旁侧位流(其可称为旁侧信道信息)。

[0042] 虽然在图2中展示为直接传输到内容消费者装置14,但内容创建者装置12可将位流21输出到位于内容创建者装置12与内容消费者装置14之间的中间装置。中间装置可存储位流21以用于稍后递送到可请求所述位流的内容消费者装置14。所述中间装置可包括文件服务器、网络服务器、台式计算机、膝上型计算机、平板计算机、移动电话、智能电话,或能够存储位流21以供音频解码器稍后检索的任何其它装置。中间装置可驻留在能够将位流21流式传输(且可能结合传输对应视频数据位流)到请求位流21的订户(例如,内容消费者装置14)的内容递送网络中。

[0043] 替代地,内容创建者装置12可将位流21存储到存储媒体,例如压缩光盘、数字视频光盘、高清视频光盘或其它存储媒体,其中大多数能够由计算机读取且因此可称为计算机可读存储媒体或非暂时性计算机可读存储媒体。在此上下文中,传输信道可指借以传输存储到媒体的内容的信道(且可包含小量存储(retail stores)和其它基于存储的(store-

based) 递送机制)。因而,在任何情况下,就此而言,本发明的技术不应限于图2的实例。

[0044] 如图2的实例中进一步展示,内容消费者装置14包含音频回放系统16。音频回放系统16可表示能够回放多信道音频数据的任何音频回放系统。音频回放系统16可包含多个不同的再现器22。再现器22可各自提供用于不同形式的再现,其中所述不同形式的再现可包含执行向量基振幅移动(VBAP)的各种方式中的一或多或少者,以及/或执行声场合成的各种方式中的一或多或少者。如本文中所使用,“A和/或B”意味着“A或B”,或“A和B”两者。

[0045] 音频回放系统16可进一步包含音频解码装置24。音频解码装置24可表示经配置以解码来自位流21的HOA系数11’的装置,其中HOA系数11’可类似于HOA系数11,但归因于有损操作(例如,量化)和/或经由传输信道的传输而不同。音频回放系统16可在解码位流21之后获得HOA系数11’并再现HOA系数11’以输出扩音器馈送25。扩音器馈送25可驱动一或多个扩音器(其为便于说明的目的在图2的实例中未示出)。

[0046] 为了选择适当再现器或在一些例子中产生适当再现器,音频回放系统16可获得指示扩音器的数目和/或扩音器的空间几何布置的扩音器信息13。在一些例子中,音频回放系统16可使用参考麦克风获得扩音器信息13且以动态地确定扩音器信息13的方式驱动扩音器。在其它例子中或结合动态确定扩音器信息13,音频回放系统16可提示用户与音频回放系统16介接并输入扩音器信息13。

[0047] 音频回放系统16接着可基于扩音器信息13选择音频再现器22中的一者。在一些例子中,当音频再现器22中无一者处于到在扩音器信息13中所指定的扩音器几何布置的某一阈值相似性量度(就扩音器几何布置来说)内时,音频回放系统16可基于扩音器信息13产生音频再现器22中的一者。音频回放系统16可在一些例子中基于扩音器信息13产生音频再现器22中的一者,而无需首先试图选择音频再现器22中的现有者。一或多个扬声器3接着可回放再现的扩音器馈送25。

[0048] 图3是更详细地说明在图2的实例中所展示的可执行本发明中描述的技术的各种方面的音频编码装置20的一个实例的框图。音频编码装置20包含内容分析单元26、基于向量的合成方法单元27、基于方向的合成方法单元28,以及去相关单元40’。尽管下文简单描述,但关于音频编码装置20和压缩或以其它方式编码HOA系数的各种方面的更多信息可在2014年5月29日申请的标题为“用于声场的经分解表示的内插(INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD)”的第WO 2014/194099号国际专利申请公开案中获得。

[0049] 内容分析单元26表示经配置以分析HOA系数11的内容以标识HOA系数11是表示从实况记录产生的内容还是表示从音频对象产生的内容的单元。内容分析单元26可确定HOA系数11是从实际声场的记录产生还是从人工音频对象产生。在一些例子中,当帧式HOA系数11是从记录产生时,内容分析单元26将HOA系数11传递到基于向量的分解单元27。在一些例子中,当帧式HOA系数11是从合成音频对象产生时,内容分析单元26将HOA系数11传递到基于方向的合成单元28。基于方向的合成单元28可表示经配置以执行HOA系数11的基于方向的合成以产生基于方向的位流21的单元。

[0050] 如在图3的实例中所展示,基于向量的分解单元27可包含线性可逆变换(LIT)单元30、参数计算单元32、重排序单元34、前景选择单元36、能量补偿单元38、心理声学音频译码器单元40、位流产生单元42、声场分析单元44、系数减少单元46、背景(BG)选择单元48、空

间-时间内插单元50以及量化单元52。

[0051] 线性可逆变换(LIT)单元30接收呈HOA信道形式的HOA系数11,所述HOA信道中的每一信道表示与球面基底函数的给定阶数、子阶数相关联的系数的块或帧(其可标示为HOA[k],其中k可标示样本的当前帧或块)。HOA系数11的矩阵可具有维度D:M×(N+1)<sup>2</sup>。

[0052] LIT单元30可表示经配置以执行被称为奇异值分解的分析形式的单元。虽然关于SVD加以描述,但可对于提供线性不相关的能量密集输出的集合的任何类似变换或分解来执行本发明中所描述的技术。而且,本发明中对“集合”的参考一般意图指非零集合(除非特定地相反陈述),且并不意图指包含所谓的“空集合”的集合的经典数学定义。替代变换可包括通常被称为“PCA”的主分量分析。取决于上下文,PCA可由若干不同名称指代,例如(仅举几例)离散卡亨南-洛维变换、霍特林变换、恰当正交分解(POD)和本征值分解(EVD)。有利于压缩音频数据的基本目标的此类操作的特性为多信道音频数据的“能量压缩”和“去相关”。

[0053] 在任何情况下,出于实例的目的,假设LIT单元30执行奇异值分解(其又可被称为“SVD”),LIT单元30可将HOA系数11变换成两个或多于两个经变换HOA系数的集合。经变换HOA系数的“集合”可包含经变换HOA系数的向量。在图3的实例中,LIT单元30可对于HOA系数11执行SVD以产生所谓的V矩阵、S矩阵及U矩阵。在线性代数中,SVD可以如下形式表示y乘z实数或复数矩阵X(其中X可表示多信道音频数据,例如HOA系数11)的因子分解:

[0054]  $X = USV^*$

[0055] U可表示y乘y实数或复数单式矩阵,其中U的y列被称为多信道音频数据的左奇异向量。S可表示在对角线上具有非负实数的y乘z矩形对角线矩阵,其中S的对角线值被称为多信道音频数据的奇异值。 $V^*$ (其可标示V的共轭转置)可表示z乘z实数或复数单式矩阵,其中 $V^*$ 的z列被称为多信道音频数据的右奇异向量。

[0056] 在一些实例中,以上提及的SVD数学表达式中的 $V^*$ 矩阵标示为V矩阵的共轭转置以反映SVD可应用于包括复数的矩阵。当应用于仅包括实数的矩阵时,V矩阵的复数共轭(或换句话说, $V^*$ 矩阵)可被视为V矩阵的转置。下文中为容易说明的目的,假设HOA系数11包括实数,结果是经由SVD而非 $V^*$ 矩阵输出V矩阵。此外,尽管在本发明中标示为V矩阵,但对V矩阵的提及应理解为在适当的情况下涉及V矩阵的转置。尽管假设为V矩阵,但所述技术可以类似方式应用于具有复数系数的HOA系数11,其中SVD的输出为 $V^*$ 矩阵。因此,就此而言,所述技术不应限于仅提供应用SVD以产生V矩阵,而是可包含将SVD应用于具有复数分量的HOA系数11以产生 $V^*$ 矩阵。

[0057] 以此方式,LIT单元30可对于HOA系数11执行SVD以输出具有维度D:M×(N+1)<sup>2</sup>的US[k]向量33(其可表示S向量和U向量的组合版本)以及具有维度D:(N+1)<sup>2</sup>×(N+1)<sup>2</sup>的V[k]向量35。US[k]矩阵中的个别向量元素也可被称为 $X_{PS}(k)$ ,而V[k]矩阵中的个别向量也可被称为v(k)。

[0058] U、S和V矩阵的分析可显示这些矩阵携载或表示上文由X表示的基本声场的空间和时间特性。U(长度为M个样本)中的N个向量中的每一者可表示随时间而变(对于由M个样本表示的时间周期)的经归一化分离音频信号,其彼此正交且已与任何空间特性(其也可被称为方向信息)解耦。表示空间形状和位置(r、θ、Φ)的空间特性可替代地由V矩阵中的个别第i向量v<sup>(i)</sup>(k)(每一者具有长度(N+1)<sup>2</sup>)表示。v<sup>(i)</sup>(k)向量中的每一者的个别元素可表示HOA系数,其描述相关联音频对象的声场的形状(包含宽度)和位置。U矩阵和V矩阵中的向量均

经归一化以使得其均方根能量等于一。U中的音频信号的能量因而由S中的对角线元素表示。将U与S相乘以形成US[k] (具有个别向量元素X<sub>PS</sub>(k)) ,因而表示具有能量的音频信号。SVD分解以使音频时间信号 (U中)、其能量 (S中) 以及其空间特性 (V中) 解耦的能力可支持本发明中所描述的技术的各种方面。另外,通过US[k] 与V[k]的向量乘法合成基本HOA[k] 系数X的模型产生贯穿此文献使用的术语“基于向量的分解”。

[0059] 尽管描述为对于HOA系数11直接执行,但LIT单元30可将线性可逆变换应用到HOA系数11的导出项。举例来说,LIT单元30可对于从HOA系数11导出的功率谱密度矩阵应用SVD。通过对于HOA系数的功率谱密度 (PSD) 而非系数本身执行SVD,LIT单元30可在处理器循环和存储空间的一或者方面潜在地降低执行SVD的计算复杂性,同时实现与将SVD直接应用于HOA系数的情况相同的源音频编码效率。

[0060] 参数计算单元32表示经配置以计算各种参数的单元,所述参数例如相关性参数 (R)、方向特性参数 ( $\theta$ 、 $\Phi$ 、r) 和能量特性 (e)。当前帧的参数中的每一者可标示为R[k]、 $\theta$ [k]、 $\varphi$ [k]、r[k] 及 e[k]。参数计算单元32可对于US[k]向量33执行能量分析和/或相关(或所谓的交叉相关)以标识这些参数。参数计算单元32还可确定前一帧的参数,其中前一帧的参数可基于具有US[k-1]向量及V[k-1]向量的前一帧而标示为R[k-1]、 $\theta$ [k-1]、 $\varphi$ [k-1]、r[k-1] 和 e[k-1]。参数计算单元32可将当前参数37和先前参数39输出到重排序单元34。

[0061] 由参数计算单元32计算的参数可供重排序单元34用以将音频对象重排序以表示其自然评估或随时间推移的连续性。重排序单元34可将来自第一US[k]向量33的参数37中的每一者与第二US[k-1]向量33的参数39中的每一者在顺序方面进行比较。重排序单元34可基于当前参数37和先前参数39对US[k]矩阵33和V[k]矩阵35内的各种向量进行重排序(作为一个实例,使用匈牙利算法)以将经重排序的US[k]矩阵33' (其可数学标示为 $\bar{US}[k]$ ) 和经重排序的V[k]矩阵35' (其可数学标示为 $\bar{V}[k]$ ) 输出到前景声音(或主导声音(PS))选择单元36 (“前景选择单元36”) 和能量补偿单元38。

[0062] 声场分析单元44可表示经配置以对于HOA系数11执行声场分析以便潜在地实现目标位速率41的单元。声场分析单元44可基于所述分析和/或基于所接收目标位速率41,确定心理声学译码器实例化的总数目(其可为环境或背景信道的总数目(BG<sub>TOT</sub>) 的函数) 以及前景信道(或换句话说,主导信道)的数目。心理声学译码器实例化的总数可标示为numHOATransportChannels。

[0063] 再次为了潜在地实现目标位速率41,声场分析单元44还可确定前景信道的总数目(nFG) 45、背景(或换句话说,环境)声场的最小阶数(N<sub>BG</sub>或替代性地,MinAmbHOAorder)、表示背景声场的最小阶数的实际信道的对应数目(nBGa = (MinAmbHOAorder+1)<sup>2</sup>) ,以及要发送的额外BG HOA信道的索引(i) (其在图3的实例中可共同地标示为背景信道信息43)。背景信道信息42也可被称为环境信道信息43。保持来自numHOATransportChannels-nBGa的信道中的每一者可为“额外背景/环境信道”、“活动的基于向量的主导信道”、“活动的基于方向的主导信号”,或为“完全非活动的”。在一个方面中,信道类型可为通过两位指示(为“ChannelType”)的语法元素(例如,00:基于方向的信号;01:基于向量的主导信号;10:额外环境信号;11:非活动信号)。可由 (MinAmbHOAorder+1)<sup>2</sup>+索引10(在以上实例中)作为信道类型在所述帧的位流中出现的次数给出背景或环境信号的总数目nBGa。

[0064] 声场分析单元44可基于目标位速率41选择背景(或换句话说,环境)信道的数目和前景(或换句话说,主导)信道的数目,从而在目标位速率41相对较高时(例如,在目标位速率41等于或大于512Kbps时)选择更多背景和/或前景信道。在一个方面中,在位流的标头部分中, numHOATransportChannels 可设置为8,而 MinAmbHOAorder 可设置为1。在此情境下,在每个帧处,四个信道可专用于表示声场的背景或环境部分,而另4个信道可在逐帧基础上随信道类型而变化,例如任一者用作额外背景/环境信道或前景/主导信道。前景/主导信号可为基于向量或基于方向的信号中的一者,如上文所描述。

[0065] 在一些例子中,帧的基于向量的主导信号的总数目可通过 ChannelType 索引在所述帧的位流中为01的次数给出。在以上方面中,对于每个额外背景/环境信道(例如,对应于 ChannelType 10),可在所述信道中表示可能的HOA系数(除前四个以外)中的何者的对应信息。对于四阶HOA内容,所述信息可为指示HOA系数5到25的索引。可在 minAmbHOAorder 设置为1时始终发送前四个环境HOA系数1到4,因此,音频编码装置可能仅需要指示具有索引5到25的额外环境HOA系数中的一者。因此,可使用5位语法元素(针对四阶内容)发送所述信息,其可标示为“CodedAmbCoeffIdx”。在任何情况下,声场分析单元44将背景信道信息43和HOA系数11输出到背景(BG)选择单元36,将背景信道信息43输出到系数减少单元46和位流产生单元42,且将nFG 45输出到前景选择单元36。

[0066] 背景选择单元48可表示经配置以基于背景信道信息(例如,背景声场( $N_{BG}$ )以及要发送的额外BG HOA信道的数目( $nBGa$ )和索引( $i$ ))确定背景或环境HOA系数47的单元。举例来说,当 $N_{BG}$ 等于一时,背景选择单元48可选择具有等于或小于一的阶数的音频帧的每一样本的HOA系数11。在此实例中,背景选择单元48接着可选择具有由索引( $i$ )中的一者标识的索引的HOA系数11作为额外BG HOA系数,其中将待于位流21中指定的nBGa提供到位流产生单元42以便使得音频解码装置(例如,图2和4的实例中所展示的音频解码装置24)能够从位流21解析背景HOA系数47。背景选择单元48接着可将环境HOA系数47输出到能量补偿单元38。环境HOA系数47可具有维度D: $M \times [(N_{BG}+1)^2 + nBGa]$ 。环境HOA系数47也可被称为“环境HOA系数47”,其中环境HOA系数47中的每一者对应于待由心理声学音频译码器单元40编码的单独环境HOA信道47。

[0067] 前景选择单元36可表示经配置以基于nFG 45(其可表示标识前景向量的一或多个索引)选择表示声场的前景或相异分量的经重排序的US[k]矩阵33'和经重排序的V[k]矩阵35'的单元。前景选择单元36可将nFG信号49(其可表示为经重排序US[k]<sub>1,...,nFG</sub>、FG<sub>1,...,nFG</sub>[k]49,或 $X_{PS}^{(1..nFG)}(k)$ 49)输出到心理声学音频译码器单元40,其中nFG信号49可具有维度D: $M \times nFG$ ,且每一者表示单声道音频对象。前景选择单元36还可将对应于声场的前景分量的经重排序的V[k]矩阵35'(或 $v^{(1..nFG)}(k)$ 35')输出到空间-时间内插单元50,其中经重排序的V[k]矩阵35'中的对应于前景分量的子集可表示为具有维度D: $((N+1)^2 \times nFG)$ 的前景V[k]矩阵51<sub>k</sub>(其可在数学上表示为 $\bar{V}_{1,...,nFG}[k]$ )。

[0068] 能量补偿单元38可表示经配置以对于环境HOA系数47执行能量补偿以补偿归因于由背景选择单元48移除HOA信道中的各者而产生的能量损失的单元。能量补偿单元38可对于经重排序的US[k]矩阵33'、经重排序的V[k]矩阵35'、nFG信号49、前景V[k]向量51<sub>k</sub>和环境HOA系数47中的一或者执行能量分析,且接着基于所述能量分析执行能量补偿以产生

经能量补偿的环境HOA系数47'。能量补偿单元38可将经能量补偿的环境HOA系数47'输出到去相关单元40'。继而,去相关单元40'可实施本发明的技术以减少或消除HOA系数47'的背景信号之间的相关性以形成一或多个经去相关的HOA系数47"。去相关单元40'可将经去相关的HOA系数47"输出到心理声学音频译码器单元40。

[0069] 空间-时间内插单元50可表示经配置以接收第k帧的前景V[k]向量51<sub>k</sub>和前一帧(因此为k-1记法)的前景V[k-1]向量51<sub>k-1</sub>且执行空间-时间内插以产生经内插的前景V[k]向量的单元。空间-时间内插单元50可将nFG信号49与前景V[k]向量51<sub>k</sub>重新组合以恢复经重排序的前景HOA系数。空间-时间内插单元50接着可将经重排序的前景HOA系数除以经内插的V[k]向量以产生经内插的nFG信号49'。空间-时间内插单元50还可输出前景V[k]向量51<sub>k</sub>,所述前景V[k]向量51<sub>k</sub>用以产生经内插的前景V[k]向量,以使得例如音频解码装置24的音频解码装置可产生经内插的前景V[k]向量且借此恢复前景V[k]向量51<sub>k</sub>。将用以产生经内插的前景V[k]向量的前景V[k]向量51<sub>k</sub>标示为剩余的前景V[k]向量53。为了确保在编码器和解码器处使用相同的V[k]和V[k-1](以创建经内插的向量V[k]),可在编码器和解码器处使用向量的经量化/经解量化版本。空间-时间内插单元50可将经内插的nFG信号49'输出到心理声学音频译码器单元46且将经内插的前景V[k]向量51<sub>k</sub>输出到系数减少单元46。

[0070] 系数减少单元46可表示经配置以基于背景信道信息43对于剩余的前景V[k]向量53执行系数减少以将经减少的前景V[k]向量55输出到量化单元52的单元。经减少的前景V[k]向量55可具有维度D: [(N+1)<sup>2</sup> - (N<sub>BG</sub>+1)<sup>2</sup> - BG<sub>TOT</sub>] × nFG。系数减少单元46可在这方面表示经配置以减少剩余的前景V[k]向量53中的系数的数目的单元。换句话说,系数减少单元46可表示经配置以消除(形成剩余的前景V[k]向量53的)前景V[k]向量中的具有极少或几乎不具有方向信息的系数的单元。在一些实例中,相异或(换句话说)前景V[k]向量的对应于一阶和零阶基底函数的系数(其可标示为N<sub>BG</sub>)提供极少方向信息,且因此可从前景V向量移除(通过可被称为“系数减少”的过程)。在此实例中,可提供较大灵活性以不仅从集合[(N<sub>BG</sub>+1)<sup>2</sup> + 1, (N+1)<sup>2</sup>]标识对应于N<sub>BG</sub>的系数而且标识额外HOA信道(其可由变量TotalOfAddAmbHOAChan标示)。

[0071] 量化单元52可表示经配置以执行任何形式的量化以压缩减少的前景V[k]向量55以产生经译码前景V[k]向量57,从而将经译码前景V[k]向量57输出到位流产生单元42的单元。在操作中,量化单元52可表示经配置以压缩声场的空间分量(即,在此实例中为经减少的前景V[k]向量55中的一或者者)的单元。量化单元52可执行如由标示为“NbtsQ”的量化模式语法元素指示的以下12种量化模式中的任一者:

- [0072] NbtsQ值 量化模式的类型
- [0073] 0-3: 保留
- [0074] 4: 向量量化
- [0075] 5: 无霍夫曼译码的标量量化
- [0076] 6: 具有霍夫曼译码的6位标量量化
- [0077] 7: 具有霍夫曼译码的7位标量量化
- [0078] 8: 具有霍夫曼译码的8位标量量化
- [0079] ...
- [0080] 16: 具有霍夫曼译码的16位标量量化

[0081] 量化单元52还可执行前述类型的量化模式中的任一者的经预测版本,其中确定前一帧的V向量的元素(或执行向量量化时的权重)与当前帧的V向量的元素(或执行向量量化时的权重)之间的差。量化单元52接着可量化当前帧与前一帧的元素或权重之间的差而非当前帧本身的V向量的元素的值。

[0082] 量化单元52可对于经减少的前景V[k]向量55中的每一者执行多种形式的量化,以获得经减少的前景V[k]向量55的多个经译码版本。量化单元52可选择经减少的前景V[k]向量55的经译码版本中的一者作为经译码前景V[k]向量57。换句话说,量化单元52可基于本发明中论述的准则的任何组合来选择未经预测的经向量量化的V向量、经预测的经向量量化的V向量、未经霍夫曼译码的经标量量化的V向量以及经霍夫曼译码的经标量量化的V向量中的一者,以用作输出的经切换量化的V向量。在一些实例中,量化单元52可从包含向量量化模式和一或多个标量量化模式的量化模式集合中选择量化模式,且基于(或根据)选定模式量化输入的V向量。量化单元52接着可将以下各者中的选定者提供到位流产生单元52以用作经译码前景V[k]向量57:未经预测的经向量量化的V向量(例如,就权重值或指示权重值的位来说)、经预测的经向量量化的V向量(例如,就误差值或指示误差值的位来说)、未经霍夫曼译码的经标量量化的V向量以及经霍夫曼译码的经标量量化的V向量。量化单元52还可提供指示量化模式的语法元素(例如,NbitsQ语法元素)和用以将V向量解量化或以其它方式重构V向量的任何其它语法元素。

[0083] 包含于音频编码装置20内的去相关单元40'可表示经配置以将一或多个去相关变换应用于HOA系数47'以获得经去相关的HOA系数47"的单元的单个或多个例子。在一些实例中,去相关单元40'可将UHJ矩阵应用于HOA系数47'。在本发明的各种例子中,UHJ矩阵还可被称作“基于相位的变换”。应用基于相位的变换在本文中也可被称作“相移去相关”。

[0084] 立体混响UHJ格式是经设计以与单声道和立体声媒体兼容的立体混响环绕声系统的发展。UHJ格式包含其中将以根据可用的信道变化的准确性程度重现所记录的声场的系统层次。在各种例子中,UHJ也被称作“C格式”。所述缩写指示并入到所述系统中的来源中的一些:来自通用的U(UD-4);来自矩阵H的H;和来自系统45J的J。

[0085] UHJ是编码和解码立体混响技术内的方向性声音信息的分层系统。取决于可用的信道的数目,系统可携带更多或更少信息。UHJ是立体声和单声道完全兼容的。可使用高达四个信道(L,R,T,Q)。

[0086] 在一种形式中,2信道(L,R)UHJ、水平(或“平面”)环绕信息可由正交立体声信号信道(CD、FM或数字无线电等)携载,所述信息可在收听端使用UHJ解码器进行恢复。将两个信道求和可产生兼容的单声道信号,其与对常规“经假立体声录音的(panpotted)单声道”源相比可为对两信道版本的更准确表示。如果第三信道(T)可用,那么当经由3信道UHJ解码器进行解码时,第三信道可用以产生对平面环绕效应的改进的定位准确性。第三信道为此目的可能并非不需要具有全音频带宽,从而导致所谓的“21/2信道”系统的可能性,其中第三信道在带宽上受限。在一个实例中,所述限值可为5kHz。第三信道可经由FM无线电例如借助于相位正交调制进行广播。将第四信道(Q)添加到UHJ系统可允许以高度n(有时被称为多声道(Periphony))对全环绕声音进行编码,其中准确性程度与4信道B格式相同。

[0087] 2信道UHJ是通常用于立体混响记录的分配的格式。2信道UHJ记录可经由所有正交立体声信道传输,且可使用正交2信道媒体中的任一者而无需更改。UHJ是立体声兼容的,因

为在无需解码的情况下,收听者可察觉立体声像,但其与常规立体声(例如,所谓的“超立体声”)相比显著更宽。也可将左信道与右信道求和以用于极高程度的单声道兼容性。经由UHJ解码器回放,可展现环绕能力。

[0088] 应用UHJ矩阵(或基于相位的变换)的去相关单元40’的实例数学表示如下:

[0089] UHJ编码:

[0090]  $S = (0.9397*W) + (0.1856*X)$ ;

[0091]  $D = \text{imag}(\text{hilbert}((-0.3420*W) + (0.5099*X))) + (0.6555*Y)$ ;

[0092]  $T = \text{imag}(\text{hilbert}((-0.1432*W) + (0.6512*X))) - (0.7071*Y)$ ;

[0093]  $Q = 0.9772*Z$ ;

[0094] S和D到左和右的转换:

[0095] 左 =  $(S+D)/2$

[0096] 右 =  $(S-D)/2$

[0097] 根据以上计算的一些实施方案,关于以上计算的假设可包含以下各项:HOA背景信道是1阶立体混响,FuMa经归一化,按照立体混响信道编号次序W(a00)、X(a11)、Y(a11-)、Z(a10)。

[0098] 在上文所列的计算中,去相关单元40’可执行各种矩阵与恒定值的标量乘法。举例来说,为获得S信号,去相关单元40’可执行W矩阵与恒定值0.9397(例如,通过标量乘法)以及X矩阵与恒定值0.1856的标量乘法。还如在上文所列的计算中所说明,去相关单元40’可在获得D和T信号中的每一者时应用希尔伯特变换(由以上UHJ编码中的“`Hilbert()`”函数标示)。以上UHJ编码中的“`imag()`”函数指示获得希尔伯特变换的结果的虚数(在数学意义上)。

[0099] 应用UHJ矩阵(或基于相位的变换)的去相关单元40’的另一实例数学表示如下:

[0100] UHJ编码:

[0101]  $S = (0.9396926*W) + (0.151520536509082*X)$ ;

[0102]  $D = \text{imag}(\text{hilbert}((-0.3420201*W) + (0.416299273350443*X))) + (0.535173990363608*Y)$ ;

[0103]  $T = 0.940604061228740 * (\text{imag}(\text{hilbert}((-0.1432*W) + (0.531702573500135*X))) - (0.577350269189626*Y))$ ;

[0104]  $Q = Z$ ;

[0105] S和D到左和右的转换:

[0106] 左 =  $(S+D)/2$ ;

[0107] 右 =  $(S-D)/2$ ;

[0108] 在以上计算的一些实例实施方案中,关于以上计算的假设可包含以下各项:HOA背景信道是1阶立体混响,N3D(或“全三维”)经归一化,按照立体混响信道编号次序W(a00)、X(a11)、Y(a11-)、Z(a10)。尽管本文中关于N3D归一化进行描述,但应了解,所述实例计算也可应用于经SN3D归一化(或“经施密特半归一化”)的HOA背景信道。N3D与SN3D归一化可在所使用的比例缩放因子方面不同。N3D归一化相对于SN3D归一化的实例表示如下表达为:

$$[0109] N_{l,m}^{N3D} = N_{l,m}^{SN3D} \sqrt{2l+1}$$

[0110] 在SN3D归一化中所使用的加权系数的实例如下表达为:

$$[0111] N_{l,m}^{SN3D} = \sqrt{\frac{2 - \delta m (l - |m|)!}{4\pi (l + |m|)!}}, \delta m \begin{cases} 1, & \text{如果 } m = 0 \\ 0, & \text{如果 } m \neq 0 \end{cases}$$

[0112] 在上文所列的计算中,去相关单元40'可执行各种矩阵与恒定值的标量乘法。举例来说,为获得S信号,去相关单元40'可执行W矩阵与恒定值0.9396926(例如,通过标量乘法)以及X矩阵与恒定值0.151520536509082的标量乘法。还如在上文所列的计算中所说明,去相关单元40'可在获得D和T信号中的每一者时应用希尔伯特变换(由以上UHJ编码中的“Hilbert()”函数或相移去相关标示)。以上UHJ编码中的“imag()”函数指示获得希尔伯特变换的结果的虚数(在数学意义上)。

[0113] 去相关单元40'可执行上文所列的计算,使得所得的S信号和D信号表示左音频信号和右音频信号(或换句话说,立体声音频信号)。在一些此类情境中,去相关单元40'可输出T信号和Q信号作为经去相关的HOA系数47"的一部分,但当T信号和Q信号再现给立体声扬声器几何布置(或换句话说,立体声扬声器配置)时,接收位流21的解码装置可不处理所述T信号和Q信号。在实例中,HOA系数47'可表示将在单声道音频再现系统上再现的声场。去相关单元40'可输出S信号和D信号作为经去相关的HOA系数47"的一部分,且接收位流21的解码装置可组合(或“混合”)S信号和D信号以形成将以单声道音频格式再现和/或输出的音频信号。在这些实例中,解码装置和/或再现装置可以各种方式恢复单声道音频信号。一个实例是通过混合左信号与右信号(由S信号和D信号表示)。另一实例是通过应用UHJ矩阵(或基于相位的变换)以解码W信号(在下文对于图5更详细地论述)。通过应用UHJ矩阵(或基于相位的变换)产生呈S信号和D信号形式的固有左信号和固有右信号,去相关单元40'可实施本发明的技术以与应用其它去相关变换(例如在MPEG-H标准中描述的模式矩阵)的技术相比提供潜在优点和/或潜在改进。

[0114] 在各种实例中,去相关单元40'可基于所接收的HOA系数47'的位速率,应用不同的去相关变换。举例来说,在其中HOA系数47'表示四信道输入的情境中,去相关单元40'可应用上文所描述的UHJ矩阵(或基于相位的变换)。更具体来说,基于HOA系数47'表示四信道输入,去相关单元40'可应用 $4 \times 4$ UHJ矩阵(或基于相位的变换)。举例来说, $4 \times 4$ 矩阵可正交于HOA系数47'的四信道输入。换句话说,在HOA系数47'表示更少数目个信道(例如,四个)的例子中,去相关单元40'可应用UHJ矩阵作为选定去相关变换,以将HOA信号47'的背景信号去相关以获得经去相关的HOA系数47"。

[0115] 根据此实例,如果HOA系数47'表示更大数目个信道(例如,九个),那么去相关单元40'可应用不同于UHJ矩阵(或基于相位的变换)的去相关变换。举例来说,在其中HOA系数47'表示九信道输入的情境中,去相关单元40'可应用模式矩阵(例如,如在MPEG-H标准中所描述),以将HOA系数47'去相关。在其中HOA系数47'表示九信道输入的实例中,去相关单元40'可应用 $9 \times 9$ 模式矩阵以获得经去相关的HOA系数47"。

[0116] 继而,音频编码装置20的各个组件(例如心理声学音频译码器40)可根据AAC或USAC对经去相关的HOA系数47"以感知方式进行译码。去相关单元40'可应用相移去相关变换(例如,在四信道输入的情况下,为UHJ矩阵或基于相位的变换),以优化针对HOA的AAC/USAC译码。在其中HOA系数47'(以及借此,经去相关的HOA系数47")表示将在立体声再现系

统上再现的音频数据的实例中,去相关单元40'可应用本发明的技术以基于AAC和USAC是经相对地定向的立体声音频数据(或针对其经优化)而改进或优化压缩。

[0117] 将理解,在其中经能量补偿的HOA系数47'包含前景信道的情境中,以及在其中经能量补偿的HOA系数47'不包含任何前景信道的情境中,去相关单元40'可应用本文中所描述的技术。作为一个实例,在其中经能量补偿的HOA系数47'包含零(0)个前景信道和四个(4)背景信道的情境(例如,更低/更小位速率的情境)中,去相关单元40'可应用上文所描述的技术和/或计算。

[0118] 在一些实例中,去相关单元40'可致使位流产生单元42用信号发送指示去相关单元40'将去相关变换应用于HOA系数47'的一或多个语法元素作为基于向量的位流21的一部分。通过将此指示提供到解码装置,去相关单元40'可使得解码装置能够对HOA域中的音频数据执行互逆去相关变换。在一些实例中,去相关单元40'可致使位流产生单元42用信号发送指示应用哪一去相关变换(例如UHJ矩阵(或其它基于相位的变换)或模式矩阵)的语法元素。

[0119] 去相关单元40'可将基于相位的变换应用于能量补偿环境HOA系数47'。用于 $C_{AMB}(k-1)$ 的第一 $0_{MIN}$  HOA系数序列的基于相位的变换定义如下

$$[0120] \begin{bmatrix} x_{AMB,LOW,1}(k-2) \\ x_{AMB,LOW,2}(k-2) \\ x_{AMB,LOW,3}(k-2) \\ x_{AMB,LOW,4}(k-2) \end{bmatrix} = \begin{bmatrix} d(9) \cdot (S(k-2) + M(k-2)) \\ d(9) \cdot (M(k-2) - S(k-2)) \\ d(8) \cdot (B_{+90}(k-2) + d(5) \cdot c_{AMB,2}(k-2)) \\ c_{AMB,3}(k-2) \end{bmatrix},$$

[0121] 其中系数d如在表1中定义,信号帧S(k-2)和M(k-2)定义如下

$$[0122] S(k-2) = A_{+90}(k-2) + d(6) \cdot c_{AMB,2}(k-2)$$

$$[0123] M(k-2) = d(4) \cdot c_{AMB,1}(k-2) + d(5) \cdot c_{AMB,4}(k-2)$$

[0124] 且 $A_{+90}(k-2)$ 和 $B_{+90}(k-2)$ 是+90度相移信号A和B的帧,定义如下

$$[0125] A(k-2) = d(0) \cdot c_{AMB,LOW,1}(k-2) + d(1) \cdot c_{AMB,4}(k-2)$$

$$B(k-2) = d(2) \cdot c_{AMB,LOW,1}(k-2) + d(3) \cdot c_{AMB,4}(k-2)^\circ$$

[0126] 因此定义针对 $C_{P,AMB}(k-1)$ 的第一 $0_{MIN}$  HOA系数序列的基于相位的变换。所描述的变换可引入一个帧的延迟。

[0127] 在上文中, $x_{AMB,LOW,1}(k-2)$ 到 $x_{AMB,LOW,4}(k-2)$ 可对应于经去相关的环境HOA系数47''。在上述方程式中,变化的 $C_{AMB,1}(k)$ 变量标示对应于具有(阶数:子阶数)为(0:0)的球形基底函数的第k帧的HOA系数,其还可被称作‘W’信道或分量。变化的 $C_{AMB,2}(k)$ 变量标示对应于具有(阶数:子阶数)为(1:-1)的球形基底函数的第k帧的HOA系数,其还可被称作‘Y’信道或分量。变化的 $C_{AMB,3}(k)$ 变量标示对应于具有(阶数:子阶数)为(1:0)的球形基底函数的第k帧的HOA系数,其还可被称作‘Z’信道或分量。变化的 $C_{AMB,4}(k)$ 变量标示对应于具有(阶数:子阶数)为(1:1)的球形基底函数的第k帧的HOA系数,其还可被称作‘X’信道或分量。 $C_{AMB,1}(k)$ 到 $C_{AMB,3}(k)$ 可对应于环境HOA系数47'。

[0128] 下文的表1说明可由去相关单元40使用以执行基于相位的变换的系数的实例。

n	d(n)
0	0.3420200999999999

1	0.41629927335044281
2	0.1431999999999999
3	0.53170257350013528
4	0.9396925999999999
5	0.15152053650908184
6	0.53517399036360758
7	0.57735026918962584
8	0.94060406122874030
9	0.5000000000000000

[0130] 表1用于基于相位的变换的系数

[0131] 在一些实例中,音频编码装置20的各个组件(例如位流产生单元42)可经配置以仅传输针对较低目标位速率(例如,128K或256K的目标位速率)的一阶HOA表示。根据一些此类实例,音频编码装置20(或其组件,例如位流产生单元42)可经配置以丢弃高阶HOA系数(例如,具有大于一阶(或换句话说,N>1)的阶数的系数)。然而,在其中音频编码装置20确定目标位速率相对较高的实例中,音频编码装置20(例如,位流产生单元42)可分离前景信道与背景信道,且可分配位(例如,以较大量)给前景信道。

[0132] 包含于音频编码装置20内的心理声学音频译码器单元40可表示心理声学音频译码器的多个例子,所述例子中的每一者用以编码经去相关的HOA系数47''和经内插的nFG信号49'中的每一者的不同音频对象或HOA信道以产生经编码的环境HOA系数59和经编码的nFG信号61。心理声学音频译码器单元40可将经编码的环境HOA系数59和经编码的nFG信号61输出到位流产生单元42。

[0133] 包含于音频编码装置20内的位流产生单元42表示将数据格式化以符合已知格式(可指解码装置已知的格式),借此产生基于向量的位流21的单元。换句话说,位流21可表示已按上文所描述的方式进行编码的经编码音频数据。在一些实例中,位流产生单元42可表示可接收经译码前景V[k]向量57、经编码环境HOA系数59、经编码nFG信号61和背景信道信息43的多路复用器。位流产生单元42接着可基于经译码前景V[k]向量57、经编码环境HOA系数59、经编码nFG信号61和背景信道信息43产生位流21。以此方式,位流产生单元42可借此规定位流中21的向量57以获得位流21。位流21可包含主要或主位流以及一或多个旁侧信道位流。

[0134] 尽管未在图3的实例中示出,但音频编码装置20还可包含位流输出单元,所述位流输出单元基于是将使用基于方向的合成还是基于向量的合成对当前帧进行编码而切换从音频编码装置20输出的位流(例如,在基于方向的位流21与基于向量的位流21之间切换)。位流输出单元可基于由内容分析单元26输出的指示执行基于方向的合成(作为检测到HOA系数11是从合成音频对象产生的结果)还是执行基于向量的合成(作为检测到HOA系数经记录的结果)的语法元素执行所述切换。位流输出单元可指定正确的标头语法以指示用于当前帧以及位流21中的相应者的切换或当前编码。

[0135] 此外,如上文所提到,声场分析单元44可标识BG<sub>TOT</sub>环境HOA系数47,所述系数可逐帧改变(但有时BG<sub>TOT</sub>可跨越两个或更多个邻近(在时间上)帧保持恒定或相同)。BG<sub>TOT</sub>的改变可导致在经减少前景V[k]向量55中表达的系数的改变。BG<sub>TOT</sub>的改变可导致背景HOA系数(其

也可被称作“环境HOA系数”)逐帧改变(但再次,BG<sub>TOT</sub>有时可跨越两个或更多个邻近(在时间上)帧保持恒定或相同)。所述改变通常导致声场的各方面的能量改变,所述能量改变由额外环境HOA系数的添加或移除以及系数从减少的前景V[k]向量55的对应移除或系数到减少的前景V[k]向量55的添加来表示。

[0136] 因此,声场分析单元44可进一步确定环境HOA系数何时从帧到帧改变,并且产生指示环境HOA系数在用于表示声场的环境分量方面的改变的标志或其它语法元素(其中所述改变也可被称作环境HOA系数的“转变”或环境HOA系数的“转变”)。特定来说,系数减少单元46可产生标志(其可表示为AmbCoeffTransition标志或AmbCoeffIdxTransition标志),从而将所述标志提供到位流产生单元42,使得可将所述标志包含在位流21中(可能地作为旁侧信道信息的部分)。

[0137] 除指定环境系数转变标志之外,系数减少单元46还可修改产生减少的前景V[k]向量55的方式。在一个实例中,在确定环境HOA环境系数中的一者在当前帧期间处于转变中时,系数减少单元46可指定减少的前景V[k]向量55的V向量中的每一者的向量系数(其也可被称作“向量元素”或“元素”),其对应于处于转变中的环境HOA系数。此外,处于转变中的环境HOA系数可添加到背景系数的BG<sub>TOT</sub>总数目或从背景系数的BG<sub>TOT</sub>总数目中移除。因此,背景系数的总数目的所得改变影响环境HOA系数是否包含于位流中,以及在上文所描述的第二和第三配置模式中是否针对位流中所指定的V向量包含V向量的对应元素。关于系数减少单元46可如何规定减少的前景V[k]向量55以克服能量改变的更多信息提供于2015年1月12日申请的标题为“环境高阶立体混响系数的转变 (TRANSITIONING OF AMBIENT HIGHER-ORDER AMBISONIC COEFFICIENTS)”的第14/594,533号美国申请案中。

[0138] 因此,音频编码装置20可表示用于压缩音频的装置的实例,所述装置经配置以将去相关变换应用于环境立体混响系数以获得环境立体混响系数的经去相关表示,环境HOA系数已从多个高阶立体混响系数中提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量,其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联。在一些实例中,为了应用去相关变换,所述装置经配置以将UHJ矩阵应用于环境立体混响系数。

[0139] 在一些实例中,所述装置经进一步配置以根据N3D(全三维)归一化对UHJ矩阵进行归一化。在一些实例中,所述装置经进一步配置以根据SN3D归一化(施密特半归一化)对UHJ矩阵进行归一化。在一些实例中,环境立体混响系数是与具有阶数零或阶数一的球形基底函数相关联,且为了将UHJ矩阵应用于环境立体混响系数,所述装置经配置以对于环境立体混响系数的至少一个子集执行UHJ矩阵的标量乘法。在一些实例中,为了应用去相关变换,所述装置经配置以将模式矩阵应用于环境立体混响系数。

[0140] 根据一些实例,为了应用去相关变换,所述装置经配置以从经去相关的环境立体混响系数获得左信号和右信号。根据一些实例,所述装置经进一步配置以用信号发送经去相关的环境立体混响系数以及一或多个前景信道。根据一些实例,为了用信号发送经去相关的环境立体混响系数以及一或多个前景信道,所述装置经配置以响应于确定目标位速率符合或超过预定阈值而用信号发送经去相关的环境立体混响系数以及一或多个前景信道。

[0141] 在一些实例中,所述装置经进一步配置以在不用信号发送任何前景信道的情况下用信号发送经去相关的环境立体混响系数。在一些实例中,为了在不用信号发送任何前景

信道的情况下用信号发送经去相关的环境立体混响系数,所述装置经配置以响应于确定目标位速率低于预定阈值而在不用信号发送任何前景信道的情况下用信号发送经去相关的环境立体混响系数。在一些实例中,所述装置经进一步配置以用信号发送对去相关变换已应用于环境立体混响系数的指示。在一些实例中,所述装置进一步包含经配置以捕获将被压缩的音频数据的麦克风阵列。

[0142] 图4是更详细地说明图2的音频解码装置24的框图。如在图4的实例中所展示,音频解码装置24可包含提取单元72、基于方向的重构单元90、基于向量的重构单元92以及再相关单元81。

[0143] 尽管下文进行描述,但关于音频解码装置24和解压缩或以其它方式解码HOA系数的各种方面的更多信息可在2014年5月29日申请的标题为“用于声场的经分解表示的内插(INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD)”的第WO 2014/194099号国际专利申请公开案中获得。

[0144] 提取单元72可表示经配置以接收位流21并提取HOA系数11的各种经编码版本(例如,基于方向的经编码版本或基于向量的经编码版本)的单元。提取单元72可从以上所述确定指示HOA系数11是经由各种基于方向的版本还是基于向量的版本经编码的语法元素。当执行基于方向的编码时,提取单元72可提取HOA系数11的基于方向的版本和与所述经编码版本相关联的语法元素(其在图4的实例中表示为基于方向的信息91),从而将基于方向的信息91传递到基于方向的重构单元90。基于方向的重构单元90可表示经配置以基于所述基于方向的信息91重构呈HOA系数11'形式的HOA系数的单元。下文描述位流内的位流和语法元素的布置。

[0145] 当语法元素指示HOA系数11使用基于向量的合成进行编码时,提取单元72可提取经译码前景V[k]向量57(其可包含经译码权重57和/或索引63或经标量量化的V向量)、经编码环境HOA系数59和对应的音频对象61(其也可被称作经编码nFG信号61)。音频对象61各自对应于向量57中的一者。提取单元72可将经译码前景V[k]向量57传递到V向量重构单元74,且将经编码环境HOA系数59以及经编码nFG信号61提供到心理声学解码单元80。

[0146] V向量重构单元74可表示经配置以从经编码前景V[k]向量57重构V向量的单元。V向量重构单元74可以与量化单元52互逆的方式操作。

[0147] 心理声学解码单元80可以与图3的实例中所展示的心理声学音频译码器单元40互逆的方式操作,以便对经编码环境HOA系数59和经编码nFG信号61进行解码且借此产生经能量补偿的环境HOA系数47'和经内插的nFG信号49'(其也可被称作经内插的nFG音频对象49')。心理声学解码单元80可将经能量补偿的环境HOA系数47'传递到再相关单元81且将nFG信号49'传递到前景制订单元78。继而,再相关单元81可将一或多个再相关变换应用于经能量补偿的环境HOA系数47'以获得一或多个再相关的HOA系数47''(或相关的HOA系数47''),并且可将相关的HOA系数47''传递到HOA系数制订单元82(任选地,通过淡化单元770)。

[0148] 类似于以上的描述,相对于音频编码装置20的去相关单元40',再相关单元81可实施本发明的技术以减少经能量补偿的环境HOA系数47'的背景信道之间的相关性,从而减少或减低噪声去掩蔽。在其中再相关单元81应用UHJ矩阵(例如,逆UHJ矩阵)作为选定再相关变换的实例中,再相关单元81可改进压缩率并且通过减少数据处理操作而节省计算资源。在一些实例中,基于向量的位流21可包含指示在编码期间应用去相关变换的一或多个语法

元素。在基于向量的位流21中包含此类语法元素可使再相关单元81能够对经能量补偿的HOA系数47'执行互逆去相关(例如,相关或再相关)变换。在一些实例中,信号语法元素可指示应用哪一去相关变换,例如UHJ矩阵或模式矩阵,借此使得再相关单元81能够选择适当再相关变换应用于经能量补偿的HOA系数47'。

[0149] 在其中基于向量的重构单元92将HOA系数11'输出到包括立体声系统的再现系统的实例中,再相关单元81可处理S信号和D信号(例如,固有左信号和固有右信号)以产生再相关的HOA系数47"。举例来说,因为S信号和D信号表示固有左信号和固有右信号,所以再现系统可使用S信号和D信号作为两个立体声输出流。在其中重构单元92将HOA系数11'输出到包括单声道音频系统的再现系统的实例中,再现系统可组合或混合S信号与D信号(如在HOA系数11'中表示)以获得单声道音频输出以用于回放。在单声道音频系统的实例中,再现系统可将经混合的单声道音频输出添加到一或多个前景信道(在存在任何前景信道的情况下)以产生音频输出。

[0150] 相对于一些现有的具有UHJ能力的编码器,以相位振幅矩阵处理信号以恢复类似于B格式的信号集合。在大多数情况下,所述信号将实际上为B格式,但在2信道UHJ的情况下,没有充足的信息可供用以能够重构正确的B格式信号,而是呈现类似于B格式信号的特性的信号。所述信息接着经由雪夫型(Shelf)滤波器集合传递到产生扬声器馈送的振幅矩阵,所述雪夫型(Shelf)滤波器集合改进解码器在较小收听环境(其在较大规模的应用中可被省略)中的准确性和性能。立体混响经设计以符合实际房间(例如,起居室)和实用扬声器位置的要求:很多此类房间是矩形的,因此基础系统经设计以对去往矩形中的四个扩音器进行解码,其中侧变长度介于1:2(宽度是长度的两倍)与2:1(长度是宽度的两倍)之间,因此符合大部分此类房间的要求。通常提供布局控制以允许解码器针对扩音器位置经配置。布局控制是不同于其它环绕声系统的立体混响回放的方面:解码器可针对扬声器阵列的大小和布局经具体配置。布局控制可呈旋纽、2路(1:2、2:1)或3路(1:2、1:1、2:1)开关的形式。四个扬声器是水平环绕解码所需的最小值,且虽然四个扬声器布局可适用于数种收听环境,但较大空间可需要更多扬声器以给出全环绕定位。

[0151] 再相关单元81可对于应用UHJ矩阵(例如,逆UHJ矩阵或基于相位的逆变换)作为再相关变换执行的计算的实例列出如下:

[0152] UHJ解码:

[0153] 左和右到S和D的转换:

[0154]  $S = \text{左} + \text{右}$

[0155]  $D = \text{左} - \text{右}$

[0156]  $W = (0.982*S) + 0.197.*\text{imag}(\text{hilbert}((0.828*D) + (0.768*T))) ;$

[0157]  $X = (0.419*S) - \text{imag}(\text{hilbert}((0.828*D) + (0.768*T))) ;$

[0158]  $Y = (0.796*D) - 0.676*T + \text{imag}(\text{hilbert}(0.187*S)) ;$

[0159]  $Z = (1.023*Q) ;$

[0160] 在以上计算的一些实例实施方案中,关于以上计算的假设可包含以下各项:HOA背景信道是1阶立体混响,FuMa经归一化,按照立体混响信道编号次序W(a00)、X(a11)、Y(a11-)、Z(a10)。

[0161] 再相关单元81可对于应用UHJ矩阵(或基于相位的逆变换)作为再相关变换执行的

计算的实例列出如下：

- [0162] UHJ解码；
- [0163] 左和右到S和D的转换；
- [0164] 左和右到S和D的转换；
- [0165]  $S = \text{左} + \text{右}$ ；
- [0166]  $D = \text{左} - \text{右}$ ；
- [0167]  $h1 = \text{imag}(\text{hilbert}(1.014088753512236*D+T))$ ；
- [0168]  $h2 = \text{imag}(\text{hilbert}(0.229027290950227*S))$ ；
- [0169]  $W = 0.982*S + 0.160849826442762*h1$ ；
- [0170]  $X = 0.513168101113076*S - h1$ ；
- [0171]  $Y = 0.974896917627705*D - 0.880208333333333*T + h2$ ；
- [0172]  $Z = Q$ ；
- [0173] 在以上计算的一些实施方案中,关于以上计算的假设可包含以下各项:HOA背景信道是1阶立体混响,N3D(或“全三维”)经归一化,按照立体混响信道编号次序W(a00)、X(a11)、Y(a11-)、Z(a10)。尽管本文中关于N3D归一化进行描述,但应了解,所述实例计算也可应用于经SN3D归一化(或“经施密特半归一化”的HOA背景信道。如上文关于图4所描述,N3D与SN3D归一化可在所使用的比例缩放因子方面不同。上文关于图4描述在N3D归一化中使用的比例缩放因子的实例表示。上文关于图4描述在SN3D归一化中使用的加权系数的实例表示。
- [0174] 在一些实例中,经能量补偿的HOA系数47'可表示仅水平布局,例如不包含任何垂直信道的音频数据。在这些实例中,再相关单元81可不对于以上的Z信号执行计算,因为Z信号表示垂直方向音频数据。替代地,在这些实例中,再相关单元81可仅对W、X和Y信号执行以上计算,因为W、X和Y信号表示水平方向数据。在其中经能量补偿的HOA系数47'表示将在单声道音频再现系统上再现的音频数据的一些实例中,再相关单元81可仅从以上计算得到W信号。更具体来说,因为所得W信号表示单声道音频数据,所以W信号可提供所必要的全部数据,其中经能量补偿的HOA系数47'表示将以单声道音频格式再现的数据,或其中再现系统包括单声道音频系统。
- [0175] 类似于如上文关于音频编码装置20的去相关单元40'所描述,在实例中,再相关单元81可在其中经能量补偿的HOA系数47'包含较少数目个背景信道的情境中应用UHJ矩阵(或逆UHJ矩阵或基于相位的逆变换),但可在经能量补偿的HOA系数47'包含较大数目个背景信道的情境中应用模式矩阵或逆模式矩阵(例如,如在MPEG-H标准中所描述)。
- [0176] 将理解,在其中经能量补偿的HOA系数47'包含前景信道的情境中,以及在其中经能量补偿的HOA系数47'不包含任何前景信道的情境中,再相关单元81可应用本文中所描述的技术。作为一个实例,在其中经能量补偿的HOA系数47'包含零(0)个前景信道和八(8)个背景信道的情境(例如,更低/更小位速率的情境)中,再相关单元81可应用上文所描述的技术和/或计算。
- [0177] 音频解码装置24的各个组件(例如再相关单元81)可为用以确定将两种处理方法中的哪一个应用于去相关的语法元素,例如标志UsePhaseShiftDecorr。在其中去相关单元40'将空间变换用于去相关的例子中,再相关单元81可确定UsePhaseShiftDecorr标志设置

为值零。

[0178] 在其中再相关单元81确定UsePhaseShiftDecorr标志设置为值一的情况下,再相关单元81可确定将使用基于相位的变换执行再相关。如果标志UsePhaseShiftDecorr具有值1,那么应用以下处理以重构环境HOA分量的前四个系数序列

$$[0179] \begin{bmatrix} \mathbf{c}_{\text{AMB},1}(k) \\ \mathbf{c}_{\text{AMB},2}(k) \\ \mathbf{c}_{\text{AMB},3}(k) \\ \mathbf{c}_{\text{AMB},4}(k) \end{bmatrix} = \begin{bmatrix} c(3) \cdot A_{+90}(k) + c(2) \cdot [\mathbf{c}_{\text{I,AMB},1}(k) + \mathbf{c}_{\text{I,AMB},2}(k)] \\ B_{+90}(k) + c(5) \cdot [\mathbf{c}_{\text{I,AMB},1}(k) - \mathbf{c}_{\text{I,AMB},2}(k)] + c(6) \cdot \mathbf{c}_{\text{I,AMB},3}(k) \\ \mathbf{c}_{\text{I,AMB},4}(k) \\ c(4) \cdot [\mathbf{c}_{\text{I,AMB},1}(k) + \mathbf{c}_{\text{I,AMB},2}(k)] - A_{+90}(k) \end{bmatrix}$$

[0180] 其中如在下文表1中定义的系数c以及 $A_{+90}(k)$ 和 $B_{+90}(k)$ 是+90度相移信号A和B的帧,定义如下

[0181]  $A(k) = c(0) \cdot [c_{\text{I,AMB},1}(k) - c_{\text{I,AMB},2}(k)]$ ,

[0182]  $B(k) = c(1) \cdot [c_{\text{I,AMB},1}(k) + c_{\text{I,AMB},2}(k)]$ 。

[0183] 下文的表2说明去相关单元40'可用以实施基于相位的变换的实例系数。

n	c(n)
0	1.0140887535122356
1	0.22902729095022714
2	0.9819999999999998
3	0.16084982644276205
4	0.51316810111307576
5	0.97489691762770481
6	-0.880208333333333337

[0185] 表2基于相位的变换的系数

[0186] 在上述方程式中,变化的 $C_{\text{AMB},1}(k)$ 变量标示对应于具有(阶数:子阶数)为(0:0)的球形基底函数的第k帧的HOA系数,其还可被称作‘W’信道或分量。变化的 $C_{\text{AMB},2}(k)$ 变量标示对应于具有(阶数:子阶数)为(1:-1)的球形基底函数的第k帧的HOA系数,其还可被称作‘Y’信道或分量。变化的 $C_{\text{AMB},3}(k)$ 变量标示对应于具有(阶数:子阶数)为(1:0)的球形基底函数的第k帧的HOA系数,其还可被称作‘Z’信道或分量。变化的 $C_{\text{AMB},4}(k)$ 变量标示对应于具有(阶数:子阶数)为(1:1)的球形基底函数的第k帧的HOA系数,其还可被称作‘X’信道或分量。 $C_{\text{AMB},1}(k)$ 到 $C_{\text{AMB},3}(k)$ 可对应于环境HOA系数47'。

[0187] 上文的记法 $[c_{\text{I,AMB},1}(k) + c_{\text{I,AMB},2}(k)]$ 标示可替代地称为‘S’的项,其等效于左信道加右信道。 $c_{\text{I,AMB},1}(k)$ 变量标示作为UHJ编码的结果产生的左信道,而 $c_{\text{I,AMB},2}(k)$ 变量标示作为UHJ编码的结果产生的右信道。下标‘I’记法标示对应信道已经与其它环境信道去相关(例如,通过应用UHJ矩阵或基于相位的变换)。 $[c_{\text{I,AMB},1}(k) - c_{\text{I,AMB},2}(k)]$ 记法标示在本发明通篇中被称为‘D’的项,其表示左信道减右信道。 $c_{\text{I,AMB},3}(k)$ 变量标示在本发明通篇中被称为变量‘T’的项。 $c_{\text{I,AMB},4}(k)$ 变量标示在本发明通篇中被称为变量‘Q’的项。

[0188]  $A_{+90}(k)$ 记法标示 $c(0)$ 乘以S的正90度相移(其还在本发明通篇中由变量‘h1’标示)。 $B_{+90}(k)$ 记法标示 $c(1)$ 乘以D的正90度相移(其还在本发明通篇中由变量‘h2’标示)。

[0189] 空间-时间内插单元76可以类似于上文关于空间-时间内插单元50所描述的方式

操作。空间-时间内插单元76可接收经减少的前景 $V[k]$ 向量 $55_k$ ，并且对于前景 $V[k]$ 向量 $55_k$ 和经减少的前景 $V[k-1]$ 向量 $55_{k-1}$ 执行空间-时间内插以产生经内插的前景 $V[k]$ 向量 $55_k''$ 。空间-时间内插单元76将经内插的前景 $V[k]$ 向量 $55_k''$ 转发到淡化单元770。

[0190] 提取单元72还可将指示环境HOA系数中的一者何时处于转变中的信号757输出到淡化单元770，所述淡化单元接着可确定SHC<sub>BG</sub> 47'（其中SHC<sub>BG</sub> 47'还可标示为“环境HOA信道47”或“环境HOA系数47”）和经内插的前景 $V[k]$ 向量 $55_k''$ 的元素中的何者将淡入或淡出。在一些实例中，淡化单元770可对于环境HOA系数47'和经内插的前景 $V[k]$ 向量 $55_k''$ 的元素中的每一者以相反方式操作。也就是说，淡化单元770可对于环境HOA系数47'中的对应环境HOA系数执行淡入或淡出或执行淡入或淡出两者，同时对于经内插的前景 $V[k]$ 向量 $55_k''$ 的元素中的对应元素执行淡入或淡出或执行淡入和淡出两者。淡化单元770可将经调整的环境HOA系数47''输出到HOA系数制订单元82且将经调整的前景 $V[k]$ 向量 $55_k'''$ 输出到前景制订单元78。在这方面，淡化单元770表示经配置以对于HOA系数或其衍生物（例如，呈环境HOA系数47'和经内插的前景 $V[k]$ 向量 $55_k''$ 的元素的形式）的各种方面执行淡化操作的单元。

[0191] 前景制订单元78可表示经配置以对于经调整的前景 $V[k]$ 向量 $55_k'''$ 和经内插的nFG信号49'执行矩阵乘法以产生前景HOA系数65的单元。在这方面，前景制订单元78可组合音频对象49'（其为借以表示经内插的nFG信号49'的另一种方式）与向量 $55_k'''$ 以重构HOA系数11'的前景（或换句话说，主导）方面。前景制订单元78可执行经内插的nFG信号49'与经调整的前景 $V[k]$ 向量 $55_k'''$ 的矩阵乘法。

[0192] HOA系数制订单元82可表示经配置以将前景HOA系数65与经调整的环境HOA系数47''组合以便获得HOA系数11'的单元。撇号记法反映HOA系数11'可与HOA系数11类似而非相同。HOA系数11与11'之间的差异可由归因于经由有损传输媒体的传输、量化或其它有损操作的损失引起。

[0193] UHJ是已经用以从一阶立体混响内容创建2信道立体声流的矩阵变换方法。UHJ在过去用以经由FM发射器发射立体声或仅水平环绕内容。然而，应了解，UHJ不限于在FM发射器中的使用。在MPEG-H HOA编码方案中，可用模式矩阵预处理HOA背景信道以将HOA背景信道转换成空间域中的正交点。接着经由USAC或AAC对经变换信道以感知方式进行译码。

[0194] 本发明的技术通常涉及在对HOA背景信道进行译码的应用中使用UHJ变换（或基于相位的变换）而非使用此模式矩阵。两种方法（(1) 经由模式矩阵到空间域中的变换，(2) UHJ变换）通常均涉及减少HOA背景信道之间的的相关性，所述相关性可引起经解码声场内的噪声去掩蔽的（潜在地非所要的）效应。

[0195] 因此，在实例中，音频解码装置24可表示经配置以进行以下操作的装置：获得具有至少一左信号和右信号的环境立体混响系数的经去相关表示，所述环境立体混响系数已从多个高阶立体混响系数提取并且表示由所述多个高阶立体混响系数描述的声场的背景分量，其中所述多个高阶立体混响系数中的至少一者与具有大于一的阶数的球形基底函数相关联；和基于所述环境立体混响系数的经去相关表示而产生扬声器馈送。在一些实例中，所述装置经进一步配置以将再相关变换应用于环境立体混响系数的经去相关表示以获得多个相关的环境立体混响系数。

[0196] 在一些实例中，为了应用再相关变换，所述装置经配置以将逆UHJ矩阵（或基于相

位的变换)应用于环境立体混响系数。根据一些实例,逆UHJ矩阵(或基于相位的逆变换)已根据N3D(全三维)归一化经归一化。根据一些实例,逆UHJ矩阵(或基于相位的逆变换)已根据SN3D归一化(施密特半归一化)经归一化。

[0197] 根据一些实例,环境立体混响系数与具有阶数零或阶数一的球形基底函数相关联,且为了应用逆UHJ矩阵(或基于相位的逆变换),所述装置经配置以对于环境立体混响系数的经去相关表示执行UHJ矩阵的标量乘法。在一些实例中,为了应用再相关变换,所述装置经配置以将逆模式矩阵应用于环境立体混响系数的经去相关表示。在一些实例中,为了产生扬声器馈送,所述装置经配置以基于左信号产生左扬声器馈送且基于右信号产生右扬声器馈送,所述左扬声器馈送和扬声器馈送通过立体声再现系统输出。

[0198] 在一些实例中,为了产生扬声器馈送,所述装置经配置以在不将再相关变换应用于所述右信号和左信号的情况下,使用左信号作为左扬声器馈送且使用右信号作为右扬声器馈送。根据一些实例,为了产生扬声器馈送,所述装置经配置以混合左信号与右信号以用于由单声道音频系统输出。根据一些实例,为了产生扬声器馈送,所述装置经配置以组合相关的环境立体混响系数与一或多个前景信道。

[0199] 根据一些实例,所述装置经进一步配置以确定没有前景信道可用于与相关的环境立体混响系数组合。在一些实例中,所述装置经进一步配置以确定将经由单声道音频再现系统输出声场,以及对包含用于由单声道音频再现系统输出的数据的经去相关的高阶立体混响系数的至少一个子集进行解码。在一些实例中,所述装置经进一步配置以获得对环境立体混响系数的经去相关表示是通过对去相关变换经去相关的指示。根据一些实例,所述装置进一步包含经配置以输出基于环境立体混响系数的经去相关表示产生的扬声器馈送的扩音器阵列。

[0200] 图5是说明音频编码装置(例如在图3的实例中展示的音频编码装置20)执行本发明中描述的基于向量的合成技术的各种方面的示范性操作的流程图。最初,音频编码装置20接收HOA系数11(106)。音频编码装置20可调用LIT单元30,其可对于HOA系数应用LIT以输出经变换的HOA系数(例如,在SVD的情况下,经变换的HOA系数可包括US[k]向量33和V[k]向量35)(107)。

[0201] 音频编码装置20接下来可调用参数计算单元32以按上文所描述的方式对于US[k]向量33、US[k-1]向量33、V[k]和/或V[k-1]向量35的任何组合执行上文所描述的分析以标识各种参数。也就是说,参数计算单元32可基于对经变换的HOA系数33/35的分析确定至少一个参数(108)。

[0202] 音频编码装置20接着可调用重排序单元34,所述重排序单元基于所述参数将经变换的HOA系数(再次在SVD的上下文中,其可指US[k]向量33和V[k]向量35)重排序,以产生经重排序的经变换HOA系数33'/35'(或换句话说,US[k]向量33'和V[k]向量35'),如上文所描述(109)。音频编码装置20可在前述操作或后续操作中的任一者期间还调用声场分析单元44。如上文所描述,声场分析单元44可对于HOA系数11和/或经变换HOA系数33/35执行声场分析,以确定前景信道(nFG)45的总数目、背景声场(NBG)的阶数以及待发送的额外BG HOA信道的数目(nBGA)和索引(i)(其可在图3的实例中共同标示为背景信道信息43)(110)。

[0203] 音频编码装置20还可调用背景选择单元48。背景选择单元48可基于背景信道信息43确定背景或环境HOA系数47(112)。音频编码装置20可进一步调用前景选择单元36,所述

前景选择单元可基于nFG 45(其可表示标识前景向量的一或多个索引)选择表示声场的前景或相异分量的经重排序US[k]向量33'和经重排序V[k]向量35'(113)。

[0204] 音频编码装置20可调用能量补偿单元38。能量补偿单元38可对于环境HOA系数47执行能量补偿,以补偿归因于由背景选择单元48移除HOA系数中的各者而导致的能量损失(114),且借此产生经能量补偿的环境HOA系数47'。

[0205] 音频编码装置20还可调用空间-时间内插单元50。空间-时间内插单元50可对于经重排序的经变换HOA系数33'/35'执行空间-时间内插,以获得经内插的前景信号49'(其也可被称作“经内插的nFG信号49'”))和剩余的前景方向信息53(其也可被称作“V[k]向量53"))(116)。音频编码装置20接着可调用系数减少单元46。系数减少单元46可基于背景信道信息43对于剩余的前景V[k]向量53执行系数减少,以获得经减少的前景方向信息55(其也可被称为经减少的前景V[k]向量55)(118)。

[0206] 音频编码装置20接着可调用量化单元52以按上文所描述的方式压缩经减少的前景V[k]向量55且产生经译码前景V[k]向量57(120)。音频编码装置20还可调用去相关单元40'以应用相移去相关,以减少或消除HOA系数47'的背景信号之间的相关性,从而形成一或多个经去相关的HOA系数47"(121)。

[0207] 音频编码装置20还可调用心理声学音频译码器单元40。心理声学音频译码器单元40可对经能量补偿的环境HOA系数47'和经内插nFG信号49'的每一向量进行心理声学译码,以产生经编码环境HOA系数59和经编码nFG信号61(122)。音频编码装置接着可调用位流产生单元42。位流产生单元42可基于经译码前景方向信息57、经译码环境HOA系数59、经译码nFG信号61和背景信道信息43产生位流21(124)。

[0208] 图6A是说明音频解码装置(例如在图4的实例中展示的音频解码装置24)执行本发明中描述的技术的各种方面的示范性操作的流程图。最初,音频解码装置24可接收位流21(130)。在接收到位流之后,音频解码装置24即可调用提取单元72。出于论述的目的假设位流21指示将执行基于向量的重建构,提取单元72可解析位流以检索上文所提到的信息,从而将所述信息传递到基于向量的重构单元92。

[0209] 换句话说,提取单元72可按上文所描述的方式从位流21中提取经译码前景方向信息57(再次,其也可被称为经译码前景V[k]向量57)、经译码环境HOA系数59和经译码前景信号(其也可被称为经译码前景nFG信号59或经译码前景音频对象59)(132)。

[0210] 音频解码装置24可进一步调用解量化单元74。解量化单元74可对经译码前景方向信息57进行熵解码和解量化以获得经减少的前景方向信息55<sub>k</sub>(136)。音频解码装置24可调用再相关单元81。再相关单元81可将一或多个再相关变换应用于经能量补偿的环境HOA系数47'以获得一或多个经再相关的HOA系数47"(或相关的HOA系数47") ,并且可将相关的HOA系数47"传递到HOA系数制订单元82(任选地,通过淡化单元770)(137)。音频解码装置24还可调用心理声学解码单元80。心理声学音频解码单元80可对经编码环境HOA系数59和经编码前景信号61进行解码以获得经能量补偿的环境HOA系数47'和经内插的前景信号49'(138)。心理声学解码单元80可将经能量补偿的环境HOA系数47'传递到淡化单元770且将nFG信号49'传递到前景制订单元78。

[0211] 音频解码装置24接下来可调用空间-时间内插单元76。空间-时间内插单元76可接收经重排序的前景方向信息55<sub>k</sub>'且对于经减少的前景方向信息55<sub>k</sub>/55<sub>k-1</sub>执行空间-时间内

插以产生经内插的前景方向信息 $55_k''$  (140)。空间-时间内插单元76可将经内插的前景 $V[k]$ 向量 $55_k''$ 转发到淡化单元770。

[0212] 音频解码装置24可调用淡化单元770。淡化单元770可(例如,从提取单元72)接收或以其它方式获得指示经能量补偿的环境HOA系数47'何时处于转变中的语法元素(例如,AmbCoeffTransition语法元素)。淡化单元770可基于转变语法元素和所维持的转变状态信息使经能量补偿的环境HOA系数47'淡入或淡出,从而将经调整的环境HOA系数47''输出到HOA系数制订单元82。淡化单元770还可基于语法元素和所维持的转变状态信息而使经内插的前景 $V[k]$ 向量 $55_k''$ 的对应一或多个元素淡出或淡入,从而将经调整的前景 $V[k]$ 向量 $55_k'''$ 输出到前景制订单元78 (142)。

[0213] 音频解码装置24可调用前景制订单元78。前景制订单元78可执行nFG信号49'与经调整前景方向信息 $55_k'''$ 的矩阵乘法以获得前景HOA系数65 (144)。音频解码装置24还可调用HOA系数制订单元82。HOA系数制订单元82可将前景HOA系数65与经调整环境HOA系数47''相加以便获得HOA系数11' (146)。

[0214] 图6B是说明音频编码装置和音频解码装置执行本发明中描述的译码技术的示范性操作的流程图。图6B是说明根据本发明的一或多个方面的实例编码和解码过程160的流程图。尽管过程160可由多种装置执行,但为便于论述,在本文中关于上文所描述的音频编码装置20和音频解码装置24来描述过程160。使用图6B中的虚线将过程160的编码区段与解码区段分界。过程160可开始于音频编码装置20的一或多个组件(例如,前景选择单元36和背景选择单元48)使用HOA空间编码从HOA输入产生前景信道164和一阶HOA背景信道166 (162)。继而,去相关单元40'可将去相关变换(例如,呈基于相位的去相关变换或矩阵形式)应用于经能量补偿的环境HOA系数47'。更具体来说,音频编码装置20可将UHJ矩阵或基于相位的去相关变换(例如,通过标量乘法)应用于经能量补偿的环境HOA系数47' (168)。

[0215] 在一些实例中,如果去相关单元40',在其中去相关单元40'确定HOA背景信道包含较少数目个信道(例如,四个)的例子中,去相关单元40'可应用UHJ矩阵(或基于相位的变换)。相反地,在这些实例中,如果去相关单元40'确定HOA背景信道包含较大数目个信道(例如,九个),那么音频编码装置20可选择不同于UHJ矩阵的去相关变换(例如,在MPEG-H标准中描述的模式矩阵)并将所述去相关变换应用于HOA背景信道。通过将去相关变换(例如,UHJ矩阵)应用于HOA背景信道,音频编码装置20可获得经去相关的HOA背景信道。

[0216] 如在图6B中展示,音频编码装置20(例如,通过调用心理声学音频译码器单元40)可将时间编码(例如,通过应用AAC和/或USAC)应用于经去相关的HOA背景信号(170)以及应用于任何前景信道(166)。应了解,在一些情境中,心理声学音频译码器单元40可确定前景信道的数目可为零(即,在这些情境中,心理声学音频译码器单元40可不从HOA输入获得任何前景信道)。因为AAC和/或USAC可能不经优化以用于或以其它方式非常适合于立体声音频数据,去相关单元40'可应用去相关矩阵以减少或消除HOA背景信道之间的相关性。经去相关的HOA背景信道中展示的经减少的相关性提供在AAC/USAC时间编码阶段减轻或消除噪声去掩蔽的潜在优点,这是因为AAC和USAC可能不针对立体声音频数据经优化。

[0217] 继而,音频解码装置24可执行对由音频编码装置20输出的经编码位流的时间解码。在过程160的实例中,音频解码装置24的一或多个组件(例如,心理声学解码单元80)可分别对于前景信道(如果有任何前景信道包含在位流中) (172) 和背景信道(174) 执行时间

解码。另外,再相关单元81可将再相关变换应用于经时间解码的HOA背景信道。作为一实例,再相关单元81可以互逆方式将去相关变换应用于去相关单元40'。举例来说,如在过程160的具体实例中所描述,再相关单元81可将UHJ矩阵或基于相位的变换应用于经时间解码的HOA背景信号(176)。

[0218] 在一些实例中,如果再相关单元81确定经时间解码的HOA背景信号包含较少数目个信道(例如,四个),那么再相关单元81可应用UHJ矩阵或基于相位的变换。相反地,在这些实例中,如果再相关单元81确定经时间解码的HOA背景信道包含较大数目个信道(例如,九个),那么再相关单元81可选择不同于UHJ矩阵的去相关变换(例如,在MPEG-H标准中描述的模式矩阵)并将所述去相关变换应用于HOA背景信道。

[0219] 另外,HOA系数制订单元82可执行对相关的HOA背景信道和任何可用的经解码前景信道的HOA空间解码(178)。继而,HOA系数制订单元82可向一或多个输出装置(例如扩音器和/或头戴式耳机(包含但不限于具有立体声或环绕声能力的输出装置)再现经解码的音频信号(180)。

[0220] 可对于任何数目个不同上下文和音频生态系统执行前述技术。下文描述数个实例上下文,但所述技术不应限于所述实例上下文。一个实例音频生态系统可包含音频内容、电影工作室、音乐工作室、游戏音频工作室、基于信道的音频内容、译码引擎、游戏音频原声(stem)、游戏音频译码/再现引擎,以及递送系统。

[0221] 电影工作室、音乐工作室和游戏音频工作室可接收音频内容。在一些实例中,音频内容可表示获取内容的输出。电影工作室可例如通过使用数字音频工作站(DAW)输出基于信道的音频内容(例如,呈2.0、5.1和7.1)。音乐工作室可例如通过使用DAW输出基于信道的音频内容(例如,呈2.0和5.1)。在一情况下,译码引擎可基于一或多个编码解码器(例如,AAC、AC3、杜比真HD(Dolby True HD)、杜比数字加(Dolby Digital Plus)以及DTS主音频)接收并编码基于信道的音频内容以用于由递送系统输出。游戏音频工作室可例如通过使用DAW输出一或多个游戏音频原声。游戏音频译码/再现引擎可译码音频原声和/或将音频原声再现成基于信道的音频内容以供递送系统输出。可执行所述技术的另一实例上下文包括音频生态系统,其可包含广播记录音频对象、专业音频系统、消费型装置上捕获、HOA音频格式、装置上再现、消费型音频、TV、和配件,以及汽车音频系统。

[0222] 广播记录音频对象、专业音频系统和消费型装置上捕获都可使用HOA音频格式对其输出进行译码。以此方式,可使用HOA音频格式将音频内容译码成单一表示,可使用装置上再现、消费型音频、TV、和配件以及汽车音频系统回放所述单一表示。换句话说,可在通用音频回放系统(即,与要求例如5.1、7.1等的特定配置相反)(例如,音频回放系统16)处回放音频内容的单一表示。

[0223] 可执行所述技术的上下文的其它实例包含可包含获取元件和回放元件的音频生态系统。获取元件可包含有线和/或无线获取装置(例如,本征麦克风)、装置上环绕声捕获,以及移动装置(例如,智能电话和平板计算机)。在一些实例中,有线和/或无线获取装置可经由有线和/或无线通信信道耦合到移动装置。

[0224] 根据本发明的一或多种技术,移动装置可用以获取声场。举例来说,移动装置可经由有线和/或无线获取装置和/或装置上环绕声捕获(例如,集成到移动装置中的多个麦克风)获取声场。移动装置接着可将所获取声场译码成HOA系数以用于由回放元件中的一或多

者回放。举例来说，移动装置的用户可记录实况事件(例如，集会、会议、比赛、音乐会等)(获取实况事件的声场)，且将所述记录内容译码成HOA系数。

[0225] 移动装置还可使用回放元件中的一或者者来回放经HOA译码声场。举例来说，移动装置可对经HOA译码声场进行解码，且将致使回放元件中的一或者者重新产生声场的信号输出到所述回放元件中的所述一或者者。作为一个实例，移动装置可使用无线和/或无线通信信道将信号输出到一或多个扬声器(例如，扬声器阵列、声棒(sound bar)等)。作为另一实例，移动装置可使用对接解决方案将信号输出到一或多个对接站和/或一或多个对接的扬声器(例如，智能汽车和/或家庭中的声音系统)。作为另一实例，移动装置可使用头戴式耳机再现将信号输出到一组头戴式耳机(例如)以创建逼真的双耳声。

[0226] 在一些实例中，特定移动装置可获取3D声场以及在稍后时间回放同一3D声场。在一些实例中，移动装置可获取3D声场，将3D声场编码为HOA，并且将经编码3D声场传输到一或多个其它装置(例如，其它移动装置和/或其它非移动装置)以用于回放。

[0227] 可执行所述技术的又一上下文包含音频生态系统，其可包含音频内容、游戏工作室、经译码音频内容、再现引擎以及递送系统。在一些实例中，游戏工作室可包含可支持HOA信号的编辑的一或多个DAW。举例来说，所述一或多个DAW可包含可经配置以与一或多个游戏音频系统一起操作(例如，工作)的HOA插件和/或工具。在一些实例中，游戏工作室可输出支持HOA的新原声格式。在任何情况下，游戏工作室可将经译码音频内容输出到再现引擎，所述再现引擎可再现声场以供递送系统回放。

[0228] 也可对于示范性音频获取装置执行所述技术。举例来说，可对于可包含共同地经配置以记录3D声场的多个麦克风的本征麦克风执行所述技术。在一些实例中，本征麦克风的所述多个麦克风可位于具有大约4cm的半径的基本呈球形的球体的表面上。在一些实例中，音频编码装置20可集成到本征麦克风中以便直接从麦克风输出位流21。

[0229] 另一示范性音频获取上下文可包含可经配置以从一或多个麦克风(例如，一或多个本征麦克风)接收信号的制作车。制作车还可包含音频编码器，例如图3的音频编码器20。

[0230] 在一些例子中，移动装置还可包含共同地经配置以记录3D声场的多个麦克风。换句话说，所述多个麦克风可具有X、Y、Z分集。在一些实例中，移动装置可包含可旋转以相对于移动装置的一或多个其它麦克风提供X、Y、Z分集的麦克风。移动装置还可包含音频编码器，例如图3的音频编码器20。

[0231] 加固型视频捕获装置可进一步经配置以记录3D声场。在一些实例中，加固型视频捕获装置可附接到参与活动的用户的头盔。举例来说，加固型视频捕获装置可在用户泛舟时附接到用户的头盔。以此方式，加固型视频捕获装置可捕获表示用户周围的动作(例如，水在用户身后的撞击、另一泛舟者在用户前方说话等)的3D声场。

[0232] 还可对于可经配置以记录3D声场的配件增强型移动装置执行所述技术。在一些实例中，移动装置可类似于上文所论述的移动装置，其中添加了一或多个配件。举例来说，本征麦克风可附接到上文所提及的移动装置以形成配件增强型移动装置。以此方式，与仅使用与配件增强型移动装置成一体的声音捕获组件相比，配件增强型移动装置可捕获3D声场的更高质量版本。

[0233] 下文进一步论述可执行本发明中所描述的技术的各种方面的实例音频回放装置。根据本发明的一或多种技术，扬声器和/或声棒在回放3D声场时可布置于任何任意的配置

中。此外,在一些实例中,头戴式耳机回放装置可经由有线或无线连接耦合到解码器24。根据本发明的一或多种技术,可使用声场的单一通用表示来在扬声器、声棒和头戴式耳机回放装置的任何组合上再现声场。

[0234] 多个不同实例音频回放环境还可适用于执行本发明中所描述的技术的各种方面。举例来说,以下环境可为用于执行本发明中所描述的技术的各种方面的合适环境:5.1扬声器回放环境、2.0(例如,立体声)扬声器回放环境、具有全高前扩音器的9.1扬声器回放环境、22.2扬声器回放环境、16.0扬声器回放环境、汽车扬声器回放环境,以及具有耳芽(ear bud)回放环境的移动装置。

[0235] 根据本发明的一或多种技术,可利用声场的单一通用表示来在前述回放环境中的任一者上再现声场。另外,本发明的技术使得再现器能够从通用表示再现声场以用于在除上文所描述的环境之外的回放环境上回放。举例来说,如果设计考虑禁止扬声器根据7.1扬声器回放环境的恰当放置(例如,如果不可能放置右环绕扬声器),那么本发明的技术使得再现器能够以其它6个扬声器进行补偿,使得可在6.1扬声器回放环境上实现回放。

[0236] 此外,用户可在佩戴头戴式耳机时观看运动比赛。根据本发明的一或多种技术,可获取运动比赛的3D声场(例如,可将一或多个本征麦克风放置在棒球场中和/或周围),可获得对应于3D声场的HOA系数且将所述HOA系数传输到解码器,所述解码器可基于HOA系数重构3D声场且将经重构3D声场输出到再现器,且所述再现器可获得关于回放环境的类型(例如,头戴式耳机)的指示,且将经重构3D声场再现为致使头戴式耳机输出运动比赛的3D声场的表示的信号。

[0237] 在上述各种例子中的每一者中,应理解,音频编码装置20可执行方法,或另外包括执行音频编码装置20经配置以执行的方法的每一步骤的装置。在一些例子中,这些装置可包括一或多个处理器。在一些例子中,所述一或多个处理器可表示借助于存储到非暂时性计算机可读存储媒体的指令配置的专用处理器。换句话说,编码实例的集合中的每一者中的技术的各种方面可提供其上存储有指令的非暂时性计算机可读存储媒体,所述指令在执行时致使一或多个处理器执行音频编码装置20已经配置以执行的方法。

[0238] 在一或多个实例中,所描述的功能可用硬件、软件、固件或其任何组合实施。如果以软件实施,那么所述功能可以作为一或多个指令或代码在计算机可读媒体上存储或传输,并且由基于硬件的处理单元来执行。计算机可读媒体可包含计算机可读存储媒体,其对应于例如数据存储媒体等有形媒体。数据存储媒体可为可由一或多个计算机或一个或多个处理器存取以检索用于实施本发明中描述的技术的指令、代码和/或数据结构的任何可用媒体。计算机程序产品可以包含计算机可读媒体。

[0239] 同样地,在上文所描述的各种例子中的每一者中,应理解,音频解码装置24可执行方法或另外包括用于执行音频解码装置24经配置以执行的方法的每一步骤的装置。在一些例子中,所述装置可包括一或多个处理器。在一些例子中,所述一或多个处理器可表示借助于存储到非暂时性计算机可读存储媒体的指令配置的专用处理器。换句话说,编码实例的集合中的每一者中的技术的各种方面可提供其上存储有指令的非暂时性计算机可读存储媒体,所述指令在执行时致使所述一或多个处理器执行音频解码装置24已经配置以执行的方法。

[0240] 借助于实例而非限制,此类计算机可读存储媒体可包括RAM、ROM、EEPROM、CD-ROM

或其它光盘存储装置、磁盘存储装置或其它磁性存储装置、快闪存储器,或可用以存储指令或数据结构的形式的期望程序代码并且可以由计算机存取的任何其它媒体。然而,应理解,所述计算机可读存储媒体和数据存储媒体并不包含连接、载波、信号或其它暂时性媒体,而是实际上针对非暂时性的有形存储媒体。如本文中所使用,磁盘和光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软性磁盘和蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘利用激光以光学方式再现数据。以上各项的组合也应包含在计算机可读媒体的范围内。

[0241] 可由例如一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其它等效集成或离散逻辑电路等一或多个处理器来执行指令。因此,如本文中所使用的术语“处理器”可指前述结构或适于实施本文中所描述的技术的任一其它结构中的任一者。另外,在一些方面中,本文中所描述的功能性可以在经配置用于编码和解码的专用硬件和/或软件模块内提供,或者并入于组合式编码解码器中。而且,所述技术可完全实施于一或多个电路或逻辑元件中。

[0242] 本发明的技术可在广泛多种装置或设备中实施,包含无线手持机、集成电路(IC)或一组IC(例如,芯片组)。本发明中描述各种组件、模块或单元是为了强调经配置以执行所揭示的技术的装置的功能方面,但未必需要由不同硬件单元实现。实际上,如上文所描述,各种单元可以结合合适的软件和/或固件组合在编码解码器硬件单元中,或者通过互操作硬件单元的集合来提供,所述硬件单元包含如上文所描述的一或多个处理器。

[0243] 已经描述了所述技术的各种方面。所述技术的这些和其它方面在所附权利要求书的范围内。

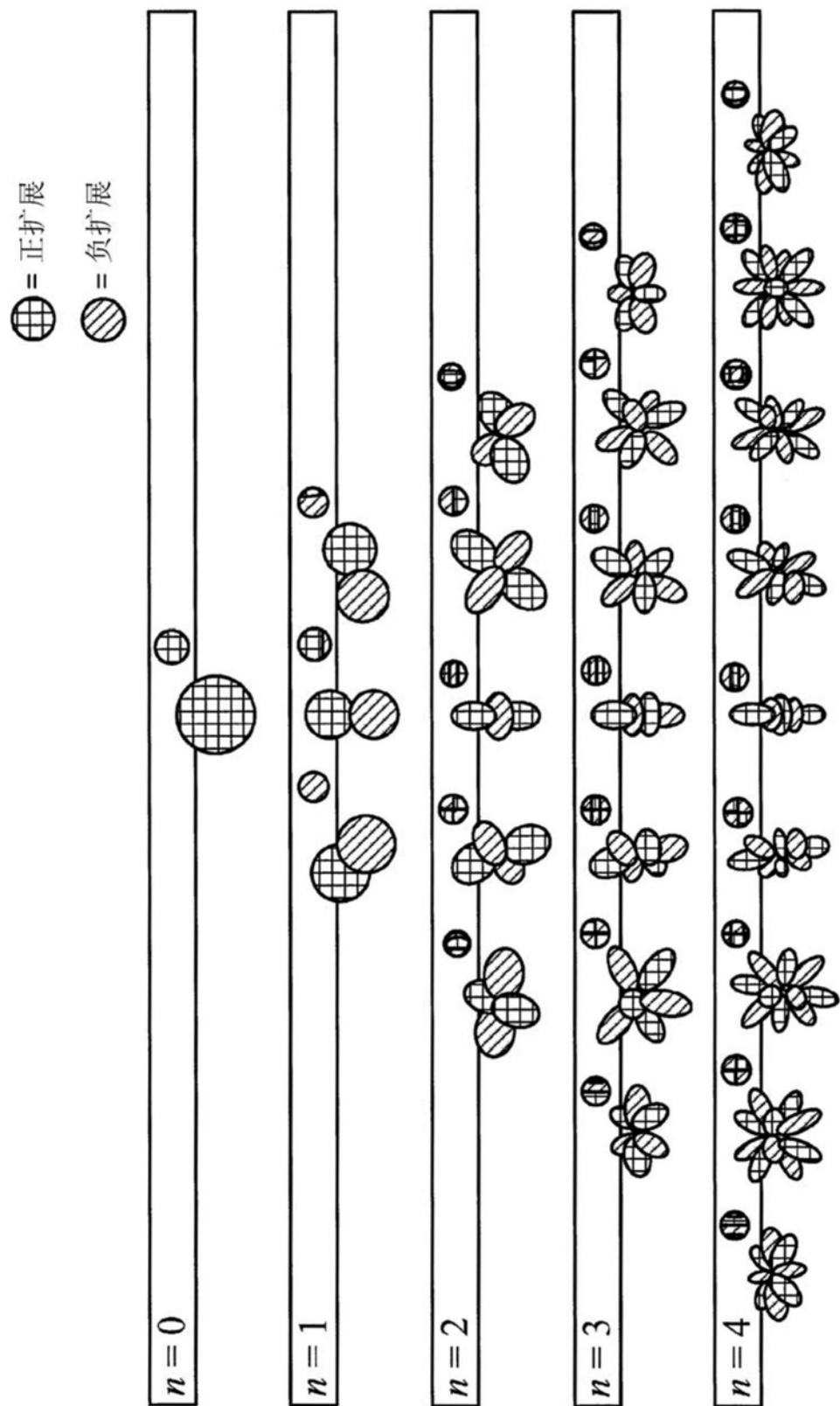


图1

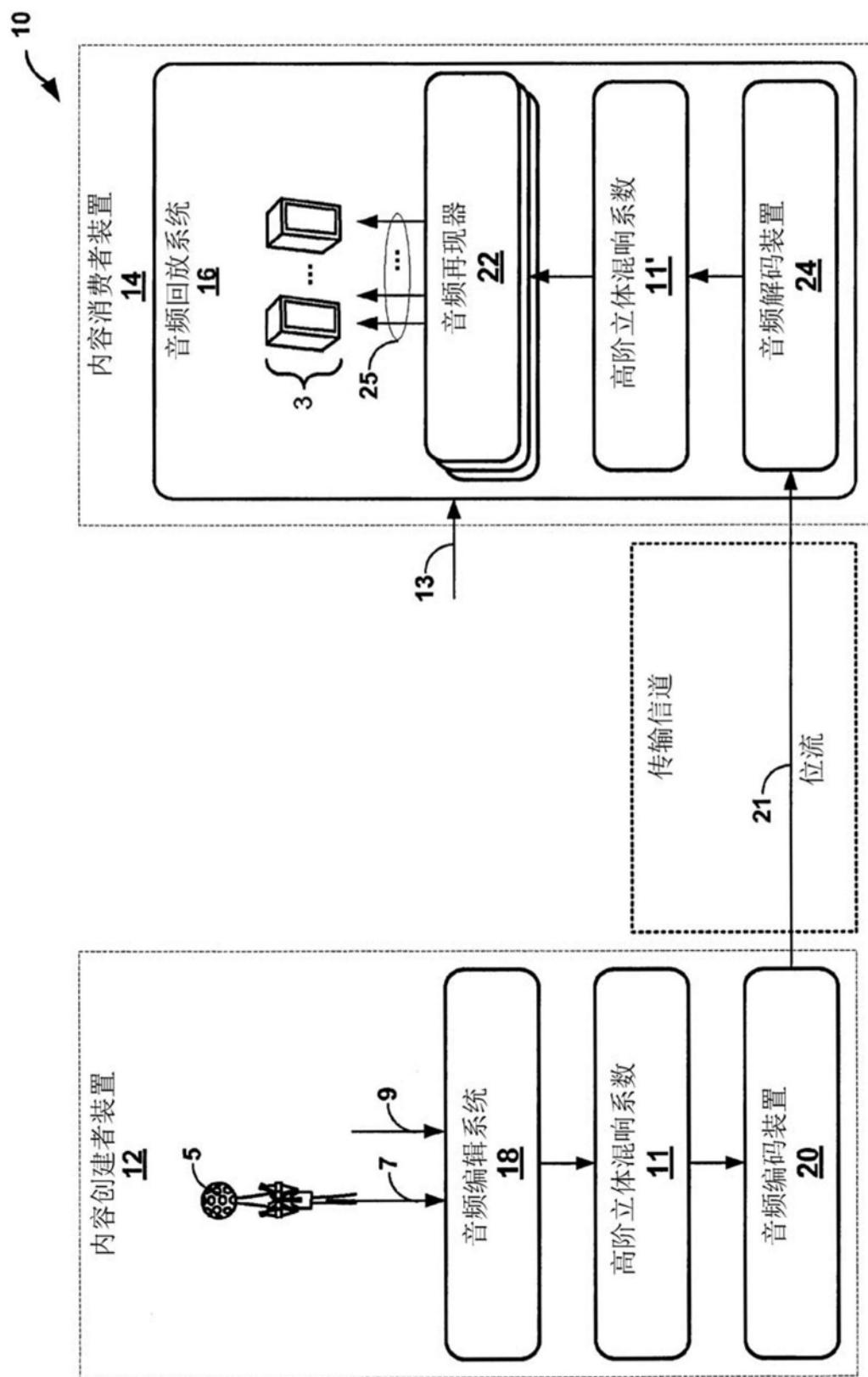


图2

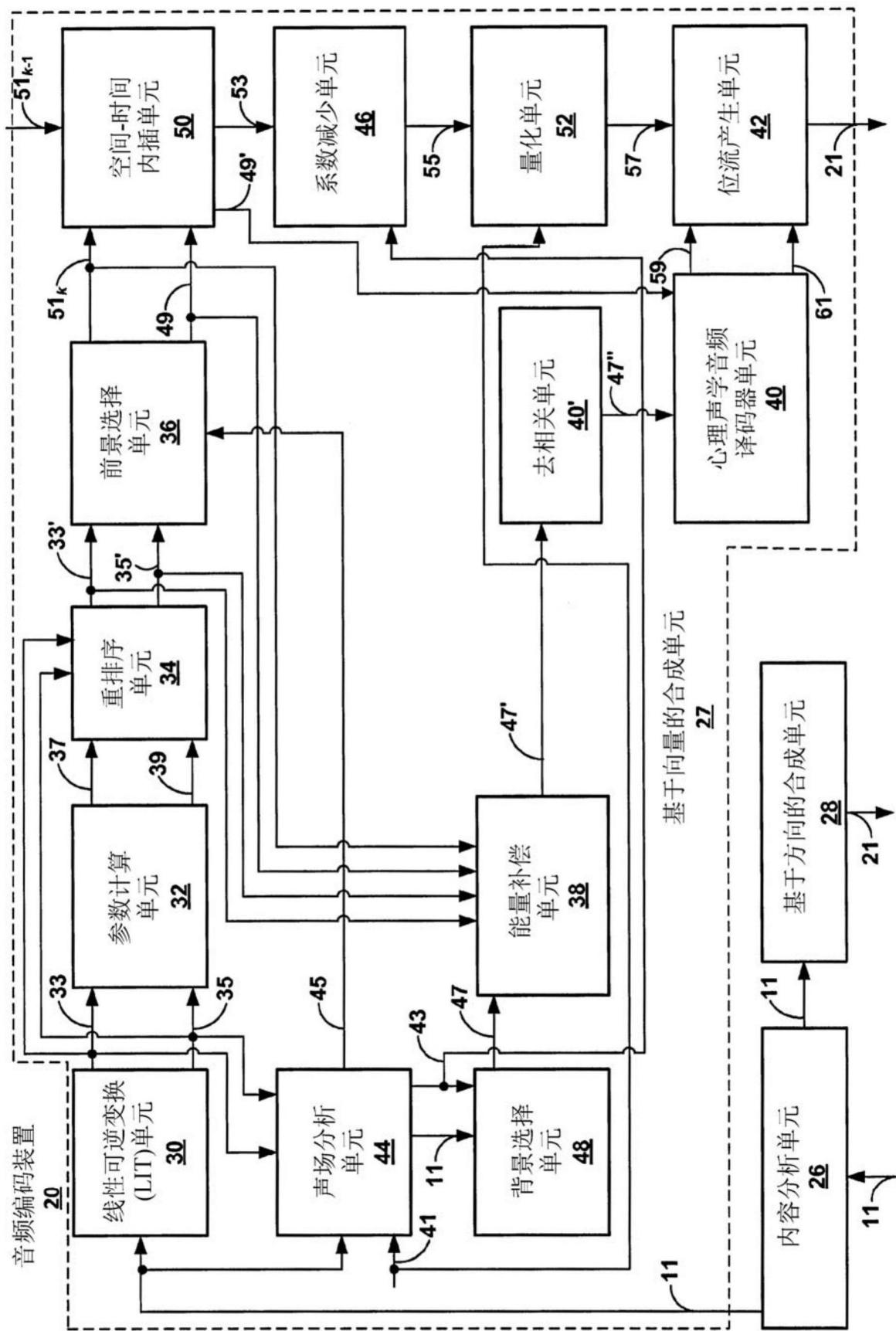


图3

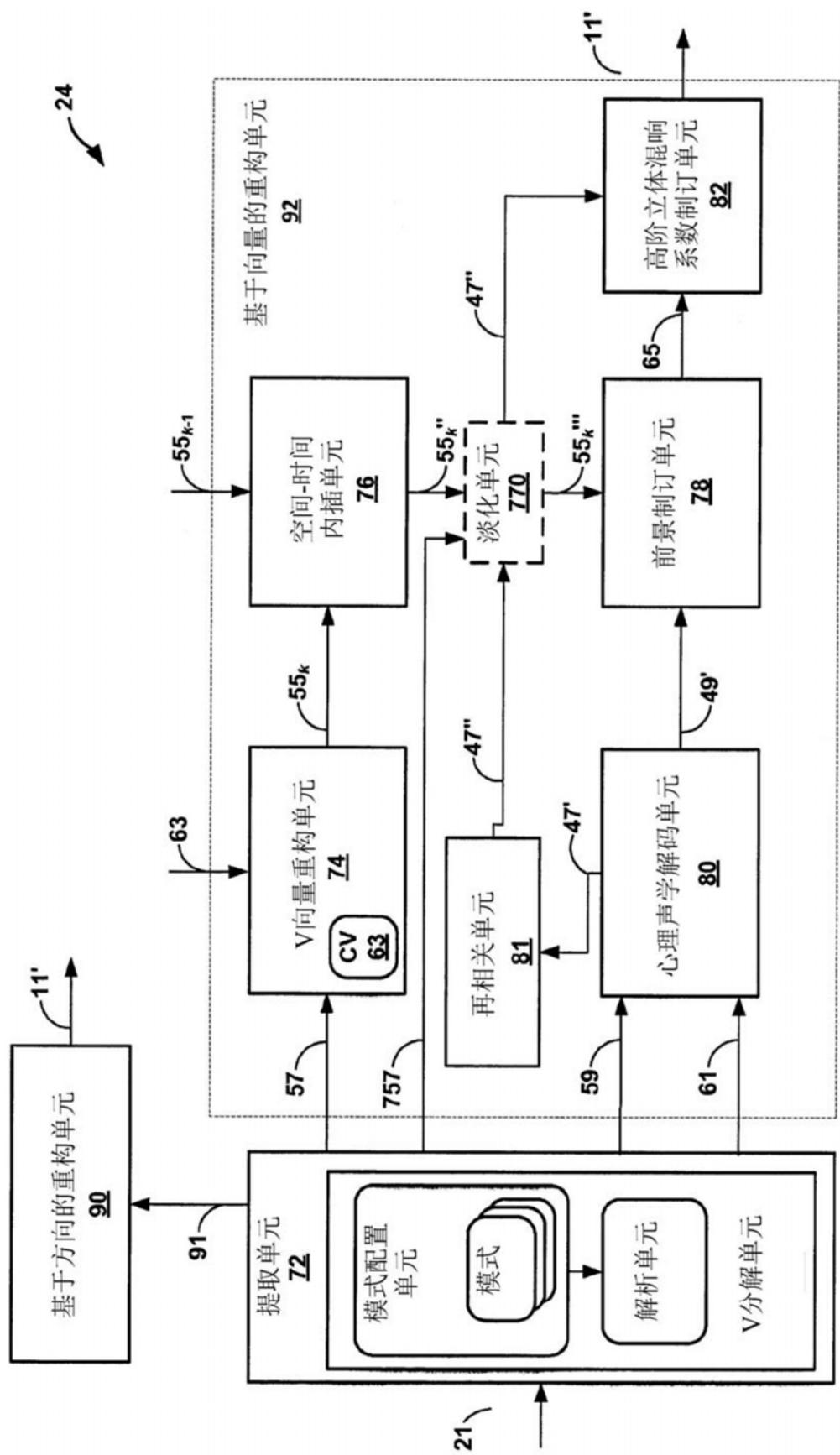


图4

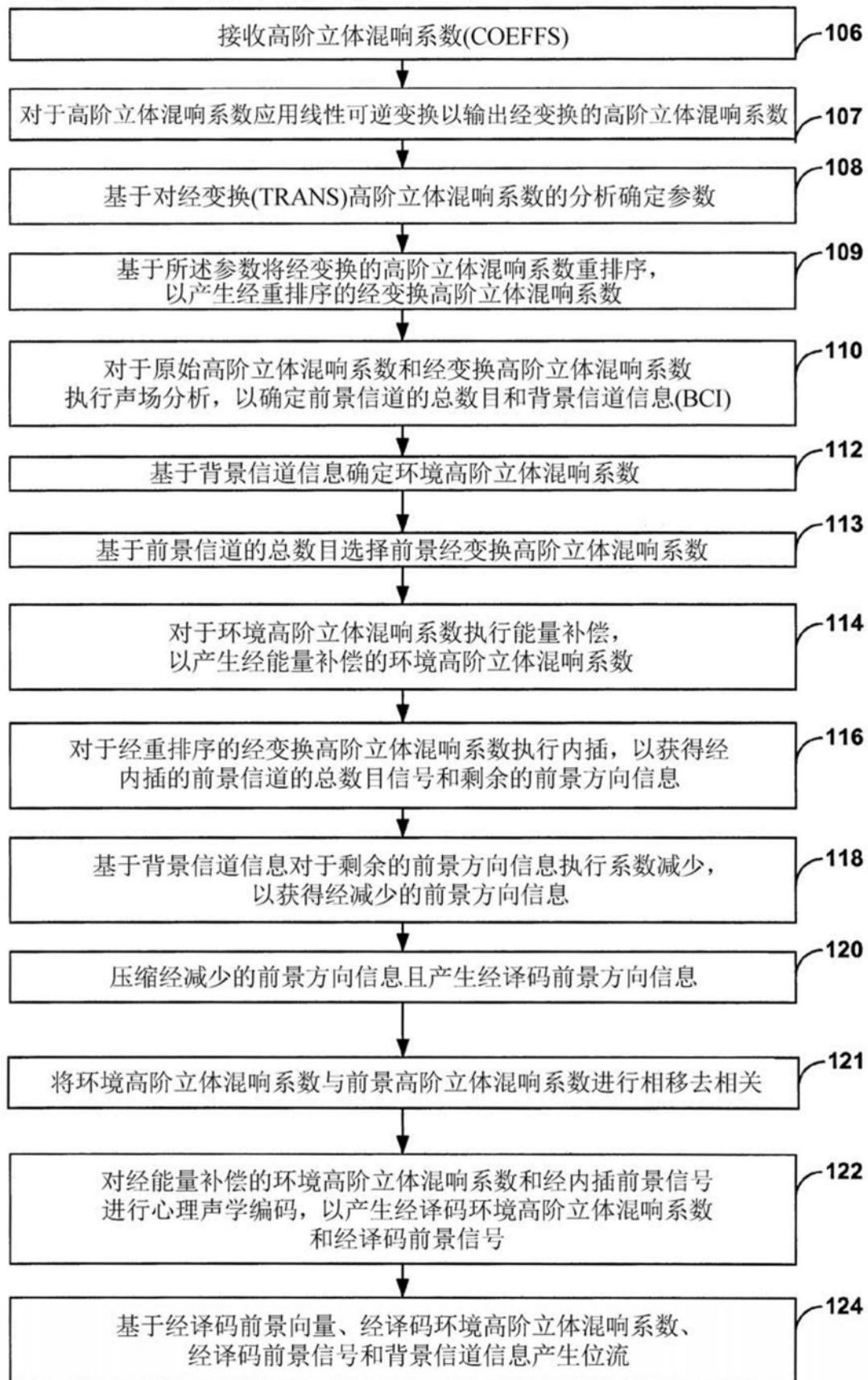


图5

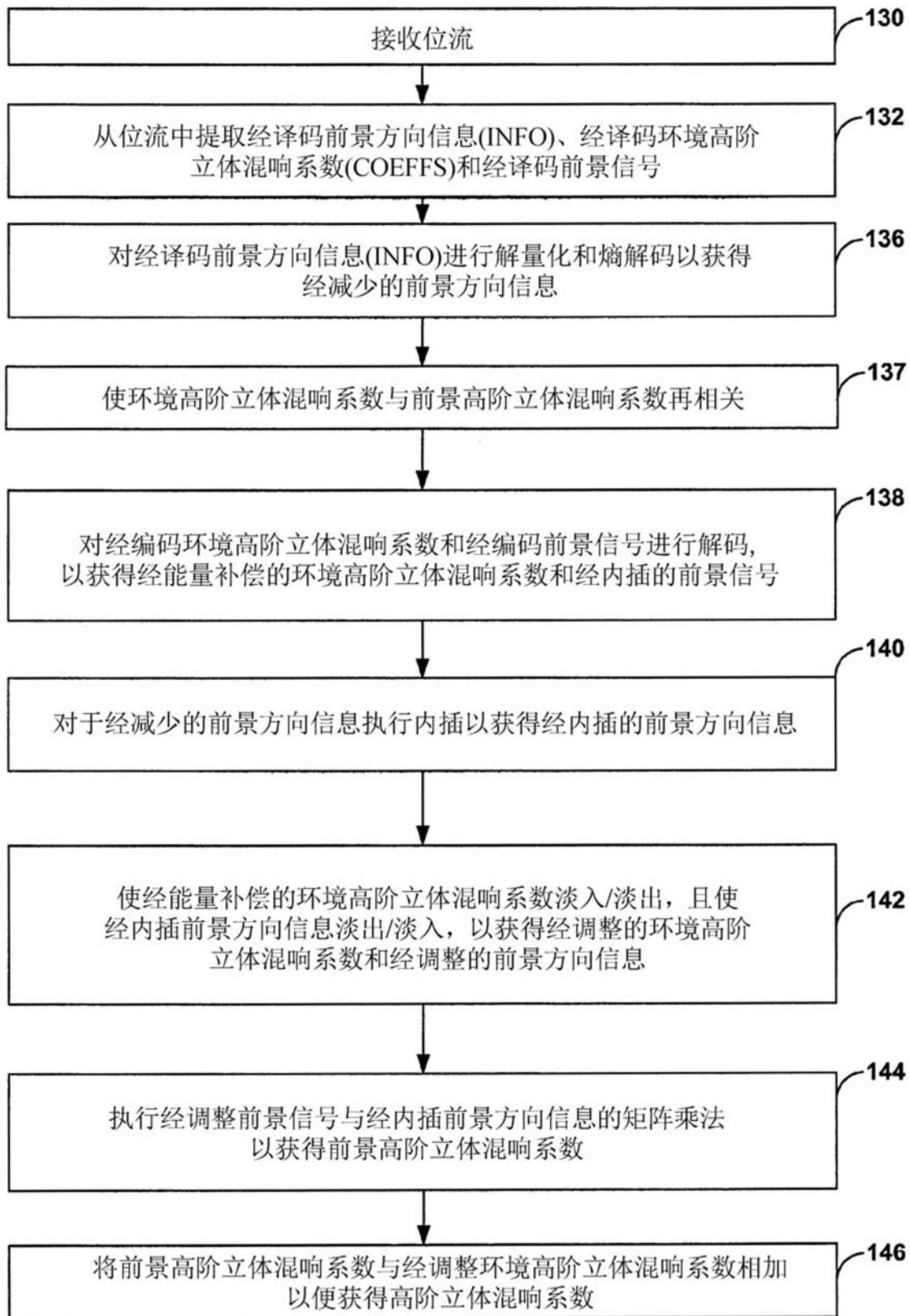


图6A

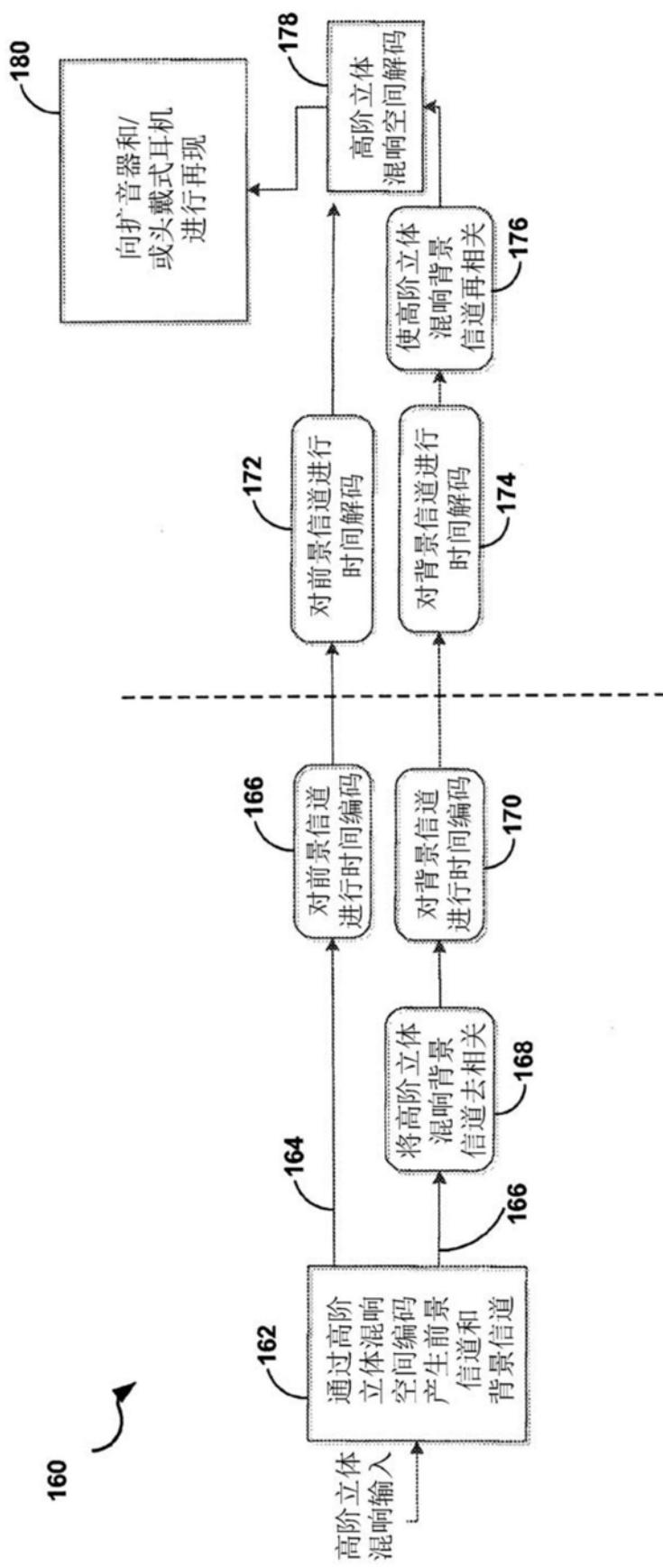


图6B