



(12) 发明专利

(10) 授权公告号 CN 106796795 B

(45) 授权公告日 2021.07.06

(21) 申请号 201580054372.5

(87) PCT国际申请的公布数据

W02016/057925 EN 2016.04.14

(22) 申请日 2015.10.09

(73) 专利权人 高通股份有限公司

地址 美国加利福尼亚州

(65) 同一申请的已公布的文献号
申请公布号 CN 106796795 A(72) 发明人 金墨永 尼尔斯·京特·彼得斯
迪潘让·森

(43) 申请公布日 2017.05.31

(74) 专利代理机构 北京市柳沈律师事务所
11105

(30) 优先权数据

代理人 安之斐

62/062,584 2014.10.10 US
 62/084,461 2014.11.25 US
 62/087,209 2014.12.03 US
 62/088,445 2014.12.05 US
 62/145,960 2015.04.10 US
 62/175,185 2015.06.12 US
 62/187,799 2015.07.01 US
 62/209,764 2015.08.25 US
 14/878,691 2015.10.08 US

(51) Int.CI.

G10L 19/008 (2013.01)
 G10L 19/16 (2013.01)

(85) PCT国际申请进入国家阶段日

(56) 对比文件

2017.04.06

US 2014016784 A1, 2014.01.16
 CN 101379554 A, 2009.03.04
 US 2014023196 A1, 2014.01.23
 US 2014016786 A1, 2014.01.16

(86) PCT国际申请的申请数据

审查员 宋玉香

PCT/US2015/054950 2015.10.09

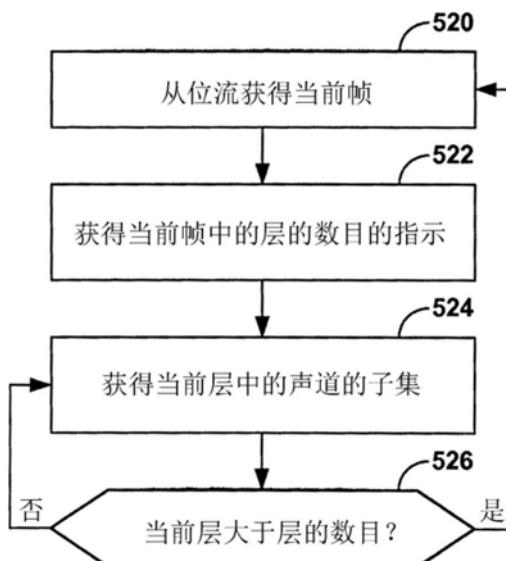
权利要求书5页 说明书61页 附图38页

(54) 发明名称

以信号表示用于高阶立体混响音频数据的
可缩放译码的层

(57) 摘要

大体来说，本发明描述用于以信号表示用于高阶立体混响音频数据的可缩放译码的层的技术。一种包括存储器和处理器的装置可经配置以执行所述技术。所述存储器可经配置以存储所述位流。所述处理器可经配置以从所述位流获得在所述位流中指定的层的数目的指示，且基于层的所述数目的所述指示获得所述位流的所述层。



1. 一种经配置以对表示高阶立体混响音频信号的位流进行解码的装置，所述装置包括：

存储器，其经配置以存储所述位流；以及

一个或多个处理器，其经配置以：

从所述位流获得在所述位流中指定的传输信道的总数目的指示；

从所述位流获得在所述位流中指定的层的数目的指示；以及

基于在所述位流中指定的层的数目的指示和在所述位流中指定的传输信道的总数目的指示获得所述位流的所述层；

其中获得在所述位流中指定的层的数目的指示包括：基于包括在所述层的位流中的语法元素，获得用于每个层的传输信道的数目的指示；

其中当剩余的传输信道的数目大于一个时，迭代获得所述层的传输信道的数目的指示，每次迭代计算出的剩余传输信道的数目为所述传输信道的总数目减去已获得的传输信道的数目的累计和；

其中在每次迭代中，使用基于剩余传输信道的数目计算出的位数，在位流中编码相应层的传输信道的数目的指示；

其中通过在每次迭代时递增计数器来获得在所述位流中指定的层的数目的指示。

2. 根据权利要求1所述的装置，

其中所述一个或多个处理器经配置以获得所述位流中针对所述层中的至少一者指定的前景声道的数目的指示，且

其中所述一个或多个处理器经配置以基于前景声道的所述数目的所述指示获得所述位流的所述层中的所述至少一者的所述前景声道。

3. 根据权利要求1所述的装置，

其中所述一个或多个处理器经配置以获得所述位流中针对所述层中的至少一者指定的背景声道的数目的指示，且

其中所述一个或多个处理器经配置以基于背景声道的所述数目的所述指示获得所述位流的所述层中的所述至少一者的所述背景声道。

4. 根据权利要求1所述的装置，

其中所述层的所述数目的所述指示指示层的所述数目为二，

其中两个层包括基础层和增强层，且

其中所述一个或多个处理器步经配置以获得前景声道的数目对于所述基础层为零且对于所述增强层为二的指示。

5. 根据权利要求1或4所述的装置，

其中所述层的所述数目的所述指示指示层的所述数目为二，

其中两个层包括基础层和增强层，且

其中所述一个或多个处理器经配置以获得背景声道的数目对于所述基础层为四且对于所述增强层为零的指示。

6. 根据权利要求1所述的装置，

其中所述层的所述数目的所述指示指示层的所述数目为三，

其中三个层包括基础层、第一增强层和第二增强层，且

其中所述一个或多个处理器经配置以获得前景声道的数目对于所述基础层为零,对于所述第一增强层为二且对于所述第二增强层为二的指示。

7. 根据权利要求1或6所述的装置,

其中所述层的所述数目的所述指示指示层的所述数目为三,

其中三个层包括基础层、第一增强层和第二增强层,且

其中所述一个或多个处理器进一步经配置以获得背景声道的数目对于所述基础层为二,对于所述第一增强层为零且对于所述第二增强层为零的指示。

8. 根据权利要求1所述的装置,

其中所述层的所述数目的所述指示指示层的所述数目为三,

其中三个层包括基础层、第一增强层和第二增强层,且

其中所述一个或多个处理器经配置以获得前景声道的数目对于所述基础层为二,对于第一增强层为二且对于第二增强层为二的指示。

9. 根据权利要求1或8所述的装置,

其中所述层的所述数目的所述指示指示层的所述数目为三,

其中三个层包括基础层、第一增强层和第二增强层,且

其中所述一个或多个处理器进一步经配置以获得指示背景声道的所述数目对于所述基础层为零,对于所述第一增强层为零且对于所述第二增强层为零的背景语法元素。

10. 根据权利要求1所述的装置,

其中层的所述数目的所述指示包括所述位流在前一帧中的层的数目的指示,且

其中所述一个或多个处理器进一步经配置以:

获得当与所述位流在所述前一帧中的层的所述数目相比时所述位流的层的数目在当前帧中是否已改变的指示;以及

基于所述位流的层的所述数目在所述当前帧中是否已改变的所述指示,获得所述位流在所述当前帧中的层的所述数目。

11. 根据权利要求10所述的装置,其中所述一个或多个处理器进一步经配置以在所述指示指示当与所述位流在所述前一帧中的层的所述数目相比时所述位流的层的所述数目在所述当前帧中尚未改变时,确定所述位流在所述当前帧中的层的所述数目等于所述位流在所述前一帧中的层的所述数目。

12. 根据权利要求10所述的装置,其中所述一个或多个处理器进一步经配置以在所述指示指示当与所述位流在所述前一帧中的层的所述数目相比时所述位流的层的所述数目在所述当前帧中尚未改变时,获得所述当前帧的所述层中的一者或多者的分量的当前数目等于所述前一帧的所述层中的一个或多个中的分量的先前数目的指示。

13. 根据权利要求1所述的装置,

其中层的所述数目的所述指示指示在所述位流中指定了三个层,且

其中所述一个或多个处理器经配置以:

获得指示所述高阶立体混响音频信号的提供立体声声道回放的背景分量的所述位流的所述层中的第一个层;

获得指示所述高阶立体混响音频信号的提供由布置在一个或多个水平面上的三个或更多个扬声器进行三维回放的背景分量的所述位流的所述层中的第二个层;以及

获得指示所述高阶立体混响音频信号的前景分量的所述位流的所述层中的第三个层。

14. 根据权利要求1所述的装置，

其中层的所述数目的所述指示指示在所述位流中指定了三个层，且

其中所述一个或多个处理器经配置以：

获得指示所述高阶立体混响音频信号的提供单声道回放的背景分量的所述位流的所述层中的第一个层；

获得指示所述高阶立体混响音频信号的提供由布置在一个或多个水平面上的三个或更多个扬声器进行三维回放的背景分量的所述位流的所述层中的第二个层；以及

获得指示所述高阶立体混响音频信号的前景分量的所述位流的所述层中的第三个层。

15. 根据权利要求1所述的装置，

其中层的所述数目的所述指示指示在所述位流中指定了四个层，且

其中所述一个或多个处理器经配置以：

获得指示所述高阶立体混响音频信号的提供立体声声道回放的背景分量的所述位流的所述层中的第一个层；

获得指示所述高阶立体混响音频信号的提供由布置在单个水平面上的三个或更多个扬声器进行多声道回放的背景分量的所述位流的所述层中的第二个层；以及

获得指示所述高阶立体混响音频信号的提供由布置在两个或更多个水平面上的三个或更多个扬声器进行三维回放的背景分量的所述位流的所述层中的第三个层；以及

获得指示所述高阶立体混响音频信号的前景分量的所述位流的所述层中的第四个层。

16. 根据权利要求1所述的装置，

其中层的所述数目的所述指示指示在所述位流中指定了四个层，且

其中所述一个或多个处理器经配置以：

获得指示所述高阶立体混响音频信号的提供单声道回放的背景分量的所述位流的所述层中的第一个层；

获得指示所述高阶立体混响音频信号的提供由布置在单个水平面上的三个或更多个扬声器进行多声道回放的背景分量的所述位流的所述层中的第二个层；

获得指示所述高阶立体混响音频信号的提供由布置在两个或更多个水平面上的三个或更多个扬声器进行三维回放的背景分量的所述位流的所述层中的第三个层；以及

获得指示所述高阶立体混响音频信号的前景分量的所述位流的所述层中的第四个层。

17. 根据权利要求1所述的装置，

其中层的所述数目的所述指示指示在所述位流中指定了两个层，且

其中所述一个或多个处理器经配置以：

获得指示所述高阶立体混响音频信号的提供立体声声道回放的背景分量的所述位流的所述层中的第一个层；以及

获得指示所述高阶立体混响音频信号的提供由布置在单个水平面上的三个或更多个扬声器进行水平多声道回放的背景分量的所述位流的所述层中的第二个层。

18. 根据权利要求1所述的装置，其进一步包括经配置以基于所述高阶立体混响音频信号再现声场的扩音器。

19. 一种对表示高阶立体混响音频信号的位流进行解码的方法，所述方法包括：

从所述位流获得在所述位流中指定的传输信道的总数目的指示；

从所述位流获得在所述位流中指定的层的数目的指示；以及

基于在所述位流中指定的层的数目的指示和在所述位流中指定的传输信道的总数目的指示获得所述位流的所述层；

其中获得在所述位流中指定的层的数目的指示包括：基于包括在所述层的位流中的语法元素，获得用于每个层的传输信道的数目的指示；

其中当剩余的传输信道的数目大于一个时，迭代获得所述层的传输信道的数目的指示，每次迭代计算出的剩余传输信道的数目为所述传输信道的总数目减去已获得的传输信道的数目的累计和；

其中在每次迭代中，使用基于剩余传输信道的数目计算出的位数，在位流中编码相应层的传输信道的数目的指示；

其中通过在每次迭代时递增计数器来获得在所述位流中指定的层的数目的指示。

20. 根据权利要求19所述的方法，其中获得在所述位流中指定的声道的所述数目的所述指示包括获得所述位流中针对所述层中的至少一者指定的前景声道的数目的指示，

其中获得所述层包括基于前景声道的所述数目的所述指示获得所述位流的所述层中的所述至少一者的所述前景声道。

21. 根据权利要求19所述的方法，其中获得在所述位流中指定的声道的所述数目的所述指示包括获得所述位流中针对所述层中的至少一者指定的背景声道的数目的指示，

其中获得所述层包括基于背景声道的所述数目的所述指示获得所述位流的所述层中的所述至少一者的所述背景声道。

22. 根据权利要求19所述的方法，其中获得在所述位流中指定的声道的所述数目的所述指示包括基于在获得所述层中的至少一者之后在所述位流中剩余的声道的数目，对所述位流中针对所述层中的所述至少一者指定的前景声道的数目的指示进行解析，

其中获得所述层包括基于前景声道的所述数目的所述指示获得所述层中的所述至少一者的所述前景声道。

23. 根据权利要求22所述的方法，其中在获得所述层中的所述至少一者之后在所述位流中剩余的声道的所述数由语法元素表示。

24. 根据权利要求19所述的方法，其中获得在所述位流中指定的声道的所述数目的所述指示包括基于在获得所述层中的至少一者之后的声道的数目，对所述位流中针对所述层中的至少一者指定的背景声道的数目的指示进行解析，

其中获得所述层包括基于背景声道的所述数目的所述指示从所述位流获得所述层中的所述至少一者的所述背景声道。

25. 根据权利要求24所述的方法，其中在获得所述层中的所述至少一者之后在所述位流中剩余的声道的所述数由语法元素表示。

26. 根据权利要求19所述的方法，

其中所述位流的所述层包括基础层和增强层，且

其中所述方法进一步包括对于所述基础层的一个或多个声道应用相关变换以获得所述高阶立体混响音频信号的背景分量的相关表示。

27. 根据权利要求26所述的方法，其中所述相关变换包括逆UHJ变换。

28. 根据权利要求26所述的方法,其中所述相关变换包括逆模式矩阵变换。

29. 根据权利要求19所述的方法,其中所述位流的所述层中的每一者的声道的数目是固定的。

30. 一种经配置以对表示高阶立体混响音频信号的位流进行解码的设备,所述设备包括:

用于存储所述位流的装置;

用于从所述位流获得在所述位流中指定的传输信道的总数目的指示的装置;

用于从所述位流获得在所述位流中指定的层的数目的指示的装置;以及

用于基于在所述位流中指定的层的数目的指示和在所述位流中指定的传输信道的总数目的指示获得所述位流的所述层的装置,包括:

基于包括在所述层的位流中的语法元素,获得用于每个层的传输信道的数目的指示;

其中当剩余的传输信道的数目大于一个时,迭代获得所述层的传输信道的数目的指示,每次迭代计算出的剩余传输信道的数目为所述传输信道的总数目减去已获得的传输信道的数目的累计和;

其中在每次迭代中,使用基于剩余传输信道的数目计算出的位数,在位流中编码相应层的传输信道的数目的指示;

其中通过在每次迭代时递增计数器来获得在所述位流中指定的层的数目的指示。

31. 一种具有存储于其上的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个或多个处理器:

从位流获得在所述位流中指定的传输信道的总数目的指示;

从位流获得在所述位流中指定的层的数目的指示;以及

基于在所述位流中指定的层的数目的指示和在所述位流中指定的传输信道的总数目的指示获得所述位流的所述层;

其中所述指令致使一个或多个处理器获得在所述位流中指定的层的数目的指示包括:

基于包括在所述层的位流中的语法元素,获得用于每个层的传输信道的数目的指示;

其中当剩余的传输信道的数目大于一个时,迭代获得所述层的传输信道的数目的指示,每次迭代计算出的剩余传输信道的数目为所述传输信道的总数目减去已获得的传输信道的数目的累计和;

其中在每次迭代中,使用基于剩余传输信道的数目计算出的位数,在位流中编码相应层的传输信道的数目的指示;

其中通过在每次迭代时递增计数器来获得在所述位流中指定的层的数目的指示。

以信号表示用于高阶立体混响音频数据的可缩放译码的层

- [0001] 本申请案主张以下临时申请案的权益：
- [0002] 2014年10月10日申请的标题为“高阶立体混响音频数据的可缩放译码 (SCALABLE CODING OF HIGHER ORDER AMBISONIC AUDIO DATA)”的第62/062,584号美国临时申请案；
- [0003] 2014年11月25日申请的标题为“高阶立体混响音频数据的可缩放译码 (SCALABLE CODING OF HIGHER ORDER AMBISONIC AUDIO DATA)”的第62/084,461号美国临时申请案；
- [0004] 2014年12月3日申请的标题为“高阶立体混响音频数据的可缩放译码 (SCALABLE CODING OF HIGHER ORDER AMBISONIC AUDIO DATA)”的第62/087,209号美国临时申请案；
- [0005] 2014年12月5日申请的标题为“高阶立体混响音频数据的可缩放译码 (SCALABLE CODING OF HIGHER ORDER AMBISONIC AUDIO DATA)”的第62/088,445号美国临时申请案；
- [0006] 2015年4月10日申请的标题为“高阶立体混响音频数据的可缩放译码 (SCALABLE CODING OF HIGHER ORDER AMBISONIC AUDIO DATA)”的第62/145,960号美国临时申请案；
- [0007] 2015年6月12日申请的标题为“高阶立体混响音频数据的可缩放译码 (SCALABLE CODING OF HIGHER ORDER AMBISONIC AUDIO DATA)”的第62/175,185号美国临时申请案；
- [0008] 2015年7月1日申请的标题为“降低高阶立体混响 (HOA) 背景声道之间的相关 (REDUCING CORRELATION BETWEEN HIGHER ORDER AMBISONIC (HOA) BACKGROUND CHANNELS)”的第62/187,799号美国临时申请案，以及
- [0009] 2015年8月25日申请的标题为“传输经译码可缩放音频数据 (TRANSPORTING CODED SCALABLE AUDIO DATA)”的第62/209,764号美国临时申请案，
- [0010] 所述申请案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

- [0011] 本发明涉及音频数据，且更具体来说，涉及高阶立体混响音频数据的可缩放译码。

背景技术

- [0012] 高阶立体混响 (HOA) 信号 (通常由多个球谐系数 (SHC) 或其它阶层元素表示) 是声场的三维表示。HOA或SHC表示可以独立于用以回放从SHC信号渲染的多声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信号渲染为众所周知的且被广泛采用的多声道格式 (例如，5.1音频声道格式或7.1音频声道格式)。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

- [0013] 一般来说，描述用于高阶立体混响音频数据的可缩放译码的技术。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谐基底函数的至少一个高阶立体混响 (HOA) 系数。所述技术可通过使用例如基础层和一个或多个增强层等多个层对HOA系数进行译码来提供HOA系数的可缩放译码。所述基础层可允许再现由所述HOA系数表示的声场，所述声场可由所述一个或多个增强层增强。换句话说，所述增强层 (与所述基础层组合) 可提供允

许与单独所述基础层相比更完整(或,更准确)地再现所述声场的额外分辨率。

[0014] 在一个方面中,一种装置经配置以对表示高阶立体混响音频信号的位流进行解码。所述装置包括:存储器,其经配置以存储所述位流;以及一个或多个处理器,其经配置以从所述位流获得在所述位流中指定的层的数目的指示,且基于层的所述数目的所述指示而获得所述位流的所述层。

[0015] 在另一方面中,一种对表示高阶立体混响音频信号的位流进行解码的方法,所述方法包括:从所述位流获得在所述位流中指定的层的数目的指示,且基于层的所述数目的所述指示而获得所述位流的所述层。

[0016] 在另一方面中,一种设备经配置以对表示高阶立体混响音频信号的位流进行解码。所述设备包括:用于存储所述位流的装置;用于从所述位流获得在所述位流中指定的层的数目的指示的装置;以及用于基于层的所述数目的所述指示而获得所述位流的所述层的装置。

[0017] 在另一方面中,一种非暂时性计算机可读存储媒体上存储有指令,所述指令在执行时致使一个或多个处理器从位流获得在所述位流中指定的层的数目的指示,且基于层的所述数目的所述指示而获得所述位流的所述层。

[0018] 在另一方面中,一种装置经配置以对高阶立体混响音频信号进行编码以产生位流。所述装置包括:存储器,其经配置以存储所述位流;以及一个或多个处理器,其经配置以在所述位流中指定层的数目的指示,且输出包含所述层的所述所指示数目的所述位流。

[0019] 在另一方面中,一种产生表示高阶立体混响音频信号的位流的方法,所述方法包括:在所述位流中指定层的数目的指示;以及输出包含所述层的所述所指示数目的所述位流。

[0020] 在另一方面中,一种装置经配置以对表示高阶立体混响音频信号的位流进行解码。所述装置包括:存储器,其经配置以存储所述位流;以及一个或多个处理器,其经配置以从所述位流获得在所述位流中的一个或多个层中指定的声音的数目的指示,且基于声音的所述数目的所述指示而获得在所述位流中的所述一个或多个层中指定的所述声音。

[0021] 在另一方面中,一种对表示高阶立体混响音频信号的位流进行解码的方法,所述方法包括:从所述位流获得在所述位流中的一个或多个层中指定的声音的数目的指示;以及基于声音的所述数目的所述指示而获得在所述位流中的所述一个或多个层中指定的所述声音。

[0022] 在另一方面中,一种装置经配置以对表示高阶立体混响音频信号的位流进行解码。所述装置包括:用于从所述位流获得在所述位流中的一个或多个层中指定的声音的数目的指示的装置;以及用于基于声音的所述数目的所述指示而获得在所述位流中的所述一个或多个层中指定的所述声音的装置。

[0023] 在另一方面中,一种非暂时性计算机可读存储媒体上存储有指令,所述指令在执行时致使一个或多个处理器从表示高阶立体混响音频信号的位流获得在所述位流中的一个或多个层中指定的声音的数目的指示,且基于声音的所述数目的所述指示而获得在所述位流中的所述一个或多个层中指定的所述声音。

[0024] 在另一方面中,一种装置经配置以对高阶立体混响音频信号进行编码以产生位流。所述装置包括:一个或多个处理器,其经配置以在所述位流中指定在所述位流的一个或

多个层中指定的声音的数目的指示,且在所述位流的所述一个或多个层中指定所述声道的所述所指示数目;以及存储器,其经配置以存储所述位流。

[0025] 在另一方面中,一种对高阶立体混响音频信号进行编码以产生位流的方法,所述方法包括:在所述位流中指定在所述位流的一个或多个层中指定的声音的数目的指示;以及在所述位流的所述一个或多个层中指定所述声道的所述所指示数目。

[0026] 在附图和以下描述中陈述所述技术的一个或多个方面的细节。所述技术的其它特征、目标和优点将从所述描述和附图以及权利要求而显而易见。

附图说明

[0027] 图1是说明具有各种阶数及子阶数的球谐基底函数的图。

[0028] 图2是说明可执行本发明中所描述的技术的各种方面的系统的图。

[0029] 图3是更详细说明可执行本发明中描述的技术的各种方面的图2的实例中所示的音频编码装置的一个实例的框图。

[0030] 图4是更详细地说明图2的音频解码装置的框图。

[0031] 图5是更详细地说明在经配置以执行本发明中描述的可缩放音频译码技术的潜在版本中的第一者时的图3的位流产生单元的图。

[0032] 图6是更详细地说明在经配置以执行本发明中描述的可缩放音频解码技术的潜在版本中的第一者时的图4的提取单元的图。

[0033] 图7A到7D是说明音频编码装置在产生高阶立体混响(HOA)系数的经编码双层表示时的实例操作的流程图。

[0034] 图8A和8B是说明音频编码装置在产生HOA系数的经编码三层表示时的实例操作的流程图。

[0035] 图9A和9B是说明音频编码装置在产生HOA系数的经编码四层表示时的实例操作的流程图。

[0036] 图10是说明根据所述技术的各种方面的在位流中指定的HOA配置对象的实例的图。

[0037] 图11是说明由位流产生单元针对第一层和第二层产生的边带信息的图。

[0038] 图12A和12B是说明根据本发明中描述的技术的可缩放译码方面产生的边带信息的图。

[0039] 图13A和13B是说明根据本发明中描述的技术的可缩放译码方面产生的边带信息的图。

[0040] 图14A和14B是说明音频编码装置在执行本发明中描述的技术的各种方面时的实例操作的流程图。

[0041] 图15A和15B是说明音频解码装置在执行本发明中描述的技术的各种方面时的实例操作的流程图。

[0042] 图16是说明根据本发明中描述的技术的各种方面的如由图16的实例中所示的位流产生单元执行的可缩放音频译码的图。

[0043] 图17是其中语法元素指示存在两个层的实例的概念图,其中在基础层中指定四个经编码环境HOA系数,且在增强层中指定两个经编码前景信号。

[0044] 图18是更详细地说明在经配置以执行本发明中描述的可缩放音频译码技术的潜在版本中的第二者时的图3的位流产生单元的图。

[0045] 图19是更详细地说明在经配置以执行本发明中描述的可缩放音频解码技术的潜在版本中的第二者时的图3的提取单元的图。

[0046] 图20是说明图18的位流产生单元和图19的提取单元可借以执行本发明中描述的技术的潜在版本中的第二者的第二使用情况的图。

[0047] 图21是其中语法元素指示存在三个层的实例的概念图,其中在基础层中指定两个经编码环境HOA系数,在第一增强层中指定两个经编码前景信号,且在第二增强层中指定两个经编码前景信号。

[0048] 图22是更详细地说明在经配置以执行本发明中描述的可缩放音频译码技术的潜在版本中的第三者时的图3的位流产生单元的图。

[0049] 图23是更详细地说明在经配置以执行本发明中描述的可缩放音频解码技术的潜在版本中的第三者时的图4的提取单元的图。

[0050] 图24是根据本发明中描述的技术的说明音频编码装置可借以在多层位流中指定多个层的第三使用情况的图。

[0051] 图25是其中语法元素指示存在三个层的实例的概念图,其中在基础层中指定两个经编码前景信号,在第一增强层中指定两个经编码前景信号,且在第二增强层中指定两个经编码前景信号。

[0052] 图26是根据本发明中描述的技术的说明音频编码装置可借以在多层位流中指定多个层的第三使用情况的图。

[0053] 图27和28是说明可经配置以执行本发明中描述的技术的各种方面的可缩放位流产生单元和可缩放位流提取单元的框图。

[0054] 图29表示表示可经配置以根据本发明中描述的技术的各种方面而操作的编码器的概念图。

[0055] 图30是更详细地说明图27的实例中所示的编码器的图。

[0056] 图31是说明可经配置以根据本发明中描述的技术的各种方面而操作的音频解码器的框图。

具体实施方式

[0057] 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“声道”式的,这是因为其以某些几何坐标隐含地指定到扩音器的馈入。消费型环绕声格式包含普遍的5.1格式(其包含以下六个声道: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的扬声器(成对称和非对称几何形状),其常常称为“环绕阵列”。此阵列的一个实例包含被安置在截角二十面体的拐角上的坐标上的32个扩音器。

[0058] 到未来MPEG编码器的输入视情况为三个可能格式中的一者:(i)传统的基于声道的音频(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(以及其它信息)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,其涉及使用球谐基底函数的系数(也称为“球谐系数”或SHC、“高阶立体混响”或HOA及“HOA系数”)来表示声场。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 3DAudio)”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。

[0059] 在市场中存在各种“环绕声”基于声道的格式。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方面已获得最大成功)到NHK(日本广播协会或日本广播公司)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音轨,而不用花费精力来针对每个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,其可调适且不知晓回放位置(涉及渲染器)处的扬声器几何布置(和数目)以及声学条件。

[0060] 为向内容创建者提供此种灵活性,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一组元素。在所述集合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。

[0061] 分层要素集合的一个实例是球谐系数(SHC)的集合。以下表达式示范使用SHC对声场的描述或表示:

$$[0062] p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t},$$

[0063] 所述表达式展示,声场的任一点 $\{r_r, \theta_r, \varphi_r\}$ 在时间t的压力 p_i 可由SHC, $A_n^m(k)$ 唯一地表示。此处, $k = \frac{\omega}{c}$, c为声速(约343m/s), $\{r_r, \theta_r, \varphi_r\}$ 为参考点(或观测点), $j_n(\cdot)$ 为阶n的球贝塞耳函数(spherical Bessel function),且 $Y_n^m(\theta_r, \varphi_r)$ 为阶n和子阶m的球谐基底函数。可辨识到,方括号中的项为信号的频域表示(即, $S(\omega, r_r, \theta_r, \varphi_r)$),其可由各种时频变换近似表示,例如离散傅里叶变换(DFT)、离散余弦变换(DCT)或小波变换。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。

[0064] 图1是说明从零阶($n=0$)到四阶($n=4$)的球谐基底函数的图。如可看出,对于每一阶,存在子阶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但未明确提到。

[0065] SHC $A_n^m(k)$ 可以通过各种麦克风阵列配置以物理方式来获取(例如,记录),或替代地,它们可以从声场的基于声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可促成更有效的发射或存储。举例来说,可使用涉及 $(1+4)^2$ 个(25,且因此为四阶)系数的四阶表示。

[0066] 如上文所指出,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1004-1025页)中。

[0067] 为说明如何可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:

[0068] $A_n^m(k) = g(\omega)(-4\pi ik)h_n^{(2)}(kr_s)Y_n^m(\theta_s, \varphi_s),$

[0069] 其中 i 为 $\sqrt{-1}$, $h_n^{(2)}(\cdot)$ 为阶数为n的(第二种)球面汉克尔函数(spherical Hankel function),且 $\{r_s, \theta_s, \varphi_s\}$ 为对象的位置。根据频率知晓对象源能量 $g(\omega)$ (例如,使用时间-频率分析技术,例如对PCM流执行快速傅立叶变换)允许我们将每一PCM对象及对应位置转换为SHC $A_n^m(k)$ 。另外,可以看出(由于上式为线性正交分解),每一对象的 $A_n^m(k)$ 系数是累加性的。以此方式,PCM对象的量值可由 $A_n^m(k)$ 系数表示(例如,表示为个别对象的系数向量的和)。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),且上式表示从个别对象到观测点 $\{r_r, \theta_r, \varphi_r\}$ 附近的总体声场表示的变换。下文在基于对象和基于SHC的音频译码的上下文中描述其余各图。

[0070] 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的实例所示,系统10包含内容创建者装置12和内容消费者装置14。虽然在所述内容创建者装置12和所述内容消费者装置14的情形中描述,但是所述技术可在声场的SHC(其也可被称作HOA系数)或任何其它分层表示被编码以形成表示音频数据的位流的任何情形中实现。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,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)。

[0071] 内容创建者装置12可通过电影工作室或可产生用于由内容消费者装置,例如内容消费者装置14的操作者消费的多声道音频内容的其它实体来操作。在一些实例中,内容创建者装置12可由希望压缩HOA系数11的个人用户操作。通常,内容创建者可产生音频内容以及视频内容。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,其可指能够渲染SHC以用于作为多声道音频内容回放的任何形式的音频回放系统。

[0072] 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可获得各种格式(包含直接作为HOA系数)的实时记录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可俘获实时记录7。在编辑过程期间,内容创建者可渲染来自音频对象9的HOA系数11,收听所渲染的扬声器馈给以试图识别需要进一步编辑的声场的各个方面。内容创建者装置12可接着编辑HOA系数11(可能经由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可使用音频编辑系统18来产生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个或多个源球谐系数的任何系统。

[0073] 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。即,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)发射(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(其可称为旁侧声道信息)。

[0074] 虽然在图2中展示为直接发射到内容消费者装置14,但内容创建者装置12可将音频位流21输出到定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。所述中间装置可存储音频位流21以用于稍后递送到可能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(且可能结合发射对应视频数据位流)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。

[0075] 或者,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,传输信道可指借以发射存储到媒体的内容的信道(且可包含零售商店及其它基于商店的递送机制)。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。

[0076] 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可表示能够回放多声道音频数据的任何音频回放系统。音频回放系统16可包含数个不同渲染器22。渲染器22可各自提供不同形式的渲染,其中不同形式的渲染可包含执行向量基振幅平移(VBAP)的各种方式中的一者或多者及/或执行声场合成的各种方式中的一者或多者。如本文所使用,“A和/或B”意味着“A或B”,或“A和B”两者。

[0077] 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可表示经配置以对来自位流21的HOA系数11'进行解码的装置,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可在对位流21进行解码之后获得HOA系数11'并渲染HOA系数11'以输出扩音器馈给25。扩音器馈给25可驱动一个或多个扩音器(其为便于说明的目的而未在图2的实例中展示)。

[0078] 为了选择适当渲染器或在一些情况下产生适当渲染器,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。在一些情况下,音频回放系统16可使用参考麦克风且以使得动态地确定扩音器信息13的方式驱动扩音器来获得扩音器信息13。在其它情况下或结合扩音器信息13的动态确定,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。

[0079] 音频回放系统16可随后基于扩音器信息13选择音频渲染器22中的一者。在一些情况下,在音频渲染器22中无一者处于到扩音器信息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产生音频渲染器22中的一者。音频回放系统16可在一些情况下基于扩音器信息13产生音频渲染器22中的一者,而无需首先试图选择音频渲染器22中的现有者。一个或多个扬声器3接着可回放经渲染扩音器馈给25。换句话说,扬声器3可经配置以基于高阶立体混响音频数据再现声场。

[0080] 图3是更详细地说明可执行本发明中描述的技术的各种方面的图2的实例中所展示的音频编码装置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、基于向量的分解单元27和基于方向的分解单元28。

[0081] 尽管下文简要描述,但关于基于向量的分解单元27以及压缩HOA系数的各个方面的较多信息在2014年5月29日申请的标题为“声场的经分解表示的内插(INTERPOLATION

FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD)"的第WO 2014/194099号国际专利申请案中可用。此外,根据MPEG-H 3D音频标准压缩HOA系数的各种方面的更多细节(包含下方概述的基于向量的分解的论述)可见于:

[0082] 标题为“信息技术-异质环境中的高效率译码和媒体递送-第3部分:3D音频(Information technology-High efficiency coding and media delivery in heterogeneous environments-Part 3:3D audio)”的ISO/IEC DIS 23008-3文档,ISO/IEC JTC 1/SC 29/WG11,2014年7月25日(可见于:<http://mpeg.chiariglione.org/standards/mpeg-h/3d-audio/dis-mpeg-h-3d-audio>,下文称为“MPEG-H 3D音频标准阶段I”);

[0083] 标题为“信息技术-异质环境中的高效率译码和媒体递送-第3部分:3D音频,修正3:MPEG-H 3D音频阶段2(Information technology-High efficiency coding and media delivery in heterogeneous environments-Part 3:3D audio,AMENDMENT 3:MPEG-H 3DAudio Phase 2)”的ISO/IEC DIS 23008-3:2015/PDAM 3文档,ISO/IEC JTC 1/SC 29/WG11,2015年7月25日(可见于:<http://mpeg.chiariglione.org/standards/mpeg-h/3d-audio/text-isoiec-23008-3201xpdam-3-mpeg-h-3d-audio-phase-2>,且下文称为“MPEG-H 3D音频标准阶段II”);以及

[0084] Jürgen Herre等人,标题为“MPEG-H 3D音频-沉浸式3D音效译码新标准(MPEG-H 3D Audio-The New Standard for Coding of Immersive Spatial Audio)”,2015年8月,于IEEE信号处理专题杂志(IEEE Journal of Selected Topics in Signal Processing)第9卷第5期发表。

[0085] 内容分析单元26表示经配置以分析HOA系数11的内容以识别HOA系数11表示从实况记录产生的内容还是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。在一些情况下,当帧式HOA系数11是从记录中产生时,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。在一些情况下,当帧式HOA系数11是从合成音频对象中产生时,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对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。

[0086] 如图3的实例中所示,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、去相关单元60(展示为“去相关单元60”)、增益控制单元62、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,和量化单元52。

[0087] 线性可逆变换(LIT)单元30接收呈HOA声道形式的HOA1111,每一声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(其可表示为HOA[k],其中k可表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)²。

[0088] LIT单元30可表示经配置以执行被称为奇异值分解的形式的分析的单元。虽然相对于SVD加以描述,但可相对于提供数组线性不相关的能量压缩输出的任何类似变换或分解执行本发明中所描述的技术。而且,本发明中对“集合”的参考一般意欲指非零集合(除非特定地相反陈述),且不意欲指包含所谓的“空集合”的集合的经典数学定义。替代变换可包括常常被称作“PCA”的主分量分析。取决于上下文,PCA可由数个不同名称指代,例如离散卡亨南-洛维变换(Karhunen-Loeve transform)、霍特林变换(Hotelling transform)、恰当

正交分解(POD)和本征值分解(EVD),仅举几例。有益于压缩音频数据的潜在基本目标中的一个的这些操作的性质可包含多声道音频数据的“能量压缩”和“去相关”中的一者或多者。

[0089] 在任何情况下,出于实例的目的,假定LIT单元30执行奇异值分解(其再次可被称作“SVD”),LIT单元30可将HOA系数11变换成两个或更多个经变换HOA系数的集合。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可按如下形式表示y乘z实数或复数矩阵X(其中X可表示多声道音频数据,例如HOA系数11)的因子分解:

[0090] $X = USV^*$

[0091] U可表示y乘y实数或复数单位矩阵,其中U的y列被称为多声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声道音频数据的奇异值。 V^* (其可表示V的共轭转置)可表示z乘z实数或复数单位矩阵,其中 V^* 的z列被称为多声道音频数据的右奇异向量。

[0092] 在一些实例中,以上提及的SVD数学表达式中的 V^* 矩阵表示为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,假定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,尽管在本发明中表示为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。尽管假定为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。

[0093] 以此方式,LIT单元30可相对于HOA系数11执行SVD,以输出具有维度D: $M \times (N+1)^2$ 的US[k]向量33(其可表示S向量与U向量的组合版本)以及具有维度D: $(N+1)^2 \times (N+1)^2$ 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)。

[0094] U、S和V矩阵的分析可展现所述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。在(具有长度M个样本的)U中的N个向量中的每一个可表示随时间(由M个样本表示的时间周期)而变的彼此正交且已从任何空间特性(其也可被称作方向信息)解耦的经正规化的分开音频信号。表示空间形状和位置(r, θ, φ)的空间特性可实际上由V矩阵(长度(N+1)²的每一者)中的个别第i向量 $v^{(i)}(k)$ 表示

[0095] $v^{(i)}(k)$ 向量中的每一个的个别元素可表示描述相关联音频对象的声场的形状(包含宽度)和位置的HOA系数。U矩阵和V矩阵中的向量均经正规化而使得其均方根能量等于单位。因此,在U中的音频信号的能量由S中的对角元素表示。将U与S相乘形成US[k](具有个别向量元素 $X_{PS}(k)$),因此表示具有能量的音频信号。SVD分解使音频时间信号(U中)、其能量(S中)与其空间特性(V中)解耦的能力可支持本发明中所描述的技术的各个方面。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”。

[0096] 尽管被描述为直接相关于HOA系数11来执行,但LIT单元30可对HOA系数11的导出项应用线性可逆变换。例如,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)而不是其自身系数来执行SVD,LIT单元30可潜在

降低按照一个或多个处理器循环和存储空间执行SVD的计算复杂度,同时实现相同源音频编码效率,好像SVD被直接应用于HOA系数一样。

[0097] 参数计算单元32表示经配置以计算各种参数的单元,所述参数例如为相关参数(R)、方向特性参数(θ, φ, r)和能量特性(e)。用于当前帧的所述参数中的每一者可表示为R[k]、 $\theta[k]$ 、 $\varphi[k]$ 、r[k]和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)以识别所述参数。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]向量和V[k-1]向量的前一帧而表示为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。

[0098] 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。重新排序单元34可逐轮地比较来自第一US[k]向量33的参数37中的每一个与第二US[k-1]向量33的参数39中的每一个。重新排序单元34可基于当前参数37和先前参数39将US[k]矩阵33和V[k]矩阵35内的各个向量重排序(作为一实例,使用匈牙利算法(Hungarian algorithm))以将重排序的US[k]矩阵33'(其可在数学上表示为 $\bar{U}[k]$)和重排序的V[k]矩阵35'(其可在数学上表示为 $\bar{V}[k]$)输出到前景声音(或占优势声音-PS)选择单元36(“前景选择单元36”)和能量补偿单元38。

[0099] 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可基于所述分析及/或所接收的目标位速率41确定心理声学译码器实例化的总数目(其可为环境或背景声道的总数目(BG_{TOT})及前景声道或换句话说占优势声道的数目的函数)。音质译码器实例化的总数可表示为numHOATransportChannels。

[0100] 再次为了潜在地实现目标位速率41,声场分析单元44还可确定前景声道的总数(nFG)45、背景(或换句话说,环境)声场的最小阶数(N_{BG}或替代地,MinAmbHOAorder)、表示背景声场的最小阶数的实际声道的对应数目(nBGA=(MinAmbHOAorder+1)²)以及要发送的额外BG HOA声道的索引(i)(其在图3的实例中可共同地被表示为背景声道信息43)。背景声道信息42也可被称作环境声道信息43。从numHOATransportChannels-nBGA保持的声音中的每一者可为“额外背景/环境声道”、“起作用的基于向量的占优势声道”、“起作用的基于方向的占优势信号”或“完全未起作用”。一方面,可由两个位将所述声道类型指示为(作为“ChannelType”)语法元素(例如,00:基于方向的信号;01:基于向量的占优势信号;10:另外环境信号;11:未起作用信号)。可由(MinAmbHoaOrder+1)²+索引10(在以上实例中)呈现为用于所述帧的位流中的声道类型的次数给出背景或环境信号的总数nBGA。

[0101] 声场分析单元44可基于所述目标位速率41选择背景(或换句话说,环境)声道的数目和前景(或换句话说,占优势)声道的数目,从而在所述目标位速率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/或前景声道。一方面,在位流的头部部分中,numHOATransportChannels可被设置为8,而MinAmbHOAorder可被设置为1。在此情况下,在每个帧,四个声道可专用于表示所述声场的背景或环境部分,而另4个声道可以在逐帧基础上随声道类型而变化,例如任一者用作另外背景/环境声道或前景/占优势声道。如上所述,所述前景/占优势信号可以为基于向量或基于方向的信号中的任一者。

[0102] 在一些情况下,用于帧的基于向量的占优势信号的总数可由ChannelType索引在

所述帧的位流中为01的次数给出。在以上方面中,对于每个另外背景/环境声道(例如,对应于Channel Type 10),可在声道中表示可能的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5-25的索引。当minAmbHOAorder被设置为1时,前四个环境HOA系数1-4可一直发送,因此,音频编码装置可能仅需要指示具有5-25的索引的另外环境HOA系数中的一者。因此,所述信息可以使用可被表示为“CodedAmbCoeffIdx”的5位语法元素(用于4阶内容)来发送。在任何情况下,声场分析单元44将背景声道信息43和HOA系数11输出到背景(BG)选择单元36,将背景声道信息43输出到系数减少单元46和位流产生单元42,且将nFG45输出到前景选择单元36。

[0103] 背景选择单元48可表示经配置以基于背景声道信息(例如,背景声场(N_{BG})以及要发送的额外BG HOA声道的数目(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可接着选择具有由索引(i)中的一者识别的索引的HOA系数11作为额外BG HOA系数,其中,将nBGa提供给位流产生单元42以在位流21中指定,以便使得音频解码装置,例如图2和4的实例中所示的音频解码装置24,能够从位流21提取背景HOA系数47。背景选择单元48可接着将环境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也可被称作“环境HOA系数47”,其中,环境HOA系数47中的每一个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声道47。

[0104] 前景选择单元36可表示经配置以基于nFG 45(其可表示识别前景向量的一个或多个索引)选择表示声场的前景或不同分量的重排序的US[k]矩阵33'及重排序的V[k]矩阵35'的单元。前景选择单元36可将nFG信号49(其可被表示为重排序的 $US[k]_1, \dots, nFG, 49, FG_{1,\dots,nFG}[k]49$,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 $D:M \times nFG$,且每个nFG信号49表示单声道音频对象。前景选择单元36也可将对应于声场的前景分量的重排序的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,对应于前景分量的重排序的V[k]矩阵35'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_k(其可被数学表示为 $\bar{V}_{1,\dots,nFG}[k]$)。

[0105] 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,以补偿由于通过背景选择单元48移除HOA声道中的各种HOA声道所造成的能力损失的单元。能量补偿单元38可对于重排序的US[k]矩阵33'、重排序的V[k]矩阵35'、nFG信号49、前景V[k]向量51_k和环境HOA系数47中的一者或者执行能量分析,并且接着基于所述能量分析执行能量补偿以产生能量补偿的环境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60。

[0106] 去相关单元60可表示经配置以实施本发明中描述的技术的各种方面以降低或消除经能量补偿的环境HOA系数47'之间的相关以形成一个或多个去相关环境HOA音频信号67的单元。去相关单元40'可将去相关HOA音频信号67输出到增益控制单元62。增益控制单元62可表示经配置以对于去相关环境HOA音频信号67执行自动增益控制(其可简称为“AGC”)以获得增益受控环境HOA音频信号67'。在应用增益控制之后,自动增益控制单元62可将增益受控环境HOA音频信号67'提供到心理声学音频译码器单元40。

[0107] 包含在音频编码装置20内的去相关单元60可表示经配置以将一个或多个去相关变换应用于经能量补偿的环境HOA系数47'以获得去相关HOA音频信号67的单元的单个或多个例项。在一些实例中,去相关单元40'可将UHJ矩阵应用于经能量补偿的环境HOA系数47'。在本发明的各种实例中,UHJ矩阵也可以被称作“基于相位的变换”。应用基于相位的变换还可以在本文中称为“相移去相关”。

[0108] 立体混响UHJ格式为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格式包含系统阶层,其中所记录声场将根据可用声道而变化的准确度再现。在各种情况下,UHJ也称为“C格式”。所述缩写指示并入系统中的源中的一些:U来自Universal (UD-4);H来自矩阵H;且J来自系统45J。

[0109] UHJ为立体混响技术内的对方向声音信息进行编码和解码的阶层式系统。取决于可用声道的数目,系统可携载较多或较少信息。UHJ与立体声和单声道完全兼容。可使用高达四个声道(L,R,T,Q)。

[0110] 在一种形式中,可通过可使用收听端处的UHJ解码器恢复的正常立体声信号声道(CD、FM或数字无线电,等)携载2声道(L,R)UHJ、水平(或“平面”)环绕信息。将两个声道求和可得到兼容的单声道信号,其可为比对常规“声相式单声道”源进行求和更准确的双声道版本表示。如果第三声道(T)可用,则在经由3声道UHJ解码器解码时,可使用第三声道来得到平面环绕效果的改善的定位准确度。可能并不需要第三声道为此目的而具有全音频带宽,从而导致所谓的“ $2\frac{1}{2}$ 声道”系统的可能性,其中第三声道是带宽受限的。在一个实例中,限制可为5kHz。第三声道可经由FM无线电(例如借助于相位正交调制)来广播。将第四声道(Q)添加到UHJ系统可允许对具有高度(有时称为n,如Periphony)的全环绕声音进行编码,其中准确度等同于4声道B格式。

[0111] 2声道UHJ为常用于分布立体混响记录的格式。可经由所有正常立体声声道来发射2声道UHJ记录,且正常2声道媒体中的任一者可以使用而不加更改。UHJ与立体声兼容,兼容之处在于,无需解码,收听者即可察觉立体声像,但其比常规立体声(例如,所谓的“超立体声”)显著较宽。左声道与右声道也可以求和以获得非常高度的单声道兼容性。经由UHJ解码器重放,可展现环绕能力。

[0112] 应用UHJ矩阵(或基于相位的变换)的去相关单元60的实例数学表示如下:

[0113] UHJ编码:

[0114] $S = (0.9397*W) + (0.1856*X)$;

[0115] $D = \text{imag}(\text{hilbert}((-0.3420*W) + (0.5099*X))) + (0.6555*Y)$;

[0116] $T = \text{imag}(\text{hilbert}((-0.1432*W) + (0.6512*X))) - (0.7071*Y)$;

[0117] $Q = 0.9772*Z$;

[0118] S与D到左与右的转换:

[0119] $\text{Left} = (S+D)/2$

[0120] $\text{Right} = (S-D)/2$

[0121] 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背景声道是1阶立体混响,FuMa经正规化,按照立体混响声道编号次序W(a00)、X(a11)、Y(a11-)、Z(a10)。

[0122] 在上文所列的计算中,去相关单元40'可执行各种矩阵乘以常数值的标量乘法。举

例来说,为获得S信号,去相关单元60可执行W矩阵乘以常数值0.9397(例如,标量乘法)以及X矩阵乘以常数值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60可应用希尔伯特变换(Hilbert transform,由以上UHJ编码中的“Hilbert()”函数表示)来获得D和T信号中的每一个。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)。

[0123] 应用UHJ矩阵(或基于相位的变换)的去相关单元60的另一实例数学表示如下:

[0124] UHJ编码:

[0125] $S = (0.9396926*W) + (0.151520536509082*X)$;

[0126] $D = \text{imag}(\text{hilbert}((-0.3420201*W) + (0.416299273350443*X))) + (0.535173990363608*Y)$;

[0127] $T = 0.940604061228740 * (\text{imag}(\text{hilbert}((-0.1432*W) + (0.531702573500135*X))) - (0.577350269189626*Y))$;

[0128] $Q = Z$;

[0129] S与D到左与右的转换:

[0130] $\text{Left} = (S+D)/2$;

[0131] $\text{Right} = (S-D)/2$;

[0132] 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背景声道是1阶立体混响,N3D(或“全三维”)经正规化,按照立体混响声道编号次序W(a00)、X(a11)、Y(a11-)、Z(a10)。尽管本文中相对于N3D正规化进行描述,但将了解,实例计算还可以应用于经SN3D正规化(或“施密特半正规化”)的HOA背景声道。N3D与SN3D正规化可在所使用的缩放因数方面不同。相对于SN3D正规化的N3D正规化的实例表示如下表达为:

$$[0133] N_{l,m}^{N3D} = N_{l,m}^{SN3D} \sqrt{2l+1}$$

[0134] 在SN3D正规化中所使用的加权系数的实例如下表达为:

$$[0135] N_{l,m}^{SN3D} = \sqrt{\frac{2 - \delta m (l - |m|)!}{4\pi (l + |m|)!}}, \delta m \begin{cases} 1 & \text{如果 } m = 0 \\ 0 & \text{如果 } m \neq 0 \end{cases}$$

[0136] 在上文所列的计算中,去相关单元60可执行各种矩阵与常数值的标量乘法。举例来说,为获得S信号,去相关单元60可执行W矩阵与常数值0.9396926(例如,通过标量乘法)以及X矩阵与常数值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60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()”函数或相移去相关表示)。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)。

[0137] 去相关单元60可执行上文所列的计算,使得所得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)。在一些此类情形中,去相关单元60可输出T信号和Q信号作为去相关环境HOA音频信号67的部分,但接收位流21的解码装置可能在显现为立体声扬声器几何布置(或,换句话说,立体声扬声器配置)时不处理T信号和Q信号。在诸实例中,环境HOA系数47'可表示待显现在单声道音频再现系统上的声场。去相关单元60可输出S信号和D信号作为去相关环境HOA音频信号67的部分,且接收位流21的解码装置可组合

(或“混合”)S信号与D信号以形成待以单声道音频格式显现及/或输出的音频信号。

[0138] 在这些实例中,解码装置及/或再现装置可以各种方式恢复单声道音频信号。一个实例为通过混合左信号与右信号(由S信号和D信号表示)。另一实例为通过应用UHJ矩阵(或基于相位的变换)来对W信号进行解码。通过应用UHJ矩阵(或基于相位的变换)产生呈S信号和D信号形式的自然左信号和自然右信号,去相关单元60可实施本发明的技术以提供优于应用其它去相关变换(例如在MPEG-H标准中描述的模式矩阵)的技术的潜在优势及/或潜在改善。

[0139] 在各种实例中,去相关单元60可基于所接收的经能量补偿的环境HOA系数47'的位速率而应用不同的去相关变换。举例来说,在经能量补偿的环境HOA系数47'表示四声道输入的情形中,去相关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,基于表示四声道输入的经能量补偿的环境HOA系数47',去相关单元60可应用 4×4 UHJ矩阵(或基于相位的变换)。举例来说, 4×4 矩阵可与经能量补偿的环境HOA系数47'的四声道输入正交。换句话说,在经能量补偿的环境HOA系数47'表示较小数目声道(例如,四个)的情况下,去相关单元60可应用UHJ矩阵作为所选去相关变换,以使经能量补偿的环境HOA信号47'的背景信号去相关以获得去相关环境HOA音频信号67。

[0140] 根据此实例,如果经能量补偿的环境HOA系数47'表示较大数目个声道(例如,九个),那么去相关单元60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,在经能量补偿的环境HOA系数47'表示九声道输入的情形中,去相关单元60可应用模式矩阵(例如,如以上提及的MPEG-H 3D音频标准阶段I中所描述),以使经能量补偿的环境HOA系数47'去相关。在经能量补偿的环境HOA系数47'表示九声道输入的实例中,去相关单元60可应用 9×9 模式矩阵以获得去相关环境HOA音频信号67。

[0141] 音频编码装置20的各种组件(例如心理声学音频译码器40)又可在感知上根据AAC或USAC对去相关环境HOA音频信号67进行译码。去相关单元60可应用相移去相关变换(例如,在四声道输入情况下,UHJ矩阵或基于相位的变换),以潜在地优化对于HOA的AAC/USAC译码。在经能量补偿的环境HOA系数47'(且由此,去相关环境HOA音频信号67)表示待显现在立体声再现系统上的音频数据的实例中,去相关单元60可基于AAC与USAC为相对定向(或针对其优化)的立体声音频数据而应用本发明的技术以改善或优化压缩。

[0142] 将理解,在经能量补偿的环境HOA系数47'包含前景声道的情境中,以及在经能量补偿的环境HOA系数47'不包含任何前景声道的情境中,去相关单元60可应用本文中所描述的技术。作为一个实例,在经能量补偿的环境HOA系数47'包含零(0)个前景声道和四个(4)个背景声道的情形(例如,更低/更小位速率的情形)中,去相关单元40'可应用上文所描述的技术及/或计算。

[0143] 在一些实例中,去相关单元60可致使位流产生单元42用信号发送指示去相关单元60将去相关变换应用于经能量补偿的环境HOA系数47'的一个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60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60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。

[0144] 去相关单元60可将基于相位的变换应用于经能量补偿的环境HOA系数47'。用于

$C_{AMB}(k-1)$ 的 O_{MIN} HOA 系数序列的基于相位的变换由下式定义：

$$[0145] \quad \begin{bmatrix} x_{AMB,LOW,1}(k-2) \\ x_{AMB,LOW,2}(k-2) \\ x_{AMB,LOW,3}(k-2) \\ x_{AMB,LOW,4}(k-2) \end{bmatrix} = \begin{bmatrix} d(9) \cdot (S(k-2) + M(k-2)) \\ d(9) \cdot (M(k-2) - S(k-2)) \\ d(8) \cdot (B_{+90}(k-2) + d(5) \cdot c_{AMB,2}(k-2)) \\ c_{AMB,3}(k-2) \end{bmatrix},$$

[0146] 其中系数 d 如表1中所定义，信号帧 $S(k-2)$ 和 $M(k-2)$ 由下式定义：

$$[0147] \quad S(k-2) = A_{+90}(k-2) + d(6) \cdot c_{AMB,2}(k-2)$$

$$[0148] \quad M(k-2) = d(4) \cdot c_{AMB,1}(k-2) + d(5) \cdot c_{AMB,4}(k-2)$$

[0149] 且 $A_{+90}(k-2)$ 和 $B_{+90}(k-2)$ 是 +90 度相移信号 A 和 B 的帧，由下式定义：

$$[0150] \quad A(k-2) = d(0) \cdot c_{AMB,LOW,1}(k-2) + d(1) \cdot c_{AMB,4}(k-2)$$

$$[0151] \quad B(k-2) = d(2) \cdot c_{AMB,LOW,1}(k-2) + d(3) \cdot c_{AMB,4}(k-2)。$$

[0152] 用于 $C_{P,AMB}(k-1)$ 的第一 O_{MIN} HOA 系数序列的基于相位的变换相应地定义。所描述变换可能引入一个帧的延迟。

[0153] 在上文中， $x_{AMB,LOW,1}(k-2)$ 到 $x_{AMB,LOW,4}(k-2)$ 可对应于去相关环境 HOA 音频信号 67。在以上方程式中，变量 $C_{AMB,1}(k)$ 变量表示对应于 (阶数: 子阶数) 为 (0:0) 的球基函数的第 k 个帧的 HOA 系数，其也可称为“W”声道或分量。变量 $C_{AMB,2}(k)$ 变量表示对应于 (阶数: 子阶数) 为 (1:-1) 的球基函数的第 k 个帧的 HOA 系数，其也可称为“Y”声道或分量。变量 $C_{AMB,3}(k)$ 变量表示对应于 (阶数: 子阶数) 为 (1:0) 的球基函数的第 k 个帧的 HOA 系数，其也可称为“Z”声道或分量。变量 $C_{AMB,4}(k)$ 变量表示对应于 (阶数: 子阶数) 为 (1:1) 的球基函数的第 k 个帧的 HOA 系数，其也可称为“X”声道或分量。 $C_{AMB,1}(k)$ 到 $C_{AMB,3}(k)$ 可对应于环境 HOA 系数 47'。

[0154] 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。

[0155]

n	d(n)
0	0.342020099999999999
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

[0156] 表1用于基于相位的变换的系数

[0157] 在一些实例中，音频编码装置 20 的各种组件（例如位流产生单元 42）可经配置以仅发射一阶 HOA 表示以获得较低目标位速率（例如，128K 或 256K 的目标位速率）。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）可经配置以舍弃高阶 HOA 系数（例如，阶数大于一阶的系数，或换句话说， $N > 1$ ）。然而，在音频编码装置 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）可分开前景与背景声道，且可将位（例如，以较大量）指派到前景声道。

[0158] 尽管描述为应用于经能量补偿的环境HOA系数47'，但音频编码装置20可不将去相关应用于经能量补偿的环境HOA系数47'。替代地，能量补偿单元38可将经能量补偿的环境HOA系数47'直接提供到增益控制单元62，所述增益控制单元可执行对于经能量补偿的环境HOA系数47'的自动增益控制。由此，去相关单元60展示为虚线以指示去相关单元可能并不始终执行去相关或包含于音频解码装置20中。

[0159] 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_k 和前一帧(因此为k-1记法)的前景V[k-1]向量 51_{k-1} ，并执行空间-时间内插以产生内插前景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_k 重新组合以恢复重排序的前景HOA系数。空间-时间内插单元50可接着将重排序的前景HOA系数除以经内插V[k]向量以产生经内插nFG信号49'。

[0160] 空间-时间内插单元50还可输出所述前景V[k]向量 51_k ，其用于产生经内插前景V[k]向量，使得例如音频解码装置24等音频解码装置可产生经内插前景V[k]向量并由此恢复前景V[k]向量 51_k 。用以产生经内插前景V[k]向量的前景V[k]向量 51_k 被表示为其余前景V[k]向量53。为确保在编码器和解码器处使用所述相同V[k]和V[k-1](以创建内插向量V[k])，可在编码器和解码器处使用向量的经量化/解量化版本。空间-时间内插单元50可将经内插nFG信号49'输出到增益控制单元62，且将经内插前景V[k]向量 51_k 输出到系数减少单元46。

[0161] 增益控制单元62还可表示经配置以对于经内插nFG信号49'执行自动增益控制(其可简称为“AGC”)以获得增益受控nFG信号49”的单元。在应用增益控制之后，自动增益控制单元62可将增益受控nFG信号49”提供到心理声学音频译码器单元40。

[0162] 系数减少单元46可表示经配置以基于背景声道信息43对于其余前景V[k]向量53执行系数减少，以将减少前景V[k]向量55输出到量化单元52的单元。减少前景V[k]向量55可具有维度D: $[(N+1)^2 - (N_{BG}+1)^2 - BG_{TOT}] \times nFG$ 。就此而言，系数减少单元46可表示经配置以减少其余前景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前景V[k]向量(其形成其余前景V[k]向量53)中具有极少或没有方向信息的系数的单元。在一些实例中，相异的系数或换句话说对应于一阶和零阶基底函数(其可被表示为 N_{BG})的前景V[k]向量提供极少方向信息，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 $[(NBG+1)^{2+1}, (N+1)^2]$ 识别对应于 N_{BG} 的系数而且识别额外HOA声道(其可由变量TotalOfAddAmbHOAChan表示)。

[0163] 量化单元52可表示经配置以执行任何形式的量化以压缩经减少的前景V[k]向量55来产生经译码前景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(即在此实例中的减少前景V[k]向量55中的一者或多者)的单元。量化单元52可执行在以上提及的MPEG-H3D音频译码标准阶段I或阶段II中所阐述的以下12个量化模式中的任一个。量化单元52还可执行前述类型的量化模式中的任一者的预测版本，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。量化单元52可将经译码前景V[k]向量57提供到位流产生单元42。量化单元52还可提供指示量化模式的语法元素(例如，NbtsQ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。

[0164] 包含在音频编码装置20内的心理声学音频译码器单元40可表示心理声学音频编码器的多个例项，所述多个例项中的每一者用以对经能量补偿的环境HOA系数47'和经内插nFG信号49'中的每一者的不同音频对象或HOA声道进行编码，以产生经编码环境HOA系数59和经编码nFG信号61。心理声学音频译码器单元40可将经编码环境HOA系数59和经编码nFG信号61输出到位流产生单元42。

[0165] 包含在音频编码装置20内的位流产生单元42表示格式化数据以符合已知格式(其可指解码装置已知的格式)，由此产生基于向量的位流21的单元。换句话说，位流21可表示已经以上文所描述的方式编码的经编码音频数据。在一些实例中，位流产生单元42可表示多路复用器，其可接收经译码前景V[k]向量57、经编码环境HOA系数59、经编码nFG信号61和背景声道信息43。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和背景声道信息43产生位流21。以此方式，位流产生单元42可由此指定位流21中的向量57以获得位流21。位流21可包含主要或主位流和一个或多个旁声道位流。

[0166] 尽管在图3的实例中未展示，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当前帧将使用基于方向的合成还是基于向量的合成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中的相应一者的所述切换或当前编码。

[0167] 此外，如上文所提及，声场分析单元44可识别 BG_{TOT} 环境HOA系数47，所述 BG_{TOT} 环境HOA系数可在逐帧基础上改变(但时常 BG_{TOT} 可跨越两个或更多个邻近(在时间上)帧保持恒定或相同)。 BG_{TOT} 的改变可产生在减少前景V[k]向量55中表达的系数的改变。 BG_{TOT} 的改变可产生逐帧改变(尽管有时 BG_{TOT} 可跨越两个或两个以上邻近(在时间上)帧保持恒定或相同)的背景HOA系数(其也可被称作“环境HOA系数”)。所述改变常常导致声场的各方面的能量改变，所述声场通过额外环境HOA系数的添加或移除及系数从减少前景V[k]向量55的对应移除或系数到所述减少前景V[k]向量的添加表示。

[0168] 结果，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，且产生指示环境HOA系数的改变的旗标或其它语法元素(就用于表示声场的环境分量来说)(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”)。确切地说，系数减少单元46可产生旗标(其可表示为AmbCoeffTransition旗标或AmbCoeffIdxTransition旗标)，从而将所述旗标提供到位流产生单元42，使得可将所述旗标包含在位流21中(可能作为旁侧声道信息的部分)。

[0169] 除指定环境系数转变旗标之外，系数减少单元46也可修改产生减少前景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。同样，处于转变中的环境HOA系数可添加至背景系数的 BG_{TOT} 总数目或从背景系数的 BG_{TOT} 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,及在上文所描述的第二及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可如何指定减少的前景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日申请的名称为“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。

[0170] 在此方面中,位流产生单元42可以多种不同的编码方案来产生位流21,所述编码方案可促进灵活的位流产生以适应大量不同的内容递送上下文。表现为在音频行业内获得动力的一个上下文为经由网络将音频数据递送(或,换句话说,“流式传输”)到数目不断增多的不同回放装置。经由带宽受限网络将音频内容递送到具有不同程度的回放能力的装置可能是困难的,尤其是在回放期间准许高度3D音频保真度(具有大带宽消耗)(相对于基于声道或对象的音频数据)的HOA音频数据的上下文中。

[0171] 根据本发明中描述的技术,位流产生单元42可利用一个或多个可缩放层以允许HOA系数11的各种重构。层中的每一者可为阶层式的。举例来说,第一层(其可称为“基础层”)可提供准许渲染立体声扩音器馈给的HOA系数的第一重构。第二层(其可称为第一“增强层”)在应用于HOA系数的第一重构时可缩放HOA系数的第一重构以准许渲染水平环绕声音扩音器馈给(例如,5.1扩音器馈给)。第三层(其可称为第二“增强层”)可在应用于HOA系数的第二重构时可提供缩放HOA系数的第一重构以准许渲染3D环绕声音扩音器馈给(例如,22.2扩音器馈给)。就此而言,层可认为阶层式缩放前一层。换句话说,层为阶层式的,使得第一层在与第二层组合时提供高阶立体混响音频信号的较高分辨率表示。

[0172] 尽管上文描述为允许缩放紧接在前的层,但在另一层上方的任何层可缩放较低层。换句话说,上文所描述的第三层可以用来缩放第一层,即使第一层尚未被第二层“缩放”。第三层在直接应用于第一层时可提供高度信息,且由此允许渲染对应于不规则布置的扬声器几何布置的不规则扬声器馈送。

[0173] 为了准许从位流21提取层,位流产生单元42可指定在位流中指定的层的数目的指示。位流产生单元42可输出包含层的所指示数目的位流21。关于图5更详细地描述位流产生单元42。在下文图7A到9B中描述产生可缩放HOA音频数据的各种不同实例,其中用于以上实例中的每一者的边带信息的实例描述于图10到13B中。

[0174] 图5是更详细地说明在经配置以执行本发明中描述的可缩放音频译码技术的潜在版本中的第一者时的图3的位流产生单元42的图。在图5的实例中,位流产生单元42包含可缩放位流产生单元1000和非可缩放位流产生单元1002。可缩放位流产生单元1000表示经配置以产生包括具有类似于下文关于图11到13B的实例展示和描述的HOAFrame()的两个或更多个层(但在一些情况下,可缩放位流可对于某些音频上下文包括单个层)的可缩放位流21的单元。非可缩放位流产生单元1002可表示经配置以产生不提供层或换句话说不提供可缩放性的非可缩放位流21的单元。

[0175] 鉴于非可缩放位流21和可缩放位流21通常在经编码环境HOA系数59、经编码nFG信号61和经译码前景V[k]向量57方面包含相同基本数据,两者皆可称为“位流21”。然而,非可缩放位流21与可缩放位流21之间的一个差异为可缩放位流21包含可表示为层21A、21B等的层。层21A可包含经编码环境HOA系数59、经编码nFG信号61经译码前景V[k]向量57的子集,如下文更详细地描述。

[0176] 尽管可缩放和非可缩放位流21可有效地为相同位流21的不同表示,但非可缩放位流21表示为非可缩放位流21'以区分可缩放位流21与非可缩放位流21'。此外,在一些情况下,可缩放位流21可包含符合非可缩放位流21的各种层。举例来说,可缩放位流21可包含符合非可缩放位流21的基础层。在这些情况下,非可缩放位流21'可表示可缩放位流21的子位流,其中此非可缩放子位流21'可利用可缩放位流21的额外层(其称为增强层)加以增强。

[0177] 位流产生单元42可获得指示调用可缩放位流产生单元1000还是非可缩放位流产生单元1002的可缩放性信息1003。换句话说,可缩放性信息1003可指示位流产生单元42输出可缩放位流21还是非可缩放位流21'。出于说明的目的,假定可缩放性信息1003指示位流产生单元42调用可缩放位流产生单元1000以输出可缩放位流21'。

[0178] 如图5的实例中进一步展示,位流产生单元42可接收经编码环境HOA系数59A到59D、经编码nFG信号61A和61B,以及经译码前景V[k]向量57A和57B。经编码环境HOA系数59A可表示与具有阶数零和子阶数零的球基函数相关联的经编码环境HOA系数。经编码环境HOA系数59B可表示与具有阶数一和阶数零的球基函数相关联的经编码环境HOA系数。经编码环境HOA系数59C可表示与具有阶数一和子阶数负一的球基函数相关联的经编码环境HOA系数。经编码环境HOA系数59D可表示与具有阶数一和子阶数正一的球基函数相关联的经编码环境HOA系数。经编码环境HOA系数59A到59D可表示上文所论述的经编码环境HOA系数59的一个实例,且因此可统称为经编码环境HOA系数59。

[0179] 经编码nFG信号61A和61B可各自表示在此实例中表示声场的两个最占优势前景方面的US音频对象。经译码前景V[k]向量57A与57B可分别表示经编码nFG信号61A与61B的方向信息(其除了方向之外还可指定宽度)。经编码nFG信号61A和61B可表示上文所描述的经编码nFG信号61的一个实例,且因此可统称为经编码nFG信号61。经译码前景V[k]向量57A和57B可表示上文所描述的经译码前景V[k]向量57的一个实例,且因此可统称为经译码前景V[k]向量57。

[0180] 一旦调用,可缩放位流产生单元1000就可以大体类似于下文关于图7A到9B描述的方式产生可缩放位流21以包含层21A和21B。可缩放位流产生单元1000可在可缩放位流21中指定层的数目的指示,且在层21A和21B中的每一者中指定前景元素和背景元素的数目。作为一个实例,可缩放位流产生单元1000可指定可指定L个层的NumberOfLayers语法元素,其中变量L可表示层的数目。可缩放位流产生单元1000可接着对于每一层(其可表示为变量i=1到L)指定对于每一层发送的Bi个经编码环境HOA系数59和Fi个经译码nFG信号61(其还可或替代地指示对应经译码前景V[k]向量57的数目)。

[0181] 在图5的实例中,可缩放位流产生单元1000可在可缩放位流21中指定已启用可缩放译码,且两个层包含于可缩放位流21中,第一层21A包含四个经编码环境HOA系数59和零个经编码nFG信号61,且第二层21A包含零个经编码环境HOA系数59和w个经编码nFG信号61。可缩放位流产生单元1000还可产生第一层21A(其也可被称作“基础层21A”)以包含经编码环境HOA系数59。可缩放位流产生单元1000可进一步产生第二层21A(其可称为“增强层21B”)以包含经编码nFG信号61和经译码前景V[k]向量57。可缩放位流产生单元1000可输出层21A和21B作为可缩放位流21。在一些实例中,可缩放位流产生单元1000可将可缩放位流21'存储到存储器(在编码器20内部或外部)。

[0182] 在一些情况下,可缩放位流产生单元1000可不指定层的数目、一个或多个层中的

前景分量的数目(例如,经编码nFG信号61和经译码前景V[k]向量57的数目)以及一个或多个层中的背景分量的数目(例如,经编码环境HOA系数59)的指示中的一者或多者或任一者。在本发明中,分量也可被称作声道。替代地,可缩放位流产生单元1000可比较当前帧的层的数目与前一帧(例如,在时间上最近的前一帧)的层的数目。在比较得出无差异(意味着当前帧中的层的数目等于前一帧中的层的数目)时,可缩放位流产生单元1000可以类似方式比较每一层中的背景分量与前景分量和数目。

[0183] 换句话说,可缩放位流产生单元1000可比较一个或多个层中在当前帧的背景分量的数目与一个或多个层中在前一帧的背景组件的数目。可缩放位流产生单元1000可进一步比较一个或多个层中在当前帧的前景分量的数目与一个或多个层中在前一帧的前景分量的数目。

[0184] 在两个基于分量的比较皆得出无差异(意味着,前景和背景分量在前一帧中的数目等于前景和背景分量在当前帧中的数目)时,可缩放位流产生单元1000可在可缩放位流21中指定当前帧中的层的数目等于前一帧中的层的数目的指示(例如,HOABaseLayerConfigurationFlag语法元素),而非指定层的数目、一个或多个层中的前景分量的数目(例如,经编码nFG信号61和经译码前景V[k]向量57的数目)和一个或多个层中的背景分量(例如,经编码环境HOA系数59)的数目的指示中的任一者。音频解码装置24可接着确定层、背景分量和前景分量的数目的前一帧指示等于层的数目、背景分量和前景分量的数目的当前帧指示,如下文更详细地描述。

[0185] 在上文提到的任一比较得出差异时,可缩放位流产生单元1000可在可缩放位流21中指定当前帧中的层的数目不等于前一帧中的层的数目的指示(例如,HOABaseLayerConfigurationFlag语法元素)。可缩放位流产生单元1000可接着指定层的数目、一个或多个层中的前景分量的数目(例如,经编码nFG信号61和经译码前景V[k]向量57的数目)和一个或多个层中的背景分量(例如,经编码环境HOA系数59)的数目的指示,如上文所指出。就此而言可缩放位流产生单元1000可在位流中指定位流的层的数目当与位流在前一帧中的层的数目相比时在当前帧中是否已改变的指示,且指定位流在当前帧中的层的所指示数目。

[0186] 在一些实例中,并非不指定前景分量的数目的指示和背景分量的数目的指示,可缩放位流产生单元1000可不在可缩放位流21中指定分量的数目的指示(例如,“NumChannels”语法元素,其可为具有[i]项的阵列,其中i等于层的数目)。可缩放位流产生单元1000可不指定分量(其中这些分量也可被称作“声道”)的数目的此指示,代替不指定前景和背景分量的数目(鉴于前景和背景分量的数目可从更通用声道数目导出)。在一些实例中,前景分量的数目的指示和背景声道的数目的指示的导出可根据下表进行:

[0187] 表—ChannelSideInfoData(i)的语法

	语法	位数	助记符
[0188]	ChannelSideInfoData(i) { ChannelType[i]	2	uimsbf

[0189]	<pre> switch ChannelType[i] { case 0: ActiveDirsIds[i]; break; case 1: if(hoaDependencyFlag){ NbitsQ(k)[i] if (NbitsQ(k)[i] == 4) { CodebkIdx(k)[i]; NumVecIndices(k)[i]++; } elseif (NbitsQ(k)[i] >= 6) { PFlag(k)[i] = 0; CbFlag(k)[i]; } } else{ bA; bB; if ((bA + bB) == 0) { NbitsQ(k)[i] = NbitsQ(k-1)[i]; PFlag(k)[i] = PFlag(k-1)[i]; CbFlag(k)[i] = CbFlag(k-1)[i]; CodebkIdx(k)[i] = CodebkIdx(k-1)[i]; NumVecIndices(k)[i] = NumVecIndices[k-1][i]; } else{ NbitsQ(k)[i] = (8*bA)+(4*bB)+uintC; if (NbitsQ(k)[i] == 4) { CodebkIdx(k)[i]; } elseif (NbitsQ(k)[i] >= 6) { PFlag(k)[i]; CbFlag(k)[i]; } } } break; case 2: AddAmbHoaInfoChannel(i); break; default: } } </pre>	NumOfBitsPerDirIdx uimsbf	
	4 uimsbf 3 uimsbf uimsbf 1 bslbf 1 bslbf 1 bslbf 2 uimsbf 3 uimsbf 1 bslbf 1 bslbf		

[0190] 其中ChannelType的描述给出如下：

[0191] ChannelType:

[0192] 0:基于方向的信号

[0193] 1:基于向量的信号(其可表示前景信号)

[0194] 2:额外环境HOA系数(其可表示背景或环境信号)

[0195] 3:空

[0196] 作为根据以上SideChannelInfo语法表用信号表示ChannelType的结果,可根据设

定为1的Channel1Type语法元素的数目而确定每层的前景分量的数目,且可根据设定为2的Channel1Type语法元素的数目确定每层的背景分量的数目。

[0197] 在一些实例中,可缩放位流产生单元1000可逐帧地指定HOADecoderConfig,其提供用于从位流21提取层的配置信息。可替代或结合上表指定HOADecoderConfig。下表可界定位流21中的HOADecoderConfig_FrameByFrame()对象的语法。

语法	位数	助记符
HOADecoderConfig_FrameByFrame(numHOATransportChannels)		
{		
HOABaseLayerPresent;	1	bslbf
if(HOABaseLayerPresent){		
HOABaseLayerConfigurationFlag;	1	bslbf
if(HOABaseLayerConfigurationFlag){		
NumLayerBits =		
ceil(log2(numHOATransportChannels-2));		
NumLayers = NumLayers +2;	NumLayerBits	uimsbf
numAvailableTransportChannels =		
numHOATransportChannels-2;		
numAvailableTransportChannelsBits =		
NumLayerBits;		
for (i=0; i<NumLayers-1; ++i) {		
NumFGchannels[i] =		
NumFGchannels[i] +1;	numAvailableTransportChannelsBit s	uimsbf
numAvailableTransportChannels =		
numAvailableTransportChannels -		
NumFGchannels[i]		
numAvailableTransportChannelsBits =		
ceil(log2(numAvailableTransportChannels));		
NumBGchannels[i] =		
NumBGchannels[i] + 1;	numAvailableTransportChannelsBit s	uimsbf
numAvailableTransportChannels =		
numAvailableTransportChannels -		
NumBGchannels[i]		
numAvailableTransportChannelsBits =		
ceil(log2(numAvailableTransportChannels));		
}		

[0198]

[0199]	<pre> } else { NumLayers=NumLayersPrevFrame; for (i=0; i<NumLayers; ++i) { NumFGchannels[i] = NumFGchannels_PrevFrame[i]; NumBGchannels[i] = NumBGchannels_PrevFrame[i]; } } MinAmbHoaOrder = escapedValue(3,5,0) alue MinNumOfCoeffsForAmbHOA = (MinAmbHoaOrder + 1)^2; . . . NumLayersPrevFrame=NumLayers; for (i=0; i<NumLayers; ++i) { NumFGchannels_PrevFrame[i] = NumFGchannels[i]; NumBGchannels_PrevFrame[i] = NumBGchannels[i]; } } </pre>	3,8	uimsbf

[0200] 在上表中,HOABaseLayerPresent语法元素可表示指示可缩放位流21的基础层是否存在的旗标。当存在时,可缩放位流产生单元1000指定HOABaseLayerConfigurationFlag语法元素,其可表示指示基础层的配置信息是否存在与位流21中的语法元素。在基础层的配置信息存在于位流21中时,可缩放位流产生单元1000指定层的数目(即,所述实例中的NumLayers语法元素)、层中的每一者的前景声道的数目(即,所述实例中的NumFGchannels语法元素),和层中的每一者的背景声道的数目(即,所述实例中的NumBGchannels语法元素)。在HOABaseLayerPresent旗标指示不存在基础层配置时,可缩放位流产生单元1000可不提供任何额外语法元素,且音频解码装置24可确定当前帧的配置数据与前一帧的配置数据相同。

[0201] 在一些实例中,可缩放位流产生单元1000可在可缩放位流21中指定HOADecoderConfig对象,但不每层指定前景和背景声道的数目,其中前景和背景声道的数目可为静态的或如上文关于ChannelSideInfo表所描述而加以确定。在此实例中,HOADecoderConfig可根据下表加以界定。

语法	位数	助记符
HOADecoderConfig(numHOATransportChannels)		
{		
HOABaseLayerPresent;	1	bslbf
if(HOABaseLayerPresent){		
HOABaseLayerChBits =		
ceil(log2(numHOATransportChannels));		
NumHOABaseLayerCh;	HOABaseLayerChBits	uimsbf
HOABaseLayerConfigurationFlag;	1	bslbf
if(HOABaseLayerConfigurationFlag){		
NumLayerBits =		
ceil(log2(numHOATransportChannels));		
NumLayers;	NumLayerBits	uimsbf
numAvailableTransportChannels =		
numHOATransportChannels		
numAvailableTransportChannelsBits =		
ceil(log2(numAvailableTransportChannels));		
for i=1:NumLayers-1 {		
NumChannels[i]	numAvailableTransportChannelsBits	
numAvailableTransportChannels =		
numAvailableTransportChannels -		
NumChannels[i]		
}		
}		
}		
}		
MinAmbHoaOrder = escapedValue(3,5,0) - 1;	3,8	uimsbf
MinNumOfCoeffsForAmbHOA =		
(MinAmbHoaOrder + 1)^2;		
.		
.		
.		
.		
.		
NumLayersPrevFrame=NumLayers;		
for i=1:NumLayers {		
NumChannels_PrevFrame[i] =		
NumChannels[i];		
}		
}		

[0203] 作为又一替代方案,用于HOADecoderConfig的前述语法表可替换为用于HOADecoderConfig的以下语法表。

语法	位数	助记符
HOADecoderConfig(numHOATransportChannels)		
{		
MinAmbHoaOrder = escapedValue(3,5,0) - 1;	3,8	uimsbf
MinNumOfCoeffsForAmbHOA = (MinAmbHoaOrder + 1)^2;		
NumOfAdditionalCoders = numHOATransportChannels -		
MinNumOfCoeffsForAmbHOA;		
SingleLayer ;	1	bslbf
if(SingleLayer==0){		
NumOfAdditionalCoders = escapeValue(5,8,16) +1 +		uimsbf
NumOfAdditionalCoders;		
HOALayerChBits = ceil(log2(NumOfAdditionalCoders));		
NumHOAChannelsLayer[0] = codedLayerCh +		HOALayerChBits uimsbf
MinNumOfCoeffsForAmbHOA;		
remainingCh = numHOATransportChannels -		
NumHOAChannelsLayer[0];		
NumLayers = 1;		
while (remainingCh>1) {		
HOALayerChBits = ceil(log2(remainingCh));		
NumHOAChannelsLayer[NumLayers]	= HOALayerChBits uimsbf	
codedLayerCh + 1;		
remainingCh = remainingCh -		
NumHOAChannelsLayer[NumLayers];		
NumLayers++;		
}		
if(remainingCh) {		
NumHOAChannelsLayer[NumLayers] = 1;	2	uimsbf
NumLayers++;		
}		
}		
MaxNoOfDirSigsForPrediction =		
MaxNoOfDirSigsForPrediction + 1;	4	uimsbf
NoOfBitsPerScalefactor = NoOfBitsPerScalefactor + 1;	3	uimsbf
CodedSpatialInterpolationTime ;	1	bslbf
SpatialInterpolationMethod ;	2	uimsbf
CodedVVecLength ;	3	uimsbf
MaxGainCorrAmpExp ;		
MaxNumAddActiveAmbCoeffs = NumOfHoaCoeffs -		
MinNumOfCoeffsForAmbHOA;		
AmbAsignmBits =		
ceil(log2(MaxNumAddActiveAmbCoeffs));		
ActivePredIdsBits = ceil(log2(NumOfHoaCoeffs));		
i = 1;		
while(i * ActivePredIdsBits + ceil(log2(i)) <		
NumOfHoaCoeffs){		
i++;		
}		

[0205]

```

NumActivePredIdsBits = ceil( log2( max( 1, i - 1 ) ) );
GainCorrPrevAmpExpBits = ceil( log2( ceil( log2(
                                         1.5 * NumOfHoaCoeffs ) )
                                         + MaxGainCorrAmpExp ) +
1 ) );
for (i=0; i<NumOfAdditionalCoders; ++i){
    AmbCoeffTransitionState[i] = 3;
}
}

```

注意：节省了 MinAmbHoaOrder = 30 ... 37。

[0206] 就此而言，可缩放位流产生单元1000可经配置以如上文所描述在位流中指定在位流的一个或多个层中指定的声音的数目的指示，且指定位流的一个或多个层中的声音的所指示数目。

[0207] 此外，可缩放位流产生单元1000可经配置以指定指示声音的数目的语法元素（例如，呈NumLayers语法元素或codedLayerCh语法元素形式，如下文更详细地所描述）。

[0208] 在一些实例中，可缩放位流产生单元1000可经配置以指定在位流中指定的声音的总数目的指示。在这些情况下，可缩放位流产生单元1000可经配置以指定位流的一个或多个层中的声音的所指示总数目。在这些情况下，可缩放位流产生单元1000可经配置以指定指示声音的总数目的语法元素（例如，numHOATransportChannels语法元素，如下文更详细地描述）。

[0209] 在这些和其它实例中，可缩放位流产生单元1000可经配置以指定在位流中的一个或多个层中指定的声音中的一者的类型的指示。在这些情况下，可缩放位流产生单元1000可经配置以指定位流的一个或多个层中的声音中的所述一者的所指示类型的所指示数目。前景声道可包括US音频对象和对应V向量。

[0210] 在这些和其它实例中，可缩放位流产生单元1000可经配置以指定在位流中的一个或多个层中指定的声音中的一者的类型的指示、指示声道中的一者为前景声道的声音中的一者的类型的指示。在这些情况下，可缩放位流产生单元1000可经配置以指定位流的一个或多个层中的前景声道。

[0211] 在这些和其它实例中，可缩放位流产生单元1000可经配置以指定在位流中的一个或多个层中指定的声音中的一者的类型的指示、指示声道中的一者为背景声道的声音中的一者的类型的指示。在这些情况下，可缩放位流产生单元1000可经配置以指定位流的一个或多个层中的背景声道。背景声道可包括环境HOA系数。

[0212] 在这些和其它实例中，可缩放位流产生单元1000可经配置以指定指示声道中的一者的类型的语法元素（例如，ChannelType语法元素）。

[0213] 在这些和其它实例中，可缩放位流产生单元1000可经配置以基于在获得层中的一者之后在位流中剩余的声音的数目指定声音的数目的指示（如例如由remainingCh语法元素或numAvailableTransportChannels语法元素界定，如下文更详细地描述）。

[0214] 图7A到7D是说明音频编码装置20在产生HOA系数11的经编码双层表示时的实例操作的流程图。首先参考图7A的实例，去相关单元60可首先对于表示为经能量补偿的背景HOA系数47A'到47D'的一阶立体混响背景（其中“立体混响背景”可指描述声场的背景分量的立体混响系数）应用UHJ去相关（300）。一阶立体混响背景47A'到47D'可包含对应于具有以下（阶数，子阶数）的球基函数的HOA系数：(0,0)、(1,0)、(1,-1)、(1,1)。

[0215] 去相关单元60可输出去相关环境HOA音频信号67作为以上所提到的Q、T、L和R音频信号。Q音频信号可提供高度信息。T音频信号可提供水平信息(包含用于表示甜点后的声道的信息)。L音频信号提供左立体声声道。R音频信号提供右立体声声道。

[0216] 在一些实例中,UHJ矩阵可至少包括与左音频声道相关联的高阶立体混响音频数据。在其它实例中UHJ矩阵可至少包括与右音频声道相关联的高阶立体混响音频数据。在另外其它实例中,UHJ矩阵可至少包括与定位声道相关联的高阶立体混响音频数据。在其它实例中,UHJ矩阵可至少包括与高度声道相关联的高阶立体混响音频数据。在其它实例中,UHJ矩阵可至少包括与用于自动增益校正的边带相关联的高阶立体混响音频数据。在其它实例中,UHJ矩阵可至少包括与左音频声道、右音频声道、定位声道以及高度声道和用于自动增益校正的边带相关联的高阶立体混响音频数据。

[0217] 增益控制单元62可将自动增益控制(AGC)应用于去相关环境HOA音频信号67(302)。增益控制单元62可将经调整环境HOA音频信号67'传递到位流产生单元42,所述位流产生单元可基于经调整环境HOA音频信号67'形成基础层,且基于高阶立体混响增益控制数据(HOAGCD)形成旁侧声道的至少部分(304)。

[0218] 增益控制单元62还可对于经内插nFG音频信号49'(其也可被称作“基于向量的占优势信号”)应用自动增益控制(306)。增益控制单元62可将经调整nFG音频信号49”连同用于经调整nFG音频信号49”的HOAGCD输出到位流产生单元42。位流产生单元42可基于经调整nFG音频信号49”形成第二层,同时基于用于经调整nFG音频信号49”的HOAGCD和对应经译码前景V[k]向量57形成边带信息的部分(308)。

[0219] 高阶立体混响音频数据的两个或更多个层中的第一层(即,基础层)可包括对应于具有等于或小于一的阶数的一个或多个球基函数的高阶立体混响系数。在一些实例中,第二层(即,增强层)包括基于向量的占优势音频数据。

[0220] 在一些实例中,基于向量的占优势音频至少包括占优势音频数据和经编码V向量。如上文所描述,经编码V向量可通过音频编码装置20的LIT单元30经由应用线性可逆变换而从高阶立体混响音频数据分解而来。在其它实例中,基于向量的占优势音频数据至少包括额外高阶立体混响声道。在另外其它实例中,基于向量的占优势音频数据至少包括自动增益校正边带。在其它实例中,基于向量的占优势音频数据至少包括占优势音频数据、经编码V向量、额外高阶立体混响声道和自动增益校正边带。

[0221] 在形成第一层和第二层时,位流产生单元42可执行错误检测、错误校正或错误检测与校正两者的错误检查过程。在一些实例中,位流产生单元42可对第一层(即,基础层)执行错误检查过程。在另一实例中,音频译码装置可对第一层(即,基础层)执行错误检查过程,且制止对第二层(即,增强层)执行错误检查过程。在又一实例中,位流产生单元42可对第一层(即,基础层)执行错误检查过程,且响应于确定第一层无错误,音频译码装置可对第二层(即,增强层)执行错误检查过程。在位流产生单元42对第一层(即,基础层)执行错误检查过程的以上实例中的任一者中,第一层可认为是极少出现错误的稳健层。

[0222] 接下来参考图7B,增益控制单元62和位流产生单元42执行与上文关于图7A所描述的增益控制单元62和位流产生单元42的操作类似的操作。然而,去相关单元60可将模式矩阵去相关而非UHJ去相关应用于一阶立体混响背景47A'到47D'(301)。

[0223] 接下来参考图7C,增益控制单元62和位流产生单元42可执行与与上文关于图7A和

7B的实例所描述的增益控制单元62和位流产生单元42的操作类似的操作。然而,在图7C的实例中,去相关单元60可不将任何变换应用于一阶立体混响背景47A'到47D'。在以下实例8A到10B中的每一者中,假定但不说明去相关单元60可替代地不对于一阶立体混响背景47A'到47D'中的一者或者应用去相关。

[0224] 接下来参考图7D,去相关单元60和位流产生单元42可执行与上文关于图7A和7B的实例所描述的增益控制单元52和位流产生单元42的操作类似的操作。然而,在图7D的实例中,增益控制单元62可不将任何增益控制应用于去相关环境HOA音频信号67。在以下实例8A到10B中的每一者中,假定但不说明增益控制单元52可替代地不对于去相关环境HOA音频信号67中的一者或者应用去相关。

[0225] 在图7A到7D的实例中的每一者中,位流产生单元42可在位流21中指定一个或多个语法元素。图10是说明在位流21中指定的HOA配置对象的实例的图。对于图7A到7D的实例中的每一者,位流产生单元42可将codedVVecLength语法元素400设定为1或2,其指示1阶背景HOA声道含有所有占优势声音的1阶分量。位流产生单元42还可设定ambienceDecorrelationMethod语法元素402,使得元素402用信号表示UHJ去相关的使用(例如,如上文关于图7A所描述),用信号表示矩阵模式去相关的使用(例如,如上文关于图7B所描述),或用信号表示不使用去相关(例如,如上文关于图7C所描述)。

[0226] 图11是说明由位流产生单元42针对第一层和第二层产生的边带信息410的图。边带信息410包含边带基础层信息412以及边带第二层信息414A和414B。仅在基础层提供到音频解码装置24时,音频编码装置20才可仅提供边带基础层信息412。边带基础层信息412包含用于基础层的HOAGCD。边带第二层信息414A包含传输信道1到4语法元素和对应HOAGCD。边带第二层信息414B包含对应于传输信道1和2的对应的两个经译码减少V[k]向量57(鉴于传输信道3和4为空,如由等于11₂或3₁₀的ChannelType语法元素所表示)。

[0227] 图8A和8B是说明音频编码装置20在产生HOA系数11的经编码三层表示时的实例操作的流程图。首先参考图8A的实例,去相关单元60和增益控制单元62可执行与上文关于图7A所描述的操作类似的操作。然而,位流产生单元42可基于经调整环境HOA音频信号67的L音频信号和R音频信号而非所有经调整环境HOA音频信号67形成基础层(310)。就此而言,在音频解码装置24显现时,基础层可提供立体声声道。位流产生单元42还可产生基础层的边带信息,包含HOAGCD。

[0228] 位流产生单元42的操作还可不同于上文关于图7A所描述的操作,不同之处在于位流产生单元42可基于经调整环境HOA音频信号67的Q和T音频信号形成第二层(312)。在图8A的实例中,第二层可在音频解码装置24处显现时提供水平声道和3D音频声道。位流产生单元42还可产生用于第二层的边带信息,包含HOAGCD。位流产生单元42还可以大体类似于上文关于在图7A的实例中形成第二层所描述的方式形成第三层。

[0229] 位流产生单元42可指定类似于上文关于图10所描述对象的用于位流21的HOA配置对象。另外,音频编码器20的位流产生单元42将MinAmbHoaOrder语法元素404设定为2以便指示发射1阶HOA背景。

[0230] 位流产生单元42还可产生类似于图12A的实例中所示的边带信息412的边带信息。图12A为说明根据本发明中描述的技术的可缩放译码方面产生的边带信息412的图。边带信息412包含边带基础层信息416、边带第二层信息418,以及边带第三层信息420A和420B。边

带基础层信息416可提供用于基础层的HOAGCD。边带第二层信息418可提供用于第二层的HOAGCD。边带第三层信息420A和420B可类似于上文关于图11所描述的边带信息414A和414B。

[0231] 类似于图7A,位流产生装置42可执行错误检查过程。在一些实例中,位流产生装置42可对第一层(即,基础层)执行错误检查过程。在另一实例中,位流产生装置42可对第一层(即,基础层)执行错误检查过程,且制止对第二层(即,增强层)执行错误检查过程。在又一实例中,位流产生装置42可对第一层(即,基础层)执行错误检查过程,且响应于确定第一层无错误,音频译码装置可对第二层(即,增强层)执行错误检查过程。在音频译码装置对第一层(即,基础层)执行错误检查过程的以上实例中的任一者中,第一层可认为是极少出现错误的稳健层。

[0232] 尽管描述为提供三个层,但在一些实例中,位流产生装置42可在位流中指定仅存在两个层的指示,且指定位流的层中的第一者指示高阶立体混响音频信号的提供立体声声道回放的背景分量,且位流的层中的第二者指示高阶立体混响音频信号的提供由布置在单个水平面上的三个或更多个扬声器进行的水平多声道回放的背景分量。换句话说,尽管展示为提供三个层,但位流产生装置42可在一些情况下仅产生三个层中的两者。应理解,可产生所述层的任何子集,尽管本文中未详细地描述。

[0233] 接下来参考图8B,增益控制单元62和位流产生单元42可执行与与上文关于图8A的实例所描述的增益控制单元62和位流产生单元42的操作类似的操作。然而,去相关单元60可将模式矩阵去相关而非UHJ去相关应用于一阶立体混响背景47A'(316)。在一些实例中,一阶立体混响背景47A'可包含零阶立体混响系数47A'。增益控制单元62可将自动增益控制应用于对应于具有一阶的球谐系数的一阶立体混响系数和去相关环境HOA音频信号67。

[0234] 位流产生单元42可基于经调整环境HOA音频信号67形成基础层,且基于对应HOAGCD形成边带的至少部分(310)。环境HOA音频信号67可于在音频解码装置24处显现时提供单声道。位流产生单元42可基于经调整环境HOA系数47B"到47D"形成第二层,且基于对应HOAGCD形成边带的至少部分(318)。经调整环境HOA系数47B"到47D"可在于音频解码装置24处显现时提供X、Y和Z(或立体声,水平和高度)声道。位流产生单元42可以类似于上文关于图8A所描述的方式形成第三层和边带信息的至少部分。位流产生单元42可产生边带信息412,如关于图12B更详细描述的(326)。

[0235] 图12B为说明根据本发明中描述的技术的可缩放译码方面产生的边带信息414的图。边带信息414包含边带基础层信息416、边带第二层信息422,以及边带第三层信息424A到424C。边带基础层信息416可提供用于基础层的HOAGCD。边带第二层信息422可提供用于第二层的HOAGCD。边带第三层信息424A到424C可类似于上文关于图11所描述的边带信息414A(只是边带信息414A指定为边带第三层信息424A和424B)和414B。

[0236] 图9A和9B是说明音频编码装置20在产生HOA系数11的经编码四层表示时的实例操作的流程图。首先参考图9A的实例,去相关单元60和增益控制单元62可执行与上文关于图8A所描述的操作类似的操作位流产生单元42可以类似于上文关于图8A的实例所描述的方式形成基础层,即,基于经调整环境HOA音频信号67的L音频信号和R音频信号而非所有经调整环境HOA音频信号67(310)。就此而言,基础层可于在音频解码装置24处显现时提供立体声声道(或,换句话说,提供立体声声道回放)。位流产生单元42还可产生用于基础层

的边带信息,包含HOAGCD。

[0237] 位流产生单元42的操作可不同于上文关于图8A所描述的操作,不同之处在于位流产生单元42可基于经调整环境HOA音频信号67的T音频信号(而非Q音频信号)形成第二层(322)。在图9A的实例中,第二层可在于音频解码装置24处显现时提供水平声道(或,换句话说,在单个水平面上的三个或更多个扩音器进行的多声道回放)。位流产生单元42还可产生用于第二层的边带信息,包含HOAGCD。位流产生单元42还可基于经调整环境HOA音频信号67的Q音频信号形成第三层(324)。第三层可提供由布置在一个或多个水平面上的三个或更多个扬声器进行的三维回放。位流产生单元42可以大体类似于上文关于在图8A的实例中形成第三层所描述的方式形成第四层(326)。

[0238] 位流产生单元42可指定类似于上文关于图10所描述对象的用于位流21的HOA配置对象。另外,音频编码器20的位流产生单元42将MinAmbHoaOrder语法元素404设定为2以便指示发射1阶HOA背景。

[0239] 位流产生单元42还可产生类似于图13A的实例中所示的边带信息412的边带信息。图13A是说明根据本发明中描述的技术的可缩放译码方面产生的边带信息430的图。边带信息430包含边带基础层信息416、边带第二层信息418、边带第三层信息432以及边带第四层信息434A和434B。边带基础层信息416可提供用于基础层的HOAGCD。边带第二层信息418可提供用于第二层的HOAGCD。边带第三层信息430可提供用于第三层的HOAGCD。边带第四层信息434A和434B可类似于上文关于图12A所描述的边带信息420A和420B。

[0240] 类似于图7A,位流产生装置42可执行错误检查过程。在一些实例中,位流产生装置42可对第一层(即,基础层)执行错误检查过程。在另一实例中,位流产生装置42可对第一层(即,基础层)执行错误检查过程,且制止对其余层(即,增强层)执行错误检查过程。在又一实例中,位流产生装置42可对第一层(即,基础层)执行错误检查过程,且响应于确定第一层无错误,音频译码装置可对第二层(即,增强层)执行错误检查过程。在音频译码装置对第一层(即,基础层)执行错误检查过程的以上实例中的任一者中,第一层可认为是极少出现错误的稳健层。

[0241] 接下来参考图9B,增益控制单元62和位流产生单元42可执行与上文关于图9A的实例所描述的增益控制单元62和位流产生单元42的操作类似的操作。然而,去相关单元60可将模式矩阵去相关而非UHJ去相关应用于一阶立体混响背景47A'(316)。在一些实例中,一阶立体混响背景47A'可包含零阶立体混响系数47A'。增益控制单元62可将自动增益控制应用于对应于具有一阶的球谐系数的一阶立体混响系数和去相关环境HOA音频信号67(302)。

[0242] 位流产生单元42可基于经调整环境HOA音频信号67形成基础层,且基于对应HOAGCD形成边带的至少部分(310)。环境HOA音频信号67可于在音频解码装置24处显现时提供单声道。位流产生单元42可基于经调整环境HOA系数47B"到47D"形成第二层,且基于对应HOAGCD形成边带的至少部分(322)。经调整环境HOA系数47B'和47C'可提供由布置在单个水平面上的三个或更多个扬声器进行的X、Y水平多声道回放。位流产生单元42可基于经调整环境HOA系数47D"形成第三层,且基于对应HOAGCD形成边带的至少部分(324)。经调整环境HOA系数47D"可提供由布置于一个或多个水平面中的三个或更多个扬声器进行的三维回放。位流产生单元42可以类似于上文关于图8A所描述的方式形成第四层和边带信息

的至少部分(326)。位流产生单元42可产生边带信息412,如关于图12B更详细描述的。

[0243] 图13B是说明根据本发明中描述的技术的可缩放译码方面产生的边带信息440的图。边带信息440包含边带基础层信息416、边带第二层信息442、边带第三层信息444以及边带第四层信息446A到446C。边带基础层信息416可提供用于基础层的HOAGCD。边带第二层信息442可提供用于第二层的HOAGCD。边带第三层信息可提供用于第三层的HOAGCD。边带第四层信息446A到446C可类似于上文关于图12B所描述的边带信息424A到424C。

[0244] 图4是更详细地说明图2的音频解码装置24的框图。如图4的实例中所示,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和基于向量的重构单元92。尽管在下文描述,但关于音频解码装置24和解压缩或以其它方式对HOA系数进行解码的各种方面的更多信息可用于2014年5月29日申请的标题为“用于声场的分解表示的内插(INTERPOLATION FOR DECOMPOSED REPRESENTATIONS OF A SOUND FIELD)”的第WO 2014/194099号国际专利申请公开案。进一步信息还可见于上文提及的MPEG-H 3D音频译码标准阶段I和阶段II以及上文提及的概述MPEG-H 3D音频译码标准阶段I的对应论文中。

[0245] 提取单元72可表示经配置以接收位流21且提取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可从以上所提到的指示HOA系数11是经由各种基于方向的还是基于向量的版本而编码的语法元素进行确定。在执行基于方向的编码时,提取单元72可提取HOA系数11的基于方向的版本以及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),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方向的信息91重构呈HOA系数11'形式的HOA系数的单元。

[0246] 在语法元素指示HOA系数11是使用基于向量的合成编码时,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及/或索引63或标量经量化V向量)、经编码环境HOA系数59和对应音频对象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连同经编码nFG信号61传递到心理声学解码单元80。关于图6的实例更详细地描述提取单元72。

[0247] 图6是更详细地说明在经配置以执行本发明中描述的可缩放音频解码技术的潜在版本中的第一者时的图4的提取单元72的图。在图6的实例中,提取单元72包含模式选择单元1010、可缩放提取单元1012和非可缩放提取单元1014。模式选择单元1010表示经配置以选择将对于位流21执行可缩放还是非可缩放提取的单元。模式选择单元1010可包含存储位流21的存储器。模式选择单元1010可基于是否已启用可缩放译码的指示确定将执行可缩放还是非可缩放提取。HOABaseLayerPresent语法元素可表示在对位流21进行编码时是否执行可缩放译码的指示。

[0248] 在HOABaseLayerPresent语法元素指示已启用可缩放译码时,模式选择单元1010可将位流21识别为可缩放位流21,且将可缩放位流21输出到可缩放提取单元1012。在HOABaseLayerPresent语法元素指示尚未启用可缩放译码时,模式选择单元1010可将位流21识别为非可缩放位流21',且将非可缩放位流21'输出到非可缩放提取单元1014。非可缩放提取单元1014表示经配置以根据MPEG-H 3D音频译码标准阶段I操作的单元。

[0249] 可缩放提取单元1012可表示经配置以基于下文更详细描述(且上文在各种

HOADecoderConfig表中展示)的各种语法元素从可缩放位流21的一个或多个层提取环境HOA系数59、经编码nFG信号61和经译码前景V[k]向量57中的一者或多者。在图6的实例中,作为一个实例,可缩放提取单元1012可从可缩放位流21的基础层21A提取四个经编码环境HOA系数59A到59D。可缩放提取单元1012还可从可缩放位流21的增强层21B提取两个经编码nFG信号61A和61B(作为一个实例)以及两个经译码前景V[k]向量57A和57B。可缩放提取单元1012可将环境HOA系数59、经编码nFG信号61和经译码前景V[k]向量57输出到图4的实例中所示的基于向量的解码单元92。

[0250] 更具体来说,音频解码装置24的提取单元72可提取L层的声音,如在以上HOADecoderConfig_FrameByFrame语法表中所阐述。

[0251] 根据以上HOADecoderConfig_FrameByFrame语法表,模式选择单元1010可首先获得可指示是否执行可缩放音频编码的HOABaseLayerPresent语法元素。在未启用时(如例如由HOABaseLayerPresent语法元素的零值指定),模式选择单元1010可确定MinAmbHoaOrder语法元素,且将非可缩放位流提供到非可缩放提取单元1014,所述非可缩放提取单元执行类似于上文所描述的过程的非可缩放提取过程。在启用时(如例如由HOABaseLayerPresent语法元素的一值所指定),模式选择单元1010将MinAmbHoaOrder语法元素值设定为负一(-1),且将可缩放位流21'提供到可缩放提取单元1012。

[0252] 可缩放提取单元1012可获得位流的层的数目当与位流在前一帧中的层的数目相比时在当前帧中是否已改变的指示。位流的层的数目当与位流在前一帧中的层的数目相比时在当前帧中是否已改变的指示可表示为前述表中的“HOABaseLayerConfigurationFlag”语法元素。

[0253] 可缩放提取单元1012可基于所述指示获得位流在当前帧中的层的数目的指示。在所述指示指示位流的层的数目当与所述位流在前一帧中的层的数目相比时在当前帧中尚未改变时,可缩放提取单元1012可根据以上语法表的部分确定位流在当前帧中的层的数目等于位流在前一帧中的层的数目,所述部分述及:

[0254] ...

[0255] } else}

[0256] NumLayers=NumLayersPrevFrame;

[0257] 其中“NumLayers”可表示表示位流在当前帧中的层的数目的语法元素,且“NumLayersPrevFrame”可表示表示位流在前一帧中的层的数目的语法元素。

[0258] 根据以上HOADecoderConfig_FrameByFrame语法表,在所述指示指示位流的层的数目当与所述位流在前一帧中的层的数目相比时在当前帧中尚未改变时,可缩放提取单元1012可确定层中的一者或多者在当前帧的前景分量的当前数目的当前前景指示等于层中的一者或多者在前一帧的前景分量的先前数目的先前前景指示。换句话说,在HOABaseLayerConfigurationFlag等于零时,可缩放提取单元1012可确定表示层中的一者或多者在当前帧的前景分量的当前数目的当前前景指示的NumFGchannels[i]语法元素等于表示一个或多个层在前一帧的前景分量的先前数目的先前前景指示的NumFGchannels_PrevFrame[i]语法元素。可缩放提取单元1012可基于当前前景指示进一步从当前帧中的一个或多个层获得前景分量。

[0259] 在所述指示指示位流的层的数目当与所述位流在前一帧中的层的数目相比时在

当前帧中尚未改变时,可缩放提取单元1012还可确定层中的一者或者在当前帧的背景分量的当前数目的当前背景指示等于层中的一者或者在前一帧的背景分量的先前数目的先前背景指示。换句话说,在HOABaseLayerConfigurationFlag等于零时,可缩放提取单元1012可确定表示层中的一者或者在当前帧的背景分量的当前数目的当前背景指示的NumBGchannels[i]语法元素等于表示一个或多个层在前一帧的背景分量的先前数目的先前背景指示的NumBGchannels_PrevFrame[i]语法元素。可缩放提取单元1012可基于当前背景指示进一步从当前帧中的一个或多个层获得背景分量。

[0260] 为实现可潜在地降低层的数目、前景分量和背景分量的各种指示的信令的前述技术,可缩放提取单元1012可将NumFGchannels_PrevFrame[i]语法元素和NumBGchannel_PrevFrame[i]语法元素设定为用于当前帧的指示(例如,NumFGchannels[i]语法元素和NumBGchannels[i]),在所有i个层迭代。此以以下语法表示:

[0261]	<pre> NumLayersPrevFrame=NumLayers; for i=1:NumLayers { NumFGchannels_PrevFrame[i] = NumFGchannels[i]; NumBGchannels_PrevFrame[i] = NumBGchannels[i]; } </pre>
--------	--

[0262] 在所述指示指示位流的层的数目当与所述位流在前一帧中的层的数目相比时在当前帧中已改变时(例如,在HOABaseLayerConfigurationFlag等于一时),可缩放提取单元1012根据numHOATransportChannels获得NumLayerBits语法元素,所述NumLayerBits语法元素传递到已根据本发明中未描述的其它语法表获得的语法表中。

[0263] 可缩放提取单元1012可获得在位流中指定的层的数目的指示(例如,NumLayers语法元素),其中所述指示可具有由NumLayerBits语法元素指示的位数。NumLayers语法元素可指定在位流中指定的层的数目,其中层的数目可表示为以上L。可缩放提取单元1012可接下来根据numHOATransportChannels确定numAvailableTransportChannels,且根据numAvailableTransportChannels确定numAvailable TransportChannelBits。

[0264] 可缩放提取单元1012可接着迭代经过NumLayers,从1到NumLayers-1,以确定为第i层指定的背景HOA声道的数目(B_i)和前景HOA声道的数目(F_i)。可缩放提取单元1012可不迭代经过最后层的编号(NumLayer)而仅经过NumLayer-1,因为在位流中发送的前景和背景HOA声道的总数目被可缩放提取单元1012知晓时(例如,在前景和背景HOA声道的总数目用信号表示为语法元素时),可确定最后层 B_L 。

[0265] 就此而言,可缩放提取单元1012可基于层的数目的指示获得位流的层。如上文所描述,可缩放提取单元1012可获得位流21中指定的声道的数目的指示(例如,numHOATransportChannels),且通过至少部分地基于层的数目的指示和声道的指示获得位流21的层而获得所述层。

[0266] 在迭代经过每一层时,可缩放提取单元1012可首先通过获得NumFGchannels[i]语法元素而确定第i层的前景声道的数目。可缩放提取单元1012可接着从numAvailableTransportChannels减去NumFGchannels[i]以更新NumAvailableTransportChannels,且反射已从位流提取前景HOA声道61的NumFGchannels[i](其也可被称作“经编码nFG信号61”)。以此方式,可缩放提取单元1012可获得用于层中

的至少一者的在位流21中指定的前景声道的数目的指示(例如,NumFGchannels),且基于前景声道的数目的指示获得位流的层中的所述至少一者的前景声道。

[0267] 同样,可缩放提取单元1012可通过获得NumBGchannels[i]语法元素确定第i层的背景声道的数目。可缩放提取单元1012可接着从numAvailableTransportChannels减去NumBGchannels[i]以反射已从位流提取背景HOA声道59的NumBGchannels[i](其也可被称作“经编码环境HOA系数59”)。以此方式,可缩放提取单元1012可获得用于层中的至少一者的在位流21中指定的背景声道的数目的指示(例如,NumBGChannels),且基于背景声道的数目的指示获得位流的层中的所述至少一者的背景声道。

[0268] 可缩放提取单元1012可通过根据numAvailableTransports获得numAvailableTransportChannelsBits而继续。根据以上语法表,可缩放提取单元1012可解析由numAvailableTransportChannelsBits指定的位数以确定NumFGchannels[i]和NumBGchannels[i]。鉴于numAvailableTransportChannelBits改变(例如,在每一迭代之后变小),用来表示NumFGchannels[i]语法元素和NumBGchannels[i]语法元素的位数减少,由此提供潜在地减小用信号表示NumFGchannels[i]语法元素和NumBGchannels[i]语法元素的额外负荷的形式的可变长度译码。

[0269] 如上文所指出,可缩放位流产生单元1000可指定NumChannels语法元素代替NumFGchannels和NumBGchannels语法元素。在此情况下,可缩放提取单元1012可经配置以根据以上展示的第二HOADecoderConfig语法表操作。

[0270] 就此而言,在所述指示指示位流的层的数目当与所述位流在前一帧中的层的数目相比时在当前帧中已改变时,可缩放提取单元1012可基于层中的一者或者在前一帧的分量的数目获得层中的一者或者在当前帧的分量的数目的指示。可缩放提取单元1012可进一步基于分量的数目的指示获得一个或多个层在当前帧的背景分量的数目的指示。可缩放提取单元1012还可基于分量的数目的指示获得一个或多个层在当前帧的前景分量的数目的指示。

[0271] 鉴于层的数目可能从帧到帧改变,前景和背景声道的数目的指示可能从帧到帧改变,因此层的数目已改变的指示可有效地还指示声道的数目已改变。结果,层的数目已改变的指示可导致可缩放提取单元1012获得位流21中的一个或多个层中指定的声道的数目当与位流中的一个或多个层在前一帧指定的声道的数目相比时在当前帧中是否已改变的指示。由此,可缩放提取单元1012可基于位流中的一个或多个层中指定的声道的数目在当前帧中是否已改变的指示获得声道中的一者。

[0272] 此外,可缩放提取单元1012可在所述指示指示在位流21的一个或多个层中指定的声道的数目当与所述位流的一个或多个层中在前一帧中指定的声道的数目相比时在当前帧中尚未改变时,确定位流21的所述一个或多个层中在所述当前帧中指定的声道的数目与位流21的一个或多个层中在前一帧中指定的声道的数目相同。

[0273] 此外,在所述指示指示在所述位流21的所述一个或多个层中指定的声道的所述数目当与所述位流的所述一个或多个层中在所述前一帧中指定的声道的所述数目相比时在所述当前帧中尚未改变时,可缩放提取单元1012可获得所述层中的一者或者在所述当前帧的声道的当前数目与所述层中的一者或者在前一帧的声道的先前数目相同的指示。

[0274] 为实现可潜在地降低层和分量(其在本发明中也可被称作“声道”的数目的各种

指示的信令的前述技术,可缩放提取单元1012可将NumChannels_PrevFrame[i]语法元素设定为当前帧的指示(例如,NumChannels[i]),迭代经过所有i个层。此以以下语法表示:

	NumLayersPrevFrame=NumLayers;
[0275]	for i=1:NumLayers {
	NumChannels_PrevFrame[i] = NumChannels[i];

[0276]	}
--------	---

[0277] 或者,可省略前述语法(NumLayersPrevFrame=NumLayers等),且可如下表中所阐述更新上文所列的语法表HOADecoderConfig(numHOATransportChannels):

语法	位数	助记符
HOADecoderConfig(numHOATransportChannels)		
{		
HOALayerPresent;	1	bslbf
if(HOALayerPresent){		
NumLayerBits =		
ceil(log2(numHOATransportChannels-2));		
NumLayers = NumLayers +2;	NumLayerBits	uimsbf
numAvailableTransportChannels =		
numHOATransportChannels-2;		
numAvailableTransportChannelsBits =		
NumLayerBits;		
for (i=0; i<NumLayers-1; ++i) {		
NumChannels[i] =		
NumChannels[i]+1;	numAvailableTransportChannelsBits	uimsbf
numAvailableTransportChannels =		
numAvailableTransportChannels -		
NumChannels[i];		
numAvailableTransportChannelsBits =		
ceil(log2(numAvailableTransportChannels));		
}		
}		
MinAmbHoaOrder = escapedValue(3,5,0) - 1;	3,8	uimsbf
MinNumOfCoeffsForAmbHOA =		
(MinAmbHoaOrder + 1)^2;		
.		
.		
.		
}		

[0279] 作为又一替代方案,提取单元72可根据上文所列的第三HOADecoder Config而操作。根据上文所列的第三HOADecoderConfig语法表,可缩放提取单元1012可经配置以从可缩放位流21获得在位流中的一个或多个层中指定的声道的数目的指示,且基于声道的数目的指示获得在位流中的一个或多个层中指定的声道(其可指声场的背景分量或前景分量)。在这些和其它实例中,可缩放提取单元1012可经配置以获得指示声道的数目的语法元素(例如,以上提及的表中的codedLayerCh)。

[0280] 在这些和其它实例中,可缩放提取单元1012可经配置以获得在位流中指定的声道的总数目的指示。可缩放提取单元1012还可经配置以基于在一个或多个层中指定的声道的

数目的指示和声道的总数目的指示获得在一个或多个层中指定的声音。在这些和其它实例中,可缩放提取单元1012可经配置以获得指示声道的总数目的语法元素(例如,以上所提到的NumHOATransportChannels语法元素)。

[0281] 在这些和其它实例中,可缩放提取单元1012可经配置以获得在位流中的一个或多个层中指定的声音中的一者的类型的指示。可缩放提取单元1012还可经配置以基于层的数目的指示和声道中的一者的类型的指示获得声道中的一者。

[0282] 在这些和其它实例中,可缩放提取单元1012可经配置以获得在位流中的一个或多个层中指定的声音中的一者的类型的指示、指示声道中的一者为前景声道的声音中的一者的类型的指示。可缩放提取单元1012可经配置以基于层的数目的指示和声道中的一者的类型为前景声道的指示获得声道中的一者。在这些情况下,声道中的一者包括US音频对象和对应V向量。

[0283] 在这些和其它实例中,可缩放提取单元1012可经配置以获得在位流中的一个或多个层中指定的声音中的一者的类型的指示、指示声道中的一者为背景声道的声音中的一者的类型的指示。在这些情况下,可缩放提取单元1012可经配置以基于层的数目的指示和声道中的一者的类型为背景声道的指示获得声道中的一者。在这些情况下,声道中的一者包括背景高阶立体混响系数。

[0284] 在这些和其它实例中,可缩放提取单元1012可经配置以获得指示声道中的一者的类型的语法元素(例如,上文关于图30所描述的ChannelType语法元素)。

[0285] 在这些和其它实例中,可缩放提取单元1012可经配置以基于在获得层中的一者之后在位流中剩余的声音的数目获得声音的数目的指示。即,在整个循环过程同时,HOALayerChBits语法元素的值随如在以上语法表中所阐述的remainingCh语法元素而变。可缩放提取单元1012可接着基于改变的HOALayerChBits语法元素解析codedLayerCh语法元素。

[0286] 返回到四个背景声道和两个前景声道的实例,可缩放提取单元1012可接收层的数目为二(即,在图6的实例中,基础层21A和增强层21B)的指示。可缩放提取单元1012可获得前景声道的数目对于基础层21A为零(例如,从NumFGchannels[0])且对于增强层21B为二(例如,从NumFGchannels[1])的指示。在此实例中,可缩放提取单元1012还获得背景声道的数目对于基础层21A为四(例如,从NumBGchannels[0])且对于增强层21B为零(例如,从NumBGchannels[1])的指示。尽管是关于特定实例描述,但可指示背景与前景声道的任何不同组合。可缩放提取单元1012可接着从基础层21A提取所指定的四个背景声道59A到59D,且从增强层21B提取两个前景声道61A和61B(连同从边带信息提取对应V向量信息57A和57B)。

[0287] 尽管上文关于NumFGchannels和NumBGchannels语法元素进行描述,但还可以使用来自以上ChannelSideInfo语法表的ChannelType语法元素执行所述技术。就此而言,NumFGchannels和NumBG声道还可表示声道中的一者的类型的指示。换句话说,NumBGchannels可表示声道中的一者的类型为背景声道的指示。NumFG声道可表示声道中的一者的类型为前景声道的指示。

[0288] 由此,不管使用ChannelType语法元素还是NumFGchannels语法元素与NumBGchannels语法元素(或可能两者或任一者的某一子集),可缩放位流提取单元1012可获得在位流中的一个或多个层中指定的声音中的一者的类型的指示。在类型的指示指示声

道中的一者为背景声道时,可缩放位流提取单元1012可基于层的数目的指示和声道中的一者的类型为背景声道的指示获得声道中的一者。在类型的指示指示声道中的一者为前景声道时,可缩放位流提取单元1012可基于层的数目的指示和声道中的一者的类型为前景声道的指示获得声道中的一者。

[0289] 所述V向量重构单元74可表示被配置成从所述编码前景V[k]向量57重构所述V向量的单元。所述V向量重构单元74可以与所述量化单元52的运行方式互反的方式运行。

[0290] 心理声学解码单元80可以与图3的实例中所示的心理声学音频译码器单元40反向的方式操作,以便对经编码环境HOA系数59和经编码nFG信号61进行解码,且由此产生经调整环境HOA音频信号67'和经调整经内插nFG信号49"(其也可被称作经调整经内插nFG音频对象49')。心理声学解码单元80可将经调整环境HOA音频信号67'和经调整经内插nFG信号49"传递到逆增益控制单元86。

[0291] 逆增益控制单元86可表示经配置以对于经调整环境HOA音频信号67'和经调整经内插nFG信号49"中的每一者执行逆增益控制的单元,其中此逆增益控制与增益控制单元62执行的增益控制反向。逆增益控制单元86可根据在上文关于图11到13B的实例所论述的边带信息中指定的对应HOAGCD执行逆增益控制。逆增益控制单元86可将去相关环境HOA音频信号67输出到再相关单元88(在图4的实例中展示为“再相关单元88”),且将经内插nFG音频信号49"输出到前景制订单元78。

[0292] 再相关单元88可实施本发明的技术以降低去相关环境HOA音频信号67的背景声道之间的相关,以降低或缓解噪声去掩蔽。在再相关单元88将UHJ矩阵(例如,逆UHJ矩阵)应用为所选再相关变换的实例中,再相关单元81可通过减少数据处理操作而改善压缩率且节省计算资源。

[0293] 在一些实例中,可缩放位流21可包含指示在编码期间应用了去相关变换的一个或多个语法元素。将这些语法元素包含在基于向量的位流21中可使得再相关单元88能够对去相关环境HOA音频信号67执行反向去相关(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了哪一去相关变换,例如UHJ矩阵或模式矩阵,由此使得再相关单元88能够选择适当再相关变换应用于去相关HOA音频信号67。

[0294] 再相关单元88可对于去相关环境HOA音频信号67执行再相关以获得经能量补偿的环境HOA系数47'。再相关单元88可将经能量补偿的环境HOA系数47'输出到衰减单元770。尽管描述为执行去相关,但在一些实例中,可能不执行去相关。由此,基于向量的重构单元92可不执行,或在一些实例中包含再相关单元88。在一些实例中,不存在再相关单元88由再相关单元88的虚线表示。

[0295] 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。

[0296] 提取单元72还可将指示环境HOA系数中的一者何时转变的信号757输出到淡化单元770,所述淡化单元可接着确定哪些SHC_{BG} 47'(其中SHC_{BG} 47'还可以表示为“环境HOA声道47”或“环境HOA系数47”)和经内插前景V[k]向量55_k"的元素将淡入或淡出。在一些实例中,淡化单元770可对于环境HOA系数47'与经内插前景向量55_k"的元素中的每一者相反

地操作。即，淡化单元770可对于环境HOA系数47'中的对应一者执行淡入或淡出或淡入或淡出两者，而对于经内插前景V[k]向量 55_k "的元素中的所述对应一者执行淡入或淡出或淡入和淡出两者。淡化单元770可将经调整环境HOA系数47"输出到HOA系数制订单元82，且将经调整前景V[k]向量 55_k "输出到前景制订单元78。就此而言，淡化单元770表示经配置以对于HOA系数或其导出项(例如，呈环境HOA系数47'和经内插前景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。

[0297] 前景制订单元78可表示经配置以对于经调整前景V[k]向量 55_k "，和经内插nFG信号49'执行矩阵乘法以产生前景HOA系数65的单元。就此而言，前景制订单元78可组合音频对象49'(其为表示经内插nFG信号49'的另一方式)与向量 55_k "以重构前景，或换句话说，重构HOA系数11'的占优势方面。前景制订单元78可执行经内插nFG信号49'乘以经调整前景V[k]向量 55_k "的矩阵乘法。

[0298] HOA系数制订单元82可表示经配置以组合前景HOA系数65与经调整环境HOA系数47"以便获得HOA系数11'的单元。撇号记法反映HOA系数11'可类似于但不不同于HOA系数11。HOA系数11与11'之间的差异可源于归因于经由有损传输媒体的传输、量化或其它有损操作而造成的损失。

[0299] 图14A和14B是说明音频编码装置20在执行本发明中描述的技术的各种方面时的实例操作的流程图。首先参考图14A的实例，音频编码装置20可以上文所描述的方式(例如，线性分解、内插，等)获得HOA系数11的当前帧的声音(500)。所述声道可包括经编码环境HOA系数59、经编码nFG信号61(和呈经译码前景V向量57形式的对应边带)或经编码环境HOA系数59和经编码nFG信号61(和呈经译码前景V向量57形式的对应边带)两者。

[0300] 音频编码装置20的位流产生单元42可接着以上文所描述的方式在可缩放位流21中指定层的数目的指示(502)。位流产生单元42可在可缩放位流21的当前层中指定声道的子集(504)。位流产生单元42可维持用于当前层的计数器，其中所述计数器提供当前层的指示。在指定当前层中的声道之后，位流产生单元42可递增计数器。

[0301] 位流产生单元42可接着确定当前层(例如，计数器)是否大于位流中指定的层的数目(506)。在当前层不大于层的数目(“否”506)时，位流产生单元42可指定当前层中的声道的不同子集(其在计数器递增时改变)(504)。位流产生单元42可以此方式继续，直到当前层大于层的数目(“是”506)。在当前层大于层的数目(“是”506)时，位流产生单元可继续进行到下一帧，其中当前帧变为前一帧，且获得可缩放位流21的现在的当前帧的声音(500)。所述过程可继续，直到到达HOA系数11的最后帧(500到506)。如上文所指出，在一些实例中，层的数目的指示可能并非明确地指示，而是隐含地指定于可缩放位流21中(例如，在层的数目从前一帧到当前帧尚未改变时)。

[0302] 接下来参考图14B的实例，音频编码装置20可以上文所描述的方式(例如，线性分解、内插，等)获得HOA系数11的当前帧的声音(510)。所述声道可包括经编码环境HOA系数59、经编码nFG信号61(和呈经译码前景V向量57形式的对应边带)或经编码环境HOA系数59和经编码nFG信号61(和呈经译码前景V向量57形式的对应边带)两者。

[0303] 音频编码装置20的位流产生单元42可接着以上文所描述的方式在可缩放位流21的层中指定声道的数目的指示(512)。位流产生单元42可在可缩放位流21的当前层中指定对应声道(514)。

[0304] 位流产生单元42可接着确定当前层(例如,计数器)是否大于层的数目(516)。即,在图14B的实例中,层的数目可为静态或固定的(而非在可缩放位流21中指定),同时可指定每层的声道数目,这不同于图14A的实例,其中声道的数目可为静态或固定的且不用信号表示。位流产生单元42可仍维持指示当前层的计数器。

[0305] 在当前层(如由计数器指示)不大于层的数目(“否”516)时,位流产生单元42可在可缩放位流21的另一层中指定现在的当前层(其归因于递增计数器而改变)的声道数目的另一指示(512)。位流产生单元42还可在位流21的额外层中指定声道的对应数目(514)。位流产生单元42可以此方式继续,直到当前层大于层的数目(“是”516)。在当前层大于层的数目时(“是”516),位流产生单元可继续进行到下一帧,其中当前帧变为前一帧,且获得可缩放位流21的现在的当前帧的声音(510)。所述过程可继续,直到到达HOA系数11的最后帧(510到516)。

[0306] 如上文所指出,在一些实例中,声道的数目的指示可能并非明确地指示,而是隐含地指定于可缩放位流21中(例如,在层的数目从前一帧到当前帧尚未改变时)。此外,尽管描述为单独过程,但关于图14A和14B描述的技术可以上文所描述的方式组合地执行。

[0307] 图15A和15B是说明音频解码装置24在执行本发明中描述的技术的各种方面时的实例操作的流程图。首先参考图15A的实例,音频解码装置24可从可缩放位流21获得当前帧(520)。当前帧可包含一个或多个层,其中的每一个可包含一个或多个声道。所述声道可包括经编码环境HOA系数59、经编码nFG信号61(和呈经译码前景V向量57形式的对应边带)或经编码环境HOA系数59和经编码nFG信号61(和呈经译码前景V向量57形式的对应边带)两者。

[0308] 音频解码装置24的提取单元72可接着以上文所描述的方式获得可缩放位流21的当前帧中的层的数目的指示(522)。提取单元72可获得可缩放位流21的当前层中的声道的子集(524)。提取单元72可维持当前层的计数器,其中所述计数器提供当前层的指示。在指定当前层中的声道之后,提取单元72可递增计数器。

[0309] 提取单元72可接着确定当前层(例如,计数器)是否大于位流中指定的层的数目(526)。在当前层不大于层的数目时(“否”526),提取单元72可获得当前层(其在计数器递增时改变)中的声道的不同子集(524)。提取单元72可以此方式继续,直到当前层大于层的数目(“是”526)。在当前层大于层的数目时(“是”526),提取单元72可继续进行到下一帧,其中当前帧变为前一帧,且获得可缩放位流21的现在的当前帧(520)。所述过程可继续,直到到达可缩放位流21的最后帧(520到526)。如上文所指出,在一些实例中,层的数目的指示可能并非明确地指示,而是隐含地指定于可缩放位流21中(例如,在层的数目从前一帧到当前帧尚未改变时)。

[0310] 接下来参考图15B的实例,音频解码装置24可从可缩放位流21获得当前帧(530)。当前帧可包含一个或多个层,其中的每一个可包含一个或多个声道。所述声道可包括经编码环境HOA系数59、经编码nFG信号61(和呈经译码前景V向量57形式的对应边带)或经编码环境HOA系数59和经编码nFG信号61(和呈经译码前景V向量57形式的对应边带)两者。

[0311] 音频解码装置24的提取单元72可接着以上文所描述的方式获得可缩放位流21的层中的声道的数目的指示(532)。位流产生单元42可从可缩放位流21的当前层获得声道的对应数目(534)。

[0312] 提取单元72可接着确定当前层(例如,计数器)是否大于层的数目(536)。即,在图15B的实例中,层的数目可为静态或固定的(而非在可缩放位流21中指定),同时可指定每层的声道数目,这不同于图15A的实例,其中声道的数目可为静态或固定的且不用信号表示。提取单元72可仍维持指示当前层的计数器。

[0313] 在当前层(如由计数器指示)不大于层的数目(“否”536)时,提取单元72可在可缩放位流21的另一层中获得现在的当前层(其归因于递增计数器而改变)的声道数目的另一指示(532)。提取单元72还可在位流21的额外层中指定声道的对应数目(514)。提取单元72可以此方式继续,直到当前层大于层的数目(“是”516)。在当前层大于层的数目时(“是”516),位流产生单元可继续进行到下一帧,其中当前帧变为前一帧,且获得可缩放位流21的现在的当前帧的声道(510)。所述过程可继续,直到到达HOA系数11的最后帧(510到516)。

[0314] 如上文所指出,在一些实例中,声道的数目的指示可能并非明确地指示,而是隐含地指定于可缩放位流21中(例如,在层的数目从前一帧到当前帧尚未改变时)此外,尽管描述为单独过程,但关于图15A和15B描述的技术可以上文所描述的方式组合地执行。

[0315] 图16是说明根据本发明中描述的技术的各种方面的如由图16的实例中所示的位流产生单元42执行的可缩放音频译码的图。在图16的实例中,例如图2和3的实例中所示的音频编码装置20等HOA音频编码器可对HOA系数11(其也可以被称作“HOA信号11”)进行编码。HOA信号11可包括24个声道,每一声道具有1024个样本。如上文所指出,每一声道包含1024个样本,其可指对应于球基函数中的一者的1024个HOA系数。如上文对于图5的实例中所示的位流产生单元42所描述,音频编码装置20执行各种操作以从HOA信号11获得经编码环境HOA系数59(其也可以被称作“背景HOA声道59”)。

[0316] 如图16的实例中进一步展示,音频编码装置20获得背景HOA声道59作为HOA信号11的前四个声道。背景HOA声道59表示为 $H_{1:4}^{BG}$,其中1:4反映出选择HOA信号11的前四个声道来表示声场的背景分量。此声道选择可在语法元素中用信号表示为B=4。音频编码装置20的可缩放位流产生单元1000可接着在基础层21A(其可称为两个或更多个层中的第一层)中指定HOA背景声道59。

[0317] 可缩放位流产生单元1000可产生基础层21A以包含背景声道59和如根据以下方程式指定的增益信息:

$$[0318] H_1^{BG} \text{ (第1 BG声道音频信号)} + G_1^{BG} \text{ (第1 BG增益)}$$

$$[0319] H_2^{BG} \text{ (第2 BG声道音频信号)} + G_2^{BG} \text{ (第2 BG增益)}$$

$$[0320] H_3^{BG} \text{ (第3 BG声道音频信号)} + G_3^{BG} \text{ (第3 BG增益)}$$

[0321] 如图16的实例中进一步展示,音频编码装置20可获得F个前景HOA声道,其可表达为US音频对象和对应V向量。出于说明的目的,假定F=2。音频编码装置20因此可选择第一和第二US音频对象61(还可以称为“经编码nFG信号61”)以及第一和第二V向量57(也可以被称作“经译码前景V[k]向量57”),其中所述选择在图5的实例中分别表示为US_{1:2}和V_{1:2}。可缩放位流产生单元1000可接着产生可缩放位流21的第二层21B以包含第一和第二US音频对象61以及第一和第二V向量57。

[0322] 可缩放位流产生单元1000还可产生增强层21B以包含前景HOA声道61和增益信息连同根据以下方程式指定的V向量57:

[0323] US_1^{FG} (第1 FG声道音频信号) + G_1^{FG} (第1 FG增益)+ V_1^{FG} (第1 V向量)

[0324] US_2^{FG} (第2 FG声道音频信号) + G_2^{FG} (第2 FG增益)+ V_2^{FG} (第2 V向量)

[0325] US_3^{FG} (第3 FG声道音频信号) + G_3^{FG} (第3 FG增益)+ V_3^{FG} (第3 V向量)

[0326] 为从可缩放位流21'获得HOA系数11',图2和3的实例中所示的音频解码装置24可调用在图6的实例中更详细地展示的提取单元72。提取单元72可以上文关于图6所描述的方式提取经编码环境HOA系数59A到59D、经编码nFG信号61A和61B,以及经译码前景V[k]向量57A和57B。提取单元72可接着将经编码环境HOA系数59A到59D、经编码nFG信号61A和61B,以及经译码前景V[k]向量57A和57B输出到基于向量的解码单元92。

[0327] 基于向量的解码单元92可接着根据以下等式将US音频对象61乘以V向量57:

$$[0328] H_{1:F}^{FG} = \sum_{i=1}^F US_i V_i^T$$

$$[0329] \text{例如, } F = 2: H_{1:2}^{FG} = \sum_{i=1}^2 US_i V_i^T$$

[0330] 第一方程式提供对于F的通用操作的数学表达式。第二方程式提供所述实例中的数学表达式,其中假定F等于二。此乘法的结果表示为前景HOA信号1020。基于向量的解码单元92接着选择较高声道(考虑到最低四个系数已经选择作为HOA背景声道59),其中这些较高声道表示为 $H_{5:25}^{FG,1:2}$ 。换句话说,基于向量的解码单元92从前景HOA信号1020获得HOA前景声道65。

[0331] 结果,所述技术可促进可变分层(相对于需要层的静态数目)以适应大量译码上下文且潜在地提供指定声场的背景和前景分量时的更大灵活性。所述技术可提供许多其它使用情况,如关于图17到26所描述。这些各种使用情况可单独地执行或在给定音频流内一起执行。此外,在可缩放音频编码技术内指定这些分量可允许更多使用情况。换句话说,所述技术不应限于下文描述的使用情况,而可包含可借以在可缩放位流一个或多个层中用信号表示背景和前景分量的任何方式。

[0332] 图17是其中语法元素指示存在两个层的实例的概念图,其中在基础层中指定四个经编码环境HOA系数,且在增强层中指定两个经编码nFG信号。图17的实例展示HOA帧,此时图5的实例中所示的可缩放位流产生单元1000可分段所述帧以形成包含经编码环境HOA系数59A到59D的边带HOA增益校正数据的基础层。可缩放位流产生单元1000还可分段HOA帧以形成增强层21,所述增强层包含两个经译码前景V[k]向量57和用于经编码环境nFG信号61的HOA增益校正数据。

[0333] 如图17的实例中进一步展示,心理声学音频编码单元40展示为划分成心理声学音频编码器40A(其可称为基础层时间编码器40A)与心理声学音频编码器40B(其可称为增强层时间编码器40B)的单独例项。基础层时间编码器40A表示处理基础层的四个分量的四个心理声学音频编码器例项。增强层时间编码器40B表示处理增强层的两个分量的两个心理声学音频编码器例项。

[0334] 图18是更详细地说明在经配置以执行本发明中描述的可缩放音频译码技术的潜在版本中的第二者时的图3的位流产生单元42的图。在此实例中,位流产生单元42大体类似

于上文关于图5的实例所描述的位流产生单元42。然而,位流产生单元42执行可缩放译码技术的第二版本以指定三个层21A到21C而非两个层21A和21B。可缩放位流产生单元1000可指定在基础层21A中指定两个经编码环境HOA系数和零个经编码nFG信号的指示、在第一增强层21B中指定零个经编码环境HOA系数和两个经编码nFG信号的指示,以及在第二增强层21C中指定零个经编码环境HOA系数和两个经编码nFG信号61的指示。可缩放位流产生单元1000可接着在基础层21A中指定两个经编码环境HOA系数59A和59B,在第一增强层21B中指定具有对应的两个经译码前景V[k]向量57A和57B的两个经编码nFG信号61A和61B,且在第二增强层21C指定具有对应的两个经译码前景V[k]向量57C和57D的两个经编码nFG信号61C和61D。可缩放位流产生单元1000可接着输出这些层作为可缩放位流21。

[0335] 图19是更详细地说明在经配置以执行本发明中描述的可缩放音频解码技术的潜在版本中的第二者的图。在此实例中,位流提取单元72大体类似于上文关于图6的实例所描述的位流提取单元72。然而,位流提取单元72对于三个层21A到21C而非两个层21A和21B执行可缩放译码技术的第二版本。可缩放位流提取单元1012可获得在基础层21A中指定两个经编码环境HOA系数和零个经编码nFG信号的指示、在第一增强层21B中指定零个经编码环境HOA系数和两个经编码nFG信号的指示,以及在第二增强层21C中指定零个经编码环境HOA系数和两个经编码nFG信号的指示。可缩放位流提取单元1012可接着从基础层21A获得两个经编码环境HOA系数59A和59B,从第一增强层21B获得具有对应的两个经译码前景V[k]向量57A和57B的两个经编码nFG信号61A和61B,且从第二增强层21C获得具有对应的两个经译码前景V[k]向量57C和57D的两个经编码nFG信号61C和61D。可缩放位流提取单元1012可将经编码环境HOA系数59、经编码nFG信号61和经译码前景V[k]向量57输出到基于向量的解码单元92。

[0336] 图20是说明图18的位流产生单元和图19的提取单元可借以执行本发明中描述的技术的潜在版本中的第二者的第二使用情况的图。举例来说,图18的实例中所示的位流产生单元42可指定NumLayer(其为易于理解而展示为“NumberOfLayers”)语法元素以指示在可缩放位流21中指定的层的数目为三。位流产生单元42可进一步指定在第一层21A(其也称为“基础层”)中指定的背景声道的数目为二,而在第一层21B中指定的前景声道的数目为零(即,在图20的实例中,B₁=2,F₁=0)。位流产生单元42可进一步指定在第二层21B(其也称为“增强层”)中指定的背景声道的数目为零,而在第二层21B中指定的前景声道的数目为二(即,在图20的实例中,B₂=0,F₂=2)。位流产生单元42可进一步指定在第二层21C(其也称为“增强层”)中指定的背景声道的数目为零,而在第二层21C中指定的前景声道的数目为二(即,在图20的实例中,B₃=0,F₃=2)。然而,在前景和背景声道的总数目已经在解码器处知晓(例如,借助于额外语法元素,例如totalNumBGchannels和totalNumFGchannels)时,音频编码装置20可能不必用信号表示第三层背景和前景声道信息。

[0337] 位流产生单元42可将这些B_i和F_i值指定为NumBGchannels[i]和NumFGchannels[i]。对于以上实例,音频编码装置20可将NumBGchannels语法元素指定为{2,0,0},且将NumFGchannels语法元素指定为{0,2,2}。位流产生单元42还可在可缩放位流21中指定背景HOA音频声道59、前景HOA声道61和V向量57。

[0338] 图2和4的实例中所示的音频解码装置24可以与音频编码装置20的操作反向的方式来操作,以解析来自位流的这些语法元素(例如,如在以上HOADecoderConfig语法表中所

阐述),如上文关于图19的位流提取单元72所描述。音频解码装置24还可根据所解析语法元素解析来自位流21的对应背景HOA音频声道1002和前景HOA声道1010,再次如上文关于图19的位流提取单元72所描述。

[0339] 图21是其中语法元素指示存在三个层的实例的概念图,其中在基础层中指定两个经编码环境HOA系数,在第一增强层中指定两个经编码nFG信号,且在第二增强层中指定两个经编码nFG信号。图21的实例展示HOA帧,此时图18的实例中所示的可缩放位流产生单元1000可分段所述帧以形成包含用于经编码环境HOA系数59A和59B的边带HOA增益校正数据的基础层。可缩放位流产生单元1000还可分段HOA帧以形成包含用于经编码环境nFG信号61的两个经译码前景V[k]向量57和HOA增益校正数据的增强层21B以及包含用于经编码环境nFG信号61的两个额外经译码前景V[k]向量57和HOA增益校正数据的增强层21C。

[0340] 如图21的实例中进一步展示,心理声学音频编码单元40展示为划分成心理声学音频编码器40A(其可称为基础层时间编码器40A)与心理声学音频编码器40B(其可称为增强层时间编码器40B)的单独实例化。基础层时间编码器40A表示处理基础层的四个分量的两个心理声学音频编码器例项。增强层时间编码器40B表示处理增强层的两个分量的四个心理声学音频编码器例项。

[0341] 图22是更详细地说明在经配置以执行本发明中描述的可缩放音频译码技术的潜在版本中的第三者时的图3的位流产生单元42的图。在此实例中,位流产生单元42大体类似于上文关于图18的实例所描述的位流产生单元42。然而,位流产生单元42执行可缩放译码技术的第三版本以指定三个层21A到21C而非两个层21A和21B。可缩放位流产生单元1000可指定在基础层21A中指定零个经编码环境HOA系数和两个经编码nFG信号的指示、在第一增强层21B中指定零个经编码环境HOA系数和两个经编码nFG信号的指示,以及在第二增强层21C中指定零个经编码环境HOA系数和两个经编码nFG信号的指示。可缩放位流产生单元1000可接着在基础层21A中指定具有对应的两个经译码前景V[k]向量57A和57B的两个经编码nFG信号61A和61B,在第一增强层21B中指定具有对应的两个经译码前景V[k]向量57C和57D的两个经编码nFG信号61C和61D,且在第二增强层21C指定具有对应的两个经译码前景V[k]向量57E和57F的两个经编码nFG信号61E和61F。可缩放位流产生单元1000可接着输出这些层作为可缩放位流21。

[0342] 图23是更详细地说明在经配置以执行本发明中描述的可缩放音频解码技术的潜在版本中的第三者时的图4的提取单元72的图。在此实例中,位流提取单元72大体类似于文关于图19的实例所描述的位流提取单元72。然而,位流提取单元72对于三个层21A到21C而非两个层21A和21B执行可缩放译码技术的第三版本。此外,可缩放位流提取单元1012可获得在基础层21A中指定零个经编码环境HOA系数和两个经编码nFG信号的指示、在第一增强层21B中指定零个经编码环境HOA系数和两个经编码nFG信号的指示,以及在第二增强层21C中指定零个经编码环境HOA系数和两个经编码nFG信号的指示。可缩放位流提取单元1012可接着从基础层21A获得具有对应的两个经译码前景V[k]向量57A和57B的两个经编码nFG信号61A和61B,从第一增强层21B获得具有对应的两个经译码前景V[k]向量57C和57D的两个经编码nFG信号61C和61D,且从第二增强层21C获得具有对应的两个经译码前景V[k]向量57E和57F的两个经编码nFG信号61E和61F。可缩放位流提取单元1012可将经编码nFG信号61和经译码前景V[k]向量57输出到基于向量的解码单元92。

[0343] 图24是根据本发明中描述的技术的说明音频编码装置可借以在多层次流中指定多个层的第三使用情况的图。举例来说,图22的位流产生单元42可指定NumLayer(其为易于理解而展示为“NumberOfLayers”)语法元素以指示在位流21中指定的层的数目为三。位流产生单元42可进一步指定在第一层(其也称为“基础层”)中指定的背景声道的数目为零,而在第一层中指定的前景声道的数目为二(即,在图24的实例中, $B_1=0, F_1=2$)。换句话说,基础层并不始终仅用于传输环境HOA系数,而是可允许指定占优势或换句话说前景HOA音频信号。

[0344] 这两个前景音频声道表示为经编码nFG信号61A/B和经译码前景V[k]向量57A/B,且可在数学上由以下方程式表示:

$$[0345] H_{1:25}^{FG,1:2} = \sum_{i=1}^2 US_i V_i^T .$$

[0346] $H_{1:25}^{FG,1:2}$ 表示两个前景音频声道,其可由第一和第二音频对象(US_1 和 US_2)连同对应V向量(V_1 和 V_2)表示。

[0347] 位流产生装置42可进一步指定在第二层(其也称为“增强层”)中指定的背景声道的数目为零,而在第二层中指定的前景声道的数目为二(即,在图24的实例中, $B_2=0, F_2=2$)。这两个前景音频声道表示为经编码nFG信号61C/D和经译码前景V[k]向量57C/D,且可在数学上由以下方程式表示:

$$[0348] H_{1:25}^{FG,3:4} = \sum_{i=3}^4 US_i V_i^T .$$

[0349] $H_{1:25}^{FG,3:4}$ 表示两个前景音频声道,其可由第三和第四音频对象(US_3 和 US_4)连同对应V向量(V_3 和 V_4)表示。

[0350] 此外,位流产生单元42可指定在第三层(其也称为“增强层”)中指定的背景声道的数目为零,而在第三层中指定的前景声道的数目为二(即,在图24的实例中, $B_3=0, F_3=2$)。这两个前景音频声道表示为前景音频声道1024,且可在数学上由以下方程式表示:

$$[0351] H_{1:25}^{FG,5:6} = \sum_{i=5}^6 US_i V_i^T .$$

[0352] $H_{1:25}^{FG,5:6}$ 表示两个前景音频声道1024,其可由第五和第六音频对象(US_5 和 US_6)连同对应V向量(V_5 和 V_6)表示。然而,在前景和背景声道的总数目在解码器处已经知晓(例如,借助于额外语法元素,例如totalNumBGchannels和totalNumFGchannels)时,位流产生单元42可能不必以信号表示此第三层背景和前景声道信息。然而,在前景和背景声道的总数目在解码器处已经知晓(例如,借助于额外语法元素,例如totalNumBGchannels和totalNumFGchannels)时,位流产生单元42可能不以信号表示第三层背景和前景声道信息。

[0353] 位流产生单元42可将这些 B_i 和 F_i 值指定为NumBGchannels[i]和NumFGchannels[i]。对于以上实例,音频编码装置20可将NumBGchannels语法元素指定为{0,0,0},且将NumFGchannels语法元素指定为{2,2,2}。音频编码装置20还可在位流21中指定前景HOA声道1020到1024。

[0354] 图2和4的实例中所示的音频解码装置24可以与音频编码装置20的操作反向的方式来操作,以解析来自位流的这些语法元素(例如,如在以上HOADecoderConfig语法表中所

阐述),如上文关于图23的位流提取单元72所描述。音频解码装置24还可根据所解析语法元素解析来自位流21的对应前景HOA音频声道1020到1024,再次如上文关于图23的位流提取单元72所描述,且经由对前景HOA音频声道1020到1024进行求和来重构HOA系数1026。

[0355] 图25是其中语法元素指示存在三个层的实例的概念图,其中在基础层中指定两个经编码nFG信号,在第一增强层中指定两个经编码nFG信号,且在第二增强层中指定两个经编码nFG信号。图25的实例展示HOA帧,此时图22的实例中所示的可缩放位流产生单元1000可分段所述帧以形成包含用于经编码nFG信号61A和61B和两个经译码前景V[k]向量57的边带HOA增益校正数据的基础层。可缩放位流产生单元1000还可分段HOA帧以形成包含用于经编码环境nFG信号61的两个经译码前景V[k]向量57和HOA增益校正数据的增强层21B以及包含用于经编码环境nFG信号61的两个额外经译码前景V[k]向量57和HOA增益校正数据的增强层21C。

[0356] 如图25的实例中进一步展示,心理声学音频编码单元40展示为划分成心理声学音频编码器40A(其可称为基础层时间编码器40A)与心理声学音频编码器40B(其可称为增强层时间编码器40B)的单独实例化。基础层时间编码器40A表示处理基础层的四个分量的两个心理声学音频编码器例项。增强层时间编码器40B表示处理增强层的两个分量的四个心理声学音频编码器例项。

[0357] 图26是根据本发明中描述的技术的说明音频编码装置可借以在多层次位流中指定多个层的第三使用情况的图。举例来说,图2和3的实例中所示的音频编码装置20可指定NumLayer(其为易于理解而展示为“NumberOfLayers”)语法元素以指示在位流21中指定的层的数目为四。音频编码装置20可进一步指定在第一层(其也称为“基础层”)中指定的背景声道的数目为一,而在第一层中指定的前景声道的数目为零(即,在图26的实例中,B₁=1,F₁=0)。

[0358] 音频编码装置20可进一步指定在第二层(其也称为“第一增强层”)中指定的背景声道的数目为一,而在第二层中指定的前景声道的数目为零(即,在图26的实例中,B₂=1,F₂=0)。音频编码装置20还可指定在第三层(其也称为“第二增强层”)中指定的背景声道的数目为一,而在第三层中指定的前景声道的数目为零(即,在图26的实例中,B₃=1,F₃=0)。此外,音频编码装置20可指定在第四层(其也称为“增强层”)中指定的背景声道的数目为一,而在第三层中指定的前景声道的数目为零(即,在图26的实例中,B₄=1,F₄=0)。然而,在前景和背景声道的总数目已经在解码器处知晓(例如,借助于额外语法元素,例如totalNumBGchannels和totalNumFGchannels)时,音频编码装置20可能不必用信号表示第四层背景和前景声道信息。

[0359] 音频编码装置20可将这些B_i和F_i值指定为NumBGchannels[i]和NumFGchannels[i]。对于以上实例,音频编码装置20可将NumBGchannels语法元素指定为{1,1,1,1},且将NumFGchannels语法元素指定为{0,0,0,0}。音频编码装置20还可在位流21中指定背景HOA音频声道1030。就此而言,所述技术可允许增强层指定环境或换句话说背景HOA声道1030,其可在指定于位流21的基础和增强层中之前去相关,如上文关于图7A到9B的实例所描述。然而,再次,本发明中所阐述的技术不必受限于去相关,且可不在位流中提供与如上文所描述的去相关有关的语法元素或任何其它指示。

[0360] 图2和4的实例中所示的音频解码装置24可以与音频编码装置20的操作反向的方

式来操作,以解析来自位流的这些语法元素(例如,如在以上HOADecoderConfig语法表中所阐述)。音频解码装置24还可根据所解析语法元素解析来自位流21的对应背景HOA音频声道1030。

[0361] 如上文所指出,在一些情况下,可缩放位流21可包含符合非可缩放位流21的各种层。举例来说,可缩放位流21可包含符合非可缩放位流21的基础层。在这些情况下,非可缩放位流21可表示可缩放位流21的子位流,其中此非可缩放子位流21可利用可缩放位流21的额外层(其称为增强层)增强。

[0362] 图27和28是说明可经配置以执行本发明中描述的技术的各种方面的可缩放位流产生单元42和可缩放位流提取单元72的框图。在图27的实例中,可缩放位流产生单元42可表示上文关于图3的实例所描述的位流产生单元42的实例。可缩放位流产生单元42可输出符合(依据语法和由不支持可缩放译码的音频解码器解码的能力)非可缩放位流21的基础层21。可缩放位流产生单元42可以上文关于前述位流产生单位42中的任一者所描述的方式而操作,只是可缩放位流产生单元42不包含非可缩放位流产生单元1002。替代地,可缩放位流产生单元42输出符合非可缩放位流的基础层21,且因此不需要单独的非可缩放位流产生单元1000。在图28的实例中,可缩放位流提取单元72可与可缩放位流产生单元42往复操作。

[0363] 图29表示表示可经配置以根据本发明中描述的技术的各种方面而操作的编码器900的概念图。编码器900可表示音频编码装置20的另一实例。编码器900可包含空间分解单元902、去相关单元904和时间编码单元906。空间分解单元902可表示经配置以输出基于向量的占优势声音(呈上文提到的音频对象形式)、与这些基于向量的占优势声音相关联的对象V向量和水平环境HOA系数903的单元。空间分解单元902可不同于基于方向的分解,不同之处在于V向量描述音频对象中的对应一者的方向和宽度(随着每一音频对象随时间推移而在声场内移动)两者。

[0364] 空间分解单元902可包含图3的实例中所示的基于向量的合成单元27的单元30到38和44到52,且通常以上文关于单元30到38和44到52所描述的方式而操作。空间分解单元902可不同于基于向量的合成单元27,不同之处在于空间分解单元902可不执行心理声学编码,或另外包含心理声学译码器单元40且可不包含位流产生单元42。此外,在可缩放音频编码上下文中,空间分解单元902可通过水平环境HOA系数903(意味着,在一些实例中,这些水平HOA系数可能不修改或以其它方式调整,且从HOA系数901解析)。

[0365] 水平环境HOA系数903可指描述声场的水平分量的HOA系数901(其也可被称作HOA音频数据901)中的任一者。举例来说,水平环境HOA系数903可包含与具有阶数零和子阶数零的球基函数相关联的HOA系数、对应于具有阶数一和子阶数负一的球基函数的高阶立体混响系数,以及对应于具有阶数一和子阶数一的球基函数的第三高阶立体混响系数。

[0366] 去相关单元904表示经配置以对于高阶立体混响音频数据903(其中环境HOA系数903为此HOA音频数据的一个实例)的两个或更多个层中的第一层执行去相关以获得所述高阶立体混响音频数据的所述两个或更多个层中的第一层的经去相关表示905。基础层903可类似于上文关于图21到26所描述的第一层、基础层或基础子层中的任一者。去相关单元904可使用以上所提到的UHJ矩阵或模式矩阵执行去相关。去相关单元904还可使用例如旋转等变换以类似于2014年2月27日申请的标题为“变换球谐系数(TRANSFORMING SPHERICAL HARMONIC COEFFICIENTS)”的第14/192,829号美国申请案中描述的方式执行去相关,只是

执行旋转以获得第一层的经去相关表示而非减少系数的数目。

[0367] 换句话说,去相关单元904可执行声场的旋转以沿着分开120度的三个不同横轴(例如0度方位角/0度仰角,120度方位角/0度仰角,和240度方位角/0度仰角)对准环境HOA系数903的能量。通过对准这些能量与三个横轴,去相关单元904可尝试使能量彼此去相关,使得去相关单元904可利用空间变换有效地渲染三个去相关音频声道905。去相关单元904可应用此空间变换以便计算0度、120度和240度方位角处的空间音频信号905。

[0368] 尽管关于0度、120度和240度方位角进行描述,但可关于均匀或接近均匀地划分圆的360度方位角的任何三个方位角来应用所述技术。举例来说,还可以关于计算60度、180度和300度的方位角处的空间音频信号905的变换执行所述技术。此外,尽管关于三个环境HOA系数901进行描述,但可更一般地对于任何水平HOA系数执行所述技术,包含如上文所描述的那些系数和任何其它水平HOA系数,例如与具有阶数二和子阶数二的球基函数、具有阶数二和子阶数负二的球基函数、具有阶数X和子阶数X的球基函数以及具有阶数X和子阶数负X的球基函数相关联的那些系数,其中X可表示任何数目,包含3、4、5、6,等。

[0369] 随着水平HOA系数的数目增大,360度圆的均匀或接近均匀部分的数目可增大。举例来说,在水平HOA系数的数目增大到五时,去相关单元904可将圆分段为五个均匀分区(例如,各自大致为72度)。作为另一实例,水平HOA系数的数目X可导致X个均匀分区,每一分区具有360度/X度。

[0370] 为识别指示由水平环境HOA系数903表示的声场旋转量,去相关单元904可执行声场分析、内容特性分析及/或空间分析。基于这些分析中的一者或更多者,去相关单元904可将旋转信息(或其它变换信息,其中旋转信息为一个实例)识别为声场的水平旋转度数,且旋转所述声场,从而有效地获得高阶立体混响音频数据的基础层的经旋转表示(其为更通用经变换表示的一个实例)。

[0371] 去相关单元904可接着将空间变换应用于高阶立体混响音频数据的基础层903(其也可被称作两个或更多个层中的第一层903)的经旋转表示。空间变换可将高阶立体混响音频数据的两个或更多个层中的基础层的经旋转表示从球面谐波域转换到空间域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示。第一层的去相关表示可包含在0度、120度和240度的三个对应方位角处渲染的空间音频信号905,如上文所指出。去相关单元904可接着将水平环境空间音频信号905传递到时间编码单元906。

[0372] 时间编码单元906可表示经配置以执行心理声学音频译码的单元。时间编码单元906可表示AAC编码器或统一话音和音频译码器(USAC),提供两个实例。例如时间编码单元906等时间音频编码单元可对于去相关音频数据(例如5.1扬声器设置的6个声道,这6个声道已渲染到去相关声道)正常操作。然而,水平环境HOA系数903在性质上是累加性的,且由此在某些方面相关。将这些水平环境HOA系数903直接提供到时间编码单元906而不首先执行某一形式的去相关可能会导致空间噪声去掩蔽,其中声音出现在不希望的位置。例如空间噪声去掩蔽的这些感知伪声可通过执行上文所描述的基于变换(或更具体来说,在图29的实例中,基于旋转)的去相关来减少。

[0373] 图30是更详细地说明图27的实例中所示的编码器900的图。在图30的实例中,编码器900可表示基础层编码器900,其仅对HOA一阶水平基础层903进行编码,且不展示空间分解单元902,因为此单元902在此通过实例中不执行除了将基础层903提供到声场分析单元

910和去相关单元904的二维(2D)旋转单元912之外的有意义操作。

[0374] 即,去相关单元904包含声场分析单元910和2D旋转单元912。声场分析单元910表示经配置以执行上文更详细描述的声场分析以获得旋转角参数911的单元。旋转角参数911表示呈旋转信息形式的变换信息的一个实例。2D旋转单元912表示经配置以基于旋转角参数911执行绕着声场的Z轴的水平旋转的单元。此旋转为二维的,因为旋转仅涉及单个旋转轴,且在此实例中不包含任何仰角旋转。2D旋转单元912可获得逆旋转信息913(作为一个实例,通过反转旋转角参数911以获得逆旋转角参数913),其可为更通用的逆变换信息的实例。2D旋转单元912可提供逆旋转角参数913,使得编码器900可在位流中指定逆旋转角参数913。

[0375] 换句话说,2D旋转单元912可基于声场分析旋转2D声场,以使得占优势能量潜在地从2D空间变换模块中使用的空间取样点($0^\circ, 120^\circ, 240^\circ$)中的一者到达。作为一个实例,2D旋转单元912可应用以下旋转矩阵:

$$[0376] \begin{bmatrix} 1 & 0 & 0 \\ \cos(\phi) & 0 & \sin(\phi) \\ -\sin(\phi) & 0 & \cos(\phi) \end{bmatrix}$$

[0377] 在一些实例中,为避免帧伪声,2D旋转单元912可应用平滑(内插)函数以确保时变旋转角的平滑转变。此平滑函数可包括线性平滑函数。然而,可使用其它平滑函数,包含非线性平滑函数。举例来说,2D旋转单元912可使用仿样平滑函数(spline smoothing function)。

[0378] 为进行说明,在声场分析单元910模块指示声场的优势方向在一个分析帧内为 70° 方位角处时,2D旋转单元912可根据 $\phi = -70^\circ$ 平滑地旋转声场,以使得优势方向现在为 0° 。作为另一可能性,2D旋转单元912可根据 $\phi = 50^\circ$ 旋转声场,以使得优势方向现在为 120° 。2D旋转单元912可接着将所应用的旋转角913以信号表示为位流内的额外边带参数,以使得解码器可应用正确的逆旋转操作。

[0379] 如图30的实例中进一步展示,去相关单元904还包含2D空间变换单元914。2D空间变换单元914表示经配置以将基础层的经旋转表示从球面谐波域转换到空间域,从而有效地将经旋转基础层915渲染到三个方位角(例如, $0, 120$ 和 240)的单元。2D空间变换单元914可将经旋转基础层915的系数与以下变换矩阵相乘,其采用HOA系数阶数“00+”、“11-”、“11+”和N3D正规化:

$$[0380] \begin{bmatrix} 1/3 & 0 & 0.384900179459750 \\ 1/3 & 1/3 & -0.192450089729875 \\ 1/3 & -1/3 & -0.192450089729875 \end{bmatrix}$$

[0381] 前述矩阵计算方位角 $0^\circ, 120^\circ$ 和 240° 处的空间音频信号905,以使得 360° 圆被均匀划分为3部分。如上文所指出,其它分离是可能的,只要每一部分涵盖120度即可,例如,计算 $60^\circ, 180^\circ$ 和 300° 处的空间信号。

[0382] 以此方式,所述技术可提供经配置以执行可缩放高阶立体混响音频数据编码的装置900。装置900可经配置以对于高阶立体混响音频数据的两个或更多个层中的第一层903执行去相关,以获得高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905。

[0383] 在这些和其它实例中,高阶立体混响音频数据的两个或更多个层中的第一层903包括对应于具有等于或小于一的阶数的一个或多个球基函数的环境高阶立体混响系数。在这些和其它实例中,高阶立体混响音频数据的两个或更多个层中的第一层903包括仅对应于描述声场的水平方面的球基函数的环境高阶立体混响系数。在这些和其它实例中,仅对应于描述声场的水平方面的球基函数的环境高阶立体混响系数可包括对应于具有阶数零和子阶数零的球基函数的第一环境高阶立体混响系数、对应于具有阶数一和子阶数负一的球基函数的第二高阶立体混响系数,以及对应于具有阶数一和子阶数一的球基函数的第三高阶立体混响系数。

[0384] 在这些和其它实例中,装置900可经配置以对于高阶立体混响音频数据的第一层903执行变换(例如,借助于2D旋转单元912)。

[0385] 在这些和其它实例中,装置900可经配置以对于高阶立体混响音频数据的第一层903执行旋转(例如,借助于2D旋转单元912)。

[0386] 在这些和其它实例中,装置900可经配置以对于高阶立体混响音频数据的两个或更多个层中的第一层903应用变换(例如,借助于2D旋转单元912)以获得高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915,且将高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915从球面谐波域转换(例如,借助于2D空间变换单元914)到空间域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905。

[0387] 在这些和其它实例中,装置900可经配置以对于高阶立体混响音频数据的两个或更多个层中的第一层903应用旋转以获得高阶立体混响音频数据的两个或更多个层中的第一层的经旋转表示915,且将高阶立体混响音频数据的两个或更多个层中的第一层的经旋转表示915从球面谐波域转换到空间域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905。

[0388] 在这些和其它实例中,装置900可经配置以获得变换信息911,基于变换信息911对于高阶立体混响音频数据的两个或更多个层中的第一层903应用变换以获得高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915,且将高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915从球面谐波域转换到空间域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905。

[0389] 在这些和其它实例中,装置900可经配置以获得旋转信息911,且基于旋转信息911对于高阶立体混响音频数据的两个或更多个层中的第一层903应用旋转以获得高阶立体混响音频数据的两个或更多个层中的第一层的经旋转表示915,且将高阶立体混响音频数据的两个或更多个层中的第一层的经旋转表示915从球面谐波域转换到空间域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905。

[0390] 在这些和其它实例中,装置900可经配置以至少部分地使用平滑函数对于高阶立体混响音频数据的两个或更多个层中的第一层903应用变换以获得高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915,且将高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915从球面谐波域转换到空间域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905。

[0391] 在这些和其它实例中,装置900可经配置以至少部分地使用平滑函数对于高阶立

体混响音频数据的两个或更多个层中的第一层903应用旋转以获得高阶立体混响音频数据的两个或更多个层中的第一层的经旋转表示915,且将高阶立体混响音频数据的两个或更多个层中的第一层的经旋转表示915从球面谐波域转换到空间域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示。

[0392] 在这些和其它实例中,装置900可经配置以指定将在应用逆变换或逆旋转时使用的平滑函数的指示。

[0393] 在这些和其它实例中,装置900可进一步经配置以将线性可逆变换应用于高阶立体混响音频数据以获得V向量,且将所述V向量指定为高阶立体混响音频数据的两个或更多个层中的第二层,如上文关于图3所描述。

[0394] 在这些和其它实例中,装置900可进一步经配置以获得与具有阶数一和子阶数零的球基函数相关联的高阶立体混响系数,且将高阶立体混响系数指定为高阶立体混响音频数据的两个或更多个层中的第二层。

[0395] 在这些和其它实例中,装置900可进一步经配置以对于高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示执行时间编码。

[0396] 图31是说明可经配置以根据本发明中描述的技术的各种方面而操作的音频解码器920的框图。在重构HOA系数、重构增强层的V向量、执行时间音频解码(如由时间音频解码单元922执行)等方面,解码器920可表示图2的实例中所示的音频解码装置24的另一实例。然而,解码器920不同,不同之处在于解码器920如位流中所指定而对于可缩放经译码高阶立体混响音频数据进行操作。

[0397] 如图31的实例中所示,音频解码器920包含时间解码单元922、逆2D空间变换单元924、基础层渲染单元928和增强层处理单元930。时间解码单元922可经配置而以与时间编码单元906的操作反向的方式进行操作。逆2D空间变换单元924可表示经配置而以与2D空间变换单元914的操作反向的方式进行操作的单元。

[0398] 换句话说,逆2D空间变换单元924可经配置以将以下矩阵应用于空间音频信号905以获得经旋转水平环境HOA系数915(其也可被称作“经旋转基础层915”)。逆2D空间变换单元924可使用以下变换矩阵将3个所发射音频信号905向后变换为HOA域,所述变换矩阵类似于以上矩阵,采用HOA系数阶数“00+”、“11-”、“11+”和N3D正规化:

$$[0399] \begin{bmatrix} 1.0 & 1.0 & 1.0 \\ 0.0 & 1.5 & -1.5 \\ 1.732050807568878 & -0.866025403784438 & -0.866025403784440 \end{bmatrix}$$

[0400] 前述矩阵为解码器中使用的变换矩阵的倒置。

[0401] 逆2D旋转单元926可经配置而以与上文关于2D旋转单元912所描述的操作反向的方式进行操作。就此而言,2D旋转单元912可基于逆旋转角参数913而非旋转角参数911根据上文提到的旋转矩阵执行旋转。换句话说,逆旋转单元926可基于用信号表示的旋转 ϕ 而应用以下矩阵,所述矩阵再次采用HOA系数阶数“00+”、“11-”、“11+”和N3D正规化:

$$[0402] \begin{bmatrix} 1 & 0 & 0 \\ \cos(\phi) & 0 & \sin(\phi) \\ -\sin(\phi) & 0 & \cos(\phi) \end{bmatrix}$$

[0403] 逆2D旋转单元926可使用与解码器中使用的相同的平滑(内插)函数来确保时变旋

转角的平滑转变,其可用信号表示于位流中或先验地配置。

[0404] 基础层渲染单元928可表示经配置以将基础层的仅水平环境HOA系数渲染到扩音器馈给的单元。增强层处理单元930可表示经配置以利用任何所接收增强层(经由涉及上文关于额外环境HOA系数和V向量连同对应于V向量的音频对象所描述的单独增强层解码路径解码)对基础层执行进一步处理以渲染扬声器馈给的单元。增强层处理单元930可有效地扩增基础层以提供声场的较高分辨率表示,其可提供使声音潜在地在声场内实际移动的更具沉浸性的音频体验。基础层可类似于上文关于图11到13B所描述的第一层、基础层或基础子层中的任一者。增强层可类似于上文关于图11到13B所描述的第二层、增强层或增强子层中的任一者。

[0405] 就此而言,所述技术提供经配置以执行可缩放高阶立体混响音频数据解码的装置920。所述装置可经配置以获得高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示(例如,空间音频信号905),所述高阶立体混响音频数据描述声场。第一层的经去相关表示通过对于高阶立体混响音频数据的第一层执行去相关而去相关。

[0406] 在一些情况下,高阶立体混响音频数据的两个或更多个层中的第一层包括对应于具有等于或小于一的阶数的一个或多个球基函数的环境高阶立体混响系数。在这些和其它实例中,高阶立体混响音频数据的两个或更多个层中的第一层包括仅对应于描述声场的水平方面的球基函数的环境高阶立体混响系数。在这些和其它实例中,仅对应于描述声场的水平方面的球基函数的环境高阶立体混响系数包括对应于具有阶数零和子阶数零的球基函数的第一环境高阶立体混响系数、对应于具有阶数一和子阶数负一的球基函数的第二高阶立体混响系数,以及对应于具有阶数一和子阶数一的球基函数的第三高阶立体混响系数。

[0407] 在这些和其它实例中,第一层的经去相关表示通过对于高阶立体混响音频数据的第一层执行变换而去相关,如上文关于编码器900所描述。

[0408] 在这些和其它实例中,装置920可经配置以对于高阶立体混响音频数据的第一层执行旋转(例如,通过逆2D旋转单元926)。

[0409] 在这些和其它实例中,装置920可经配置以使高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示重新相关,以获得高阶立体混响音频数据的两个或更多个层中的第一层,如上文例如关于逆2D空间变换单元924和逆2D旋转单元926所描述。

[0410] 在这些和其它实例中,装置920可经配置以将高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905从空间域转换到球面谐波域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915,且对于高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915应用逆变换(例如,如上文关于逆2D旋转单元926所描述)以获得高阶立体混响音频数据的两个或更多个层中的第一层。

[0411] 在这些和其它实例中,装置920可经配置以将高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905从空间域转换到球面谐波域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915,且对于高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915应用逆旋转以获得高阶立体混响音频数据的两个或更多个层中的第一层。

[0412] 在这些和其它实例中,装置920可经配置以将高阶立体混响音频数据的两个或更

多个层中的第一层的经去相关表示905从空间域转换到球面谐波域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915,获得变换信息913,且基于变换信息913对于高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915应用逆变换以获得高阶立体混响音频数据的两个或更多个层中的第一层。

[0413] 在这些和其它实例中,装置920可经配置以将高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905从空间域转换到球面谐波域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915,获得旋转信息913,且基于旋转信息913对于高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915应用逆旋转以获得高阶立体混响音频数据的两个或更多个层中的第一层。

[0414] 在这些和其它实例中,装置920可经配置以将高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905从空间域转换到球面谐波域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915,且至少部分地使用平滑函数对于高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915应用逆变换以获得高阶立体混响音频数据的两个或更多个层中的第一层。

[0415] 在这些和其它实例中,装置920可经配置以将高阶立体混响音频数据的两个或更多个层中的第一层的经去相关表示905从空间域转换到球面谐波域以获得高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915,且至少部分地使用平滑函数对于高阶立体混响音频数据的两个或更多个层中的第一层的经变换表示915应用逆旋转以获得高阶立体混响音频数据的两个或更多个层中的第一层。

[0416] 在这些和其它实例中,装置920可进一步经配置以获得将在应用逆变换或逆旋转时使用的平滑函数的指示。

[0417] 在这些和其它实例中,装置920可进一步经配置以获得高阶立体混响音频数据的两个或更多个层中的第二层的表示,其中第二层的所述表示包括基于向量的占优势音频数据,基于向量的音频数据至少包括占优势音频数据和经编码V向量,且经编码V向量是经由应用线性可逆变换而从高阶立体混响音频数据分解而来,如上文关于图3的实例所描述。

[0418] 在这些和其它实例中,装置920可进一步经配置以获得高阶立体混响音频数据的两个或更多个层中的第二层的表示,其中第二层的所述表示包括与具有阶数一和子阶数零的球基函数相关联的高阶立体混响系数。

[0419] 以此方式,所述技术可使得装置能够经配置以执行以下条项中所阐述的方法,或提供包括用于执行以下条项中所阐述的方法的装置的设备或其上存储的指令在执行时致使一个或多个处理器执行以下条项中所阐述的方法的非暂时性计算机可读媒体。

[0420] 条项1A。一种对高阶立体混响音频信号进行编码以产生位流的方法,所述方法包括在所述位流中指定层的数目的指示,以及输出包含所述层的所述所指示数目的所述位流。

[0421] 条项2A。根据条项1A所述的方法,其进一步包括指定包含于所述位流中的声道的数目的指示。

[0422] 条项3A。根据条项1A所述的方法,其中层的所述数目的所述指示包括所述位流中用于前一帧的层数目的指示,且其中所述方法进一步包括在所述位流中指定所述位流的层的数目当与所述位流的用于所述前一帧的层的所述数目相比时是否已改变的指示,以及指

定所述位流在所述当前帧中的层的所述所指示数目。

[0423] 条项4A。根据条项3A所述的装置,其中指定层的所述所指示数目包括,在所述指示指示所述位流的层的所述数目当与所述位流在所述前一帧中的层的所述数目相比时在所述当前帧中尚未改变时,指定层的所述所指示数目而不在所述位流中指定所述当前帧的所述层中的一者或者中的背景分量的当前数目等于所述前一帧的所述层中的一者或者中的背景分量的先前数目的指示。

[0424] 条项5A。根据条项1A所述的方法,其中所述层为阶层式的,使得第一层在与第二层组合时提供所述高阶立体混响音频信号的较高分辨率表示。

[0425] 条项6A。根据条项1A所述的方法,其中所述位流的所述层包括基础层和增强层,且其中所述方法进一步包括对于所述基础层的一个或多个声道应用去相关变换以获得所述高阶立体混响音频信号的背景分量的经去相关表示。

[0426] 条项7A。根据条项6A所述的方法,其中所述去相关变换包括UHJ变换。

[0427] 条项8A。根据条项6A所述的方法,其中所述去相关变换包括模式矩阵变换。

[0428] 此外,所述技术可使得装置能够经配置以执行以下条项中所阐述的方法,或提供包括用于执行以下条项中所阐述的方法的装置的设备或其上存储的指令在执行时致使一个或多个处理器执行以下条项中所阐述的方法的非暂时性计算机可读媒体。

[0429] 条项1B。一种对高阶立体混响音频信号进行编码以产生位流的方法,所述方法包括:在所述位流中指定在所述位流的一个或多个层中指定的声音的数目的指示;以及在所述位流的所述一个或多个层中指定所述声道的所述所指示数目。

[0430] 条项2B。根据条项1B所述的方法,其进一步包括指定在所述位流中指定的声音的总数目的指示,其中指定声道的所述所指示数目包括在所述位流的所述一个或多个层中指定所述声道的所述所指示总数目。

[0431] 条项3B。根据条项1B所述的方法,其进一步包括指定在所述位流中的所述一个或多个层中指定的所述声道中的一者的类型的指示,且指定声道的所述所指示数目包括在所述位流的所述一个或多个层中指定所述声道中的所述一者的所述所指示类型的所指示数目。

[0432] 条项4B。根据条项1B所述的方法,其进一步包括指定在所述位流中的所述一个或多个层指定的所述声道中的一者的类型的指示,所述声道中的所述一者的所述类型的所述指示指示所述声道中的所述一者为前景声道,且其中指定声道的所述所指示数目包括在所述位流的所述一个或多个层中指定所述前景声道。

[0433] 条项5B。根据条项1B所述的方法,其进一步包括在所述位流中指定在所述位流中指定的层的数目的指示。

[0434] 条项6B。根据条项1B所述的方法,其进一步包括指定在所述位流中的所述一个或多个层中指定的所述声道中的一者的类型的指示,所述声道中的所述一者的所述类型的所述指示指示所述声道中的所述一者为背景声道,其中指定声道的所述所指示数目包括在所述位流的所述一个或多个层中指定所述背景声道。

[0435] 条项7B。根据条项6B所述的方法,其中所述声道中的所述一者包括背景高阶立体混响系数。

[0436] 条项1B。根据条项1B所述的方法,其中指定声道的所述数目的所述指示包括基于

在指定所述层中的一者之后在所述位流中剩余的声道的数目指定声道的所述数目的所述指示。

[0437] 以此方式,所述技术可使得装置能够经配置以执行以下条项中所阐述的方法,或提供包括用于执行以下条项中所阐述的方法的装置的设备或其上存储的指令在执行时致使一个或多个处理器执行以下条项中所阐述的方法的非暂时性计算机可读媒体。

[0438] 条项1C。一种对表示高阶立体混响音频信号的位流进行解码的方法,所述方法包括:从所述位流获得在所述位流中指定的层的数目的指示,且基于层的所述数目的所述指示而获得所述位流的所述层。

[0439] 条项2C。根据条项1C所述的方法,其进一步包括获得在所述位流中指定的声道的总数目的指示,且其中获得所述层包括基于层的所述数目的所述指示和声道的所述数目的所述指示获得所述位流的所述层。

[0440] 条项3C。根据条项1C所述的方法,其进一步包括获得所述位流中针对所述层中的至少一个指定的前景声道的数目的指示,且其中获得所述层包括基于前景声道的所述数目的所述指示获得所述位流的所述层中的所述至少一个的所述前景声道。

[0441] 条项4C。根据条项1C所述的方法,其进一步包括获得所述位流中针对所述层中的至少一者指定的背景声道的数目的指示,且其中获得所述层包括基于背景声道的所述数目的所述指示获得所述位流的所述层中的所述至少一者的所述背景声道。

[0442] 条项5C。根据条项1C所述的方法,其中所述层的所述数目的所述指示指示层的所述数目为二,其中所述两个层包括基础层和增强层,且其中获得所述层包括获得前景声道的数目对于所述基础层为零且对于所述增强层为二的指示。

[0443] 条项6C。根据条项1C或5C所述的方法,其中所述层的所述数目的所述指示指示层的所述数目为二,其中所述两个层包括基础层和增强层,且其中所述方法进一步包括获得背景声道的数目对于所述基础层为四且对于所述增强层为零的指示。

[0444] 条项7。根据条项1C所述的方法,其中所述层的所述数目的所述指示指示层的所述数目为三,其中所述三个层包括基础层、第一增强层和第二增强层,且其中所述方法进一步包括获得前景声道的数目对于所述基础层为零、对于所述第一增强层为二且对于所述第三增强层为二的指示。

[0445] 条项8C。根据条项1C或7C所述的方法,其中所述层的所述数目的所述指示指示层的所述数目为三,其中所述三个层包括基础层、第一增强层和第二增强层,且其中所述方法进一步包括获得背景声道的数目对于所述基础层为二、对于所述第一增强层为零且对于所述第三增强层为零的指示。

[0446] 条项9C。根据条项1C所述的方法,其中所述层的所述数目的所述指示指示层的所述数目为三,其中所述三个层包括基础层、第一增强层和第二增强层,且其中所述方法进一步包括获得前景声道的数目对于所述基础层为零、对于所述第一增强层为二且对于第三增强层为二的指示。

[0447] 条项10C。根据条项1C或9C所述的方法,其中所述层的所述数目的所述指示指示层的所述数目为三,其中所述三个层包括基础层、第一增强层和第二增强层,且其中所述方法进一步包括获得指示背景声道的数目对于所述基础层为零、对于所述第一增强层为零且对于所述第三增强层为零的背景语法元素。

[0448] 条项11C。根据条项1C所述的方法,其中层的所述数目的所述指示包括所述位流在前一帧中的层的数目的指示,且其中所述方法进一步包括获得所述位流的层的数目当与所述位流在所述前一帧中的层的所述数目相比时在当前帧中是否已改变的指示,以及基于所述位流的层的所述数目在所述当前帧中是否已改变的所述指示获得所述位流在所述当前帧中的层的所述数目。

[0449] 条项12C。根据条项11C所述的方法,其进一步包括在所述指示指示所述位流的层的所述数目当与所述位流在所述前一帧中的层的所述数目相比时在所述当前帧中尚未改变时,确定所述位流在所述当前帧中的层的所述数目等于所述位流在所述前一帧中的层的所述数目。

[0450] 条项13C。根据条项11C所述的方法,其中方法进一步包括,在所述指示指示所述位流的层的所述数目当与所述位流在所述前一帧中的层的所述数目相比时在所述当前帧中尚未改变时,获得所述当前帧的所述层中的一者或多者的分量的当前数目等于所述前一帧的所述层中的一个或多个中的分量的先前数目的指示。

[0451] 条项14C。根据条项1C所述的方法,其中层的所述数目的所述指示指示在所述位流中指定了三个层,且其中获得所述层包括获得指示所述高阶立体混响音频信号的提供立体声声道回放的背景分量的所述位流的所述层中的第一者,获得指示所述高阶立体混响音频信号的提供由布置在一个或多个水平面上的三个或更多个扬声器进行三维回放的背景分量的所述位流的所述层中的第二者,且获得指示所述高阶立体混响音频信号的前景分量的所述位流的所述层中的第三者。

[0452] 条项15C。根据条项1C所述的方法,其中层的所述数目的所述指示指示在所述位流中指定了三个层,且其中获得所述层包括获得指示所述高阶立体混响音频信号的提供单声道回放的背景分量的所述位流的所述层中的第一者,获得指示所述高阶立体混响音频信号的提供由布置在一个或多个水平面上的三个或更多个扬声器进行三维回放的背景分量的所述位流的所述层中的第二者,且获得指示所述高阶立体混响音频信号的前景分量的所述位流的所述层中的第三者。

[0453] 条项16C。根据条项1C所述的方法,其中层的所述数目的所述指示指示在所述位流中指定了三个层,且其中获得所述层包括获得指示所述高阶立体混响音频信号的提供立体声声道回放的背景分量的所述位流的所述层中的第一者,获得指示所述高阶立体混响音频信号的提供由布置在单个水平面上的三个或更多个扬声器进行多声道回放的背景分量的所述位流的所述层中的第二者,获得指示所述高阶立体混响音频信号的提供由布置在两个或更多个水平面上的三个或更多个扬声器进行三维回放的背景分量的所述位流的所述层中的第三者,且获得指示所述高阶立体混响音频信号的前景分量的所述位流的所述层中的第四者。

[0454] 条项17C。根据条项1C所述的方法,其中层的所述数目的所述指示指示在所述位流中指定了三个层,且其中获得所述层包括获得指示所述高阶立体混响音频信号的提供单声道回放的背景分量的所述位流的所述层中的第一者,获得指示所述高阶立体混响音频信号的提供由布置在单个水平面上的三个或更多个扬声器进行多声道回放的背景分量的所述位流的所述层中的第二者,且获得指示所述高阶立体混响音频信号的提供由布置在两个或更多个水平面上的三个或更多个扬声器进行三维回放的背景分量的所述位流的所述层中

的第三者,且获得指示所述高阶立体混响音频信号的前景分量的所述位流的所述层中的第四者。

[0455] 条项18C。根据条项1C所述的方法,其中层的所述数目的所述指示指示在所述位流中指定了两个层,且其中获得所述层包括获得指示所述高阶立体混响音频信号的提供立体声声道回放的背景分量的所述位流的所述层中的第一者,且获得指示所述高阶立体混响音频信号的提供由布置在单个水平面上的三个或更多个扬声器进行水平多声道回放的背景分量的所述位流的所述层中的第二者。

[0456] 条项19C。根据条项1C所述的方法,其进一步包括获得在所述位流中指定的声音的总数目的指示,其中获得所述层包括基于层的所述数目的所述指示和声道的所述数目的所述指示获得所述位流的所述层。

[0457] 条项20C。根据条项1C所述的方法,其进一步包括获得所述位流中针对所述层中的至少一个指定的前景声道的数目的指示,其中获得所述层包括基于前景声道的所述数目的所述指示获得所述位流的所述层中的所述至少一个的所述前景声道。

[0458] 条项21C。根据条项1C所述的方法,其进一步包括获得所述位流中针对所述层中的至少一者指定的背景声道的数目的指示,其中获得所述层包括基于背景声道的所述数目的所述指示获得所述位流的所述层中的所述至少一者的所述背景声道。

[0459] 条项22C。根据条项1C所述的方法,其进一步包括基于在获得所述层中的至少一者之后在所述位流中剩余的声音的数目解析所述位流中针对所述层中的所述至少一者指定的前景声道的数目的指示,其中获得所述层包括基于前景声道的所述数目的所述指示获得所述层中的所述至少一者的所述前景声道。

[0460] 条项23C。根据条项22C所述的方法,其中在获得所述层中的所述至少一者之后在所述位流中剩余的声音的所述数目由语法元素表示。

[0461] 条项24C。根据条项1C所述的方法,其进一步包括基于在获得所述层中的至少一者之后的声音的数目解析所述位流中针对所述层中的所述至少一者指定的背景声道的数目的指示,其中获得所述背景声道包括基于背景声道的所述数目的所述指示获得所述层中的所述至少一者的所述背景声道。

[0462] 条项25C。根据条项24C所述的方法,其中在获得所述层中的所述至少一者之后在所述位流中剩余的声音的所述数目由语法元素表示。

[0463] 条项26C。根据条项1C所述的方法,其中所述位流的所述层包括基础层和增强层,且其中所述方法进一步包括对于所述基础层的一个或多个声道应用相关变换以获得所述高阶立体混响音频信号的背景分量的相关表示。

[0464] 条项27C。根据条项26C所述的方法,其中所述相关变换包括逆UHJ变换。

[0465] 条项28C。根据条项26C所述的方法,其中所述相关变换包括逆模式矩阵变换。

[0466] 条项29C。根据条项1C所述的方法,其中所述位流的所述层中的每一者的声道的数目是固定的。

[0467] 此外,所述技术可使得装置能够经配置以执行以下条项中所阐述的方法,或提供包括用于执行以下条项中所阐述的方法的装置的设备或其上存储的指令在执行时致使一个或多个处理器执行以下条项中所阐述的方法的非暂时性计算机可读媒体。

[0468] 条项1D。一种对表示高阶立体混响音频信号的位流进行解码的方法,所述方法包

括：从所述位流获得在所述位流中的一个或多个层中指定的信道的数目的指示；以及基于信道的所述数目的所述指示而获得在所述位流中的所述一个或多个层中指定的所述信道。

[0469] 条项2D。根据条项1D所述的方法，其进一步包括获得在所述位流中指定的声道的总数目的指示，且其中获得所述声道包括基于在所述一个或多个层中指定的声道的所述数目的所述指示和声道的所述总数目的所述指示获得所述一个或多个层中指定的所述声道。

[0470] 条项3D。根据条项1D所述的方法，其进一步包括获得在所述位流中的所述一个或多个层中指定的所述声道中的一者的类型的指示，且其中获得所述声道包括基于声道的所述数目的所述指示和所述声道中的所述一者的所述类型的所述指示获得所述声道中的所述一者。

[0471] 条项4D。根据条项1D所述的方法，其进一步包括获得在所述位流中的所述一个或多个层中指定的所述声道中的一者的类型的指示，所述声道中的所述一者的所述类型的所述指示指示所述声道中的所述一者为前景声道，且其中获得所述声道包括基于声道的所述数目的所述指示和所述声道中的所述一者的所述类型为所述前景声道的所述指示获得所述声道中的所述一者。

[0472] 条项5D。根据条项1D所述的方法，其进一步包括获得在所述位流中指定的层的数目的指示，且其中获得所述声道包括基于声道的所述数目的所述指示和层的所述数目的所述指示获得所述声道中的所述一者。

[0473] 条项6D。根据条项5D所述的方法，其中层的所述数目的所述指示包括所述位流在前一帧中的层的数目的指示，其中所述方法进一步包括获得在所述位流中的一个或多个层中指定的声道的所述数目当与所述位流中的一个或多个层中在所述前一帧中指定的声道的数目相比时在当前帧中是否已改变的指示，且其中获得所述声道包括基于在所述位流中的一个或多个层中指定的声道的所述数目在所述当前帧中是否已改变的所述指示获得所述声道中的所述一者。

[0474] 条项7D。根据条项5D所述的方法，其进一步包括在所述指示指示在所述位流的所述一个或多个层中指定的声道的所述数目当与所述位流的所述一个或多个层中在所述前一帧中指定的声道的所述数目相比时在所述当前帧中尚未改变时，确定所述位流的所述一个或多个层中在所述当前帧中指定的声道的所述数目与所述位流的所述一个或多个层中在所述前一帧中指定的声道的所述数目相同。

[0475] 条项8D。根据条项5D所述的方法，其中所述一个或多个处理器进一步经配置以在所述指示指示在所述位流的所述一个或多个层中指定的声道的所述数目当与所述位流的所述一个或多个层中在所述前一帧中指定的声道的所述数目相比时在所述当前帧中尚未改变时，获得所述层中的一者或者在所述当前帧的声道的当前数目与所述层中的一者或者在前一帧的声道的先前数目相同的指示。

[0476] 条项9D。根据条项1D所述的方法，其进一步包括获得在所述位流中的所述一个或多个层中指定的所述声道中的一者的类型的指示，所述声道中的所述一者的所述类型的所述指示指示所述声道中的所述一者为背景声道，且其中获得所述声道包括基于层的所述数目的所述指示和所述声道中的所述一者的所述类型为所述背景声道的所述指示获得所述声道中的所述一者。

[0477] 条项10D。根据条项9D所述的方法，其进一步包括获得在所述位流中的所述一个或

多个层中指定的所述声道中的一者的类型的指示，所述声道中的所述一者的所述类型的所述指示指示所述声道中的所述一者为背景声道，且其中获得所述声道包括基于层的所述数目的所述指示和所述声道中的所述一者的所述类型为所述背景声道的所述指示获得所述声道中的所述一者。

[0478] 条项11D。根据条项9D所述的方法，其中所述声道中的所述一者包括背景高阶立体混响系数。

[0479] 条项12D。根据条项9D所述的方法，其中获得所述声道中的所述一者的所述类型的所述指示包括获得指示所述声道中的所述一者的所述类型的语法元素。

[0480] 条项13D。根据条项1D所述的方法，其中获得声道的所述数目的所述指示包括基于在获得所述层中的一者之后在所述位流中剩余的声道的数目获得声道的所述数目的所述指示。

[0481] 条项14D。根据条项1D所述的方法，其中所述层包括基础层。

[0482] 条项15D。根据条项1D所述的方法，其中所述层包括基础层和一个或多个增强层。

[0483] 条项16D。根据条项1D所述的方法，其中所述一个或多个层的数目是固定的。

[0484] 前述技术可对于任何数目的不同上下文和音频生态系统来执行。下文描述多个实例上下文，但所述技术应限于实例上下文。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、基于声道的音频内容、译码引擎、游戏音频原声、游戏音频编码/渲染引擎和递送系统。

[0485] 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)输出基于声道的音频内容(例如，呈2.0、5.1和7.1)。音乐工作室可例如通过使用DAW输出基于声道的音频内容(例如，呈2.0和5.1)。在任一情况下，译码引擎可基于一个或多个编解码器(例如，AAC、AC3、杜比真HD、杜比数字加和DTS主音频)接收并编码基于声道的音频内容以供传递系统输出。游戏音频工作室可例如通过使用DAW输出一个或多个游戏音频原声。游戏音频译码/渲染引擎可译码音频原声和/或将音频原声渲染到基于声道的音频内容中以供递送系统输出。可执行所述技术的另一实例情形包括音频生态系统，其可包含广播录音音频对象、专业音频系统、消费型装置上俘获、HOA音频格式、装置上渲染、消费型音频、TV和配件以及汽车音频系统。

[0486] 广播录音音频对象、专业音频系统和消费型装置上俘获都可使用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，可使用装置上渲染、消费型音频、TV和配件以及汽车音频系统来回放所述单一表示。换句话说，可在通用音频播放系统(即，与需要例如5.1、7.1等的特定配置相比)(例如，音频播放系统16)处播放音频内容的单一表示。

[0487] 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)、装置上环绕声俘获和移动装置(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及/或无线通信声道耦合到移动装置。

[0488] 根据本发明的一个或多个技术，所述移动装置可用于获取声场。举例来说，移动装置可经由有线和/或无线获取装置和/或装置上环绕声俘获(例如，集成到移动装置中的多

个麦克风)获取声场。移动装置可接着将所获取的声场译码成HOA系数以用于由回放元件中的一者或者回放。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),且将记录译码成HOA系数。

[0489] 移动装置还可利用回放元件中的一者或者来回放经HOA译码的声场。举例来说,移动装置可对经HOA译码的声场进行解码,且将致使回放元件中的一者或者重新产生声场的信号输出到回放元件中的一者或者。作为一个实例,移动装置可利用无线和/或无线通信声道将信号输出到一个或多个扬声器(例如,扬声器阵列、声棒等)。作为另一实例,移动装置可利用对接解决方案将信号输出到一个或多个对接站和/或一个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,移动装置可利用头戴式耳机渲染将信号输出到一组头戴式耳机(例如)以创建真实的双耳声音。

[0490] 在一些实例中,特定移动装置可获取3D声场并且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,且将经编码的3D声场发射到一个或多个其它装置(例如,其它移动装置和/或其它非移动装置)以用于回放。

[0491] 可执行所述技术的又一情形包含音频生态系统,其可包含音频内容、游戏工作室、经译码音频内容、渲染引擎和递送系统。在一些实例中,游戏工作室可包含可支持HOA信号的编辑的一个或多个DAW。举例来说,所述一个或多个DAW可包含可经配置以与一个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渲染引擎,所述渲染引擎可渲染声场以供递送系统回放。

[0492] 也可相对于示范性音频获取装置执行所述技术。举例来说,可相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的半径的实质上球面球的表面上。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。

[0493] 另一示范性音频获取情形可包含可经配置以从一个或多个麦克风(例如,一个或多个本征麦克风)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。

[0494] 在一些情况下,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个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。

[0495] 加固型视频俘获装置可进一步经配置以记录3D声场。在一些实例中,加固型视频俘获装置可附接到参与活动的用户的头盔。举例来说,加固型视频俘获装置可在用户泛舟时附接到用户的头盔。以此方式,加固型视频俘获装置可俘获表示用户环境的动作(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。

[0496] 还可相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,其中添加了一个或多个配件。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,配件增强型移动装置可俘获3D声场的较高质量版本,而不是仅使用与附件增强型移动装置成一体式的音频俘获组件。

[0497] 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,可利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渲染声场。

[0498] 多个不同实例音频回放环境还可适合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,以及具有耳机回放环境的移动装置。

[0499] 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渲染声场。另外,本发明的技术使得渲染器能够从通用表示渲染声场以供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置放(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渲染器能够用其它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。

[0500] 此外,用户可在佩戴头戴式耳机时观看体育比赛。根据本发明的一或多种技术,可获取体育比赛的3D声场(例如,可将一个或多个本征麦克风放置在棒球场中和/或环境),可获得对应于3D声场的HOA系数,且将所述HOA系数发射到解码器,所述解码器可基于HOA系数重构3D声场且将经重构3D声场输出到渲染器,所述渲染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声场渲染为致使头戴式耳机输出所述体育比赛的3D声场的表示的信号。

[0501] 在上述各种实例中的每一者中,应理解,音频编码装置20可执行一种方法,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。在一些情况下,这些装置可包括一个或多个处理器。在一些情况下,所述一个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个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。

[0502] 在一个或多个实例中,所描述功能可用硬件、软件、固件或其任何组合来实施。如果以软件实施,那么所述功能可作为一个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个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。计算机程序产品可包含计算机可读媒体。

[0503] 同样地,在上文所述的各种实例中的每一者中,应理解,音频解码装置24可执行方法或以其它方式包括执行所述音频解码装置24经配置以执行的所述方法的每一步骤的装置。在一些情况下,所述装置可包括一个或多个处理器。在一些情况下,所述一个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个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。

[0504] 作为实例而非限制,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、快闪存储器或可用来存储指令或数据结构的形式的所要程序代码并且可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。

[0505] 可由例如一个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个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”可指前述结构或适合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在经配置以用于编码和解码的专用硬件和/或软件模块内提供,或者并入在组合式编解码器中。而且,所述技术可完全实施于一个或多个电路或逻辑元件中。

[0506] 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公开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个或多个处理器。

[0507] 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。

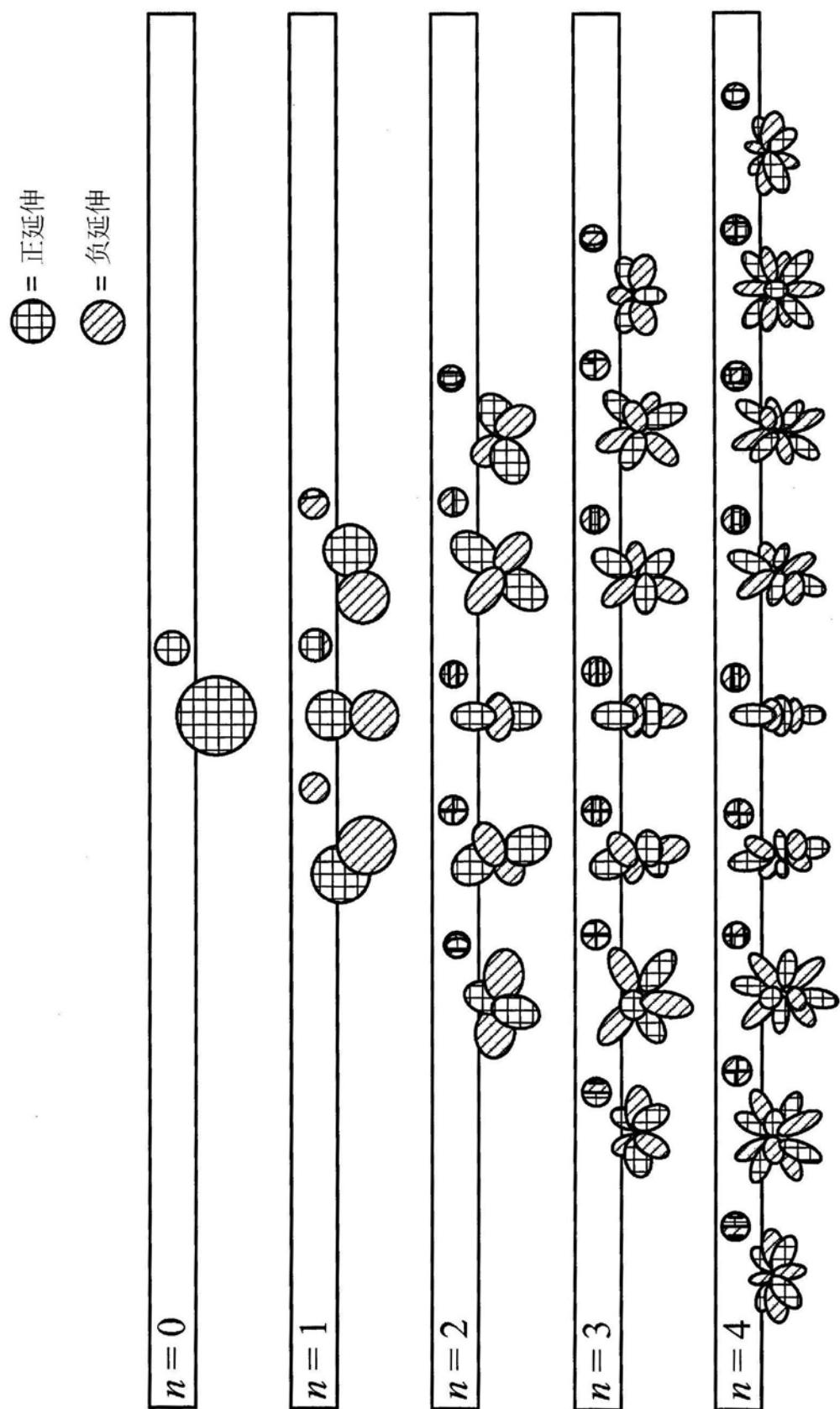


图1

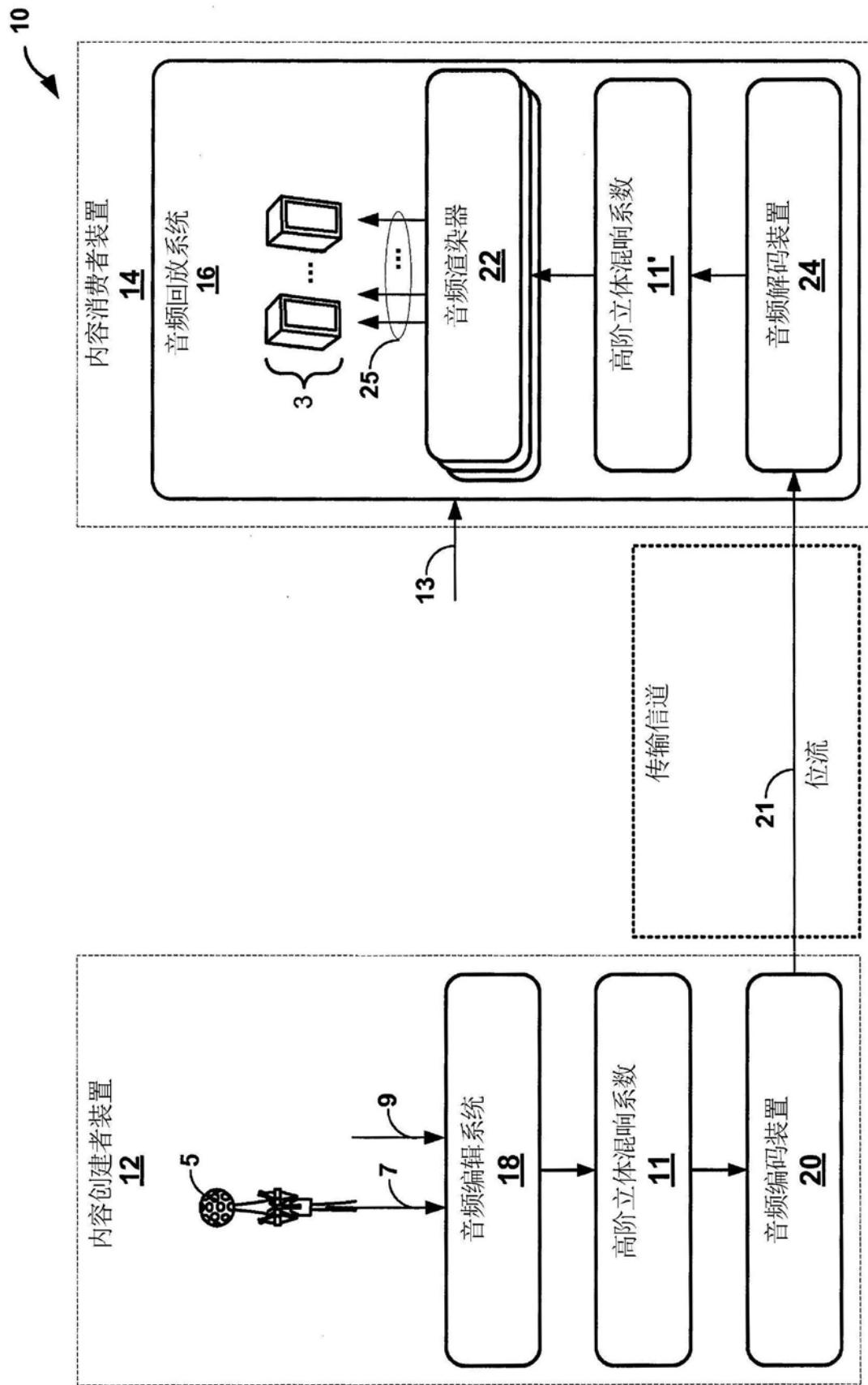


图2

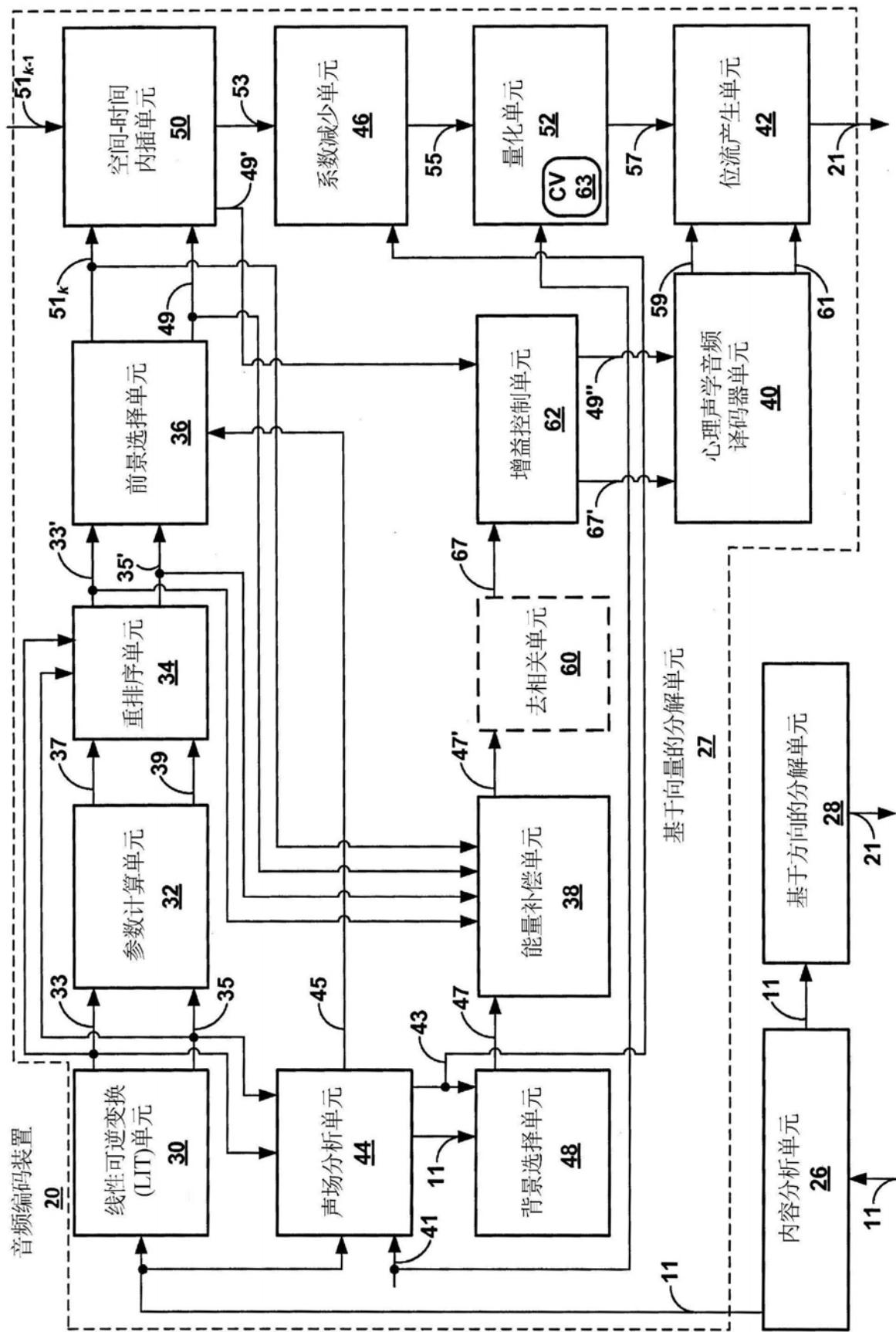


图3

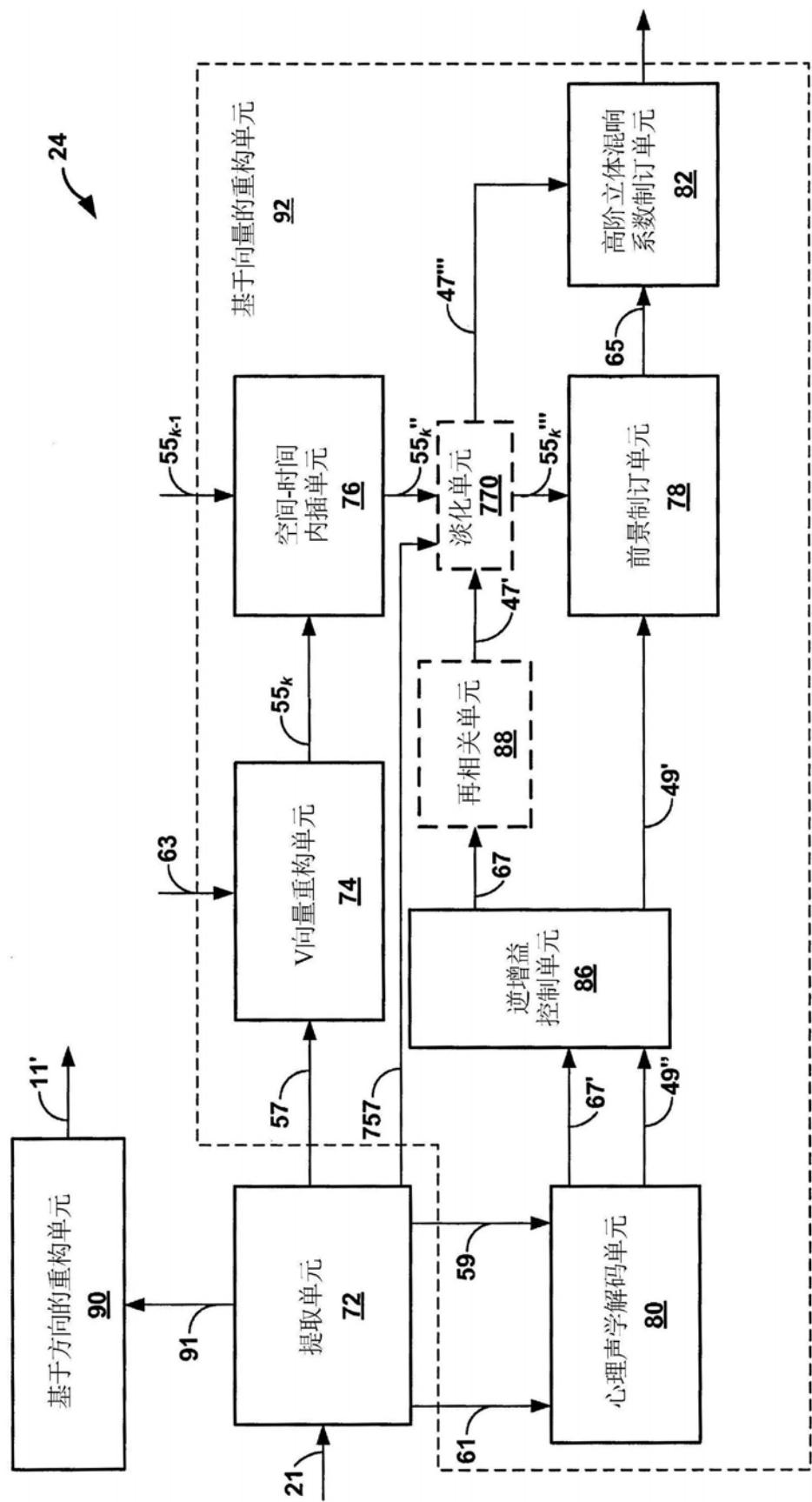


图4

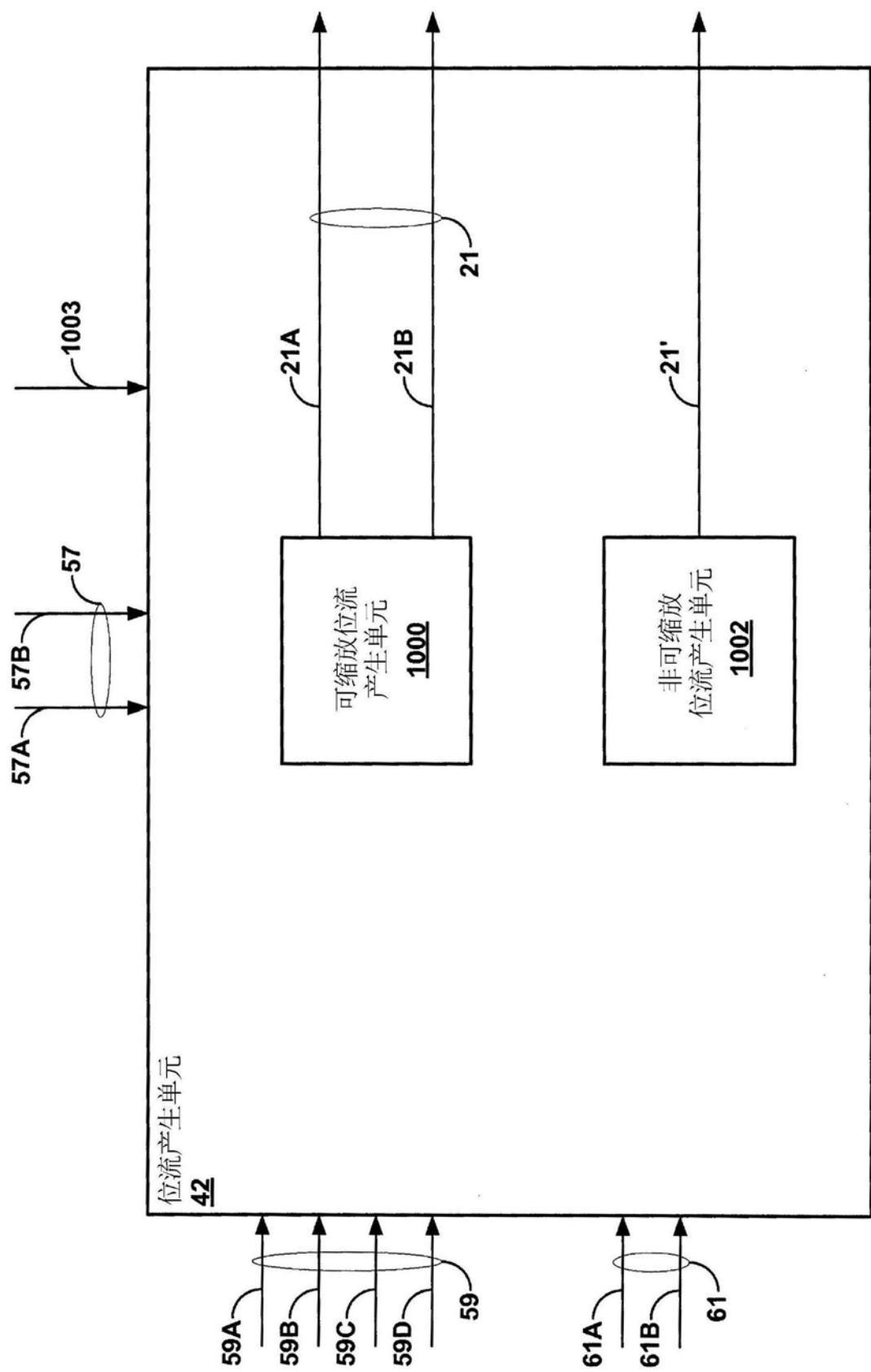


图5

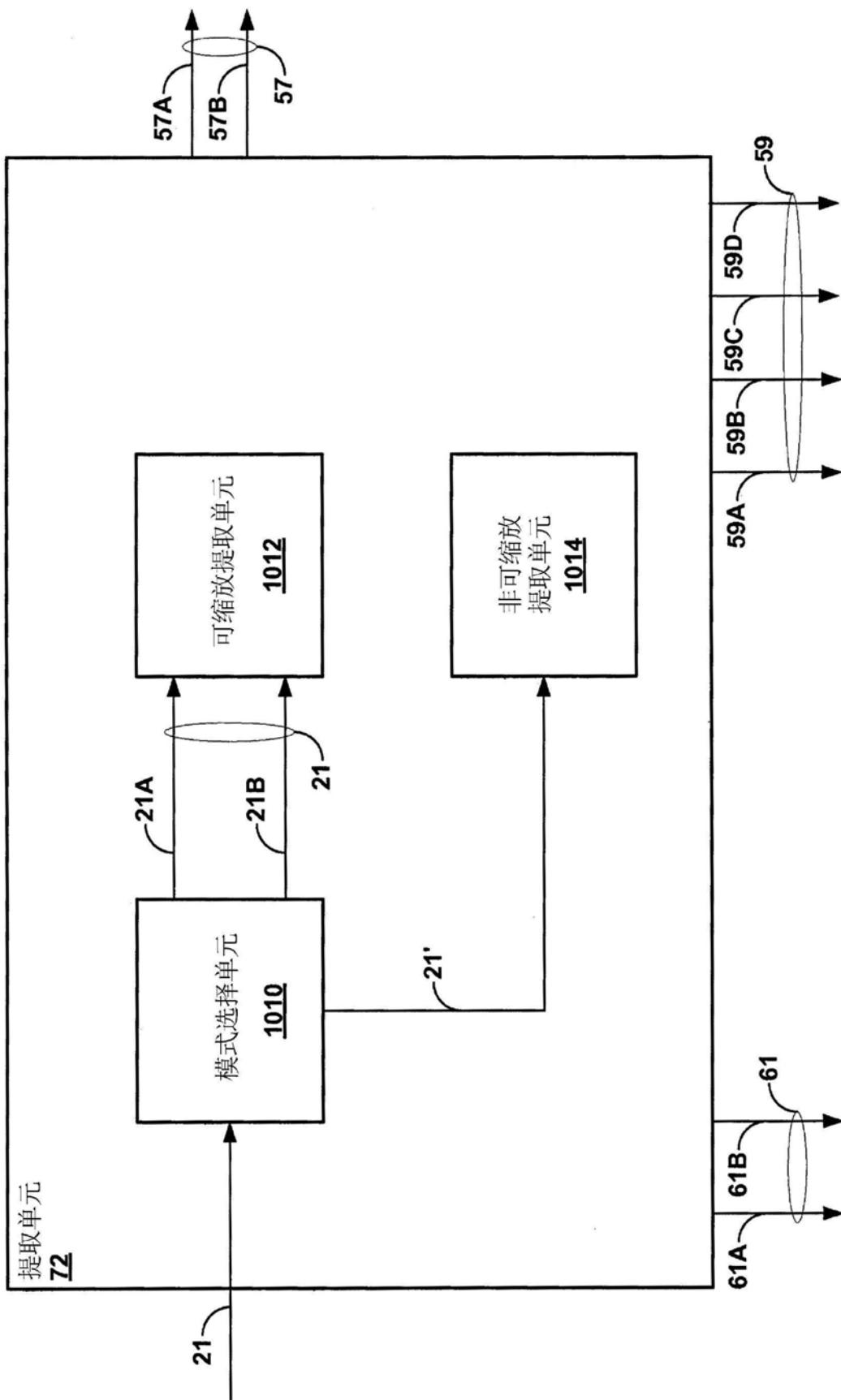


图6

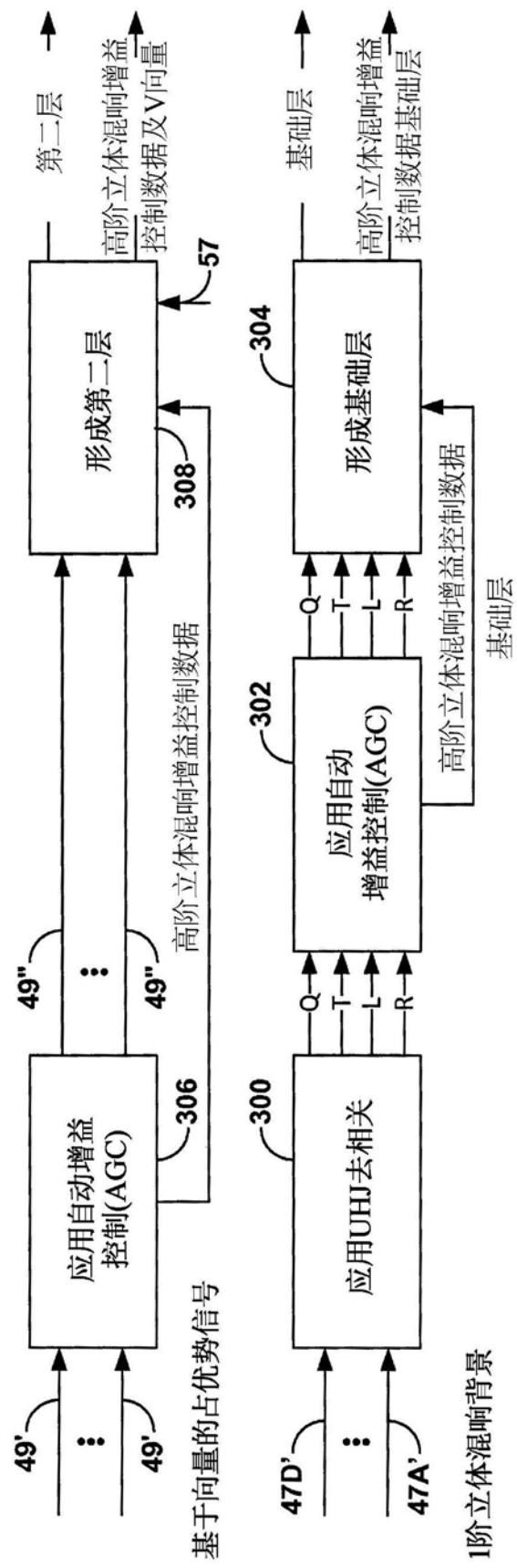


图7A

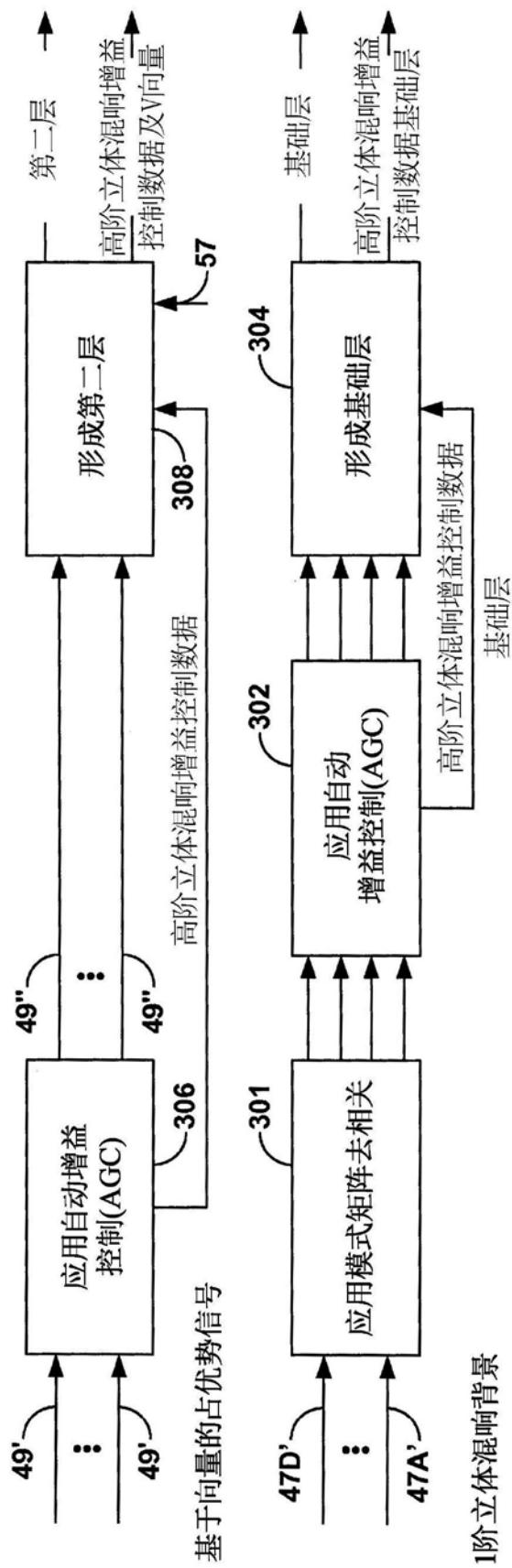


图7B

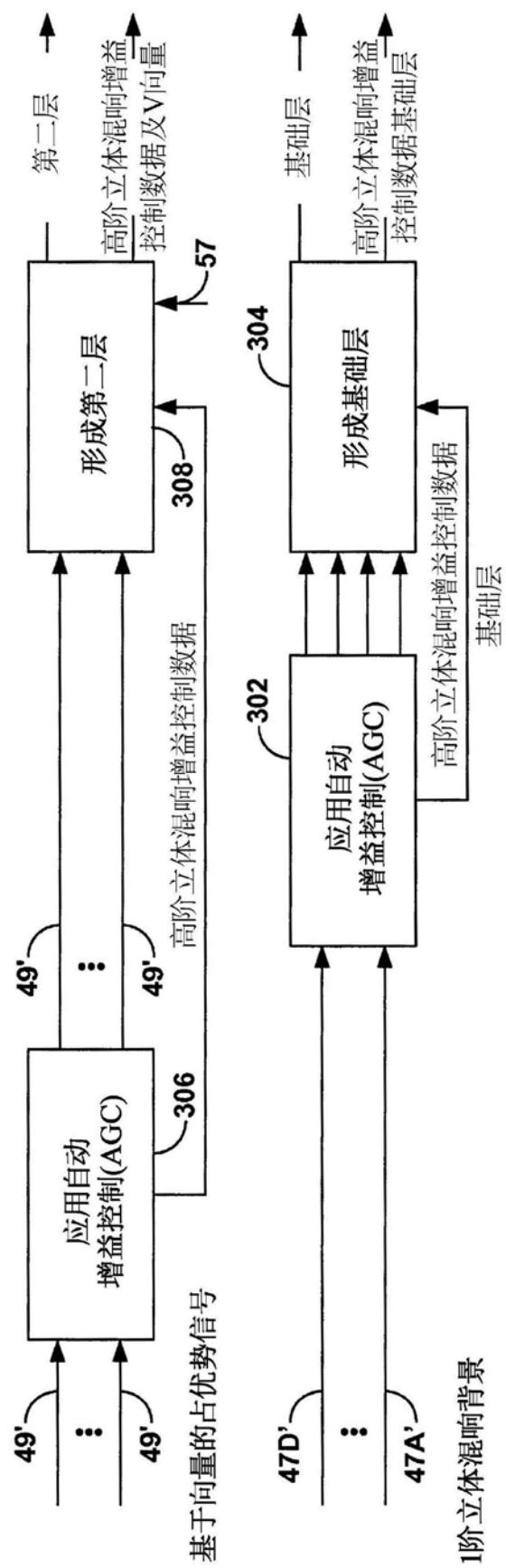


图7C

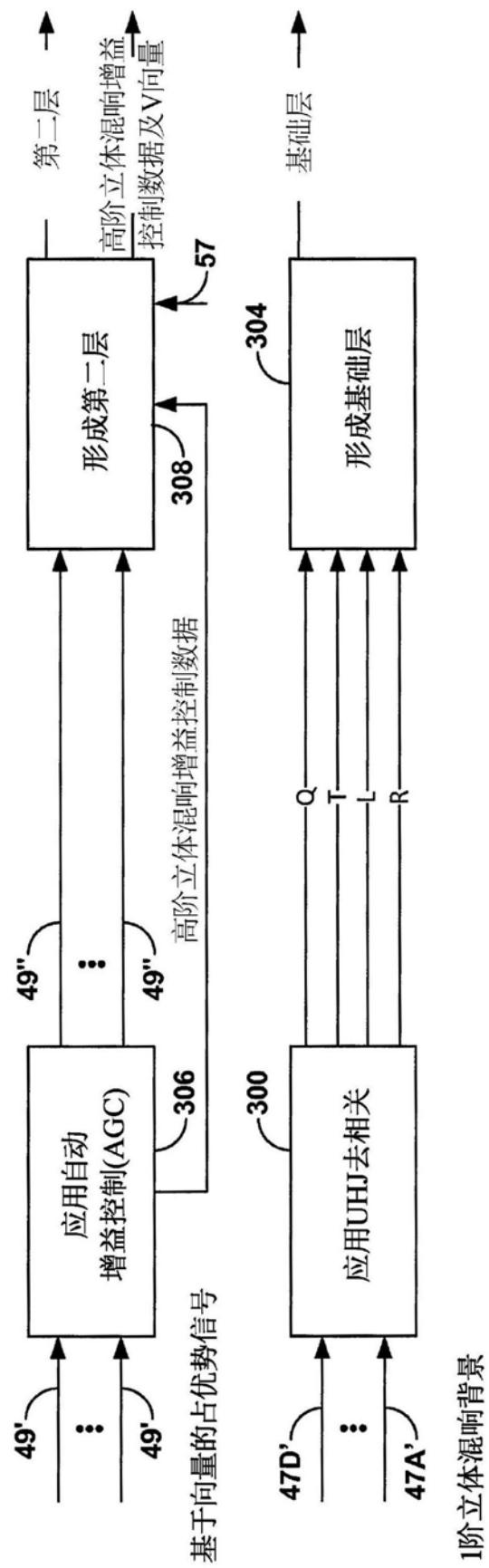


图 7D

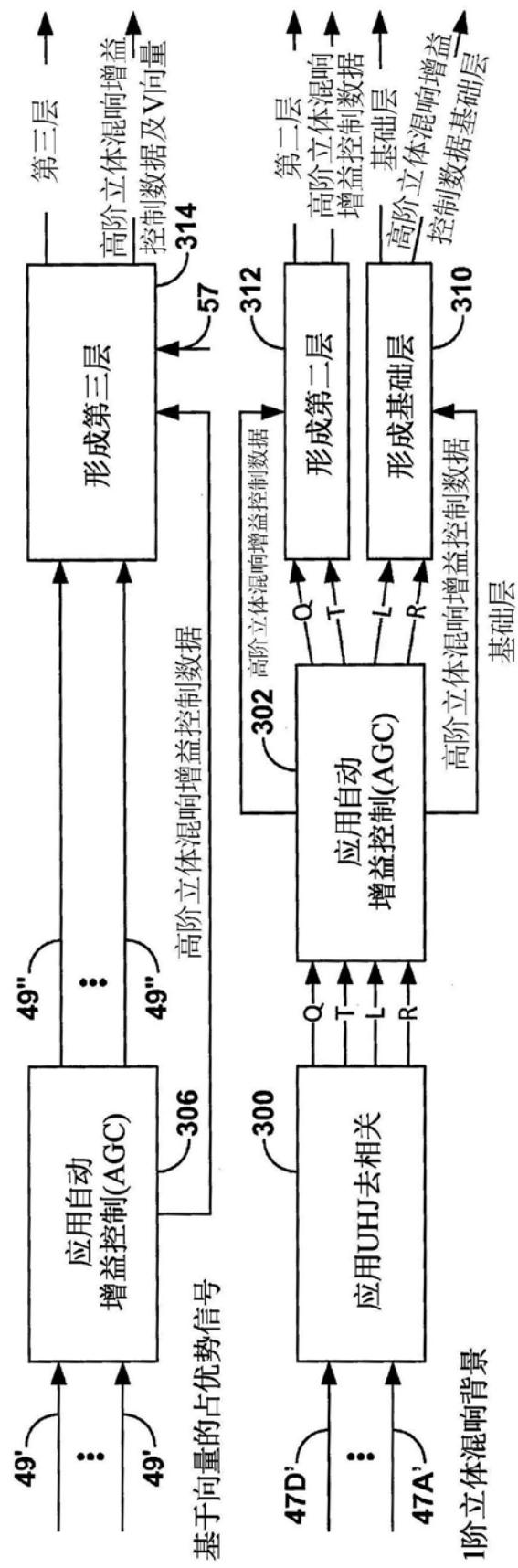


图8A

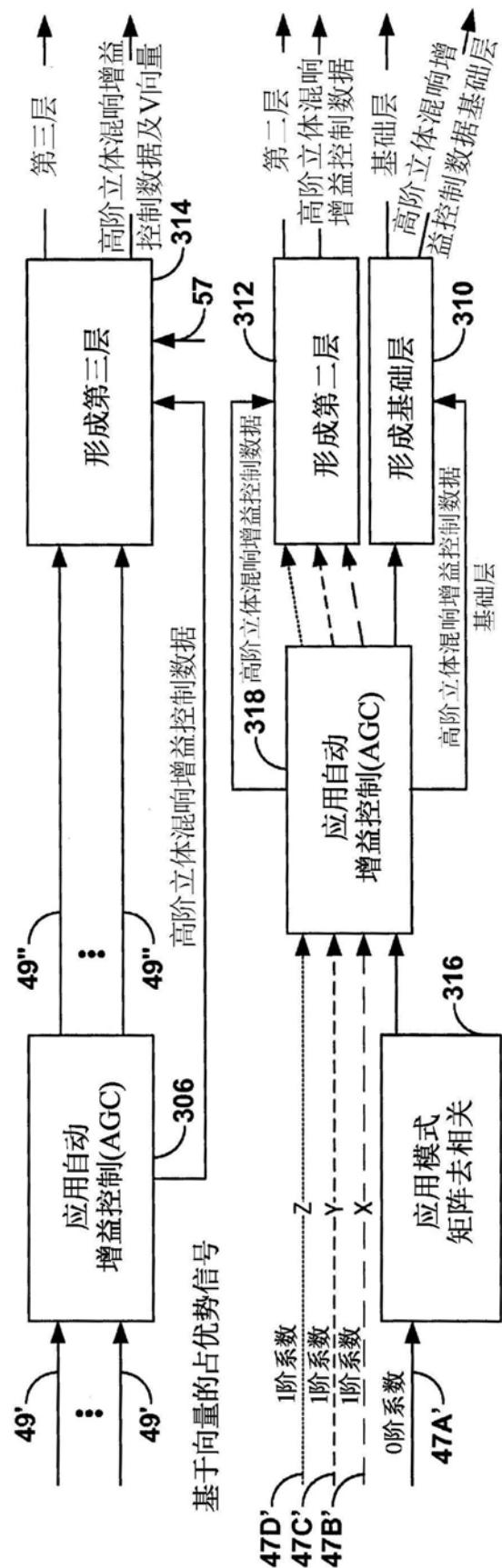


图 8B

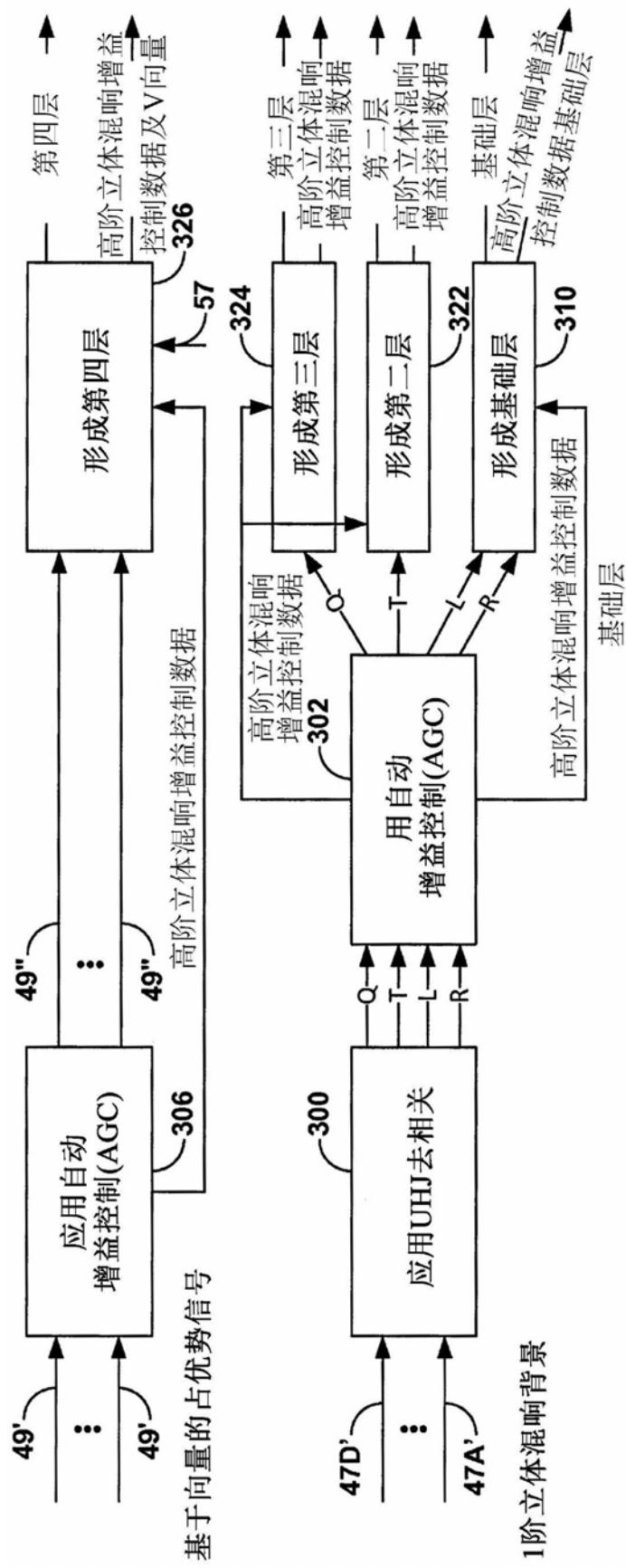
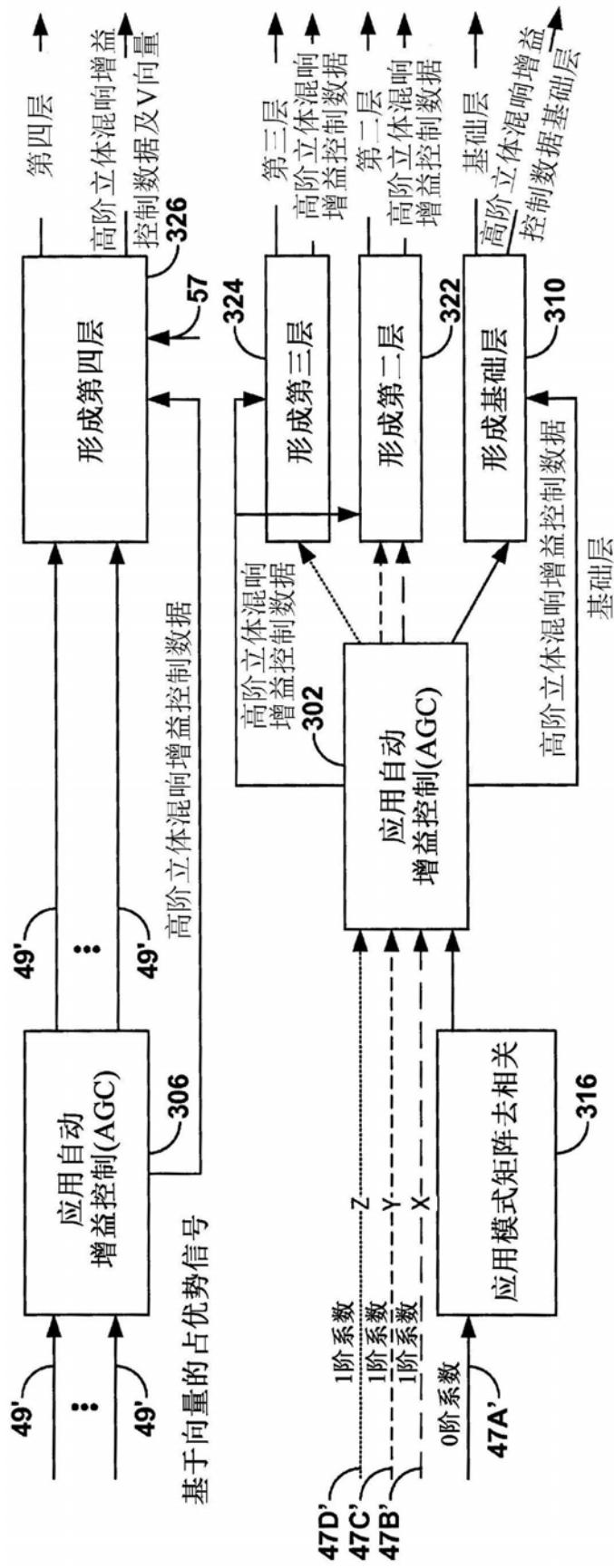


图9A



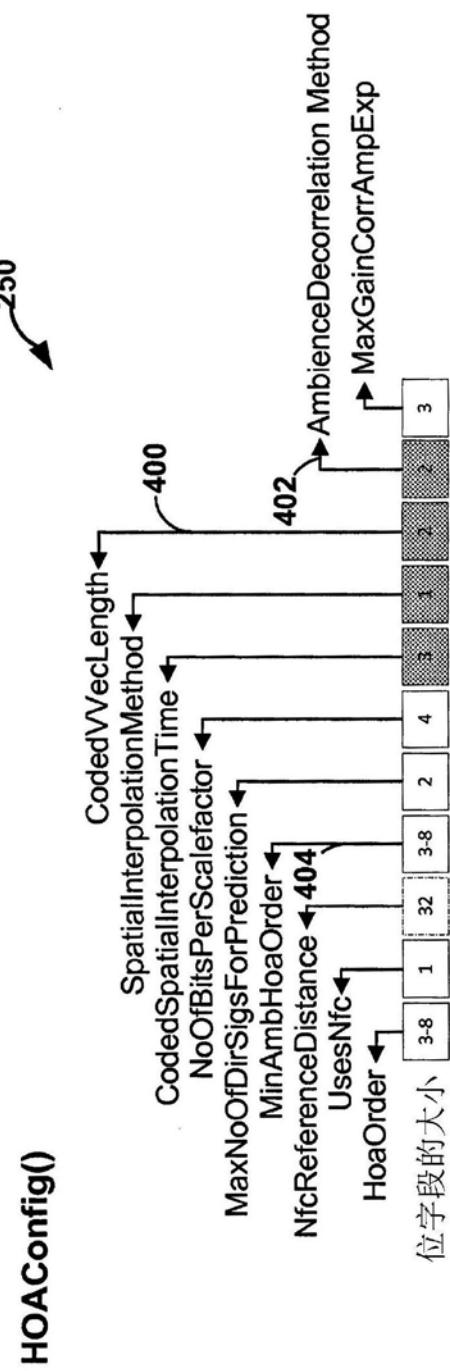


图10

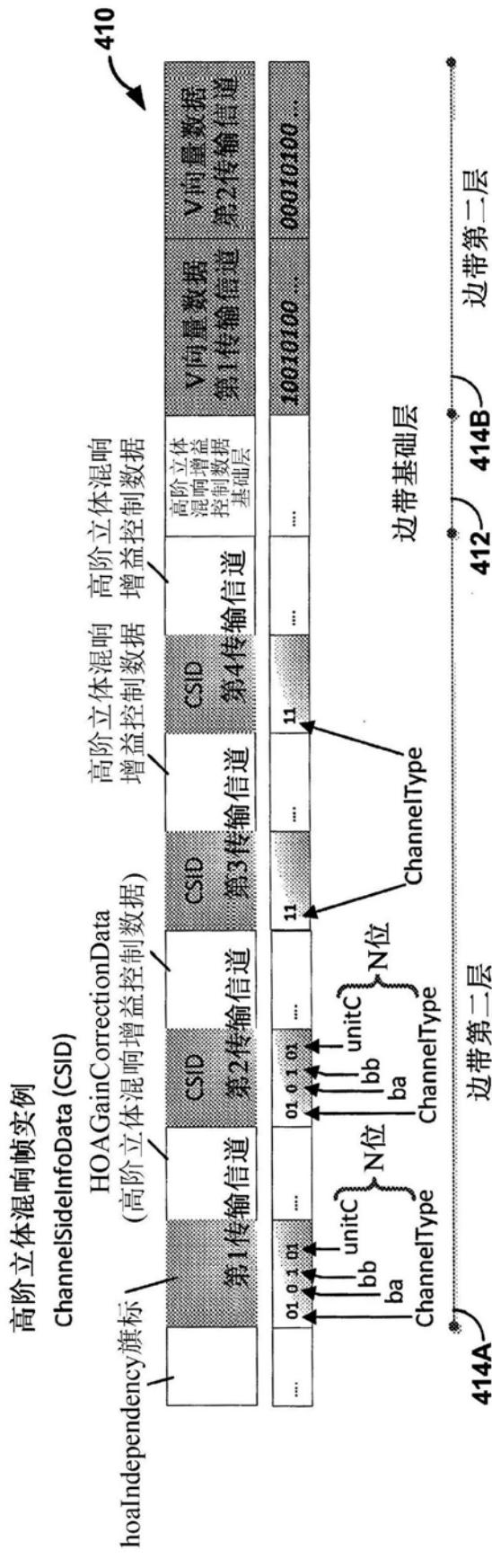


图11

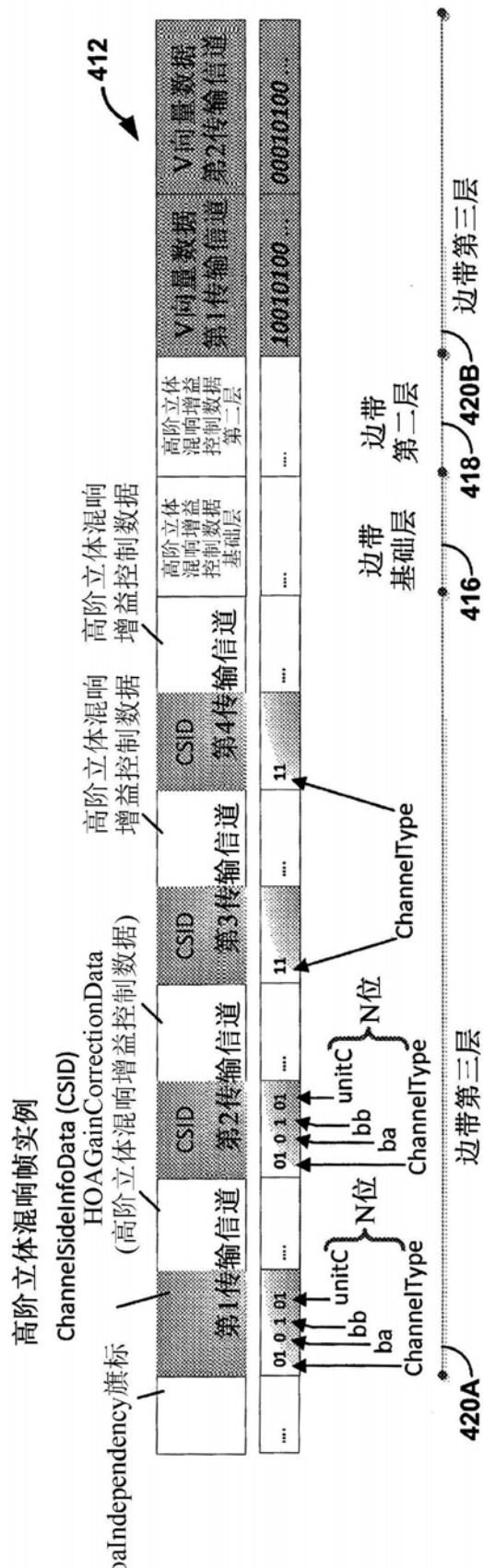


图12A

高阶立体混响帧实例
localDependency 旗标

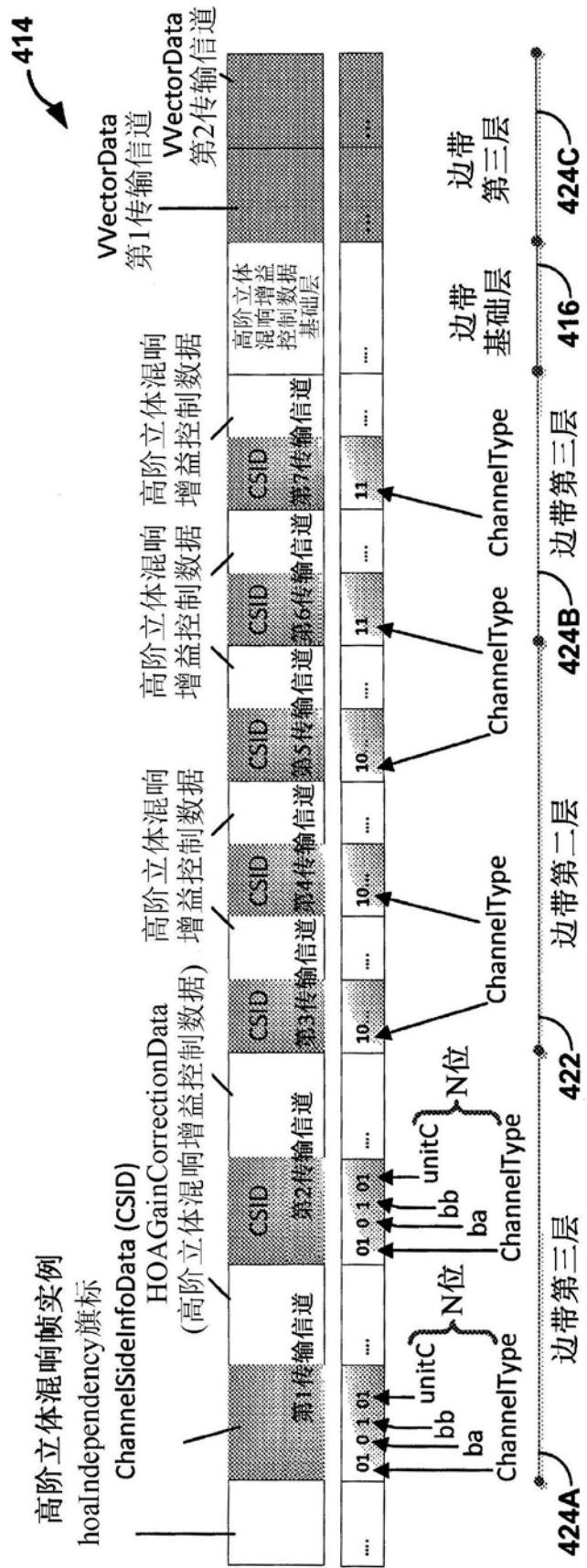


图12B

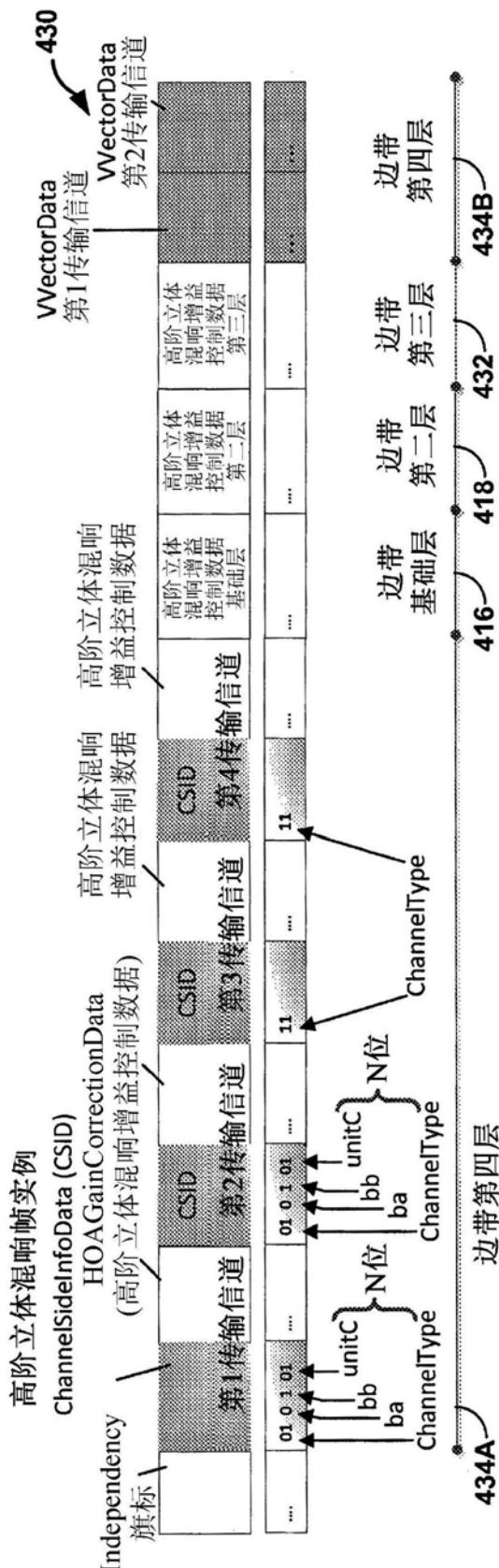


图 13A

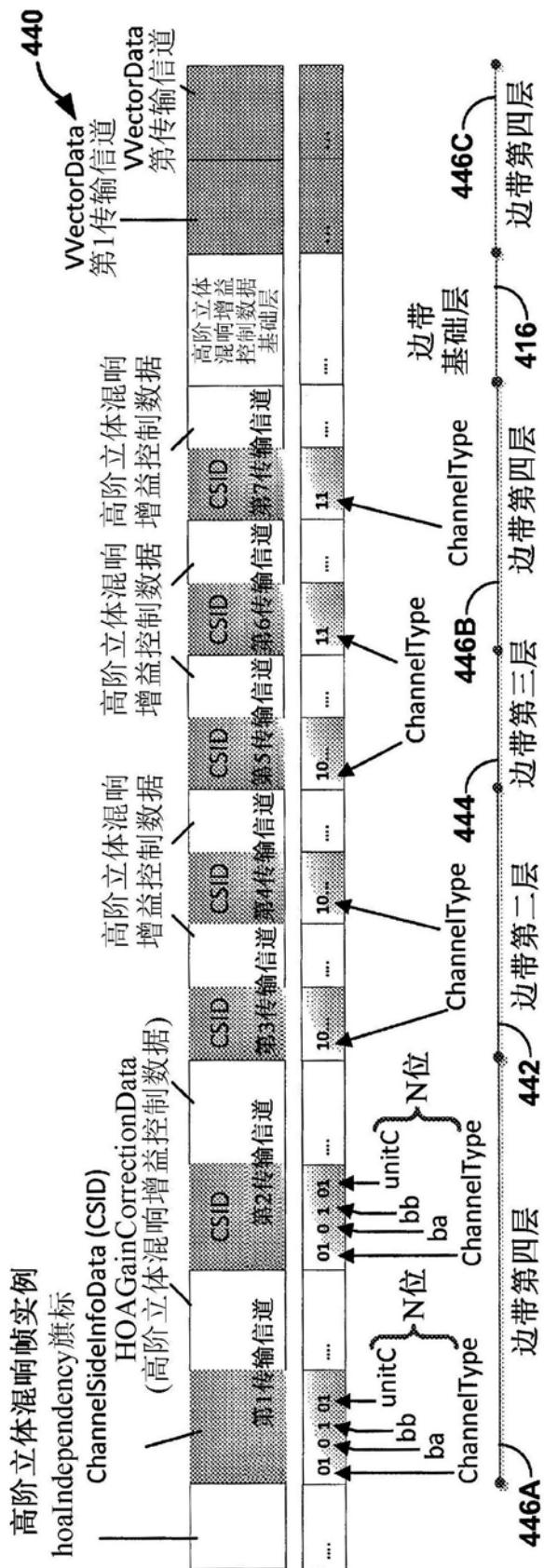


图13B

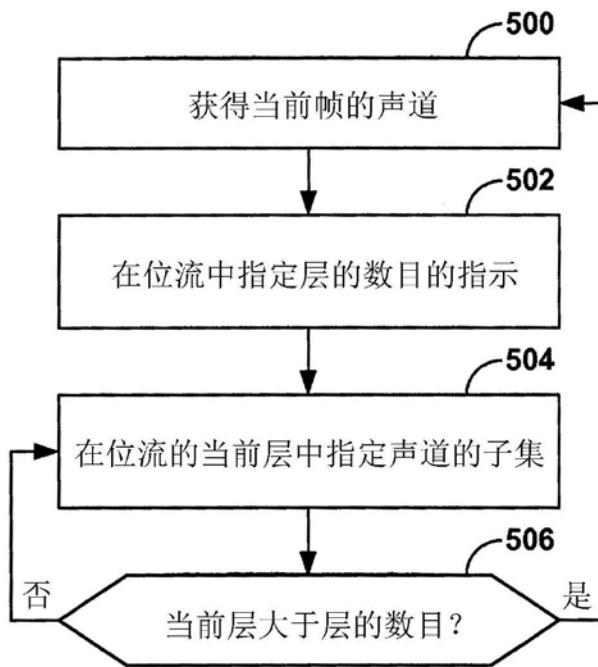


图14A

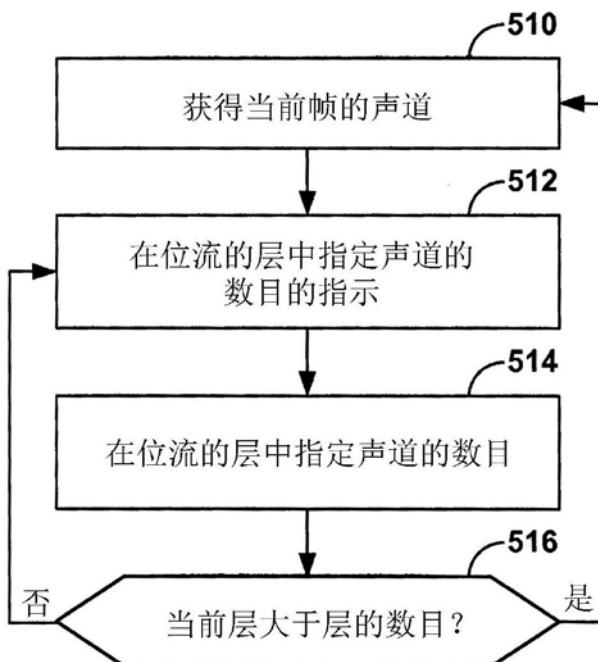


图14B

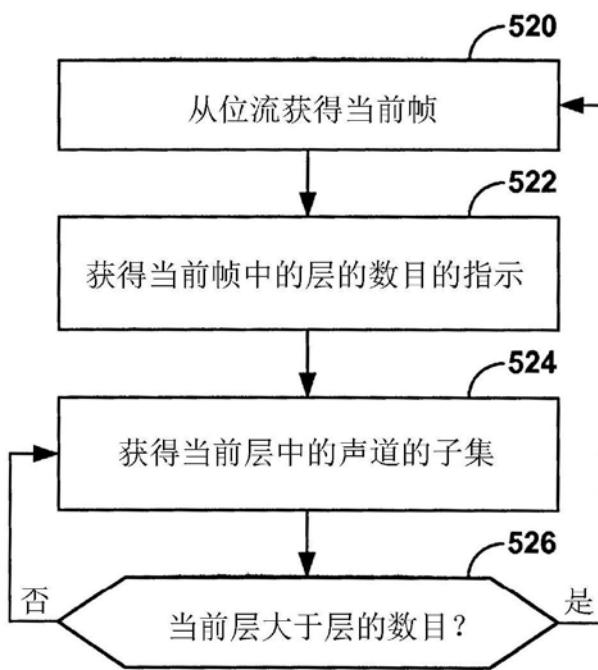


图15A

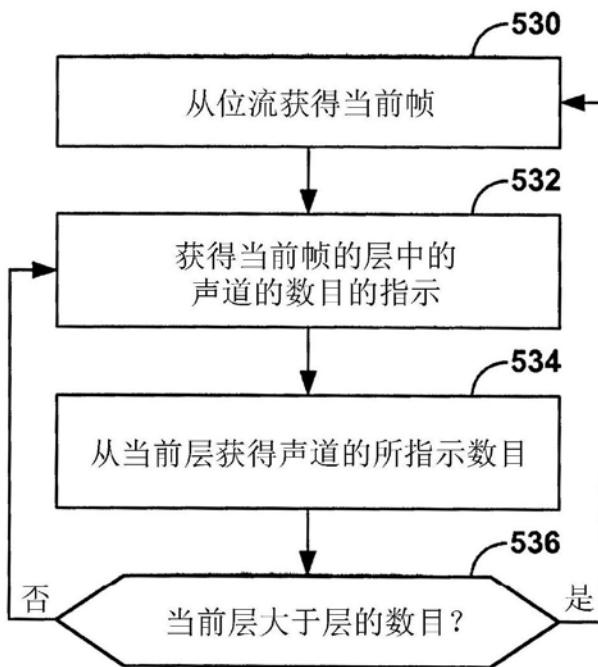


图15B

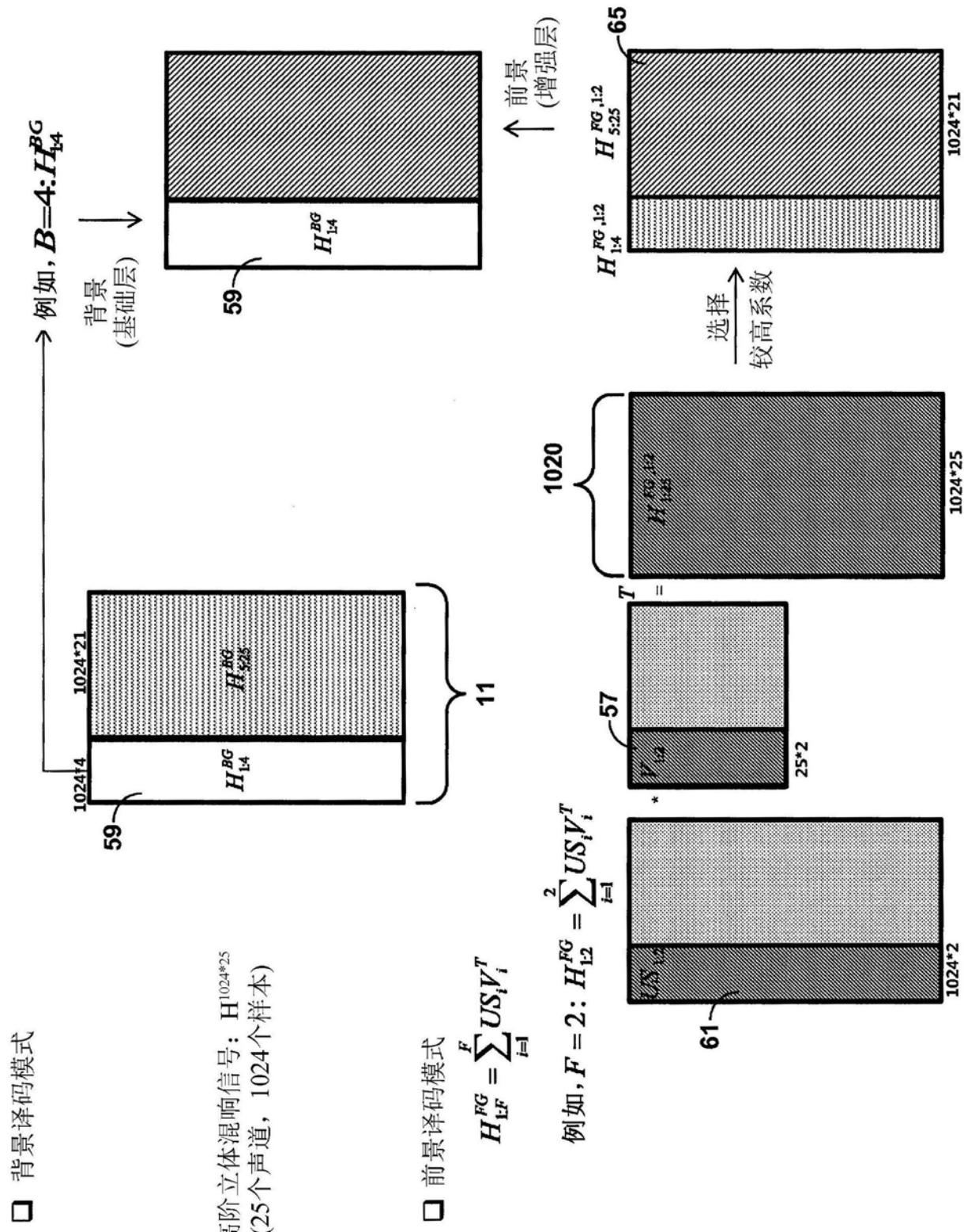


图16

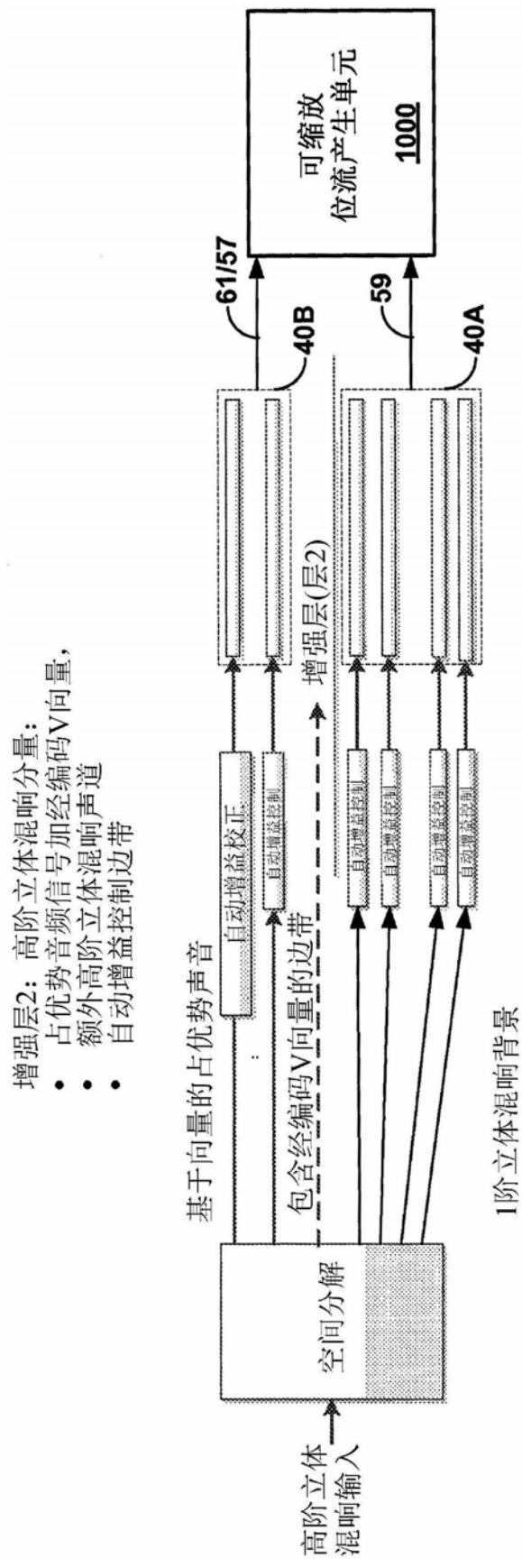


图17

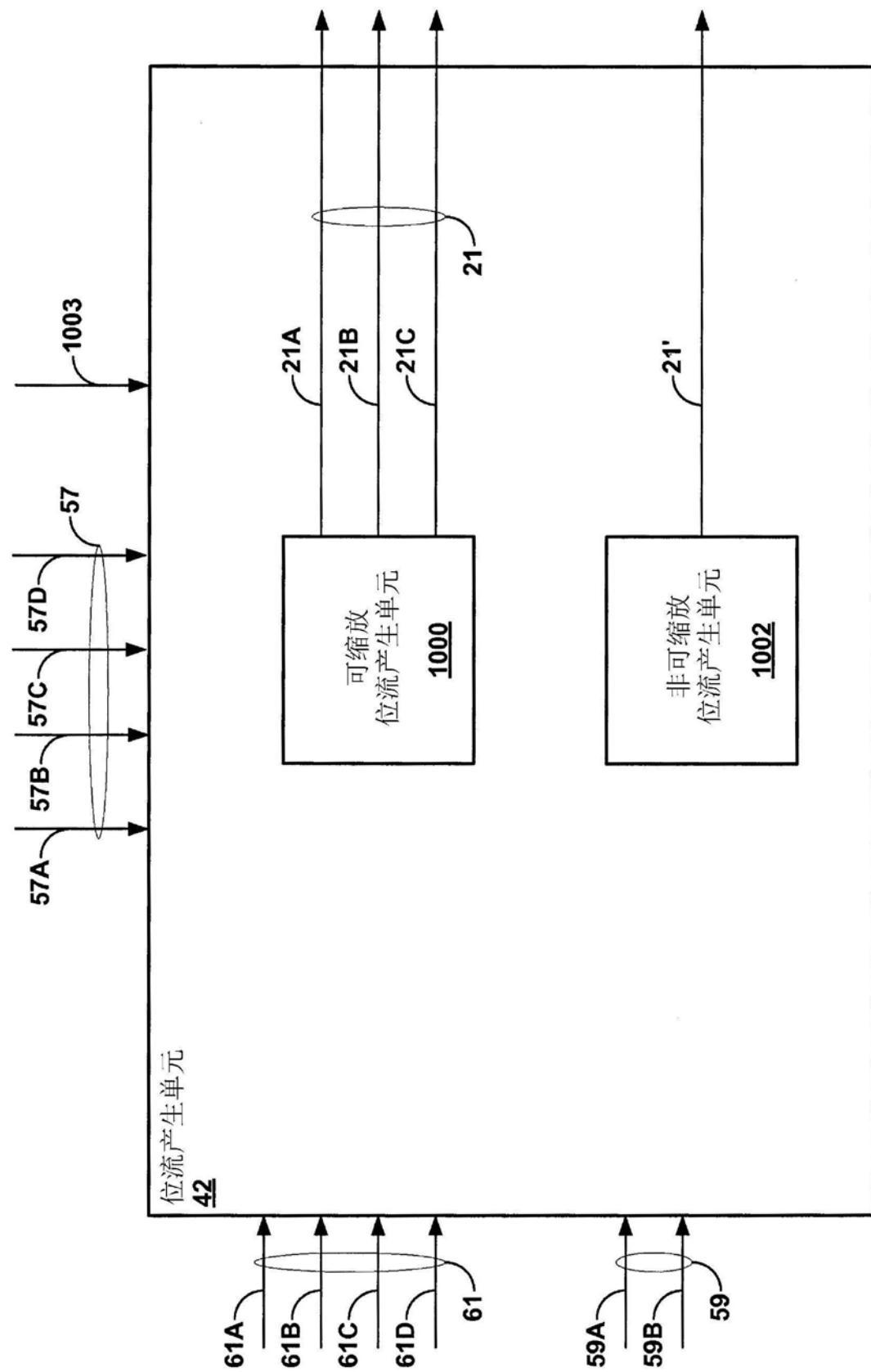


图18

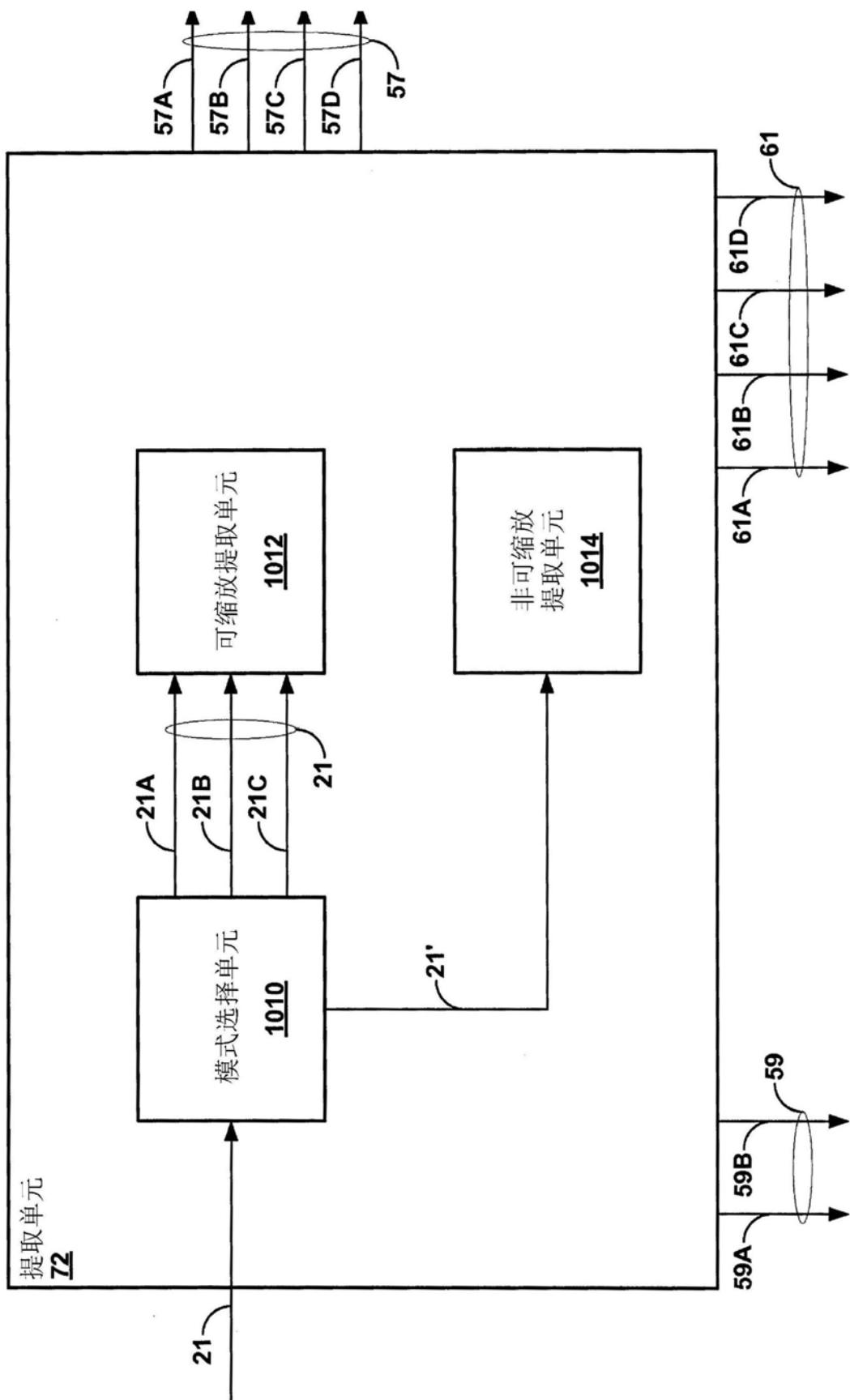
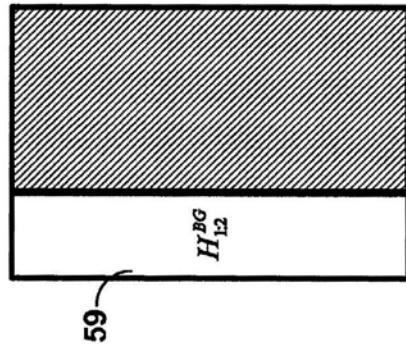


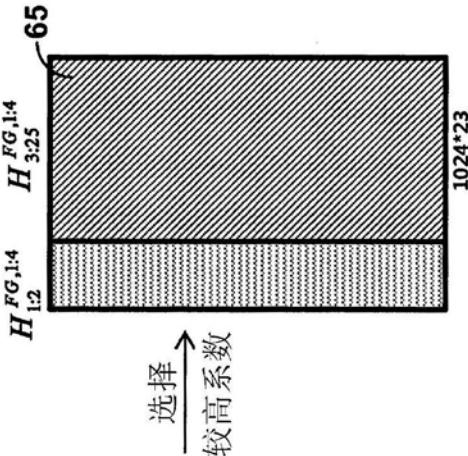
图19

例如, $B=2: H_{\text{L2}}^{\text{BG}}$

背景
(基础层) ↓



↑
前景
(增强层)



□ 用于分层译码的位流

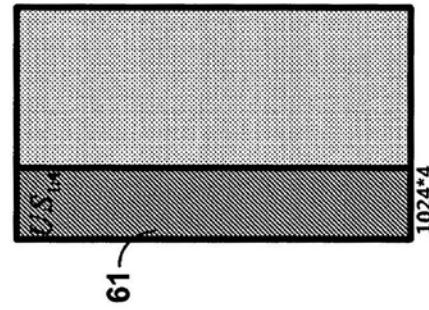
- NumberOfLayers $L = 2$
- $\{B_1=2, F_1=0, \{B_3=0, F_2=2\}, \{B_3=0, F_3=2\}\}$

如果前景与背景声道的总数目在解码器处已经知晓, 那么不必发射此信息。

□ 前景译码模式

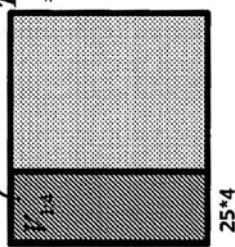
$$H_{\text{L,F}}^{\text{FG}} = \sum_{i=1}^F US_i V_i^T$$

例如, $F=4: H_{\text{L4}}^{\text{FG}} = \sum_{i=1}^4 US_i V_i^T$



1020

T



61

图20

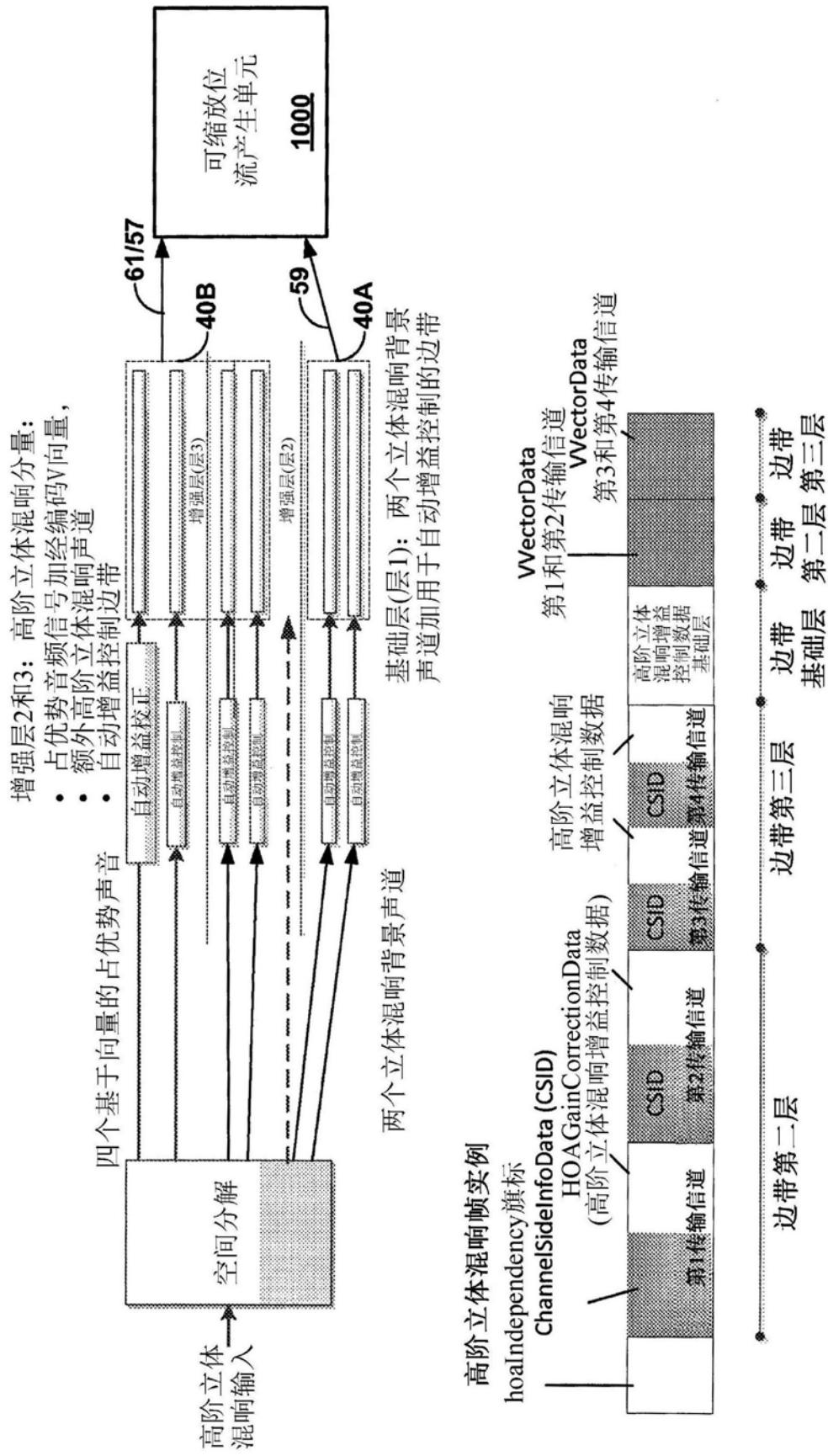


图 21

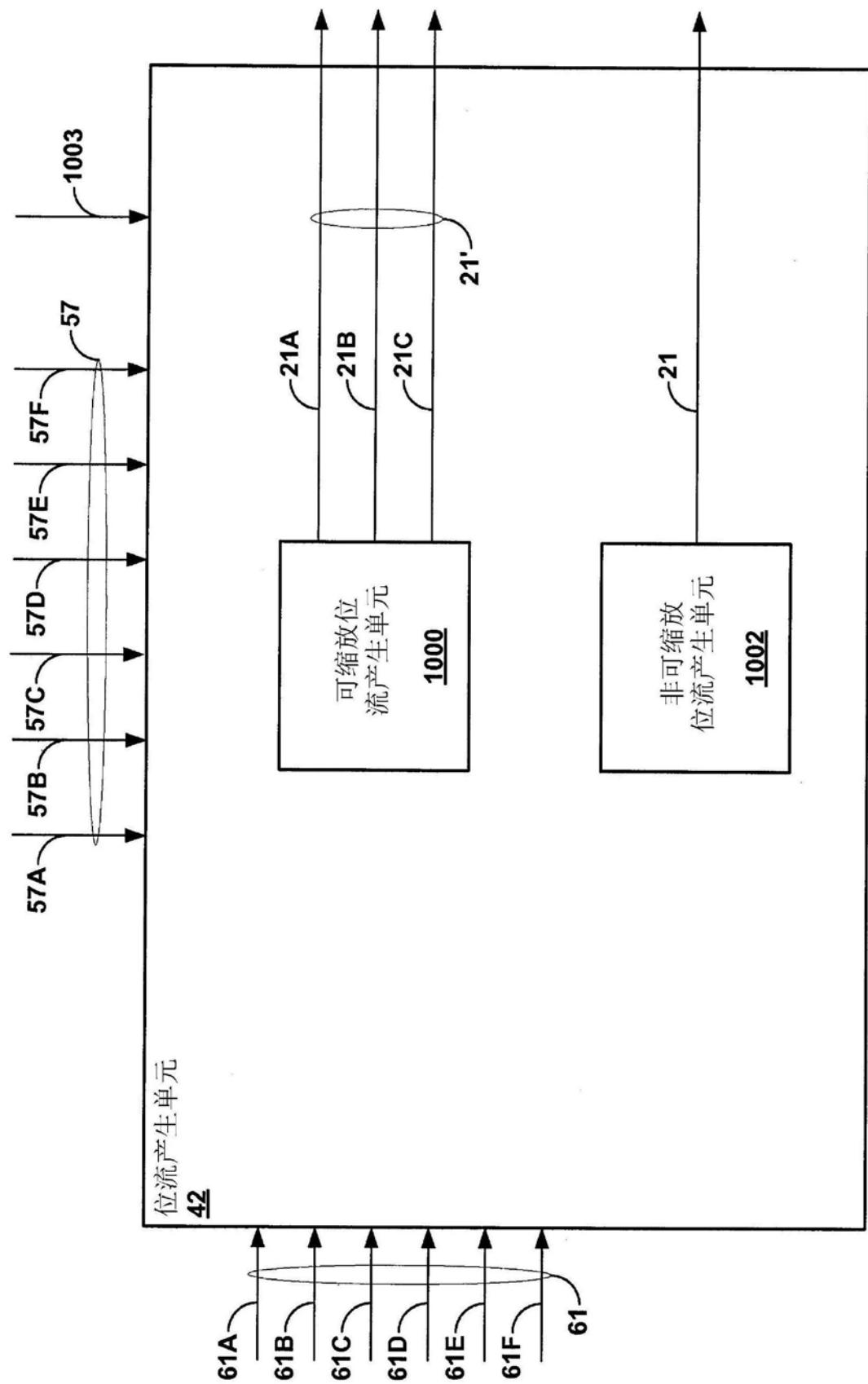


图22

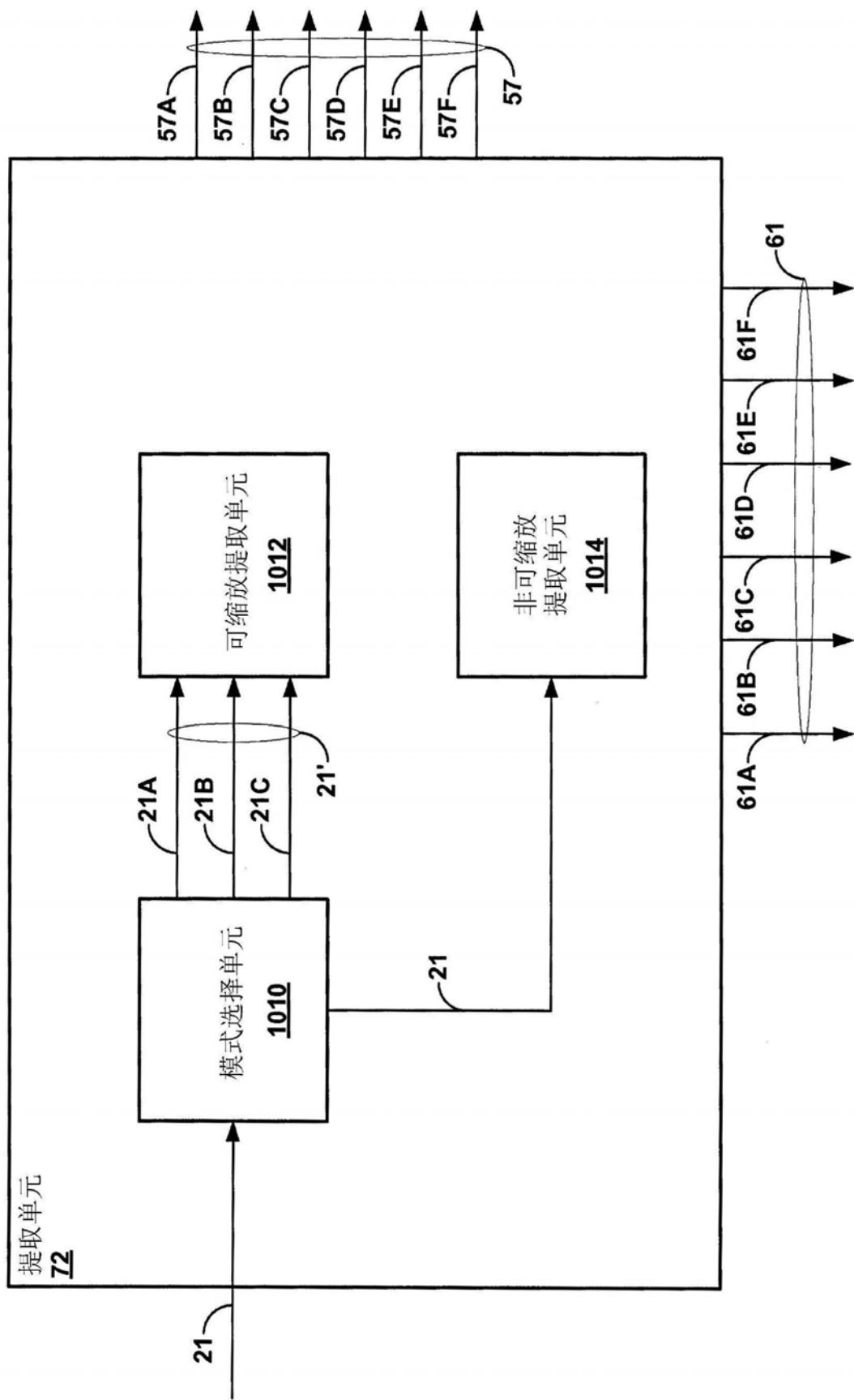


图23

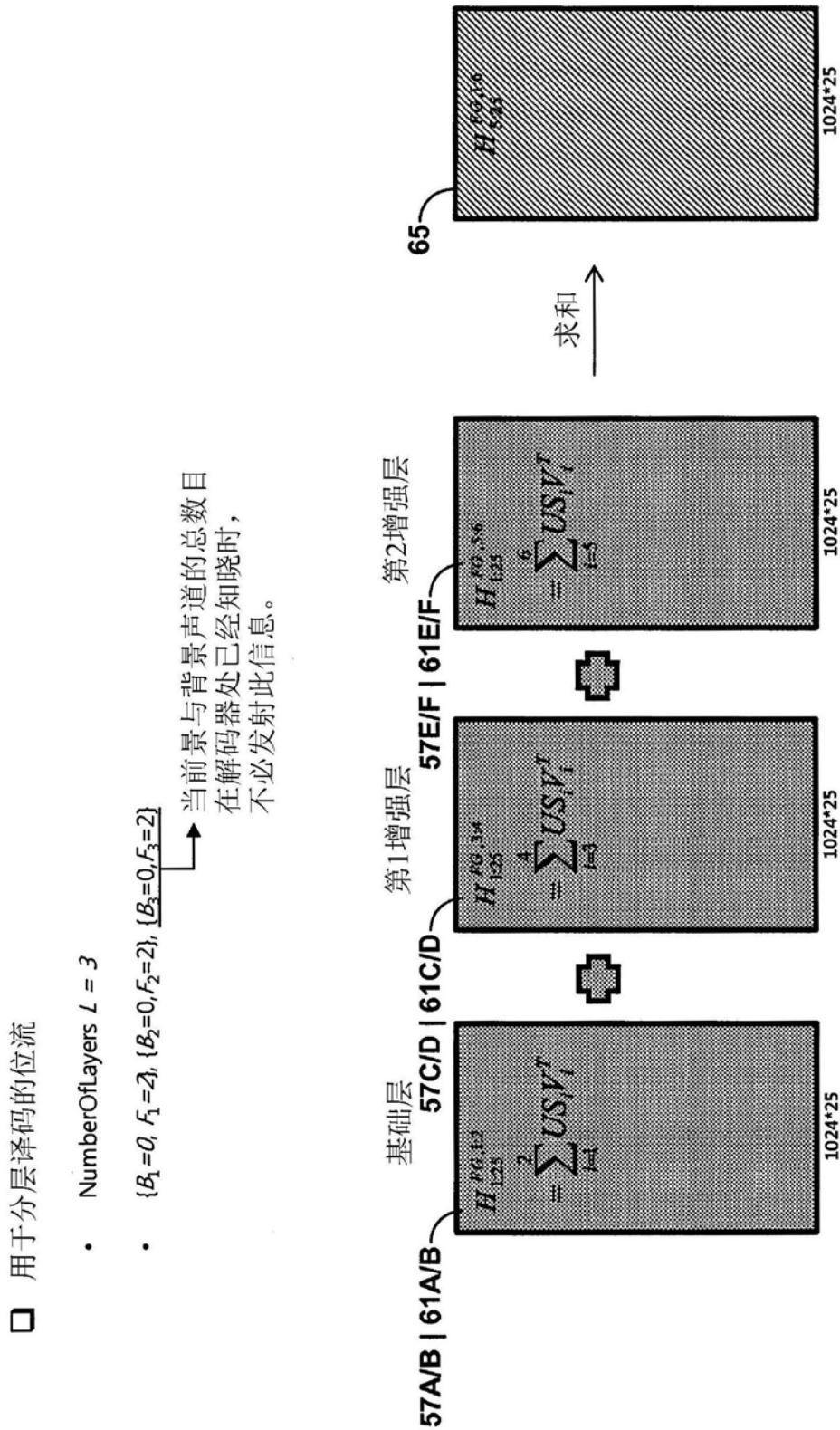
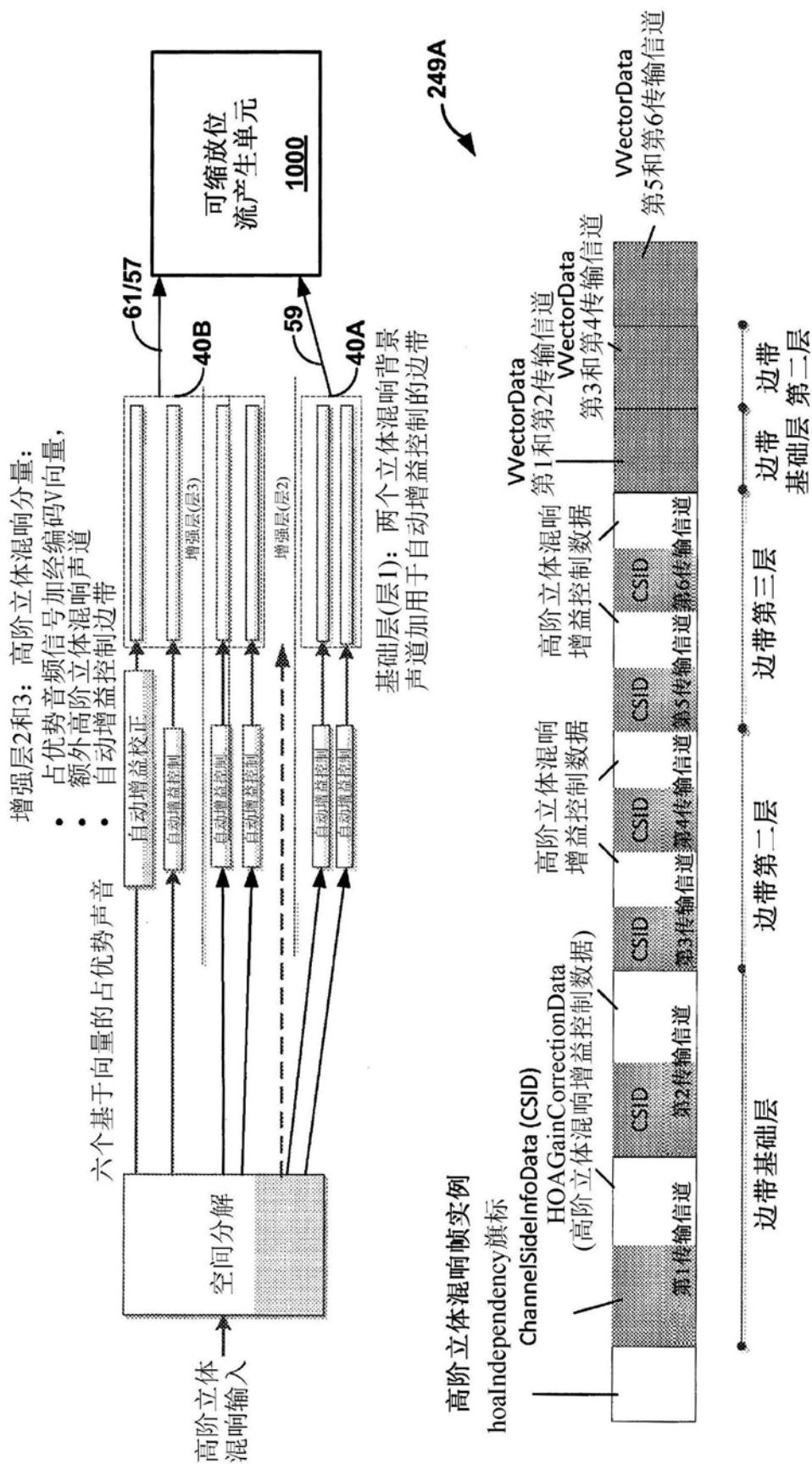


图24



- 用于分层译码的位流
 - NumberOfLayers $L = 4$
 - $\{B_1=1, F_1=0\}, \{B_2=1, F_2=0\}, \{B_3=1, F_3=0\}, \{B_4=1, F_4=0\}$

当前景与背景声道的总数目
在解码器处已经知晓时，
不必发射此信息。



图26

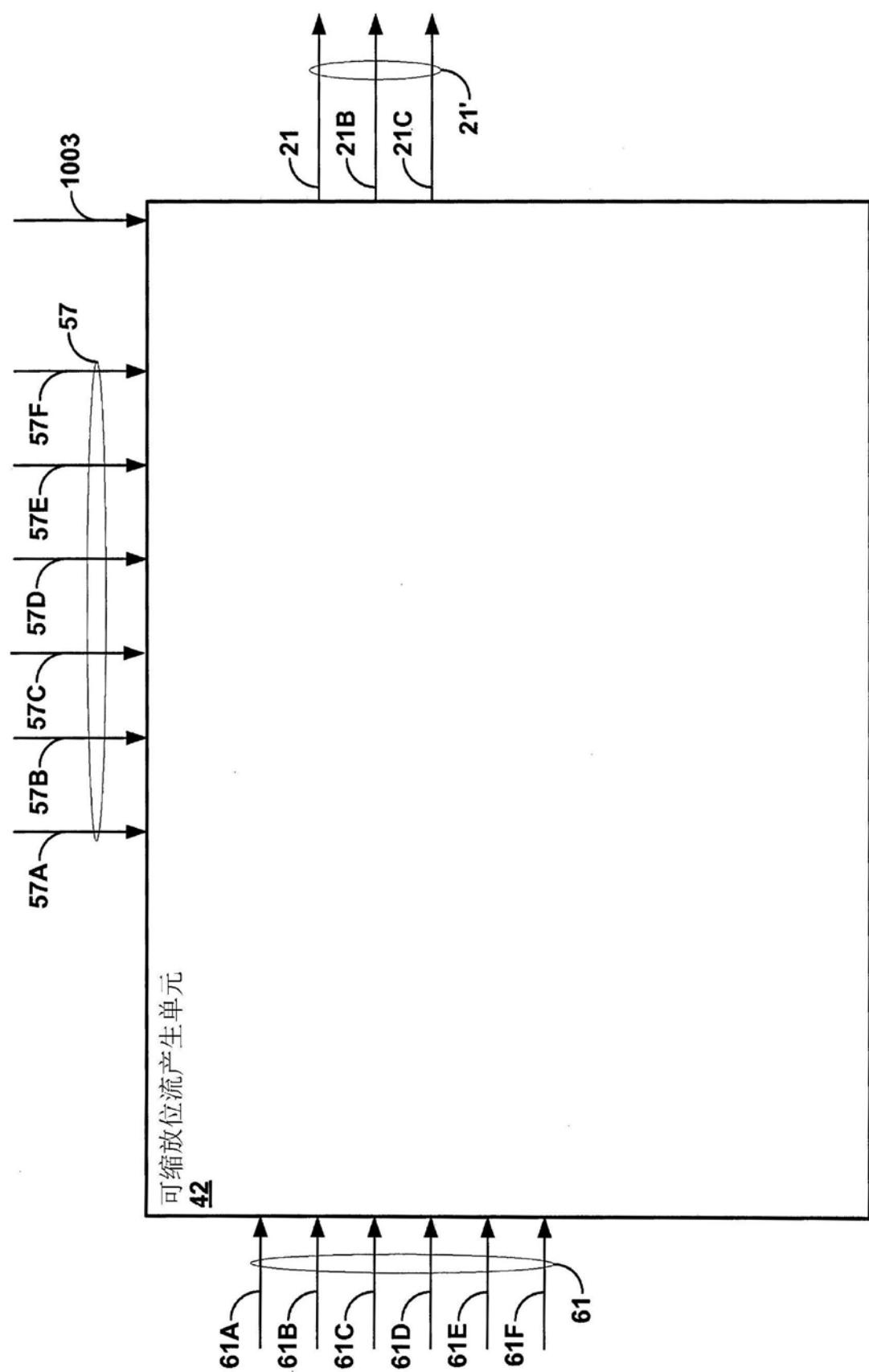


图27

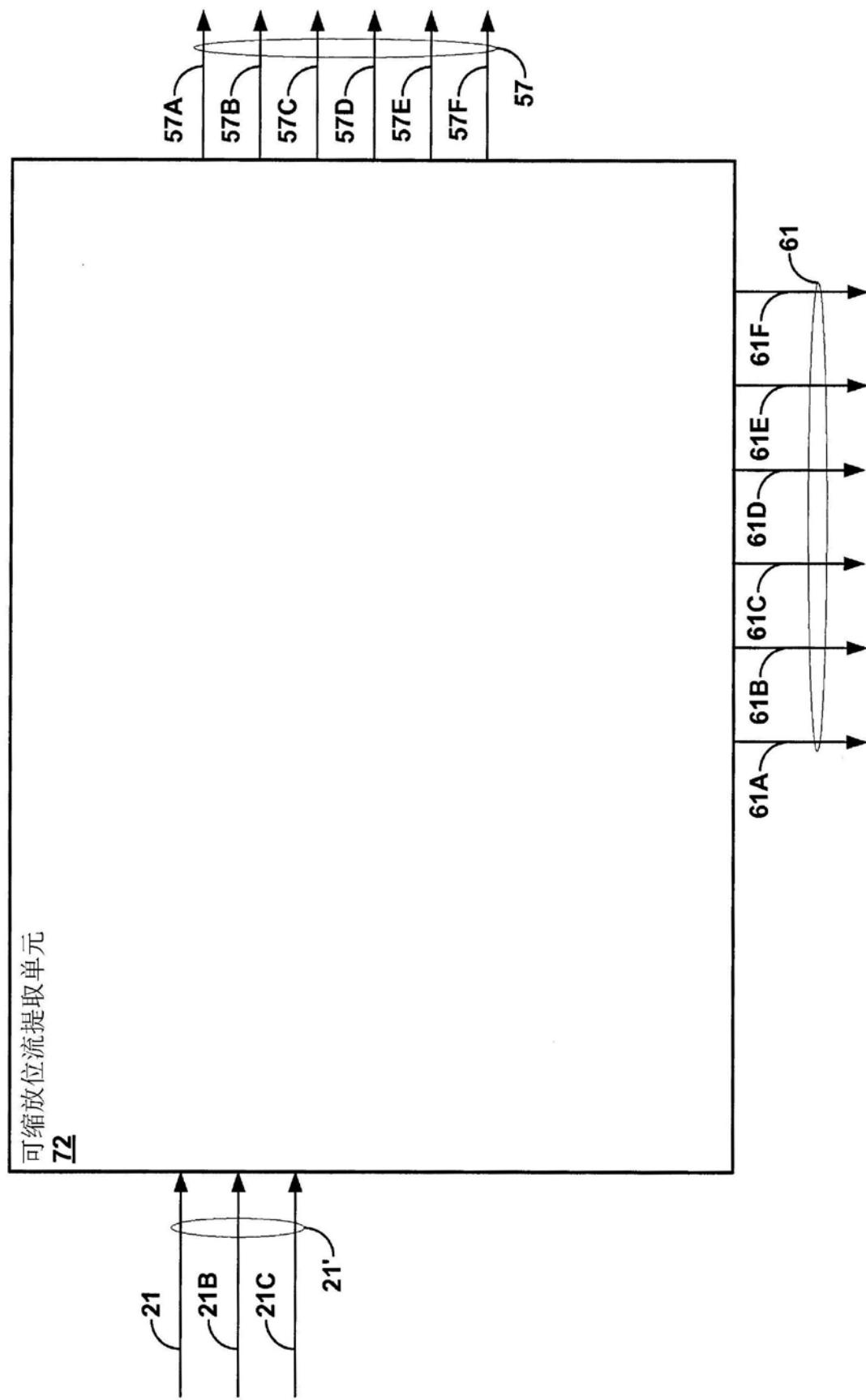


图28

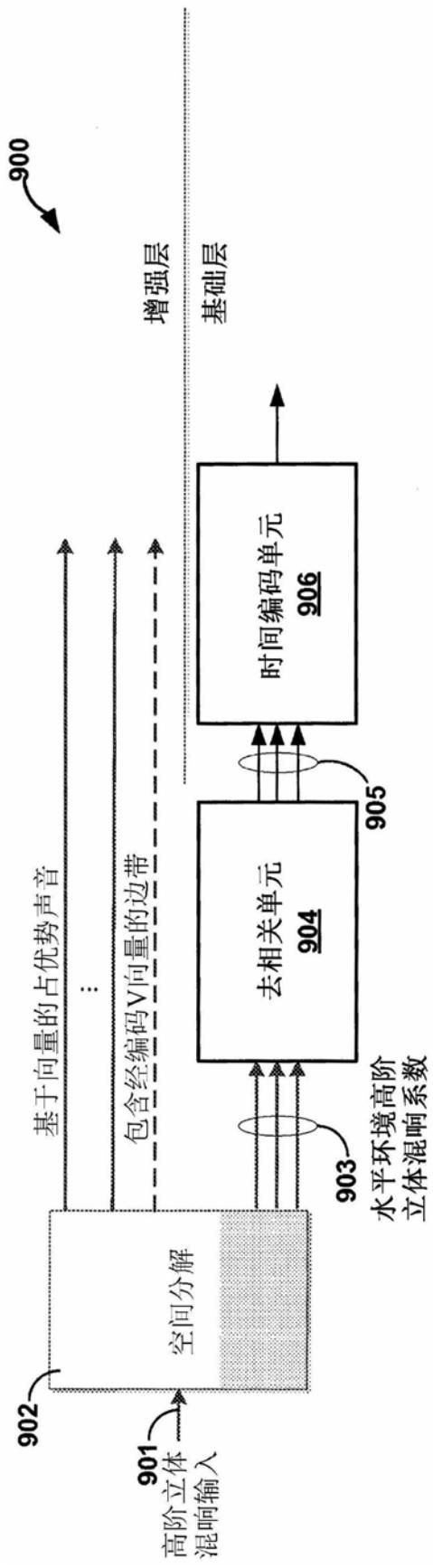


图29

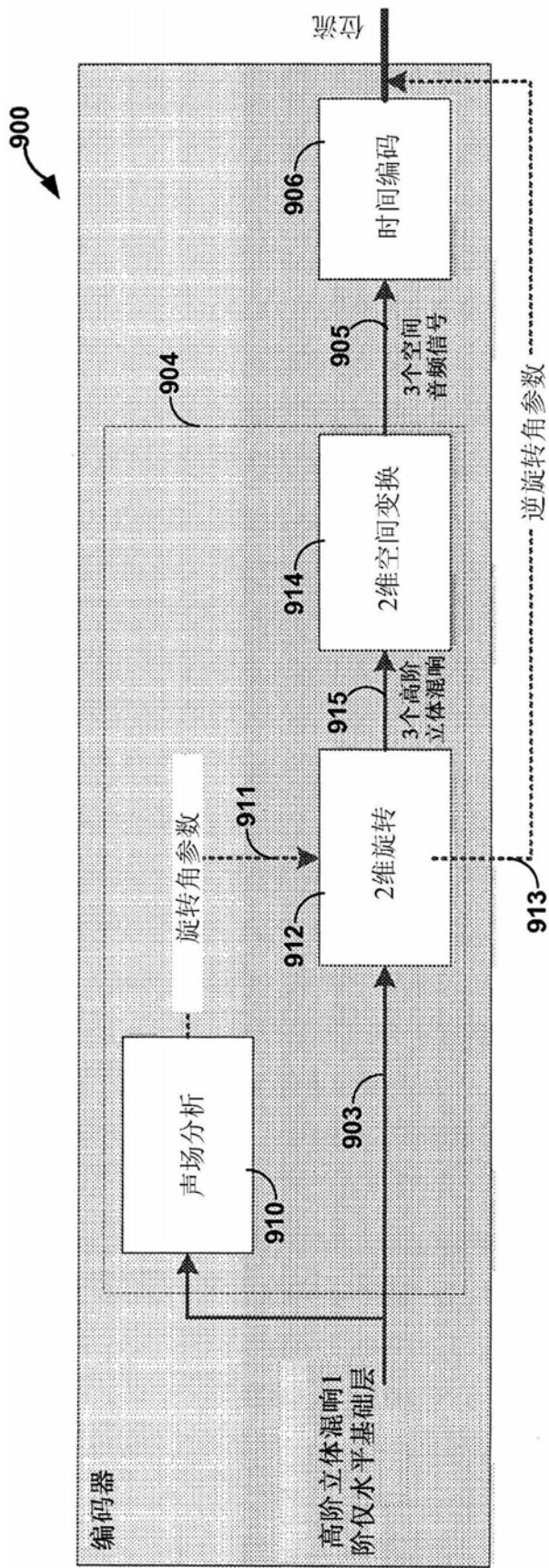


图30

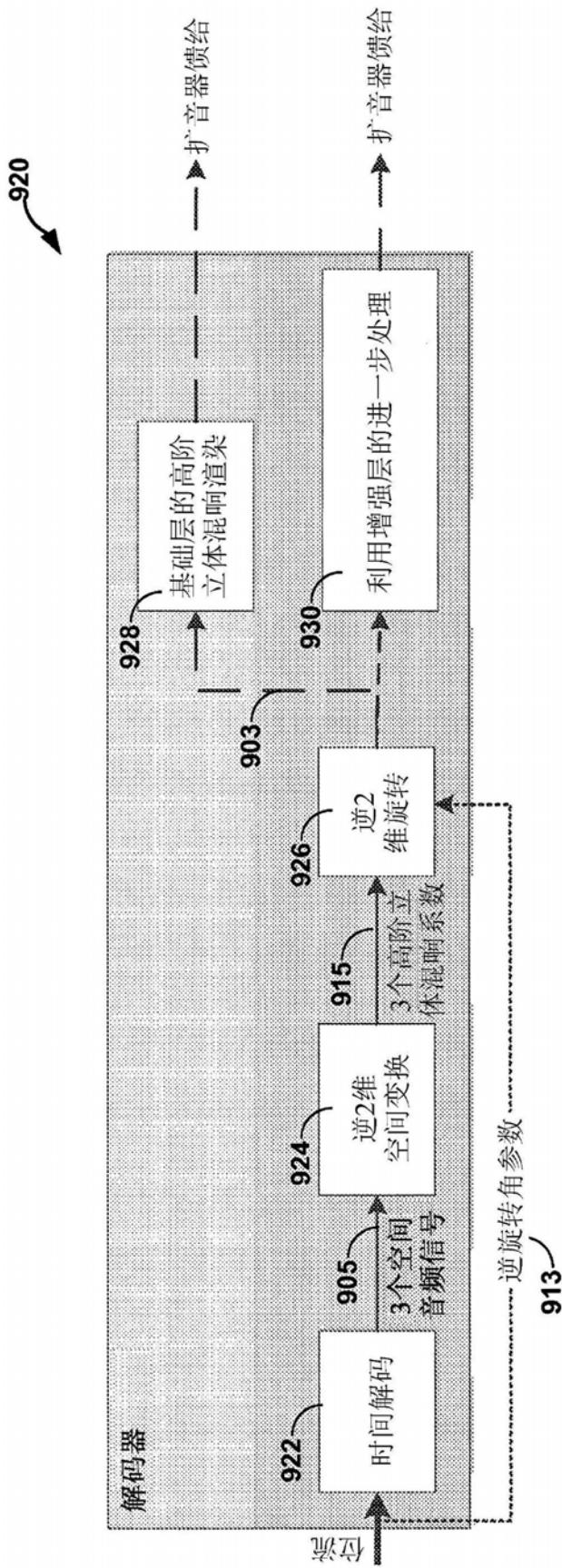


图31