



# (12)发明专利

(10)授权公告号 CN 105340298 B

(45)授权公告日 2017.05.31

(21)申请号 201480035597.1

(22)申请日 2014.05.28

(65)同一申请的已公布的文献号  
申请公布号 CN 105340298 A

(43)申请公布日 2016.02.17

(30)优先权数据

61/828,620 2013.05.29 US

61/847,543 2013.07.17 US

61/886,593 2013.10.03 US

61/886,620 2013.10.03 US

14/288,276 2014.05.27 US

(85)PCT国际申请进入国家阶段日  
2015.12.22

(86)PCT国际申请的申请数据  
PCT/US2014/039863 2014.05.28

(87)PCT国际申请的公布数据  
W02014/194004 EN 2014.12.04

(73)专利权人 高通股份有限公司  
地址 美国加利福尼亚州

(72)发明人 马丁·詹姆斯·莫雷尔

尼尔斯·京特·彼得斯

迪潘让·森

(74)专利代理机构 北京律盟知识产权代理有限公司  
11287

代理人 宋献涛

(51)Int.Cl.

H04S 5/00(2006.01)

H04S 3/00(2006.01)

H04S 1/00(2006.01)

(56)对比文件

CN 1643982 A,2005.07.20,

CN 102318372 A,2012.01.11,

CN 1735922 A,2006.02.15,

US 2009292544 A1,2009.11.26,

Dylan Menzies.NEARFIELD SYNTHESIS OF  
COMPLEX SOURCES WITH HIGH-ORDER  
AMBISONICS,AND BINAURAL RENDERING.  
《Proceedings of the 13th International  
Conference on Auditory Display》.2007,

审查员 李敏

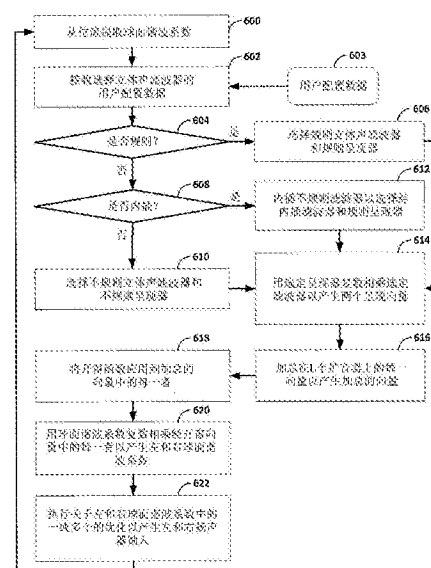
权利要求书5页 说明书27页 附图17页

(54)发明名称

球面谐波系数的立体声呈现

(57)摘要

一种装置包括一或多个处理器,所述一或多个处理器经配置以将立体声房间脉冲响应滤波器应用到以三维形式表示声场的球面谐波系数以便呈现所述声场。



1. 一种立体声音频呈现的方法,其包括:

将多个不规则立体声房间脉冲响应BRIR滤波器应用到较高阶环境立体混合声系数以便呈现声场,作为多个扬声器馈入,其中:

应用所述多个不规则BRIR滤波器包括将左和右立体声呈现矩阵与所述较高阶环境立体混合声系数卷积,所述左和右立体声呈现矩阵自将所述不规则BRIR滤波器转换到球面谐波域而产生,

所述较高阶环境立体混合声系数以三维形式表示所述声场,

所述多个不规则BRIR滤波器中的每一相应的不规则BRIR滤波器表示对在多个虚拟扩音器中的相应的虚拟扩音器的脉冲位置处产生的脉冲的响应,且

所述多个虚拟扩音器并不同等地间隔。

2. 根据权利要求1所述的方法,其中所述较高阶环境立体混合声系数为第一组较高阶环境立体混合声系数且所述声场为第一声场,所述多个虚拟扩音器为第一多个虚拟扩音器,所述方法进一步包括:

响应于接收到指定多个规则BRIR滤波器的使用的用户配置数据且在将所述多个不规则BRIR滤波器应用到所述第一组较高阶环境立体混合声系数后,将所述多个规则BRIR滤波器应用到第二组较高阶环境立体混合声系数以便呈现第二声场,其中:

所述多个规则BRIR滤波器中的每一相应的规则BRIR滤波器表示对在第二多个虚拟扩音器中的相应的虚拟扩音器的脉冲位置处产生的脉冲的响应,且

所述第二多个虚拟扩音器同等地间隔。

3. 根据权利要求1所述的方法,其中所述卷积产生左和右修改的较高阶环境立体混合声系数,所述多个扬声器馈入包含第一频域扬声器馈入和第二频域扬声器馈入,所述方法进一步包括:

加总在与所述较高阶环境立体混合声系数相关联的许多阶次和次阶范围内的第一修改的较高阶环境立体混合声系数以产生所述第一频域扬声器馈入,所述第一修改的较高阶环境立体混合声系数包括所述左修改的较高阶环境立体混合声系数或所述右修改的较高阶环境立体混合声系数;

倒转与负次阶相关联的所述第一修改的较高阶环境立体混合声系数的较高阶环境立体混合声系数以产生经倒转较高阶环境立体混合声系数;以及

加总在所述许多阶次和次阶范围内的所述经倒转较高阶环境立体混合声系数以产生所述第二频域扬声器馈入。

4. 根据权利要求1所述的方法,其中所述较高阶环境立体混合声系数对应于的球面基底函数的阶次大于一。

5. 根据权利要求1所述的方法,其进一步包括:

内插所述多个不规则BRIR滤波器以产生用于扬声器的规则布置的一或多个规则BRIR滤波器,且

其中应用所述多个不规则BRIR滤波器包括将所述多个规则BRIR滤波器应用到所述较高阶环境立体混合声系数以便呈现所述声场。

6. 根据权利要求1所述的方法,其进一步包括:

将开窗函数应用到所述多个不规则BRIR滤波器以产生经开窗BRIR滤波器,

其中应用所述多个不规则BRIR滤波器包括将所述经开窗BRIR滤波器应用到所述较高阶环境立体混合声系数以便呈现所述声场。

7. 根据权利要求1所述的方法,其进一步包括:

将所述多个不规则BRIR滤波器从时域变换到频域以便产生经变换不规则BRIR滤波器,其中应用所述多个不规则BRIR滤波器包括将所述经变换不规则BRIR滤波器应用到所述较高阶环境立体混合声系数以便呈现所述声场。

8. 根据权利要求1所述的方法,其进一步包括:

将所述多个不规则BRIR滤波器从时域变换到频域以便产生经变换不规则BRIR滤波器;以及

将所述较高阶环境立体混合声系数从所述时域变换到所述频域以便产生经变换较高阶环境立体混合声系数,

其中应用所述多个不规则BRIR滤波器包括将所述经变换不规则BRIR滤波器应用到所述经变换较高阶环境立体混合声系数以便呈现所述声场的频域表示,且

其中所述方法进一步包括将反变换应用到所述声场的所述频域表示以呈现所述声场。

9. 一种立体声音频呈现的装置,其包括:

一或多个处理器,其经配置以将多个不规则立体声房间脉冲响应BRIR滤波器应用到较高阶环境立体混合声系数以便呈现声场,作为多个扬声器馈入,其中:

所述一或多个处理器经配置使得作为应用所述多个不规则BRIR滤波器的部分,所述一或多个处理器将左和右立体声呈现矩阵与所述较高阶环境立体混合声系数卷积,所述左和右立体声呈现矩阵自将所述不规则BRIR滤波器转换到球面谐波域而产生,

所述较高阶环境立体混合声系数以三维形式表示所述声场,

所述多个不规则BRIR滤波器中的每一相应的不规则BRIR滤波器表示对在多个虚拟扩音器中的相应的虚拟扩音器的脉冲位置处产生的脉冲的响应,且

所述多个虚拟扩音器并不同等地间隔。

10. 根据权利要求9所述的装置,其中所述较高阶环境立体混合声系数为第一组较高阶环境立体混合声系数,所述声场为第一声场,所述多个虚拟扩音器为第一多个虚拟扩音器,且所述一或多个处理器经进一步配置以响应于接收到指定多个规则BRIR滤波器的使用的用户配置数据,将所述多个规则BRIR滤波器应用到第二组较高阶环境立体混合声系数以便呈现第二声场,其中:

所述多个规则BRIR滤波器中的每一相应的规则BRIR滤波器表示对在第二多个虚拟扩音器中的相应的虚拟扩音器的脉冲位置处产生的脉冲的响应,且

所述第二多个虚拟扩音器同等地间隔。

11. 根据权利要求9所述的装置,其中所述卷积产生左和右修改的较高阶环境立体混合声系数,所述多个扬声器馈入包含第一频域扬声器馈入和第二频域扬声器馈入,且

所述一或多个处理器经配置以:

加总在与所述较高阶环境立体混合声系数相关联的许多阶次和次阶范围内的第一修改的较高阶环境立体混合声系数以产生所述第一频域扬声器馈入,所述第一修改的较高阶环境立体混合声系数包括所述左修改的较高阶环境立体混合声系数或所述右修改的较高阶环境立体混合声系数;

倒转与负次阶相关联的所述第一修改的较高阶环境立体混合声系数的较高阶环境立体混合声系数以产生经倒转较高阶环境立体混合声系数;以及

加总在所述许多阶次和次阶范围内的所述经倒转较高阶环境立体混合声系数以产生所述第二频域扬声器馈入。

12. 根据权利要求9所述的装置,其中所述较高阶环境立体混合声系数对应于的球面基底函数的阶次大于一。

13. 根据权利要求9所述的装置,

其中所述一或多个处理器经进一步配置以内插所述多个不规则BRIR滤波器以产生多个规则BRIR滤波器,其中所述多个规则BRIR滤波器包括用于扬声器的规则布置的多个BRIR滤波器,且

其中所述一或多个处理器经进一步配置以为了应用所述多个不规则BRIR滤波器,将所述多个规则BRIR滤波器应用到所述较高阶环境立体混合声系数以便呈现所述声场。

14. 根据权利要求9所述的装置,

其中所述一或多个处理器经进一步配置以将开窗函数应用到所述多个不规则BRIR滤波器以产生经开窗BRIR滤波器,且

其中所述一或多个处理器经进一步配置以当应用所述多个不规则BRIR滤波器时,将所述经开窗BRIR滤波器应用到所述较高阶环境立体混合声系数以便呈现所述声场。

15. 根据权利要求9所述的装置,

其中所述一或多个处理器经进一步配置以将所述多个不规则BRIR滤波器从时域变换到频域以便产生经变换不规则BRIR滤波器,且

其中所述一或多个处理器经进一步配置以当应用所述多个不规则BRIR滤波器时,将所述经变换不规则BRIR滤波器应用到所述较高阶环境立体混合声系数以便呈现所述声场。

16. 根据权利要求9所述的装置,

其中所述一或多个处理器经进一步配置以将所述多个不规则BRIR滤波器从时域变换到频域以便产生经变换不规则BRIR滤波器,和将所述较高阶环境立体混合声系数从所述时域变换到所述频域以便产生经变换较高阶环境立体混合声系数,

其中所述一或多个处理器经进一步配置以当应用所述多个不规则BRIR滤波器时,将所述经变换不规则BRIR滤波器应用到所述经变换较高阶环境立体混合声系数以便呈现所述声场的频域表示,且

其中所述一或多个处理器经进一步配置以将反变换应用到所述声场的所述频域表示以呈现所述声场。

17. 一种立体声音频呈现的设备,其包括:

用于确定以三维形式表示声场的较高阶环境立体混合声系数的装置;以及

用于将多个不规则立体声房间脉冲响应BRIR滤波器应用到所述较高阶环境立体混合声系数以便呈现所述声场作为多个扬声器馈入的装置,其中:

所述用于应用所述多个不规则BRIR滤波器的装置包括用于将左和右立体声呈现矩阵与所述较高阶环境立体混合声系数卷积的装置,所述左和右立体声呈现矩阵自将所述不规则BRIR滤波器转换到球面谐波域而产生,

所述多个不规则BRIR滤波器中的每一相应的不规则BRIR滤波器表示对在多个虚拟扩

音器中的相应的虚拟扩音器的脉冲位置处产生的脉冲的响应,且

所述多个虚拟扩音器并不同等地间隔。

18. 根据权利要求17所述的设备,其中所述较高阶环境立体混合声系数为第一组较高阶环境立体混合声系数且所述声场为第一声场,所述多个虚拟扩音器为第一多个虚拟扩音器,所述设备进一步包括:

用于接收指定多个规则BRIR滤波器的使用的用户配置数据的装置;以及

用于将所述多个规则BRIR滤波器应用到第二组较高阶环境立体混合声系数以便呈现第二声场的装置,其中:

所述多个规则BRIR滤波器中的每一相应的规则BRIR滤波器表示对在第二多个虚拟扩音器中的相应的虚拟扩音器的脉冲位置处产生的脉冲的响应,且

所述第二多个虚拟扩音器同等地间隔。

19. 根据权利要求17所述的设备,其中所述卷积产生左和右修改的较高阶环境立体混合声系数,所述多个扬声器馈入包含第一频域扬声器馈入和第二频域扬声器馈入,所述设备进一步包括:

用于加总在与所述较高阶环境立体混合声系数相关联的许多阶次和次阶范围内的第一修改的较高阶环境立体混合声系数以产生所述第一频域扬声器馈入的装置,所述第一修改的较高阶环境立体混合声系数包括所述左修改的较高阶环境立体混合声系数或所述右修改的较高阶环境立体混合声系数;

用于倒转与负次阶相关联的所述第一修改的较高阶环境立体混合声系数的较高阶环境立体混合声系数以产生经倒转较高阶环境立体混合声系数的装置;以及

用于加总在所述许多阶次和次阶范围内的所述经倒转较高阶环境立体混合声系数以产生所述第二频域扬声器馈入的装置。

20. 根据权利要求17所述的设备,其中所述较高阶环境立体混合声系数对应于的球面基底函数的阶次大于一。

21. 根据权利要求17所述的设备,其进一步包括用于内插所述多个不规则BRIR滤波器以产生多个规则BRIR滤波器的装置,其中所述多个规则BRIR滤波器包括用于扬声器的规则布置的多个BRIR滤波器,且

其中所述用于应用所述多个不规则BRIR滤波器的装置包括用于将所述多个规则BRIR滤波器应用到所述较高阶环境立体混合声系数以便呈现所述声场的装置。

22. 根据权利要求17所述的设备,其进一步包括:

用于将开窗函数应用到所述多个不规则BRIR滤波器以产生经开窗BRIR滤波器的装置,

其中所述用于应用所述多个不规则BRIR滤波器的装置包括用于将所述经开窗BRIR滤波器应用到所述较高阶环境立体混合声系数以便呈现所述声场的装置。

23. 根据权利要求17所述的设备,其进一步包括用于将所述多个不规则BRIR滤波器从时域变换到频域以便产生经变换不规则立体声房间脉冲响应滤波器的装置,

其中所述用于应用所述多个不规则BRIR滤波器的装置包括用于将所述经变换不规则BRIR滤波器应用到所述较高阶环境立体混合声系数以便呈现所述声场的装置。

24. 根据权利要求17所述的设备,其进一步包括:

用于将所述多个不规则BRIR滤波器从时域变换到频域以便产生经变换不规则BRIR滤

波器的装置;以及

用于将所述较高阶环境立体混合声系数从所述时域变换到所述频域以便产生经变换较高阶环境立体混合声系数的装置,

其中所述用于应用所述多个不规则BRIR滤波器的装置包括用于将所述经变换不规则BRIR滤波器应用到所述经变换较高阶环境立体混合声系数以便呈现所述声场的频域表示的装置,且

其中所述设备进一步包括用于将反变换应用到所述声场的所述频域表示以呈现所述声场的装置。

## 球面谐波系数的立体声呈现

### [0001] 优先权主张

[0002] 本申请案主张2013年5月29日申请的美国临时专利申请案第61/828,620号、2013年7月17日申请的美国临时专利申请案第61/847,543号、2013年10月3日申请的美国临时专利申请案第61/886,593号和2013年10月3日申请的美国临时申请案第61/886,620号的权益。

### 技术领域

[0003] 本发明涉及音频呈现,且更具体地说,涉及音频数据的立体声呈现。

### 发明内容

[0004] 一般来说,描述用于具有大于一的阶次的球面谐波系数(其可被称作较高阶环境立体混合声(HOA)系数)的立体声音频呈现的技术。

[0005] 作为一个实例,一种立体声音频呈现的方法包括将立体声房间脉冲响应滤波器应用到以三维形式表示声场的球面谐波系数以便呈现所述声场。

[0006] 在另一实例中,一种装置包括一或多个处理器,其经配置以将立体声房间脉冲响应滤波器应用到以三维形式表示声场的球面谐波系数以便呈现所述声场。

[0007] 在另一实例中,一种装置包括用于确定以三维形式表示声场的球面谐波系数的装置,和用于将立体声房间脉冲响应滤波器应用到表示声场的球面谐波系数以便呈现所述声场的装置。

[0008] 在另一实例中,一种非暂时性计算机可读存储媒体具有存储于其上的指令,所述指令在经执行时使一或多个处理器将立体声房间脉冲响应滤波器应用到以三维形式表示声场的球面谐波系数以便呈现所述声场。

[0009] 技术的一或多个方面的细节阐述于随附图式和以下描述中。这些技术的其它特征、目标和优势将从描述和图式以及从权利要求书而显而易见。

### 附图说明

[0010] 图1和2为说明各种阶次和次阶的球面谐波基底函数的图。

[0011] 图3为说明可执行本发明中所描述的技术以更有效率地呈现音频信号信息的系统的图。

[0012] 图4为说明实例立体声房间脉冲响应(BRIR)的框图。

[0013] 图5为说明用于产生房间中的BRIR的实例系统模型的框图。

[0014] 图6为说明用于产生房间中的BRIR的更深层系统模型的框图。

[0015] 图7为说明可执行本发明中所描述的立体声音频呈现技术的各种方面的音频回放装置的实例的框图。

[0016] 图8为说明可执行本发明中所描述的立体声音频呈现技术的各种方面的音频回放装置的实例的框图。

[0017] 图9为说明根据本发明中所描述的技术的各种方面的用于立体声呈现装置呈现球

面谐波系数的实例操作模式的流程图。

[0018] 图10A、10B描绘说明根据本发明中所描述的技术的各种方面的可由图7和8的音频回放装置执行的替代操作模式的流程图。

[0019] 图11为说明可执行本发明中所描述的立体声音频呈现技术的各种方面的音频回放装置的实例的框图。

[0020] 图12为说明根据本发明中所描述的技术的各种方面的可由图11的音频回放装置执行的过程的流程图。

[0021] 图13为说明可执行本发明中描述的立体声音频呈现技术的各种方面的音频回放装置的实例的框图。

[0022] 图14为说明可执行本发明中描述的立体声音频呈现的各种方面的音频回放装置的实例的框图。

[0023] 图15为说明根据本发明中描述的技术的各种方面的用于立体声呈现装置呈现球面谐波系数的实例操作模式的流程图。

[0024] 图16A、16B描绘各说明根据本发明中描述的技术的各种方面的可由图13、14的音频回放装置执行的概念过程的图。

[0025] 相似参考字符贯穿诸图和文本表示相似元件。

## 具体实施方式

[0026] 环绕声的演化现今已使得许多输出格式可用于娱乐。这些环绕声格式的实例包含流行5.1格式(其包含以下六个信道:左前(FL)、右前(FR)、中心或前心、左后或左环绕、右后或右环绕,和低频效应(LFE))、发展中的7.1格式,和即将到来的22.2格式(例如,供超高清晰度电视标准使用)。空间音频格式的另一实例为球面谐波系数(也被称为较高阶环境立体混合声)。

[0027] 到未来标准化音频编码器(将PCM音频表示转换成位流的装置——节省每时间样本所需的位的数目)的输入可任选地为三种可能格式中的一者:(i)基于传统信道的音频,其意味着通过预先指定位置处的扩音器进行播放;(ii)基于物件的音频,其涉及用于单一音频物件的离散脉码调制(PCM)数据与含有其位置坐标(以及其它信息)的相关联的元数据;和(iii)基于场景的音频,其涉及使用球面谐波系数(SHC)表示声场—其中系数表示球面谐波基底函数的线性加总的“权重”。关于这点,SHC可包含根据较高阶环境立体混合声(HoA)模型的HoA信号。球面谐波系数可替代地或另外包含平面模型和球面模型。

[0028] 市场中存在各种“环绕声”格式。举例来说,其范围为从5.1家庭影院系统(其为除了立体声系统以外在严重消耗起居室方面最成功的)到由NHK(日本广播协会或日本广播公司)开发的22.2系统。内容创建者(例如,好莱坞影城)将愿意产生用于电影的信道一次,但并不花费精力来针对每一扬声器配置进行重新混合。近来,标准委员会已在考虑提供编码成标准化位流和提供对于呈现器的位置处的扬声器几何形状和声学条件可适应和不可知的后续解码的方式。

[0029] 为了为内容创建者提供此灵活性,可使用元素的阶层集合来表示声场。元素的阶层集合可指元素经排序以使得较低阶元素的基本集合提供对模型化声场的完全表示的元素集合。当扩展集合以包含较高阶元素时,表示变得更详细。

[0030] 元素的阶层集合的一个实例为球面谐波系数 (SHC) 的集合。以下表达式表明使用 SHC 进行的声场的描述或表示:

$$[0031] \quad p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[ 4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t},$$

[0032] 此表达式展示了声场的任何点  $\{r_r, \theta_r, \varphi_r\}$  (在此实例中, 其以相对于俘获声场的麦克风的球面坐标来表达) 处的压力  $p_i$  可通过 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  $n$  阶球面贝塞尔函数, 且  $Y_n^m(\theta_r, \varphi_r)$  为  $n$  阶和  $m$  次阶球面谐波基底函数。可认识到, 方括号中的项为信号的频域表示 (即,  $S(\omega, r_r, \theta_r, \varphi_r)$ ), 其可通过例如离散傅立叶变换 (DFT)、离散余弦变换 (DCT) 或小波变换的各种时间频率变换来趋近。阶层集合的其它实例包含小波变换系数的集合和多重分辨率基底函数系数的其它集合。

[0033] 图1为说明从零阶 ( $n=0$ ) 到第四阶 ( $n=4$ ) 的球面谐波基底函数的图。如可见的, 对于每一阶, 存在  $m$  次阶的展开, 为了便于说明的目的, 在图1的实例中展示但未明确注释所述展开。

[0034] 图2为说明从零阶 ( $n=0$ ) 到第四阶 ( $n=4$ ) 的球面谐波基底函数的另一图。在图2中, 在三维坐标空间中展示球面谐波基底函数, 其中展示了阶次和次阶两者。

[0035] 在任何情况下, SHC  $A_n^m(k)$  可通过各种麦克风阵列配置来物理地获取 (例如, 记录), 或替代地, 其可从声场的基于信道或基于物件的描述而导出。SHC 表示基于场景的音频。举例来说, 第四阶 SHC 表示涉及每一时间样本  $(1+4)^2 = 25$  个系数。

[0036] 为了说明可如何从基于物件的描述导出这些 SHC, 考虑以下等式。对应于个别音频物件的声场的系数  $A_n^m(k)$  可表达为:

$$[0037] \quad A_n^m(k) = g(\omega)(-4\pi i k) h_n^{(2)}(kr_s) Y_n^{m*}(\theta_s, \varphi_s),$$

[0038] 其中  $i$  为  $\sqrt{-1}$ ,  $h_n^{(2)}(\cdot)$  为  $n$  阶球面汉克尔函数 (第二种), 且  $\{r_s, \theta_s, \varphi_s\}$  为物件的位置。知晓依据频率的源能量  $g(\omega)$  (例如, 使用时间频率分析技术, 例如, 对 PCM 流式传输执行快速傅立叶变换) 允许将每一 PCM 物件和其位置转换成 SHC  $A_n^m(k)$ 。此外, 可展示 (因为上述等式为线性和正交分解) 每一物件的  $A_n^m(k)$  系数为加成性的。以此方式, 大量 PCM 物件可由  $A_n^m(k)$  系数来表示 (例如, 作为个别物件的系数向量的总和)。基本上, 这些系数含有关于声场的信息 (依据 3D 坐标的压力), 且上述等式表示在观测点  $\{r_r, \theta_r, \varphi_r\}$  附近从个别物件到整个声场的表示的变换。

[0039] 也可从麦克风阵列记录导出 SHC 如下:

$$[0040] \quad a_n^m(t) = b_n(r_i, t) * \langle Y_n^m(\theta_i, \varphi_i), m_i(t) \rangle$$

[0041] 其中,  $a_n^m(t)$  为  $A_n^m(k)$  的时域等效于 (SHC),  $*$  表示卷积运算,  $\langle, \rangle$  表示内积,  $b_n(r_i, t)$

表示相依赖于 $r_i$ 的时域滤波函数, $m_i(t)$ 为第 $i$ 个麦克风信号,其中第 $i$ 个麦克风换能器位于半径 $r_i$ 、仰角 $\theta_i$ 和方位角 $\varphi_i$ 处。因此,如果麦克风阵列中存在32个换能器且每一麦克风定位于球体上以使得 $r_i=a$ 为常数(例如,来自mhAcoustics的Eigenmike EM32装置上的麦克风),那么可使用矩阵运算导出25个SHC如下:

$$[0042] \quad \begin{bmatrix} a_0^0(t) \\ a_1^{-1}(t) \\ \vdots \\ a_4^4(t) \end{bmatrix} = \begin{bmatrix} b_0(a,t) \\ b_1(a,t) \\ \vdots \\ b_4(a,t) \end{bmatrix} * \begin{bmatrix} Y_0^0(\theta_1, \varphi_1) & Y_0^0(\theta_2, \varphi_2) & \dots & Y_0^0(\theta_{32}, \varphi_{32}) \\ Y_1^{-1}(\theta_1, \varphi_1) & Y_1^{-1}(\theta_2, \varphi_2) & \dots & Y_1^{-1}(\theta_{32}, \varphi_{32}) \\ \vdots & \vdots & \ddots & \vdots \\ Y_4^4(\theta_1, \varphi_1) & Y_4^4(\theta_2, \varphi_2) & \dots & Y_4^4(\theta_{32}, \varphi_{32}) \end{bmatrix} \begin{bmatrix} m_1(a,t) \\ m_2(a,t) \\ \vdots \\ m_{32}(a,t) \end{bmatrix}.$$

[0043] 上述等式中的矩阵可更一般地被称作 $E_s(\theta, \varphi)$ ,其中下标 $s$ 可指示矩阵是针对某一换能器几何形状集合 $s$ 。上述等式中的卷积(通过 $*$ 指示)是基于逐列的,使得(例如)输出 $a_0^0(t)$ 为由 $E_s(\theta, \varphi)$ 矩阵的第一列与麦克风信号的行的向量乘法产生的 $b_0(a, t)$ 与时间序列之间的卷积的结果(其依据时间而变化——考虑向量乘法的结果为时间序列的事实)。当麦克风阵列的换能器位置在所谓的T-设计几何形状(其极接近于Eigenmike换能器几何形状)中时,计算可为最准确的。T-设计几何形状的一个特性可为:由几何形状产生的 $E_s(\theta, \varphi)$ 矩阵具有表现极好的逆(或伪逆),且另外,所述逆常常可通过矩阵 $E_s(\theta, \varphi)$ 的转置来极好地趋近。如果将忽略 $b_n(a, t)$ 的滤波运算,那么此性质将允许从SHC恢复麦克风信号(即,在此实例中, $[m_i(t)] = [E_s(\theta, \varphi)]^{-1} [SHC]$ )。下文在基于物件和基于SHC的音频译码的情境中描述剩余图。

[0044] 图3为说明可执行本发明中所描述的技术以更有效率地呈现音频信号信息的系统20的图。如图3的实例中所示,系统20包含内容创建者22和内容消费者24。虽然在内容创建者22和内容消费者24的情境中加以描述,但可在利用SHC或界定声场的阶层表示的任何其它阶层元素的任何情境中实施所述技术。

[0045] 内容创建者22可表示可产生用于供例如内容消费者24的内容消费者消费的多信道音频内容的电影工作室或其它物理。此内容创建者常常结合视频内容产生音频内容。内容消费者24可表示拥有或能够存取音频回放系统的个体,所述音频回放系统可指能够回放多信道音频内容的任何形式的音频回放系统。在图3的实例中,内容消费者24拥有或能够存取音频回放系统32以用于呈现界定声场的阶层表示的阶层元素。

[0046] 内容创建者22包含音频呈现器28和音频编辑系统30。音频呈现器28可表示呈现或以其它方式产生扬声器馈入(其也可被称作“扩音器馈入”、“扬声器信号”或“扩音器信号”)的音频处理单元。每一扬声器馈入可对应于再现多信道音频系统的特定信道的声音的扬声器馈入或对应于希望与匹配扬声器位置的头部相关转移函数(HRTF)滤波器进行卷积的虚拟扩音器馈入。每一扬声器馈入可对应于球面谐波系数信道(其中信道可通过球面谐波系数所对应于的相关联的球面基底函数的阶次和/或次阶表示),其使用SHC的多个信道来表示定向声场。

[0047] 在图3的实例中,音频呈现器28可呈现用于常规5.1、7.1或22.2环绕声格式的扬声器馈入,产生用于5.1、7.1或22.2环绕声扬声器系统中的5个、7个或22个扬声器中的每一者

的扬声器馈入。替代地,在给定上文所论述的源球面谐波系数的性质的情况下,音频呈现器28可经配置以呈现具有任何数目个扬声器的任何扬声器配置的来自源球面谐波系数的扬声器馈入。音频呈现器28可以此方式产生数个扬声器馈入,所述扬声器馈入在图3中经表示为扬声器馈入29。

[0048] 内容创建者可在编辑过程期间呈现球面谐波系数27 (“SHC 27”),从而在识别声场的不具有高保真度或不提供令人信服的环境声体验的方面的尝试中收听所呈现的扬声器馈入。内容创建者22接着可编辑源球面谐波系数(常常间接地通过操纵以上文所描述的方式导出源球面谐波系数所自的不同物件)。内容创建者22可使用音频编辑系统30来编辑球面谐波系数27。音频编辑系统30表示能够编辑音频数据和输出此音频数据作为一或多个源球面谐波系数的任何系统。

[0049] 当编辑过程完成时,内容创建者22可基于球面谐波系数27产生位流31。即,内容创建者22包含位流产生装置36,所述位流产生装置可表示能够产生位流31的任何装置。在一些例子中,位流产生装置36可表示编码器,所述编码器频宽压缩(作为一个实例,通过熵编码)球面谐波系数27且将球面谐波系数27的经熵编码版本布置成接受的格式以形成位流31。在其它例子中,位流产生装置36可表示音频编码器(可能为遵照例如MPEG环绕或其导出形式的已知音频译码标准的音频编码器),所述音频编码器使用(作为一个实例)类似于用以压缩多信道音频内容或其导出形式的常规音频环绕声编码过程的过程来编码多信道音频内容29。接着可以某种其它方式熵编码或译码经压缩的多信道音频内容29以频宽压缩内容29,且将内容根据商定的格式进行布置以形成位流31。不管是直接经压缩以形成位流31抑或经呈现且接着经压缩以形成位流31,内容创建者22均可将位流31传输到内容消费者24。

[0050] 虽然图3中经展示为直接传输到内容消费者24,但内容创建者22可将位流31输出到定位于内容创建者22与内容消费者24之间的中间装置。此中间装置可存储位流31以供稍后递送到可请求此位流的内容消费者24。所述中间装置可包括文件服务器、web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31以供稍后由音频解码器检索的任何其它装置。此中间装置可驻留于能够将位流31流式传输传输(和可能结合传输对应视频数据位流)到请求位流31的用户(例如,内容消费者24)的内容递送网络中。替代地,内容创建者22可将位流31存储到存储媒体(例如,紧密光盘、数字视频光盘、高清晰度视频光盘或其它存储媒体),大部分存储媒体能够由计算机读取且因此可被称作计算机可读存储媒体或非暂时性计算机可读存储媒体。关于这点,传输信道可指藉以传输存储到这些媒体的内容的那些信道(且可包含零售店和其它基于店铺的递送机构)。在任何情况下,本发明的技术因此不应在此方面限于图3的实例。

[0051] 如图3的实例中进一步展示,内容消费者24拥有或能够以另外方式存取音频回放系统32。音频回放系统32可表示能够回放多信道音频数据的任何音频回放系统。音频回放系统32包含立体声音频呈现器34,其呈现SHC 27'以供输出作为立体声扬声器馈入35A到35B(统称为“扬声器馈入35”)。立体声音频呈现器34可提供不同形式的呈现,例如执行向量基振幅水平移动(VBAP)的各种方式中的一或多者,和/或执行声场合成的各种方式中的一或多者。

[0052] 音频回放系统32可进一步包含提取装置38。提取装置38可表示能够通过大体上可

与位流产生装置36的过程互逆的过程提取球面谐波系数27' ("SHC 27'"), 其可表示球面谐波系数27的经修改形式或副本)的任何装置。在任何情况下, 音频回放系统32可接收球面谐波系数27' 且使用立体声音频呈现器34来呈现球面谐波系数27' 且由此产生扬声器馈入35 (对应于电耦合或可能无线耦合到音频回放系统32的扩音器的数目, 为了便于说明的目的, 在图3的实例中并未加以展示)。扬声器馈入35的数目可为二, 且音频回放系统可无线耦合到包含两个对应扩音器的一对头戴式耳机。然而, 在各种例子中, 立体声音频呈现器34可输出比所说明的且主要关于图3描述的扬声器馈入多或少的扬声器馈入。

[0053] 音频回放系统的二进制房间脉冲响应 (BRIR) 滤波器37各自表示在一位置处的对在脉冲位置处产生的脉冲的响应。BRIR滤波器37为“立体声”, 这是因为其各自经产生以表示如人耳在所述位置处将会体验到的脉冲响应。因此, 常常产生脉冲的BRIR滤波器, 且将其用于成对地声音呈现, 其中所述对中的一个元素是针对左耳, 且另一元素是针对右耳。在所说明实例中, 立体声音频呈现器34使用左BRIR滤波器33A和右BRIR滤波器33B来呈现相应立体声音频输出35A和35B。

[0054] 举例来说, 可通过对声源信号与经测量为脉冲响应 (IR) 的头部相关转移函数 (HRTF) 进行卷积来产生BRIR滤波器37。对应于BRIR滤波器37中的每一者的脉冲位置可表示虚拟空间中的虚拟扩音器的位置。在一些实例中, 立体声音频呈现器34对SHC 27' 与对应于虚拟扩音器的BRIR滤波器37进行卷积, 接着将所得卷积累加 (即, 加总) 以呈现由SHC 27' 界定的声场以供输出作为扬声器馈入35。如本文中所描述, 立体声音频呈现器34可应用用于通过操纵BRIR滤波器37同时呈现SHC 27' 作为扬声器馈入35来缩减呈现计算的技术。

[0055] 在一些例子中, 所述技术包含将BRIR滤波器37分段成表示房间内的一位置处的脉冲响应的不同阶段的数个片段。这些片段对应于在声场上的任何点处产生压力 (或缺少压力) 的不同物理现象。举例来说, 因为BRIR滤波器37中的每一者与脉冲一致地进行计时, 所以第一或“初始”片段可表示直到来自脉冲位置的压力波到达测量脉冲响应所在的位置为止的时间。除了时序信息以外, 相应初始片段的BRIR滤波器37值可能为无关紧要的, 且可被排除与描述声场的阶层元素进行卷积。类似地, 例如, BRIR滤波器37中的每一者可包含最后或“尾部”片段, 所述片段包含衰减到低于人类听觉的动态范围或衰减到低于指明阈值的脉冲响应信号。相应尾部片段的BRIR滤波器37值也可能为无关紧要的, 且可被排除与描述声场的阶层元素进行卷积。在一些实例中, 技术可包含通过执行与指明阈值的施罗德反向积分确定尾部片段, 和舍弃来自反向积分超过指明阈值的尾部片段的元素。在一些实例中, 混响时间 $RT_{60}$ 的指明阈值为-60dB。

[0056] BRIR滤波器37中的每一者的额外片段可表示在不包含来自房间的回波效应的情况下由脉冲产生的压力波引起的脉冲响应。这些片段可经表示和描述为BRIR滤波器37的头部相关转移函数 (HRTF), 其中HRTF俘获在压力波朝向鼓膜行进时归因于压力波围绕头部、肩膀/躯干和外耳的绕射和反射产生的脉冲响应。HRTF脉冲响应为线性和非时变系统 (LTI) 的结果且可经模型化为最小相位滤波器。在一些实例中, 用以缩减在呈现期间的HRTF片段计算的技术可包含最小相位重建和使用无限脉冲响应 (IIR) 滤波器来缩减原始有限脉冲响应 (FIR) 滤波器 (例如, HRTF滤波器片段) 的阶次。

[0057] 实施为IIR滤波器的最小相位滤波器可用以趋近具有缩减的滤波器阶次的BRIR滤波器37的HRTF滤波器。缩减阶次导致频域中的时间步长的计算的数目的伴随缩减。另外, 由

最小相位滤波器的构造产生的残余/过量滤波器可用以估计表示由声音压力波从源行进到每一耳朵的距离引起的时间或相位距离的耳间时间差 (ITD)。在计算一或多个BRIR滤波器37与描述声场的阶层元素的卷积 (即, 确定双耳立体声) 之后, 接着可使用ITD将一或两个耳朵的声音局部化模型化。

[0058] BRIR滤波器37中的每一者的又一片段是在HRTF片段之后且可考虑到房间对脉冲响应的影响。此房间片段可进一步分解成早期回波 (或“早期反射”) 片段和晚期混响片段 (即, 早期回波和晚期混响可各自由BRIR滤波器37中的每一者的单独片段表示)。在HRTF数据可用于BRIR滤波器37的情况下, 可通过BRIR滤波器37与HRTF的解卷积识别早期回波片段的开始, 以识别HRTF片段。在HRTF片段之后为早期回波片段。不同于残余房间响应, HRTF和早期回波片段为方向相依的, 这是因为对应虚拟扬声器的位置在显著方面确定信号。

[0059] 在一些实例中, 立体声音频呈现器34使用准备用于球面谐波域 ( $\theta, \phi$ ) 或描述声场的阶层元素的其它域的BRIR滤波器37。即, 可在球面谐波域 (SHD) 中将BRIR滤波器37界定为经变换的BRIR滤波器37, 以允许立体声音频呈现器34执行快速卷积, 同时利用数据集的某些性质, 包含BRIR滤波器37 (例如, 左/右) 和SHC 27' 的对称性。在这些实例中, 可通过将SHC呈现矩阵与原始BRIR滤波器相乘 (或在时域中进行卷积) 而产生经变换的BRIR滤波器37。在数学上, 可根据以下等式 (1) 到 (5) 来表达此情形:

$$[0060] \quad BRIR'_{(N+1)^2, L, left} = SHC_{(N+1)^2, L} * BRIR_{L, left} \quad (1)$$

$$[0061] \quad BRIR'_{(N+1)^2, L, right} = SHC_{(N+1)^2, L} * BRIR_{L, right} \quad (2)$$

[0062] 或

$$[0063] \quad BRIR'_{(N+1)^2, L, right} = \begin{bmatrix} Y_0^0(\theta_1, \phi_1) & Y_0^0(\theta_2, \phi_2) & \dots & Y_0^0(\theta_L, \phi_L) \\ Y_1^{-1}(\theta_1, \phi_1) & Y_1^{-1}(\theta_2, \phi_2) & \dots & Y_1^{-1}(\theta_L, \phi_L) \\ \vdots & \vdots & \ddots & \vdots \\ Y_4^4(\theta_1, \phi_1) & Y_4^4(\theta_2, \phi_2) & \dots & Y_4^4(\theta_L, \phi_L) \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_L \end{bmatrix}^T \quad (3)$$

$$[0064] \quad BRIR''_{(N+1)^2, left} = \sum_{k=0}^{L-1} [BRIR'_{(N+1)^2, k, left}] \quad (4)$$

$$[0065] \quad BRIR''_{(N+1)^2, right} = \sum_{k=0}^{L-1} [BRIR'_{(N+1)^2, k, right}] \quad (5)$$

[0066] 此处, (3) 以四阶球面谐波系数的矩阵形式描绘 (1) 或 (2) (其可为提及与四阶或更低阶的球面基底函数相关联的球面谐波系数的系数的替代方式)。当然, 可修改等式 (3) 以用于较高阶或较低阶球面谐波系数。等式 (4) 到 (5) 描绘扩音器维度L范围内的经变换的左和右BRIR滤波器37的加总, 以产生经加总的SHC立体声呈现矩阵 (BRIR'')。组合来说, 经加载的SHC立体声呈现矩阵具有维数  $[(N+1)^2, \text{Length}, 2]$ , 其中Length为可应用等式 (1) 到 (5) 的任何组合的脉冲响应向量的长度。在等式 (1) 和 (2) 的一些例子中, 可将呈现矩阵SHC立体声化以使得可将等式 (1) 修改成  $BRIR'_{(N+1)^2, L, left} = SHC_{(N+1)^2, L, left} * BRIR_{L, left}$ , 且可将等式 (2) 修改成  $BRIR'_{(N+1)^2, L, right} = SHC_{(N+1)^2, L} * BRIR_{L, right}$ 。

[0067] 上述等式 (1) 到 (3) 中所提出的SHC呈现矩阵SHC包含用于SHC 27' 的每一阶/次阶组合的元素, 其有效地界定单独SHC信道, 其中在球面谐波域中设定针对扬声器的位置L的元素值。BRIR<sub>L, left</sub>表示在左耳或在扬声器的位置L处产生的脉冲的位置处的BRIR响应, 且在

(3) 中使用脉冲响应向量 $B_i$ 来描绘,其中 $\{i | i \in [0, L]\}$ 。 $BRIR'_{(N+1)^2, L, left}$ 表示“SHC立体声呈现矩阵”的一半,即,经变换成球面谐波域的在左耳或在扬声器的位置 $L$ 处产生的脉冲的位置处的SHC立体声呈现矩阵。 $BRIR'_{(N+1)^2, L, right}$ 表示SHC立体声呈现矩阵的另一半。

[0068] 在一些实例中,技术可包含将SHC呈现矩阵仅应用于HRTF和相应原始BRIR滤波器37的早期反射片段以产生经变换的BRIR滤波器37和SHC立体声呈现矩阵。此情形可缩减与SHC 27'的卷积的长度。

[0069] 在一些实例中,如等式(4)到(5)中所描绘,可将具有将各种扩音器并入于球面谐波域中的维数的SHC立体声呈现矩阵加总以产生组合SHC呈现与BRIR呈现/混合的 $(N+1)^2 * Length * 2$ 滤波器矩阵。即,可通过(例如)将 $L$ 维度范围内的系数加总来组合 $L$ 个扩音器中的每一者的SHC立体声呈现矩阵。对于长度为 $Length$ 的SHC立体声呈现矩阵,此情形产生 $(N+1)^2 * Length * 2$ 加总的SHC立体声呈现矩阵,可将所述矩阵应用于球面谐波系数的音频信号以将信号立体声化。 $Length$ 可为根据本文中所描述的技术分段的BRIR滤波器的片段的长度。

[0070] 用于模型缩减的技术也可应用于经变更的呈现滤波器,其允许直接用新滤波器矩阵(经加总的SHC立体声呈现矩阵)来对SHC 27'(例如,SHC内容)进行滤波。立体声音频呈现器34接着可通过将经滤波的阵列加总以获得立体声输出信号35A、35B来转换成立体声音频。

[0071] 在一些实例中,音频回放系统32的BRIR滤波器37表示先前根据上文所描述的技术中的任何一或多者计算的球面谐波域中的经变换的BRIR滤波器。在一些实例中,可在执行时间执行原始BRIR滤波器37的变换。

[0072] 在一些实例中,因为BRIR滤波器37通常为对称的,所以技术可通过仅使用用于左耳或右耳的SHC立体声呈现矩阵来促成立体声输出35A、35B的计算的进一步缩减。当将由滤波器矩阵进行滤波的SHC 27'加总时,立体声音频呈现器34可作出关于在呈现最终输出时输出信号35A或35B作为第二信道的条件决策。如本文中所描述,对处理内容或修改关于左耳或右耳所描述的呈现矩阵的参考应被理解为可类似地应用于另一耳朵。

[0073] 以此方式,技术可提供多种途径来缩减BRIR滤波器37的长度以便可能地避免被排除的BRIR滤波器样本与多个信道的直接卷积。因此,立体声音频呈现器34可提供来自SHC 27'的立体声输出信号35A、35B的有效率的呈现。

[0074] 图4为说明实例立体声房间脉冲响应(BRIR)的框图。BRIR 40说明五个片段42A到42E。初始片段42A和尾部片段42E两者包含可能为无关紧要的且被排除进行呈现计算的静寂样本。头部相关转移函数(HRTF)片段42B包含归因于头部相关转移产生的脉冲响应且可使用本文中所描述的技术来识别。早期回波(替代地,“早期反射”)片段42C和晚期房间混响片段42D组合HRTF与房间效应,即,早期回波片段42C的脉冲响应匹配通过房间的早期回波和晚期混响进行滤波的BRIR 40的HRTF的脉冲响应。然而,早期回波片段42C可包含与晚期房间混响片段42D相比较来说更离散的回波。混合时间为早期回波片段42C与晚期房间混响片段42D之间的时间,且指示早期回波变为密集混响的时间。混合时间经说明为出现在到HRTF中的大约 $1.5 \times 10^4$ 个样本或从HRTF片段42B的开始的大约 $7.0 \times 10^4$ 个样本处。在一些实例中,技术包含使用来自房间音量的统计数据 and 估计计算混合时间。在一些实例中,具有50%信赖区间 $t_{mp50}$ 的感知混合时间大约为36毫秒(ms),且具有95%信赖区间 $t_{mp95}$ 的感知混合时间大约为80ms。在一些实例中,可使用相干性匹配噪声尾部来合成对应于BRIR 40的滤

波器的晚期房间混响片段42D。

[0075] 图5为说明用于产生房间中的BRIR(例如,图4的BRIR 40)的实例系统模型50的框图。模型包含串接系统,此处为房间52A和HRTF 52B。在将HRTF 52B应用于脉冲之后,脉冲响应匹配通过房间52A的早期回波进行滤波的HRTF的脉冲响应。

[0076] 图6为说明用于产生房间中的BRIR(例如,图4的BRIR 40)的更深层系统模型60的框图。此模型60也包含串接系统,此处为HRTF 62A、早期回波62B和残余房间62C(其组合HRTF与房间回波)。模型60描绘将房间52A分解成早期回波62B和残余房间62C且将每一系统62A、62B、62C视为线性非时变的。

[0077] 早期回波62B包含比残余房间62C更离散的回波。因此,早期回波62B可依据虚拟扬声器信道而变化,而具有较长尾部的残余房间62C可经合成为单一立体声复本。对于用以获得BRIR的一些测量人体模型,HRTF数据可为可用的,如在消声腔室中所测量。可通过将BRIR与HRTF数据解卷积以识别早期回波(其可被称作“反射”)的位置来确定早期回波62B。在一些实例中,HRTF数据并非容易得到的,且用于识别早期回波62B的技术包含盲估计。然而,简单明了的途径可包含将前几毫秒(例如,前5、10、15或20ms)视为通过HRTF进行滤波的直接脉冲。如上文所注明,技术可包含使用来自房间音量的统计数据和估计计算混合时间。

[0078] 在一些实例中,技术可包含合成残余房间62C的一或多个BRIR滤波器。在混合时间之后,在一些例子中,可互换BRIR混响尾部(表示为图6中的系统残余房间62C)而无感知惩罚。另外,可将BRIR混响尾部与匹配能量衰变减缓(EDR)和频率相依耳间相干性(FDIC)的高斯白噪声合成。在一些实例中,可产生BRIR滤波器的共同合成BRIR混响尾部。在一些实例中,共同EDR可为所有扬声器的EDR的平均值,或可为能量匹配平均能量的前零度EDR。在一些实例中,FDIC可为横跨所有扬声器的平均FDIC,或可为针对宽敞度的最大去相关测量的横跨所有扬声器的最小值。在一些实例中,也可通过反馈延迟网络(FDN)使用伪音混响来模拟混响尾部。

[0079] 在共同混响尾部的情况下,对应BRIR滤波器的晚期部分可被排除与每一扬声器馈入进行单独卷积,而是可一次应用于所有扬声器馈入的混合上。如上文所描述和下文更详细描述,可用球面谐波系数信号呈现进一步简化所有扬声器馈入的混合。

[0080] 图7为说明可执行本发明中所描述的立体声音频呈现技术的各种方面的音频回放装置的实例的框图。虽然经说明为单一装置(即,图7的实例中的音频回放装置100),但技术可由一或多个装置来执行。因此,技术在此方面应不受限制。

[0081] 如图7的实例中所示,音频回放装置100可包含提取单元104和立体声呈现单元102。提取单元104可表示经配置以从位流120中提取经编码音频数据的单元。提取单元104可将呈球面谐波系数(SHC) 122(其也可被称作较高阶环境立体混合声(HOA),这是因为SHC 122可包含与大于一的阶次相关联的至少一个系数)形式的经提取的经编码音频数据转递到立体声呈现单元146。

[0082] 在一些实例中,音频回放装置100包含经配置以解码经编码音频数据以便产生SHC122的音频解码单元。音频解码单元可执行在一些方面中与用以编码SHC 122的音频编码过程互逆的音频解码过程。音频解码单元可包含经配置以将经编码音频数据的SHC从时域变换到频域由此产生SHC 122的时间频率分析单元。即,当经编码音频数据表示未被从时域转换到频域的SHC 122的经压缩形式时,音频解码单元可调用时间频率分析单元将SHC从

时域转换到频域以便产生SHC 122(在频域中指定)。时间频率分析单元可应用任何形式的基于傅立叶的变换(包含快速傅立叶变换(FFT)、离散余弦变换(DCT)、经修改的离散余弦变换(MDCT)和离散正弦变换(DST))以提供将SHC从时域变换到频域中的SHC 122的几个实例。在一些例子中,SHC 122可能已在频域中在位流120中指定。在这些例子中,时间频率分析单元可将SHC 122传递到立体声呈现单元102而不应用变换或以其它方式变换所接收的SHC 122。虽然关于频域中指定的SHC 122加以描述,但可关于时域中指定的SHC 122执行技术。

[0083] 立体声呈现单元102表示经配置以将SHC 122立体声化的单元。换句话说,立体声呈现单元102可表示经配置以将SHC 122呈现到左和右信道的单元,其可以空间化为特征,从而将记录SHC 122的房间中的收听者将听到左和右信道的方式模型化。立体声呈现单元102可呈现SHC 122以产生适于经由例如头戴式耳机的耳机回放的左信道136A和右信道136B(其可被统称作“信道136”)。如图7的实例中所示,立体声呈现单元102包含BRIR滤波器108、BRIR调节单元106、残余房间响应单元110、BRIR SHC域转换单元112、卷积单元114和组合单元116。

[0084] BRIR滤波器108包含一或多个BRIR滤波器且可表示图3的BRIR滤波器37的实例。BRIR滤波器108可包含表示左和右HRTF对相应BRIR的效应的单独BRIR滤波器126A、126B。

[0085] BRIR调节单元106接收BRIR滤波器126A、126B的L个执行个体,每一虚拟扩音器L一个执行个体且其中每一BRIR滤波器具有长度N。BRIR滤波器126A、126B可能已经经调节以去除静寂样本。BRIR调节单元106可将上文所描述的技术应用于片段BRIR滤波器126A、126B以识别相应HRTF、早期反射和残余房间片段。BRIR调节单元106将HRTF和早期反射片段提供到BRIR SHC域转换单元112作为表示大小为[a, L]的左和右矩阵的矩阵129A、129B,其中a为HRTF和早期反射片段的串接的长度,且L为扩音器(虚拟的或真实的)的数目。BRIR调节单元106将BRIR滤波器126A、126B的残余房间片段提供到残余房间响应单元110作为大小为[b, L]的左和右残余房间矩阵128A、128B,其中b为残余房间片段的长度且L为扩音器(虚拟的或真实的)的数目。

[0086] 残余房间响应单元110可应用上文所描述的技术以计算或以其它方式确定用于与描述声场的阶层元素(例如,球面谐波系数)的至少某一部分进行卷积的左和右共同残余房间响应片段,如图7中通过SHC 122表示。即,残余房间响应单元110可接收左和右残余房间矩阵128A、128B且组合L范围内的相应左和右残余房间矩阵128A、128B以产生左和右共同残余房间响应片段。在一些例子中,残余房间响应单元110可通过对L范围内的左和右残余房间矩阵128A、128B求平均值来执行组合。

[0087] 残余房间响应单元110接着可计算左和右共同残余房间响应片段与SHC 122的至少一个信道(在图7中说明为信道124B)的快速卷积。在一些实例中,因为左和右共同残余房间响应片段表示环境的无方向性声音,所以信道124B为SHC 122信道中的W信道(即,第0阶),其编码声场的无方向性部分。在这些实例中,对于长度为Length的W信道样本,由残余房间响应单元110进行的与左和右共同残余房间响应片段的快速卷积产生长度为Length的左和右输出信号134A、134B。

[0088] 如本文中所使用,术语“快速卷积”和“卷积”可指时域中的卷积运算以及频域中的逐点乘法运算。换句话说,且如熟习信号处理的技术者众所周知,时域中的卷积等效于频域中的逐点乘法,其中时域和频域为彼此的变换。输出变换为输入变换与转移函数的逐点乘

积。因此,卷积和逐点乘法(或简称为“乘法”)可指关于相应域(本文中为时域和频域)作出的概念上类似的运算。卷积单元114、214、230;残余房间响应单元210、354;滤波器384和混响386可替代地应用频域中的乘法,其中在频域中而非在时域中提供到这些组件的输入。本文中描述为“快速卷积”或“卷积”的其它运算类似地也可指频域中的乘法,其中在频域中而非在时域中提供到这些运算的输入。

[0089] 在一些实例中,残余房间响应单元110可从BRIR调节单元106接收共同残余房间响应片段的开始时间的值。残余房间响应单元110可零填补或以其它方式延迟输出信号134A、134B,以预期与BRIR滤波器108的较早期片段组合。

[0090] BRIR SHC域转换单元112(下文中为“域转换单元112”)将SHC呈现矩阵应用于BRIR矩阵以可能地将左和右BRIR滤波器126A、126B转换成球面谐波域且接着可能地将L范围内的滤波器加总。域转换单元112输出转换结果分别作为左和右SHC立体声呈现矩阵130A、130B。在矩阵129A、129B的大小为 $[a, L]$ 的情况下,在对L范围内的滤波器加总之后,SHC立体声呈现矩阵130A、130B中的每一者的大小为 $[(N+1)^2, a]$ (参见(例如)等式(4)到(5))。在一些实例中,SHC立体声呈现矩阵130A、130B是在音频回放装置100中加以配置而不是在执行时间或设定时间加以计算。在一些实例中,SHC立体声呈现矩阵130A、130B的多个执行个体是在音频回放装置100中加以配置,且音频回放装置100选择多个执行个体的左/右对来应用于SHC 124A。

[0091] 卷积单元114将左和右立体声呈现矩阵130A、130B与SHC 124A进行卷积,所述SHC在一些实例中可按阶次从SHC 122的阶次缩减。对于频域(例如,SHC)中的SHC124A,卷积单元114可计算SHC 124A与左和右立体声呈现矩阵130A、130B的相应逐点乘法。对于长度为Length的SHC信号,卷积产生大小为 $[Length, (N+1)^2]$ 的左和右经滤波的SHC信道132A、132B,对于球面谐波域的每一阶/次阶组合通常存在每一输出信号矩阵的列。

[0092] 组合单元116可组合左和右经滤波的SHC信道132A、132B与输出信号134A、134B以产生立体声输出信号136A、136B。组合单元116接着可单独地对L范围内的每一左和右经滤波的SHC信道132A、132B加总以在组合左和右立体声输出信号与左和右输出信号134A、134B以产生立体声输出信号136A、136B之前产生HRTF和早期回波(反射)片段的左和右立体声输出信号。

[0093] 图8为说明可执行本发明中所描述的立体声音频呈现技术的各种方面的音频回放装置的实例的框图。音频回放装置200可进一步详细地表示图7的音频回放装置100的实例执行个体。

[0094] 音频回放装置200可包含可选SHC阶次缩减单元204,其处理来自位流240的入端口SHC 242以缩减SHC 242的阶次。可选SHC阶次缩减将SHC 242的最高阶(例如,第0阶)信道262(例如,W信道)提供到残余房间响应单元210,且将阶数缩减的SHC 242提供到卷积单元230。在SHC阶次缩减单元204不缩减SHC 242的阶次的例子中,卷积单元230接收与SHC 242相同的SHC 272。在任一状况下,SHC 272具有维度 $[Length, (N+1)^2]$ ,其中N为SHC 272的阶次。

[0095] BRIR调节单元206和BRIR滤波器208可表示图7的BRIR调节单元106和BRIR滤波器108的实例执行个体。残余响应单元210的卷积单元214接收由BRIR调节单元206使用上文所描述的技术调节的共同左和右残余房间片段244A、244B,且卷积单元214将共同左和右残余

房间片段244A、244B与最高阶信道262进行卷积以产生左和右残余房间信号262A、262B。延迟单元216可用到共同左和右残余房间片段244A、244B的样本的开始数目零填补左和右残余房间信号262A、262B以产生左和右残余房间输出信号268A、268B。

[0096] BRIR SHC域转换单元220(下文中为域转换单元220)可表示图7的域转换单元112的实例执行个体。在所说明实例中,变换单元222将具有 $(N+1)^2$ 维度的SHC呈现矩阵224应用于表示大小为 $[a, L]$ 的左和右矩阵的矩阵248A、248B,其中 $a$ 为HRTF和早期反射片段的串接的长度,且 $L$ 为扩音器(例如,虚拟扩音器)的数目。变换单元222输出SHC域中维度为 $[(N+1)^2, a, L]$ 的左和右矩阵252A、252B。加总单元226可将 $L$ 范围内的左和右矩阵252A、252B中的每一者加总以产生维度为 $[(N+1)^2, a]$ 的左和右中间SHC呈现矩阵254A、254B。缩减单元228可应用上文所描述的技术以进一步缩减将SHC呈现矩阵应用于SHC 272的计算复杂性,例如最小相位缩减,和使用平衡模型截断方法来设计IIR滤波器以趋近已应用最小相位缩减的中间SHC呈现矩阵254A、254B的相应最小相位部分的频率响应。缩减单元228输出左和右SHC呈现矩阵256A、256B。

[0097] 卷积单元230对呈SHC 272的形式的SHC内容进行滤波以产生中间信号258A、258B,加总单元232将所述中间信号加总以产生左和右信号260A、260B。组合单元234组合左和右残余房间输出信号268A、268B以及左和右信号260A、260B以产生左和右立体声输出信号270A、270B。

[0098] 在一些实例中,立体声呈现单元202可通过仅使用由变换单元222产生的SHC立体声呈现矩阵252A、252B中的一者实施对计算的进一步缩减。因此,卷积单元230可对左或右信号中的仅一者进行运算,从而将卷积运算缩减一半。在这些实例中,加总单元232作出在呈现输出260A、260B时关于第二信道的条件决策。

[0099] 图9为说明根据本发明中所描述的技术的用于立体声呈现装置呈现球面谐波系数的实例操作模式的流程图。为了说明的目的,关于图7的音频回放装置200描述实例操作模式。立体声房间脉冲响应(BRIR)调节单元206通过从BRIR滤波器246A、246B中提取方向相依分量/片段(具体地说,头部相关转移函数和早期回波片段)分别调节左和右BRIR滤波器246A、246B(300)。左和右BRIR滤波器126A、126B中的每一者可包含用于一或多个对应扩音器的BRIR滤波器。BRIR调节单元106将经提取的头部相关转移函数和早期回波片段的串接提供到BRIR SHC域转换单元220作为左和右矩阵248A、248B。

[0100] BRIR SHC域转换单元220应用H0A呈现矩阵224以变换包含经提取的头部相关转移函数和早期回波片段的左和右滤波器矩阵248A、248B以产生在球面谐波(例如,H0A)域中的左和右滤波器矩阵252A、252B(302)。在一些实例中,音频回放装置200可经配置而具有左和右滤波器矩阵252A、252B。在一些实例中,音频回放装置200接收位流240的频带外或频带内信号中的BRIR滤波器208,在所述状况下,音频回放装置200产生左和右滤波器矩阵252A、252B。加总单元226将扩音器维度范围内的相应左和右滤波器矩阵252A、252B加总以产生在SHC域中的立体声呈现矩阵,所述立体声呈现矩阵包含左和右中间SHC呈现矩阵254A、254B(304)。缩减单元228可进一步缩减中间SHC呈现矩阵254A、254B以产生左和右SHC呈现矩阵256A、256B。

[0101] 立体声呈现单元202的卷积单元230将左和右中间SHC呈现矩阵256A、256B应用于SHC内容(例如,球面谐波系数272)以产生左和右经滤波的SHC(例如,H0A)信道258A、258B

(306)。

[0102] 加总单元232将SHC维度  $(N+1)^2$  范围内的左和右经滤波的SHC信道258A、258B中的每一者加总以产生方向相依片段的左和右信号260A、260B (308)。组合单元116接着可组合左和右信号260A、260B与左和右残余房间输出信号268A、268B以产生包含左和右立体声输出信号270A、270B的立体声输出信号。

[0103] 图10A为说明根据本发明中所描述的技术的各种方面的可由图7和8的音频回放装置执行的实例操作模式310的图。下文关于图8的音频回放装置200描述操作模式310。音频回放装置200的立体声呈现单元202可经配置而具有可为BRIR滤波器208的实例执行个体的BRIR数据312,和可为HOA呈现矩阵224的实例执行个体的HOA呈现矩阵314。音频回放装置200可接收在关于位流240的频带内或频带外传讯信道中的BRIR数据312和HOA呈现矩阵314。BRIR数据312在此实例中具有表示例如L个真实或虚拟扩音器的L个滤波器,L个滤波器中的每一者具有长度K。L个滤波器中的每一者可包含左和右分量(“x 2”)。在一些状况下,L个滤波器中的每一者可包含用于左或右的单一分量,左或右与其对应物对称:右或左。此情形可缩减快速卷积的成本。

[0104] 音频回放装置200的BRIR调节单元206可通过应用分段和组合运算来调节BRIR数据312。具体地说,在实例操作模式310中,BRIR调节单元206根据本文中所描述的技术将L个滤波器中的每一者分段成HRTF加上具有组合长度a的早期回波片段以产生矩阵315(维数  $[a, 2, L]$ )和分段成残余房间响应片段以产生残余矩阵339(维数  $[b, 2, L]$ ) (324)。BRIR数据312的L个滤波器的长度K大约为a和b的总和。变换单元222可将维数为  $(N+1)^2$  的HOA/SHC呈现矩阵314应用于矩阵315的L个滤波器以产生维数为  $[(N+1)^2, a, 2, L]$  的矩阵317(其可为左和右矩阵252A、252B的组合适例执行个体)。加总单元226可将L范围内的左和右矩阵252A、252B中的每一者加总以产生维数为  $[(N+1)^2, a, 2]$  的中间SHC呈现矩阵335(第三维度具有表示左和右分量的值2;中间SHC呈现矩阵335可表示为左和右中间SHC呈现矩阵254A、254B两者的实例执行个体) (326)。在一些实例中,音频回放装置200可经配置而具有用于应用于HOA内容316(或其经缩减的版本,例如HOA内容321)的中间SHC呈现矩阵335。在一些实例中,缩减单元228可通过仅使用矩阵317的左或右分量中的一者而应用对计算的进一步缩减 (328)。

[0105] 音频回放装置200接收 $N_i$ 阶和长度为Length的HOA内容316,且在一些方面中,应用阶次缩减运算以将其中的球面谐波系数(SHC)的阶次缩减到N (330)。 $N_i$ 指示输入HOA内容321的阶次。阶次缩减运算 (330) 的HOA内容321如同SHC域中的HOA内容316。可选阶次缩减运算也产生最高阶(例如,第0阶)信号319并将其提供到残余响应单元210以用于进行快速卷积运算 (338)。在HOA阶次缩减单元204不缩减HOA内容316的阶次的例子中,应用快速卷积运算 (332) 对并不具有缩减的阶次的输入进行运算。在任一状况下,到快速卷积运算 (332) 的HOA内容321输入具有维度  $[Length, (N+1)^2]$ ,其中N为阶次。

[0106] 音频回放装置200可应用HOA内容321与矩阵335的快速卷积以产生具有左和右分量且因此具有维度  $[Length, (N+1)^2, 2]$  的HOA信号323 (332)。此外,快速卷积可指频域中的HOA内容321与矩阵335的逐点乘法或时域中的卷积。音频回放装置200可进一步将  $(N+1)^2$  范围内的HOA信号323加总以产生维度为  $[Length, 2]$  的经加总的信号325 (334)。

[0107] 现返回到残余矩阵339,音频回放装置200可根据本文中所描述的技术组合L个残

余房间响应片段,以产生维度为 $[b, 2]$ 的共同残余房间响应矩阵327 (336)。音频回放装置200可应用第0阶HOA信号319与共同残余房间响应矩阵327的快速卷积以产生维度为 $[Length, 2]$ 的房间响应信号329 (338)。因为为了产生残余矩阵339的 $L$ 个残余响应房间响应片段,音频回放装置200获得在BRIR数据312的 $L$ 个滤波器的第 $(a+1)$ 个样本处开始的残余响应房间响应片段,所以音频回放装置200通过延迟(例如,填补) $a$ 个样本以产生维度为 $[Length, 2]$ 的房间响应信号311而考虑到初始 $a$ 个样本 (340)。

[0108] 音频回放装置200通过将元素相加以产生维度为 $[Length, 2]$ 的输出信号318来组合经加总的信号325与房间响应信号311 (342)。以此方式,音频回放装置可避免应用 $L$ 个残余房间响应片段中的每一者的快速卷积。对于用于转换到立体声音频输出信号的22信道输入,此情形可将用于产生残余房间响应的快速卷积的数目从22缩减到2。

[0109] 图10B为说明根据本发明中所描述的技术的各种方面的可由图7和8的音频回放装置执行的实例操作模式350的图。下文关于图8的音频回放装置200描述操作模式350,且所述操作模式类似于操作模式310。然而,根据本文中所描述的技术,操作模式350包含首先将HOA内容呈现到 $L$ 个真实或虚拟扩音器的时域中的多信道扬声器信号中,且接着将有效率的BRIR滤波应用于扬声器馈入中的每一者。为此,音频回放装置200将HOA内容321变换到维度为 $[Length, L]$ 的多信道音频信号333 (344)。另外,音频回放装置不会将BRIR数据312变换到SHC域。因此,通过音频回放装置200对信号314应用缩减产生维度为 $[a, 2, L]$ 的矩阵337 (328)。

[0110] 音频回放装置200接着应用多信道音频信号333与矩阵337的快速卷积332以产生维度为 $[Length, L, 2]$  (具有左和右分量)的多信道音频信号341 (348)。音频回放装置200接着可通过 $L$ 个信道/扬声器将多信道音频信号341加总以产生维度为 $[Length, 2]$ 的信号325 (346)。

[0111] 图11为说明可执行本发明中所描述的立体声音频呈现技术的各种方面的音频回放装置350的实例的框图。虽然经说明为单一装置(即,图11的实例中的音频回放装置350),但所述技术可由一或多个装置来执行。因此,技术在此方面应不受限制。

[0112] 此外,虽然上文关于图1到10B的实例大体上描述为在球面谐波域中加以应用,但也可关于任何形式的音频信号实施技术,所述音频信号包含遵照上文所注明的环境声格式(例如,5.1环绕声格式、7.1环绕声格式,和/或22.2环绕声格式)的基于信道的信号。因此,技术也不应限于球面谐波域中所指定的音频信号,而是可关于任何形式的音频信号来应用。

[0113] 如图11的实例中所示,音频回放装置350可类似于图7的实例中所示的音频回放装置100。然而,音频回放装置350可操作或以其它方式执行关于一般基于信道的音频信号的技术,作为一个实例,所述信号遵照22.2环绕声格式。提取单元104可提取音频信道352,其中音频信道352大体上可包含“ $n$ ”个信道,且假定在此实例中包含遵照22.2环绕声格式的22个信道。将这些信道352提供到立体声呈现单元351的残余房间响应单元354和按信道截断的滤波器单元356两者。

[0114] 如上文所描述,BRIR滤波器108包含一或多个BRIR滤波器且可表示图3的BRIR滤波器37的实例。BRIR滤波器108可包含表示左和右HRTF对相应BRIR的效应的单独BRIR滤波器126A、126B。

[0115] BRIR调节单元106接收BRIR滤波器126A、126B的 $n$ 个执行个体,每一信道 $n$ 一个执行个体,且其中每一BRIR滤波器具有长度 $N$ 。BRIR滤波器126A、126B可能已经经调节以去除静寂样本。BRIR调节单元106可应用上文所描述的技术以将BRIR滤波器126A、126B分段以识别相应HRTF、早期反射和残余房间片段。BRIR调节单元106将HRTF和早期反射片段提供到按信道截断的滤波器单元356作为表示大小为 $[a, L]$ 的左和右矩阵的矩阵129A、129B,其中 $a$ 为HRTF和早期反射片段的串接的长度,且 $n$ 为扩音器(虚拟的或真实的)的数目。BRIR调节单元106将BRIR滤波器126A、126B的残余房间片段提供到残余房间响应单元354作为大小为 $[b, L]$ 的左和右残余房间矩阵128A、128B,其中 $b$ 为残余房间片段的长度且 $n$ 为扩音器(虚拟的或真实的)的数目。

[0116] 残余房间响应单元354可应用上文所描述的技术以计算或以其它方式确定用于与音频信道352进行卷积的左和右共同残余房间响应片段。即,残余房间响应单元110可接收左和右残余房间矩阵128A、128B且组合 $n$ 范围内的相应左和右残余房间矩阵128A、128B以产生左和右共同残余房间响应片段。在一些例子中,残余房间响应单元354可通过对 $n$ 范围内的左和右残余房间矩阵128A、128B求平均值来执行组合。

[0117] 残余房间响应单元354接着可计算左和右共同残余房间响应片段与音频信道352中的至少一者的快速卷积。在一些实例中,残余房间响应单元352可从BRIR调节单元106接收共同残余房间响应片段的开始时间的值。残余房间响应单元354可零填补或以其它方式延迟输出信号134A、134B,以预期与BRIR滤波器108的较早期片段组合。输出信号134A可表示左音频信号,而输出信号134B可表示右音频信号。

[0118] 按信道截断的滤波器单元356(在下文中为“经截断的滤波器单元356”)可将BRIR滤波器的HRTF和早期反射片段应用于信道352。更具体地说,按信道截断的滤波器单元356可将表示BRIR滤波器的HRTF和早期反射片段的矩阵129A和129B应用于信道352中的每一者。在一些例子中,可将矩阵129A和129B组合以形成单一矩阵129。此外,通常存在HRTF和早期反射矩阵129A和129B中的每一者中的左者以及HRTF和早期反射矩阵129A和129B中的每一者中的右者。即,通常存在用于左耳和右耳的HRTF和早期反射矩阵。按信道方向单元356可应用左和右矩阵129A、129B中的每一者以输出左和右经滤波的信道358A和358B。组合单元116可组合(或换句话说,混合)左经滤波的信道358A与输出信号134A,同时组合(或换句话说,混合)右经滤波的信道358B与输出信号134B,以产生立体声输出信号136A、136B。立体声输出信号136A可对应于左音频信道,且立体声输出信号136B可对应于右音频信道。

[0119] 在一些实例中,立体声呈现单元351可彼此同时发生地调用残余房间响应单元354和按信道截断的滤波器单元356,以使得残余房间响应单元354与按信道截断的滤波器单元356的操作同时发生地操作。即,在一些实例中,残余房间响应单元354可与按信道截断的滤波器单元356并行地(但经常并非同时地)操作,常常改善可产生立体声输出信号136A、136B的速度。虽然在上文的各图中经展示为可能地以串接方式操作,但除非以其它方式特别地指示,否则技术可提供本发明中所描述的单元或模块中的任一者的同时发生的或并行操作。

[0120] 图12为说明根据本发明中所描述的技术的各种方面的可由图11的音频回放装置350执行的过程380的图。过程380达成将每一BRIR分解成两个部分:(a)并有由左滤波器384AL到384NL和由右滤波器384AR到384NR(统称为“滤波器384”)表示的HRTF和早期反射的

效应的较小分量和 (b) 从原始BRIR的所有尾部的性质产生和由左混响滤波器386L和右混响滤波器386R(统称为“共同滤波器386”)表示的共同“混响尾部”。在过程380中展示的按信道滤波器384可表示上文所注明的部分 (a), 而过程380中所展示的共同滤波器386可表示上文所注明的部分 (b)。

[0121] 过程380通过分析BRIR以消除听不见的分量和确定包括HRTF/早期反射的分量和归因于晚期反射/漫射产生的分量来执行此分解。对于部分 (a), 此情形导致长度(作为一个实例) 为2704个分接头的FIR滤波器, 和对于部分 (b), 此情形导致长度(作为另一实例) 为15232个分接头的FIR滤波器。根据过程380, 在操作396中, 音频回放装置350可仅将较短FIR滤波器应用于个别 $n$ 个信道中的每一者, 出于说明的目的, 假定其为22。此运算的复数可表示于下文再现的等式 (8) 中的计算的第一部分中(使用4096点FFT)。在过程380中, 音频回放装置350可能不将共同“混响尾部”应用于22个信道中的每一者, 而是在操作398中将其全部应用于其加成性混合。此复数表示于等式 (8) 中的复数计算的后一半中, 其再次展示于随附附录中。

[0122] 在此方面, 过程380可表示基于来自 $N$ 个信道中的多个的混合音频内容产生复合音频信号的立体声音频呈现方法。另外, 过程380可进一步通过延迟将复合音频信号与 $N$ 个信道滤波器的输出对准, 其中每一信道滤波器包含经截断的BRIR滤波器。此外, 在过程380中, 音频回放装置350接着可在操作398中用共同合成残余房间脉冲响应对经对准的复合音频信号进行滤波, 且在立体声音频输出388L、388R的左和右分量的操作390L和390R中将每一信道滤波器的输出与经滤波的经对准的复合音频信号混合。

[0123] 在一些实例中, 经截断的BRIR滤波器和共同合成残余脉冲响应经预先载入于存储器中。

[0124] 在一些实例中, 在时间频域中执行经对准的复合音频信号的滤波。

[0125] 在一些实例中, 在时域中通过卷积执行经对准的复合音频信号的滤波。

[0126] 在一些实例中, 经截断的BRIR滤波器和共同合成残余脉冲响应是基于分解分析。

[0127] 在一些实例中, 对 $N$ 个房间脉冲响应中的每一者执行分解分析, 且其导致 $N$ 个经截断的房间脉冲响应和 $N$ 个残余脉冲响应(其中 $N$ 可在上文中表示为 $n$ 或 $n$ )。

[0128] 在一些实例中, 经截断的脉冲响应表示每一房间脉冲响应的总长度的小于百分的四十。

[0129] 在一些实例中, 经截断的脉冲响应包含在111与17,830之间的分接头范围。

[0130] 在一些实例中,  $N$ 个残余脉冲响应中的每一者组合成缩减复杂性的共同合成残余房间响应。

[0131] 在一些实例中, 将每一信道滤波器的输出与经滤波的经对准的复合音频信号混合包含用于左扬声器输出的混合的第一集合和用于右扬声器输出的混合的第二集合。

[0132] 在各种实例中, 上文所描述的过程380的各种实例或其任何组合的方法可由以下各者来执行: 包括存储器和一或多个处理器的装置、包括用于执行方法的每一步骤的装置的设备, 和通过执行存储于非暂时性计算机可读存储媒体上的指令执行所述方法的每一步骤的一或多个处理器。

[0133] 此外, 上文所描述的实例中的任一者中所阐述的特定特征中的任一者可组合成所描述的技术的有益实例。即, 特定特征中的任一者大体上适用于技术的所有实例。已描述技

术的各种实例。

[0134] 在一些情况下,本发明中所描述的技术可仅识别横跨BRIR集合的可听见的样本111到17830。从实例房间的体积计算混合时间 $T_{mp95}$ ,技术接着可使所有BRIR在53.6ms之后共享共同混响尾部,从而导致15232样本长的共同混响尾部和剩余2704样本HRTF+反射脉冲,其间具有3ms淡入淡出。在计算成本减轻方面,可出现以下情况:

[0135] (a) 共同混响尾部: $10*6*\log_2(2*15232/10)$ 。

[0136] (b) 剩余脉冲: $22*6*\log_2(2*4096)$ ,使用4096FFT来在一个帧中进行。

[0137] (c) 额外22个添加。

[0138] 因此,最终优值因此可大致等于 $C_{mod} = \max(100*(C_{conv}-C)/C_{conv}, 0) = 88.0$ ,其中:

[0139]  $C_{mod} = \max(100*(C_{conv}-C)/C_{conv}, 0)$ , (6)

[0140] 其中 $C_{conv}$ 为对未经优化的实施方案的估计:

[0141]  $C_{conv} = (22+2)*(10)*(6*\log_2(2*48000/10))$ , (7)

[0142] 在一些方面中,C可由两个加成性因子确定:

[0143]  $C = 22*6*\log_2(2*4096) + 10*6*\log_2(2*\frac{15232}{10})$ 。 (8)

[0144] 因此,在一些方面中,优值 $C_{mod} = 87.35$ 。

[0145] 可将表示为 $B_n(z)$ 的BRIR滤波器分解成两个函数 $BT_n(z)$ 和 $BR_n(z)$ ,其分别表示经截断的BRIR滤波器和混响BRIR滤波器。上文所注明的部分(a)可指此经截断的BRIR滤波器,而上文的部分(b)可指混响BRIR滤波器。 $B_n(z)$ 接着可等于 $BT_n(z) + (z^{-m}*BR_n(z))$ ,其中m表示延迟。输出信号 $Y(z)$ 因此可计算为:

[0146]  $\sum_{n=0}^{N-1} [X_n(z) \cdot BT_n(z) + z^{-m} \cdot X_n(z) \cdot BR_n(z)]$  (9)

[0147] 过程380可分析 $BR_n(z)$ 以导出共同合成混响尾部片段,其中可应用此共同BR(z)而不是信道特定 $BR_n(z)$ 。当使用此共同(或信道通用)合成BR(z)时, $Y(z)$ 可计算为:

[0148]  $\sum_{n=0}^{N-1} [X_n(z) \cdot BT_n(z) + z^{-m} BR_n(z)] \cdot \sum_{n=0}^{N-1} X_n(z)$  (10)

[0149] 图13为说明可执行本发明中描述的立体声音频呈现技术的各种方面的音频回放装置的实例的框图。虽然说明为单一装置(即,图13的实例中的音频回放装置400),但所述技术可由一或多个装置执行。因此,所述技术在此方面不应受限。此外,音频回放装置400可表示音频回放系统62的一个实例。

[0150] 如图13的实例中展示,音频回放装置400可包含提取单元404、BRIR选择单元424和立体声呈现单元402。提取单元404可表示经配置以从位流420提取经编码音频数据的单元。提取单元404可以球面谐波系数(SHC)422(其也可被称作较高阶环境立体混合声(HOA),其中SHC 422可包含与大于一的阶次相关联的至少一个系数)的形式将提取的经编码音频数据转发到立体声呈现单元146。BRIR选择单元424表示用户、用户代理或其它外部实体可提供用户输入425以根据本文中描述的技术选择将使用一组规则还是不规则的BRIR来立体声化SHC 422所借的接口。BRIR选择单元424可包含命令行或图形用户接口、应用程序编程接口、网络接口、应用程序接口(例如,简单对象访问协议)、远程程序调用或外部实体可配置将使用一组规则还是不规则的BRIR所借的任何其它接口。信号426表示引导或配置立体声呈现单元402使用一组规则还是不规则的BRIR用于立体声化SHC 422。信号426可表示旗标、

函数参数、信号或音频回放装置400可引导立体声呈现单元402选择一组规则还是不规则的BRIR用于立体声化SHC 422的任何其它方式。

[0151] 在一些实例中,音频回放装置400包含经配置以解码经编码音频数据以便产生SHC422的音频解码单元。音频解码单元可执行在一些方面与用以编码SHC 422的音频编码过程互逆的音频解码过程。音频解码单元可包含时频分析单元,其经配置以将经编码音频数据的SHC从时域变换到频域,由此产生SHC 422。即,当经编码音频数据表示未从时域转换到频域的SHC 422的经压缩形式时,音频解码单元可调用时频分析单元来将SHC从时域转换到频域,以便产生SHC 422(在频域中指定)。

[0152] 时频分析单元可应用任何形式的基于傅立叶的变换,包含快速傅里叶变换(FFT)、离散余弦变换(DCT)、经修改离散余弦变换(MDCT)和离散正弦变换(DST)(提供几个实例),以将SHC从时域变换到频域中的SHC 422。在一些情况下,SHC 422可已在位流420中指定于频域中。在这些情况下,时频分析单元可将SHC 422传送到立体声呈现单元402,而不应用变换或另外变换接收到的SHC 422。虽然关于在频域中指定的SHC 422来描述,但可关于在时域中指定的SHC 422来执行所述技术。

[0153] 立体声呈现单元402表示经配置以立体声化SHC 422的单元。换句话说,立体声呈现单元402可表示经配置以将SHC 422呈现到左信道和右信道的单元,左信道和右信道可表征空间化以模型化左信道和右信道将由在记录SHC 422的房间中的收听者听到的程度。立体声呈现单元402可呈现SHC 422,以产生左信道436A和右信道436B(其可共同地被称作“信道436”),其适合于经由头戴装置(例如,头戴式耳机)回放。如图13的实例中所展示,立体声呈现单元402包含内插单元406、时间频率分析单元408、复数BRIR单元410、加总单元442、复数乘法单元416、对称优化单元418、非对称优化单元420和反时间频率分析单元422。

[0154] 立体声呈现单元402可调用内插单元406内插不规则BRIR滤波器407A以便产生内插的规则BRIR滤波器407C,其中在BRIR滤波器的情况下,对“规则”或“不规则”的参考可表示扬声器相对于彼此的间距的规则性或不规则性。不规则BRIR滤波器407A可具有等于 $L \times 2$ (其中L表示扩音器的数目)的大小。规则BRIR滤波器407A可包括L个扩音器 $\times 2$ (假定这些扩音器规则地成对布置)。音频回放装置400的用户或其他操作者可指示或另外配置在SHC 422的立体声化期间将使用不规则BRIR滤波器407A还是规则BRIR滤波器407B。

[0155] 此外,音频回放装置400的用户或其他操作者可指示或另外配置当在SHC 422的立体声化期间将使用不规则BRIR滤波器407A时,是否将关于不规则BRIR滤波器407A执行内插以产生规则BRIR滤波器407C。内插单元406可使用基于向量的振幅水平移动或其它水平移动技术内插不规则BRIR滤波器407B以形成B数目个扩音器对,从而导致规则BRIR滤波器407C具有 $L \times 2$ 的大小(再次假定此规则且因此关于轴线对称)。虽然在图13的实例中未展示,但用户或其他操作者可经由用户接口与音频回放装置400介接,不管用户接口是经由图形用户接口以图形方式呈现还是物理呈现(例如,作为一系列按钮或其它输入)以选择在立体声化SHC 422时是否将使用不规则BRIR滤波器407A、规则BRIR滤波器407B和/或规则BRIR滤波器407C。

[0156] 无论如何,当BRIR滤波器407A到407C(取决于选择哪一者来立体声化SHC 422)在时域中呈现时,立体声呈现单元402可调用时间-频率分析单元408将BRIR滤波器407A到407C(“BRIR滤波器407”)中的选定者从时域变换到频域,从而分别导致经变换BRIR滤波器

409A到409C(“BRIR滤波器409”)。复数BRIR单元410表示经配置以关于不规则呈现器405A(具有 $L \times (N+1)^2$ 的大小)或规则呈现器405B(具有 $L \times (N+1)^2$ 的大小)和一或多个BRIR滤波器409执行逐元素复数乘法和加总以产生各自为 $L \times (N+1)^2$ 的大小的两个BRIR呈现向量411A和411B的单元,其中N再次表示SHC 422中的一或多个者对应于的球面基底函数的最高阶次。

[0157] 取决于BRIR滤波器407中的选定者规则还是不规则,复数BRIR单元410可选择不规则呈现器405A或规则呈现器405B。即,作为一个实例,当BRIR滤波器407中的选定者规则(例如,BRIR滤波器407B或407C),复数BRIR单元410选择规则呈现器405B。当BRIR滤波器407中的选定者不规则(例如,BRIR滤波器407A)时,复数BRIR单元410选择不规则呈现器405A。在一些实例中,音频回放装置400的用户或其他操作者可指示或另外选择使用不规则呈现器405A还是规则呈现器405B。在一些实例中,音频回放装置400的用户或其他操作者可指示或另外选择使用不规则呈现器405A还是规则呈现器405B,而非选择使用BRIR滤波器407中的一者(其中呈现器405A或405B的选择实现BRIR滤波器407中的所述一者的选择,例如,选择规则呈现器405B导致BRIR滤波器407B和/或407C的选择,且选择不规则呈现器405A导致BRIR滤波器407A的选择)。

[0158] 求和单元442可表示对L范围内的BRIR呈现向量411A和411B中的每一者加总以产生加总的BRIR呈现向量413A和413B的单元。开窗单元可表示将开窗函数应用到加总的BRIR呈现向量413A和413B中的每一者以产生经开窗BRIR呈现向量415A和415B的单元。开窗函数的实例可包含maxRE开窗函数、同相开窗函数和凯泽(Kaiser)开窗函数。复数乘法单元416表示用向量415A和415B中的每一者执行SHC 422的逐元素复数乘法以产生左修改的SHC 417A和右修改的SHC 417B的单元。

[0159] 立体声呈现单元402可接着潜在地基于由音频回放装置400的用户或其他操作者键入的配置数据调用对称优化单元418或非对称优化单元420中的任一者。即,当用户指定在SHC 422的立体声化期间将使用不规则BRIR滤波器407A时,立体声呈现单元402可确定不规则BRIR滤波器407A对称还是不对称。即,并非所有不规则BRIR滤波器407A都是不对称的,而是可对称。当不规则BRIR滤波器407A对称但并非规则地间隔时,立体声呈现单元402调用对称优化单元418以优化左修改的SHC 417A和右修改的SHC 417B的呈现。当不规则BRIR滤波器407A不对称时,立体声呈现单元402调用非对称优化单元420以优化左修改的SHC 417A和右修改的SHC 417B的呈现。当选择规则BRIR滤波器407B或407C时,立体声呈现单元402调用对称优化单元420以优化左修改的SHC 417A和右修改的SHC 417B的呈现。

[0160] 对称优化单元418当被调用时可加总在n阶次和m次阶上的左修改的SHC 417A和右修改的SHC 417B中的任一者。即,对称优化单元418可加总在n阶次和m次阶上的SHC 417A以产生频域左扬声器馈入419A。对称优化单元418可接着倒转具有负次阶的与球面基底函数相关联的SHC 417A中的那些,且接着将在n阶次和m次阶上的SHC417A的此经倒转型式加总以产生频域右扬声器馈入419B。非对称优化单元420当被调用时,将在n阶次和m次阶上的左修改的SHC 417A和右修改的SHC 417B中的每一者加总以分别产生频域左扬声器馈入421A和频域右扬声器馈入421B。反时间频率分析单元422可表示将频域左扬声器馈入419A或421A和对应的频域右扬声器馈入419B或421A从频域变换到时域以便产生左扬声器馈入436A和右扬声器馈入436B的单元。

[0161] 以此方式,所述技术使包括一或多个处理器的装置400能够将立体声房间脉冲响

应滤波器应用到以三维形式表示声场的球面谐波系数以便呈现声场。

[0162] 在一些实例中,所述一或多个处理器经进一步配置以当应用立体声房间脉冲响应滤波器时,将不规则立体声房间脉冲响应滤波器应用到球面谐波系数以便呈现所述声场,其中不规则立体声房间脉冲响应滤波器包括用于扬声器的不规则布置的一或多个立体声房间脉冲响应滤波器。

[0163] 在一些实例中,所述一或多个处理器经进一步配置以当应用立体声房间脉冲响应滤波器时,将规则立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场,其中所述规则立体声房间脉冲响应滤波器包括用于扬声器的规则布置的一或多个立体声房间脉冲响应滤波器。

[0164] 在一些实例中,所述一或多个处理器经进一步配置以内插不规则立体声房间脉冲响应滤波器以产生规则立体声房间脉冲响应滤波器。在这些和其它实例中,所述不规则立体声房间脉冲响应滤波器包括用于扬声器的不规则布置的一或多个立体声房间脉冲响应滤波器且所述规则立体声房间脉冲响应滤波器包括用于扬声器的规则布置的一或多个立体声房间脉冲响应滤波器。在这些和其它实例中,所述一或多个处理器经进一步配置以当应用所述立体声房间脉冲响应滤波器时,将所述规则立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场。

[0165] 在一些实例中,所述一或多个处理器经进一步配置以将开窗函数应用到所述立体声房间脉冲响应滤波器以产生经开窗立体声房间脉冲响应滤波器。在这些和其它实例中,所述一或多个处理器经进一步配置以当应用所述立体声房间脉冲响应滤波器时将所述经开窗立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场。

[0166] 在一些实例中,所述一或多个处理器经进一步配置以将所述立体声房间脉冲响应滤波器从时域变换到频域以便产生经变换立体声房间脉冲响应滤波器。在这些和其它实例中,所述一或多个处理器经进一步配置以当应用所述立体声房间脉冲响应滤波器时将所述经变换立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场。

[0167] 在一些实例中,所述一或多个处理器经进一步配置以将所述立体声房间脉冲响应滤波器从时域变换到频域以便产生经变换立体声房间脉冲响应滤波器,且将所述球面谐波系数从所述时域变换到所述频域以便产生经变换球面谐波系数。在这些和其它实例中,所述一或多个处理器经进一步配置以当应用所述立体声房间脉冲响应滤波器时,将所述经变换立体声房间脉冲响应滤波器应用到所述经变换球面谐波系数以便呈现所述声场的频域表示。在这些和其它实例中,所述一或多个处理器经进一步配置以将反变换应用到所述声场的所述频域表示以呈现所述声场。

[0168] 图14为说明可执行本发明中描述的立体声音频呈现的各种方面的音频回放装置的实例的框图。音频回放装置500可进一步详细地表示图1的音频回放系统62的另一实例执行个体。音频回放装置500可类似于图13的音频回放装置400,其中音频回放装置500包含执行类似于以上关于图13的音频回放装置400描述的操作的提取单元404、BRIR选择单元424和立体声呈现单元402。

[0169] 然而,音频回放装置500还可包含阶次缩减单元504,其处理入站的SHC 422以缩减SHC 422的阶次或次阶以产生阶次缩减的SHC 502。阶次缩减单元504可基于SHC 422的分析(例如,能量分析、方向性分析和其它形式的分析或其组合)执行此阶次缩减以从SHC 422去

除一或多个次阶 $m$ 或阶次 $n$ 。能量分析可涉及执行关于SHC 422的奇异值分解。方向性分析也可涉及执行关于SHC 422的奇异值分解。SHC 502可因此包含比SHC422少的阶次和/或次阶。

[0170] 阶次缩减单元504也可产生识别SHC 422的阶次和/或次阶的经去除以产生SHC 502的阶次缩减数据506。阶次缩减单元504可将此阶次缩减数据506和阶次缩减的SHC 502提供到立体声呈现单元402。音频回放装置500的立体声呈现单元402可功能上实质类似于音频回放装置400的立体声呈现单元402,除了音频回放装置500的立体声呈现单元402可基于阶次缩减的SHC 502更改呈现器405中的各者同时还关于阶次缩减的SHC502(而非非阶次缩减的SHC 422)操作之外。音频回放装置500的立体声呈现单元402可通过至少部分去除负责呈现SHC 422的经去除阶次和/或次阶的呈现器405的那些部分来基于阶次缩减数据506更改、修改或确定呈现器405。执行阶次缩减可缩减与SHC422的立体声化相关联的计算复杂性(就处理器循环和/或存储器消耗来说),通常不显著影响音频回放(就如所希望引入明显伪音或另外使声场的回放失真来说)。

[0171] 本发明中描述且图13到14的实例中展示的技术可提供通过频域中的一组规则或不规则BRIR使3D声场立体声化所借的有效率方式。如果一组不规则的BRIR 407A将由立体声呈现单元402用以呈现(例如)SHC 422,那么在一些情况下,立体声呈现单元402可将BRIR集合内插到一组规则间隔的BRIR 407C。此内插可经由线性内插法、向量基础振幅平移(VBAP)等进行。如果尚未在频域中,那么可使用(例如)快速傅立叶变换(FFT)、离散傅立叶变换(DFT)、离散余弦变换(DCT)、修改的DCT(MDCT)和抽取信号对角化(DSD)将待使用的BRIR集合(或“选定BRIR集合”)变换成频域。立体声呈现单元402可接着分别取决于规则BRIR滤波器407B或不规则BRIR滤波器407A的先前选择将待使用的BRIR集合与规则呈现器405B或不规则呈现器405A复数相乘。规则呈现器405B或不规则呈现器405A的阶次 $N$ 可由选择来确定以使用传入的HOA信号(例如,SHC 422)的完全阶次,使得 $N \leq N_I$ ,其中 $N_I$ 为传入的HOA信号的输入阶次或完全阶次。在图14的实例中应用阶次缩减操作的阶次缩减单元504也可影响在呈现器405A、406B还有BRIR内插中需要的扩音器的数目 $L$ 。然而,如果未选择BRIR集合的规则化,那么可将来自待使用的BRIR集合的 $L$ 的值向后馈入到阶次缩减504还有呈现器405A、406B。

[0172] 在呈现器405A、406B中的适当呈现器与待使用的BRIR集合的复数相乘后,可在 $L$ 范围上将输出的信号411A、411B加总以产生立体声化的HOA呈现器信号413A、413B。为了进一步增强呈现,可包含窗块,使得可使用例如 $\max_{Re}$ 、同相或凯泽的开窗函数改变 $n$ 、 $m$ ( $m$ 为HOA次阶)在频率上的加权。那些窗可帮助符合由盖尔宗(Gerzon)阐述的传统环境立体混合声准则,其给出了目标测量以符合音质准则。在此任选窗后,立体声呈现单元402将HOA信号与立体声化的HOA呈现器信号415A、415B复数相乘以产生立体声HOA信号417A、417B(这些是在本发明中其它处被描述为左修改的SHC 417A、右修改的SHC 417B的实例)。在一些情况下,所述技术也可允许对称的BRIR优化。如果立体声呈现单元402应用不对称优化,那么立体声呈现单元402将用于左信道和右信道的 $n$ 、 $m$  HOA系数加总。然而,如果立体声呈现单元402应用对称优化,那么立体声呈现单元402将用于左信道的 $n$ 、 $m$  HOA系数加总且输出。但归因于球面谐波基底函数的对称性,在加总前使 $m < 0$ 的值取倒数。贯穿以上描述的技术,可反向应用此对称性,其中仅确定BRIR集合的左侧。对于立体声输出436A、436B,立体声呈现单元402可将左和右信号变换回到时域(反变换)。

[0173] 以此方式,技术可a)包含3D(不仅仅2D),b)较高阶环境立体混合声(不仅仅一阶环境立体混合声)的立体声化,c)规则或不规则BRIR集合的应用,d)来自不规则到规则BRIR集合的BRIR的内插,e)BRIR信号的开窗以较好地匹配环境立体混合声再现准则;和f)至少部分通过利用频域计算而非时域计算在计算上潜在地改善效率。

[0174] 图15为说明根据本发明中描述的技术的用于立体声呈现装置呈现球面谐波系数的实例操作模式的流程图。出于说明的目的,关于图13的音频回放装置400描述所述实例操作模式。

[0175] 提取单元404可从位流420提取经编码音频数据。提取单元404可按球面谐波系数(SHC)422(其也可被称作较高阶环境立体混合声(HOA),其中SHC 422可包含与大于一的阶次相关联的至少一个系数)的形式将提取的经编码音频数据转发到立体声呈现单元146(600)。假定SHC 422已在位流420中在频域中指定,那么时间-频率分析单元可将SHC 422传送到立体声呈现单元402,而不应用变换或另外变换接收到的SHC 422。虽然关于在频域中指定的SHC 422描述,但可关于在时域中指定的SHC 422执行所述技术。

[0176] 无论如何,换句话说,立体声呈现单元402可表示经配置以将SHC 422对左信道和右信道呈现的单元,左信道和右信道可表征空间化以模型化左信道和右信道将由在记录SHC 422的房间中的收听者听到的程度。立体声呈现单元402可呈现SHC 422,以产生左信道436A和右信道436B(其可共同地被称作“信道436”),其适合于经由头戴装置(例如,头戴式耳机)回放。

[0177] 立体声呈现单元402可接收用户配置数据603以确定是否关于不规则BRIR滤波器407A、规则BRIR滤波器407B和/或内插的BRIR滤波器407C执行立体声呈现。换句话说,立体声呈现单元402可接收选择当执行SHC 422的立体声化时应使用滤波器407中的哪一者的用户配置数据603(602)。用户配置数据603可表示图13到14的信号426的实例。当用户配置数据603指定将使用规则BRIR滤波器407B(“是”604)时,立体声呈现单元402选择规则BRIR滤波器407B和规则呈现器405B(606)。当用户配置数据603指示将使用不规则BRIR滤波器407A(“否”604)而不内插此滤波器407A(“否”608)时,立体声呈现单元402选择不规则BRIR滤波器407A和不规则呈现器405A(610)。当用户配置数据603指示将使用不规则BRIR滤波器407A(“否”604)但将内插此滤波器407A(“是”608)时,立体声呈现单元402选择内插的BRIR滤波器407C(在调用内插单元406以内插选定滤波器407A以产生滤波器407C后)和规则呈现器405B(612)。

[0178] 无论如何,当BRIR滤波器407A到407C(取决于选择哪一者来立体声化SHC 422)在时域中呈现时,立体声呈现单元402可调用时间-频率分析单元408将BRIR滤波器407A到407C(“BRIR滤波器407”)中的选定者从时域变换到频域,从而分别导致经变换BRIR滤波器409A到409C(“BRIR滤波器409”)。复数BRIR单元410可执行关于呈现器405中的选定者和BRIR滤波器409中的选定者的逐元素复数相乘和加总以产生两个BRIR呈现向量411A和411B(614)。

[0179] 求和单元442可将L上的BRIR呈现向量411A和411B中的每一者加总以产生经加总的BRIR呈现向量413A和413B(616)。开窗单元可将开窗函数应用到加总的BRIR呈现向量413A和413B中的每一者以产生经开窗BRIR呈现向量415A和415B(618)。复数相乘单元416可接着用向量415A和415B中的每一者执行SHC 422的逐元素复数相乘以产生左修改的SHC

417A和右修改的SHC 417B(620)。

[0180] 如上所述,立体声呈现单元402可接着潜在地基于由音频回放装置400的用户或其他操作者键入的配置数据603调用对称优化单元418或非对称优化单元420中的任一者。

[0181] 对称优化单元418当被调用时可加总在n阶次和m次阶上的左修改的SHC 417A和右修改的SHC 417B中的仅一者。即,对称优化单元418可加总在n阶次和m次阶上的SHC 417A以产生频域左扬声器馈入419A。对称优化单元418可接着倒转具有负次阶的与球面基底函数相关联的SHC 417A中的那些,且接着将在n阶次和m次阶上的SHC417A的此型式加总以产生频域右扬声器馈入419A。

[0182] 非对称优化单元420当被调用时,将在n阶次和m次阶上的左修改的SHC 417A和右修改的SHC 417B中的每一者加总以分别产生频域左扬声器馈入421A和频域右扬声器馈入421B。反时间频率分析单元422可表示将频域左扬声器馈入419A或421A和对应的频域右扬声器馈入419B或421A从频域变换到时域以便产生左扬声器馈入436A和右扬声器馈入436B的单元。以此方式,立体声呈现单元402可执行关于左SHC 417A和右SHC 417B中的一或多者的优化以产生左扬声器馈入436A和右扬声器馈入436B(622)。音频回放装置400可继续上文所描述的方式操作,提取和立体声化SHC 422以呈现左扬声器馈入436A和右扬声器馈入436B(600到622)。

[0183] 图16A、16B描绘各说明根据本发明中描述的技术的各种方面的可由图13的音频回放装置400和图14的音频回放装置500执行的概念过程的图。由较高阶环境立体混合声(HOA)系数组成的空间声场的立体声化传统地涉及将HOA信号对扩音器信号呈现且接着将扩音器信号与针对那个扩音器位置选取的BRIR的左和右型式卷积。此传统方法可在计算上代价大,因为此传统方法通常产生的(L个扩音器中的)每个扩音器信号需要两个卷积,其中必须存在比存在的HOA系数多的扩音器。换句话说, $L > (N+1)^2$ ——对于多声道扩音器阵列,其中N为环境立体混合声阶次。用于定义两个维度上的声场的经典一阶环境立体混合声的方法处理规则(在一些情况下,意味着等间隔)虚拟扩音器布置,以用于再现一阶环境立体混合声内容。可将此方法考虑为过分简单,倘若此方法假定最佳情况情境且不将关于较高阶环境立体混合声或其应用的信息提供到三个维度。此方法也未提到频域计算,但依赖于时域内的卷积。

[0184] 本发明中描述且图8的实例中展示的技术可提供通过频域中的一组规则或不规则BRIR使3D声场立体声化所借的有效率方式。如果使用一组不规则的BRIR,那么可存在将所述BRIR集合内插到一组规则间隔的BRIR的选择。此内插可经由线性内插法、向量基础振幅平移(VBAP)等进行。如图16A中所描绘,如果尚未在频域中,那么在一些实例中,可使用快速傅立叶变换(FFT)、离散傅立叶变换(DFT)、离散余弦变换(DCT)、MDCT和DSD(仅举几个实例)将待使用的BRIR集合变换成频域。取决于先前规则/不规则选择,可接着将BRIR集合与规则或不规则呈现器复数相乘。规则或不规则呈现器的阶次N可由选择控管以使用传入的HOA信号的完全阶次,使得 $N \leq N_I$ 。图16A、16B的实例中的‘阶次缩减’也可影响在呈现器还有BRIR内插两者中需要的扩音器的数目L。然而,如果未选择BRIR集合的规则化,那么可将来自BRIR集合的L的值反向馈入到阶次缩减还有呈现器内。

[0185] 在正确呈现器与正确BRIR信号集合的复数相乘后,可在L维度上将输出的信号加总以产生立体声化的HOA呈现器信号。为了进一步增强呈现,可包含窗块,使得可使用例如

maxRe、同相或凯泽的开窗函数改变 $n$ 、 $m$ 在频率上的加权。那些窗可帮助符合由盖尔宗阐述的传统环境立体混合声准则,其给出了目标测量以符合音质准则。在此任选窗后,将HOA(如果在如图16A中所描绘的频域中)与立体声化的HOA呈现器信号复数相乘。如果HOA在时域中,那么可将HOA与立体声化的HOA呈现器信号快速卷积,如图16B中所描绘。

[0186] 在一些情况下,所述技术也可允许对称的BRIR优化。如果执行非优化的路线,那么可针对左信道和右信道加总 $n$ 、 $m$  HOA系数。如果选择对称路径,那么用于左边的输出的信号为 $n$ 、 $m$ 值的总和,但归因于球面谐波基底函数的对称性,在加总前,将 $m < 0$ 的值取倒数。贯穿以上描述的技术,可反向应用此对称性,其中仅确定BRIR集合的左侧。可接着将左和右信号变换回到时域(反变换)以用于立体声输出。

[0187] 以此方式,技术可a)包含3D(不仅仅2D),b)立体声化较高阶环境立体混合声(不仅仅一阶环境立体混合声),c)应用规则或不规则BRIR集合,d)执行来自不规则到规则BRIR集合的BRIR的内插,e)执行BRIR信号的开窗以较好地匹配环境立体混合声再现准则;和f)至少部分通过利用频域计算而非时域计算在计算上潜在地改善效率(再次,如图16A中所描绘)。

[0188] 除了以上之外或作为对以上的替代,还描述以下实例。在以下实例中的任一者中描述的特征可与本文中描述的其它实例中的任一者一起利用。

[0189] 一个实例是针对一种立体声音频呈现的方法,其包括将立体声房间脉冲响应滤波器应用到以三维形式表示声场的球面谐波系数以便呈现所述声场。

[0190] 在一些实例中,应用所述立体声房间脉冲响应滤波器包括将不规则立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场,其中所述不规则立体声房间脉冲响应滤波器包括用于扬声器的不规则布置的一或多个立体声房间脉冲响应滤波器。

[0191] 在一些实例中,应用所述立体声房间脉冲响应滤波器包括将规则立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场,其中所述规则立体声房间脉冲响应滤波器包括用于扬声器的规则布置的一或多个立体声房间脉冲响应滤波器。

[0192] 在一些实例中,球面谐波系数对应于的球面基底函数的阶次大于一。

[0193] 在一些实例中,所述方法进一步包括内插不规则立体声房间脉冲响应滤波器以产生规则立体声房间脉冲响应滤波器,其中所述不规则立体声房间脉冲响应滤波器包括用于扬声器的不规则布置的一或多个立体声房间脉冲响应滤波器,且所述规则立体声房间脉冲响应滤波器包括用于扬声器的规则布置的一或多个立体声房间脉冲响应滤波器,且应用所述立体声房间脉冲响应滤波器包括将所述规则立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场。

[0194] 在一些实例中,所述方法进一步包括将开窗函数应用到所述立体声房间脉冲响应滤波器以产生经开窗立体声房间脉冲响应滤波器,且应用所述立体声房间脉冲响应滤波器包括将所述经开窗立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场。

[0195] 在一些实例中,所述方法进一步包括将所述立体声房间脉冲响应滤波器从时域变换到频域以便产生经变换立体声房间脉冲响应滤波器,且应用所述立体声房间脉冲响应滤波器包括将所述经变换立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场。

[0196] 在一些实例中,所述方法进一步包括将所述立体声房间脉冲响应滤波器从时域变换到频域以便产生经变换立体声房间脉冲响应滤波器;以及将所述球面谐波系数从所述时域变换到所述频域以便产生经变换球面谐波系数,其中应用所述立体声房间脉冲响应滤波器包括将所述经变换立体声房间脉冲响应滤波器应用到所述经变换球面谐波系数以便呈现所述声场的频域表示,且其中所述方法进一步包括将反变换应用到所述声场的所述频域表示以呈现所述声场。

[0197] 一个实例是针对一种装置,其包括一或多个处理器,所述一或多个处理器经配置以将立体声房间脉冲响应滤波器应用到以三维形式表示声场的球面谐波系数以便呈现所述声场。

[0198] 在一些实例中,所述一或多个处理器经进一步配置以当应用立体声房间脉冲响应滤波器时,将不规则立体声房间脉冲响应滤波器应用到球面谐波系数以便呈现所述声场,其中不规则立体声房间脉冲响应滤波器包括用于扬声器的不规则布置的一或多个立体声房间脉冲响应滤波器。

[0199] 在一些实例中,所述一或多个处理器经进一步配置以当应用立体声房间脉冲响应滤波器时,将规则立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场,其中所述规则立体声房间脉冲响应滤波器包括用于扬声器的规则布置的一或多个立体声房间脉冲响应滤波器。

[0200] 在一些实例中,球面谐波系数对应于的球面基底函数的阶次大于一。

[0201] 在一些实例中,所述一或多个处理器经进一步配置以内插不规则立体声房间脉冲响应滤波器以产生规则立体声房间脉冲响应滤波器,其中所述不规则立体声房间脉冲响应滤波器包括用于扬声器的不规则布置的一或多个且所述规则立体声房间脉冲响应滤波器包括用于扬声器的规则布置的一或多个立体声房间脉冲响应滤波器,且所述一或多个处理器经进一步配置以当应用所述立体声房间脉冲响应滤波器时将所述规则立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场。

[0202] 在一些实例中,所述一或多个处理器经进一步配置以将开窗函数应用到所述立体声房间脉冲响应滤波器以产生经开窗立体声房间脉冲响应滤波器,且所述一或多个处理器经进一步配置以当应用所述立体声房间脉冲响应滤波器时,将所述经开窗立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场。

[0203] 在一些实例中,所述一或多个处理器经进一步配置以将所述立体声房间脉冲响应滤波器从时域变换到频域以便产生经变换立体声房间脉冲响应滤波器,且所述一或多个处理器经进一步配置以当应用所述立体声房间脉冲响应滤波器时,将所述经变换立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场。

[0204] 在一些实例中,所述一或多个处理器经进一步配置以将所述立体声房间脉冲响应滤波器从时域变换到频域以便产生经变换立体声房间脉冲响应滤波器,和将所述球面谐波系数从所述时域变换到所述频域以便产生经变换球面谐波系数,所述一或多个处理器经进一步配置以当应用所述立体声房间脉冲响应滤波器时,将所述经变换立体声房间脉冲响应滤波器应用到所述经变换球面谐波系数以便呈现所述声场的频域表示,且所述一或多个处理器经进一步配置以将反变换应用到所述声场的所述频域表示以呈现所述声场。

[0205] 一个实例是针对一种装置,其包括用于确定以三维形式表示声场的球面谐波系数

的装置,和用于将立体声房间脉冲响应滤波器应用到表示声场的球面谐波系数以便呈现所述声场的装置。

[0206] 在一些实例中,所述用于应用所述立体声房间脉冲响应滤波器的装置包括用于将不规则立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场的装置,其中所述不规则立体声房间脉冲响应滤波器包括用于扬声器的不规则布置的一或多个立体声房间脉冲响应滤波器。

[0207] 在一些实例中,所述用于应用所述立体声房间脉冲响应滤波器的装置包括用于将规则立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场的装置,且所述规则立体声房间脉冲响应滤波器包括用于扬声器的规则布置的一或多个立体声房间脉冲响应滤波器。

[0208] 在一些实例中,球面谐波系数对应于的球面基底函数的阶次大于一。

[0209] 在一些实例中,所述装置进一步包括用于内插不规则立体声房间脉冲响应滤波器以产生规则立体声房间脉冲响应滤波器的装置,所述不规则立体声房间脉冲响应滤波器包括用于扬声器的不规则布置的一或多个立体声房间脉冲响应滤波器,且所述规则立体声房间脉冲响应滤波器包括用于扬声器的规则布置的一或多个立体声房间脉冲响应滤波器,且所述用于应用所述立体声房间脉冲响应滤波器的装置包括用于将所述规则立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场的装置。

[0210] 在一些实例中,所述装置进一步包括用于将开窗函数应用到所述立体声房间脉冲响应滤波器以产生经开窗立体声房间脉冲响应滤波器的装置,且所述用于应用所述立体声房间脉冲响应滤波器的装置包括用于将所述经开窗立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场的装置。

[0211] 在一些实例中,所述装置进一步包括用于将所述立体声房间脉冲响应滤波器从时域变换到频域以便产生经变换立体声房间脉冲响应滤波器的装置,且用于应用所述立体声房间脉冲响应滤波器的装置包括用于将所述经变换立体声房间脉冲响应滤波器应用到所述球面谐波系数以便呈现所述声场的装置。

[0212] 在一些实例中,所述装置进一步包括用于将所述立体声房间脉冲响应滤波器从时域变换到频域以便产生经变换立体声房间脉冲响应滤波器的装置;以及用于将所述球面谐波系数从所述时域变换到所述频域以便产生经变换球面谐波系数的装置,且所述用于应用所述立体声房间脉冲响应滤波器的装置包括用于将所述经变换立体声房间脉冲响应滤波器应用到所述经变换球面谐波系数以便呈现所述声场的频域表示的装置,且所述装置进一步包括用于将反变换应用到所述声场的所述频域表示以呈现所述声场的装置。

[0213] 一个实例是针对一种非暂时性计算机可读存储媒体,其具有存储于其上的指令,所述指令在经执行时使一或多个处理器将立体声房间脉冲响应滤波器应用到以三维形式表示声场的球面谐波系数以便呈现所述声场。

[0214] 此外,以上描述的实例中的任一者中所阐明的特定特征中的任一者可组合到所描述技术的有益实例中。即,所述特定特征中的任一者通常适用于本发明的所有实例。已描述本发明的各种实例。

[0215] 应理解,取决于实例,本文中所描述的方法中的任一者的某些动作或事件可按不同序列来执行,可经添加、合并或一起省去(例如,并非所有所描述的动作或事件为达成方

法的实践所必要的)。此外,在某些实例中,动作或事件可(例如)通过多线程处理、中断处理或多个处理器而同时执行而非顺序执行。另外,虽然出于清晰的目的,本发明的某些方面经描述为由单一装置、模块或单元执行,但应理解,本发明的技术可由装置、单元或模块的组合执行。

[0216] 在一或多个实例中,可以硬件、软件、固件或其任何组合来实施所描述的功能。如果以软件来实施,那么功能可作为一或多个指令或代码而存储于计算机可读媒体上或通过计算机可读媒体予以传输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或通信媒体,计算机可读存储媒体对应于例如数据存储媒体的有形媒体,通信媒体包含促进计算机程序(例如)根据通信协议从一处传送到另一处的任何媒体。

[0217] 以此方式,计算机可读媒体大体上可对应于(1)非暂时性的有形计算机可读存储媒体,或(2)例如信号或载波的通信媒体。数据存储媒体可为可由一或多个计算机或一或多个处理器存取以检索用于实施本发明中所描述的技术的指令、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。

[0218] 通过实例且非限制,这些计算机可读存储媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置、磁盘存储装置,或其它磁性存储装置、快闪存储器,或可用以存储呈指令或数据结构的形式且可由计算机存取的所要程序代码的任何其它媒体。又,将任何连接恰当地称为计算机可读媒体。举例来说,如果使用同轴电缆、光缆、双绞线、数字用户线(DSL)或例如红外线、无线电和微波的无线技术从网站、服务器或其它远程源传输指令,那么同轴电缆、光缆、双绞线、DSL或例如红外线、无线电和微波的无线技术包含于媒体的定义中。

[0219] 然而,应理解,计算机可读存储媒体和数据存储媒体不包含连接、载波、信号或其它暂时性媒体,而是针对非暂时性有形存储媒体。如本文中所使用,磁盘和光盘包含紧密光盘(CD)、激光光盘、光学光盘、数字影音光盘(DVD)、软盘和蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘通过激光以光学方式再现数据。以上各者的组合也应包含于计算机可读媒体的范围内。

[0220] 可由例如一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其它等效集成或离散逻辑电路的一或多个处理器来执行指令。因此,如本文中所使用,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何其它结构中的任一者。另外,在一些方面中,可将本文所描述的功能性提供于经配置以用于编码和解码的专用硬件和/或软件模块内,或并入于组合式编码解码器中。又,所述技术可完全实施于一或多个电路或逻辑元件中。

[0221] 本发明的技术可实施于广泛多种装置或设备中,包含无线手机、集成电路(IC)或IC的集合(例如,芯片集)。本发明中描述各种组件、模块或单元以强调经配置以执行所揭示的技术的装置的功能方面,但未必需要由不同硬件单元来实现。相反地,如上文所描述,各种单元可组合于编码解码器硬件单元中或由互操作硬件单元的集合(包含如上文所描述的一或多个处理器)结合合适的软件和/或固件来提供。

[0222] 已描述技术的各种实施例。这些和其它实施例在以下权利要求书的范围内。

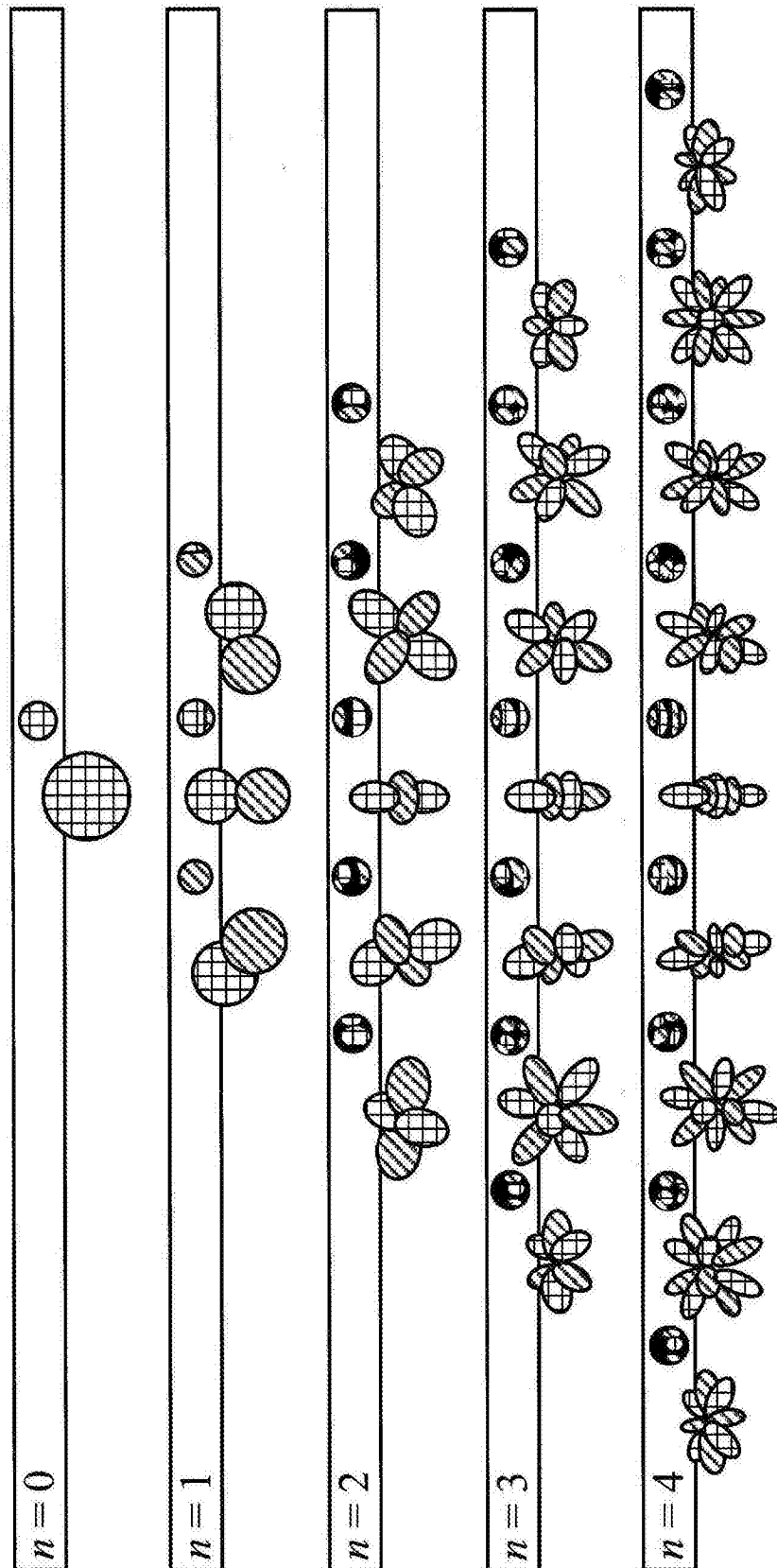


图1

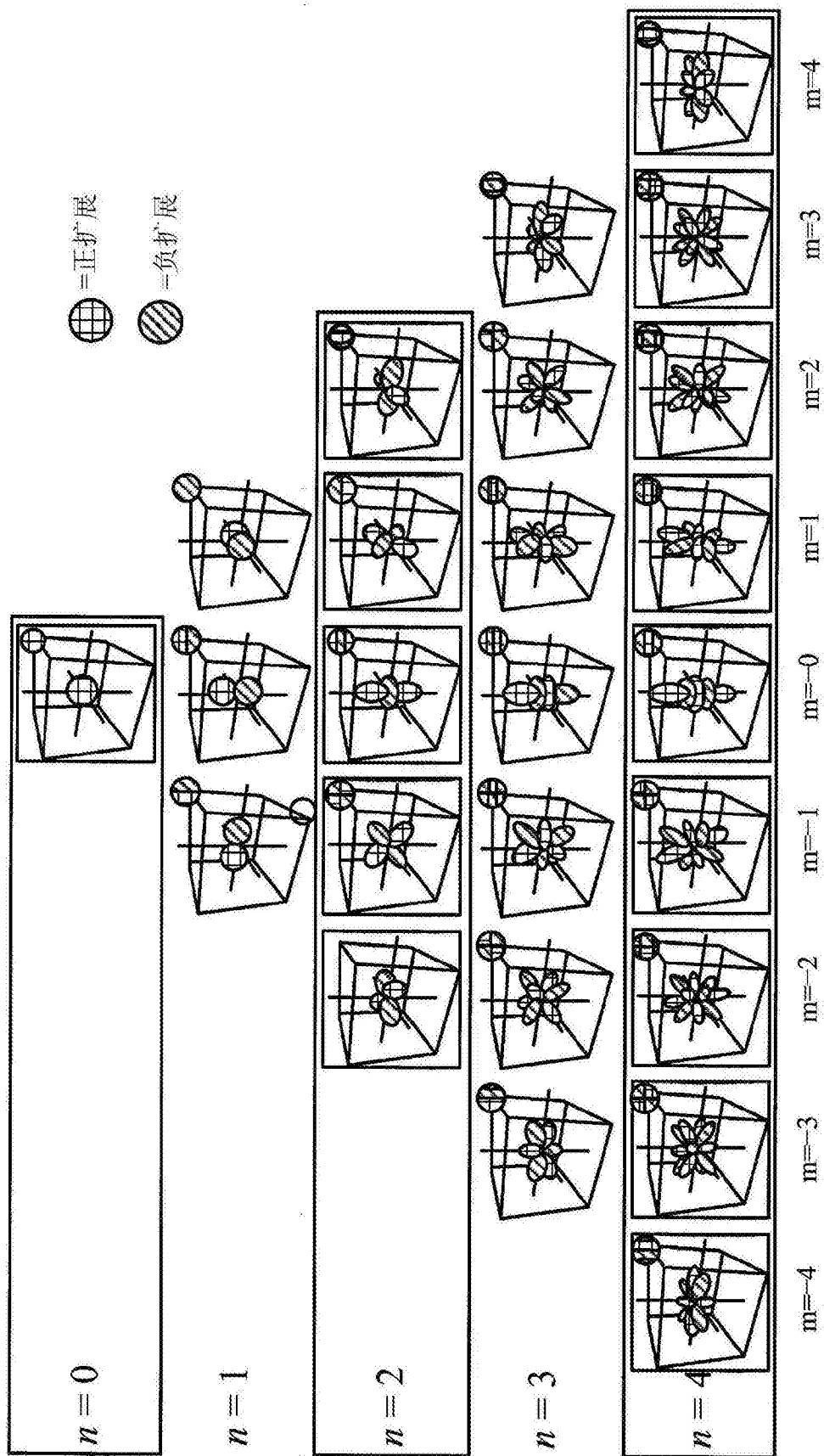


图2

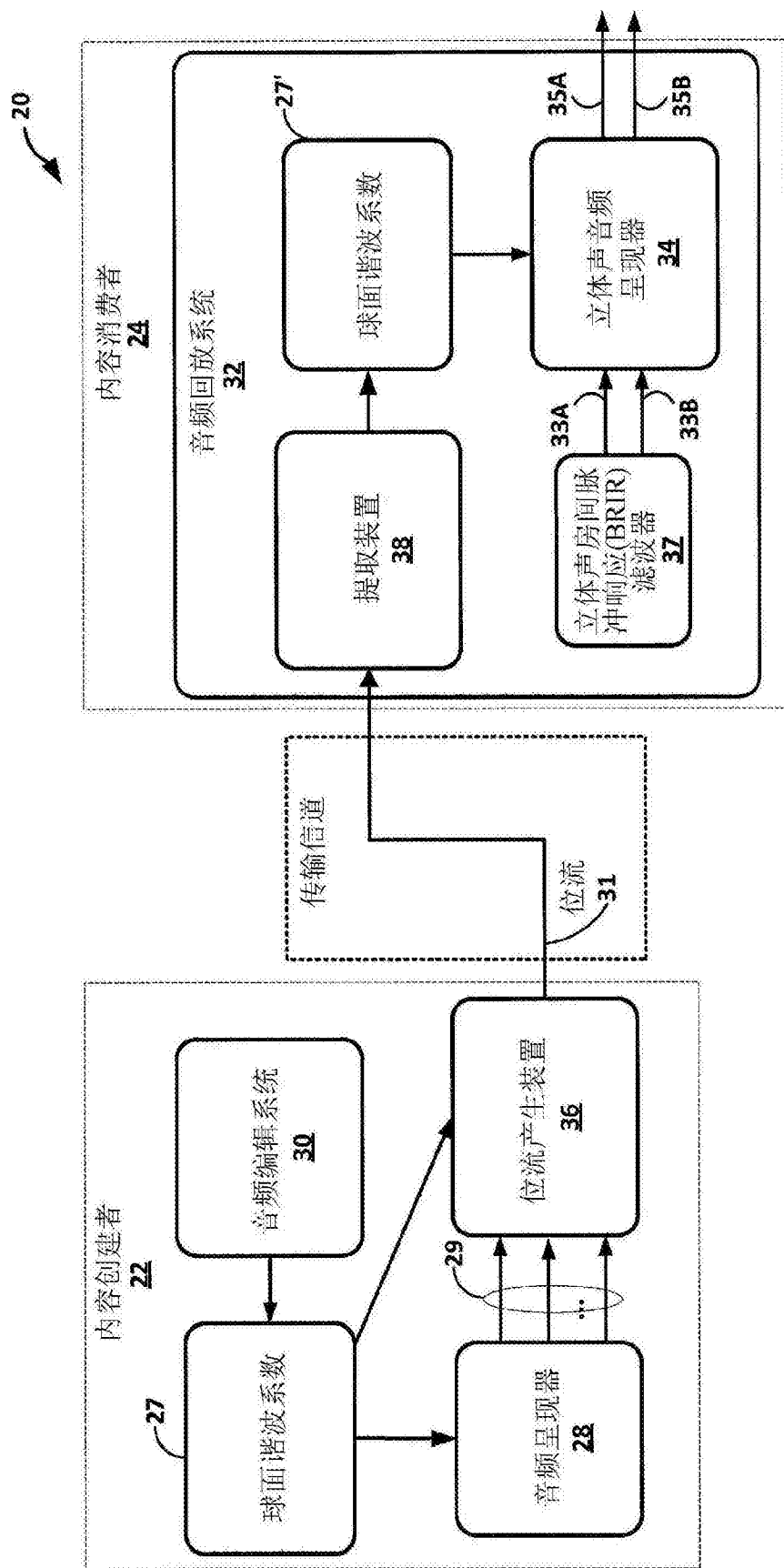


图3

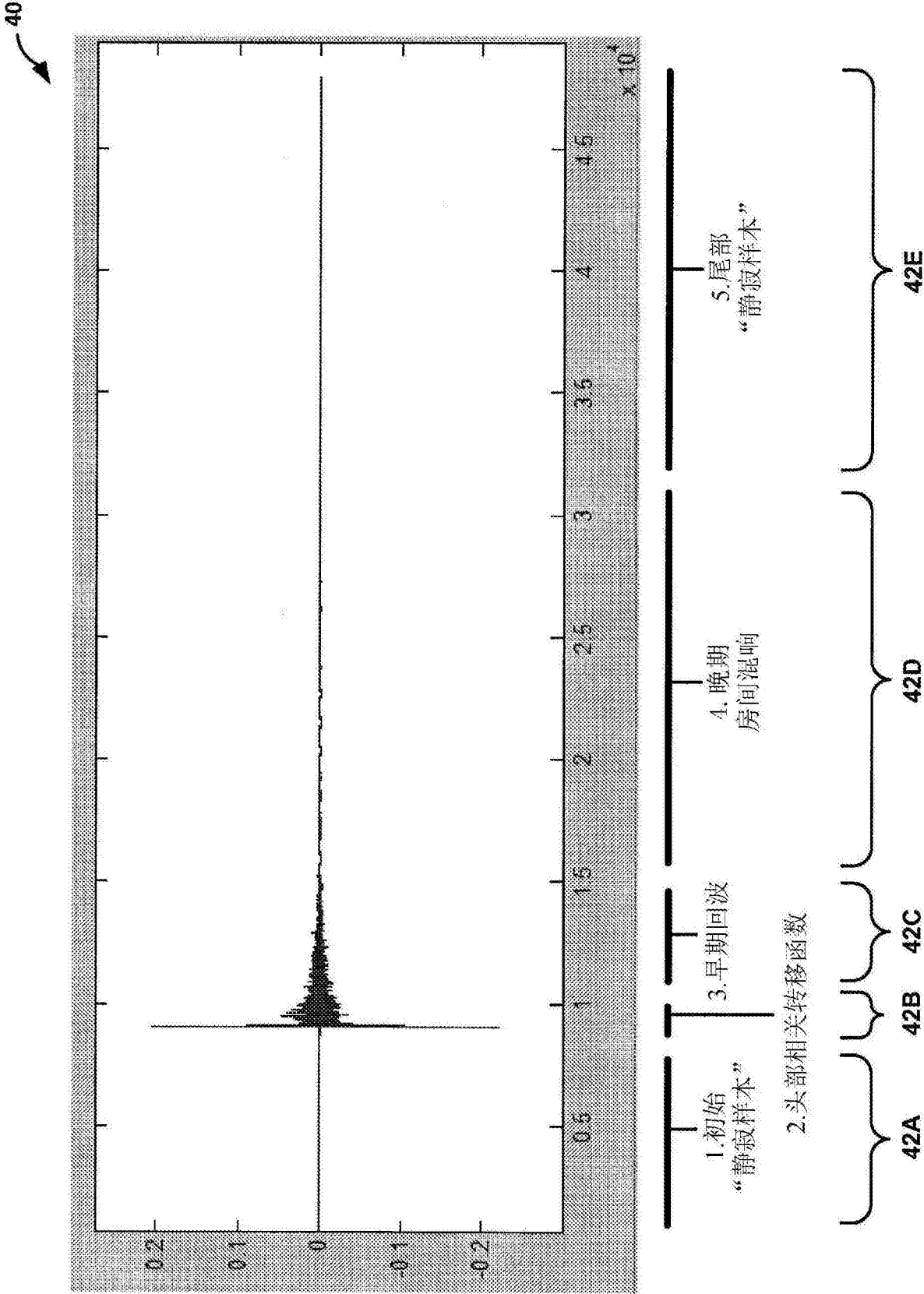


图4

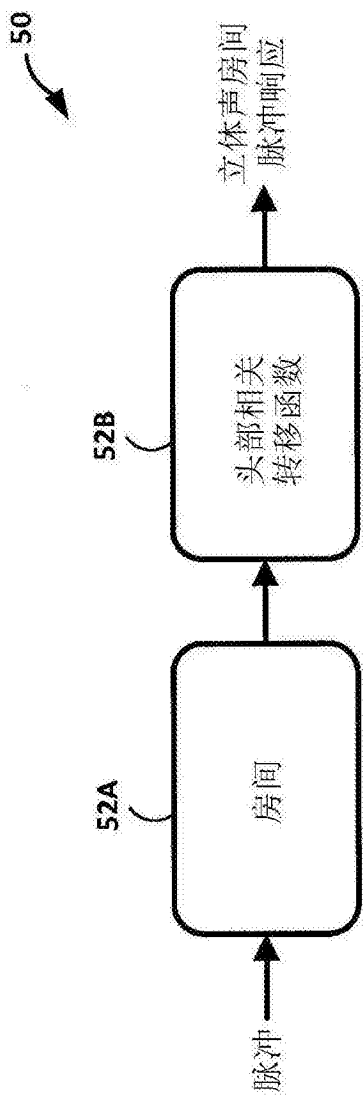


图5

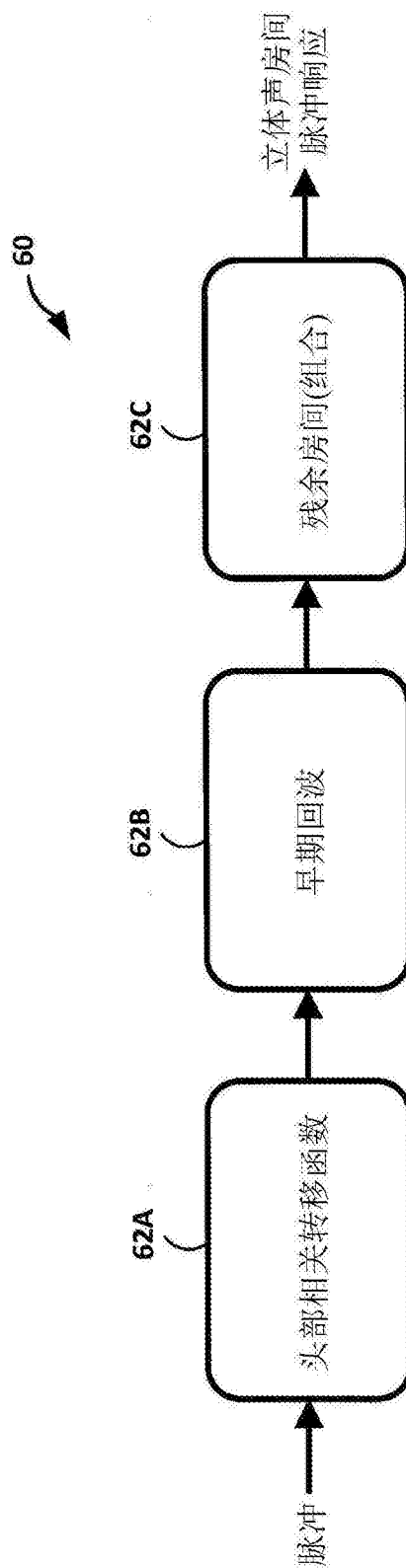


图6

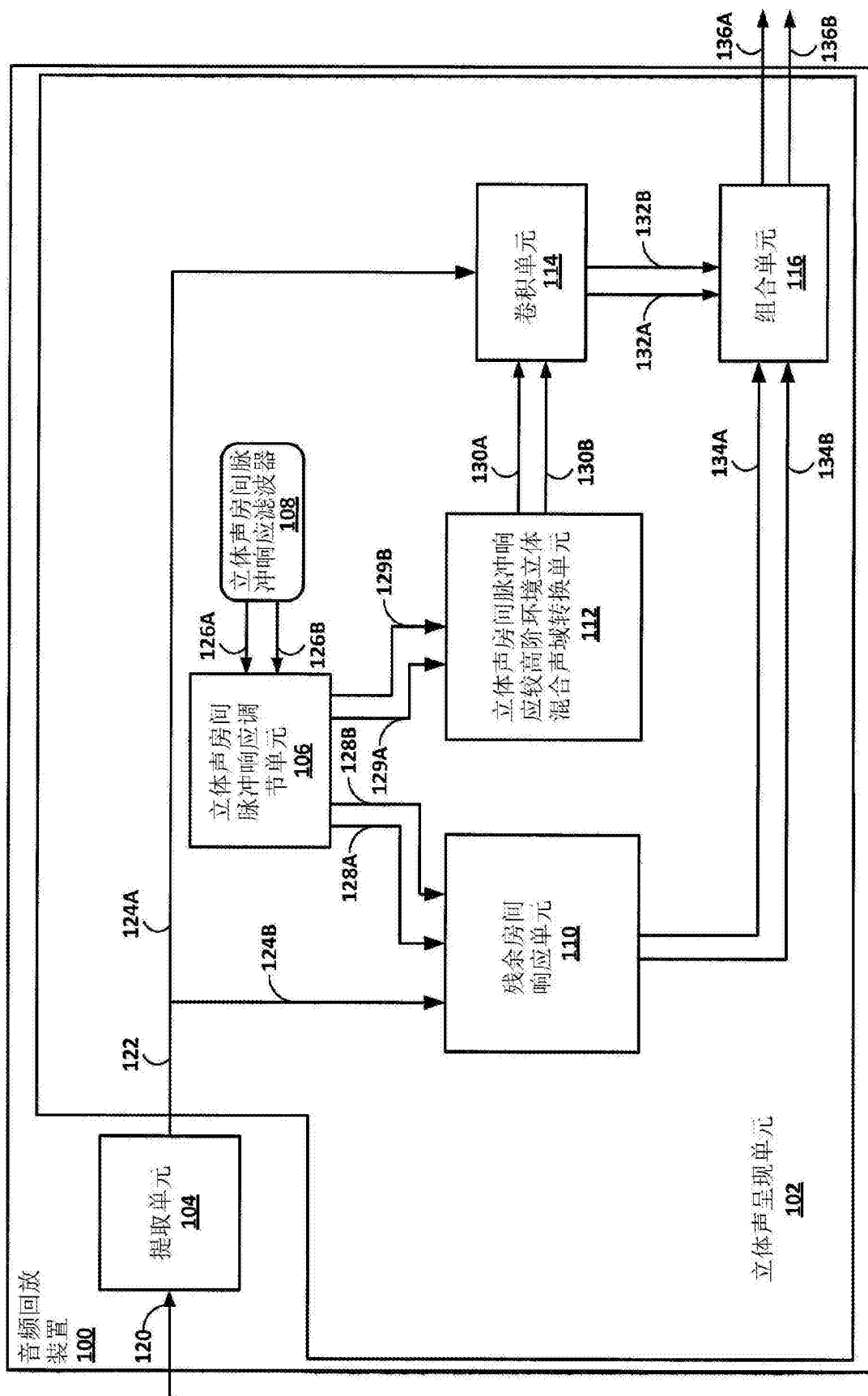


图7

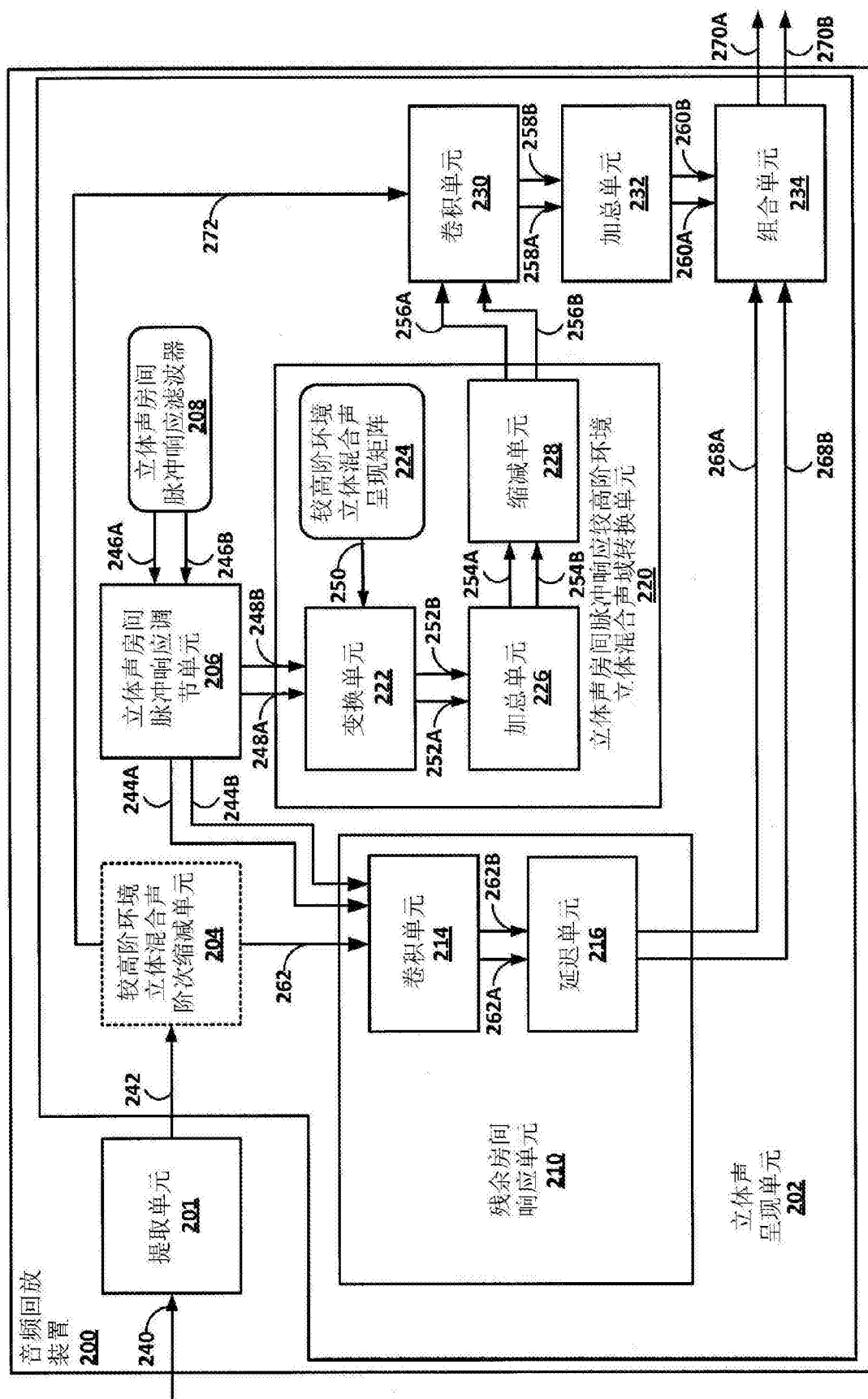


图8

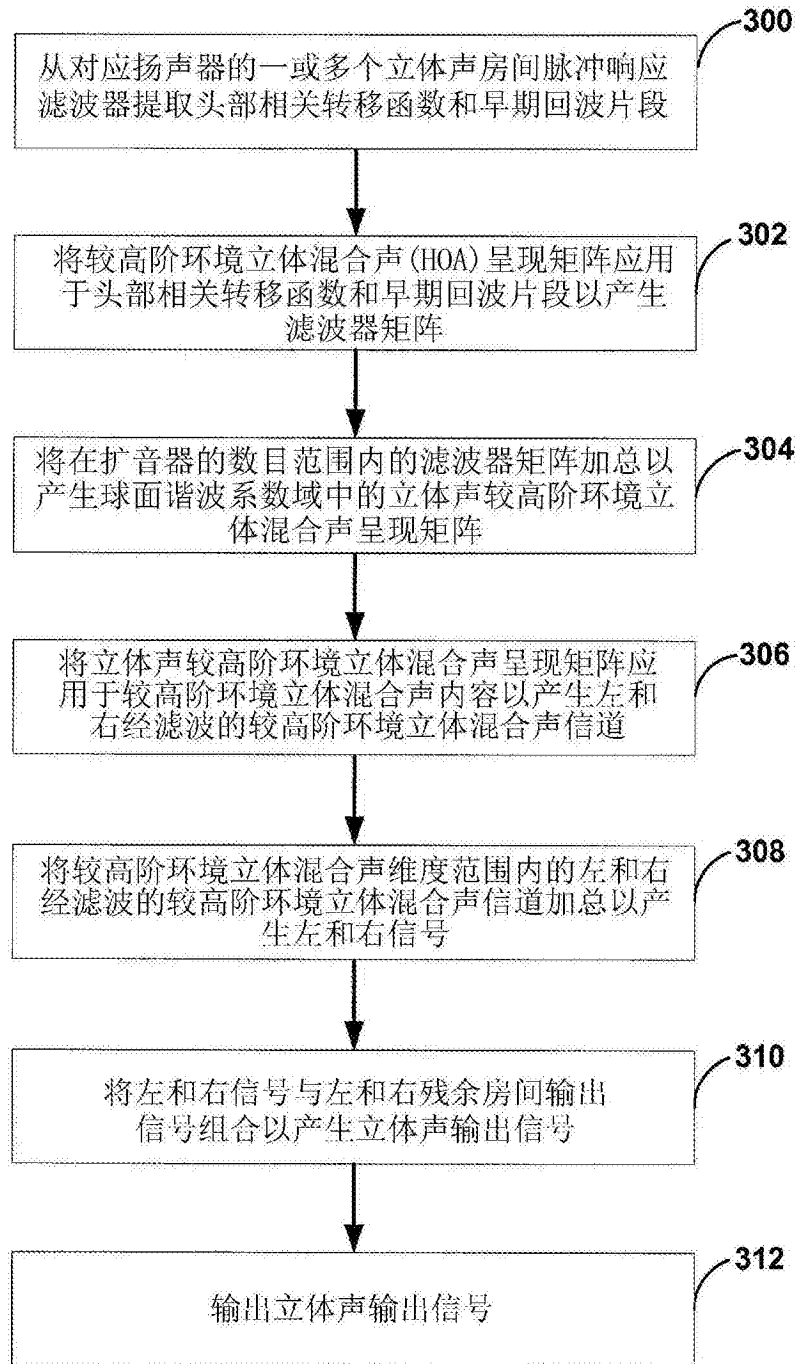


图9

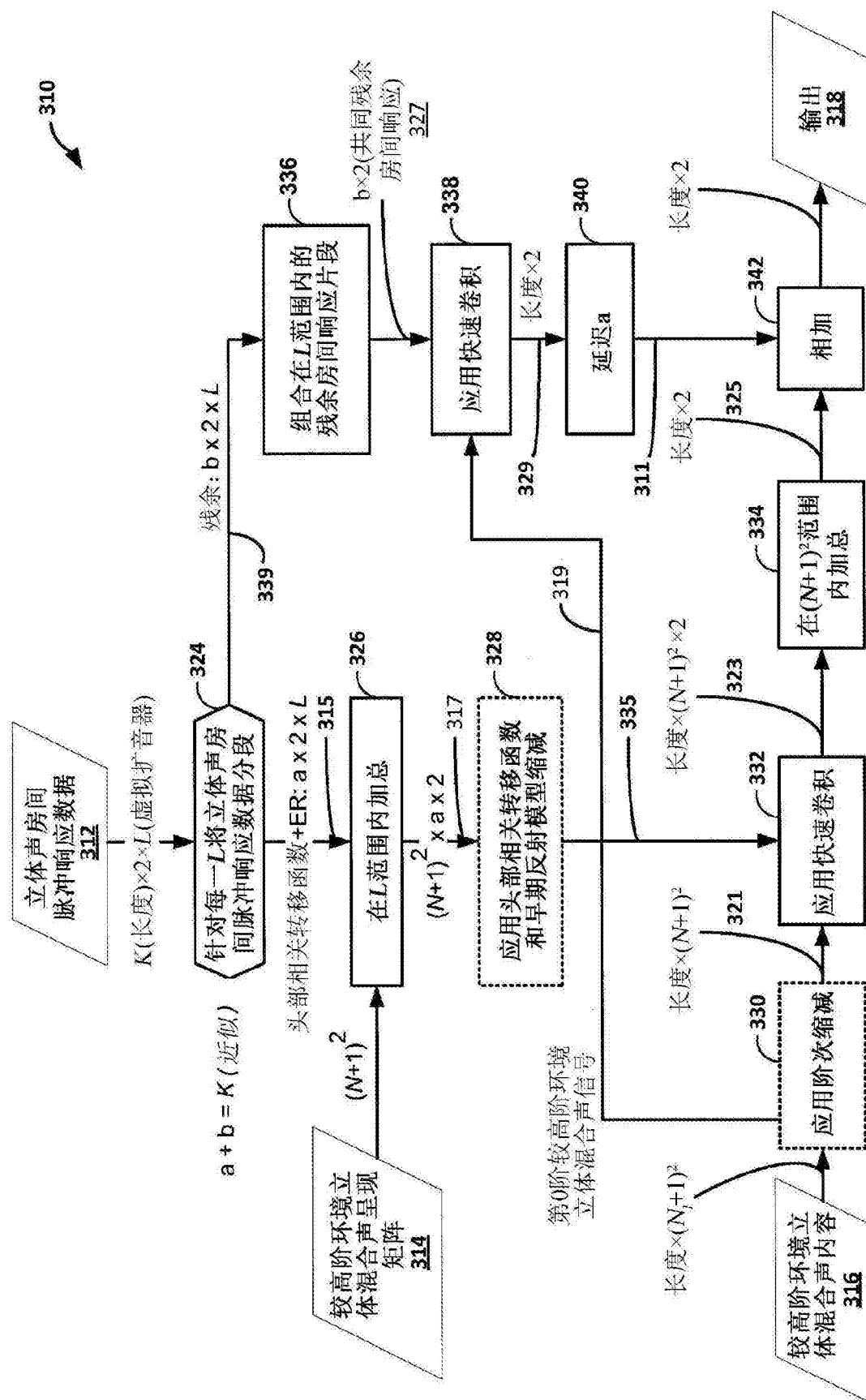


图10A

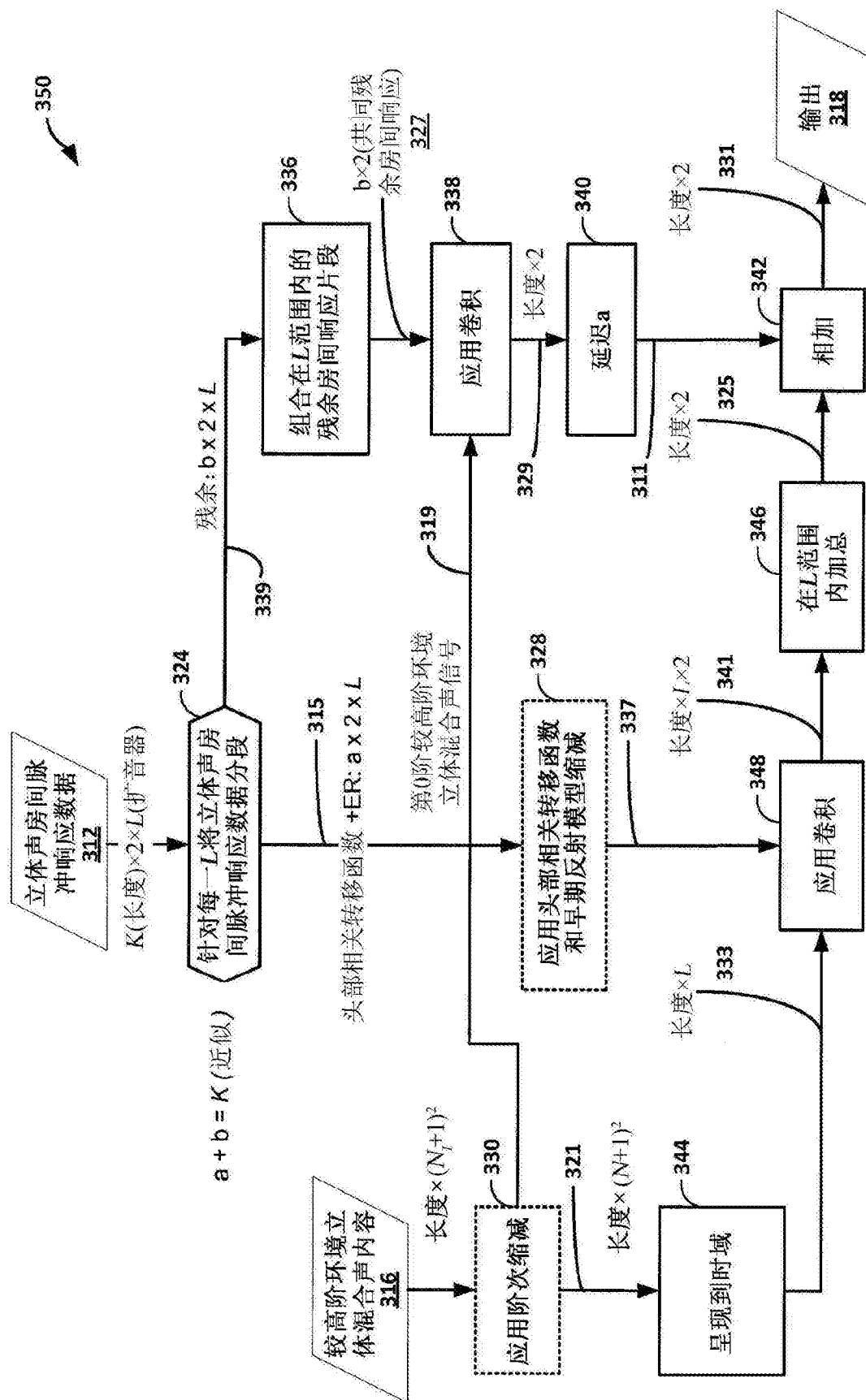


图10B

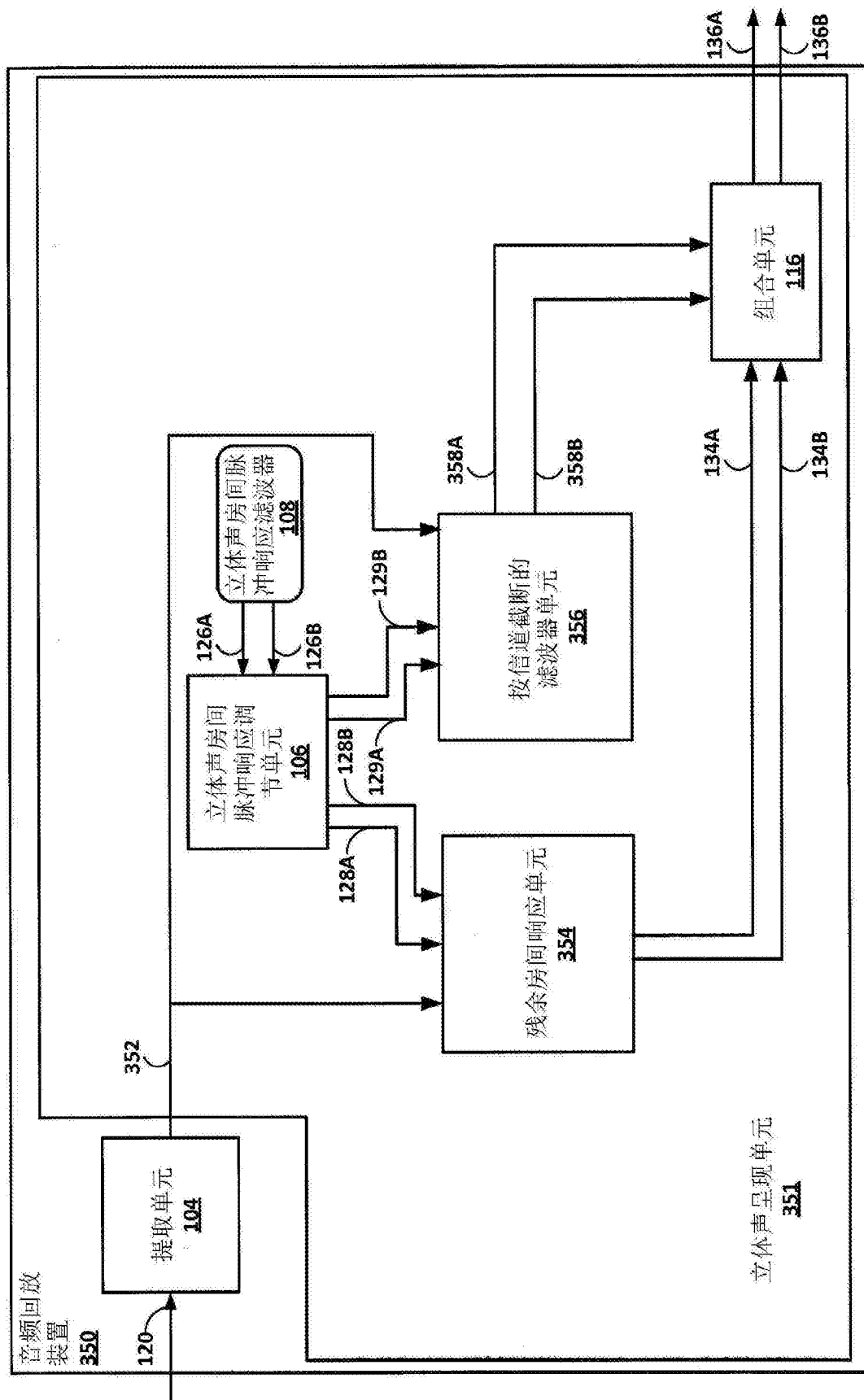


图11

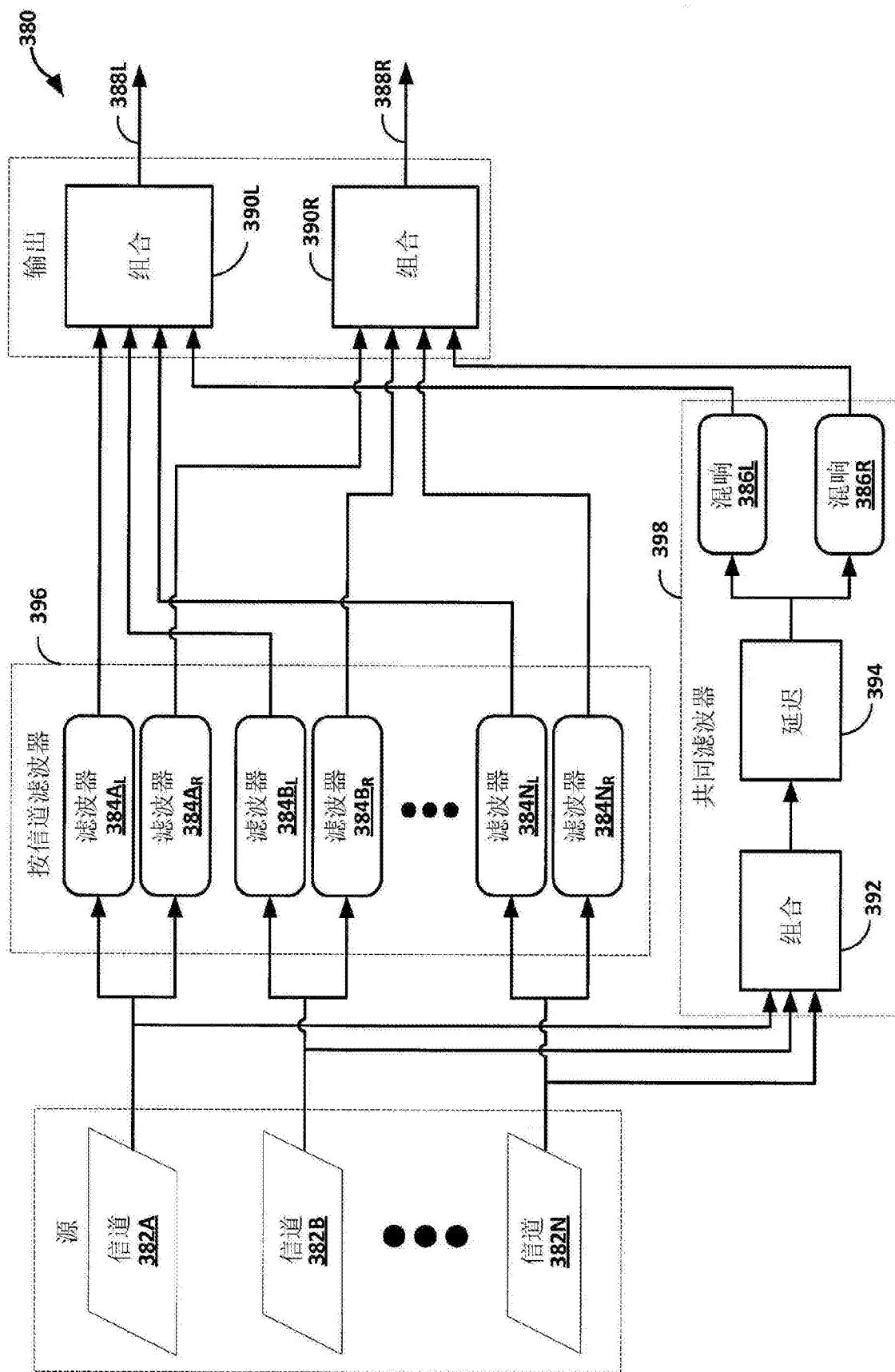


图12

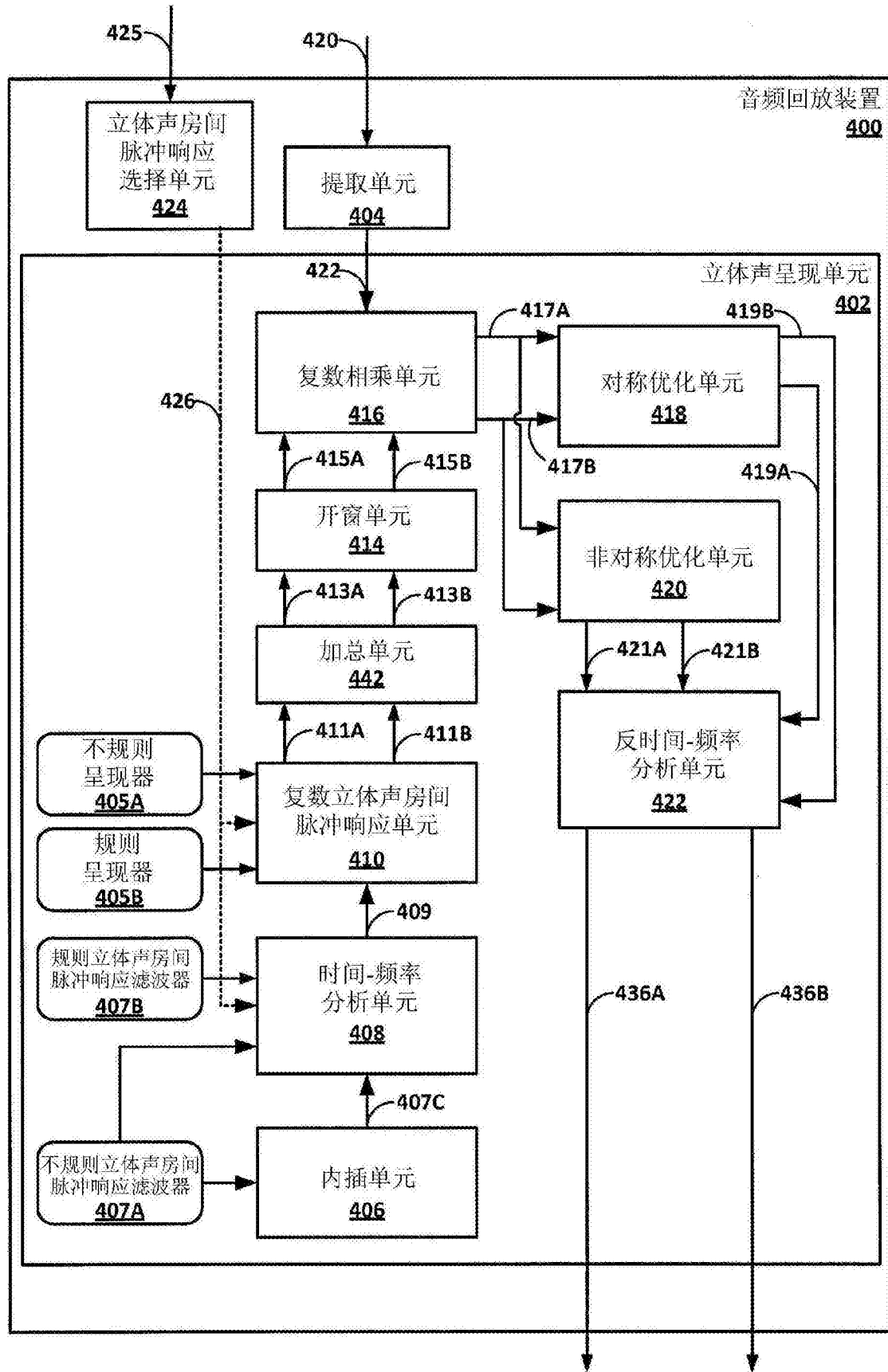


图13

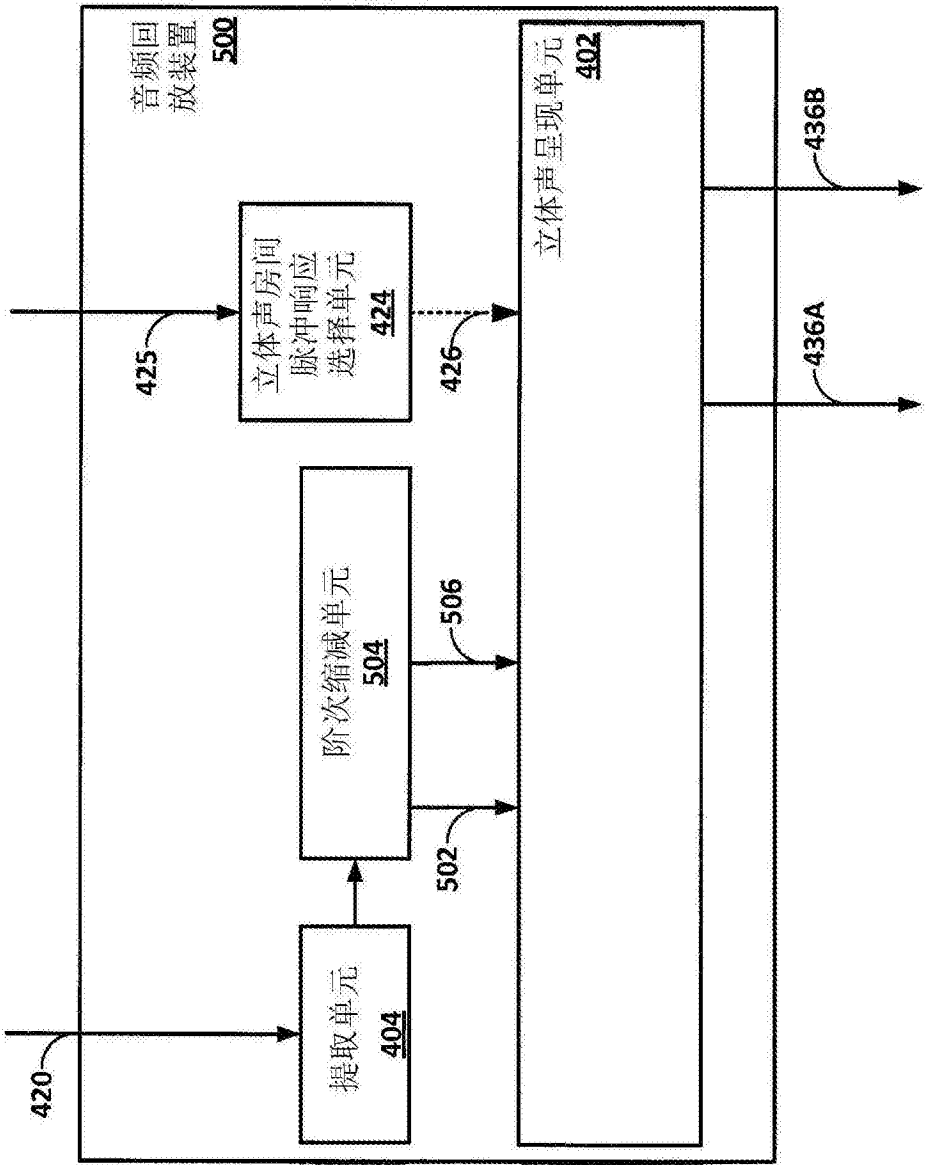


图14

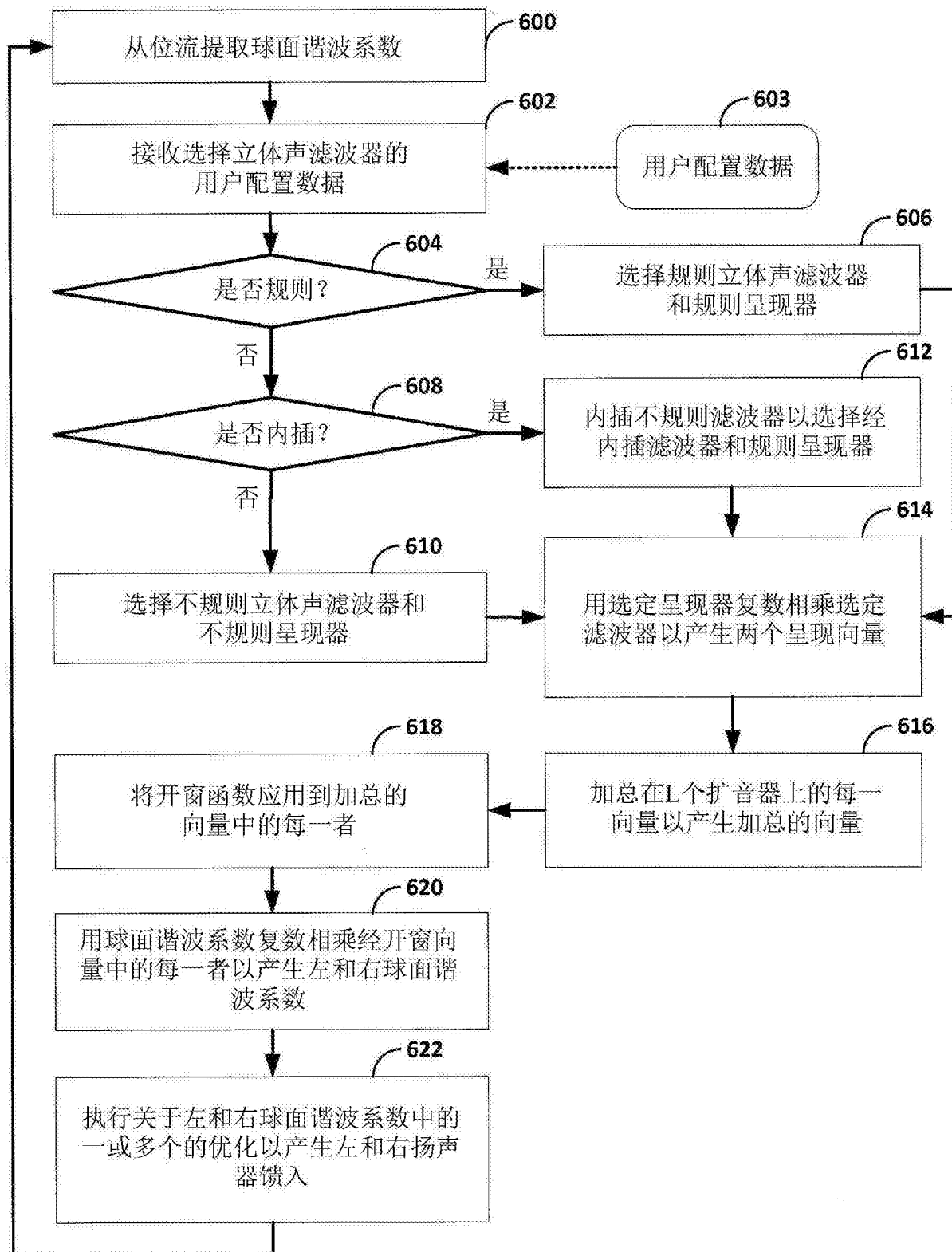


图15

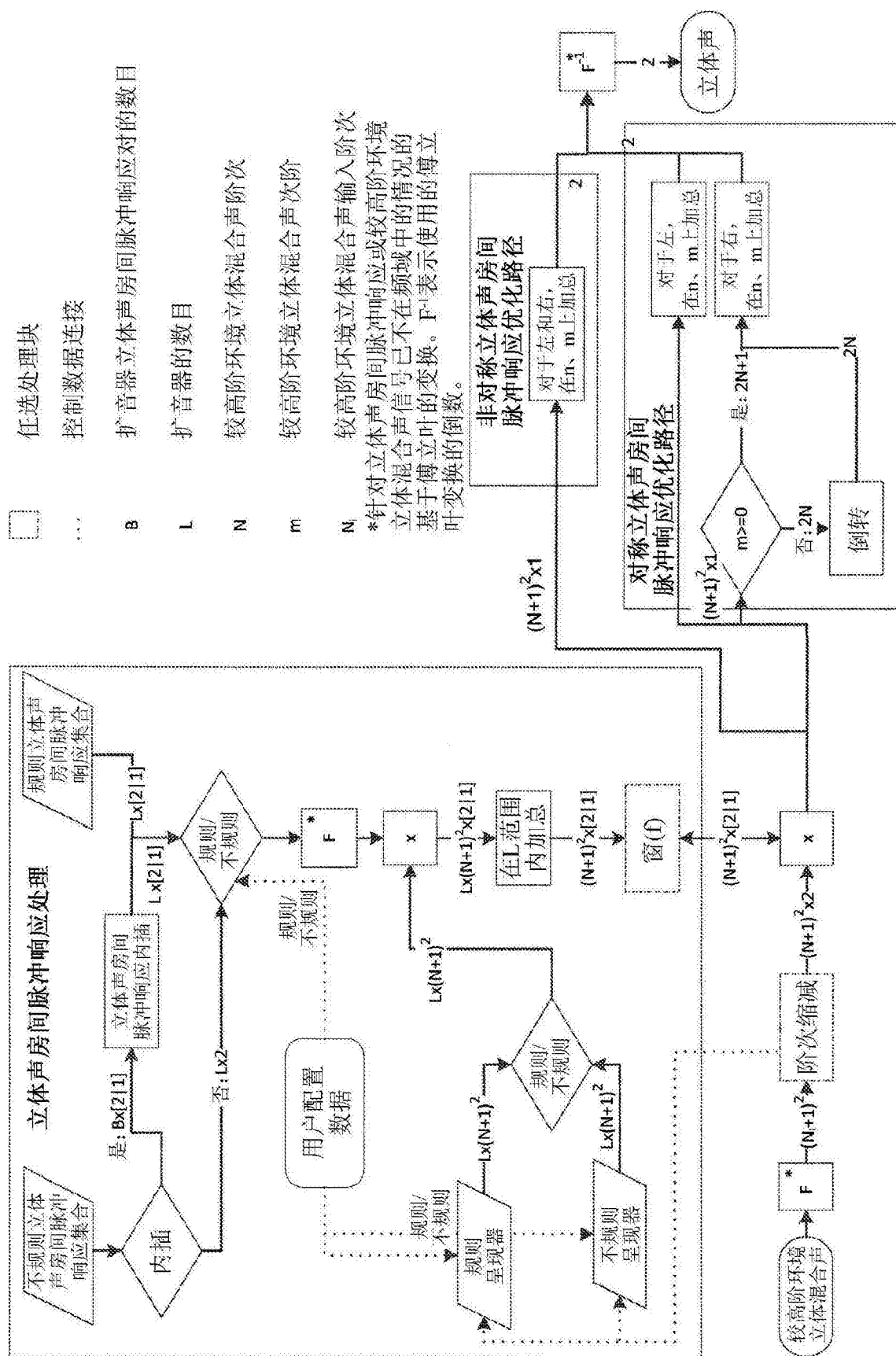


图 16A

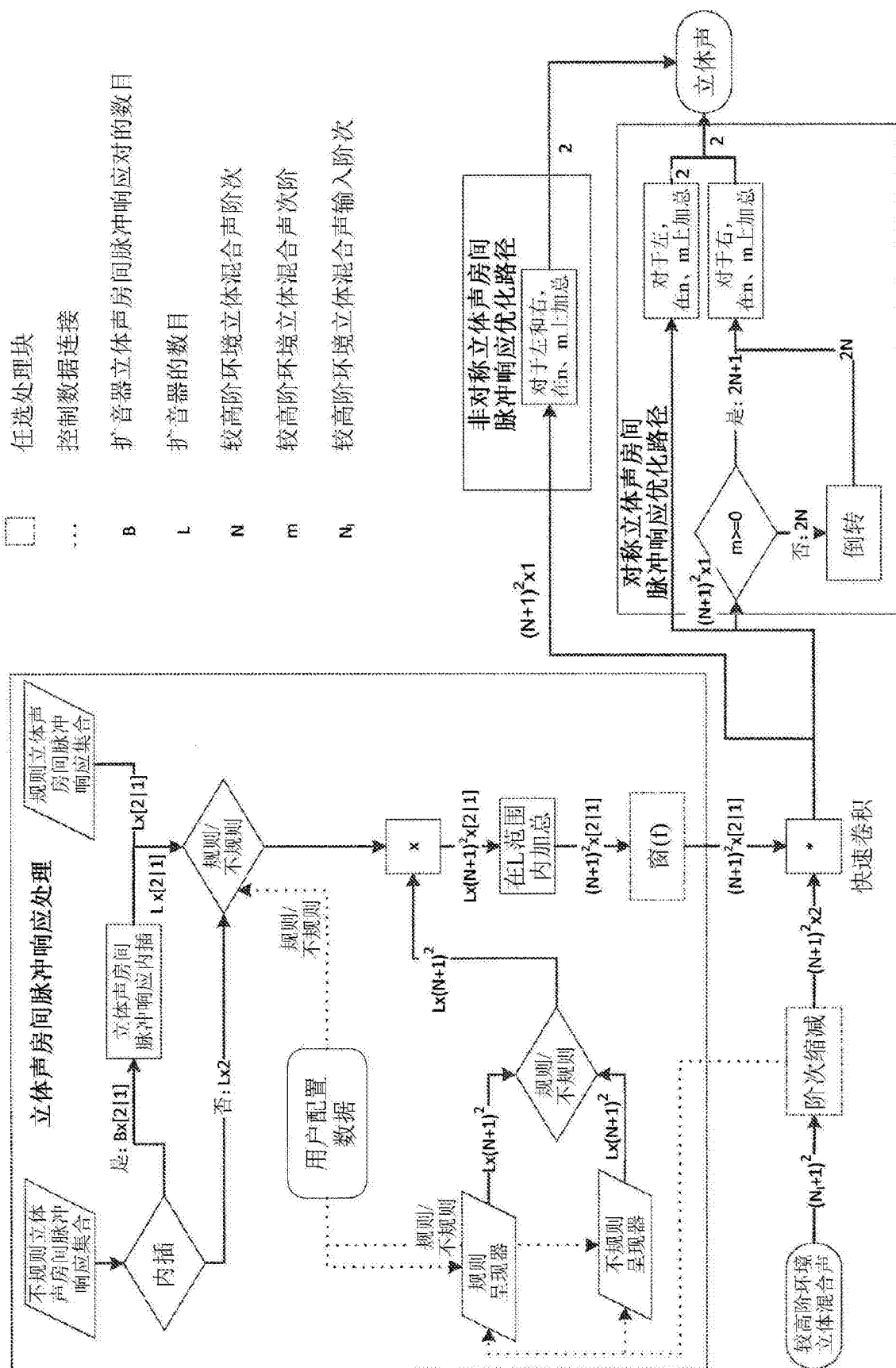


图 16B