



(12) 发明专利

(10) 授权公告号 CN 113316943 B

(45) 授权公告日 2023. 06. 06

(21) 申请号 201980084851.X

(22) 申请日 2019.12.17

(65) 同一申请的已公布的文献号
申请公布号 CN 113316943 A

(43) 申请公布日 2021.08.27

(30) 优先权数据
18214182.0 2018.12.19 EP

(85) PCT国际申请进入国家阶段日
2021.06.18

(86) PCT国际申请的申请数据
PCT/EP2019/085733 2019.12.17

(87) PCT国际申请的公布数据
W02020/127329 EN 2020.06.25

(73) 专利权人 弗劳恩霍夫应用研究促进协会
地址 德国慕尼黑

(72) 发明人 尤根·埃雷拉 伊曼纽尔·哈贝茨
塞巴斯蒂安·施莱赫特
亚历山大·阿达米

(74) 专利代理机构 北京清亦华知识产权代理事务
所(普通合伙) 11201
专利代理师 宋融冰

(51) Int.Cl.
H04S 7/00 (2006.01)

(56) 对比文件
JP 2006503491 A, 2006.01.26
JP 2007003989 A, 2007.01.11
JP H08149600 A, 1996.06.07

审查员 陈艳萍

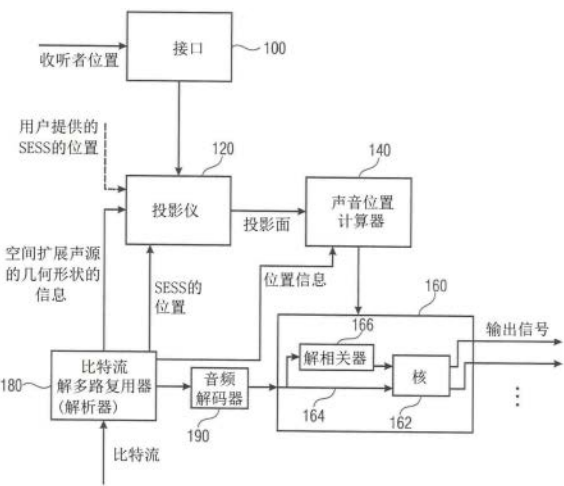
权利要求书6页 说明书18页 附图11页

(54) 发明名称

再现空间扩展声源的设备与方法、或从空间扩展声源生成比特流的设备与方法

(57) 摘要

一种再现空间扩展声源的设备,空间扩展声源在空间中具有限定位置及几何形状,该设备包含:接口(100),用以接收收听者位置;投影仪(120),用于使用收听者位置、有关空间扩展声源的几何形状的信息、及有关空间扩展声源的位置的信息,计算投影至投影面上的与空间扩展声源相关联的二维或三维外壳的投影;声音位置计算器(140),用于使用投影面计算用于空间扩展声源的至少两个声源的位置;以及渲染器(160),用于在该类位置处渲染至少两个声源,以获得具有两个或更多个输出信号的空间扩展声源的再现,其中,渲染器(160)用以对不同位置使用不同声音信号,其中不同声音信号与空间扩展声源相关联。



1. 一种用于再现空间扩展声源的设备,所述空间扩展声源在空间中具有限定位置和几何形状,所述设备包括:

接口(100),用于接收收听者位置;

投影仪(120),用于使用所述收听者位置、有关空间扩展声源的几何形状的信息、以及有关空间扩展声源的位置的信息,计算与所述空间扩展声源相关联的二维或三维外壳至投影面上的投影;

声音位置计算器(140),用于使用所述投影面计算用于所述空间扩展声源的至少两个声源的位置;以及

渲染器(160),用于在所述至少两个声源的位置处渲染所述至少两个声源,以获得具有两个或更多个输出信号的所述空间扩展声源的再现,其中,所述渲染器(160)被配置为对于所述至少两个声源的不同位置使用不同声音信号,其中所述不同声音信号与所述空间扩展声源相关联,

其中检测器被配置为使用追踪系统检测在所述空间中的瞬时收听者位置,或者其中所述接口(100)被配置为使用经由所述接口(100)输入的位置数据。

2. 如权利要求1所述的设备,被配置为接收场景描述以及至少一个声音信号,所述场景描述包括有关空间扩展声源的限定位置的信息以及有关空间扩展声源的限定几何形状的信息,至少一个基础声音信号(301,302)与所述空间扩展声源相关联,

其中所述设备还包括场景解析器(180),所述场景解析器用于解析所述场景描述以取得有关限定位置的信息(341)、有关限定几何形状的信息(331)以及所述至少一个基础声音信号(301,302),或

其中针对所述空间扩展声源,所述场景描述包括至少两个基础声音信号(301,302)以及所述至少两个基础声音信号(301,302)每个基础声音信号关于所述有关空间扩展声源的几何形状的信息(331)的位置信息(321),以及其中所述声音位置计算器(140)被配置为当使用所述投影面计算所述至少两个声源的位置时使用所述至少两个基础声音信号(301,302)的位置信息。

3. 如权利要求1所述的设备,

其中所述投影仪(120)被配置为使用所述有关空间扩展声源的几何形状的信息(331)计算所述空间扩展声源的外壳,并使用所述收听者位置将所述外壳沿朝向收听者的方向投影以获得所述二维或三维外壳至所述投影面上的投影,或

其中所述投影仪(120)被配置为将由所述有关空间扩展声源的几何形状的信息(331)所限定的所述空间扩展声源的几何形状沿朝向所述收听者位置的方向投影,并计算投影几何形状的外壳以获得所述二维或三维外壳至所述投影面上的投影。

4. 如权利要求1所述的设备,

其中所述声音位置计算器(140)被配置为从外壳投影数据以及所述收听者位置计算所述空间中的所述至少两个声源的位置。

5. 如权利要求1所述的设备,

其中所述声音位置计算器(140)被配置为计算所述位置,以使得所述至少两个声源为多个周边声源并位于所述投影面上,或

其中所述声音位置计算器(140)被配置为计算,以使得多个周边声源中的周边声源的

位置关于收听者位于所述投影面的右侧、和/或关于收听者位于所述投影面的左侧、和/或关于收听者位于所述投影面的顶部、和/或关于收听者位于所述投影面的底部。

6. 如权利要求1所述的设备，

其中所述渲染器 (160) 被配置为使用以下渲染所述至少两个声源：

依据所述至少两个声源的位置的平移操作，以获得用于预定义扬声器设置的扬声器信号，或

依据所述至少两个声源的位置使用头相关传递函数的双耳渲染操作，以获得耳机信号。

7. 如权利要求1所述的设备，

其中第一数量的基础声音信号 (301, 302) 与所述空间扩展声源相关联，所述第一数量为一或大于一，其中所述第一数量的基础声音信号 (301, 302) 与相同的空间扩展声源相关，

其中所述声音位置计算器 (140) 确定用于渲染所述空间扩展声源的所述第二数量的声源，所述第二数量大于一，以及

其中所述渲染器 (160) 包括一个或多个解相关器 (166)，所述一个或多个解相关器用于在所述第二数量大于所述第一数量时，从第一数量的一个或多个基础声音信号 (164, 301, 302) 生成解相关信号。

8. 如权利要求1所述的设备，

其中所述接口 (100) 被配置为接收收听者在所述空间中的时变位置，

其中所述投影仪 (120) 被配置为计算所述空间中的时变投影，

其中所述声音位置计算器 (140) 被配置为计算所述空间中的声源的时变数量或所述至少两个声源的时变位置，以及

其中所述渲染器 (160) 被配置为在所述空间中渲染时变数量的声源或在所述空间中的时变位置处渲染所述至少两个声源。

9. 如权利要求1所述的设备，

其中所述接口 (100) 被配置为接收六个自由度的收听者位置，以及

其中所述投影仪 (120) 被配置为依据所述六个自由度计算所述投影。

10. 如权利要求1所述的设备，其中所述投影仪 (120) 被配置为：

将所述投影计算为图片平面，如垂直于收听者的视线的平面，或

将所述投影计算为围绕收听者的头部的球面，或

将所述投影计算为投影面，所述投影面位于距收听者的头部的中心预定距离处，或

从相对于收听者的头部的视角的方位角和仰角计算空间扩展声源的所述外壳的投影，所述方位角和所述仰角从球坐标得出，所述外壳为凸壳。

11. 如权利要求1所述的设备，

其中所述声音位置计算器 (140) 被配置为计算所述至少两个声源的位置，以使得所述位置绕所述外壳的投影均匀分布，或者使得所述位置置于外壳投影的极点或边缘点，或者使得所述位置位于所述外壳的投影的水平或垂直极点或边缘点。

12. 如权利要求1所述的设备，

其中所述声音位置计算器 (140) 被配置为除周边声源的位置以外还确定辅助声源的位置，所述辅助声源的位置关于收听者位于所述外壳的投影之上、之前、之后或之内。

13. 如权利要求1所述的设备，

其中所述投影仪 (120) 被配置为在如水平方向和垂直方向的不同方向上以变量或预定量或以不同变量或不同预定量例如朝向所述外壳或所述投影的重心额外地收缩所述外壳的投影。

14. 如权利要求1所述的设备，其中所述声音位置计算器 (140) 被配置为计算，以使得至少一个额外辅助声源位于所述投影面上、关于所述收听者位置在左周边声源和右周边声源之间，或

其中所述声音位置计算器 (140) 被配置为计算，以使得至少一个额外辅助声源位于所述投影面上、关于所述收听者位置在左周边声源和右周边声源之间，其中在所述左周边声源和所述右周边声源之间的中间放置单个额外辅助源，或者在所述左周边声源和所述右周边声源之间等距地放置两个或更多个额外辅助源。

15. 如权利要求1所述的设备，

其中所述声音位置计算器 (140) 被配置为在经由所述接口接收收听者绕所述空间扩展声源的圆周运动的情况下，或者在经由所述接口接收所述空间扩展声源关于固定收听者的旋转的情况下，优选地绕所述投影的重心执行所述空间扩展声源的至少两个声源的位置的旋转。

16. 如权利要求1所述的设备，

其中所述渲染器 (160) 被配置为，针对每个声源，依据收听者和所述声源之间的距离接收张开角，并且依据所述张开角渲染所述声源。

17. 如权利要求1所述的设备，

其中所述渲染器 (160) 被配置为接收每个声源的距离信息，以及

其中所述渲染器 (160) 被配置为依据所述距离渲染所述声源，以使得相较于放置地不太靠近收听者且具有相同音量的声源，放置地更靠近所述收听者的声源被渲染为具有较大音量。

18. 如权利要求1所述的设备，其中所述声音位置计算器 (140) 被配置为：

为每个声源确定距离，所述距离等于所述空间扩展声源关于收听者的距离，或

通过投影上的声源的位置至所述空间扩展声源的几何形状上的反向投影，确定每个声源的距离；

其中所述渲染器 (160) 被配置为使用有关所述距离的信息渲染所述至少两个声源。

19. 如权利要求1所述的设备，

其中有关几何形状的信息 (331) 被限定为一维直线或曲线、二维面、或三维体，其中所述二维面例如为椭圆形、矩形、多边形，或多边形组，所述三维体例如为椭球体、立方体或多面体，和/或

其中所述信息被限定为参数化描述或多边形描述或多边形描述的参数化表示。

20. 如权利要求1所述的设备，

其中所述声音位置计算器 (140) 被配置为依据收听者至所述空间扩展声源的距离确定声源的数量，其中相较于针对所述收听者与所述空间扩展声源之间的较大距离的较少数量，针对较小距离，声源的数量较大。

21. 如权利要求1所述的设备，被配置为接收有关空间扩展声源所引入的扩展的信息，

以及

其中所述投影仪 (120) 被配置为使用有关扩展的信息对所述外壳或所述投影施加收缩操作,以至少部分地补偿所述扩展。

22. 如权利要求1所述的设备,

其中所述渲染器 (160) 被配置为在所述至少两个声源的位置在限定容差范围内彼此相同的情况下,通过例如使用吉文斯旋转来组合与所述空间扩展声源相关联的基础信号以获得经旋转的基础信号并在所述至少两个声源的位置处渲染所述经旋转的基础信号,来渲染所述声源。

23. 如权利要求1所述的设备,

其中所述渲染器 (160) 被配置为在依据位置或方向相关特性生成所述至少两个声源时,执行预处理或后处理。

24. 如权利要求1所述的设备,

其中所述空间扩展声源具有所述空间扩展声源是球状、以及椭球状、线状、立方体状或钢琴形状空间扩展声源的信息作为有关几何形状的信息 (331)。

25. 如权利要求1所述的设备,被配置为:

接收表示所述空间扩展声源的压缩描述的比特流,所述比特流包括比特流元素 (311),所述比特流元素指示包括在由所述设备接收的所述比特流或经编码的音频信号中的用于所述空间扩展声源的不同声音信号的第一数量,所述第一数量为一或大于一,

读取所述比特流元素 (311),并取得包括在所述比特流或所述经编码的音频信号中的用于所述空间扩展声源的不同声音信号的所述第一数量,以及

其中,所述声音位置计算器 (140) 确定用于渲染所述空间扩展声源的第二数量的声源,所述第二数量大于1,以及

其中,所述渲染器 (160) 被配置为依据从所述比特流中提取的所述第一数量生成第三数量的一个或多个解相关信号,所述第三数量从所述第二数量与所述第一数量之间的差得出。

26. 一种用于生成比特流的设备,所述比特流表示空间扩展声源的压缩描述,所述设备包括:

声音提供器 (200),用于提供用于所述空间扩展声源的至少两个不同声音信号 (301, 302);

几何形状提供器 (220),用于计算有关空间扩展声源的几何形状的信息 (331,341);以及

输出数据形成器 (240),用于生成表示压缩描述的比特流,所述比特流包括所述至少两个不同声音信号 (301,302)、有关几何形状的信息 (331,341)、以及所述至少两个不同声音信号的每个声音信号的各自位置信息 (321),其中所述各自位置信息 (321) 指示对应声音信号关于有关所述空间扩展声源的几何形状的信息 (331) 的位置,

其中所述声音提供器 (200) 被配置为:

在单个麦克风位置或方向或多个麦克风位置或方向上执行自然声源的记录,或通过一个或多个解相关滤波器从单个基础信号或多个基础信号中得出声音信号。

27. 如权利要求26所述的设备,其中所述有关几何形状的信息 (331,341) 包括位置信息

(341),所述位置信息指示所述空间扩展声源在空间中的位置。

28.如权利要求26所述的设备,

其中所述声音提供器(200)被配置为使用音频信号编码器(260)对至少两个声音信号进行比特率压缩,以及

其中所述输出数据形成器(240)被配置为将经比特率压缩的至少两个不同声音信号(301,302)用于所述空间扩展声源。

29.如权利要求26所述的设备,其中所述几何形状提供器(220)被配置为从所述空间扩展声源的几何形状得出参数化描述或多边形描述或多边形描述的参数化表示,并且其中所述输出数据形成器(240)被配置为将所述参数化描述或所述多边形描述或所述多边形描述的参数化表示引入所述比特流中作为所述有关几何形状的信息(331,341)。

30.如权利要求26所述的设备,其中所述输出数据形成器(240)被配置为将比特流元素(311)引入所述比特流中,所述比特流元素指示包括在所述比特流或与所述比特流相关联的经编码的音频信号中的用于所述空间扩展声源的至少两个不同声音信号(301,302)的数量,所述数量为二或大于二。

31.一种用于再现空间扩展声源的方法,所述空间扩展声源在空间中具有限定位置及几何形状,所述方法包括:

接收收听者位置;

使用所述收听者位置、有关空间扩展声源的几何形状的信息(331)、以及有关空间扩展声源的位置的信息(341),计算与所述空间扩展声源相关联的二维或三维外壳至投影面上的投影;

使用所述投影面计算用于所述空间扩展声源的至少两个声源的位置;以及

在所述至少两个声源的位置处渲染所述至少两个声源,以获得具有两个或更多个输出信号的所述空间扩展声源的再现,其中,所述渲染包括对于所述至少两个声源的不同位置使用不同声音信号,其中所述不同声音信号与所述空间扩展声源相关联,

其中使用追踪系统检测在所述空间中的瞬时收听者位置,或者其中所述接收收听者位置包括使用经由接口(100)输入的位置数据。

32.一种生成比特流的方法,所述比特流表示空间扩展声源的压缩描述,所述方法包括:

提供用于所述空间扩展声源的至少两个不同声音信号(301,302);

提供有关空间扩展声源的几何形状的信息(331,341);以及

生成表示压缩描述的比特流,所述比特流包括所述至少两个不同声音信号(301,302)、所述有关空间扩展声源的几何形状的信息(331,341)、以及所述至少两个不同声音信号的每个声音信号的各自位置信息(321),其中所述各自位置信息(321)指示对应声音信号关于有关所述空间扩展声源的几何形状的信息(331)的位置,

其中所述提供用于所述空间扩展声源的至少两个不同声音信号包括在单个麦克风位置或方向或多个麦克风位置或方向上执行自然声源的记录,或者通过一个或多个解相关滤波器从单个基础信号或多个基础信号中得出声音信号。

33.如权利要求32所述的方法,其中所述有关空间扩展声源的几何形状的信息(331,341)包括所述空间扩展声源在空间中的位置信息(341)。

34. 如权利要求32所述的方法,其中所述生成所述比特流包括将比特流元素(311)引入所述比特流中,所述比特流元素指示包括在所述比特流或与所述比特流相关联的经编码的音频信号中的用于所述空间扩展声源的所述至少两个不同声音信号(301,302)的数量,所述数量为二或大于二。

35. 一种存储介质,存储有计算机程序,所述计算机程序包括指令,当所述计算机程序由计算机或处理器执行时,使得所述计算机进行如权利要求31所述的方法。

36. 一种存储介质,存储有计算机程序,所述计算机程序包括指令,当所述计算机程序由计算机或处理器执行时,使得所述计算机进行如权利要求32所述的方法。

再现空间扩展声源的设备与方法、或从空间扩展声源生成比特流的设备与方法

[0001] 本发明涉及音频信号处理,以及特别涉及空间扩展声源的编码、解码或再现。

[0002] 长期以来,一直在研究在多个扬声器或耳机上的声源的再现。在此类设置上再现声源的最简单方法是将它们渲染为点源,即非常小(理想为无限小)的声源。但是,这种理论概念几乎无法以现实的方式建立现有物理声源的模型。例如,一架三角钢琴有很大的振动木质封闭体,其内部有许多空间分布的琴弦,并且因此在听觉感知上显得比点源要大得多(尤其是当收听者(和麦克风)靠近三角钢琴时)。许多现实世界中的声源具有相当大的尺寸(“空间范围”),例如乐器、机器、管弦乐队、合唱团或环境声音(瀑布声)。

[0003] 此类声源的正确/逼真再现已成为许多声音再现方法的目标,无论是使用耳机的双耳方式(即使用所谓的头相关传递函数(HRTF)或双耳房间脉冲响应(BRIR)),还是使用扬声器设置(从两个扬声器(“立体声”)到在水平面上排列的许多扬声器(“环绕声”)以及在所有三个维度上围绕收听者的许多扬声器(“3D音频”))的传统方式均是如此。

[0004] 本发明的目的是提供一种用于编码或再现具有可能复杂的几何形状的空间扩展声源的概念。

[0005] 2D源宽度

[0006] 本节描述关于在从收听者的视角所面对的2D表面上渲染扩展声源的方法,例如在仰角为零的某个方位角范围(如传统立体声/环绕声中的情况)或某些方位角及仰角范围(如3D音频或具有用户运动的3个自由度(3DoF)的虚拟现实中的情况,即头部在俯仰/偏航/翻滚轴上旋转)中。

[0007] 可以通过减小参与声道信号的相关性来实现在两个或更多个扬声器之间平移的音频对象的视在宽度的增大(生成所谓的幻像或幻像源)(Blauert,2001,S.241-257)。随着相关性的减小,幻像源的扩展增大,直到对于接近零的相关性值(且张开角不太宽),其覆盖扬声器之间的整个范围。

[0008] 通过得出并应用适当的解相关滤波器,获得源信号的解相关版本,Lauridsen(Lauridsen,1954)提出对源信号本身加上/减去源信号的时间延迟和缩放的版本,以获得信号的两个解相关版本。另外,Kendall(Kendall,1995)提出了更复杂的方法,他基于随机数序列的组合迭代地得出成对的解相关全通滤波器。Faller等人在研究(Baumgarte、Faller,2003)(Faller、Baumgarte,2003)中提出合适的解相关滤波器(扩散器)。此外,Zotter等人(Zotter、Frank,2013)也得出滤波器对,其中使用频率相关相位或幅度差来实现幻像源的加宽。又有,Alary、Politis、Välimäki等人(Alary、Politis、Välimäki,2017)提出了基于天鹅绒噪声的解相关滤波器,这被Schlecht、Alary、Välimäki、Habets等人(Schlecht、Alary、Välimäki、Habets,2018)进一步优化。

[0009] 除了减少幻像源的对应声道信号的相关性之外,还可以通过增加归因于音频对象的幻像源的数量来增大源宽度。在Pulkki的研究(Pulkki,1999)中,通过将同一源信号平移到(略微)不同方向来控制源宽度。最初提出该方法是为了在VBAP平移(Pulkki,1997)源信号在声音场景中移动时,稳定所感知的VBAP平移源信号的幻像源扩展。这因为取决于源的

方向而是有利的,渲染源由两个或更多个扬声器再现,这可能导致所感知的源宽度的非期望改变。

[0010] 虚拟世界DirAC (Pulkki、Laitinen、Erkut,2009) 是对传统定向音频编码 (DirAC) (Pulkki,2007) 方法的扩展以用于虚拟世界中的声音合成。为了渲染空间范围,将源的定向声音分量在源的原始方向周围的一定范围内随机平移,其中平移方向随时间和频率而变化。

[0011] 在Pihlajamäki、Santala、Pulkki的研究 (Pihlajamäki、Santala、Pulkki,2014) 中也采用类似方法,其中空间范围是通过将源信号的频带随机分布到不同的空间方向来达成。这是一种旨在从所有方向同等地产生空间分布且包络的声音而非控制范围的精确程度的方法。

[0012] Verron等人并非使用平移的相关信号,而是通过合成源信号的多个非相干版本、将它们均匀地分布在收听者周围的圆上、并在它们之间进行混合 (Verron、Aramaki、Kronland-Martinet、Pallone,2010) 来达成源的空间范围。同时激活的源的数量和增益决定了加宽效果的强度。此方法被实现为对环境声音合成器的空间扩展。

[0013] 3D源宽度

[0014] 本节描述关于在3D空间中渲染扩展声源的方法,即以具有6个自由度 (6DoF) 的虚拟现实所需的体积法。这意味着用户运动的6个自由度,即头部在俯仰/偏航/翻滚轴上的旋转再加上3个平移运动方向x/y/z。

[0015] Potard等人通过研究源形状的感知,将源范围的概念扩展为源的一维参数,即在两个扬声器之间的宽度 (Potard,2003)。他们通过将(时变)解相关技术应用于原始源信号,然后将非相干源放置在不同的空间位置并借此产生了三维范围 (Potard、Burnett,2004),从而生成了多个非相干点源。

[0016] 在MPEG-4高级音频BIFS (Schmidt、Schröder,2004年) 中,可以用若干个均匀分布且解相关的声源填充体积对象/形状(壳状、盒状、椭球体和圆柱体),以引致三维源范围。

[0017] 为了使用高保真度立体声响复制 (Ambisonics) 增大和控制源范围,Schmele等人 (Schmele、Sayin,2018) 提出了一种降低输入信号的Ambisonics阶数(这固有地增大视在源宽度)并在聆听空间周围分布源信号的解相关副本的混合法。

[0018] Zotter等人介绍了另一种方法,他们将在Zotter、Frank的研究 (Zotter、Frank,2013) 中提出的原理(即,得出引入频率相关相位和幅度差的滤波器对,以在立体声再设置中达成源范围)用于Ambisonics (Zotter F.、Frank、Kronlachner、Choi,2014)。

[0019] 基于平移的方法的常见缺点(例如 (Pulkki,1997)、(Pulkki,1999)、(Pulkki,2007)、(Pulkki、Laitinen、Erkut,2009)) 是对收听者位置的依赖性。即使与最佳位置的偏差很小,也会导致空间图像崩解为最靠近收听者的扬声器。这极大地限制了它们在虚拟现实和其中支持收听者到处自由移动的具有6个自由度 (6DoF) 的增强现实环境中的应用。此外,在基于DirAC的方法中分布时频点(例如 (Pulkki,2007)、(Pulkki、Laitinen、Erkut,2009)) 并不总能保证幻像源的空间范围的正确渲染。此外,它通常会显著降低源信号的音质。

[0020] 通常通过以下方法之一实现源信号的解相关:i) 得出具有互补幅度的滤波器对(例如 (Lauridsen,1954));ii) 使用具有恒定幅度但(随机)加扰的相位的全通滤波器(例如

(Kendall,1995)、(Potard&Burnett,2004);或iii)在空间上随机分布源信号的时频点(例如(Pihlajamäki、Santala、Pulkki,2014))。

[0021] 所有方法都有其各自的可能后果:根据i)对源信号进行互补滤波通常会导致解相关信号的感知音质发生变化;尽管ii)中的全通滤波保留源信号的音质,但加扰的相位会破坏原始相位关系,尤其是对于瞬态信号,会导致严重的时间离散和拖影。空间分布时频点被证明对于某些信号是有效的,但也会改变信号的感知音质。此外,它显示出高度信号依赖,并为脉冲信号引入了严重的伪像。

[0022] 如高级音频BIFS((Schmidt、Schröder,2004)、(Potard,2003)、(Potard&Burnett,2004))中提出的以源信号的多个解相关版本来填充体积形状假定大量的产生相互解相关的输出信号(通常每体积形状使用十个以上的点源)的滤波器是可获得的。然而,找到这样的滤波器并不是容易的任务,并且越需要更多的这样的滤波器就变得更加困难。此外,如果源信号没有完全解相关,并且收听者围绕这种形状移动,例如在(虚拟现实)场景中,则到收听者的各个源距离对应于源信号的不同延迟,且其在收听者的耳朵处的叠加导致位置相关梳状滤波,从而可能引入源信号的烦人的不稳定音质变化。

[0023] 通过利用Schmele、Sayin的研究(Schmele、Sayin,2018)中的基于Ambisonics的技术降低Ambisonics的阶数来控制源宽度显示出仅对从2阶到1阶或到0阶的过渡具有可听见的效果。此外,这些过渡不仅被感知为源加宽,而且经常被感知为幻像源的移动。虽然添加源信号的解相关版本可以帮助稳定对视在源宽度的感知,但也会引入梳状滤波器效果,从而改变幻像源的音质。

[0024] 本发明的目的是提供一种再现空间扩展声源或从空间扩展声源生成比特流的改进概念。

[0025] 此目的可通过权利要求1的用于再现空间扩展声源的设备、权利要求27的用于生成比特流的设备、权利要求35的用于再现空间扩展声源的方法、权利要求36的用于生成比特流的方法、权利要求41的比特流或权利要求47的计算机程序来达成。

[0026] 本发明基于以下发现,即通过使用收听者位置计算与空间扩展声源相关联的二维或三维外壳至投影面上的投影,可以实现以及特别地甚至可能渲染空间扩展声源的再现。该投影用于计算用于空间扩展声源的至少两个声源的位置,并且在该位置处渲染至少两个声源以获得空间扩展声源的再现,其中渲染导致两个或更多个输出信号,并且对不同位置使用不同声音信号,但是不同声音信号都与一个相同的空间扩展声源相关联。

[0027] 获得高质量的二维或三维音频再现,因为,一方面,考虑了空间扩展声源和(虚拟)收听者位置之间的时变相对位置。另一方面,空间扩展声源由关于所感知的声源范围的几何形状信息以及多个至少两个声源(例如周边点源)有效地表示,这些声源可以由现有技术已知的渲染器轻松处理。特别地,本领域中简单的渲染器始终能够关于特定输出格式或扬声器设置在特定位置处渲染声源。例如,在特定位置处的由声音位置计算器计算出的两个声源可以例如通过振幅平移在这些位置处被渲染。

[0028] 例如,当声音位置在5.1输出格式的左左环绕之间时,并且当其他声源在输出格式的右右环绕之间时,渲染器执行的振幅平移过程将对于一个声源导致用于左左环绕声道的非常相似信号,以及对于另一声源导致用于右右环绕的相应的非常相似的信号,从而用户将声源感知为来自声音位置计算器计算出的位置。但是,由于所有四个信号最终都与空间

扩展声源相关联并相关的事实,用户不会简单地感知到与声音位置计算器计算出的位置相关联的两个幻像源,而是收听者感知到单个空间扩展声源。

[0029] 一种用于再现在空间以几何形状具有限定位置的空间扩展声源的设备,该设备包括接口、投影仪、声音位置计算器以及渲染器。本发明允许解释例如在钢琴内发生的声音增强情况。钢琴是一种大型设备,并且到目前为止,钢琴声音可能已被渲染为来自单个点源。然而,这不能完全代表钢琴的真实声音特征。根据本发明,作为空间扩展声源的示例的钢琴通过至少两个声音信号反映,其中一个声音信号可以由靠近钢琴的左部,即靠近低音弦的麦克风所记录,而另一个声源可以由靠近钢琴右部,即位于产生高音的高音弦附近的不同的第二麦克风所记录。自然地,两个麦克风都将记录彼此不同的声音,这归因于钢琴内的反射情况,以及当然也归因于低音弦相比右麦克风更靠近左麦克风以及反之亦然的事实。但是,另一方面,两个麦克风信号都将具有大量相似的声音分量,最终构成钢琴的独特声音。

[0030] 根据本发明,通过也记录空间扩展声源的几何形状信息以及可选地记录与不同麦克风位置相关的位置信息(或通常与两个不同声源相关的两个不同位置)或提供对(钢琴)声音的所感知几何形状的描述,通过记录信号,来生成表示诸如钢琴的空间扩展声源的比特流。为了反映关于声源的收听者位置,即收听者可以在虚拟现实或增强实境或任何其他声音场景中“到处走动”,与空间扩展声源(诸如钢琴)相关联的外壳的投影使用收听者位置来计算,并且使用投影面来计算至少两个声源的位置,其中,特别地,优选实施例涉及声源在投影面的边缘点处的定位。

[0031] 可以利用减少的计算总量和减少的渲染总量来在二维或三维情况下实际表示示例性钢琴声音,从而,例如,当收听者更靠近声源(诸如钢琴)的左侧部分时收听者所感知到的声音不同于当收听者位于靠近声源(诸如钢琴)的右侧部分或甚至在声源(诸如钢琴)的后方时所听到的声音。

[0032] 鉴于以上,本发明构思的独特之处在于,在编码器侧,提供一种表征空间扩展声源的方式,该方式允许在声音再现情况下将空间扩展声源用于真实的二维或三维设置。此外,通过使用收听者位置计算二维或三维外壳至投影面上的投影,可以以有效方式在空间扩展声源的高度灵活描述中使用收听者位置。使用投影面计算用于空间扩展声源的至少两个声源的声音位置,并在声音位置计算器计算出的位置渲染至少两个声源以获得空间扩展声源的再现,其具有两个或更多个输出信号用于立体声再现设置或具有两个以上声道(诸如五个、七个或甚至更多个声道)的再现设置中的两个或更多个声道的耳机或多声道输出信号。

[0033] 与通过在待填充体积的所有部分中放置许多不同点源来用声音填充3D体积的现有技术方法相比,投影避免了需对许多声源进行建模,并且通过仅填充外壳的投影,即2D空间,而大幅减少所应用的点源的数量。此外,通过优选地仅对在投影的外壳上的源进行建模,在极端情况下,这些源可以是仅空间扩展声源左边界处的一个声源以及空间扩展声源右边界处的一个声源,甚至更多地减少所需点源的数量。两种减少步骤均基于两种心理声学观察:

[0034] 1、与声源的方位角(和仰角)相比,声源的距离不能被非常可靠地感知。因此,原始体积在垂直于收听者的平面上的投影不会明显改变感知(但可以帮助减少渲染所需的点源的数量)。

[0035] 2、作为点源分别分布在左侧和右侧的两个解相关声音倾向于在感知上用声音填

充它们之间的空间。

[0036] 此外,编码器侧不仅允许表征单个空间扩展声源,而且其灵活性在于作为表示生成的比特流可以包括两个或更多个空间扩展声源的所有数据,其优选地关于这些空间扩展声源的几何形状信息和到单个坐标系的位置是相关的。在解码器侧,不仅可以针对单个空间扩展声源进行再现,还可以针对若干个空间扩展声源进行再现,其中投影仪使用(虚拟)收听者位置为每个声源计算投影。另外,声音位置计算器为每个空间扩展声源计算至少两个声源的位置,并且渲染器例如通过将来自每个空间扩展声源的两个或更多个输出信号以逐信号方式或逐声道方式相加并通过将相加的声道提供至用于双耳再现的相应耳机或提供至扬声器相关再现设置中的相应扬声器或可选地提供至用于存储(组合的)两个或更多个输出信号以供后续使用或传输的存储装置,来为每个空间扩展声源渲染所有计算出的声源。

[0037] 在生成器侧或编码器侧,使用用于生成表示空间扩展声源的压缩描述的比特流的设备来生成比特流,其中该设备包括用于为空间扩展声源提供一个或多个不同声音信号的声音提供器,以及生成表示压缩声音场景的比特流的输出数据形成器,比特流包含优选地以经压缩的方式诸如被比特率压缩编码器(例如MP3、AAC、USAC或MPEG-H编码器)压缩的一个或多个不同声音信号。此外,输出数据形成器被配置为在两个或更多个不同声音信号(即第一信号是在以上示例中的钢琴左部记录的信号,以及在钢琴右部记录的信号)的情况下,将两个或更多个不同声音信号中的每个声音信号的可选的各自位置信息引入比特流中,各自位置信息指示优选地关于空间扩展声源的几何形状的信息的对应声音信号的位置。

[0038] 但是,可替代地,位置信息不必一定与空间扩展声源的几何形状有关,而是也可以与一般坐标原点有关,然而与空间扩展声源的几何形状的有关是优选的。

[0039] 此外,用于生成压缩比特流的设备还包括几何形状提供器,几何形状提供器用于计算关于空间扩展声源的几何形状的信息,并且输出数据形成器被配置为除至少两个声音信号(诸如由麦克风记录的声音信号)之外将关于几何形状的信息、每个声音信号的各自位置信息的信息引入比特流中。但是,声音提供器不一定需要实际地拾取麦克风信号,而是视情况而定也可以使用解相关处理在编码器侧生成声音信号。同时,仅少数声音信号或甚至单个声音信号可以被传输用于空间扩展声音信号,并且剩余的声音信号使用解相关处理在再现侧生成。优选地,这通过比特流中的比特流元素来发信号通知,使得声音再现器始终知道每个空间扩展声源包括多少个声音信号,从而使得再现器可以决定特别是在声音位置计算器中多少个声音信号可获得,以及应该在解码器侧诸如通过信号合成或相关处理得出多少个声音信号。

[0040] 在本实施例中,再生器将比特流元素写入比特流,该比特流元素指示针对空间扩展声源所包括的声音信号的数量,并且在解码器侧,声音再现器从比特流引导比特流元素、读取比特流元素,并且基于比特流元素决定需要基于比特流中的至少一个接收到的声音信号计算多少个用于优选地周边点源或位于周边声源之间的辅助源的信号。

[0041] 随后,将参考附图讨论本发明的优选实施例,其中:

[0042] 图1是再现侧的优选实施例的框图的概览;

[0043] 图2示出具有不同数量的周边点源的球状空间扩展声源;

[0044] 图3示出具有若干个周边点源的椭球状空间扩展声源;

- [0045] 图4示出利用不同方法来分布周边点源的位置的线状空间扩展声源；
- [0046] 图5示出利用不同过程来分布周边点源的立方体状空间扩展声源；
- [0047] 图6示出在不同距离处的球状空间扩展声源；
- [0048] 图7示出在近似参数化椭球形状内的钢琴形状空间扩展声源；
- [0049] 图8示出钢琴形状空间扩展声源，其具有三个周边点源分布在投影凸壳的极点上；
- [0050] 图9示出用于再现空间扩展声源的设备或方法的优选实施方式；
- [0051] 图10示出用于生成表示空间扩展声源的压缩描述的比特流的设备或方法的优选实施方式；以及
- [0052] 图11示出如图10所示的设备或方法所生成的比特流的优选实施方式。
- [0053] 图9示出用于再现在空间中具有限定位置和几何形状的空间扩展声源的设备的优选实施方式。设备包括接口100、投影仪120、声音位置计算器140以及渲染器160。接口被配置用以接收收听者位置。此外，投影仪120被配置为使用接口100接收的收听者位置并且额外地使用有关空间扩展声源的几何形状的信息及额外地使用有关空间扩展声源在空间中的位置的信息，计算与空间扩展声源相关联的二维或三维外壳在投影面上的投影。优选地，空间扩展声源在空间中的限定位置以及额外的空间扩展声源在空间中的几何形状经由到达比特流解多路复用器或场景解析器180的比特流被接收以用于再现空间扩展声源。比特流解多路复用器180从比特流中提取空间扩展声源的几何形状的信息，并将该信息提供给投影仪。此外，比特流解多路复用器还从比特流中提取空间扩展声源的位置，并将此信息转发给投影仪。优选地，比特流还包括至少两个不同声源的位置信息，并且优选地，比特流解多路复用器还从比特流中提取至少两个声源的压缩表示，并且通过解码器（如音频解码器190）对至少两个声源进行解压缩/解码。经解码的至少两个声源最终被转发到渲染器160，并且渲染器在声音位置计算器140提供给渲染器160的位置处渲染至少两个声源。
- [0054] 虽然图9示出具有比特流解多路复用器180和音频解码器190的比特流相关再现设备，但是再现也可以在不同于编码器/解码器场景的情况下进行。例如，在空间中的限定位置和几何形状可能已经存在于再现设备中，诸如在虚拟现实或增强现实场景中，其中数据在现场生成并在同一现场消耗。比特流解多路复用器180和音频解码器190实际上不是必需的，并且空间扩展声源的几何形状和空间扩展声源的位置的信息是可获得的，而无需从比特流的任何提取。此外，还可以预先固定地协商将至少两个声源的位置与空间扩展声源的几何形状信息相关的位置信息，并且因此不必将该信息从编码器传输到解码器，或可替代地，该数据是再次在现场生成的。
- [0055] 因此，需要注意地，在本实施例中仅提供位置信息，并且甚至在两个或更多个声源信号的情况下无需传输该信息。例如，解码器或再现器始终可以将比特流中的第一声源信号当作投影上的被放置地更靠左的声源。类似地，比特流中的第二声源信号可以被当作投影上的被放置地更靠右的声源。
- [0056] 此外，尽管声音位置计算器使用投影面计算用于空间扩展声源的至少两个声源的位置，但是不必从比特流接收至少两个声源。取而代之的是，至少两个声源中的仅单个声源可以经由比特流被接收，以及另外的声源和因此的另外的位置或位置信息实际上也可以仅在再现侧生成而无需从比特流生成器传输此类信息至再现器。然而，在其他实施例中，可以在比特流中传输所有该些信息，并且当比特率要求不严格时，额外地可以在比特流中传输

比一个或两个更多数量的声音信号,并且,音频解码器190可以对表示至少两个声源的两个、三个或甚至更多个声音信号进行解码,至少两个声源的位置由声音位置计算器140计算。

[0057] 图10示出在编码器/解码器应用内应用再现时的这种场景的编码器侧。图10示出用于生成表示空间扩展声源的压缩描述的比特流的设备。特别地,提供声音提供器200以及输出数据形成器240。在本实施方式中,空间扩展声源由具有一个或多个不同声音信号的压缩描述表示,并且输出数据形成器生成表示压缩声音场景的比特流,其中该比特流至少包括一个或多个不同声音信号以及与空间扩展声源有关的几何形状信息。这表示关于图9所示的情况,其中,所有其他信息,诸如空间扩展声源的位置(参见图9的框120中的虚线箭头),可由再现侧的用户自由选择。因此,可提供具有用于此空间扩展声源的至少一个或多个不同声音信号的空间扩展声源的唯一描述,其中该些声音信号仅仅是点源信号。

[0058] 用于生成的设备额外地包括几何形状提供器220,几何形状提供器220用于提供诸如计算有关空间扩展声源的几何形状的信息。与计算不同的提供几何形状信息的其他方式包括接收用户输入,诸如由用户手动绘制的图形或由用户例如通过语音、音调、手势或任何其他用户动作提供的任何其他信息。除了一个或多个不同声音信号之外,有关几何形状的信息也被引入比特流中。

[0059] 可选地,一个或多个不同声音信号中的每个声音信号的有关各自位置信息的信息也被引入比特流中,和/或空间扩展声源的位置信息也被引入比特流中。声源的位置信息可以与几何形状信息分开,或者可以包含在几何形状信息中。在第一种情况下,可以相对于位置信息给出几何形状信息;在第二种情况下,几何形状信息可以包括,例如对于球体,坐标中心点和半径或直径。对于盒状空间扩展声源,可以在绝对坐标中给出八个或至少一个拐角点。

[0060] 一个或多个不同声音信号中的每一个的位置信息优选地与空间扩展声源的几何形状信息有关。然而,可替代地,与相同坐标系有关的绝对位置信息也是有用的,其中给出空间扩展声源的位置或几何形状信息,以及,可替代地,也可以在具有绝对坐标的绝对坐标系内给出几何形状信息,而不是以相对方式给出几何形状信息。然而,以与一般坐标系无关的相对方式提供该数据允许用户自行在再现设置中定位空间扩展声源,如图9中指向投影仪120的虚线所示。

[0061] 在另一实施例中,图10的声音提供器200被配置用于为空间扩展声源提供至少两个不同声音信号,并且输出数据形成器被配置用于生成比特流,使得比特流包括(优选地为编码格式的)至少两个不同声音信号以及可选地至少两个不同声音信号中的每个声音信号的以绝对坐标或关于空间扩展声源的几何形状的各自位置信息。

[0062] 在实施例中,声音提供器被配置为在各个多个麦克风位置或方向上执行自然声源的记录,或者通过一个或多个解相关滤波器从单个基础信号或若干个基础信号中得出声音信号,例如,关于图1所讨论,项164和166。在发生器中使用的基础信号可以与在再现现场提供或从发生器传输至再现器的基础信号相同或不同。

[0063] 特别地,在在用于再现空间扩展声源的设备中,第一数量的基础声音信号301、302与空间扩展声源相关联,第一数量为一或大于一,其中第一数量的基础声音信号301和302与相同的空间扩展声源相关。声音位置计算器140确定用于渲染空间扩展声源的第二数量

的声源,第二数量大于1。并且,渲染器160包括一个或多个解相关器166,用于从第一数量的一个或多个基础声音信号164、301、302生成去相关信号,其中第二数量大于第一数量。

[0064] 在用于再现空间扩展声源的设备的另一实施例中,设备被配置为接收表示空间扩展声源的压缩描述的比特流,比特流包括比特流元素311,比特流元素311指示包括在装置接收的比特流或经编码的音频信号中的空间扩展声源的不同声音信号的第一数量,设备读取比特流元素311并检索包括在比特流或经编码的音频信号中的空间扩展声源的不同声音信号的第一数量。并且,声音位置计算器140确定用于渲染空间扩展声源的第二数量的声源,第二数量大于1,并且渲染器160被配置为基于从比特流提取的第一数量生成第三数量的一个或多个去相关信号,所述第三数量是从第二数量和第三数量之间的差得出的。

[0065] 在另一实施例中,几何形状提供器220被配置为从空间扩展声源的几何形状中得出参数化描述或多边形描述,并且输出数据形成器被配置为将该参数化描述或多边形描述引入比特流中。

[0066] 此外,在优选实施例中,输出数据形成器被配置为将比特流元素引入比特流中,其中,该比特流元素指示包括在比特流中或包括在与该比特流相关联的经编码的音频信号中的用于空间扩展声源的至少一个不同声音信号的数量,该数量为1或大于1。由输出数据形成器生成的比特流不一定必须是一方面具有音频波形数据以及另一方面具有元数据的完整比特流。相反地,比特流也可以仅是单独的元数据比特流,其例如包括比特流字段,比特流字段用于每个空间扩展声源的声音信号的数量、空间扩展声源的几何形状信息以及在一实施例中还用于空间扩展声源的位置信息以及可选的每个声音信号和每个空间扩展声源的位置信息、空间扩展声源的几何形状信息以及在一实施例中还用于空间扩展声源的位置信息。通常以压缩形式可用的波形音频信号通过单独的数据流或单独的传输通道被传输至再现器,以便再现器从一个源接收经编码的元数据,并从不同源接收(经编码的)波形信号。

[0067] 此外,比特流生成器的实施例包括控制器250。控制器250被配置为关于待被声音提供器提供的声音信号的数量来控制声音提供器200。与该过程一致,控制器250如由表示可选特征的阴影线指示地还将比特流元素信息提供至输出数据形成器240。输出数据形成器将有关如控制器250控制并由声音提供器200提供的声音信号的数量特定信息引入到比特流元素中。优选地,控制声音信号的数量,以便包含经编码的音频声音信号的输出比特流满足外部比特率要求。当允许的比特率高时,与允许的比特率较小的情况相比,声音提供器将提供更多的声音信号。在极端情况下,当比特率要求严格时,声音提供器将为空间扩展声源提供仅单个声音信号。

[0068] 再现器将读取相应设置的比特流元素,并将在渲染器160内继续以在解码器侧并使用所传输的声音信号合成对应数量的其他声音信号,从而最终生成所需数量的周边点源和可选的辅助源。

[0069] 但是,当比特率要求不是很严格时,控制器250将控制声音提供器提供大量的不同声音信号,例如,由对应数量的麦克风或麦克风方向记录的声音信号。然后,在再现侧,根本不需要任何解相关处理,或者仅在很小的程度上需要解相关处理,以便最终由于减少了或不需要进行解相关处理,再现器获得更好的再现质量。优选地,通过指示每个空间扩展声源的声音信号的数量比特流元素的功能,获得一方面在比特率和另一方面在质量之间的折衷。

[0070] 图11示出由图10所示的比特流生成设备生成的比特流的优选实施例。比特流包括例如具有对应数据的第二空间扩展声源401,其如SESS₂所指示。

[0071] 此外,图11示出与空间扩展声源编号1有关的每个空间扩展声源的详细数据。在图11的示例中,针对在空间扩展声源存在已经在比特流生成器中从例如拾取自放置在空间扩展声源的两个不同位置处的麦克风的麦克风输出数据中生成的两个声音信号。第一声音信号是由301所指示的声音信号1,以及第二声音信号是由302所指示的声音信号2,并且两个声音信号优选地经由音频编码器被编码以用于比特率压缩。此外,项311表示比特流元素,该比特流元素指示例如由图10的控制器250控制的用于空间扩展声源1的声音信号的数量。

[0072] 如框331所示,引入空间扩展声源的几何形状信息。项321指示优选地关于几何形状信息的声音信号的可选位置信息,例如关于钢琴示例,该位置信息对于声音信号1指示“靠近低音弦”而对于由302所指示的声音信号2指示“靠近高音弦”。几何形状信息例如可以是钢琴模型参数化表示或多边形表示,并且例如,该钢琴模型对于三角钢琴或(小)钢琴会有所不同。项341额外地示出有关空间扩展声源在空间内的位置信息的可选数据。如上所述,当用户如图9中的指向投影仪的虚线所示地提供位置信息时,该位置信息341不是必需的。然而,即使在比特流中包括位置信息341时,用户仍然可以通过用户交互来替换或修改位置信息。

[0073] 随后将讨论本发明的优选实施例。实施例涉及在6DoF VR/AR(虚拟现实/增强现实)中的空间扩展声源的渲染。

[0074] 本发明的优选实施例针对被设计用以增强空间扩展声源(SESS)的再现的方法、设备或计算机程序。特别地,本发明的方法或设备的实施例考虑了空间扩展声源与虚拟收听者位置之间的时变相对位置。换句话说,本发明的方法或设备的实施例允许听觉源宽度匹配在关于收听者的任何相对位置处的所表示的声音对象的空间范围。如此,本发明的方法或设备的实施例尤其适用于六个自由度(6DoF)虚拟、混合和增强现实应用,其中空间扩展声源补充传统采用的点源。

[0075] 本发明的方法或设备的实施例通过使用被馈送(优选且显著地)解相关信号的若干个周边点源来渲染空间扩展声源。与其他方法相比,这些周边点源的位置取决于收听者相对于空间扩展声源的位置。图1描绘了根据本发明的方法或设备的实施例的空间扩展声源渲染器的概览框图。

[0076] 框图的主要组成部分是:

[0077] 1、收听者位置:此框提供收听者的瞬时位置,例如由虚拟现实追踪系统测量出。该框可以被实现为用于检测的检测器100或用于接收收听者位置的接口100。

[0078] 2、空间扩展声源的位置和几何形状:此框提供待被渲染的空间扩展声源的位置和几何形状数据,例如作为虚拟现实场景表示的部分。

[0079] 3、投影和凸壳计算:此框120计算空间扩展声源几何形状的凸壳,然后将其在朝向收听者位置的方向上投影(例如,像平面,见下文)。可替代地,可以通过首先朝向收听者位置投影几何形状以及然后计算其凸壳来实现相同的功能。

[0080] 4、周边点源的位置:该框140根据由前一框计算出的凸壳投影数据来计算所使用的周边点源的位置。在此计算中,也可以考虑收听者位置以及因此考虑收听者的接近度/距离(请参见下文)。输出为n个周边点源位置。

[0081] 5、渲染器核：渲染器核162通过将 n 个周边点源放置在指定的目标位置来将 n 个周边点源可听化。这可以是例如使用头相关传递函数的双耳渲染器或用于扬声器再现的渲染器（例如基于矢量的振幅平移）。渲染器核从 k 个输入音频基础信号（例如乐器录音的解相关信号）以及 $m \geq (n-k)$ 个额外解相关音频信号中产生1个扬声器或耳机输出信号。

[0082] 6、源基础信号：此框164是 k 个基础音频信号的输入， k 个基础音频信号相互（充分）解相关并表示待被渲染的声源（例如，单声道- $k=1$ -或立体声- $k=2$ -乐器的录音）。例如从接收自解码器侧生成器的比特流（例如，参见图11的元素301、302）获取 k 个基础音频信号，或者可以从外部源在再现现场处提供 k 个基础音频信号。

[0083] 7、解相关器：此可选框166根据渲染 n 个周边点源的需求生成额外解相关音频信号。

[0084] 8、信号输出：渲染器为扬声器（例如 $n=5.1$ ）或双耳（通常地 $n=2$ ）渲染提供1个输出信号。

[0085] 图1示出本发明的方法或设备的实施例的框图的概览。虚线指示诸如几何形状和位置的元数据的传输。实线指示音频的传输，其中 k 、 l 及 m 指示多个音频声道。渲染器核162可以接收 $k+m$ 个音频信号和 n （ $\leq k+m$ ）个位置数据。框162、164、166共同形成一般渲染器160的实施例。

[0086] 周边点源的位置取决于空间扩展声源的几何形状（尤其是空间范围）以及收听者相对于空间扩展声源的相对位置。特别地，周边点源可以位于空间扩展声源的凸壳至投影面上的投影上。投影面可以是图片平面，即与从收听者到空间扩展声源的视线垂直的平面，也可以是围绕收听者头部的球面。投影面位于距收听者头部中心任意小距离处。可替代地，可以从相对于收听者头部的视角而言的作为球坐标的子集的方位角和仰角计算空间扩展声源的投影凸壳。在下面的说明性示例中，投影面是优选的，因为其具有更直观的特征。在投影凸壳的计算的实现中，由于正规化更简单且计算复杂度较低，角度表示是优选的。需注意地，空间扩展声源的凸壳的投影与投影空间扩展声源几何形状的凸壳是相同的，即凸壳计算和在图片平面上的投影可以按任意顺序使用。

[0087] 周边点源位置可以通过各种方式分布在空间扩展声源的凸壳的投影上，包括：

[0088] ●它们可在外壳投影周围均匀受到干扰

[0089] ●它们可分布在外壳投影的极点处

[0090] ●它们可位于外壳投影的水平或/或垂直极点处（请参见实践示例部分中的图）。

[0091] 除了周边点源之外，还可以以额外的计算复杂性为代价，使用其他辅助点源来产生增强的声学填充感。此外，可以在定位周边点源之前修改投影凸壳。例如，可以将投影凸壳朝向投影凸壳的重心收缩，这种收缩的投影凸壳可以解决渲染方法引入的各个周边点源的额外空间扩展。凸壳的修改可以进一步区分水平方向和垂直方向的缩放。

[0092] 当相对于空间扩展声源的收听者位置改变时，则空间扩展声源至投影面上的投影相应地改变。转而，周边点源的位置相应地改变。周边点源位置应优选地被选择成使得其对于空间扩展声源和收听者的连续运动平滑地改变。此外，当改变空间扩展声源的几何形状时，投影凸壳改变。这包括在3D空间中旋转空间扩展声源几何形状，从而改变投影凸壳。几何形状的旋转等于收听者位置相对于空间扩展声源的角位移，并且例如以包含性方式称为收听者和空间扩展声源的相对位置。例如，通过绕重心旋转周边点源来表示收听者绕球状

空间扩展声源的圆周运动。同样地,空间扩展声源与固定收听者的旋转会导致周边点源位置的相同改变。

[0093] 由本发明的方法或设备的实施例所生成的空间范围对于空间扩展声源与收听者之间的任何距离固有地正确地再现。自然地,当用户接近空间扩展声源时,周边点源之间的张开角增大,因为其适合于对物理现实进行建模。

[0094] 尽管周边点源的角安排(angular placement)由投影面上投影凸壳的位置唯一地确定,但是周边点源的距离可以通过各种方式进一步选择,包括:

[0095] ●所有周边点源具有等于整个空间扩展声源的距离(例如,该距离是通过空间扩展声源的重心相对于收听者头部限定)的相同距离。

[0096] ●每个周边点源的距离通过投影凸壳上的位置至空间扩展声源的几何形状的反向投影来确定,例如,周边点源投影在投影面上结果为同一点。从投影凸壳至空间扩展声源的周边点源的反向投影可能无法始终唯一确定,从而需要应用额外的投影规则(请参见实践示例部分)。

[0097] ●如果周边点源的渲染不需要距离属性,而仅需要方位角和仰角中的相对角安排,则可能根本不确定周边点源的距离。

[0098] 为了指定空间扩展声源的几何形状/凸壳,近似被使用(以及可能地被传输到渲染器或渲染器核),包括简化的一维(例如线,曲线)、二维(例如椭圆形、矩形、多边形)、或三维形状(例如椭球体、立方体、多面体),可以以各种方式分别描述空间扩展声源的几何形状或对应近似形状,包括:

[0099] ●参数化描述,即通过接受额外参数的数学表达式来对几何形状进行正规化,例如,可以通过笛卡尔坐标系上的隐式函数来描述3D的椭球形状,并且额外参数是主轴在所有三个方向上的延伸。其他参数可包括3D旋转、椭球表面的变形函数。

[0100] ●多边形描述,即原始几何形状的集合,诸如线、三角形、正方形、四面体和立方体。灵长类动物(primate)多边形和多面体可以连接到更大、更复杂的几何形状。

[0101] 周边点源信号是从空间扩展声源的基础信号中得出的。基础信号可以通过各种方式获取,诸如:1)在单个或多个麦克风位置和方向上记录自然声源(示例:如实践示例中可见,记录钢琴声音);2)人工声源的合成(示例:利用变化参数的声音合成);3)任何音频信号的组合(例如:汽车的各种机械声音,诸如发动机、轮胎、门等)。此外,可以通过多个解相关滤波器从基础信号中人为地生成额外周边点源信号(请参见前面的部分)。

[0102] 在某些应用场景中,焦点是6DoF VR/AR内容的紧凑且可互操作的存储/传输。在这种情况下,整个链包括三个步骤:

[0103] 1、将期望空间扩展声源创作/编码为比特流。

[0104] 2、传输/存储所生成的比特流。根据本发明,比特流除其他元素外还包含对空间扩展声源几何形状(参数化或多边形)的描述以及相关源的源基础信号,如单音或立体声钢琴录音。可以使用诸如mp3或MPEG-2/4高级音频编码(AAC)之类的感知音频编码算法来压缩波形(参见图10中的项260)。

[0105] 3、如前所述,基于所传输的比特流对空间扩展声源进行解码/渲染。

[0106] 除了前面描述的核心方法之外,还存在一些进一步处理的选项:

[0107] 选项1-动态选择周边点源数量和位置

[0108] 根据收听者到空间扩展声源的距离,周边点声源的数量可以变化。例如,当空间扩展声源和收听者彼此远离时,投影凸壳的张开角(孔径)变小,并且因此可以有利地选择较少的周边点源,从而节省计算和内存复杂性。在极端情况下,所有周边点源都被简化为单个剩余点源。可以应用适当的降混技术以确保基础信号与得出信号之间的干扰不会降低所得周边点源信号的音频质量。如果根据收听者的相对视点,空间扩展声源的几何形状高度不规则,则类似技术也可以应用于空间扩展声源到收听者位置的近距离。例如,作为有限长度的线的空间扩展声源几何形状可能在投影面上朝单个点退化。通常,如果周边点源在投影凸壳上的角度范围小,则空间扩展声源可由较少的周边点源表示。在极端情况下,所有周边点源都被简化为单个剩余点源。

[0109] 选项2-扩展补偿

[0110] 由于每个周边点源还展现朝凸壳投影的外部的空间扩展,因此经渲染的空间扩展声源的感知听觉图像宽度比用于渲染的凸壳稍大。为了使其与期望目标几何形状对齐,有两种可能作法:

[0111] 1、创作期间的补偿:在内容创作期间考虑渲染过程的额外扩展。具体地,在内容创作期间选择稍小的空间扩展声源几何形状,使得实际渲染的尺寸是所期望的。这可以通过监视渲染器或渲染器核在创作环境(例如制作工作室)中的效果来检查。在这种情况下,与目标尺寸相比,所传输的比特流和渲染器或渲染器核使用减小的目标几何形状。

[0112] 2、渲染期间的补偿:通过渲染过程,可以使空间扩展声源渲染器或渲染器核了解到额外感知扩展,并且因此可以补偿这种效果。作为简单的示例,在其被应用以放置周边点源之前,用于渲染的几何形状可以

[0113] ●以恒定因子 $a < 1.0$ (例如 $a = 0.9$)被减小,或

[0114] ●以恒定张开角 $\alpha = 5^\circ$ 被减小

[0115] 在这种情况下,所传输的比特流包含空间扩展声源几何形状的最终目标尺寸。

[0116] 此外,这些方法的组合是可行的。

[0117] 选项3-生成周边点源波形

[0118] 此外,可以通过考虑相对于空间扩展声源的用户位置,从所记录的音频信号中生成用于馈送周边点源的实际信号,以便建模具有几何形状相关的声音贡献的空间扩展声源,如左侧带有低音的钢琴以及反之亦然。

[0119] 示例:立式钢琴的声音以其声学行为为特征,这由(至少)两个音频基础信号建模,一个靠近钢琴键盘下端(“低音符”),以及另一个靠近键盘上端(“高音符”)。这些基础信号可以通过在记录钢琴声音时正确使用麦克风来获取,并被传输到6DoF渲染器或渲染器核,以确保它们之间有充分的解相关性。

[0120] 然后,通过考虑用户相对于空间扩展声源的位置,从这些基础信号中得出周边点源信号:

[0121] ●当用户从正面(键盘)一侧面对钢琴时,两个周边点源分别靠近钢琴键盘的左端和右端而相距甚远。在这种情况下,低音键的基础信号可以直接馈入左周边点源,以及高音键的基础信号可以直接用于驱动右周边点源。

[0122] ●由于从侧面看钢琴体积模型的投影(例如,椭圆形)很小,随着收听者绕着钢琴向右走约90度,两个周边点声源被平移非常靠近彼此。如果基础信号继续用于直接驱动周

边点源信号,则一个周边点源将主要包含高音符,而另一个将主要携带低音符。由于从物理角度来看这是非期望的,因此可以通过以与相对于钢琴重心的用户移动相同的角度进行吉文斯旋转来旋转两个基础信号以形成周边点源信号,从而改进渲染。这样,两个信号都包含相似谱内容的信号,同时仍然是解相关的(假定基础信号已经被解相关)。

[0123] 选项4:经渲染的空间扩展声源的后处理

[0124] 考虑到位置和方向相关影响,例如空间扩展声源的方向性图案,可以对实际信号进行预处理或后处理。换句话说,如前所述,从空间扩展声源发出的整个声音可以被修改以展现例如方向相关声音辐射图案。在钢琴信号的情况下,这可能意味着朝向钢琴背面的辐射具有的高频成分少于朝向钢琴前面的辐射具有的高频成分。此外,可以针对每个周边点源单独地调整周边点源信号的预处理和后处理。例如,对于每个周边点源,可以不同地选择方向性图案。在表示钢琴的空间扩展声源的给定示例中,低键范围和高键范围的方向性图案可能与上面描述的类似,但是诸如踏板噪声之类的额外信号具有更全方向的方向性图案。

[0125] 随后,总结优选实施例的若干优点。

[0126] 与用点声源完全填充空间扩展声源内部(例如,在高级音频BIFS中使用)相比的较低计算复杂度。

[0127] ●点源信号之间的破坏性干扰的低可能性

[0128] ●紧凑尺寸的比特流信息(几何形状近似、一个或多个波形)

[0129] ●使能使用为音乐消费已制作的传统录音(例如钢琴的立体声录音),以用于VR/AR渲染

[0130] 随后,提供各种实践实现示例:

[0131] ●球状空间扩展声源

[0132] ●椭球状空间扩展声源

[0133] ●线状空间扩展声源

[0134] ●立方体状空间扩展声源

[0135] ●距离相关周边点源

[0136] ●钢琴形状空间扩展声源

[0137] 如以上在本发明的方法或设备的实施例中所述,可以应用用于确定周边点源的位置的各种方法。以下实践示例说明在特定情况下的一些孤立方法。在本发明的方法或设备的实施例的完整实现中,可以考虑计算复杂性、应用目的、音频质量和实现的容易性来适当地组合各种方法。

[0138] 空间扩展声源几何形状以绿色表面网格指示。需注意地,网格可视化并不意味着通过多边形方法描述空间扩展声源几何形状,因为实际上空间扩展声源几何形状可能是从参数化规范生成的。收听者位置由蓝色三角形指示。在以下示例中,图片平面被选择为投影面,并被描绘为指示投影面的有限子集的透明灰色平面。利用相同的表面网格以绿色描绘空间扩展声源至投影面上的投影几何形状。投影凸壳上的周边点源在投影面上用红叉描绘。至空间扩展声源几何形状上的反向投影的周边点源被描绘为红点。投影凸壳上的相应周边点源和空间扩展声源几何形状上的反向投影的周边点源通过红线连接,以帮助识别视觉对应。在笛卡尔坐标系中以米为单位描述所涉及的所有对象的位置。所描述的坐标系的

选择并不意味着所涉及的计算是利用笛卡尔坐标进行的。

[0139] 图2中的第一个示例考虑了球状空间扩展声源。球状空间扩展声源相对于收听者具有固定尺寸和固定位置。在投影凸壳上选择三组不同的三个、五个和八个周边点源。在凸壳曲线上以均匀距离选择所有三组周边点源。刻意选择周边点源在凸壳曲线上的偏移位置,以使得空间扩展声源几何形状的水平范围被良好地表示。

[0140] 图2示出球状空间扩展声源,其具有不同数量(即3个(顶部)、5个(中间)和8个(底部))的均匀分布在凸壳上的周边点源。

[0141] 图3中的下一个示例考虑了椭球状空间扩展声源。椭球状空间扩展声源在3D空间中具有固定形状、位置和旋转。在本示例中,选择了四个周边点源。以下举例说明确定周边点源位置的三种不同方法:

[0142] 图3中的顶部(子图a)示出两个周边点源放置在两个水平极点处,以及两个周边点源放置在两个垂直极点处。其中,极点定位是简单并且通常是适当的。本示例表明,本方法可能会产生彼此相对靠近的周边点源位置。

[0143] 图3中的中间部分(子图b)示出所有四个周边点源均均匀地分布在投影凸壳上。选择周边点源位置的偏移量,以使得最顶部的周边点源位置与图3的子图a)中的最顶部周边点源位置重合。可以看出,周边点源位置偏移的选择对经由周边点源的几何形状的代表具有相当大的影响。

[0144] 图3中的底部(子图c)示出所有四个周边点源均均匀地分布在收缩的投影凸壳上。周边点源位置的偏移位置等于图3的子图b)中选择的偏移位置,投影凸壳的收缩操作利用方向无关拉伸因子朝投影凸壳的重心执行。

[0145] 图3示出在确定周边点源位置的三种不同方法下具有四个周边点源的椭球状空间扩展声源:a/顶部)水平和垂直极点,b/中间)凸壳上的均匀分布的点,c/底部)收缩的凸壳上均匀分布的点。

[0146] 图4中的下一个示例考虑了线状空间扩展声源。尽管先前示例考虑了体积式空间扩展声源几何形状,但本示例说明空间扩展声源几何形状可以很好地被选择作为3D空间内的一维对象。图4中的顶部(子图a)描绘了放置在有限线状空间扩展声源几何形状的极点上的两个周边点源。图4中的中间部分(子图b)示出两个周边点源放置在有限线状空间扩展声源几何形状的极点处,并且一个额外点源放置在该线的中间。如在本发明的方法或设备的实施例中所描述的,将额外点源放置在空间扩展声源几何形状内可以帮助填充大的空间扩展声源几何形状中的大间隙。图4中的底部(子图c)示出考虑了与图4的子图a)和子图b)中相同的线状空间扩展声源几何形状,但是更改了朝向收听者的相对角度,以使得线状几何形状的投影长度明显较小。如以上本发明的方法或设备的实施例中所描述的,减小尺寸的投影凸壳可以由减少数量的周边点源表示,在本特定示例中,可以由位于线状几何形状的中心的一个周边点源表示。

[0147] 图4示出线状空间扩展声源,其利用三种不同方法来分布周边点源的位置:a/顶部)投影凸壳上的两个极点;b/中间)投影凸壳上的两个极点,具有额外点源在线的中心;c/底部)一个周边点源在凸起的中心,因为经旋转的线的投影凸壳太小而无法容纳一个以上的周边点源。

[0148] 图5中的下一个示例考虑了立方体状空间扩展声源。立方体状空间扩展声源具有

固定尺寸和固定位置,但是收听者的相对位置会改变。图5的顶部(子图a)和中间部分(子图b)描绘了在投影凸壳上放置四个周边点源的不同方法。反向投影的周边点源位置由投影凸壳的选择唯一确定。图5的底部(子图c)描绘了四个周边点源,其没有良好分离的反向投影位置,取而代之的是,周边点源位置的距离被选择为等于空间扩展声源几何形状的重心的距离。

[0149] 图5示出立方体状空间扩展声源,其利用三种不同方法来分布周边点源:a/顶部)两个周边点源在水平轴线上以及两个周边点源在垂直轴线上;b/中间)两个周边点源在投影凸壳的水平极点上以及两个周边点源在投影凸壳的垂直极点上;c/底部)反向投影的周边点源距离被选择为等于空间扩展声源几何形状的重心的距离。

[0150] 图6中的下一个示例考虑了具有固定尺寸和形状的球状空间扩展声源,但相对于收听者位置的具有三种不同距离。周边点源均匀分布在凸壳曲线上。周边点源的数量是根据凸壳曲线的长度和可能的周边点源位置之间的最小距离动态确定的。图6的顶部(子图a)示出球状空间扩展声源位于近距离处,使得在投影凸壳上选择了四个周边点源。图6的中间部分(子图b)示出球状空间扩展声源位于中等距离处,使得在投影凸壳上选择了三个周边点源。图6的底部(子图c)示出球状空间扩展声源位于远距离处,使得在投影凸壳上仅选择了两个周边点源。如以上在本发明的方法或设备的实施例中描述的,周边点源的数量也可以根据球形角坐标中表示的范围来确定。

[0151] 图6示出等尺寸但位于不同距离处的球状空间扩展声源:a/顶部)近距离,具有四个周边点源均匀地分布在投影凸壳上;b/中间)中等距离,具有三个周边点源均匀分布在投影凸壳上;c/底部)远距离,具有两个周边点源均匀分布在投影凸壳上。

[0152] 图7和8中的最后一个示例考虑了放置在虚拟世界内的钢琴形状空间扩展声源。用户佩戴头戴式显示器(HMD)和耳机。向用户呈现虚拟现实场景,该场景由开放词画布和站立在自由移动区域内的地板上的3D立式钢琴模型组成(请参见图7)。开放世界画布是投影至用户周围的球体上的球状静态图像。在这种特殊情况下,开放世界画布描绘蓝天白云。用户能够从各个角度走动并观看和收听钢琴。在此场景中,钢琴被渲染为放置在重心的单个点源,或者被渲染为在投影凸壳上具有三个周边点源的空间扩展声源(请参见图8)。渲染实验表明周边点源渲染方法的现实性远优于单个点源的渲染方法。

[0153] 为了简化周边点源位置的计算,将钢琴几何形状抽象为具有类似维度的椭球体形状,请参见图7。此外,两个替代点源放置在赤道线的左、右极点上,而第三个替代点仍保留在北极,请参见图8。这种配置可以以大幅降低的计算成本从所有角度保证适当的水平源宽度。

[0154] 图7示出具有近似参数化椭球形状(指示为红色网格)的钢琴形状空间扩展声源(以绿色描绘)。

[0155] 图8示出具有分布在投影凸壳的垂直极点和投影凸壳的垂直顶部位置上的三个周边点源的钢琴形状空间扩展声源。需注意地,为了更好的可视化,将周边点源放置在拉伸的投影凸壳上。

[0156] 随后,提供本发明的实施例的特定特征。所呈现的实施例的特征如下:

[0157] ●为了填充空间扩展声源的感知声学空间,较佳地不将整个内部以解相关的点源(周边点源)填充,而只填充其面对收听者的周围(例如,朝向收听者的空间扩展声源的凸壳

的投影)。具体地,这意味着周边点源位置未依附于空间扩展声源几何形状,而是在考虑空间扩展声源相对于收听者位置的相对位置动态计算的。

[0158] o动态计算周边点源(数量和位置)

[0159] ●使用空间扩展声源形状的近似(对于使用压缩表示的情况:作为比特流的部分传输)。

[0160] 所描述技术的应用可以作为音频6DoF VR/AR标准的部分。在此上下文下,具有经典的编码/比特流/解码器(+渲染器)场景:

[0161] ●在编码器中,空间扩展声源的形状将与空间扩展声源的“基础”波形一起被编码作为辅助信息,可以是

[0162] o单声道信号,或

[0163] o立体声信号(较佳地充分解相关),或

[0164] o甚至更多记录的信号(较佳地也充分解相关)用以表征空间扩展声源。这些波形可以是低比特率编码的。

[0165] ●在解码器/渲染器中,如前所述,从比特流中检索空间扩展声源形状和相应波形,并用于渲染空间扩展声源。

[0166] 取决于所使用的实施例以及作为所描述的实施例的替代,需注意地,接口可以被实现为用于检测收听者位置的实际追踪器或检测器。然而,收听者位置通常从外部追踪器设备接收,并通过接口被馈送到再现设备中。但是,接口可以仅表示针对来自外部追踪器的输出数据的数据输入,也可以表示追踪器本身。

[0167] 此外,如上所概述,在周边声源之间可能需要额外辅助音频源。

[0168] 此外,已经发现,对于感知印象,左/右周边源以及可选地水平(相对于收听者)间隔的辅助源比垂直间隔的周边声源(即空间扩展声源的顶部和底部的周边声源)更为重要,例如,当资源稀缺时,优选至少使用水平间隔的周边(和可选的辅助)声源,而可以省略垂直间隔的周边声源以节省处理资源。

[0169] 此外,如上所概述,比特流生成器可以被实现为生成具有用于空间扩展声源的仅一个声音信号的比特流,并且剩余的声音信号可借助于解相关在解码器侧或再现侧生成。当仅存在单个信号,并且要用该单个信号平均地填充整个空间时,不需要任何位置信息。然而,在这种情况下,至少具有关于由诸如图10中的220所示的几何形状信息计算器所计算的空间扩展声源的几何形状的额外信息可能是有用的。

[0170] 这里要提到的是,如前所述的所有替代方案或方面以及由所附权利要求中的独立权利要求所限定的所有方面可以单独使用,即除了所设想的替代方案、目的或独立权利要求外,没有任何其他替代方案或目的。然而,在其他实施例中,两个或更多个替代方案或方面或独立权利要求可以彼此组合,并且在其他实施例中,所有方面或替代方案和所有独立权利要求可以彼此组合。

[0171] 本发明的编码声场描述可以存储在数字存储介质或非暂时性存储介质上,或者可以在诸如无线传输介质或有线传输介质(如因特网)的传输介质上传输。

[0172] 尽管已经在设备的上下文中描述了一些方面,但是很明显的,这些方面也表示对应方法的描述,其中框或装置对应于方法步骤或方法步骤的特征。类似地,在方法步骤的上下文中描述的方面也表示对应设备的相应框或项或特征的描述。

[0173] 依据某些实施要求,本发明的实施例可以利用硬件或软件来实现。可以使用例如软盘、DVD、CD、ROM、PROM、EPROM、EEPROM或FLASH内存的数字存储介质执行实施,数字存储介质具有电子可读控制信号存储于其上,电子可读控制信号可以与可编程计算机系统协作(或能够协作),以便执行相应的方法。

[0174] 根据本发明的一些实施例包括具有电子可读控制信号的数据载体,电子可读控制信号能够与可编程计算机系统协作,从而执行本文描述的方法之一。

[0175] 一般而言,本发明的实施例可以被实现为具有程序代码的计算机程序产品,当计算机程序产品在计算机上运行时,程序代码可操作用于执行方法之一。程序代码可以例如被存储在机器可读载体上。

[0176] 其他实施例包括存储在机器可读载体或非暂时性存储介质上的用于执行本文描述的方法之一的计算机程序。

[0177] 换句话说,因此,本发明的方法的实施例是计算机程序,该计算机程序具有当计算机程序在计算机上运行时用于执行本文描述的方法之一的程序代码。

[0178] 因此,本发明的方法的另一实施例是数据载体(或数字存储介质,或计算机可读介质),其包括记录在其上的用于执行本文所述方法之一的计算机程序。

[0179] 因此,本发明的方法的另一实施例是表示用于执行本文描述的方法之一的计算机程序的数据流或信号序列。数据流或信号序列可以例如被配置为经由数据通信连接,例如经由因特网来传输。

[0180] 另一实施例包括处理装置,例如计算机或可编程逻辑设备,被配置为或适于执行本文描述的方法之一。

[0181] 另一实施例包括计算机,该计算机上安装了用于执行本文描述的方法之一的计算机程序。

[0182] 在一些实施例中,可编程逻辑设备(例如现场可编程逻辑门阵列)可以用于执行本文描述的方法的一些或全部功能。在一些实施例中,现场可编程逻辑门阵列可以与微处理器协作以便执行本文描述的方法之一。通常,这些方法较佳地由任何硬设备执行。

[0183] 以上所述的实施例仅用于说明本发明的原理。应当理解,本文描述的配置和细节的修改和变化对于本领域的其他技术人员是显而易见的。因此,本发明的意图仅由后续的权利要求的范围所限制,而限于通过本文的实施例的描述和解释而给出的具体细节。

[0184] 参考文献

[0185] Alary,B.,Politis,A.,&Välimäki,V. (2017).Velvet Noise Decorrelator.

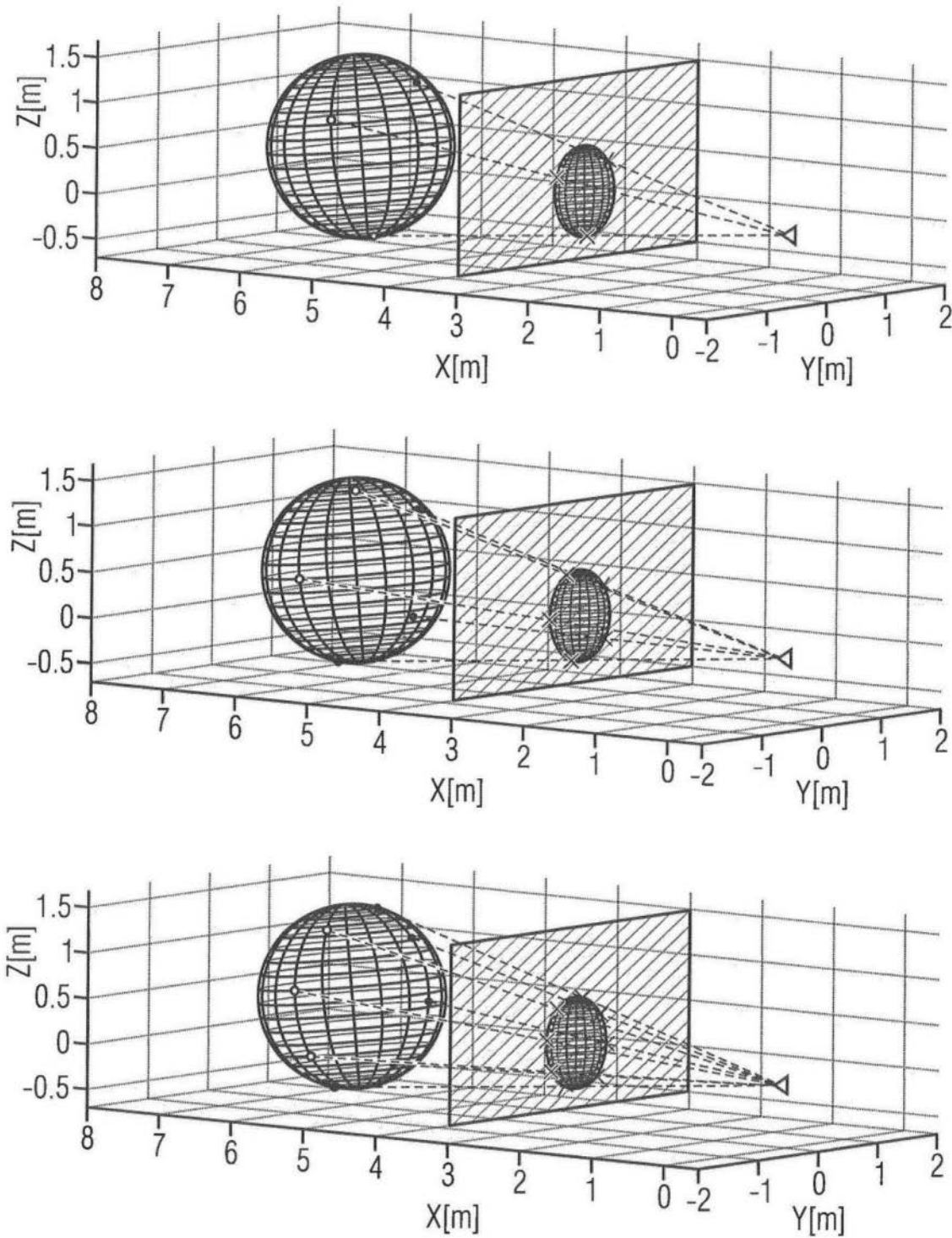
[0186] Baumgarte,F.,&Faller,C. (2003).Binaural Cue Coding-Part I: Psychoacoustic Fundamentals and Design Principles.Speech and Audio Processing,IEEE Transactions on,11 (6),S.509-519.

[0187] Blauert,J. (2001).Spatial hearing (3Ausg.).Cambridge;Mass:MIT Press.

[0188] Faller,C.,&Baumgarte,F. (2003).Binaural Cue Coding-Part II:Schemes and Applications.Speech and Audio Processing,IEEE Transactions on,11 (6),S.520-531.

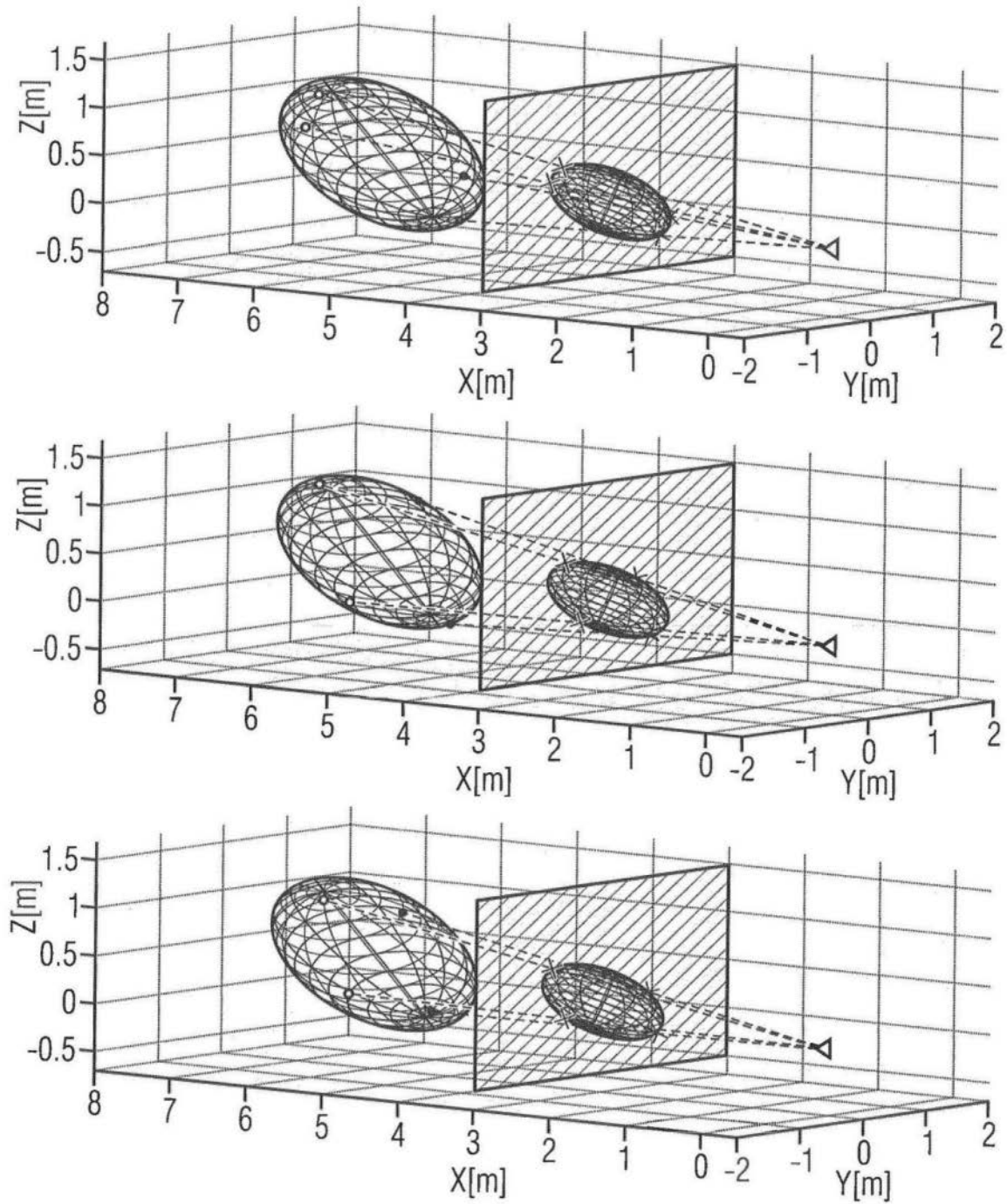
[0189] Kendall,G.S. (1995).The Decorrelation of Audio Signals and Its Impact on Spatial Imagery.

- [0190] Computer Music Journal,19(4),S.p 71-87.
- [0191] Lauridsen,H.(1954).Experiments Concerning Different Kinds of Room-Acoustics Recording.
- [0192] Ingenioren,47.
- [0193] Pihlajamäki,T.,Santala,O.,&Pulkki,V.(2014).Synthesis of Spatially Extended Virtual Source with Time-Frequency Decomposition of Mono Signals.Journal of the Audio Engineering Society,62(7/8),S.467-484.
- [0194] Potard,G.(2003).A study on sound source apparent shape and wideness.
- [0195] Potard,G.,&Burnett,I.(2004).Decorrelation Techniques for the Rendering of Apparent Sound Source Width in 3D Audio Displays.
- [0196] Pulkki,V.(1997).Virtual Sound Source Positioning Using Vector Base Amplitude Panning.Journal of the Audio Engineering Society,45(6),S.456-466.
- [0197] Pulkki,V.(1999).Uniform spreading of amplitude panned virtual sources.
- [0198] Pulkki,V.(2007).Spatial Sound Reproduction with Directional Audio Coding.J.Audio Eng.Soc,55(6),S.503-516.
- [0199] Pulkki,V.,Laitinen,M.-V.,&Erkut,C.(2009).Efficient Spatial Sound Synthesis for Virtual Worlds.Schlecht,S.J.,Alary,B.,Välimäki,V.,&Habets,E.A.(2018).Optimized Velvet-Noise Decorrelator.
- [0200] Schmele,T.,&Sayin,U.(2018).Controlling the Apparent Source Size in Ambisonics Using Decorrelation Filters.
- [0201] Schmidt,J.,&Schröder,E.F.(2004).New and Advanced Features for Audio Presentation in the MPEG-4 Standard.
- [0202] Verron,C.,Aramaki,M.,Kronland-Martinet,R.,&Pallone,G.(2010).A 3-D Immersive Synthesizer for Environmental Sounds.Audio,Speech,and Language Processing,IEEE Transactions on,title=A Backward-Compatible Multichannel Audio Codec,18(6),S.1550-1561.Zotter,F.,&Frank,M.(2013).Efficient Phantom Source Widening.Archives of Acoustics,38(1),S.27-37.
- [0203] Zotter,F.,Frank,M.,Kronlachner,M.,&Choi,J.-W.(2014).Efficient Phantom Source Widening and Diffuseness in Ambisonics.



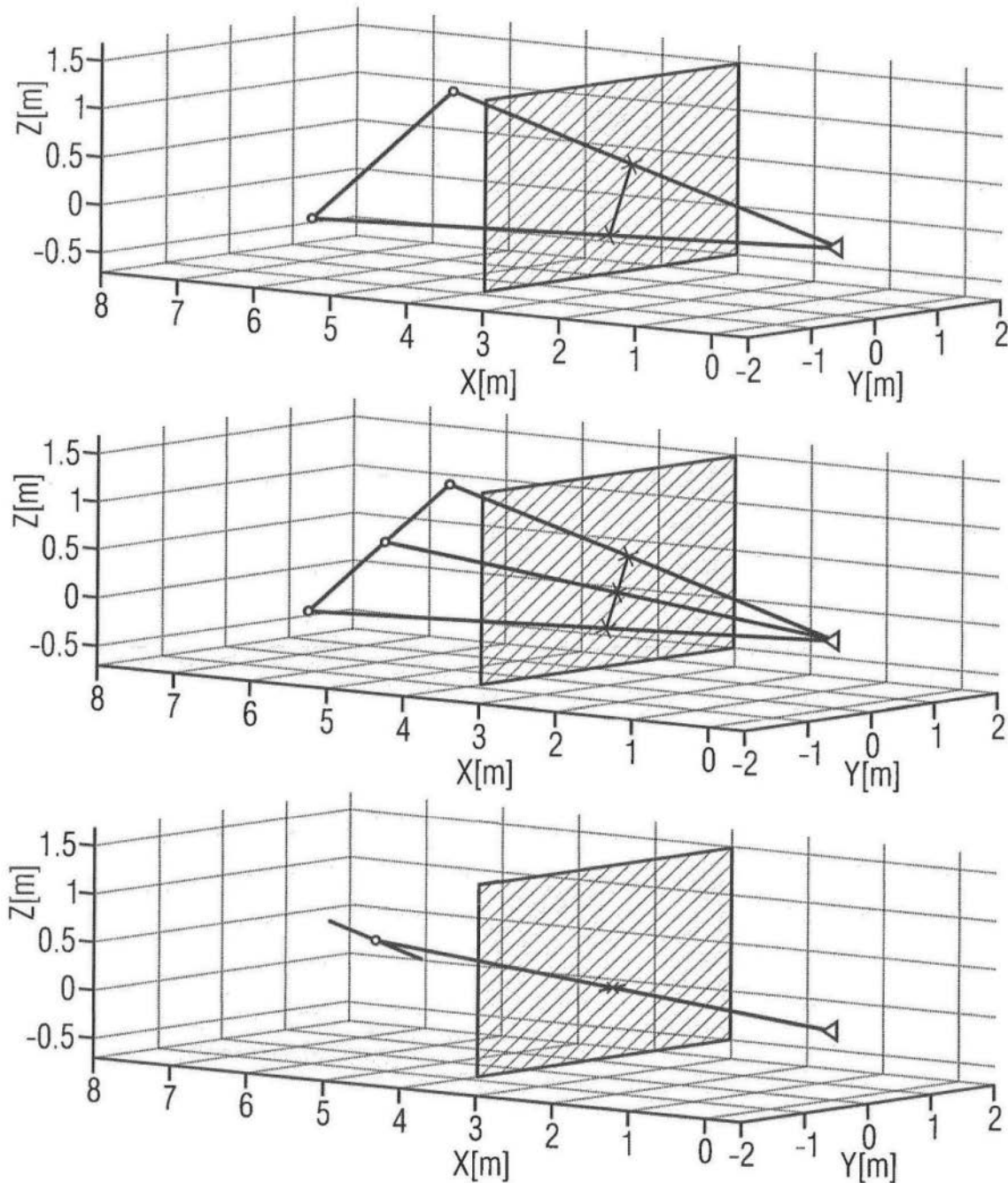
球状SESS具有不同数量(即, 3(顶部)、5(中间)、以及8(底部)个)的PPS均匀分布在凸壳上

图2



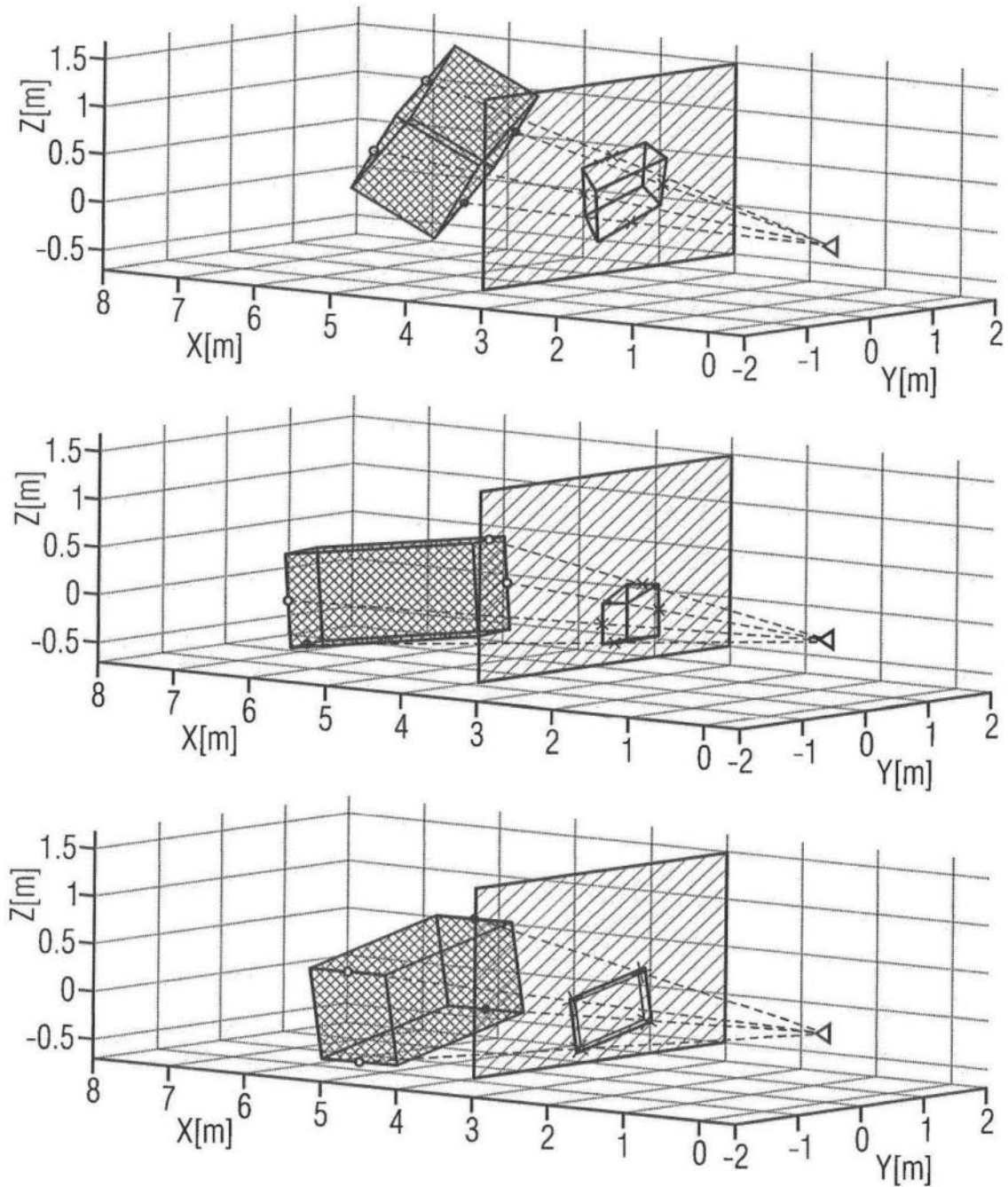
椭球状SESS在确定PPS的位置的三种不同方法下具有四个PPP：a/顶部) 水平和垂直极点；b/中间) 凸壳上的均匀分布的点；c/底部) 收缩的凸壳上的均匀分布的点

图3



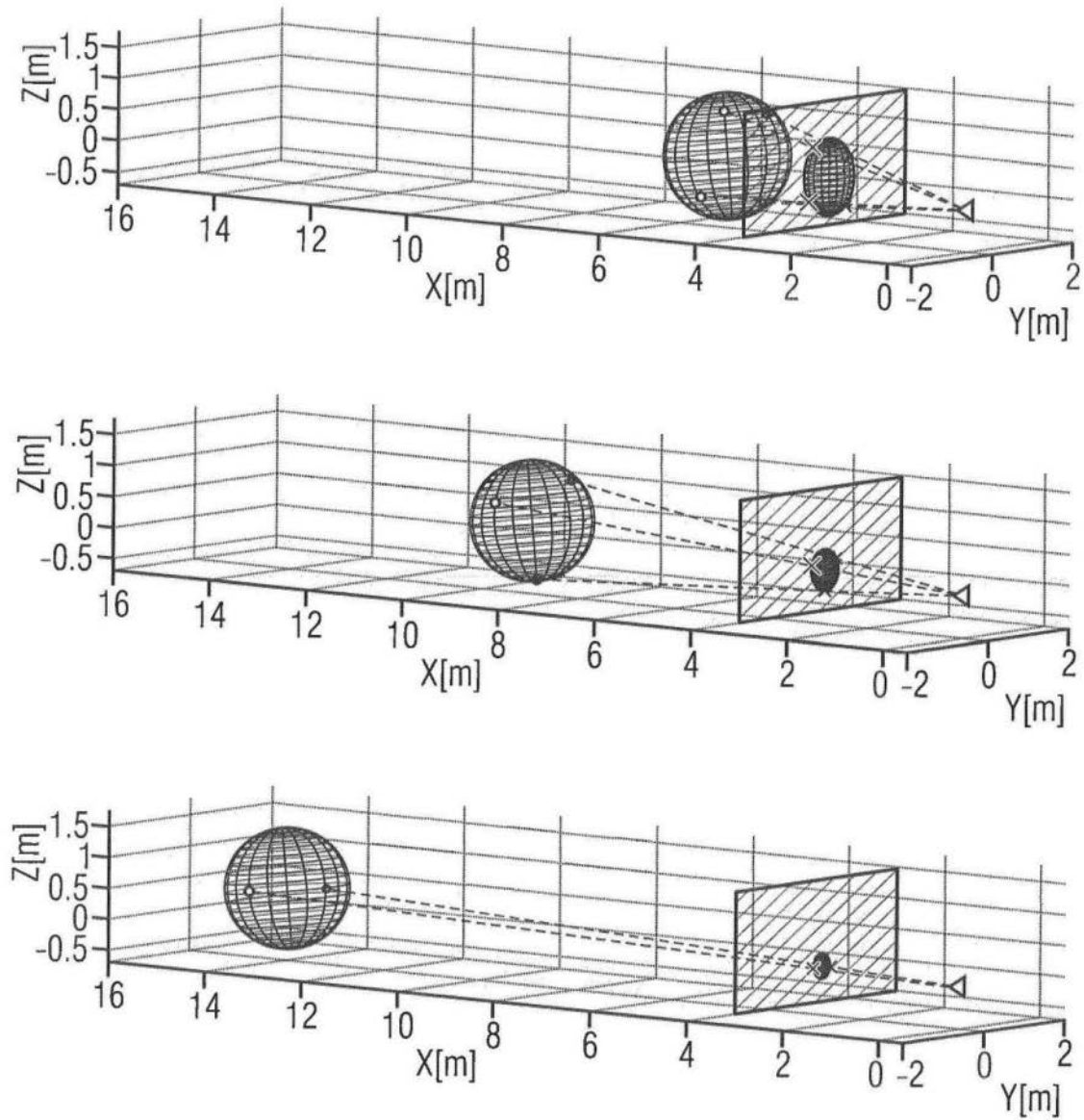
线状SESS利用三种不同方法来分布PPS的位置：a/顶部)投影凸壳上的两个极点；b/中间)投影凸壳上的两个极点、具有额外点源在线的中心；c/底部)一个PPS在凸起的中心，因为经旋转的线的投影凸壳太小而无法容纳多于一个PPS

图4



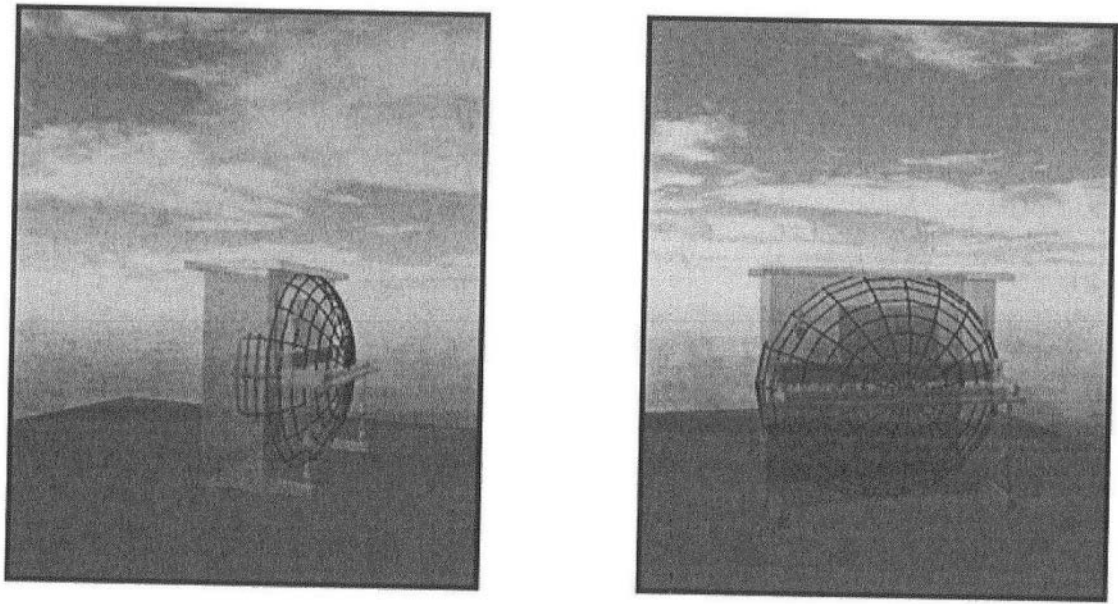
立方体状SESS利用三种不同方法来分布PPS：a/顶部)两个PPS在水平轴线上以及两个PPS在垂直轴线上；b/中间)两个PPS在投影凸壳的水平极点上以及两个PPS在投影凸壳的垂直极点上；c/底部)反向投影的PPS距离被选择为等于SESS几何形状的重心的距离

图5



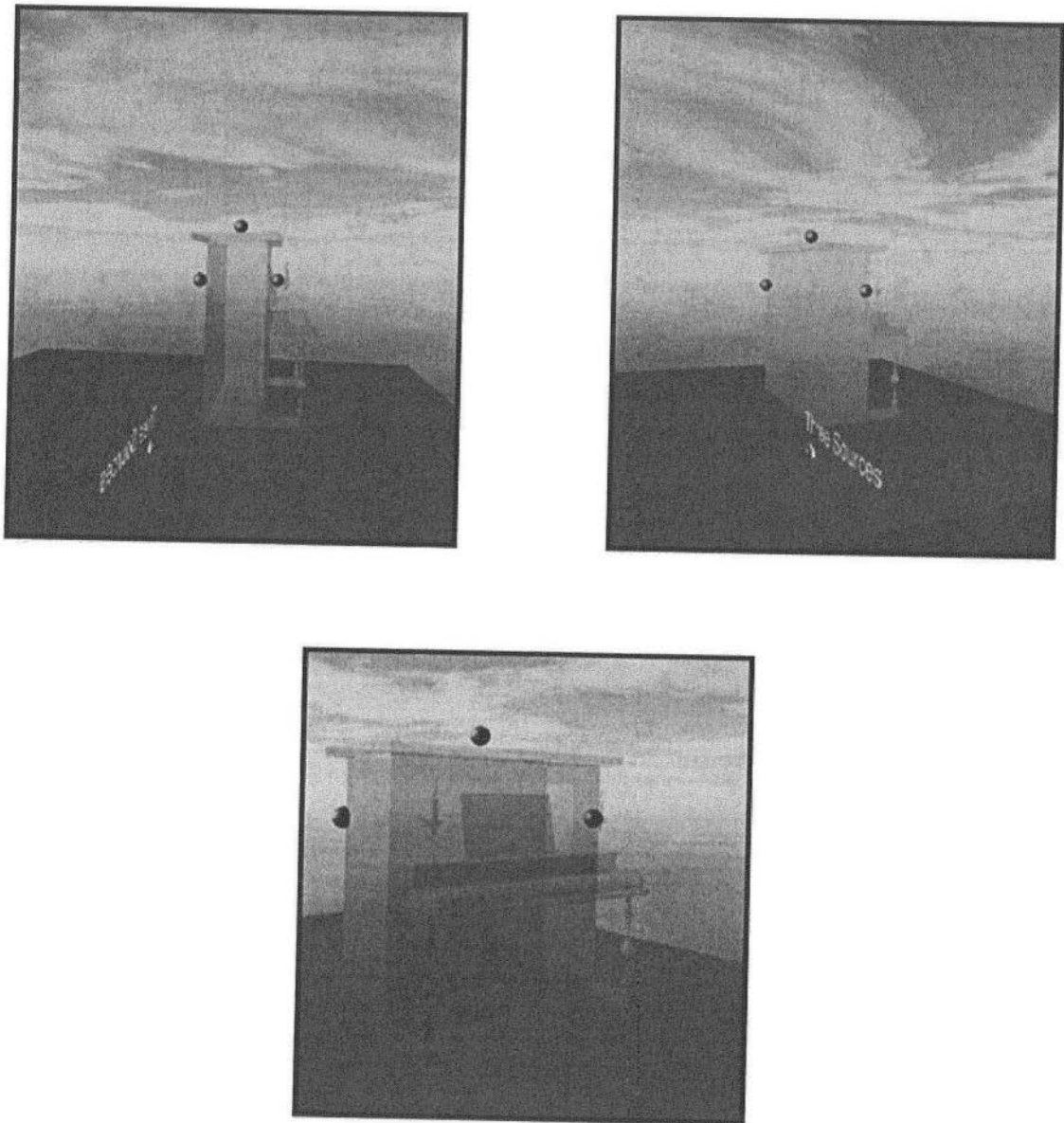
相同尺寸但在不同距离处的球状SESS：a/顶部)近距离，具有在投影凸壳上均匀分布的四个PPS；b/中间)中等距离，具有在投影凸壳上均匀分布的三个PPS；c/底部)远距离，具有在投影凸壳上均匀分布的两个PPS

图6



钢琴形状SESS(以绿色描绘)具有近似参数化椭球形状(被指示为红色网格)

图7



钢琴形状SESS具有在投影凸壳的垂直极点和投影凸壳的垂直顶部位置上分布的三个PPS。注意，为了更好的可视化，PPS被放置在收缩的投影凸壳上。

图8

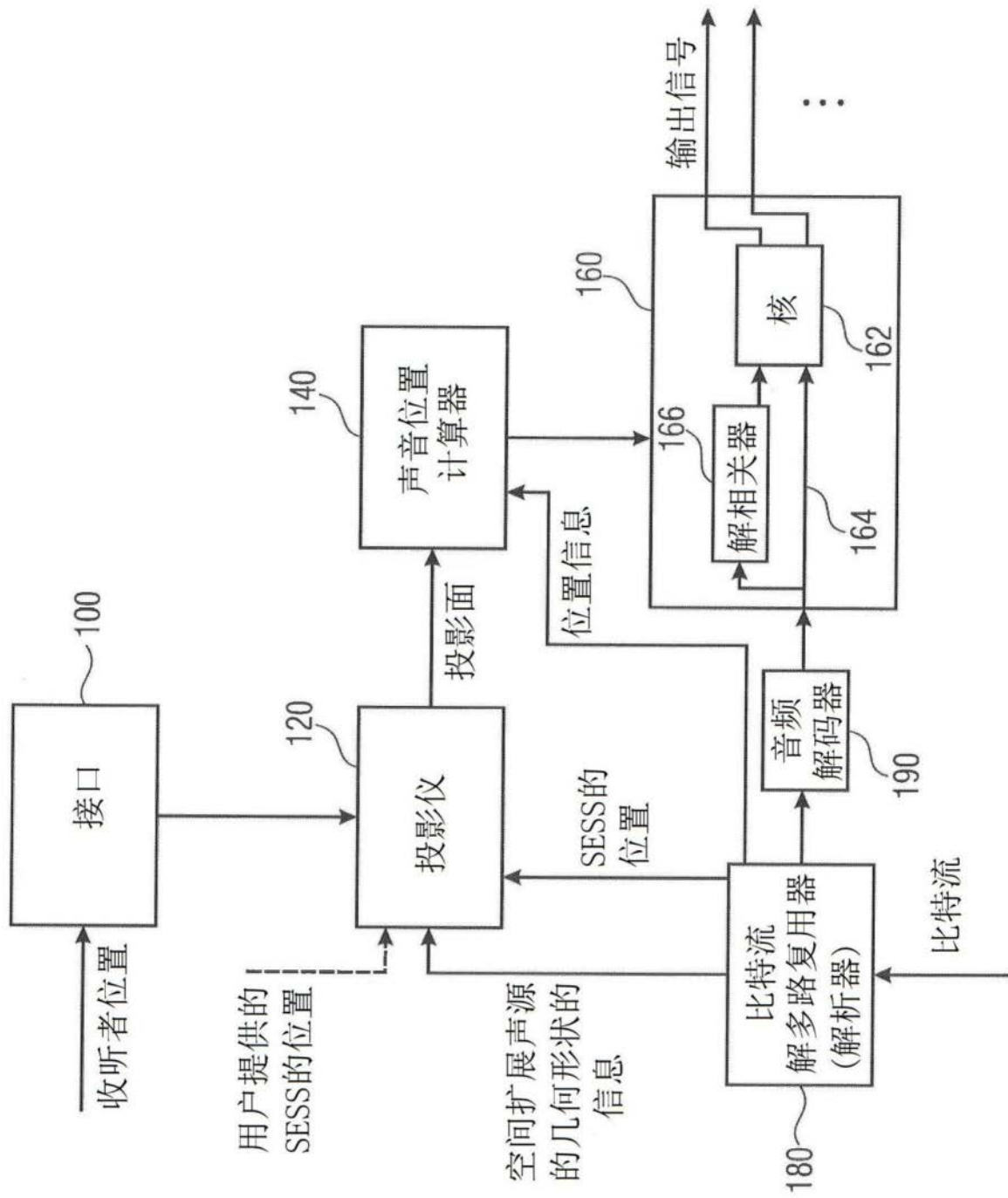


图9

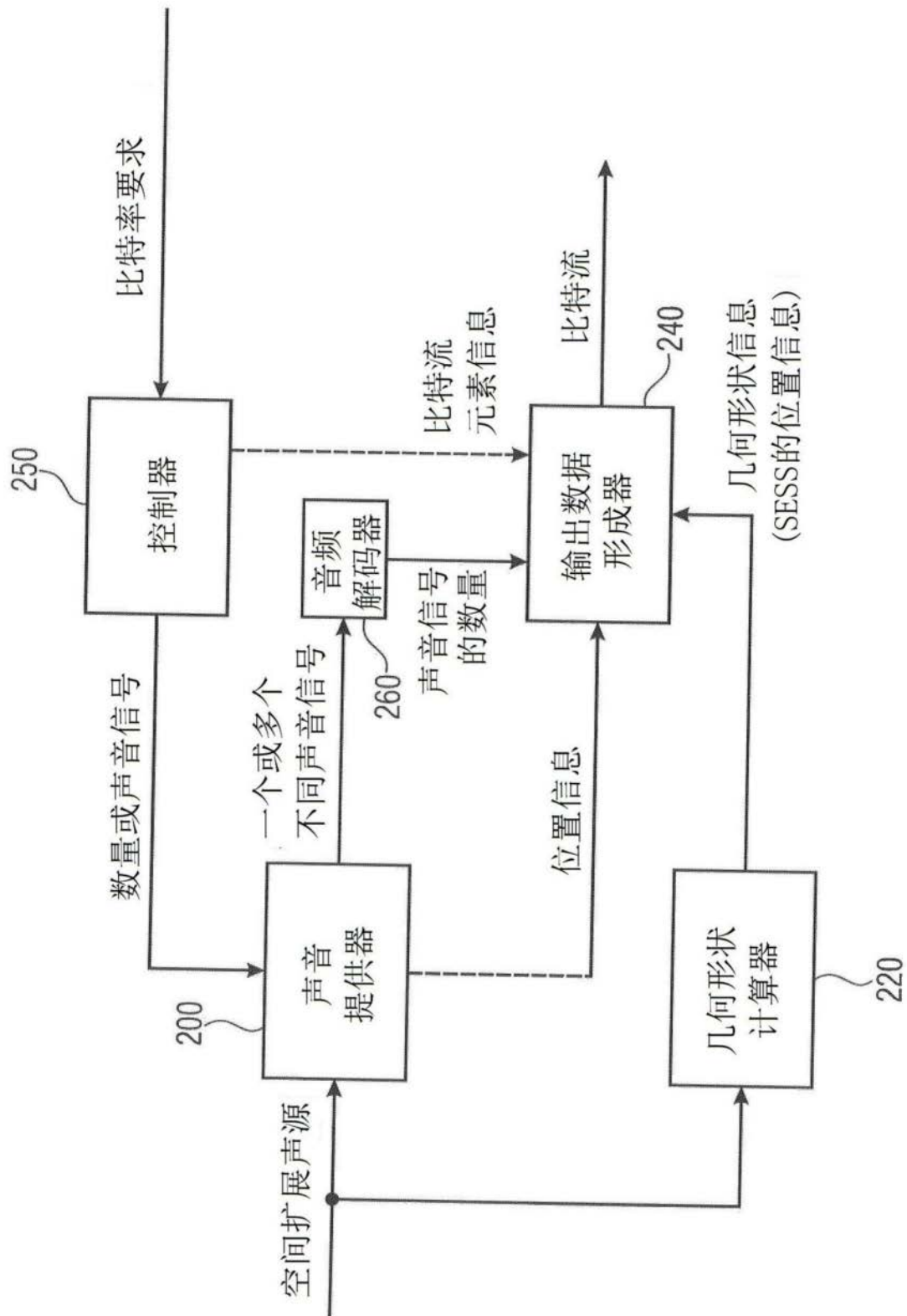


图10

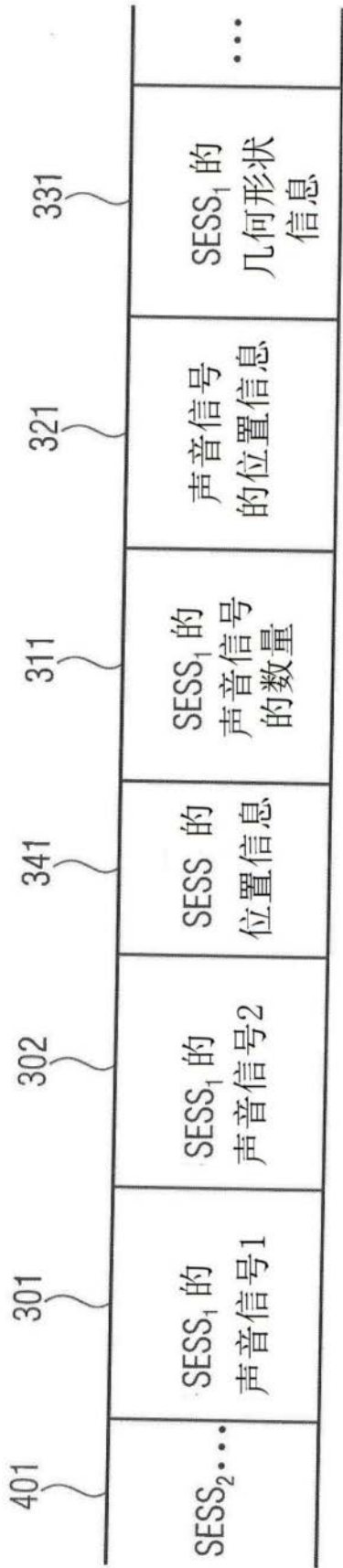


图11