(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 111276153 B
(45)授权公告日 2024.03.29

(21)申请号 202010076120.5

(22)申请日 2015.03.25

(65)同一申请的已公布的文献号
申请公布号 CN 111276153 A

(43)申请公布日 2020.06.12

(30)优先权数据
14161819.9 2014.03.26 EP
14196769.5 2014.12.08 EP

(62)分案原申请数据
201580026227.6 2015.03.25

(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 西蒙·福格 珍·普洛斯提斯
萨沙·迪克 约翰内斯·希尔珀特
朱利安·罗比亚尔 亚琴·昆兹
安德烈斯·霍瑟

(74)专利代理机构 中科专利商标代理有限责任公司 11021
专利代理师 李敬文

(51)Int.Cl.
G10L 19/20(2013.01)
G10L 19/008(2013.01)
G10L 19/16(2013.01)
H04N 21/233(2011.01)
H04N 21/431(2011.01)
H04N 21/439(2011.01)
H04N 21/45(2011.01)
H04N 21/81(2011.01)
H04N 21/84(2011.01)
H04S 3/00(2006.01)
H04S 7/00(2006.01)

(56)对比文件
CN 103313182 A,2013.09.18
CN 103650539 A,2014.03.19
CN 102667919 A,2012.09.12
US 2012183162 A1,2012.07.19
CN 102099854 A,2011.06.15
CN 103250207 A,2013.08.14
US 2014023196 A1,2014.01.23

审查员 颜博

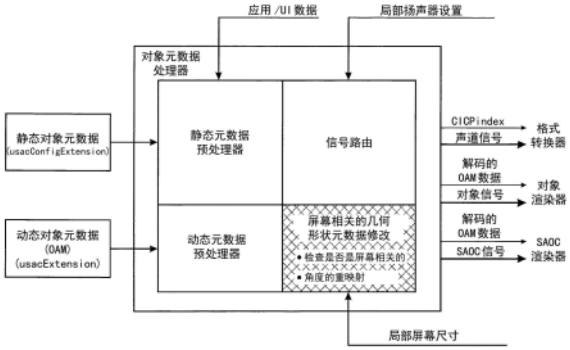
权利要求书4页 说明书26页 附图17页

(54)发明名称

屏幕相关的音频对象重映射的设备和方法

(57)摘要

提出了一种用于产生扬声器信号的设备。所述设备包括:对象元数据处理器和对象渲染器,其中对象渲染器配置为接收音频对象,且对象元数据处理器配置为接收元数据,所述元数据包括关于所述音频对象是否是屏幕相关的指示,且还包括所述音频对象的第一位置。对象元数据处理器配置为如果元数据指示音频对象是屏幕相关的,则依赖于所述音频对象的第一位置并依赖于屏幕的尺寸来计算所述音频对象的第二位置。对象渲染器配置为依赖于所述音频对象并依赖于位置信息来产生扬声器信号。对象元数据处理器配置为如果元数据指示音频对象不是屏幕相关的,则将所述音频对象的第一位置作为位置信息馈送至对象渲染器中。



CN 111276153 B

1.一种用于产生扬声器信号的设备,包括:

对象元数据处理器(110),以及

对象渲染器(120),

其中所述对象渲染器(120)配置为接收音频对象,

其中所述对象元数据处理器(110)配置为接收元数据,所述元数据包括与所述音频对象是否依赖于屏幕的尺寸有关的指示,并且还包括所述音频对象的第一位置,

其中所述对象渲染器(120)配置为:如果在所述元数据中指示所述音频对象依赖于屏幕的尺寸,则依赖于所述音频对象和依赖于第二位置来产生所述扬声器信号,其中所述音频对象的所述第二位置依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸,以及

其中所述对象渲染器(120)配置为:如果在所述元数据中指示所述音频对象不依赖于屏幕的尺寸,则依赖于所述音频对象并且依赖于所述第一位置来产生所述扬声器信号。

2.根据权利要求1所述的设备,其中所述对象元数据处理器(110)配置为如果在所述元数据中指示所述音频对象不依赖于屏幕的尺寸,则不计算所述音频对象的第二位置。

3.根据权利要求1所述的设备,其中所述对象渲染器(120)配置为进一步依赖于重放环境的扬声器的个数来产生扬声器信号。

4.根据权利要求3所述的设备,其中所述对象渲染器(120)配置为进一步依赖于所述重放环境的每一个扬声器的扬声器位置来产生扬声器信号。

5.根据权利要求1所述的设备,其中所述对象元数据处理器(110)配置为如果在所述元数据中指示所述音频对象依赖于屏幕的尺寸,则依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置,其中所述第一位置指示三维空间中的第一位置,并且其中所述第二位置指示三维空间中的第二位置。

6.根据权利要求5所述的设备,其中所述对象元数据处理器(110)配置为如果在所述元数据中指示所述音频对象依赖于屏幕的尺寸,则依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置,其中所述第一位置指示第一方位角、第一仰角和第一距离,并且其中所述第二位置指示第二方位角、第二仰角和第二距离。

7.根据权利要求1所述的设备,

其中所述对象元数据处理器(110)配置为接收所述元数据,所述元数据包括与所述音频对象是否依赖于屏幕的尺寸有关的指示作为第一指示,并且还包括第二指示,如果所述音频对象是依赖于屏幕的尺寸的,则所述第二指示表示所述音频对象是否是屏幕上对象,以及

其中所述对象元数据处理器(110)配置为依赖于所述音频对象的第一位置并且依赖于所述屏幕的尺寸来计算所述音频对象的第二位置,使得如果所述第二指示表示所述音频对象是屏幕上对象,则所述第二位置取所述屏幕的屏幕区域上的第一值。

8.根据权利要求7所述的设备,其中所述对象元数据处理器(110)配置为依赖于所述音频对象的第一位置并且依赖于所述屏幕的尺寸来计算所述音频对象的第二位置,使得如果所述第二指示表示所述音频对象不是屏幕上对象,则所述第二位置取在屏幕区域上或者不在屏幕区域上的第二值。

9.根据权利要求1所述的设备,

其中所述对象元数据处理器(110)配置为接收元数据,所述元数据包括与所述音频对

象是否依赖于屏幕的尺寸有关的指示作为第一指示,并且还包括第二指示,如果所述音频对象依赖于屏幕的尺寸,则所述第二指示表示所述音频对象是否是屏幕上对象,

其中所述对象元数据处理器(110)配置为如果所述第二指示表示所述音频对象是屏幕上对象,则依赖于所述音频对象的第一位置、依赖于所述屏幕的尺寸并且依赖于作为映射曲线的第一映射曲线来计算所述音频对象的第二位置,其中所述第一映射曲线限定了第一值区间中的原始对象位置与第二值区间中的重映射对象位置的映射,以及

其中所述对象元数据处理器(110)配置为如果所述第二指示表示所述音频对象不是屏幕上对象,则依赖于所述音频对象的第一位置、依赖于所述屏幕的尺寸并且依赖于第二映射曲线作为映射曲线来计算所述音频对象的第二位置,其中所述第二映射曲线限定了第一值区间中的原始对象位置与第三值区间中的重映射对象位置的映射,以及其中所述第二值区间小于所述第三值区间。

10.根据权利要求9所述的设备,

其中所述第一值区间、所述第二值区间和所述第三值区间的每一个是方位角的值区间,或者

其中所述第一值区间、所述第二值区间和所述第三值区间的每一个是仰角的值区间。

11.根据权利要求1所述的设备,

其中所述对象元数据处理器(110)配置为依赖于第一线性映射函数和第二线性映射函数的至少一个来计算所述音频对象的第二位置,

其中所述第一线性映射函数被定义为将第一方位角值映射至第二方位角值,

其中所述第二线性映射函数被定义为将第一仰角值映射至第二仰角值,

其中$\varphi_{left}^{nominal}$指示左方位角屏幕边缘基准,

其中$\varphi_{right}^{nominal}$指示右方位角屏幕边缘基准,

其中$\theta_{top}^{nominal}$指示顶部仰角屏幕边缘基准,

其中$\theta_{bottom}^{nominal}$指示仰角屏幕边缘基准,

其中φ_{left}^{repro}指示所述屏幕的左方位角屏幕边缘,

其中φ_{right}^{repro}指示所述屏幕的右方位角屏幕边缘,

其中θ_{top}^{repro}指示所述屏幕的顶部仰角屏幕边缘,

其中θ_{bottom}^{repro}指示所述屏幕的底部仰角屏幕边缘,

其中φ指示所述第一方位角值,

其中φ'指示所述第二方位角值,

其中θ指示所述第一仰角值,

其中θ指示所述第二仰角值,

其中所述第二方位角值φ'根据第一线性映射函数

$$\varphi' = \begin{cases} \dfrac{\varphi_{right}^{repro}+180^\circ}{\varphi_{right}^{\text{nominal}}+180^\circ}\cdot(\varphi+180^\circ)-180^\circ & for \quad -180^\circ \le \varphi < \varphi_{right}^{\text{nominal}} \\ \dfrac{\varphi_{left}^{repro}-\varphi_{right}^{repro}}{\varphi_{left}^{\text{nominal}}-\varphi_{right}^{\text{nominal}}}\cdot(\varphi-\varphi_{right}^{\text{nominal}})+\varphi_{right}^{repro} & for \quad \varphi_{right}^{\text{nominal}} \le \varphi < \varphi_{left}^{\text{nominal}} \\ \dfrac{180^\circ-\varphi_{left}^{repro}}{180^\circ-\varphi_{left}^{\text{nominal}}}\cdot(\varphi-\varphi_{left}^{\text{nominal}})+\varphi_{left}^{repro} & for \quad \varphi_{left}^{\text{nominal}} \le \varphi < 180^\circ \end{cases}$$

由第一方位角φ的第一映射产生,以及

其中所述第二仰角值θ′根据第二线性映射函数

$$\theta' = \begin{cases} \dfrac{\theta_{bottom}^{repro}+90^\circ}{\theta_{bottom}^{\text{nominal}}+90^\circ}\cdot(\theta+90^\circ)-90^\circ & for \quad -90^\circ \le \theta < \theta_{bottom}^{\text{nominal}} \\ \dfrac{\theta_{top}^{repro}-\theta_{bottom}^{repro}}{\theta_{top}^{\text{nominal}}-\theta_{bottom}^{\text{nominal}}}\cdot(\theta-\theta_{bottom}^{\text{nominal}})+\theta_{bottom}^{repro} & for \quad \theta_{bottom}^{\text{nominal}} \le \theta < \theta_{top}^{\text{nominal}} \\ \dfrac{90^\circ-\theta_{top}^{repro}}{90^\circ-\theta_{top}^{\text{nominal}}}\cdot(\theta-\theta_{top}^{\text{nominal}})+\theta_{top}^{repro} & for \quad \theta_{top}^{\text{nominal}} \le \theta < 90^\circ \end{cases}$$

由第一仰角值θ的第二映射产生。

12.一种解码器装置,包括:

USAC解码器(910),用于对比特流进行解码以获得一个或多个音频输入声道,以获得一个或多个输入音频对象,从而获得压缩的对象元数据,并且获得一个或多个SAOC传输声道,

SAOC解码器(915),用于对一个或多个SAOC传输声道进行解码,以获得第一组一个或多个渲染的音频对象,

设备(917),其中所述设备(917)包括对象元数据解码器(918)以及对象渲染器(920),其中所述对象元数据解码器(918)是根据前述权利要求之一所述的设备的对象元数据处理器(110),并且实现为对压缩的对象元数据进行解码以获得未压缩的元数据,以及所述对象渲染器(920)是根据前述权利要求之一所述的设备的对象渲染器(120),并且用于依赖于未压缩的元数据对一个或多个输入音频对象进行渲染以获得第二组一个或多个渲染的音频对象,

格式转换器(922),用于转换所述一个或多个音频输入声道以获得一个或多个转换的声道,以及

混合器(930),用于将所述第一组一个或多个渲染的音频对象的一个或多个音频对象、所述第二组一个或多个渲染的音频对象的一个或多个音频对象以及所述一个或多个转换的声道进行混合,以获得一个或多个解码的音频声道。

13.一种产生扬声器信号的方法,包括:

接收音频对象,

接收元数据,所述元数据包括与所述音频对象是否依赖于屏幕的尺寸有关的指示,并且还包括所述音频对象的第一位置,

如果在所述元数据中指示所述音频对象依赖于屏幕的尺寸,则依赖于所述音频对象并且依赖于第二位置来产生所述扬声器信号,其中所述音频对象的所述第二位置依赖于所述音频对象的所述第一位置并且依赖于屏幕的尺寸,以及

如果在所述元数据中指示所述音频对象不依赖于屏幕的尺寸,则依赖于所述音频对象并且依赖于所述第一位置来生成所述扬声器信号。

14.一种其上存储有计算机程序的计算机可读介质,所述计算机程序当在计算机或信号处理器上被执行时,用于实现权利要求13所述的方法。

屏幕相关的音频对象重映射的设备和方法

[0001] 本申请是申请日为2015年03月25日且题为“屏幕相关的音频对象重映射的设备和方法”的国际申请PCT/EP2015/056417所对应的中国国家申请(申请号:201580026227.6,进入中国国家阶段日期:2016年11月17日)的分案申请。

技术领域

[0002] 本发明涉及音频信号处理,具体地涉及音频对象重映射的设备和方法,更具体地涉及屏幕相关音频对象重映射的设备和方法。

背景技术

[0003] 随着日常生活中增加的多媒体内容消费,用于复杂多媒体解决方案的需求稳步增加。在这种场景下,将视觉内容与音频内容结合起到了重要的作用。视觉和音频多媒体内容对于可用的视觉和音频重放装置的的最优调节将是需要的。

[0004] 在现有技术中,音频对象是已知的。例如,可以将音频对象看作是具有相关联的元数据的声轨。例如,所述元数据可以描述原始音频数据的特性,例如所需的播放费而至或者音量级别。基于对象的音频的优势在于可以按照对于所有的再现扬声器层面可能的最佳方式,通过播放侧上的特定渲染处理来再现预定的移动。

[0005] 几何元数据可以用于限定应该在何处对音频对象进行渲染,例如方位角、仰角或者相对于基准点(例如,收听者)的绝对位置。将元数据与对象音频信号一起存储或传输。

[0006] 在MPEG-H的场景中,在第105界MPEG会议时,音频组评论了不同应用标准的要求和时间表(MPEG=移动图片专家组)。根据所述评论,必要的是满足时间上的某些点和针对下一代广播系统的具体要求。据此,系统应该能够在编码器输入处接收音频对象。此外,系统应该支持音频对象的信令传输、传送和渲染,并且应该使能用户控制对象,例如对话增强、交替语言跟踪和音频描述语言。

[0007] 在现有技术中,提出了不同的概念。根据“高等级高保真音频信号的回放方法和设备(Method and apparatus for playback of a higher-order ambisonics audio signal)”(参见参考文献1)中提出的第一现有技术,通过应用空间扭曲处理,使空间声音场为导向的音频重放适应所连接的视觉对象。在所述现有技术中,解码器对声音场进行扭曲,使得根据目标和基准屏幕的比率对所有声音对象沿屏幕的方向进行压缩或伸展。包括这样一种可能性:将内容和作为元数据在内容生产时使用的屏幕的基准尺寸(或者从基准收听位置的观看角度)进行编码和传输。替代地,假设固定的基准屏幕尺寸用于编码和解码,并且解码器知晓目标屏幕的实际尺寸。在这种现有技术中,解码器按照这样的方式对声音场进行翘曲,使得根据目标屏幕的尺寸与基准屏幕的尺寸的比率对所有声音对象沿屏幕的方向进行翘曲。使用所谓的“两段逐段线性”翘曲函数。伸展局限于声音项目的角向位置。在这种现有技术中,对于中心屏幕,翘曲函数的定义与用于屏幕相关重映射的映射函数的定义类似。可以将三段逐段线性映射函数的第一段和第三段定义为两段逐段线性函数。然而,利用这种现有技术,该申请局限于空间域中的HOA(HOA=高阶高保真)(声音场定向)的信号。

此外,翘曲函数只依赖于基准屏幕和再现屏幕的比率,不提供针对非中心屏幕的定义。

[0008] 在另一个现有技术中,“Vorrichtung und Verfahren zum Bestimmen einer Wiedergabeposition”(参见参考文献2),描述了一种采用声音源的位置用于视频再现的方法。依赖于与基准点的方向和距离以及照相机的参数,针对每一个声音对象来单独地确定声音源的播放位置。所述现有技术还描述了假设具有固定基准尺寸的屏幕。(在笛卡尔坐标中)对所有位置参数的线性缩放进行研究,将所述场景适应于比基准屏幕更大或更小的再现屏幕。然而根据所述现有技术,物理照相机与投影参数的集成是复杂的,并且这些参数并非总是可用的。此外,所述技术的方法工作于笛卡尔坐标(x,y,z),因为不但所述位置而且对象的距离随着屏幕缩放而改变。另外,这种现有技术不可应用于在角向坐标中将对象的位置相对于相对屏幕尺寸的改变(空间孔径角、观看角度)而进行适配。

[0009] 在另一个现有技术“Verfahren zur Audiocodierung”(参见参考文献3),描述了一种方法,所述方法包括在数据流中传输当前的(随时间变化的)水平和垂直观看角度(相对于原始场景中的收听者位置的基准观看角度)。在再现一侧,对再现的尺寸和位置进行分析,并且对声音对象的重放单独地进行优化以与基准屏幕相匹配。

[0010] 在另一个现有技术“Acoustical Zooming Based on a parametric Sound Field Representation”(参见参考文献4),描述了一种方法,所述方法提供跟随视觉场景的移动的音频渲染(听觉缩放)。将听觉缩放处理定义为虚拟记录位置的偏移。用于缩放算法的场景模型将所有的声音源放置于具有任意但是固定半径的圆上。然而,所述现有技术的方法工作于DirAC参数域,当改变距离和角度(到达方向)时,映射函数是非线性的并且依赖于缩放因子/参数,并且不支持非中心屏幕。

发明内容

[0011] 本发明的目的是提供一种改进的概念,用于采用现有多媒体再现装置的音频和视觉多媒体内容集成。本发明的目的是通过本申请所要求保护的设备、解码器装置、方法以及计算机程序来解决的。

[0012] 提出了一种音频对象重映射的设备,所述设备包括对象元数据处理器和对象渲染器。所述对象渲染器配置为接收音频对象。所述对象元数据处理器配置为接收元数据,所述元数据包括关于所述音频对象是否是屏幕相关的指示,并且还包括所述音频对象的第一位置。此外,所述对象元数据处理器配置为如果所述元数据指示所述音频对象是屏幕相关的,则依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置。所述对象渲染器配置为依赖于所述音频对象并且依赖于位置信息来产生所述扬声器信号。所述对象元数据处理器配置为如果在所述元数据中指示所述音频对象不是屏幕相关的,则将所述音频对象的第一位置作为所述位置信息馈送至所述对象渲染器。另外,所述对象元数据处理器配置为如果在所述元数据中指示所述音频对象是屏幕相关的,则将所述音频对象的第二位置作为所述位置信息馈送至所述对象渲染器。

[0013] 根据实施例,所述对象元数据处理器可以例如配置为如果在所述元数据中指示所述音频对象不是屏幕相关的,则不计算所述音频对象的第二位置。

[0014] 在实施例中,所述对象渲染器可以例如配置为不对所述位置信息是所述音频对象的第一位置还是所述音频对象的第二位置进行确定。

[0015] 根据实施例,所述对象渲染器可以例如配置为进一步依赖于重放环境的扬声器的个数来产生扬声器信号。

[0016] 在实施例中,所述对象渲染器可以例如配置为进一步依赖于所述重放环境的每一个扬声器的扬声器位置来产生扬声器信号。

[0017] 根据实施例,所述对象元数据处理器配置为如果在所述元数据中指示所述音频对象是屏幕相关的,则依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置,其中所述第一位置表示三维空间中的第一位置,并且其中所述第二位置表示三维空间中的第二位置。

[0018] 在实施例中,所述对象元数据处理器可以例如配置为如果在所述元数据中指示所述音频对象是屏幕相关的,则依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置,其中所述第一位置表示第一方位角、第一仰角和第一距离,并且其中所述第二位置表示第二方位角、第二仰角和第二距离。

[0019] 根据实施例,所述对象元数据处理器可以例如配置为接收所述元数据,所述元数据包括表示所述音频对象是否是屏幕相关的指示作为第一指示,并且还包括第二指示,如果所述音频对象是屏幕相关的,所述第二指示表示所述音频对象是否是屏幕上对象。所述对象元数据处理器可以例如配置为依赖于所述音频对象的第一位置并且依赖于所述屏幕的尺寸来计算所述音频对象的第二位置,使得如果所述第二指示表示所述音频对象是屏幕上对象,则所述第二位置取所述屏幕的屏幕区域上的第一值。

[0020] 在实施例中,所述对象元数据处理器可以例如配置为依赖于所述音频对象的第一位置并且依赖于所述屏幕的尺寸来计算所述音频对象的第二位置,使得如果所述第二指示表示所述音频对象不是屏幕上对象,所述第二位置取在屏幕区域上或者不在屏幕区域上的第二值。

[0021] 根据实施例,所述对象元数据处理器可以例如配置为接收元数据,所述元数据包括所述音频对象是屏幕相关的指示作为第一指示,并且还包括第二指示,如果所述音频对象是屏幕相关的,则所述第二指示表示所述音频对象是否是屏幕上对象。所述对象元数据处理器可以例如配置为如果所述第二指示表示所述音频对象是屏幕上对象,则依赖于所述音频对象的第一位置、依赖于所述屏幕的尺寸并且依赖于第一映射曲线作为映射曲线来计算所述音频对象的第二位置,其中所述第一映射曲线限定了第一值区间中的原始对象位置与第二值区间中的重映射对象位置的映射。此外,所述对象元数据处理器可以例如配置为如果所述第二指示表示所述音频对象不是屏幕上对象,依赖于所述音频对象的第一位置、依赖于所述屏幕的尺寸并且依赖于第二映射曲线作为映射曲线来计算所述音频对象的第二位置,其中所述第二映射曲线限定了第一值区间中的原始对象位置与第三值区间中的重映射对象位置的映射,以及其中所述第二值区间小于所述第三值区间。

[0022] 在实施例中,所述第一值区间、所述第二值区间和所述第三值区间的每一个可以例如是方位角的值区间,或者所述第一值区间、所述第二值区间和所述第三值区间的每一个可以例如是仰角的值区间。

[0023] 根据实施例,所述对象元数据处理器可以例如配置为依赖于第一线性映射函数和第二线性映射函数的至少一个来计算所述音频对象的第二位置,其中所述第一线性映射函数限定为将第一方位角值映射至第二方位角值,其中所述第二线性映射函数限定为将第一

仰角值映射至第二仰角值,其中$\varphi_{left}^{nominal}$指示左方位角屏幕边缘基准,其中$\varphi_{right}^{nominal}$指示右方位角屏幕边缘基准,其中$\theta_{top}^{nominal}$指示顶部仰角屏幕边缘基准,其中$\theta_{bottom}^{nominal}$指示底部仰角屏幕边缘基准,其中$\varphi_{left}^{repro}$指示所述屏幕的左方位角屏幕边缘,其中$\varphi_{right}^{repro}$指示所述屏幕的右方位角屏幕边缘,其中$\theta_{top}^{repro}$指示所述屏幕的顶部仰角屏幕边缘,其中$\theta_{bottom}^{repro}$指示所述屏幕的底部仰角屏幕边缘,其由$\varphi$指示所述第一方位角值,其中$\varphi'$指示所述第二方位角值,其中θ指示所述第一仰角值,其中θ′指示所述第二仰角值,其中所述第二方位角值φ'限据第一线性映射函数

[0024]
$$\varphi' = \begin{cases} \frac{\varphi_{right}^{repro}+180^\circ}{\varphi_{right}^{nominal}+180^\circ}\cdot(\varphi+180^\circ)-180^\circ & for \quad -180^\circ \le \varphi < \varphi_{right}^{nominal} \\ \frac{\varphi_{left}^{repro}-\varphi_{right}^{repro}}{\varphi_{left}^{nominal}-\varphi_{right}^{nominal}}\cdot(\varphi-\varphi_{right}^{nominal})+\varphi_{right}^{repro} & for \quad \varphi_{right}^{nominal} \le \varphi < \varphi_{left}^{nominal} \\ \frac{180^\circ-\varphi_{left}^{repro}}{180^\circ-\varphi_{left}^{nominal}}\cdot(\varphi-\varphi_{left}^{nominal})+\varphi_{left}^{repro} & for \quad \varphi_{left}^{nominal} \le \varphi < 180^\circ \end{cases}$$

[0025] 由第一方位角φ的第一映射产生,以及

[0026] 其中所述第二仰角值θ′根据第二线性映射函数

[0027]
$$\theta' = \begin{cases} \frac{\theta_{bottom}^{repro}+90^\circ}{\theta_{bottom}^{nominal}+90^\circ}\cdot(\theta+90^\circ)-90^\circ & for \quad -90^\circ \le \theta < \theta_{bottom}^{nominal} \\ \frac{\theta_{top}^{repro}-\theta_{bottom}^{repro}}{\theta_{top}^{nominal}-\theta_{bottom}^{nominal}}\cdot(\theta-\theta_{bottom}^{nominal})+\theta_{bottom}^{repro} & for \quad \theta_{bottom}^{nominal} \le \theta < \theta_{top}^{nominal} \\ \frac{90^\circ-\theta_{top}^{repro}}{90^\circ-\theta_{top}^{nominal}}\cdot(\theta-\theta_{top}^{nominal})+\theta_{top}^{repro} & for \quad \theta_{top}^{nominal} \le \theta < 90^\circ \end{cases}$$

[0028] 由第一仰角值θ的第二映射产生。

[0029] 此外,提出了一种解码器装置。所述解码器装置包括USAC解码器,用于对比特流进行解码以获得一个或多个音频输入声道,以获得一个或多个输入音频对象,以获得压缩的对象元数据,并且以获得一个或多个SAOC传输声道。另外,所述解码器装置包括SAOC解码器,用于对一个或多个SAOC传输声道进行解码,以获得第一组一个或多个渲染的音频对象。此外,所述解码器装置包括根据上述实施例所述的设备。所述设备包括对象元数据解码器,是根据上述实施例之一所述的设备的对象元数据处理器,并且实现为对压缩的对象元数据进行解码以获得未压缩的元数据,并且所述设备还包括根据上述实施例所述的设备的对象渲染器,用于依赖于未压缩的元数据对一个或多个输入音频数据进行渲染以获得第二组一个或多个渲染的音频对象。另外,所述解码器装置包括格式转换器,用于转换所述一个或多个音频输入声道以获得一个或多个转换的声道。此外,所述解码器装置包括混合器,用于将

所述第一组一个或多个渲染的音频对象的一个或多个音频对象、所述第二组一个或多个渲染的音频对象的一个或多个音频对象以及所述一个或多个转换的声道进行混合,以获得一个或多个解码的音频声道。

[0030] 另外,提出了一种产生扬声器信号的方法。所述方法包括:

[0031] -接收音频对象。

[0032] -接收元数据,所述元数据包括所述音频对象是屏幕相关的指示,且还包括所述音频对象的第一位置。

[0033] -如果在所述元数据中指示所述音频对象是屏幕相关的,依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置。

[0034] -依赖于所述音频对象并且依赖于位置信息来产生所述扬声器信号。

[0035] 如果在所述元数据中指示所述音频对象不是屏幕相关的,所述位置信息是所述音频对象的第一位置。如果在所述元数据中指示所述音频对象是屏幕相关的,所述位置信息是所述音频对象的第二位置。

[0036] 此外,提出了一种计算机程序,其中所述计算机程序配置为当在计算机或信号处理器上执行时,用于实现如上所述的方法。

附图说明

[0037] 在下文中,参考附图更详细地描述本发明的实施例,其中:

[0038] 图1是根据实施例的产生扬声器信号的设备,

[0039] 图2示出根据实施例的对象渲染器,

[0040] 图3图示了根据实施例的对象元数据处理器,

[0041] 图4示出了根据实施例的方位角重映射,

[0042] 图5示出了根据实施例的仰角重映射,

[0043] 图6示出了根据实施例的方位角重映射,

[0044] 图7示出了根据其他实施例的仰角重映射,

[0045] 图8示出了3D-音频编码器的总览图,

[0046] 图9示出了根据实施例的3D-音频解码器的总览图,

[0047] 图10示意了格式转换器的结构,

[0048] 图11示出了根据实施例的基于对象的音频渲染,

[0049] 图12示出了根据实施例的对象元数据预处理器,

[0050] 图13示出了根据实施例的方位角重映射,

[0051] 图14示出了根据实施例的仰角重映射,

[0052] 图15示出了根据实施例的方位角重映射,

[0053] 图16示出了根据其他实施例的仰角重映射,以及

[0054] 图17示出了根据另外实施例的仰角重映射。

具体实施方式

[0055] 图1示出了根据实施例的音频对象重映射的设备。所述设备包括对象元数据处理器110和对象渲染器120。

[0056] 所述对象渲染器120配置为接收音频对象。

[0057] 所述对象元数据处理器110配置为接收元数据,所述元数据包括关于所述音频对象是否是屏幕相关的指示,并且还包括所述音频对象的第一位置。此外,所述对象元数据处理器110配置为如果在所述元数据将所述音频对象表示为是屏幕相关的,则依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置。

[0058] 所述对象渲染器120配置为依赖于所述音频对象并且依赖于位置信息来产生所述扬声器信号。

[0059] 所述对象元数据处理器110配置为如果在所述元数据中指示所述音频对象不是屏幕相关的,将所述音频对象的第一位置作为所述位置信息馈送至所述对象渲染器120中。

[0060] 另外,所述对象元数据处理器110配置为如果在所述元数据中指示所述音频对象是屏幕相关的,将所述音频对象的第二位置作为所述位置信息馈送至所述对象渲染器120中。

[0061] 根据实施例,所述对象元数据处理器110可以例如配置为如果在所述元数据中指示所述音频对象不是屏幕相关的,则不计算所述音频对象的第二位置。

[0062] 在实施例中,所述对象渲染器120可以例如配置为不去确定所述位置信息是所述音频对象的第一位置还是所述音频对象的第二位置。

[0063] 根据实施例,所述对象渲染器120可以例如配置为进一步依赖于重放环境的扬声器的个数来产生扬声器信号。

[0064] 在实施例中,所述对象渲染器120可以例如配置为进一步依赖于所述重放环境的每一个扬声器的扬声器位置来产生扬声器信号。

[0065] 根据实施例,所述对象元数据处理器110配置为如果在所述元数据中指示所述音频对象是屏幕相关的,依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置,其中所述第一位置表示三维空间中的第一位置,并且其中所述第二位置表示三维空间中的第二位置。

[0066] 在实施例中,所述对象元数据处理器110可以例如配置为如果在所述元数据中指示所述音频对象是屏幕相关的,依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置,其中所述第一位置表示第一方位角、第一仰角和第一距离,并且其中所述第二位置表示第二方位角、第二仰角和第二距离。

[0067] 根据实施例,所述对象元数据处理器110可以例如配置为接收所述元数据,包括表示所述音频对象是否是屏幕相关的指示作为第一指示,并且还包括第二指示,如果所述音频对象是屏幕相关的,所述第二指示表示所述音频对象是否是屏幕上对象。所述对象元数据处理器110可以例如配置为依赖于所述音频对象的第一位置并且依赖于所述屏幕的尺寸来计算所述音频对象的第二位置,使得如果所述第二指示表示所述音频对象是屏幕上对象,则所述第二位置取所述屏幕的屏幕区域上的第一值。

[0068] 在实施例中,所述对象元数据处理器110可以例如配置为依赖于所述音频对象的第一位置并且依赖于所述屏幕的尺寸来计算所述音频对象的第二位置,使得如果所述第二指示表示所述音频对象不是屏幕上对象,所述第二位置取或者在屏幕区域上或者不在屏幕区域上的第二值。

[0069] 根据实施例,所述对象元数据处理器110可以例如配置为接收元数据,所述元数据

包括所述音频对象是屏幕相关的指示作为第一指示,并且还包括第二指示,如果所述音频对象是屏幕相关的,所述第二指示表示所述音频对象是否是屏幕上对象。所述对象元数据处理器110可以例如配置为如果所述第二指示表示所述音频对象是屏幕上对象,则依赖于所述音频对象的第一位置、依赖于所述屏幕的尺寸并且依赖于第一映射曲线作为映射曲线来计算所述音频对象的第二位置,其中所述第一映射曲线限定了第一值区间中的原始对象位置与第二值区间中的重映射对象位置的映射。此外,所述对象元数据处理器110可以例如配置为如果所述第二指示表示所述音频对象不是屏幕上对象,依赖于所述音频对象的第一位置、依赖于所述屏幕的尺寸并且依赖于第二映射曲线作为映射曲线来计算所述音频对象的第二位置,其中所述第二映射曲线限定了第一值区间中的原始对象位置与第三值区间中的重映射对象位置的映射,以及其中所述第二值区间小于所述第三值区间。

[0070] 在实施例中,所述第一值区间、所述第二值区间和所述第三值区间的每一个可以例如是方位角的值区间,或者所述第一值区间、所述第二值区间和所述第三值区间的每一个可以例如是仰角的值区间。

[0071] 在下文中,描述了本发明的具体实施例以及本发明的多个实施例的可选特征。

[0072] 可以存在并非针对固定位置而是其位置应该随着再现装置中的屏幕尺寸而变化的音频对象(与3D空间中的位置相关联的音频信号,例如给定的方位角、仰角和距离)。

[0073] 如果信令通知屏幕相关的对象(例如,通过元数据中的标志),根据具体的规则相对于屏幕尺寸来重映射/重新计算对象的位置。

[0074] 图2示出根据实施例的对象渲染器。

[0075] 作为介绍,注意以下内容:

[0076] 按照基于对象的音频格式,与对象信号一起存储或传输元数据。使用所述元数据和与播放环境有关的信息在播放侧渲染所述音频对象。这种信息是例如扬声器的个数或屏幕的尺寸。

[0077] 表1:示例元数据:

[0078]

	对象 ID
动态 OAM	方位角 仰角 增益 距离
交互性	AllowOnOff AllowPositionInteractivity AllowGainInteractivity DefaultOnOff DefaultGain InteractivityMinGain InteractivtiyMaxGain InteractivityMinAzOffset InteractivityMaxAzOffset InteractivityMinElOffset InteractivityMaxElOffset InteractivityMinDist InteractivityMaxDist
布局	IsSpeakerRelatedGroup SpeakerConfig3D AzimuthScreenRelated ElevationScreenRelated ClosestSpeakerPlayout
内容	ContentKind ContentLanguage
组	GroupID GroupDescription GroupNumMembers GroupMembers Priority
切换组	SwitchGroupID SwitchGroupDescription SwitchGroupDefault SwitchGroupNumMembers SwitchGroupMembers
音频场景	NumGroupsTotal IsMainScene NumGroupsPresent NumSwitchGroups

[0079] 对于对象,几何元数据可以用于定义应该如何渲染它们,例如按照方位角或仰角或者相对于基准点(例如收听者)的绝对位置。渲染器基于几何数据和可用的扬声器及其位

置来计算扬声器信号。

[0080] 根据本发明的实施例按照以下方式从上文中显现出来。

[0081] 为了控制屏幕相关的渲染,附加的元数据字段控制如何解释所述几何元数据:

[0082] 如果将所述字段设置为OFF,则通过渲染器对几何元数据进行解释以计算扬声器信号。

[0083] 如果将所述字段设置为ON,将所述几何元数据额定数据映射至其他值。基于几何元数据进行重映射,使得遵循对象元数据处理器的渲染器不知晓对象元数据的预处理并且操作不变。在下表中给出了这种元数据字段的示例。

[0084] 表2:用于控制屏幕相关渲染的示例元数据及其含义:

[0085]

AzimuthScreenRelated	将方位角调节为屏幕尺寸
ElevationScreenRelated	将仰角调节为屏幕尺寸

[0086]

isScreenRelatedObject	将方位角和仰角重映射以渲染与屏幕相关的对象
isOnScreenObject	对象信号涉及定位于屏幕上的对象

[0087] 此外,额定屏幕尺寸或者在音频内容的生产期间使用的屏幕尺寸可以作为元数据信息来发送。

[0088] |NominalScreenSize|在音频内容的生产期间使用的屏幕尺寸

[0089] 下表展示了如何对这些元数据进行有效地编码的示例。

[0090] 表3-根据实施例的ObjectMetadataConfig()的语义:

[0091]

语义	比特的个数	记忆值

ObjectMetadataConfig()
{
...
hasScreenRelatedObjects;
if(hasScreenRelatedObjects) {
for (o = 1; o <= num_objects; o++)
{
isScreenRelativeObject[o];
if(!isScreenRelativeObject){
isOnScreenObject[o];
}
}
}
}
	1
	1
	1
		bslbf
		bslbf
		bslbf

[0092] hasOnScreenObjects这一标志规定了是否存在屏幕相关的对象。

[0093] isScreenRelatedObject这一标志定义了对象位置是否是屏幕相关的(应该不同地渲染所述位置使得对它们的位置重映射,但是可以仍然包含所有有效的角度值)。

[0094] isOnScreenObject这一标志限定了相应的对象是“屏幕上的”。应该不同地渲染这一标志等于1的对象,使得它们的位置只取屏幕区域上的值。根据替代实施例,没有使用所述标志,但是限定了基准屏幕角度。如果isScreenRelativeObject=1,那么所有的角度都与这一基准角度相关。可能存在需要知晓所述音频对象是屏幕上的其他使用情况。

[0095] 需要注意的是,相对于isScreenRelativeObject,根据实施例存在另种可能性:重映射位置、但是仍然取所有的值(屏幕相关的),以及进行映射使其只包含屏幕区域上(屏幕上)的值。

[0096] 在对象元数据处理器中进行所述重映射,所述对象元数据处理器考虑局部屏幕尺寸并且执行几何元数据的映射。

[0097] 图3示出了根据实施例的对象元数据处理器。

[0098] 关于屏幕相关的几何元数据改进,描述一下内容。

[0099] 依赖于信息isScreenRelativeObject和isOnScreenObject,对于信号屏幕相关音频元件存在两种可能性:

[0100] a)屏幕相关的音频元件

[0101] b)屏幕上的音频元件

[0102] 在这两种情况下,通过对象元数据处理器重映射音频元件的位置数据。应用将位置的原始方位角和仰角映射到重映射的方位角和仰角的曲线。

[0103] 基准是元数据中的额定屏幕尺寸或者假设的缺省屏幕尺寸。

[0104] 例如,可以使用在ITU-R REC-BT.2022中定义的观看角度(针对平板显示器上的SDTV和HDTV电视图像的主观评估的普通观看条件)。

[0105] 两种类型的屏幕相关之间的差异在于重映射曲线的定义。

[0106] 在情况a)重映射的方位角可以取-180°和180°之间的值,并且重映射的仰角可以取-90°和90°之间的值。定义所述曲线,使得将缺省左边缘方位角和缺省右边缘方位角之间的方位角值映射(压缩或者扩大)至给定的屏幕左边缘和给定的屏幕右边缘之间的间隔(并且相应地用于仰角)。将其他方位角和仰角值相应地进行压缩或者扩大,使得覆盖整个范围的值。

[0107] 图4示出了根据实施例的方位角重映射。

[0108] 在情况b)重映射的方位角和仰角可以只取描述屏幕区域的位置的值(方位角(屏幕左边缘)、方位角(重映射)、方位角(屏幕右边缘)以及仰角(屏幕下边缘)、仰角(重映射)、仰角(屏幕上边缘))。

[0109] 对于处理这些范围外部的值存在不同的可能性:可以将它们映射到屏幕的边缘,使得-180°方位角和屏幕左边缘之间的所有对象终止于屏幕左边缘,并且屏幕右边缘和180°方位角之间的所有对象终止于屏幕右边缘。另一种可能性是将后半球的值映射到前半球。在左半球上,将-180°+方位角(屏幕左边缘)和方位角(屏幕左边缘)之间的位置映射到屏幕左边缘。将-180°和-180°+方位角(屏幕左边缘)之间的值映射到0°和方位角(屏幕左边缘)之间的值。按照相同的方式处理右半球和仰角。

[0110] 图5示出了根据实施例的仰角重映射。

[0111] 可以将曲线的梯度改变的点一x和+x2(可能与+x1不同或者相同)设置为缺省值(缺省假设的标准屏幕尺寸+位置),或者它们可以存在于元数据中(例如由随后可以将生产屏幕尺寸放置于元数据中的生产者)。

[0112] 还存在不包括线性部分而是弯曲的映射函数。

[0113] 附加的元数据可以控制重映射的方式,例如限定偏移或者非线性系数以考虑听觉的摇摄行为或者分辨率。

[0114] 同样,可以信令通知如何执行映射,例如通过将用于背面的所有对象“投影”到屏幕上。

[0115] 在以下的附图中倾听到这种替代的映射方法。

[0116] 这里,图6示出了根据实施例的方位角重映射。

[0117] 图7示出了根据实施例的仰角重映射。

[0118] 关于未知的屏幕尺寸行为:

[0119] 如果没有给出再现屏幕尺寸,那么:

[0120] -假设缺省的屏幕尺寸,或者

[0121] -即使将对象标记为屏幕相关或者屏幕上,不施加映射。

[0122] 回到图4,在另一个实施例中,在情况b),重映射的方位角和仰角可以只取描述屏幕区域上的(方位角(屏幕左边缘)≤方位角(重映射)≤方位角(屏幕右边缘)并且仰角(屏

幕下边缘)≤仰角(重映射)≤仰角(屏幕上边缘))的位置。存在两种处理这些范围之外的值的可能性:在一些实施例中,可以将它们映射到屏幕的边缘,使得+180°方位角和屏幕左边缘之间的所有对象终止于屏幕左边缘,并且屏幕右边缘和-180°方位角之间的所有对象终止于屏幕右边缘。另一种可能性是将后半球的值映射到前半球。

[0123] 在左半球上,将+180°-方位角(屏幕左边缘)和方位角(屏幕右边缘)之间的位置映射到屏幕左边缘。将+180°和+180°-方位角(屏幕左边缘)之间的值映射到0°和方位角(屏幕左边缘)之间的值。按照相同的方式处理右半球和仰角。

[0124] 图16示出了与图5类似的图。在图16所示的实施例中,在两幅图中,示出了横轴上从-90°到+90°的值区间和纵轴上从-90°至+90°的值区间。

[0125] 图17示出了与图7类似的图。在图17所示的实施例中,在两幅图中,示出了横轴上从-90°到+90°的值区间和纵轴上从-90°至+90°的值区间。

[0126] 在下文中,将参考图8-图15描述本发明的另外实施例和另外实施例的可选特征。

[0127] 根据一些实施例,如果比特流包含由OAM数据(OAM数据=相关联的对象元数据)补充的屏幕相关元件(对于至少一个音频元件,isScreenRelativeObject flag==1)并且如果经由LocalScreenSize()接口将局部屏幕尺寸信号传送至解码器,例如可以只处理屏幕相关元件的重映射。

[0128] 例如,几何位置数据(在通过用户相互作用进行的任意位置改变已经发生之前的OAM数据)可以通过映射函数的定义和利用映射到不同范围的值。例如,重映射可以改变几何位置数据作为渲染的预处理步骤,使得渲染器不知晓重映射并且操作不变。

[0129] 例如,可以考虑(混合和监测过程中使用的)额定基准屏幕的屏幕尺寸和/或播放室中的局部屏幕尺寸信息用于重映射。

[0130] 如果没有给出额定基准屏幕尺寸,可以使用缺省的基准值,例如假设4k显示器和最优的观看距离。

[0131] 如果没有给出局部屏幕尺寸信息,那么例如不应该施加重映射。

[0132] 例如,可以定义两种线性映射函数用于仰角和方位角值的重映射:

[0133] 例如,额定屏幕尺寸的屏幕边缘可以由下式给出:

[0134] $\varphi_{left}^{nominal}, \varphi_{right}^{nominal}, \theta_{top}^{nominal}, \theta_{bottom}^{nominal}$。

[0135] 例如,再现屏幕边缘可以简化为:

[0136] $\varphi_{left}^{repro}, \varphi_{right}^{repro}, \theta_{top}^{repro}, \theta_{bottom}^{repro}$。

[0137] 例如,可以通过以下的线性映射函数来定义方位角和仰角位置数据的重映射:

[0138]
$$\varphi' = \begin{cases} \dfrac{\varphi_{right}^{repro}+180^\circ}{\varphi_{right}^{nominal}+180^\circ}\cdot(\varphi+180^\circ)-180^\circ & for \quad -180^\circ \le \varphi < \varphi_{right}^{nominal} \\ \dfrac{\varphi_{left}^{repro}-\varphi_{right}^{repro}}{\varphi_{left}^{nominal}-\varphi_{right}^{nominal}}\cdot(\varphi-\varphi_{right}^{nominal})+\varphi_{right}^{repro} & for \quad \varphi_{right}^{nominal} \le \varphi < \varphi_{left}^{nominal} \\ \dfrac{180^\circ-\varphi_{left}^{repro}}{180^\circ-\varphi_{left}^{nominal}}\cdot(\varphi-\varphi_{left}^{nominal})+\varphi_{left}^{repro} & for \quad \varphi_{left}^{nominal} \le \varphi < 180^\circ \end{cases}$$

[0139] $$\theta' = \begin{cases} \dfrac{\theta_{bottom}^{repro}+90^\circ}{\theta_{bottom}^{\text{nominal}}+90^\circ}\cdot(\theta+90^\circ)-90^\circ & for \quad -90^\circ \le \theta < \theta_{bottom}^{\text{nominal}} \\ \dfrac{\theta_{top}^{repro}-\theta_{bottom}^{repro}}{\theta_{top}^{\text{nominal}}-\theta_{bottom}^{\text{nominal}}}\cdot\left(\theta-\theta_{bottom}^{\text{nominal}}\right)+\theta_{bottom}^{repro} & for \quad \theta_{bottom}^{\text{nominal}} \le \theta < \theta_{top}^{\text{nominal}} \\ \dfrac{90^\circ-\theta_{top}^{repro}}{90^\circ-\theta_{top}^{\text{nominal}}}\cdot\left(\theta-\theta_{top}^{\text{nominal}}\right)+\theta_{top}^{repro} & for \quad \theta_{top}^{\text{nominal}} \le \theta < 90^\circ \end{cases}$$

[0140] 图13示出了根据实施例的位置数据的重映射函数。具体地在图13中,描述了用于方位角映射的映射函数。在图13中,定义所述曲线,使得将额定基准左边缘方位角和额定基准右边缘方位角之间的方位角值映射(压缩或者扩大)至给定的局部屏幕左边缘和给定的局部屏幕右边缘之间的间隔。对其他方位角值相应地进行压缩或扩大,使得覆盖整个范围的值。

[0141] 例如,重映射的方位角可以取-180°和180°之间的值,并且重映射的仰角可以取-90°和90°之间的值。

[0142] 根据实施例,例如如果将isScreenRelativeObject标志设置为零,那么不会对相应的元件施加屏幕相关元件映射,并且通过渲染器直接使用几何位置数据(OAM数据加上用户相互作用引起的位置变化)以计算重放信号。

[0143] 根据一些实施例,例如可以根据再现屏幕尺寸对所有屏幕相关元件的位置重映射作为对于再现室的适应。例如,如果没有给出再现屏幕尺寸信息或者不存在屏幕相关元件,不施加重映射。

[0144] 例如,可以通过考虑重放室中的再现屏幕尺寸信息以及例如在混合和监测过程中使用的基准屏幕的屏幕尺寸信息来定义重映射。

[0145] 在图13中描述了根据实施例的方位角映射函数。在所述图13中,示出了方位角的映射函数。如图13所示,例如可以对映射函数进行定义,使得将基准屏幕的左边缘和右边缘之间的方位角值映射(压缩或扩大)至再现屏幕的左边缘和右边缘之间的间隔。可以对其他方位角值压缩或扩大,使得覆盖整个范围的值。

[0146] 可以相应地定义仰角映射函数(参见图14)。例如,屏幕相关处理也可以考虑放大为高分辨率视频内容的作坊区域。例如,可以只针对动态位置数据所伴随的并且标记为屏幕相关的元件来限定屏幕相关处理。

[0147] 在下文中,提出了3D音频编解码器系统的系统概览。可以在这种3D音频编解码系统中采用本发明的实施例。例如,3D音频编解码系统可以基于用于对声道和对象信号进行编码的MPEG-D USAC编解码器。

[0148] 根据实施例,为了增加对大量对象进行编码的效率,已经采用了MPEG SAOC技术(SAOC=空间音频对象编码)。例如根据一些实施例,三种类型的渲染器可以例如执行以下任务:渲染至声道的对象、渲染至耳机的声道或者渲染至不同扬声器装置的声道。

[0149] 当使用SAOC对对象信号进行明确地传输或者参量编码时,将相应的对象元数据信息压缩并且复用至3D-音频比特流。

[0150] 当使用SAOC显式地发送对象信号或对对象进行参数化编码时,将对应的对象元数

据信息进行压缩且复用到3D-音频比特流中。

[0151] 图8和图9示出了三维音频系统的不同算法块。

[0152] 图8和图9示出了3D-音频系统的不同算法块。具体地,图8示出了3D-音频编码器的总览图。图9示出了根据实施例的3D-音频解码器的总览图。

[0153] 现在描述图8和图8的模块的可能实施例。

[0154] 在图8中,示出了预渲染器810(也称作混合器)。在图8的结构中,预渲染器810(混合器)是可选的。在编码之前,可以可选地将预渲染器810用于将声道加对象输入场景转换为声道场景。功能上地,编码器一侧上的预渲染器810可以例如与随后描述的解码器一侧上的对象渲染器/混合器920的功能相关。对象的预渲染例如在编码器输入上确保决定性的信号熵,该信号熵基本上独立于同时活跃的对象信号的数目。利用对象的预渲染,不再要求对象元数据的传输。将离散对象信号渲染至编码器配置用于使用的声道布局。从相关联的对象元数据(OAM)获取针对每一个声道的对象权重。

[0155] 用于扬声器声道信号、离散对象信号、对象下混频信号以及预渲染信号的核心编解码基于MPEG-D USAC技术(USAC核心编码器)。通过基于对象指派以及输入声道的几何与语义信息来创建声道与对象映射信息,USAC编码器820(例如图8所示)对大量信号的编码进行处理。该映射信息描述如何将输入声道和对象映射到USAC声道元件(CPE、SCE、LFE)以及如何向解码器发送对应信息。

[0156] 所有附加的有效载荷(如SAOC数据或对象元数据)已通过扩展元素且已在编码器速率控制中考虑。

[0157] 取决于对于渲染器的速率/失真需求以及互动需求,对象的编码可以通过不同的方式进行。以下对象编码变型是可能的:

[0158] -预渲染对象:在编码前,将对象信号预渲染并混频为22.2声道信号。随后的编码链看到22.2声道信号。

[0159] -离散对象波形:将对象作为单声道波形供应至USAC编码器820。除了声道信号之外,SAC编码器820使用单个声道元件SCE来发送对象。在接收机侧对解码对象进行渲染和混合。将压缩的对象元数据信息一起发送至接收机/渲染器。

[0160] -参数化对象波形:通过SAOC参数的方式来描述对象属性及他们彼此间关系。通过USAC编码器820将对象信号与USAC下混频。将参数化信息一起发送。取决于对象的数目以及整体数据速率来选择下混频声道的数目。向SAOC渲染器发送压缩对象元数据信息。

[0161] 在解码器一侧上,USAC解码器910进行USAC解码。

[0162] 此外根据实施例,提出了一种解码器设备,参见图9。解码器设备包括USAC解码器910,用于对比特流进行解码以获得一个或多个音频输入声道,以获得一个或多个输入音频对象,以获得压缩的对象元数据并且获得一个或多个SAOC传输声道。

[0163] 另外,所述解码器设备包括SAOC解码器915,用于对一个或多个SAOC传输声道进行解码以获得第一组一个或多个已渲染的音频对象。

[0164] 此外,所述解码器设备包括上文参考图1至图7所述或者下面参考图11至图15所述的根据实施例的设备917。设备915包括对象元数据解码器918,例如图1的设备的对象元数据处理器110,并且实现用于对压缩的对象元数据进行解码以获得未压缩的元数据。

[0165] 另外,如上所述根据实施例的设备917包括对象渲染器920,类似图1的设备的对象

渲染器120,用于依赖于未压缩的元数据对一个或多个输入音频对象进行渲染以获得第二组一个或多个已渲染的音频对象。

[0166] 另外,所述解码器设备包括格式转换器922,用于对一个或多个音频输入声道进行转换以获得一个或多个已转换的声道。

[0167] 此外,所述解码器设备包括混合器930,用于将第一组的一个或多个已渲染的音频对象的一个或多个音频对象、第二组的一个或多个已渲染的音频对象的一个或多个音频对象以及一个或多个已转换的声道进行混合以获得一个或多个已解码的音频信道。

[0168] 在图9中,示出了解码器设备的具体实施例。用于对象信号的SAOC编码器815(SAOC编码器815是可选的,参见图8)和SAOC解码器915(参见图9)是基于MPEG SAOC技术。系统能够基于少量已发送声道和附加的参数数据(OLD,IOC,DMG)(OLD=对象级别差异,IOC=相互对象相关性,DMG=下混频增益)来重新创建、改进和渲染大量音频对象。附加的参数数据表现出比单独地的发送所有对象所要求的明显更低的数据率,使得编码非常有效。

[0169] SAOC编码器815采取单声道波形的对象/声道信号作为输入,并且输出参数信息(封装到3D-音频比特流中)和SAOC传输声道(使用单一声道元件进行编码并且发送)。

[0170] SAOC解码器915根据已解码的SAOC传输声道和参数信息来重构对象/声道信号,并且基于再现布局、解压缩的对象元数据信息以及可选地用户相互作用信息来产生输出音频场景。

[0171] 对于对象元数据编解码器,针对每一个对象,通过按照时间和空间对对象性质的量化对规定了几何位置和对象在3D空间中的展开的相关联元数据进行有效地编码,例如通过图8的元数据编码器818。将压缩的对象元数据cOAM(cOAM=压缩的音频对象元数据)作为侧面信息发送至接收机。在接收机处,通过对象元数据解码器918对cOAM进行解码。

[0172] 例如在图9中,对象元数据解码器918可以例如实现根据上述实施例之一的对象元数据处理器。

[0173] 对象渲染器(例如图9的对象渲染器920)根据给定的再现格式,利用压缩的对象元数据来产生对象波形。根据对象的元数据将每一个对象渲染至某个输出声道。这种块的输出由部分结果之和来产生。

[0174] 例如在图9中,对象渲染器920例如可以根据上述实施例之一来实现。

[0175] 在图9中,对象元数据解码器918可以例如实现为参考图1至图7和图11至图15所述、根据上述或下述实施例之一所述的对象元数据处理器,并且对象渲染器920可以例如实现为参考图1至图7和图11至图15所述、根据上述或下述实施例之一所述的对象渲染器。对象元数据解码器918和对象渲染器920可以例如一起实现用于参考图1至图7和图11至图15所述的如上所述或者如下所述产生扬声器信号的设备917。

[0176] 如果对基于声道的内容和离散/参数对象两者进行解码,在输出结果波形之前对基于声道的波形和已渲染的对象波形进行混合,例如通过图9的混合器930(或者像双耳渲染器或者扬声器渲染器模块那样,在将其馈送至后处理器模块之前)。

[0177] 双耳渲染器模块940可以例如产生多声道音频材料的双耳下混频,使得通过虚拟声音源来表示每一个输入信道。进行所述处理是QMF域中的帧状的。例如,双耳化可以是基于测量的双耳室内脉冲响应。

[0178] 扬声器渲染器922可以例如在发射声道结构和所需的再现格式之间转换。因此在

下文中称作格式转换器922。格式转换器922执行至少量输出声道的转换,例如其创建下混频。系统针对输入和输出格式的给定组合产生优化的下混频矩阵,并且将这些矩阵应用于下混频过程。格式转换器922允许标准的扬声器结构以及具有非标准扬声器位置的随机结构。

[0179] 图10示出了格式转换器的结构。图10示出了下混频配置器1010和用于在QMF域(QMF域=积分镜像滤波器域)处理下混频的下混频处理器。

[0180] 根据一些实施例,对象渲染器920可以配置为实现屏幕相关的音频对象重映射,如相对于已经参考图1至图7描述的上述多个实施例之一所描述的或者相对于如将参考图11至图15描述的多个下述实施例所描述的屏幕相关音频对象重映射。

[0181] 在下文中,描述了本发明实施例的另外实施例和概念。

[0182] 根据一些实施例,对用的用户控制可以例如采用描述性元数据,例如与比特流内部的对象存在和对象的高级别性质有关的信息,并且可以例如采用限制性元数据,例如与内容创建者如何进行或者使能相互作用有关的信息。

[0183] 根据一些实施例,音频对象的信令通知、传输和渲染可以例如采用位置性元数据、结构性元数据,例如对象的分组和分层、渲染特定扬声器以及作为对象信令通知声道内容的能力、以及将对象场景适应屏幕尺寸的装置。

[0184] 实施例提出:除了已经定义的3D空间中对象的几何位置和水平之外还开发的新元数据字段。

[0185] 如果在不同的再现装置中再现了基于对象的音频场景,根据一些实施例,已渲染的声音源的位置可以例如自动地缩放至再现的尺寸。在展示音频-视觉内容的情况下,音频对象再现的标准渲染可以例如导致位置型的音频-视觉一致性的扰乱,因为声音源位置和声音的视觉发起者的位置可能例如不再是恒定的。

[0186] 为了避免这种效果,例如可以采用一种可能情况来信令通知音频对象并非用于3D空间的固定位置,但是其位置应该随着再现装置中的屏幕尺寸而改变。根据一些实施例,这些音频对象的特殊处理以及针对场景缩放算法的定义可以例如允许更加逼真的体验,因为可以基于重放环境的局部特征来优化重放。

[0187] 在一些实施例中,渲染器或预处理模块可以例如考虑再现房间中的局部屏幕尺寸,并且因此可以例如保持电影或游戏内容中的音频和视频之间的关系。在这些实施例中,然后例如可以根据再现装置自动地缩放音频场景,使得视觉元件的位置和相应的声音源的位置一致。例如,可以保持对于尺寸变换的屏幕的位置音频-视觉一致性。

[0188] 例如根据实施例,然后可以与再现屏幕尺寸无关地从屏幕上扬声器的方向察觉对话或者演讲。然后,这可以用于固定源和其中声音轨迹和视觉元件的移动必须相对应的移动源。

[0189] 为了控制屏幕相关渲染,引入附加的元数据字段,所述附加的元数据字段允许将对象标记为屏幕相关的。如果将对象标记为屏幕相关的,在渲染之前将对象的几何位置元数据重映射至其他值。例如,图13示出了针对方位角的示范性(重新)映射函数。

[0190] 此外,一些实施例可以例如实现在角度域(方位角、仰角)工作来定义的简单映射函数。

[0191] 此外,一些实施例可以例如实现不改变对象的距离、不会进行朝向屏幕或者远离

屏幕的“缩放”或“视觉移动”,但是只是对象的位置进行缩放。

[0192] 另外,当映射函数不但基于屏幕比而且考虑屏幕边缘的方位角和仰角时,一些实施例可以例如处理处理非中心再现屏幕($\left|\varphi_{left}^{repro}\right| \neq \left|\varphi_{right}^{repro}\right|$和/或$\left|\theta_{top}^{repro}\right| \neq \left|\theta_{bottom}^{repro}\right|$)。

[0193] 此外,一些实施例可以例如定义用于屏幕上对象的特定映射函数。根据一些实施例,用于方位角和仰角的映射函数可以例如是独立的,因此可以所述映射函数以只对方位角或仰角值进行重映射。

[0194] 在下文中,提出了另外的实施例。

[0195] 图11示出了根据实施例的基于对象的音频的渲染。例如,可以使用元数据和与重放环境有关的信息在重放侧对音频对象进行渲染。这些信息例如是扬声器的个数或屏幕的尺寸。渲染器1110可以例如基于几何数据和可用扬声器及其位置来计算扬声器信号。

[0196] 现在,参考图12描述根据实施例的对象元数据(预)处理器1210。

[0197] 在图12中,对象元数据处理器1210配置为进行重映射,所述重映射考虑局部屏幕尺寸并且执行几何元数据的映射。

[0198] 通过对象元数据处理器1210对屏幕相关对象的位置数据进行重映射。例如,可以应用曲线,所述曲线将所述位置的原始方位角和仰角映射至重映射的方位角和重映射的仰角。

[0199] 例如可以考虑在混合和监测过程中采用的额定基准屏幕的屏幕尺寸以及重放室中局部屏幕尺寸信息用于重映射。

[0200] 可以例如在元数据中传输被称作再现屏幕尺寸的基准屏幕尺寸。

[0201] 在一些实施例中,如果没有给出额定屏幕尺寸,例如可以假设缺省屏幕尺寸。

[0202] 例如,可以使用在ITU-R REC-BT.2022(参见:用于屏幕显示器上的SDTV和HDTV电视图像质量的主观评估的通用观看条件)中定义的观看角度。

[0203] 在一些实施例中,例如可以限定两个线性映射函数用于仰角和方位角值的重映射。

[0204] 在下文中,将参考图13-图15来描述根据一些实施例的屏幕相关几何元数据改进。

[0205] 重映射的方位角可以取-180°和180°之间的值,并且重映射的仰角可以取-90°和90°之间的值。映射曲线通常地定义,使得将缺省左边缘方位角和缺省右边缘方位角之间的方位角值映射(压缩或扩大)到给定的屏幕左边缘和给定的屏幕右边缘(相应地用于仰角)。相应地压缩或扩大其他方位角和仰角,使得覆盖整个范围的值。

[0206] 如上文已经描述的,额定屏幕尺寸的屏幕边缘可以例如通过下式给出:

[0207] $\varphi_{left}^{\text{nominal}}, \varphi_{right}^{\text{nominal}}, \theta_{top}^{\text{nominal}}, \theta_{bottom}^{\text{nominal}}$。

[0208] 再现屏幕边缘可以例如通过下式来简化:

[0209] $\varphi_{left}^{repro}, \varphi_{right}^{repro}, \theta_{top}^{repro}, \theta_{bottom}^{repro}$。

[0210] 方位角和仰角位置数据的重映射可以例如通过以下的线性映射函数来定义:

[0211] $$\varphi' = \begin{cases} \dfrac{\varphi_{right}^{repro} + 180°}{\varphi_{right}^{\text{nominal}} + 180°} \cdot (\varphi + 180°) - 180° & for \quad -180° \le \varphi < \varphi_{right}^{\text{nominal}} \\ \dfrac{\varphi_{left}^{repro} - \varphi_{right}^{repro}}{\varphi_{left}^{\text{nominal}} - \varphi_{right}^{\text{nominal}}} \cdot (\varphi - \varphi_{right}^{\text{nominal}}) + \varphi_{right}^{repro} & for \quad \varphi_{right}^{\text{nominal}} \le \varphi < \varphi_{left}^{\text{nominal}} \\ \dfrac{180° - \varphi_{left}^{repro}}{180° - \varphi_{left}^{\text{nominal}}} \cdot (\varphi - \varphi_{left}^{\text{nominal}}) + \varphi_{left}^{repro} & for \quad \varphi_{left}^{\text{nominal}} \le \varphi < 180° \end{cases},$$

[0212] $$\theta' = \begin{cases} \dfrac{\theta_{bottom}^{repro} + 90°}{\theta_{bottom}^{\text{nominal}} + 90°} \cdot (\theta + 90°) - 90° & for \quad -90° \le \theta < \theta_{bottom}^{\text{nominal}} \\ \dfrac{\theta_{top}^{repro} - \theta_{bottom}^{repro}}{\theta_{top}^{\text{nominal}} - \theta_{bottom}^{\text{nominal}}} \cdot (\theta - \theta_{bottom}^{\text{nominal}}) + \theta_{bottom}^{repro} & for \quad \theta_{bottom}^{\text{nominal}} \le \theta < \theta_{top}^{\text{nominal}} \\ \dfrac{90° - \theta_{top}^{repro}}{90° - \theta_{top}^{\text{nominal}}} \cdot (\theta - \theta_{top}^{\text{nominal}}) + \theta_{top}^{repro} & for \quad \theta_{top}^{\text{nominal}} \le \theta < 90° \end{cases}。$$

[0213] 在图13中描述了针对方位角的映射函数,以及在图14中描述了针对仰角的映射函数。

[0214] 可以将曲线的梯度变化的点 $\varphi_{left}^{\text{nominal}}, \varphi_{right}^{\text{nominal}}, \theta_{top}^{\text{nominal}}, \theta_{bottom}^{\text{nominal}}$ 定义为缺省值(缺省的假设标准屏幕尺寸和缺省的假设标准屏幕位置),或者可以在元数据中出现所述点(例如,通过可以给出生产/监测屏幕尺寸的生产者)。

[0215] 关于用于屏幕相关重映射的对象元数据的定义,为了控制屏幕相关的渲染,定义了称作“isScreenRelativeObject”的附加元数据标志。这种标志可以例如限定是否应该与局部再现屏幕尺寸相关地处理/渲染音频对象。

[0216] 如果在音频场景中存在屏幕相关的元件,那么提供一种可能性来提供额定基准屏幕的屏幕尺寸信息,所述屏幕尺寸信息用于混合和监测(在音频内容的生产期间使用的屏幕尺寸)。

[0217] 表4-根据实施例的ObjectMetadataConfig()的语义:

[0218]

语义	比特个数	记忆值
ObjectMetadataConfig()		
{		
...		
hasScreenRelativeObjects;	**1**	**bslbf**
if(hasScreenRelatedObjects) {		
hasScreenSize;	**1**	**bslbf**
if(hasScreenSize) {		
bsScreenSizeAz;	**9**	**uimsbf**
bsScreenSizeTopEl;	**9**	**uimsbf**
bsScreenSizeBottomEl;	**9**	**uimsbf**
}		
for (o = 0; o <= num_objects-1; o++) {		
isScreenRelativeObject[o];	**1**	**bslbf**
}		
}		
}		

[0219] hasScreenRelativeObjects这一标志规定了是否存在屏幕相关的对象。

[0220] hasScreenSize这一标志规定了是否定义了额定屏幕尺寸。经由与平面边缘相对应的观看角度来进行这种定义。在hasScreenSize为零的情况下,缺省地使用以下的值:

[0221] $\varphi_{left}^{nominal} = 29.0°,$

[0222] $\varphi_{right}^{nominal} = -29.0°,$

[0223] $\theta_{top}^{nominal} = 17.5°,$

[0224] $\theta_{bottom}^{nominal} = -17.5°。$

[0225] bsScreenSizeAz这一字段定义了与屏幕左右边缘相对应的方位角:

[0226] $\varphi_{left}^{nominal} = 0.5 \cdot \text{bsScreenSizeAz}$

[0227] $\varphi_{left}^{nominal} = \min\left(\max\left(\varphi_{left}^{nominal}, 0\right), 180\right);$

[0228] $\varphi_{right}^{nominal} = -0.5 \cdot \text{bsScreenSizeAz},$

[0229] $\varphi_{right}^{nominal} = \min\left(\max\left(\varphi_{right}^{nominal}, -180\right), 0\right)。$

[0230] bsScreenSizeTopEl这一字段定义了与顶部平面边缘相对应的仰角:

[0231] $\theta_{top}^{nominal} = 0.5 \cdot \text{bsScreenSizeTopEl} - 255,$

[0232] $\theta_{top}^{nominal} = \min\left(\max\left(\theta_{top}^{nominal}, -90\right), 90\right);$

[0233] bsScreenSizeBottomEl这一字段定义了与屏幕底部边缘相对应的仰角:

[0234] $\theta_{bottom}^{nominal} = 0.5 \cdot \text{bsScreenSizeBottomEl} - 255$

[0235] $\theta_{bottom}^{nominal} = \min\ (\max\ (\theta_{bottom}^{nominal},\ -90),\ 90);$

[0236] isScreenRelativeObject这一标志定义了对象位置是屏幕相关的(应该不同地渲染所述位置,使得对它们的位置重映射,但是可以仍然包含所有有效的角度值)。

[0237] 根据实施例,如果没有给出再现屏幕尺寸,或者假设缺省的再现屏幕尺寸和缺省的再现屏幕位置,或者即使将对象标记为屏幕相关的也不施加映射。

[0238] 所述实施例的一些实现了可能的变体。

[0239] 在一些实施例中,采用非线性映射函数。这些映射函数可能不包括线性部分,而代替地是曲线的。在一些实施例中,附加的元数据控制重映射的方式,例如限定偏移或非线性系数以考虑听觉的摇摄行为或者分辨率。

[0240] 一些实施例实现了方位角和仰角的独立处理。可以将方位角和仰角独立地标记和处理为屏幕相关的。

[0241] 表5:示出了根据这种实施例的ObjectMetadataConfig()的语义:

[0242]

语义	比特个数	记忆值
ObjectMetadataConfig() { ...		
hasScreenRelatedObjects;	**1**	**bslbf**
if(hasScreenRelatedObjects) { ...		
for (o = 0; o <= num_objects-1; o++) {		
AzimuthScreenRelated[o];	**1**	**bslbf**
ElevationScreenRelated[o];	**1**	**bslbf**
} } }		

[0243] 一些实施例采用屏幕上对象的定义。可以在屏幕相关的对象和屏幕上对象之间进行区分。然后一种可能的语义是表6的以下内容:

[0244] 表6-根据实施例的ObjectMetadataConfig()的语义:

[0245]

语义	比特个数	记忆值
ObjectMetadataConfig() { ...		
hasScreenRelatedObjects;	1	bslbf
if(hasScreenRelatedObjects) { ...		
for (o = 0; o <= num_objects-1; o++) { **isScreenRelativeObject**[o];	1	bslbf
if(!isScreenRelativeObject){ **isOnScreenObject**[o];	1	bslbf
} }		
}		
}		

[0246]

[0247] hasOnScreenObjects这一标志规定了是否存在屏幕相关的对象。

[0248] isScreenRelatedObject这一标志限定了对象位置是否是屏幕相关的(应该不同地渲染所述位置,使得重映射它们的位置,但是所述位置仍然包含所有的有效角度值)。

[0249] isOnScreenObject这一标志限定了相应的对象是否是“屏幕上的”。应该不同地渲染这一标志等于1的对象,使得它们的位置值可以取屏幕区域上的值。

[0250] 对于屏幕上对象,重映射的方位角和仰角只可以取描述屏幕上的位置的值($\varphi_{left}^{repro} \le \varphi' \le \varphi_{right}^{repro}$ 并且 $\theta_{top}^{repro} \ge \theta' \ge \theta_{bottom}^{repro}$)。

[0251] 如通过一些实施例实现的,存在处理这些范围之外的值的不同可能性:可以将这些范围之外的值映射到屏幕的边缘。在左半球上,将180°和$180° - \varphi_{left}^{nominal}$之间的位置映射到屏幕左边缘$\varphi_{left}^{nominal}$。按照相同的方式处理右半球和仰角(图15中的非虚线映射函数1510)。

[0252] 由所述实施例的一些实现的另一种可能性是将后半球的值映射到前半球。将180°和$180° - \varphi_{left}^{nominal}$之间的值映射到0°和$\varphi_{left}^{repro}$之间的值。按照相同的方式处理右半球和仰角(图15总的虚线映射函数1520)。

[0253] 图15示出了根据这些实施例的(屏幕上对象)方位角的重映射。

[0254] 可以通过附加的元数据(例如,用于将后部[180°和$180° - \varphi_{left}^{nominal}$]以及[-180°and-$180° - \varphi_{right}^{nominal}$]所需的所有屏幕上对象“投影”到屏幕上)信道进行所需行为的选择。

[0255] 虽然已经在装置的上下文中描述了一些方面,但是将清楚的是,这些方面还表示对相应方法的描述,其中框或设备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案也表示对相应块或项或者相应装置的特征的描述。

[0256] 新颖的分解信号可以存储在数字存储介质上,或者可以在诸如无线传输介质或有线传输介质(例如,互联网)等的传输介质上传输。

[0257] 取决于某些实现要求,可以在硬件中或在软件中实现本发明的实施例。可以使用其上存储有电子可读控制信号的数字存储介质(例如,软盘、DVD、CD、ROM、PROM、EPROM、EEPROM或闪存)来执行该实现,该电子可读控制信号与可编程计算机系统协作(或者能够与之协作)从而执行相应方法。

[0258] 根据本发明的一些实施例包括具有电子可读控制信号的非临时数据载体,该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所述的方法之一。

[0259] 通常,本发明的实施例可以实现为具有程序代码的计算机程序产品,程序代码可操作以在计算机程序产品在计算机上运行时执行方法之一。程序代码可以例如存储在机器可读载体上。

[0260] 其他实施例包括存储在机器可读载体上的计算机程序,该计算机程序用于执行本文所述的方法之一。

[0261] 换言之,本发明方法的实施例因此是具有程序代码的计算机程序,该程序代码用于在计算机程序在计算机上运行时执行本文所述的方法之一。

[0262] 因此,本发明方法的另一实施例是其上记录有计算机程序的数据载体(或者数字存储介质或计算机可读介质),该计算机程序用于执行本文所述的方法之一。

[0263] 因此,本发明方法的另一实施例是表示计算机程序的数据流或信号序列,所述计算机程序用于执行本文所述的方法之一。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,经由互联网)传递。

[0264] 另一实施例包括处理装置,例如,计算机或可编程逻辑器件,所述处理装置被配置为或适于执行本文所述的方法之一。

[0265] 另一实施例包括其上安装有计算机程序的计算机,该计算机程序用于执行本文所述的方法之一。

[0266] 在一些实施例中,可编程逻辑器件(例如,现场可编程门阵列)可以用于执行本文所述的方法的功能中的一些或全部。在一些实施例中,现场可编程门阵列可以与微处理器协作以执行本文所述的方法之一。通常,方法优选地由任意硬件装置来执行。

[0267] 上述实施例对于本发明的原理仅是说明性的。应当理解的是:本文所述的布置和细节的修改和变形对于本领域其他技术人员将是显而易见的。因此,旨在仅由所附专利权利要求的范围来限制而不是由借助对本文的实施例的描述和解释所给出的具体细节来限制。

[0268] 【实施方式1】.一种用于产生扬声器信号的设备,包括:

[0269] 对象元数据处理器(110),以及

[0270] 对象渲染器(120),

[0271] 其中所述对象渲染器(120)配置为接收音频对象,

[0272] 其中所述对象元数据处理器(110)配置为接收元数据,所述元数据包括与所述音频对象是否是屏幕相关的有关的指示,并且还包括所述音频对象的第一位置,

[0273] 其中所述对象元数据处理器(110)配置为如果在所述元数据中指示所述音频对象是屏幕相关的,则依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频

对象的第二位置，

[0274] 其中所述对象渲染器(120)配置为依赖于所述音频对象并且依赖于位置信息来产生所述扬声器信号，

[0275] 其中所述对象元数据处理器(110)配置为如果在所述元数据中指示所述音频对象不是屏幕相关的，则将所述音频对象的第一位置作为所述位置信息馈送至所述对象渲染器(120)，以及

[0276] 其中所述对象元数据处理器(110)配置为如果在所述元数据中指示所述音频对象是屏幕相关的，则将所述音频对象的第二位置作为所述位置信息馈送至所述对象渲染器(120)中。

[0277] 【实施方式2】.根据实施方式1所述的设备，其中所述对象元数据处理器(110)配置为如果在所述元数据中指示所述音频对象不是屏幕相关的，则不计算所述音频对象的第二位置。

[0278] 【实施方式3】.根据实施方式1所述的设备，其中所述对象渲染器(120)配置为不对所述位置信息是所述音频对象的第一位置还是所述音频对象的第二位置进行确定。

[0279] 【实施方式4】.根据实施方式1所述的设备，其中所述对象渲染器(120)配置为进一步依赖于重放环境的扬声器的个数来产生扬声器信号。

[0280] 【实施方式5】.根据实施方式4所述的设备，其中所述对象渲染器(120)配置为进一步依赖于所述重放环境的每一个扬声器的扬声器位置来产生扬声器信号。

[0281] 【实施方式6】.根据实施方式1所述的设备，其中所述对象元数据处理器(110)配置为如果在所述元数据中指示所述音频对象是屏幕相关的，则依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置，其中所述第一位置指示三维空间中的第一位置，并且其中所述第二位置指示三维空间中的第二位置。

[0282] 【实施方式7】.根据实施方式6所述的设备，其中所述对象元数据处理器(110)配置为如果在所述元数据中指示所述音频对象是屏幕相关的，则依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置，其中所述第一位置指示第一方位角、第一仰角和第一距离，并且其中所述第二位置指示第二方位角、第二仰角和第二距离。

[0283] 【实施方式8】.根据实施方式1所述的设备，

[0284] 其中所述对象元数据处理器(110)配置为接收所述元数据，所述元数据包括表示所述音频对象是否是屏幕相关的指示作为第一指示，并且还包括第二指示，如果所述音频对象是屏幕相关的，则所述第二指示表示所述音频对象是否是屏幕上对象，以及

[0285] 其中所述对象元数据处理器(110)配置为依赖于所述音频对象的第一位置并且依赖于所述屏幕的尺寸来计算所述音频对象的第二位置，使得如果所述第二指示表示所述音频对象是屏幕上对象，则所述第二位置取所述屏幕的屏幕区域上的第一值。

[0286] 【实施方式9】.根据实施方式8所述的设备，其中所述对象元数据处理器(110)配置为依赖于所述音频对象的第一位置并且依赖于所述屏幕的尺寸来计算所述音频对象的第二位置，使得如果所述第二指示表示所述音频对象不是屏幕上对象，则所述第二位置取在屏幕区域上或者不在屏幕区域上的第二值。

[0287] 【实施方式10】.根据实施方式1所述的设备，

[0288] 其中所述对象元数据处理器(110)配置为接收元数据,所述元数据包括所述音频对象是屏幕相关的指示作为第一指示,并且还包括第二指示,如果所述音频对象是屏幕相关的,则所述第二指示表示所述音频对象是否是屏幕上对象,

[0289] 其中所述对象元数据处理器(110)配置为如果所述第二指示表示所述音频对象是屏幕上对象,则依赖于所述音频对象的第一位置、依赖于所述屏幕的尺寸并且依赖于作为映射曲线的第一映射曲线来计算所述音频对象的第二位置,其中所述第一映射曲线限定了第一值区间中的原始对象位置与第二值区间中的重映射对象位置的映射,以及

[0290] 其中所述对象元数据处理器(110)配置为如果所述第二指示表示所述音频对象不是屏幕上对象,则依赖于所述音频对象的第一位置、依赖于所述屏幕的尺寸并且依赖于第二映射曲线作为映射曲线来计算所述音频对象的第二位置,其中所述第二映射曲线限定了第一值区间中的原始对象位置与第三值区间中的重映射对象位置的映射,以及其中所述第二值区间小于所述第三值区间。

[0291] 【实施方式11】.根据实施方式10所述的设备,

[0292] 其中所述第一值区间、所述第二值区间和所述第三值区间的每一个是方位角的值区间,或者

[0293] 其中所述第一值区间、所述第二值区间和所述第三值区间的每一个是仰角的值区间。

[0294] 【实施方式12】.根据实施方式1所述的设备,

[0295] 其中所述对象元数据处理器(110)配置为依赖于第一线性映射函数和第二线性映射函数的至少一个来计算所述音频对象的第二位置,

[0296] 其中所述第一线性映射函数被定义为将第一方位角值映射至第二方位角值,

[0297] 其中所述第二线性映射函数被定义为将第一仰角值映射至第二仰角值,

[0298] 其中 $\varphi_{left}^{nominal}$ 指示左方位角屏幕边缘基准,

[0299] 其中 $\varphi_{right}^{nominal}$ 指示右方位角屏幕边缘基准,

[0300] 其中 $\theta_{top}^{nominal}$ 指示顶部仰角屏幕边缘基准,

[0301] 其中 $\theta_{bottom}^{nominal}$ 指示仰角屏幕边缘基准,

[0302] 其中 φ_{left}^{repro} 指示所述屏幕的左方位角屏幕边缘,

[0303] 其中 φ_{right}^{repro} 指示所述屏幕的右方位角屏幕边缘,

[0304] 其中 θ_{top}^{repro} 指示所述屏幕的顶部仰角屏幕边缘,

[0305] 其中 θ_{bottom}^{repro} 指示所述屏幕的底部仰角屏幕边缘,

[0306] 其中 φ 指示所述第一方位角值,

[0307] 其中 φ' 指示所述第二方位角值,

[0308] 其中θ指示所述第一仰角值,

[0309] 其中θ′指示所述第二仰角值,

[0310] 其中所述第二方位角值φ'根据第一线性映射函数

[0311] $$\varphi' = \begin{cases} \frac{\varphi_{right}^{repro}+180^\circ}{\varphi_{right}^{nominal}+180^\circ}\cdot(\varphi+180^\circ)-180^\circ & for \quad -180^\circ \le \varphi < \varphi_{right}^{nominal} \\ \frac{\varphi_{left}^{repro}-\varphi_{right}^{repro}}{\varphi_{left}^{nominal}-\varphi_{right}^{nominal}}\cdot(\varphi-\varphi_{right}^{nominal})+\varphi_{right}^{repro} & for \quad \varphi_{right}^{nominal} \le \varphi < \varphi_{left}^{nominal} \\ \frac{180^\circ-\varphi_{left}^{repro}}{180^\circ-\varphi_{left}^{nominal}}\cdot(\varphi-\varphi_{left}^{nominal})+\varphi_{left}^{repro} & for \quad \varphi_{left}^{nominal} \le \varphi < 180^\circ \end{cases}$$

[0312] 由第一方位角φ的第一映射产生,以及

[0313] 其中所述第二仰角值θ′根据第二线性映射函数

[0314] $$\theta' = \begin{cases} \frac{\theta_{bottom}^{repro}+90^\circ}{\theta_{bottom}^{nominal}+90^\circ}\cdot(\theta+90^\circ)-90^\circ & for \quad -90^\circ \le \theta < \theta_{bottom}^{nominal} \\ \frac{\theta_{top}^{repro}-\theta_{bottom}^{repro}}{\theta_{top}^{nominal}-\theta_{bottom}^{nominal}}\cdot(\theta-\theta_{bottom}^{nominal})+\theta_{bottom}^{repro} & for \quad \theta_{bottom}^{nominal} \le \theta < \theta_{top}^{nominal} \\ \frac{90^\circ-\theta_{top}^{repro}}{90^\circ-\theta_{top}^{nominal}}\cdot(\theta-\theta_{top}^{nominal})+\theta_{top}^{repro} & for \quad \theta_{top}^{nominal} \le \theta < 90^\circ \end{cases}$$

[0315] 由第一仰角值θ的第二映射产生。

[0316] 【实施方式13】.一种解码器装置,包括:

[0317] USAC解码器(910),用于对比特流进行解码以获得一个或多个音频输入声道,以获得一个或多个输入音频对象,从而获得压缩的对象元数据,并且获得一个或多个SAOC传输声道,

[0318] SAOC解码器(915),用于对一个或多个SAOC传输声道进行解码,以获得第一组一个或多个渲染的音频对象,

[0319] 根据前述实施方式之一所述的设备(917),其中所述设备包括:

[0320] 对象元数据解码器(918),是根据前述实施方式之一所述的设备的对象元数据处理器(110),并且实现为对压缩的对象元数据进行解码以获得未压缩的元数据,以及

[0321] 根据前述实施方式之一所述的设备的对象渲染器(920;120),用于依赖于未压缩的元数据对一个或多个输入音频数据进行渲染以获得第二组一个或多个渲染的音频对象,

[0322] 格式转换器(922),用于转换所述一个或多个音频输入声道以获得一个或多个转换的声道,以及

[0323] 混合器(930),用于将所述第一组一个或多个渲染的音频对象的一个或多个音频对象、所述第二组一个或多个渲染的音频对象的一个或多个音频对象以及所述一个或多个转换的声道进行混合,以获得一个或多个解码的音频声道。

[0324] 【实施方式14】.一种产生扬声器信号的方法,包括:

[0325] 接收音频对象,

[0326] 接收元数据,所述元数据包括所述音频对象是屏幕相关的指示,并且还包括所述

音频对象的第一位置，

[0327] 如果在所述元数据中指示所述音频对象是屏幕相关的，依赖于所述音频对象的第一位置并且依赖于屏幕的尺寸来计算所述音频对象的第二位置，

[0328] 依赖于所述音频对象并且依赖于位置信息来产生所述扬声器信号，

[0329] 其中如果在所述元数据中指示所述音频对象不是屏幕相关的，则所述位置信息是所述音频对象的第一位置，以及

[0330] 如果所述元数据指示所述音频对象是屏幕相关的，则所述位置信息是所述音频对象的第二位置。

[0331] 【实施方式15】.一种计算机程序，当在计算机或信号处理器上执行时，用于实现实施方式14所述的方法。

[0332] 文献

[0333] [1]“Method and apparatus for playback of a higher-order ambisonics audio signal”，专利申请号EP20120305271；

[0334] [2]“Vorrichtung und Verfahren zum Bestimmen einer Wiedergabeposition“，专利申请号WO2004073352A1；

[0335] [3]“Verfahren zur Audiocodierung”，专利申请号EP20020024643；

[0336] [4]“Acoustical Zooming Based on a Parametric Sound Field Representation”

[0337] http://www.aes.org/tmpFiles/elib/20140814/15417.pdf。

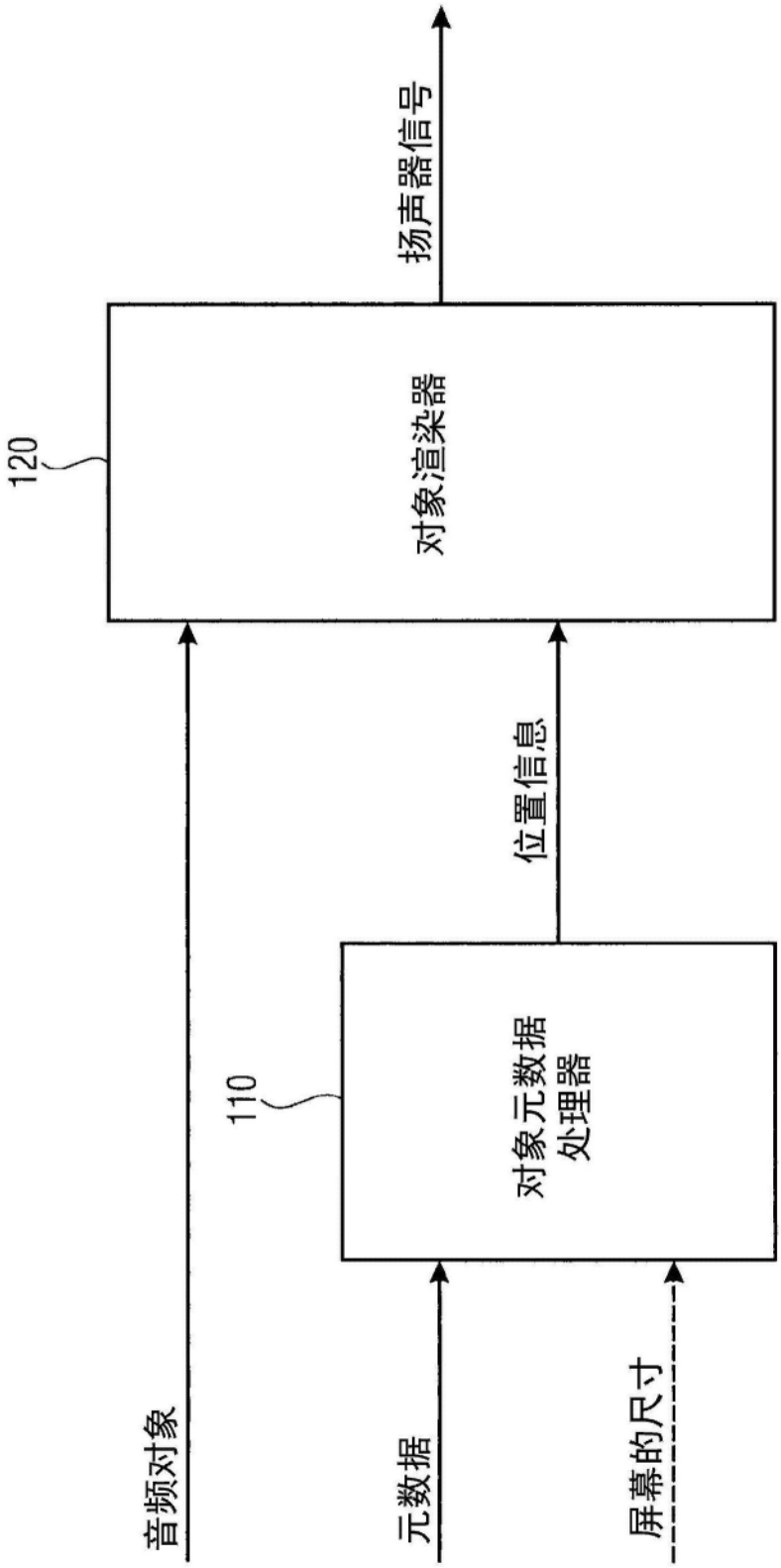


图1

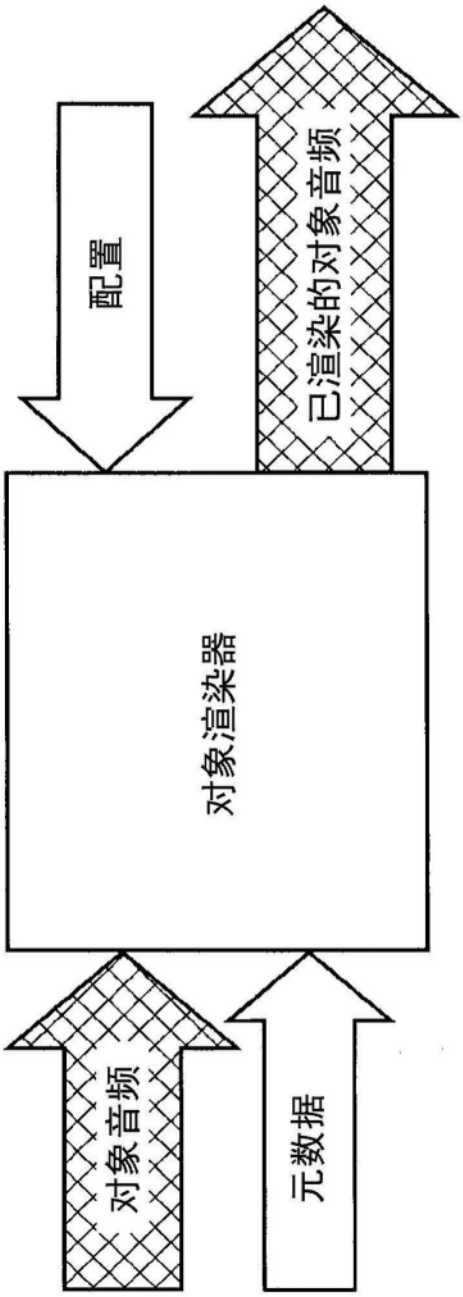


图2

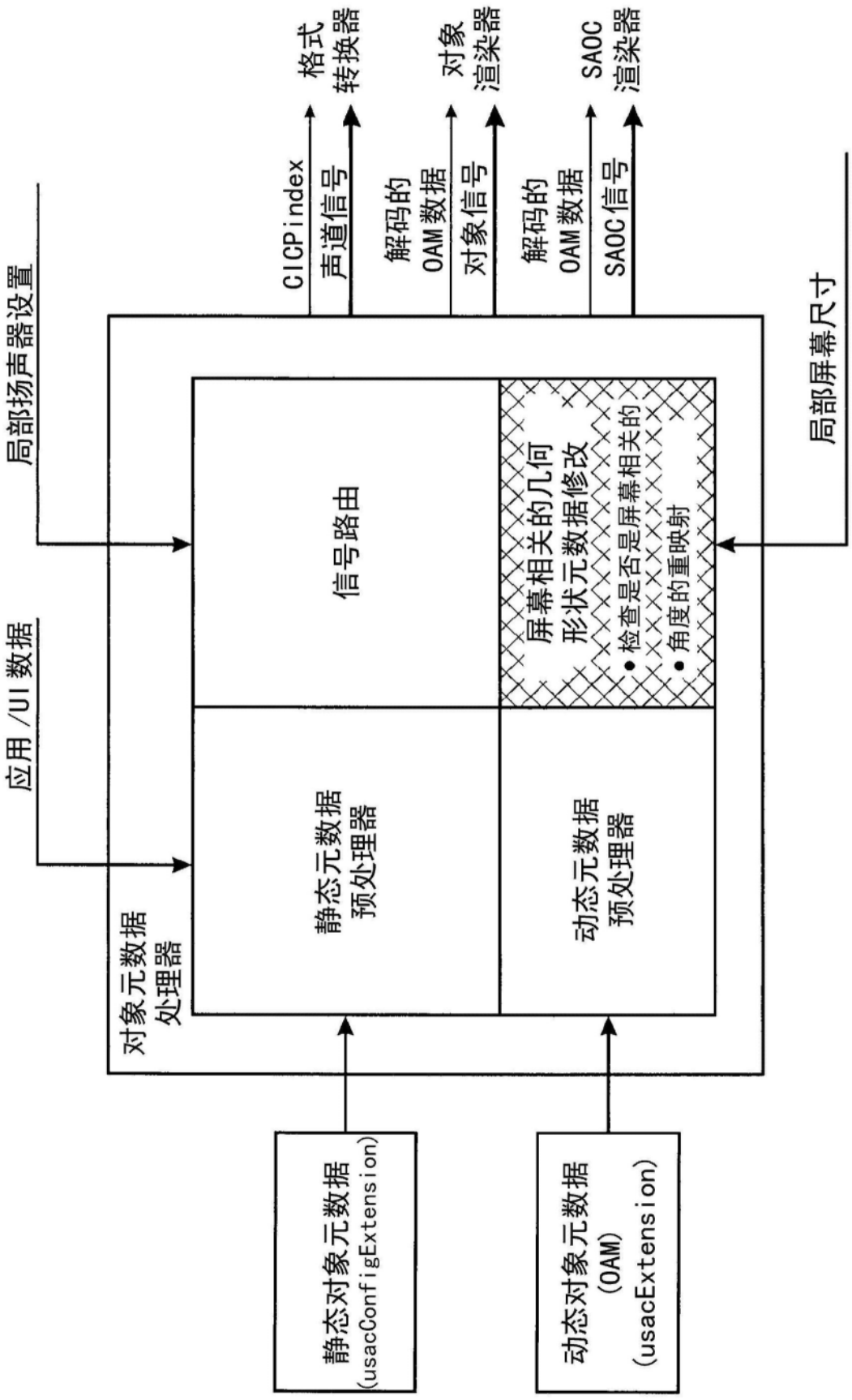

静态对象元数据
(usacConfigExtension)
动态对象元数据
(OAM)
(usacExtension)
对象元数据
处理器
静态元数据
预处理器
动态元数据
预处理器
信号路由
屏幕相关的几何
形状元数据修改
• 检查是否是屏幕相关的
• 角度的重映射
应用 /UI 数据
局部扬声器设置
局部屏幕尺寸
CICPindex
声道信号
格式
转换器
解码的
OAM数据
对象信号
对象
渲染器
解码的
OAM数据
SAOC信号
SAOC
渲染器

图3

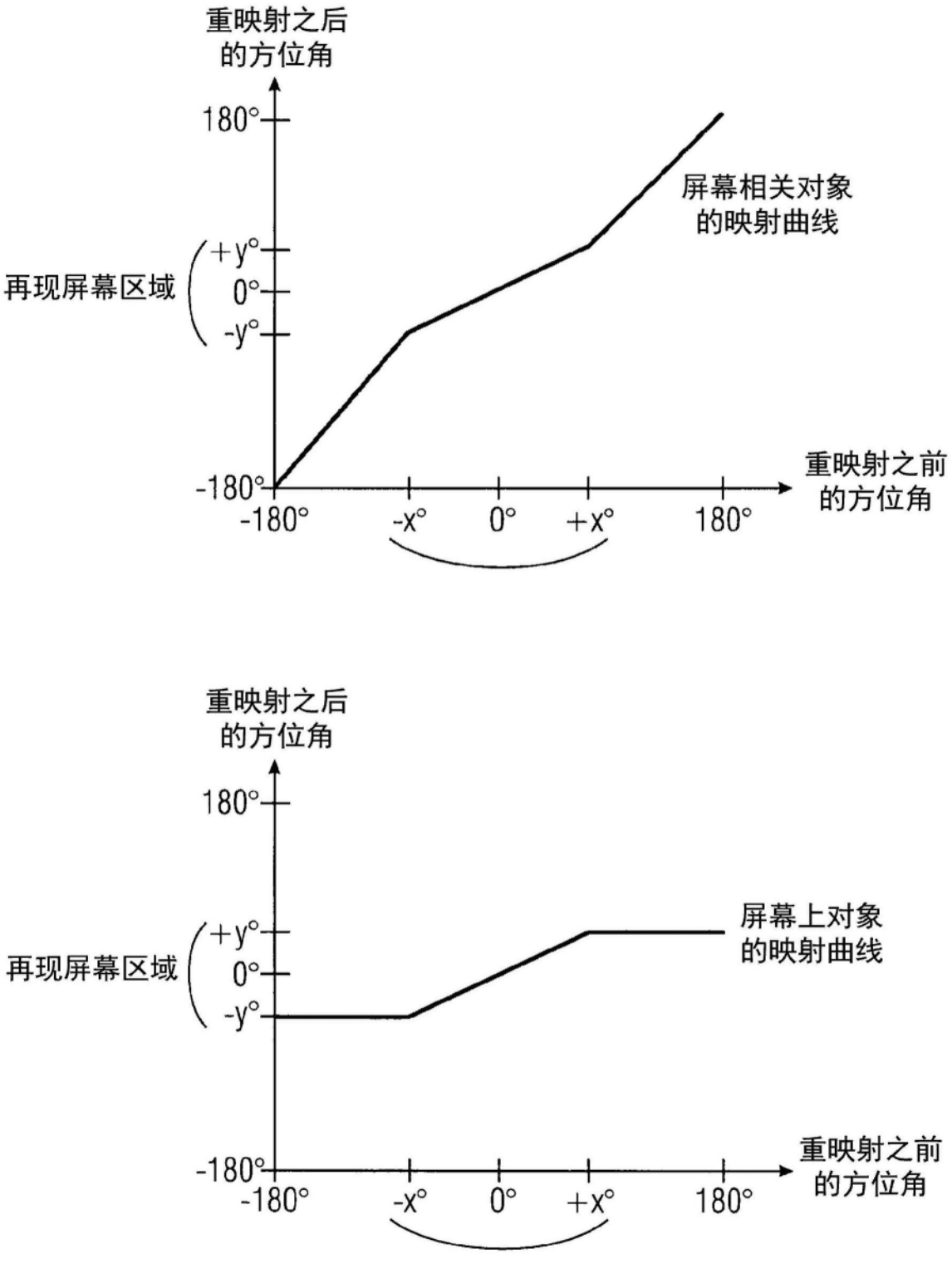

重映射之后
的方位角
180°
+y°
0°
-y°
再现屏幕区域
屏幕相关对象
的映射曲线
-180°
-180°
-x°
0°
+x°
180°
重映射之前
的方位角
重映射之后
的方位角
180°
+y°
0°
-y°
再现屏幕区域
屏幕上对象
的映射曲线
-180°
-180°
-x°
0°
+x°
180°
重映射之前
的方位角

图4

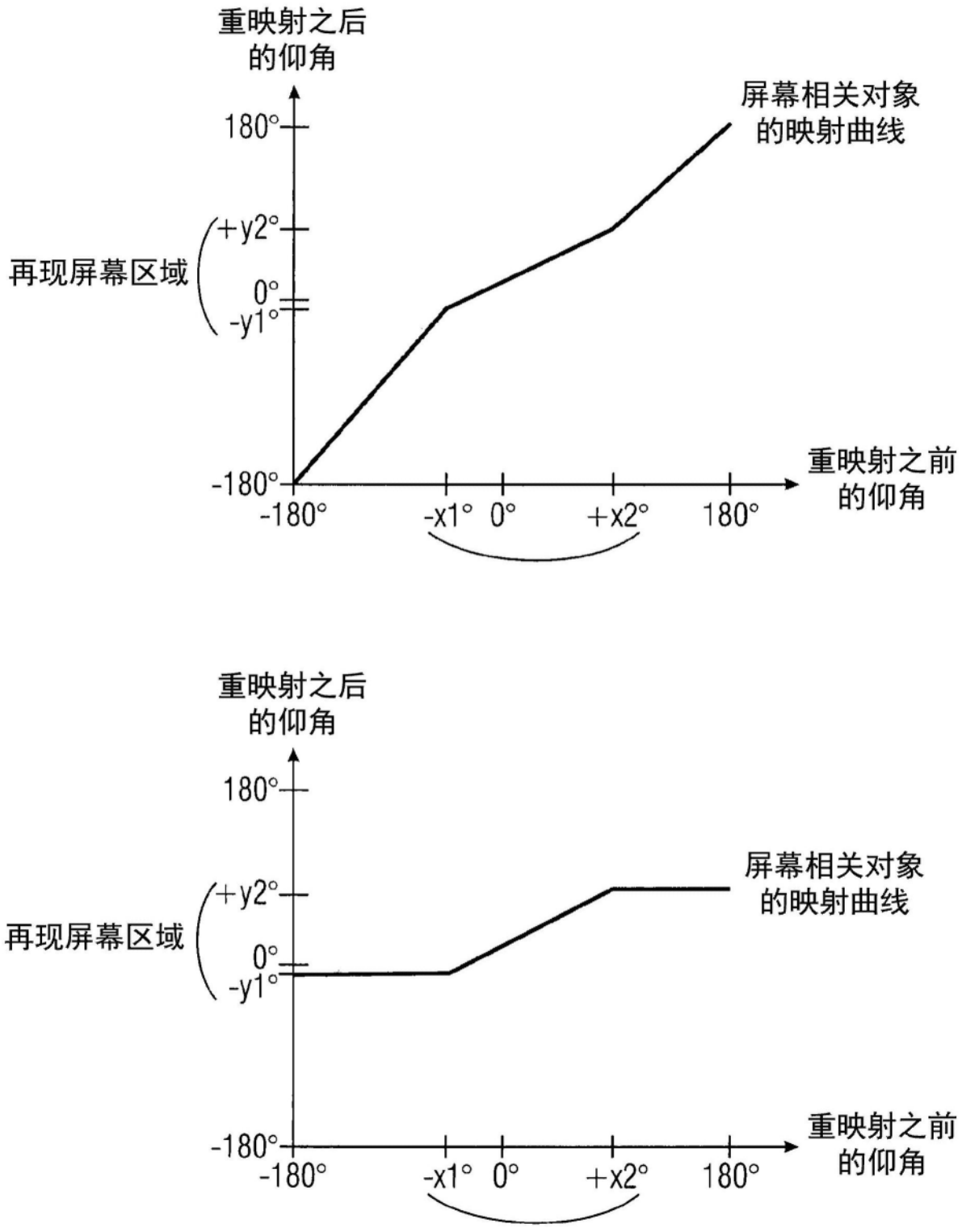

重映射之后
的仰角
180°
+y2°
再现屏幕区域
0°
-y1°
-180°
屏幕相关对象
的映射曲线
重映射之前
的仰角
-180°
-x1°
0°
+x2°
180°
重映射之后
的仰角
180°
+y2°
再现屏幕区域
0°
-y1°
-180°
屏幕相关对象
的映射曲线
重映射之前
的仰角
-180°
-x1°
0°
+x2°
180°

图5

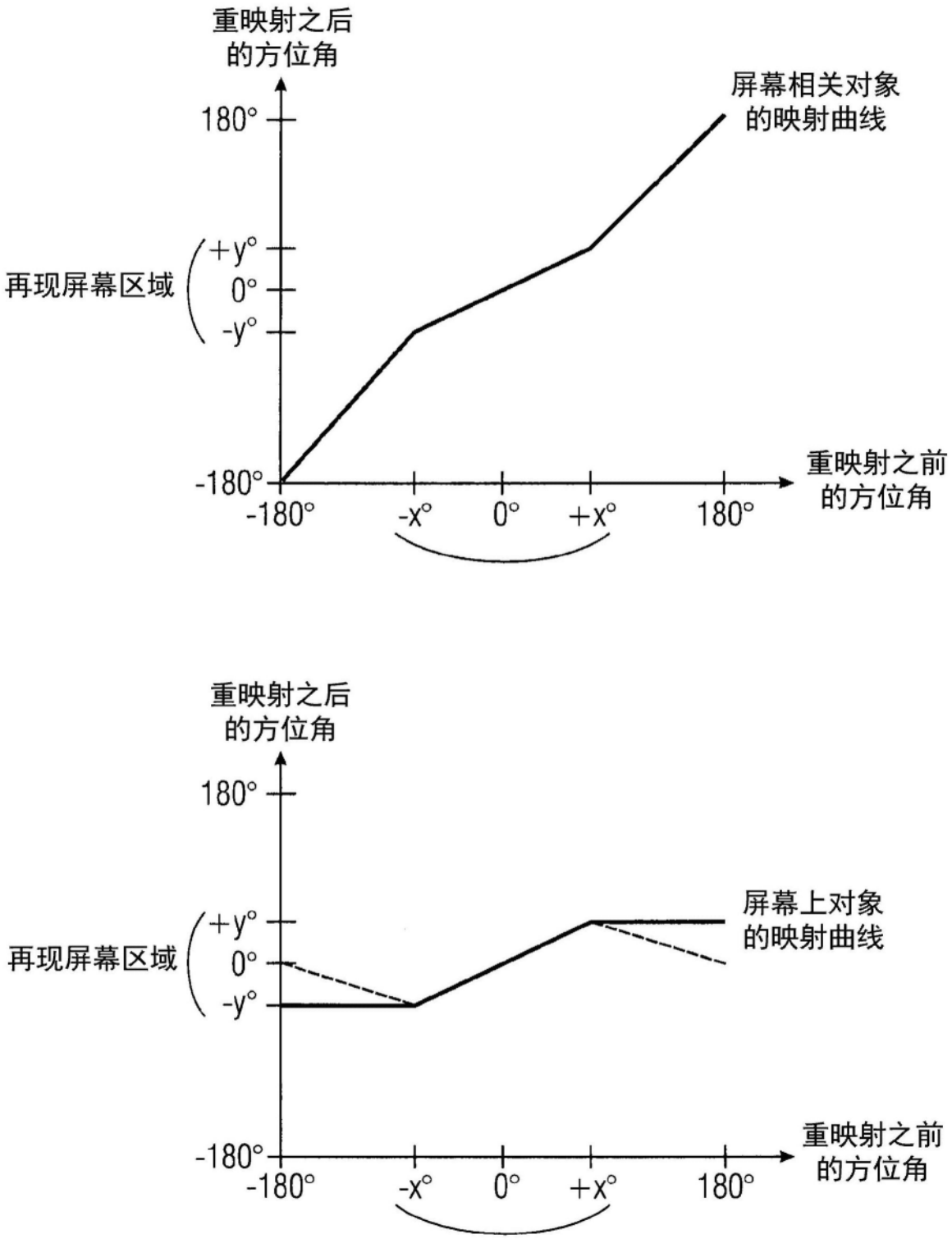

重映射之后
的方位角
180°
+y°
0°
-y°
再现屏幕区域
-180°
屏幕相关对象
的映射曲线
重映射之前
的方位角
-180°
-x°
0°
+x°
180°
重映射之后
的方位角
180°
+y°
0°
-y°
再现屏幕区域
屏幕上对象
的映射曲线
-180°
重映射之前
的方位角
-180°
-x°
0°
+x°
180°

图6

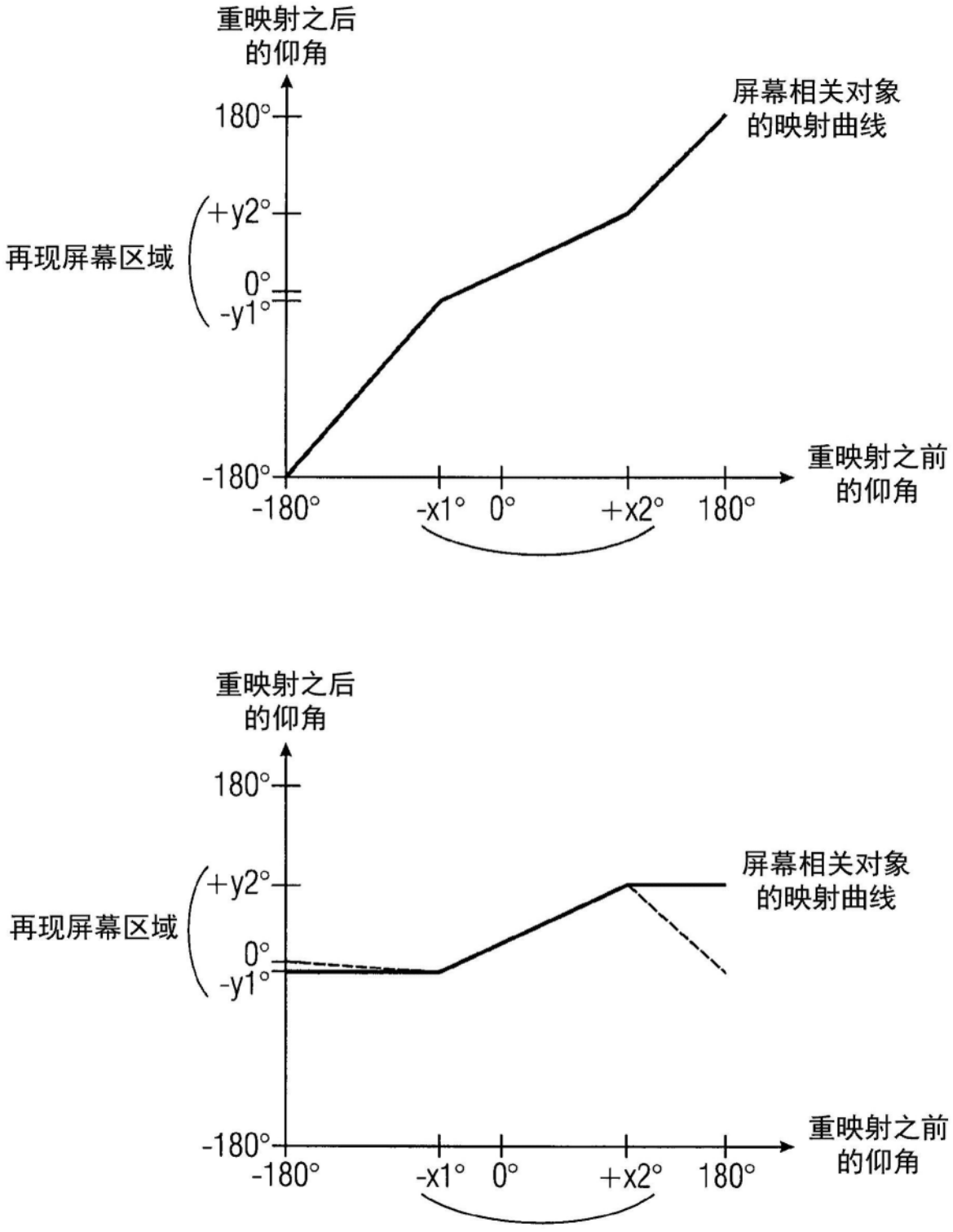

重映射之后
的仰角
180°
+y2°
0°
-y1°
再现屏幕区域
-180°
屏幕相关对象
的映射曲线
重映射之前
的仰角
-180°
-x1°
0°
+x2°
180°
重映射之后
的仰角
180°
+y2°
0°
-y1°
再现屏幕区域
屏幕相关对象
的映射曲线
-180°
重映射之前
的仰角
-180°
-x1°
0°
+x2°
180°

图7

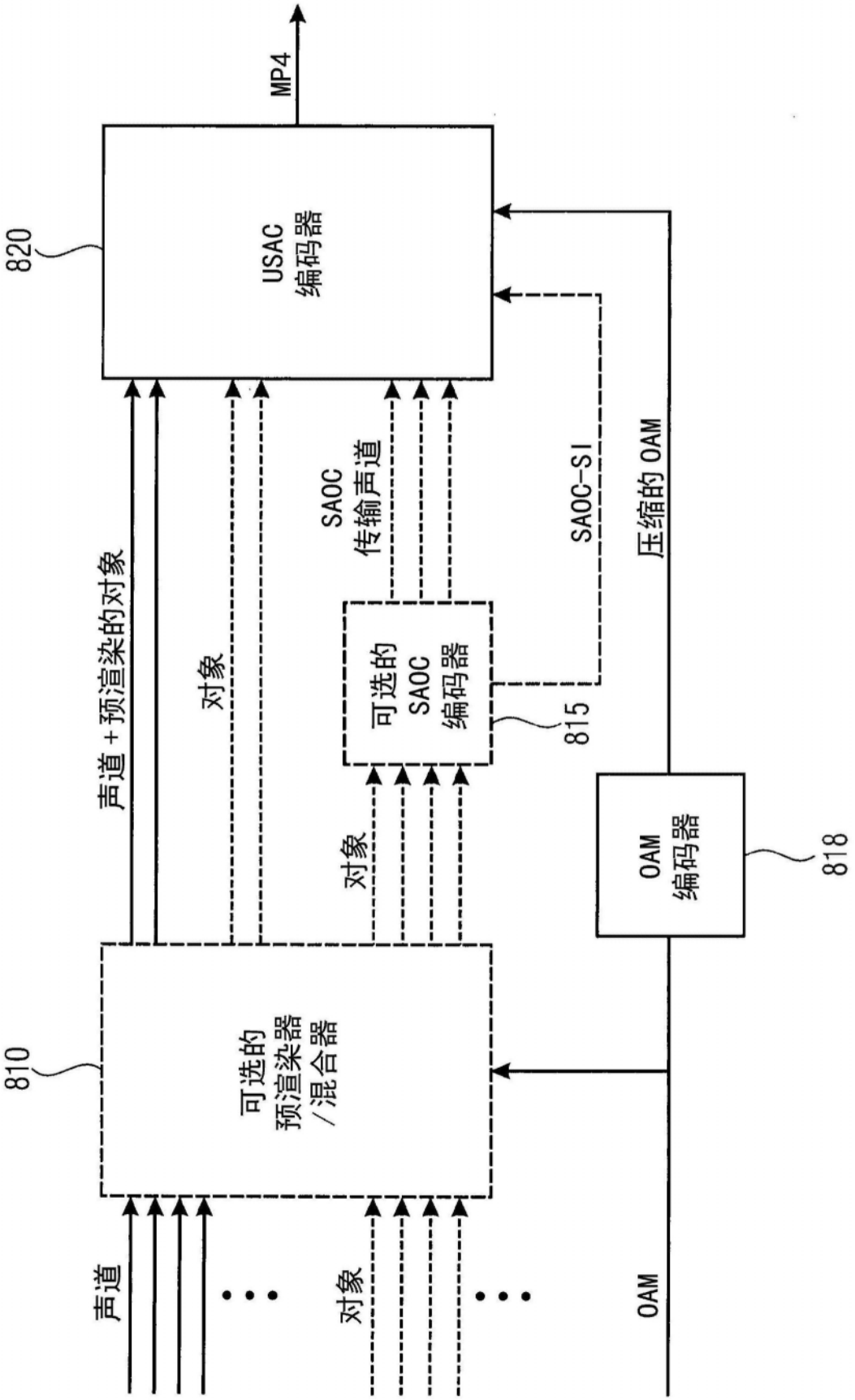

MP4
820
USAC
编码器
声道 + 预渲染的对象
对象
SAOC
传输声道
SAOC-SI
压缩的 OAM
可选的
SAOC
编码器
815
对象
OAM
编码器
818
810
可选的
预渲染器
/混合器
声道
对象
OAM

图8

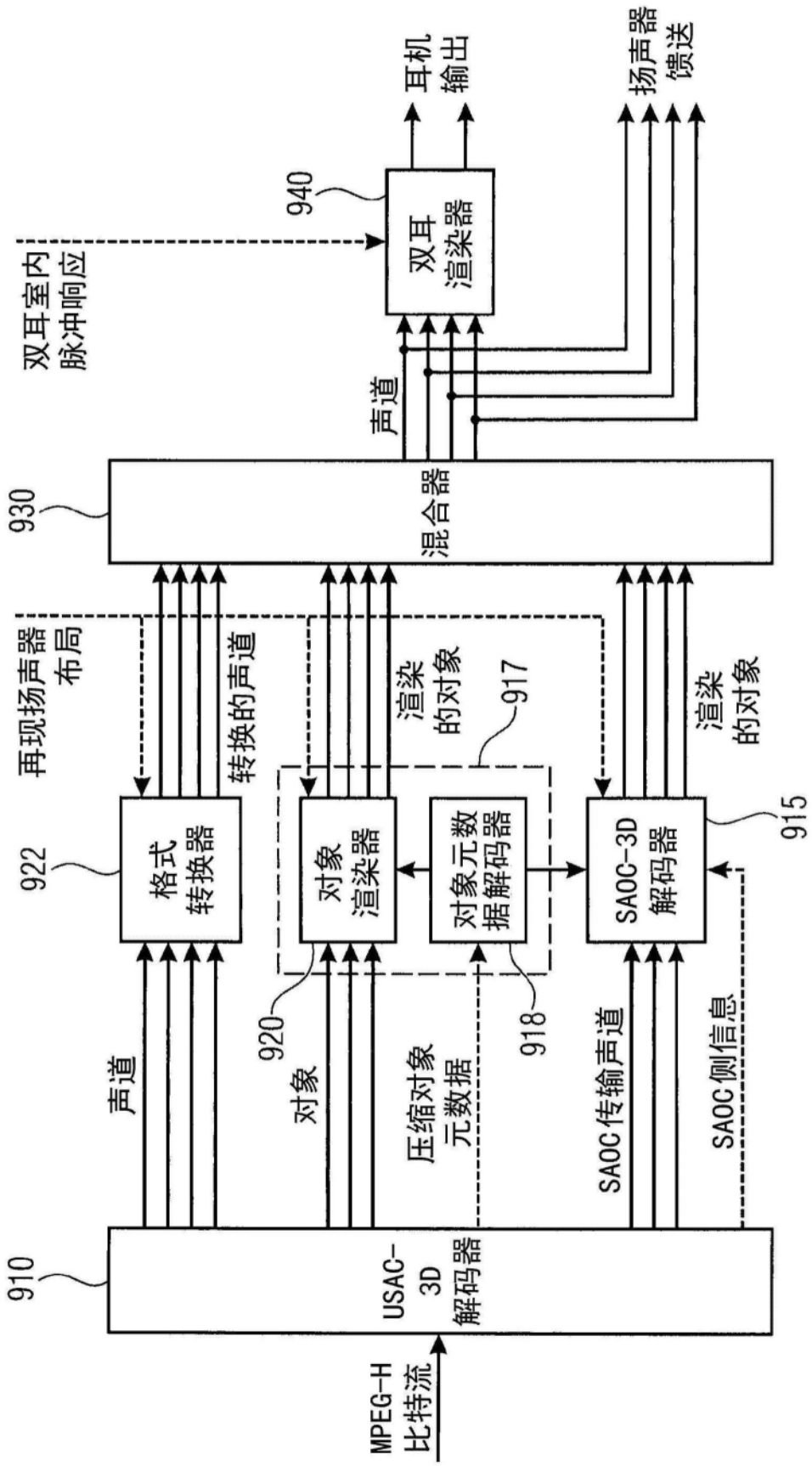

MPEG-H
比特流
910
USAC-
3D
解码器
声道
对象
920
压缩对象
元数据
918
SAOC传输声道
SAOC侧信息
922
格式
转换器
对象
渲染器
对象元数
据解码器
SAOC-3D
解码器
915
转换的声道
渲染
的对象
917
渲染
的对象
再现扬声器
布局
930
混合器
声道
双耳室内
脉冲响应
940
双耳
渲染器
耳机
输出
扬声器
馈送

图9

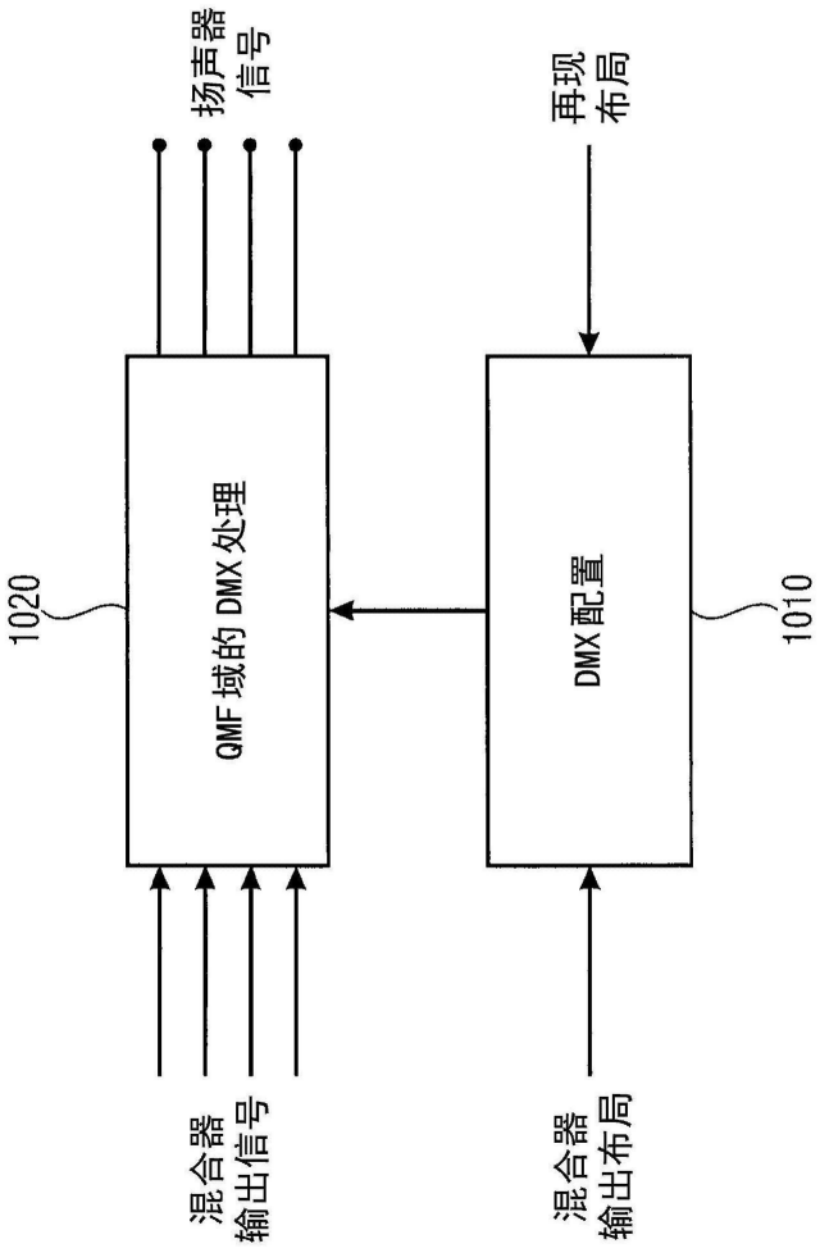


图10

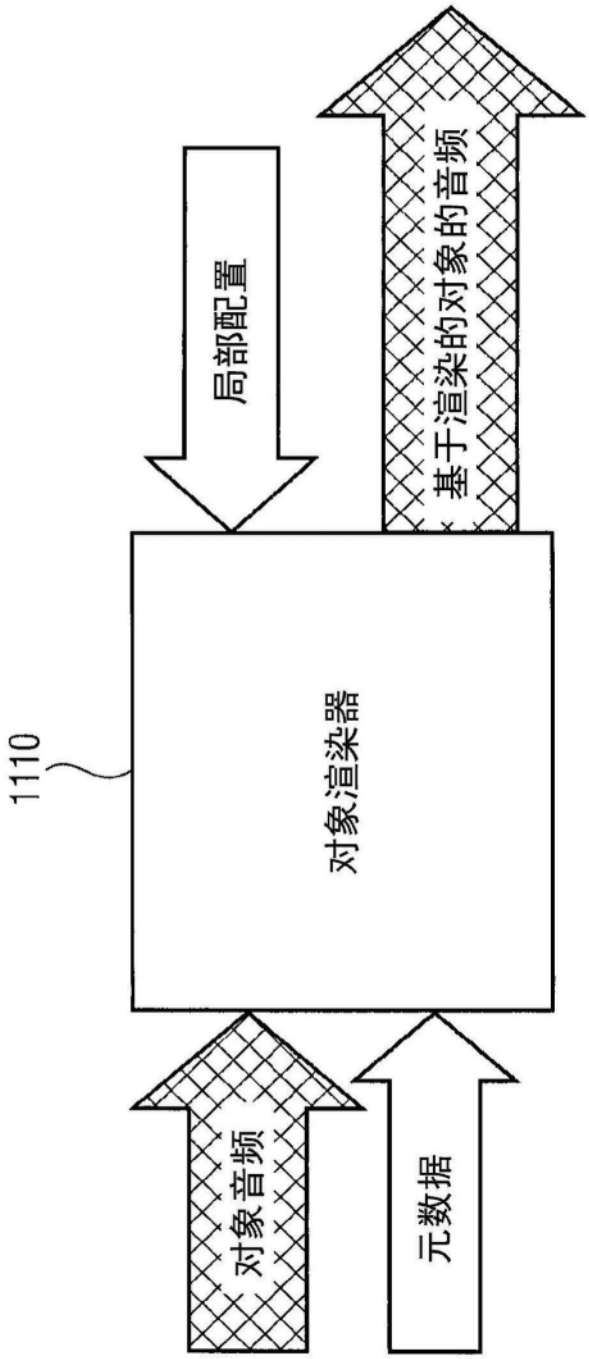


图11

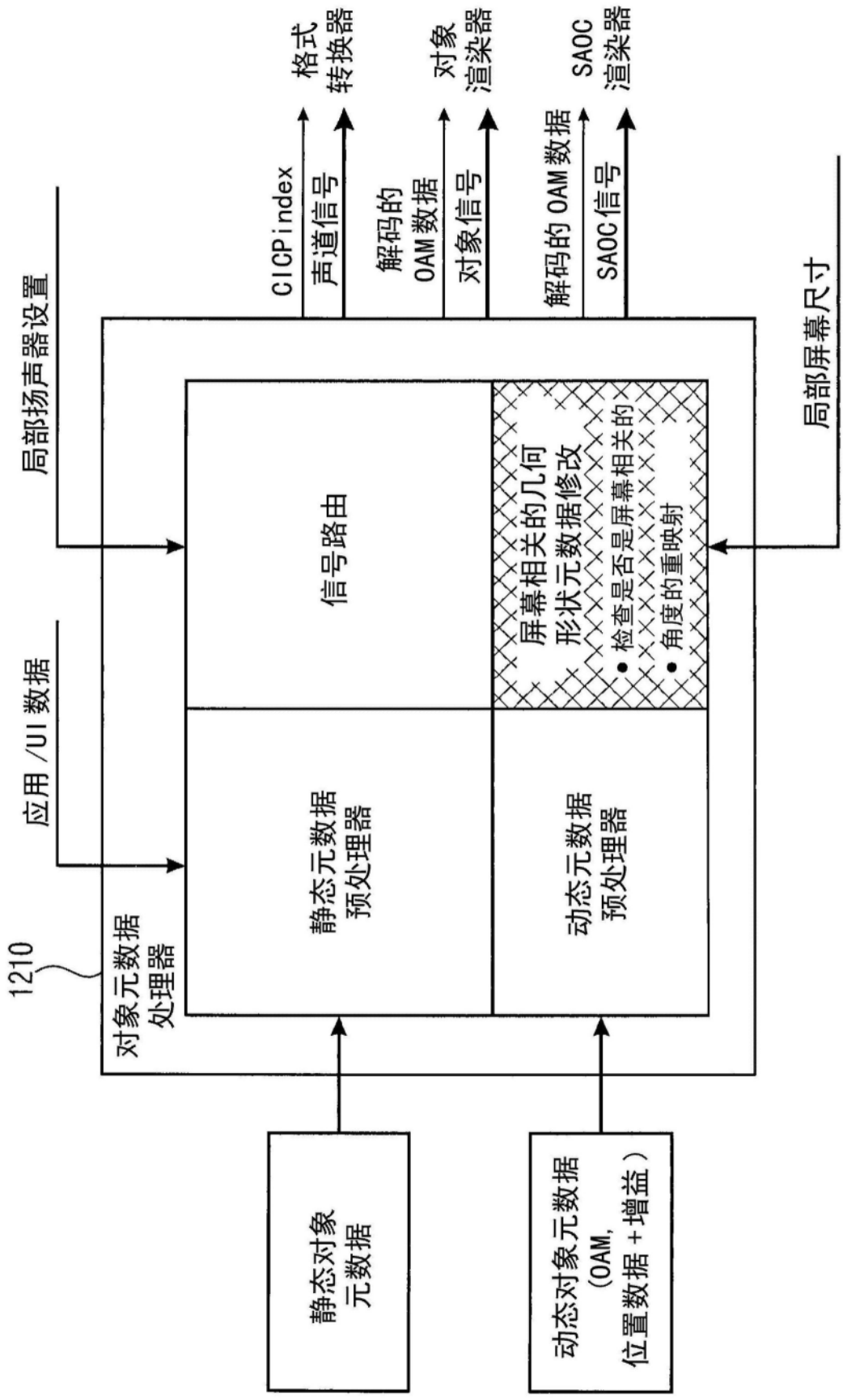

格式转换器
对象渲染器
SAOC渲染器
CICPindex
声道信号
解码的OAM数据
对象信号
解码的 OAM 数据
SAOC信号
局部扬声器设置
局部屏幕尺寸
应用 /UI 数据
信号路由
屏幕相关的几何形状元数据修改
• 检查是否是屏幕相关的
• 角度的重映射
静态元数据预处理器
动态元数据预处理器
1210
对象元数据处理器
静态对象元数据
动态对象元数据
(OAM,
位置数据 + 增益)

图12

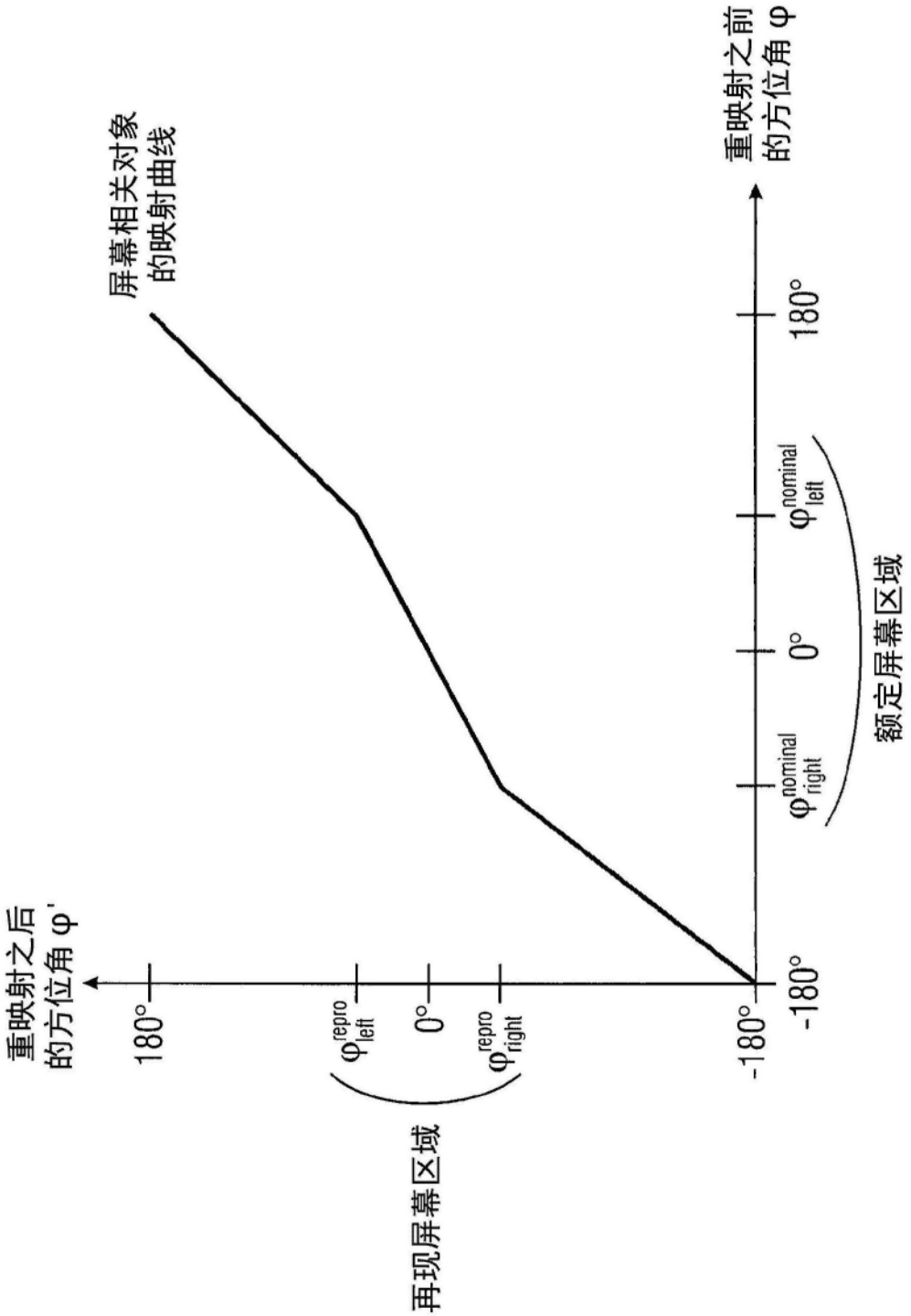

屏幕相关对象
的映射曲线
重映射之后
的方位角 φ'
180°
φ_{left}^{repro}
0°
φ_{right}^{repro}
再现屏幕区域
-180°
-180°
$\varphi_{right}^{nominal}$
0°
$\varphi_{left}^{nominal}$
180°
重映射之前
的方位角 φ
额定屏幕区域

图13

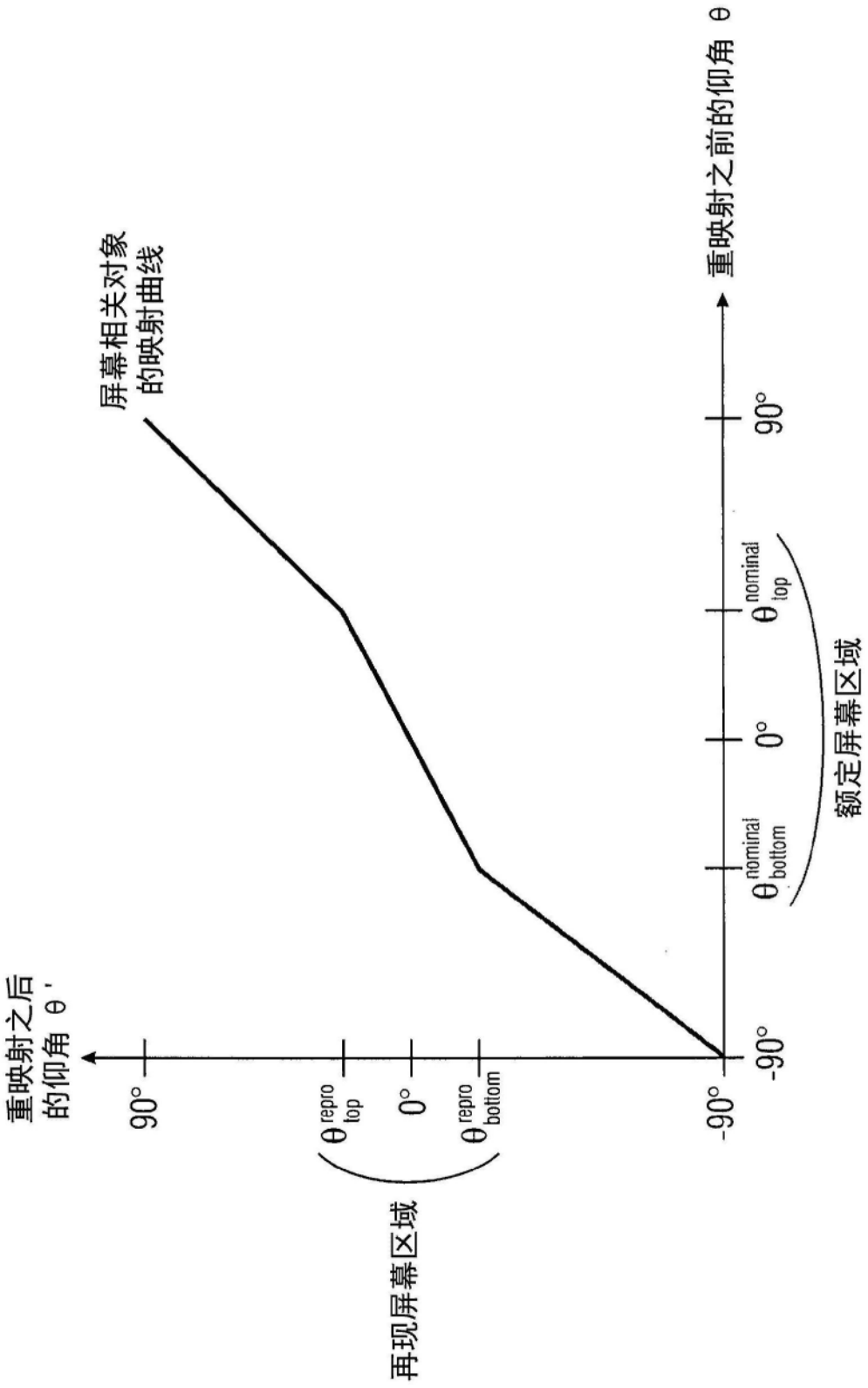

重映射之后
的仰角 θ'
90°
θ repro top
0°
θ repro bottom
-90°
-90°
θ nominal bottom
0°
θ nominal top
90°
重映射之前的仰角 θ
屏幕相关对象
的映射曲线
再现屏幕区域
额定屏幕区域

图14

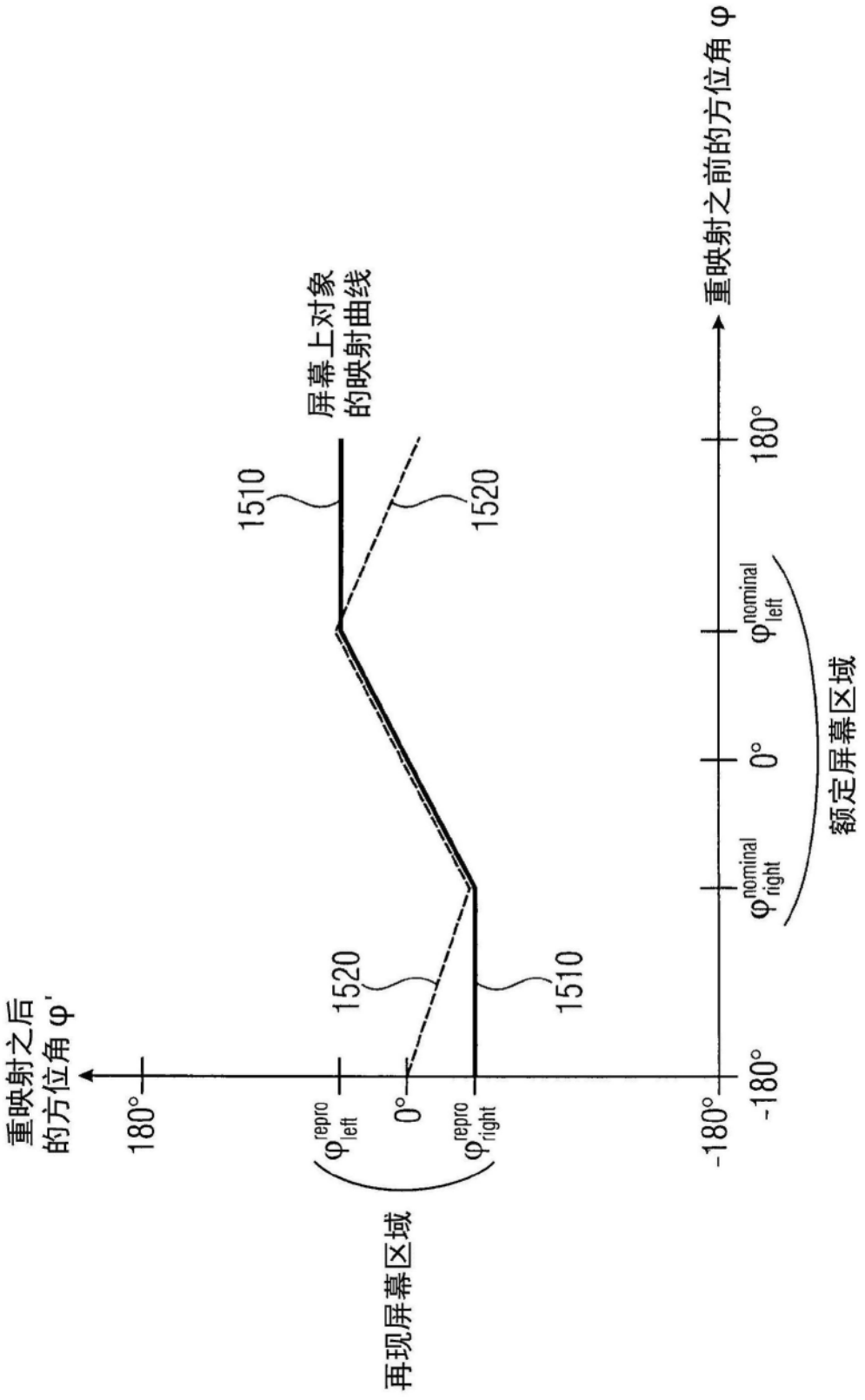

重映射之后
的方位角 φ'
180°
-180°
0°
φ_{left}^{repro}
φ_{right}^{repro}
再现屏幕区域
1520
1510
屏幕上对象
的映射曲线
1510
1520
重映射之前的方位角 φ
-180°
180°
0°
$\varphi_{right}^{nominal}$
$\varphi_{left}^{nominal}$
额定屏幕区域

图15

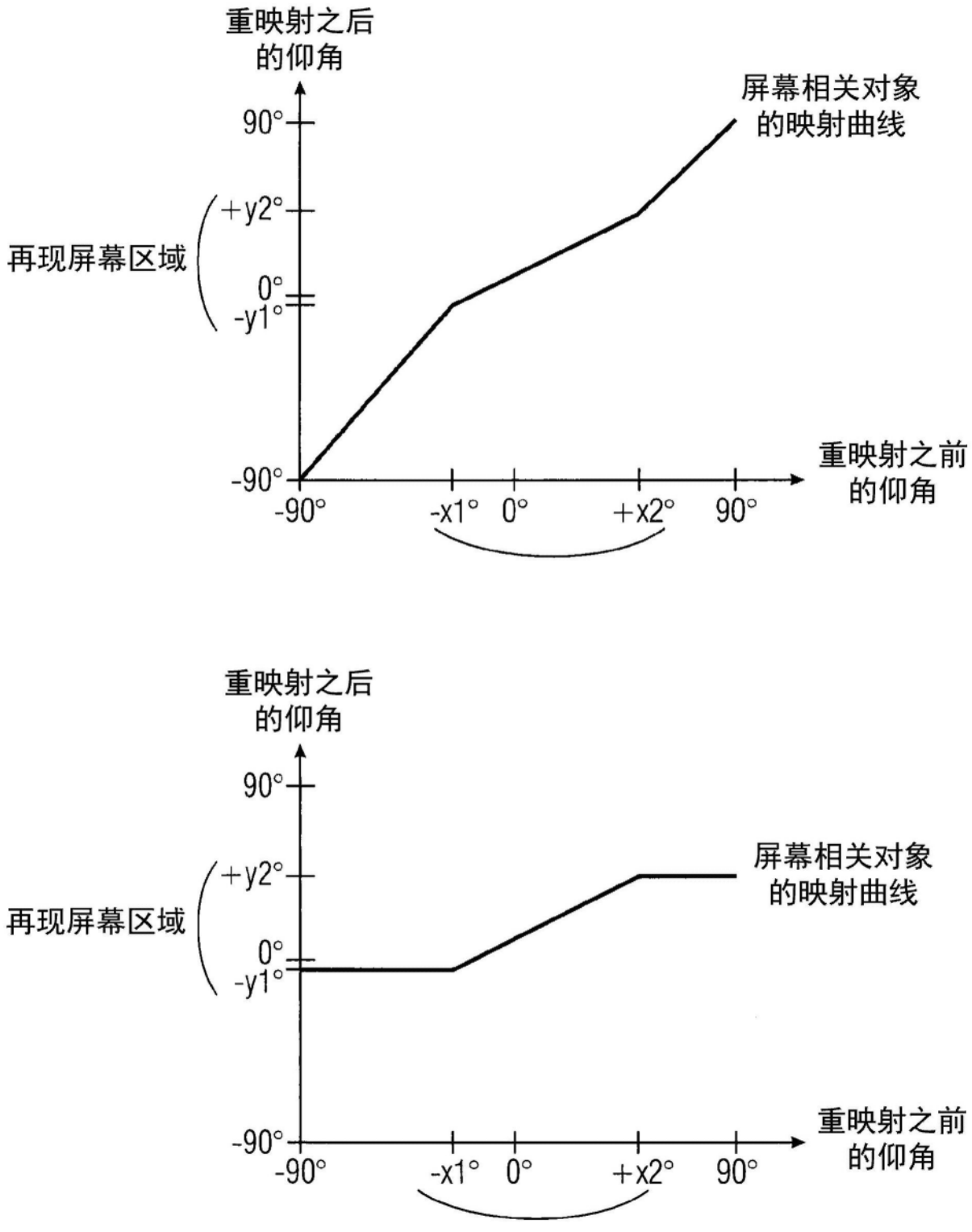

重映射之后
的仰角
90°
+y2°
0°
-y1°
-90°
再现屏幕区域
屏幕相关对象
的映射曲线
重映射之前
的仰角
-90°
-x1°
0°
+x2°
90°
重映射之后
的仰角
90°
+y2°
0°
-y1°
-90°
再现屏幕区域
屏幕相关对象
的映射曲线
重映射之前
的仰角
-90°
-x1°
0°
+x2°
90°

图16

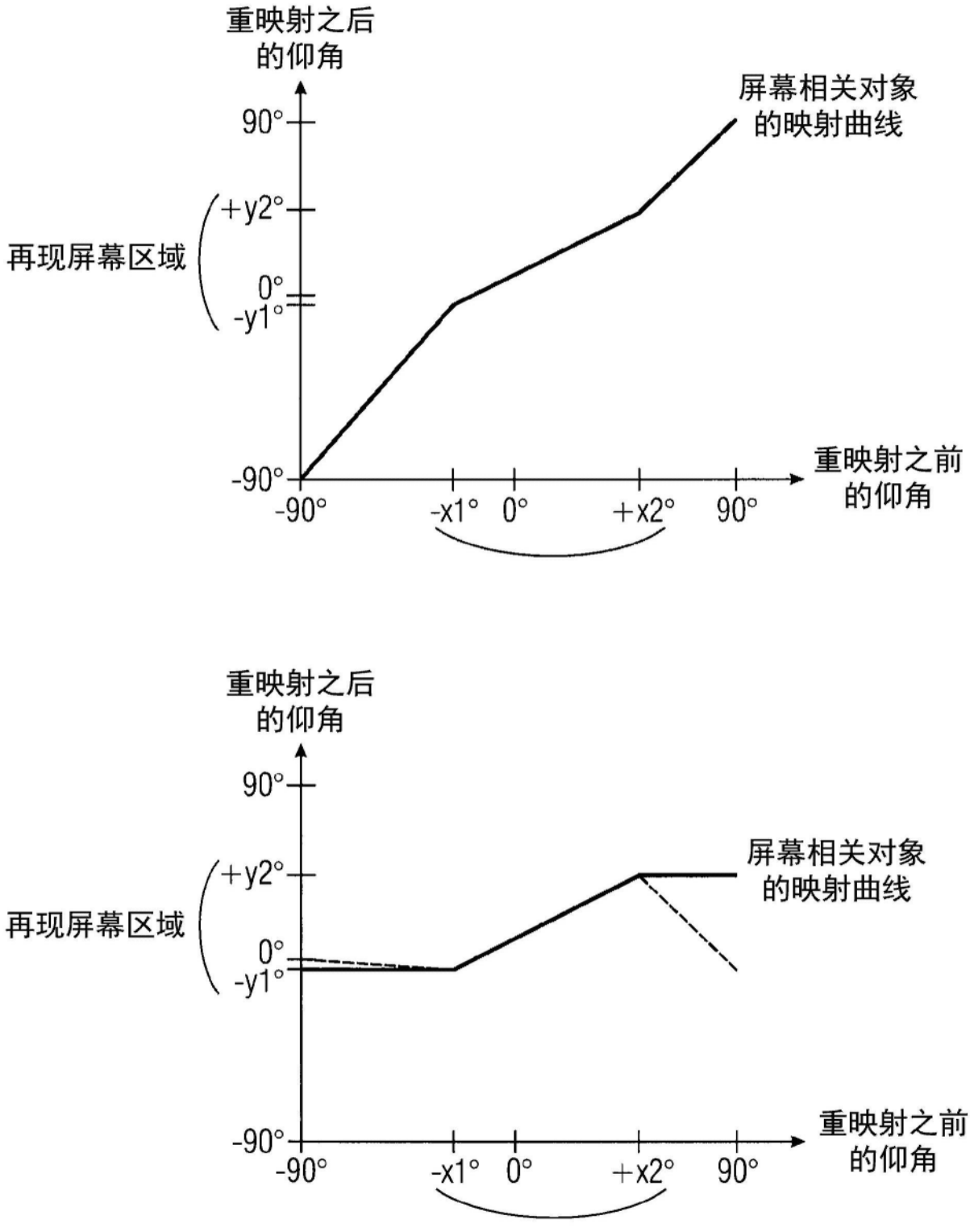

重映射之后
的仰角
屏幕相关对象
的映射曲线
90°
+y2°
再现屏幕区域
0°
-y1°
-90°
重映射之前
的仰角
-90°
-x1°
0°
+x2°
90°
重映射之后
的仰角
90°
屏幕相关对象
的映射曲线
+y2°
再现屏幕区域
0°
-y1°
-90°
重映射之前
的仰角
-90°
-x1°
0°
+x2°
90°

图17