(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 109644279 B
(45)授权公告日 2023.09.22

(21)申请号 201780053692.8
(22)申请日 2017.08.30
(65)同一申请的已公布的文献号
申请公布号 CN 109644279 A
(43)申请公布日 2019.04.16
(30)优先权数据
62/383,367 2016.09.02 US
62/407,337 2016.10.12 US
(85)PCT国际申请进入国家阶段日
2019.03.01
(86)PCT国际申请的申请数据
PCT/US2017/049495 2017.08.30
(87)PCT国际申请的公布数据
WO2018/045108 EN 2018.03.08
(73)专利权人 VID拓展公司
地址 美国特拉华州威明顿市
(72)发明人 菲利普·汉哈特 贺玉文 叶艳
(74)专利代理机构 北京润平知识产权代理有限公司 11283
专利代理师 肖冰滨 王晓晓
(51)Int.Cl.
H04N 19/597(2006.01)
H04N 19/46(2006.01)
H04N 19/167(2006.01)
G06T 15/04(2006.01)
H04N 19/117(2006.01)
(56)对比文件
WO 2016064862 A1,2016.04.28
US 2016012855 A1,2016.01.14
CN 105812759 A,2016.07.27
US 2016142697 A1,2016.05.19
CN 105681805 A,2016.06.15
JP 2006309802 A,2006.11.09
审查员 胡帆

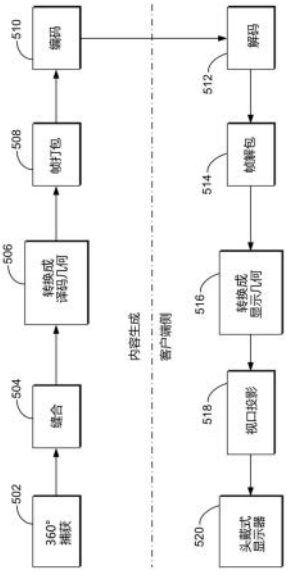
权利要求书3页 说明书25页 附图24页

(54)发明名称

用于用信号通知360度视频信息的方法和系统

(57)摘要

描述用于360度视频的译码技术。编码器选择投影格式，并使用所选择的投影格式将360度视频映射到2D平面视频。编码器对比特流中的2D平面视频进行编码，并且在比特流中进一步的用信号通知标识投影格式的参数。标识投影格式的参数可以在比特流的视频参数集、序列参数集和/或图片参数集中被用信号通知。可能用信号通知的不同投影格式包括使用如等距柱状、立方体贴图、等面积、八面体、二十面体、圆柱体和用户指定的多边形的几何的格式。可能用信号通知的其他参数包括用于几何面的不同布置或不同面的不同编码质量。还描述了相应的解码器。在一些实施方式中，投影参数还可以包括限定投影几何形状的方位的相对几何旋转参数。



CN 109644279 B

1.一种用于解码在比特流中编码的360度视频的方法,该方法包括:

接收编码了2D平面视频的比特流,所述2D平面视频表示360度视频通过针对所述2D平面视频中的至少两个图片使用不同投影几何格式而在2D平面上的投影,其中,所述比特流包括标识所述投影几何格式的参数,该参数包括标识所述2D平面视频中的当前图片的第一投影几何格式和所述当前图片的至少一个参考图片的第二投影几何格式的参数;

执行几何转换以将所述参考图片的投影几何与所述当前图片的所述投影几何对齐;以及

使用所对齐的参考图片对所述当前图片执行帧间预测。

2.根据权利要求1所述的方法,还包括使用所识别的投影几何格式将所述当前图片映射到360度视频。

3.根据权利要求1所述的方法,其中,所述当前图片的所述帧间预测是时间运动补偿预测。

4.根据权利要求3所述的方法,其中,多于一个的时间参考图片被用在运动补偿预测中,以及其中几何转换针对所述当前图片的所有参考图片被执行。

5.根据权利要求3所述的方法,其中,所述时间运动补偿预测使用仿射的运动模型。

6.根据权利要求1所述的方法,其中,所述当前图片的所述帧间预测是层间预测。

7.根据权利要求1所述的方法,其中,标识所述投影几何格式的所述参数标识从以下中的一者或多者中选择的几何类型:等距柱状、立方体贴图、八面体、二十面体、圆柱体和用户指定的多边形。

8.根据权利要求1所述的方法,其中,所标识的投影几何格式中的至少一者具有多个面,以及其中标识所述投影几何格式的所述参数包括关于面的数量的指示。

9.根据权利要求1所述的方法,其中,所标识的投影几何格式中的至少一者具有多个面,以及其中标识所述投影几何格式的所述参数包括标识所述2D平面视频中的所述面的布置的帧打包参数。

10.根据权利要求1所述的方法,其中,所标识的投影几何格式中的至少一者具有多个面,以及其中所述比特流还包括标识所述2D平面视频中的所述面的质量等级的参数。

11.根据权利要求1所述的方法,其中,所述投影几何格式中的至少一者包括投影几何的几何方位,以及其中标识所述投影几何格式的所述参数包括标识所述几何方位的参数。

12.根据权利要求11所述的方法,其中,标识所述几何方位的所述参数包括以下至少一者:偏航参数、俯仰参数和滚动参数。

13.一种包括一个或多个处理器的360度视频解码装置,所述一个或多个处理器被配置为至少执行:

接收编码了2D平面视频的比特流,所述2D平面视频表示360度视频通过针对所述2D平面视频中的至少两个图片使用不同的投影几何格式而在2D平面上的投影,其中,所述比特流包括标识所述投影几何格式的参数,该参数包括标识所述2D平面视频中的当前图片的第一投影几何格式和所述当前图片的至少一个参考图片的第二投影几何格式的参数;

执行几何转换以将所述参考图片的投影几何与所述当前图片的所述投影几何对齐;以及

使用所对齐的参考图片对所述当前图片执行帧间预测。

14.根据权利要求13所述的装置，其中所述处理器还被配置为使用所识别的投影几何格式将所述当前图片映射到360度视频。

15.根据权利要求13所述的装置，其中，所述当前图片的所述帧间预测是时间运动补偿预测。

16.一种用于在比特流中编码360度视频的方法，该方法包括：

编码2D平面视频，所述2D平面视频表示360度视频通过针对所述2D平面视频中的至少两个图片使用不同投影几何格式而在2D平面上的投影；

编码标识所述投影几何格式的参数，其中所述参数标识所述2D平面视频中的当前图片的第一投影几何格式和所述当前图片的至少一个参考图片的第二投影几何格式；

执行几何转换以将所述参考图片的投影几何与所述当前图片的所述投影几何对齐；以及

使用所对齐的参考图片对所述当前图片执行帧间预测。

17.根据权利要求16所述的方法，其中，所述当前图片的所述帧间预测是时间运动补偿预测。

18.根据权利要求17所述的方法，其中，多于一个的时间参考图片被用在运动补偿预测中，以及其中几何转换针对所述当前图片的所有参考图片被执行。

19.根据权利要求16的方法，其中，所述当前图片的所述帧间预测是层间预测。

20.根据权利要求16的方法，其中，标识所述投影几何格式的所述参数标识从以下中的一者或多者中选择的几何类型：等距柱状、立方体贴图、八面体、二十面体、圆柱体和用户指定的多边形。

21.根据权利要求16的方法，其中，所标识的投影几何格式中的至少一者具有多个面，以及其中标识所述投影几何格式的所述参数包括关于面的数量的指示。

22.根据权利要求16的方法，其中，所标识的投影几何格式中的至少一者具有多个面，以及其中标识所述投影几何格式的所述参数包括标识所述2D平面视频中的所述面的布置的帧打包参数。

23.根据权利要求16的方法，其中，所标识的投影几何格式中的至少一者具有多个面，以及其中所述比特流还包括标识所述2D平面视频中的所述面的质量等级的参数。

24.根据权利要求16的方法，其中，所述投影几何格式中的至少一者包括投影几何的几何方位，以及其中标识所述投影几何格式的所述参数包括标识所述几何方位的参数。

25.根据权利要求24的方法，其中，标识所述几何方位的所述参数包括以下至少一者：偏航参数、俯仰参数和滚动参数。

26.一种包括一个或多个处理器的360度视频编码装置，所述一个或多个处理器被配置为至少执行：

编码2D平面视频，所述2D平面视频表示360度视频通过针对所述2D平面视频中的至少两个图片使用不同投影几何格式而在2D平面上的投影；

编码标识所述投影几何格式的参数，其中所述参数标识所述2D平面视频中的当前图片的第一投影几何格式和所述当前图片的至少一个参考图片的第二投影几何格式；

执行几何转换以将所述参考图片的投影几何与所述当前图片的所述投影几何对齐；以及

使用所对齐的参考图片对所述当前图片执行帧间预测。

27.根据权利要求26所述的装置,其中,所述当前图片的所述帧间预测是层间预测。

28.根据权利要求26的装置,其中,所标识的投影几何格式中的至少一者具有多个面,以及其中标识所述投影几何格式的所述参数包括关于面的数量的指示。

29.根据权利要求26的装置,其中,所标识的投影几何格式中的至少一者具有多个面,以及其中标识所述投影几何格式的所述参数包括标识所述2D平面视频中的所述面的布置的帧打包参数。

用于用信号通知360度视频信息的方法和系统

[0001] 相关申请的交叉引用

[0002] 本申请是根据35U.S.C.§119(e)要求于2016年9月2日提交、名称为“用于用信号通知360度视频信息的方法和系统”的美国临时专利申请序列号62/383,367和于2016年10月12日提交、名称为“用于用信号通知360度视频信息的方法和系统”的美国临时专利申请序列号62/407,337的权益,其全部内容通过引用并入本文。

背景技术

[0003] 虚拟现实(VR)正在从研究实验室走出并进入我们的日常生活中。VR有很多应用领域:医疗保健、教育、社交网络、行业设计/培训、游戏、电影、购物、娱乐等。它受到行业和消费者的高度关注,因为VR能够带来身临其境的观看体验。它创建一个围绕着观看者的虚拟环境,并可能产生一种“正在存在”的真实感觉。如何在VR环境中提供完整的真实感觉对于用户的体验很重要。例如,VR系统应该通过姿态、手势、眼睛凝视、声音等来支持交互。为了允许用户以自然的方式与VR世界中的对象交互,系统还可以向用户提供触觉反馈。

[0004] 现今的VR系统使用360度视频来为用户提供从水平方向360度角度和垂直方向180度角度观看场景的能力。同时,VR和360度视频被视为超出超高清(UHD)服务的媒体消费的未来方向。为了提高VR中360度视频的质量,规范VR客户端互操作性的处理链,2016年初,属于MPEG-A(多媒体应用程序格式)部分-19的ad hoc组已在ISO/IEC/MPEG建立,以致力于全方位媒体应用格式的要求和潜在技术。另一个ad hoc组,免费观看电视(FTV),发布了360度3D视频应用的探索实验。FTV的一个主要目标是测试两种解决方案的性能:(1)基于360度视频(全向视频)的系统;(2)基于多视图的系统。来自MPEG和ITU-T的联合视频探索团队(JVET)正在探索下一代视频译码标准的新技术,发布了包括VR在内的测试序列的召集。在2016年6月的会议上,ad hoc组(AHG8)成立,AHG8组的任务是制定360视频译码的常用测试条件、测试序列格式和评估标准。AHG8还将研究应用不同投影方法时对压缩的影响,以及转换软件时对压缩的影响。

[0005] 行业正在努力提高VR处理链中各个方面的质量和用户体验,包括捕获、处理、显示和应用。在捕获方面,VR系统使用多个摄像机系统从不同的视图捕获场景(例如,在某些情况下,大约6到12个视图)。将这些视图缝合(stitch)在一起形成高分辨率(例如4K或8K)的360度视频。在客户端或用户侧,当前的虚拟现实系统通常包括计算平台、头戴式显示器(HMD)和头部跟踪传感器。计算平台负责接收和解码360度视频,并生成视口进行显示。两张照片(每个眼睛一个)针对视口而呈现。这两张照片显示在HMD中,用于立体观看。可以使用透镜来放大显示在HMD中的图像以便更好地观看。头跟踪传感器可以不断地跟踪观看者的头部方位(orientation),并且将方位信息馈送到系统以显示用于该方位的视口图片。

[0006] 一些VR系统可以为观看者提供专门的触摸设备来与虚拟世界中的对象进行交互。存在市场上可获得的现有的VR系统。一种是Oculus提供的Rift,以及来自三星和Oculus的Gear VR。Rift由具有良好的GPU支持的强大的工作站驱动。Gear VR是一种轻型VR系统,它使用作为计算平台的智能手机、HMD显示器和头部跟踪传感器。第二种VR系统是HTC Vive系

统。Rift和Vive具有相似的性能。空间HMD分辨率为2160×1200，刷新率为90Hz，视场(FOV)约为110度。头跟踪传感器的采样率为1000Hz，可以捕获非常快的运动。谷歌也有一个简单的VR系统叫做纸盒(cardboard)。Google纸盒有一个镜头和纸盒组件，类似于Gear VR，它是由智能手机驱动的。索尼也提供了用于游戏的PlayStation VR。在360度视频流服务方面，YouTube和Facebook跻身于早期的供应商之中。

[0007] 在这些当前的VR系统中，诸如互动和触觉反馈等体验的质量仍然需要进一步改进。例如，现今的HMD仍然太大，不方便穿戴。此外，由HMD提供的立体视图的当前分辨率2160×1200是不够的，并且可能导致一些用户头晕和不适。因此，增加分辨率将是有益的。此外，将VR环境中的视觉感觉与现实世界中的力量反馈相结合，是增强VR体验的一个选择。VR过山车是一示例应用。

[0008] 许多公司正在开展360度视频压缩和传送系统，他们有自己的解决方案。例如，Google YouTube为基于DASH的360度视频流提供了渠道。Facebook还拥有360度视频传送解决方案。

[0009] 本文的系统和方法旨在解决与360度视频数据编码和解码有关的问题。

发明内容

[0010] 在编码360度视频的示例性方法中，编码器选择投影格式，其中投影格式包括诸如几何类型和/或几何方位的信息。编码器使用所选投影格式将360度视频映射到2D平面视频。编码器对比特流中的2D平面视频进行编码，并且在比特流中进一步的用信号通知标识投影格式的参数。可以使用各种几何类型并且可以在比特流中用信号通知，包括等距柱状(equirectangular)、立方体贴图贴图(cubemap)、等面积、八面体、二十面体、圆柱体和用户指定的多边形。对于与多个面相关联的几何类型，可以用信号通知帧打包参数以标识2D平面视频中那些面的位置和/或方位。可以用不同的尺寸和/或不同的质量水平对不同的面进行编码。标识几何方位的参数可以包括偏航参数、俯仰参数和滚动参数中的至少一者。

[0011] 标识投影格式的参数可以在比特流的视频参数集、序列参数集和/或图片参数集中被用信号通知。可以基于速率失真优化来选择投影参数。视频中的不同图片或不同序列可以使用不同的投影格式进行编码(例如，当不同的图片或序列对于不同投影格式具有较高的速率失真性能时)，在适当的参数集处提供用信号通知投影格式参数。还描述了相应的解码技术。

[0012] 在本公开中描述的是用于360度视频译码的示例性语法。语法元素可以用于指定投影几何和/或使用网格系统来指定在帧打包图片中的面的布置。面可以具有不同的尺寸和/或方位。在一些实施方式中，2-D平面上的面布置可以具有各种特征，例如沿着每列/行的恒定的面宽度/高度。在一些实施方式中，使用任何基于多边形的表示来描述用户指定几何体的示例性语法。在一些实施方式中使用的附加特征可以包括：使用标志来跳过用于填充帧打包图片的样本、在逐个面的级别上用信号通知增量量化参数(QP)、使能/禁用特定面之间的环路滤波器的标志、和/或仅对360视频的特定区域进行译码语法。

[0013] 在一些实施方式中，投影参数还可以包括相对几何旋转参数。这样的参数可以限定投影几何的方位。投影几何可以被选择性地定向，使得感兴趣的对象基本上完全包括在投影几何的单个面内。在以不同质量水平(例如不同QP值)编码不同面的实施方式中，投影

几何可以被定向成使得感兴趣对象基本上完全包含在用相对高质量水平编码的面内。

附图说明

[0014] 从以下结合附图通过示例提供的描述可以具有更详细地理解,其中:

[0015] 图1A示出了球体几何上的使用在经度和纬度上的球体采样的等距柱状投影。

[0016] 图1B示出了针对图1A中的采样的2D平面等距柱状投影,其中在图1A中的球体上的点P被投影到2D平面中的点q。

[0017] 图1C是具有等距柱状投影的示例图片的示意图。

[0018] 图2A示出了在具有面PX(0)、NX(1)、PY(2)、NY(3)、PZ(4)、NZ(5)的3D几何结构上的立方体贴图投影。

[0019] 图2B示出了图2A中限定的六个面的2D平面。

[0020] 图2C示意性地示出了具有立方体贴图投影的示例图。

[0021] 图3A示出了用于等面积投影的以等面积方式进行的球体采样。

[0022] 图3B示出了图3A的等面积投影的2D平面,其中球体上的点p投影到2D平面中的点q,并且水平线(A0,A1,A2等)的纬度是不等间隔。

[0023] 图3C示意性地示出了具有等面积投影的示例图。

[0024] 图4A示出了具有3D几何结构的八面体投影。

[0025] 图4B示出了图4A的3D结构的2D平面打包。

[0026] 图4C示意性地示出了具有八面体投影的示例图。

[0027] 图5示出了360度视频处理工作流程的一个实施方式。

[0028] 图6示出了基于块的视频编码器的功能框图的一个实施方式。

[0029] 图7示出了视频解码器的功能框图的一个实施方式。

[0030] 图8A示出立方体贴图投影格式的物理布局的一个实施方式。

[0031] 图8B示出了八面体投影格式的物理布局的一个实施方式。

[0032] 图9A示出了以4×3格式表示的立方体贴图。

[0033] 图9B示出了以3×2格式表示的立方体贴图。

[0034] 图9C示出了以3×3格式表示的立方体贴图,前面是其他面的尺寸的两倍(面积的四倍)(在这种情况下,前面扩展了两行和两列)。

[0035] 图10A-10H示出了三角形面的面旋转的定义:图10A:0°旋转;图10B:90°旋转;图10C:180°旋转;图10D:270°旋转;图10E:0°旋转,然后垂直翻转;图10F:90°旋转,然后垂直翻转;图10G:180°旋转,然后垂直翻转;图10H:270°旋转,然后垂直翻转。

[0036] 图11A示出了用于八面体的非紧凑帧打包格式。

[0037] 图11B示出了用于八面体的紧凑帧打包格式。

[0038] 图12A示出了用于二十面体的非紧凑帧打包格式。

[0039] 图12B示出了用于二十面体的紧凑的帧打包格式。

[0040] 图13A示出了用于立方体贴图的有限球体范围译码,其中全图片表示帧打包图片,以及矩形限定译码的区域。

[0041] 图13B示出了等边的有限球体范围译码,其中全图片表示帧打包图片,以及矩形限定译码的区域。

[0042] 图14A和14B示出了帧打包图片中的面的示例性替代布置，每个示出了六个面的布置，诸如可以与立方体贴图投影结合使用。

[0043] 图15示出了在一些实施方式中可以用作编码器或解码器的示例性无线发射/接收单元(WTRU)。

[0044] 图16示出了在一些实施方式中可以用作编码器或解码器的示例性网络实体。

具体实施方式

[0045] 现在将参考各附图提供说明性实施方式的详细描述。尽管该描述提供了可能的实现的具体示例，但是应当注意，提供的细节旨在作为示例而不是限制应用的范围。

[0046] 360度视频编码和解码

[0047] 360度视频传送的一种技术是使用球体几何结构来表示360度信息。例如，由多个摄像机捕获的同步的多个视图被缝合在球体上以作为一个整体结构。然后，使用给定的几何转换过程将球体信息投影到2D平面表面，例如，等距柱状投影(ERP)方法。图1A示出了经度(φ)和纬度(θ)上的球体采样，图1B示出了使用等距柱状投影将球体投影到2D平面。在航空学中，在[-π,π]范围内的经度φ被称为偏航，在[-π/2,π/2]范围内的纬度θ称为俯仰，其中π是圆的周长与其直径的比率。为了便于说明，(x,y,z)用于表示3D空间中的点的坐标，以及(ue,ve)用于表示具有等距柱状投影的2D平面中的点的坐标。等距柱状投影可以在等式(1)和(2)中以数学方法表示：

[0048] $ue = (\phi/(2*\pi)+0.5)*W$ (1)

[0049] $ve = (0.5-\theta/\pi)*H$ (2)

[0050] 其中W和H是2D平面图片的宽度和高度。如图1A所示，使用等式(1)和(2)将球体上的经度L4和纬度A1之间的交叉点，点P，映射到2D平面中的唯一点q(图1B)。2D平面中的点q可以通过反向投影而投影回球体上的点P。图1B中的视场(FOV)示出了将球体中的FOV映射到2D平面的示例，其中沿X轴的视角为约110度。

[0051] 通过ERP，360度视频可以映射到常规的2D视频。它可以使用现有的视频编解码器(如H.264或HEVC)进行编码，然后传送到客户端。在客户端侧，通过在HMD内投影和显示属于等距柱状图片中的FOV的部分，基于用户的视口来解码和渲染等距柱状视频。虽然球体视频可以转换到2D平面图片以使用等距柱状投影进行编码，但是等距柱状2D图片的特征与传统2D图片(也称为直线视频)的特征不同。图1C是房间内部的示例等距柱状图片的示意图。由于在2D空间域中的等距柱状采样是不均匀的，与对应于赤道的图片的中间部分相比，对应于北极的图片的顶部和对应于南极的底部被拉伸。与常规2D视频中的运动相比，2D等距柱状图片中的运动场在时间方向中变得复杂。

[0052] 诸如MPEG-2、H.264和HEVC的视频编解码器使用平移模型来描述运动场，并且不能有效地表示等距柱状投影的2D平面图片中的形状变化运动。等距柱状投影的另一个缺点是，与更靠近赤道的区域相比，靠近极点的区域对于观看者和/或内容提供商来说可能不那么有趣。例如，观看者可能不会在任何相当长的时间段内专注于顶部和底部区域。然而，基于翘曲效应，这些区域在等距柱状投影之后被拉伸成2D平面的大部分，并且压缩这些区域可能需要大量的比特。

[0053] 基于这些观察，正在研究一些处理方法来改进等距柱状图片译码，例如通过对这

些极点区域应用诸如平滑的预处理来减少译码它们所需的带宽。此外，已经提出了用于表示360度视频的不同几何投影，诸如立方体贴图、等面积、圆柱体、金字塔、八面体等。在这些投影方法中，最容易压缩的几何可以是立方体贴图，其共有6个面，每个面都是一个平面方形。图2A显示立方体贴图几何的示例。立方体贴图由6个方形面组成。假设相切球体的半径为1，则立方体贴图的每个面(方形)的横向长度为2。图2B示出了将6个面放置成矩形的一种打包方法，其可以用于编码和传递。具有立方体贴图投影的示例图片的示意图在图2C示出。空白区域(20)是填充区域以填充矩形图片。对于每个面，图片看起来与常规2D图片相同。但是，每个面的边界不连续。穿过两个相邻的面的直线，例如表示墙壁和天花板之间的连接处的线22将在这两个面的边界处弯曲。这意味着在面边界处的运动也将是不连续的。

[0054] 图3A-3B示出了等面积投影的几何结构示例。与等距柱状投影不同，球体上的垂直采样不是基于间距(pitch)的偶数间隔。每个采样纬度的Y轴上的投影均匀分布，以便为球体上的每个样本实现相同的面积。对于靠近极点区域的那些区域，垂直方向的采样变得更加稀疏。这也意味着赤道附近有更多的样本。在实际情况下，这是较为理想的，因为与靠近极点的区域相比，用户通常更频繁地观看赤道附近的区域。图3C是具有等面积投影的示例图片的示意图。与图1C相比，在图3C中在赤道附近的区域被放大，而极点附近的区域被挤压。

[0055] 图4A示出了八面体投影的几何结构的示例。八面体由8个等边三角形面组成。如果相切球体的半径为1，则每个三角形的横向长度为√6。图4B示出了将八个三角形布置成一个矩形的一种打包方法。图4C示意性地示出了具有八面体投影的一个示例图。在两个相邻三角形的共享边界的角落处观察到翘曲失真，例如所看到的在门口402的失真。

[0056] 为了比较不同几何投影方法的译码效率，Yu等在M.Yu,H.Lakshman,B.Girod,“A Framework to Evaluate Omnidirectional Video Coding Scheme”,IEEE International Symposium on Mixed and Augmented Reality(IEEE混合与增强现实国际会议),2015,中提出了基于纬度的PSNR(L-PSNR)。它考虑了两个因素：(1)在球体上均匀采样；(2)观看者的观看行为。它定义了在球体上均匀分布的一些样本，并且还根据其纬度来定义样本的权重。通过考虑所有这些均匀分布的样本，通过加权均方误差(MSE)测量失真。通过在观看者观察这些训练序列时跟踪观看者的视角来导出权重。如果被更频繁地观察，则权重会更大。根据这些统计，赤道附近的权重大于极点附近的权重，因为最感兴趣的内容位于赤道附近。使用球体上这些均匀分布的样本提供了一种比较不同投影方法的性能的措施。然而，当应用不同的投影时，那些预定义的球体样本可能不会投影到整数采样位置。如果采用基于内插滤波器的重采样方法，则会引入额外的内插误差。如果应用最近邻采样，则不能再保证均匀采样。因此，客观和主观质量评估方法仍然是360度视频译码的开放主题。

[0057] 360度相机和缝合软件广泛支持等距柱状格式。为了在立方体贴图几何中编码360度视频，必须将等距柱状格式转换为立方体贴图格式。等距柱状和立方体贴图之间的关系如下。如图2A所示，每个面涉及从球体的中心到面中心的三个轴中的每一个。“P”表示正，“N”表示负，PX表示从球体中心沿正X轴的方向，NX为PX的反方向，PY、NY、PZ和NZ类似标记。然后，存在分别对应于正面、背面、顶部、底部、右侧和左侧面的6个面(PX,NX,PY,NY,PZ,NZ)，并且这些面从0到5进行索引。令Ps(X_s,Y_s,Z_s)是半径为1的球体上的点。它可以偏航φ和俯仰θ表示如下：

[0058] X_s＝cos(θ)cos(Φ) (3)

[0059] Y_s＝sin(θ) (4)

[0060] Z_s＝-cos(θ)sin(Φ) (5)

[0061] 设Pf为延伸从球从中心到Ps的线时的立方体贴图上的点。在不失一般性的情况下,设Pf在面NZ上。Pf(X_f,Y_f,Z_f)的坐标可以计算为:

[0062] X_f＝X_s/|Z_s| (6)

[0063] Y_f＝Y_s/|Z_s| (7)

[0064] Z_f＝-1 (8)

[0065] 其中|x|是变量x的绝对值。然后,面NZ的2D平面中的Pf(uc,vc)的坐标计算为:

[0066] uc＝W*(1-X_f)/2 (9)

[0067] vc＝H*(1-Y_f)/2 (10)

[0068] 从等式(3)到(10),可以建立立方体贴图中在特定面上的坐标(uc,vc)和球体上的坐标(Φ,θ)之间的关系。并且从等式(1)和(2)可知等距柱状点(ue,ve)与球体上的点(Φ,θ)之间的关系。因此,可以找到等距柱状几何和立方体贴图几何体之间的关系。从立方体贴图到等距柱状的几何映射可以总结如下。给定立方体贴图中一个面上的点(uc,vc),等距柱状平面上的输出(ue,ve)可以计算为:

[0069] 1)根据等式(9)和(10)中的关系,用(uc,vc)计算面上3D点P_f的坐标;

[0070] 2)根据等式(6)、(7)和(8)中的关系,用P_f计算球体上3D点P_s的坐标;

[0071] 3)根据等式(3)、(4)和(5)中的关系,用P_s计算球体上的(Φ,θ);

[0072] 4)根据等式(1)和(2)中的关系,从(Φ,θ)中计算等距柱状图片上点(ue,ve)的坐标。

[0073] 为了使用立方体贴图表示一个2D图片中的360度视频,立方体贴图的6个面可以被打包成一个矩形区域,这称为帧打包。然后将帧打包的图片作为一个常规的2D图片进行处理(例如,译码)。存在不同的帧打包配置,如3x2和4x3。在3x2配置中,6个面被打包成2行,一个行中有3个面。在4x3配置中,4个面PX,NZ,NX,PZ被打包在一行(例如,中心行)中,并且面PY和NY被分别打包在不同的两行(例如顶部和底部行)中。图2C利用对应于图1C中的等距柱状图片的4×3帧打包。

[0074] 在示例性场景中,采用等距柱状格式的360度视频作为输入,并且希望将输入转换成立方体贴图格式。应用以下步骤:

[0075] 1)对于立方体贴图格式的每个样本位置(uc,vc),通过上面介绍的方法计算等距柱状格式的相应的坐标(ue,ve)。

[0076] 2)如果如此计算的等距柱状的坐标(ue,ve)不在整数样本位置,则可以应用内插滤波器,以使用其相邻整数位置处的样本获得在其分数位置处的样本值。

[0077] 图5中描绘了用于360度视频系统的一个工作流程。它包括360度视频捕获502,例如使用多个摄像头捕获覆盖整个球体空间的视频。然后将这些视频缝合在一起(504),例如以等距柱状几何结构。等距柱状几何结构可以转换为另一个几何结构(506),例如立方体贴图,以用于编码,例如使用现有视频编解码器进行编码。帧打包508可以在编码510之前执行。经编码的视频通过例如动态流媒体或广播传递给客户端。在接收机处,视频被解码(512),解压缩的帧被解包(514)并转换(516)成显示几何(例如,等距柱状)。然后,它可以用

于根据用户的视角通过视口投影渲染518并显示在头戴式显示器520上。

[0078] 在专业和/或消费者视频应用中，色度分量通常被二次采样为比亮度分量更小的分辨率。色度二次采样减少了要编码的视频数据的量(并因此节省带宽和计算能力)，而不会显着影响视频质量。例如，广泛使用的色度格式之一称为4:2:0色度格式，其中两个色度分量被二次采样为亮度分辨率的1/4(水平1/2和垂直1/2)。在色度二次采样之后，色度采样网格可能已经变得不同于亮度采样网格。如图5所示，在整个处理流程中，在每个阶段处理的360度视频可以是色度分量已被二次采样的色度格式。

[0079] 图6是通用的基于块的混合视频编码系统的一个实施方式的框图。输入视频信号102被逐块处理。在HEVC中，使用扩展块大小(称为“译码单元”或CU)来有效地压缩高分辨率(例如，1080p及更高)视频信号。在HEVC中，CU可以达到64x64像素。CU可以进一步划分为预测单元或PU，针对其应用单独的预测方法。对于每个输入视频块(MB或CU)，可以执行空间预测(160)和/或时间预测(162)。空间预测(或“帧内预测”)使用来自相同视频图片/切片中的已经译码的相邻块的像素来预测当前视频块。空间预测减少视频信号固有的空间冗余。时间预测(也称为“帧间预测”或“运动补偿预测”)使用来自已经译码的视频图片的像素来预测当前视频块。时间预测减少视频信号固有的时间冗余。给定视频块的时间预测信号通常由指示当前块与其参考块之间的运动的量和方向的一个或多个运动向量进行信号通知。此外，如果支持多个参考图片(如H.264/AVC或HEVC等最近的视频编码标准的情况)，则对于每个视频块，也发送其参考图片索引；并且参考索引用于标识时间预测信号来自参考图片存储器(164)中的哪个参考图片。在空间和/或时间预测之后，编码器中的模式决定块(180)例如基于速率失真优化方法选择最佳预测模式。然后从当前视频块(116)中减去预测块；并使用变换(104)和量化(106)将预测残差去相关，以实现目标比特率。量化的残差系数被逆量化(110)和逆变换(112)以形成重构的残差，然后将其加回到预测块(126)以形成重构的视频块。在重构的视频块被放入参考图片存储器(164)中并被用于对未来的视频块进行译码之前，可以对重构的视频块应用(166)进一步的环路滤波，如解块滤波器和自适应环路滤波器。为了形成输出视频比特流120，编码模式(帧间或帧内)、预测模式信息、运动信息和量化的残差系数都被发送到熵译码单元(108)，以被进一步压缩和打包来形成比特流。

[0080] 图7是基于块的视频解码器的一般框图。视频比特流202首先在熵解码单元208处解包并进行熵解码。译码模式和预测信息被发送到空间预测单元260(如果帧内编码)或时间预测单元262(如果帧间编码)以形成预测块。残差变换系数被发送到逆量化单元210和逆变换单元212以重构残差块。然后，在226处将预测块和残差块相加在一起。重构的块在被存储在参考图片存储器264中之前可以进一步经过环路滤波。然后将参考图片存储器中的重构的视频发送出去以驱动显示设备，以及用于预测未来的视频块。

[0081] 示例性实施方式概述

[0082] 可以将360度视频数据投影到2D平面上，以使用传统2D平面视频译码对信息进行编码。由于可以使用许多种几何投影来表示360度数据，并且可以将投影数据打包成不同的配置，这会导致各种问题。

[0083] 一个问题是，为了能够从解码的2D平面视频中适当地重构360视频，几何和帧打包参数应该可用于解码器来解包数据并将其从2D空间投射回3D空间。例如，立方体贴图格式可以使用具有不同的面顺序、不同的面旋转或不同的面尺寸的不同的布置来表示，例如

3x2、4x3、1x6或6x1。另外，如果在接收机侧使用与译码格式不同的格式，则还需要将几何和帧打包参数从译码格式转换为所需格式。例如，如果译码格式是立方体贴图，但是显示格式是等距柱状，则必须进行转换。实际上，当文件格式多路复用器复用这些基本的流时，从视频本身提取这些帧打包布置信息更好，而不是依赖外部元数据。

[0084] 另一个问题是，对于一些帧打包配置，填充展开的面以使得所得到的帧打包图片形成矩形区域，对于存储或压缩目的可能是有益的。例如，在立方体贴图4x3格式中，必须在右上和右下边缘添加附加像素(见图2B和2C)。编码这些附加像素会消耗比特，但不会传达任何有意义的信息。因此，如果编码器跳过这些像素，则可以实现比特率节省。在这种情况下，应该向解码器用信号通知紧凑配置被用于360度视频的正确重构。此外，与传统的2D平面视频不同，在回放视频的任何时候，只有一部分360视频(例如，视口)被呈现并显示给用户(见图5)。统计显示，赤道周围的观察概率通常高于极点周围，靠近前视图的观察概率高于靠近后视图。因此，标识投影格式的信息将允许编码器在投影的2D视频中标识这些区域(即赤道vs极点和前面vs后面)，并应用不同的编码策略(例如，对与赤道和/或前面的区域相对应的区域花费更多的比特和/或应用更复杂的优化策略，并且对与极点和/或背面的区域相对应的区域花费较少的比特和/或应用更简单的优化策略)来利用用户的观看行为以更智能的方式分配比特和/或计算资源。

[0085] 另一个问题是现有的编解码器，例如MPEG-2，H.264和HEVC，是专为传统2D视频设计的，不考虑360度数据表示的任何属性。为了获得更好的压缩效率，先进的360视频译码工具可以利用完整的3D表示，但是由于在投影的2D平面视频上执行译码，这些工具可能会受益于关于几何和帧打包的信息。因此，关于几何和帧打包参数的信息可以对编码器和解码器都可用，以便能够适当且更有效地编码和解码360视频。例如，在立方体贴图格式中，展开的面只有2D平面视频上几个正确定位的相邻面，这限制了编解码器利用相邻面之间的冗余信息的能力。然而，如果编解码器具有关于3D表示的信息，其中立方体的每个面恰好具有4个相邻面，则可以利用更多的冗余信息来减少必须被编码的数据量。

[0086] 另外一个问题是几何和帧打包参数可以在360度视频的持续时间内变化。因此，如果几何和帧打包参数随着时间的推移而变化，则这些参数应该对用于360视频的每一帧的编码器和解码器都可用。例如，译码格式可以在特定时刻从立方体贴图改变到等距柱状，以实现更好的压缩性能，或者特定的立方体贴图面的集合的尺寸可以改变以适应特定视频段中较低或更高的带宽要求。

[0087] 本文公开的系统和方法解决了这些问题和其他问题。

[0088] 在一些实施方式中，通过借助于附加的高级语法元素来用信号通知比特流中的几何和帧打包参数，来解决针对360度视频编码的上述问题中的一个或多个问题。特别地，可以指定投影几何类型，包括用于几何面的不同参数，以将其定位在2D平面视频上。360视频参数可以在不同的级别被用信号通知。下面的一部分描述了如何在视频级别(例如，视频参数集或VPS级别)存储投影格式参数，以最小化当不同层和/或序列和/或图片使用相同的投影格式时必须传输的信息量。下面的另一部分介绍如何在序列级别(例如，序列参数集或SPS级别)发送投影格式，允许相同视频的不同序列使用不同的投影格式或改变与给定投影格式相关的参数。下面的另一部分介绍如何在图片级别(例如，图片参数集或PPS级别)用信号通知投影格式，允许相同序列的不同图片使用不同的投影格式或改变与给定投影格式相

关的参数。本文公开的系统和方法的另一方面是使得能够以不同的品质因子编码不同的几何面。例如，在立方体贴图格式中，可以以更高的质量对前、后、左和右面进行译码，而顶面和底面可以以较低质量进行译码。这是因为观看者更有可能观看地平线附近的区域，而不是靠近两极的区域。以这种方式，可以更有效地译码360视频。

[0089] 在一些实施方式中，提出了指定几何坐标系相对于绝对坐标系的旋转的系统和方法。可以使用这些系统和方法来旋转3D几何形状，使得感兴趣的对象或区域投影到可以以更高质量编码的面或面的集合中。类似地，如果感兴趣的物体或区域在几个面上分割，这可以减少每个面内的冗余，则几何旋转可以用于限定不同的方位，使得一个或多个重要物体可以放置在一个面内，使得可以实现更好的压缩效率。在某些情况下，如果这是不可能的，例如，如果一个对象是大的和/或足够近的，它在水平和垂直方向中的一个或两个中跨过90度，那么这些面可以被旋转，尽可能地将重要物体的部分放置在一个面内。由于3D几何的固有特性，当物体跨越多于一个面时，其几何结构在从一个面转换到另一个面时会“失真”，从而降低了相关性和译码效率。能够指定投影方向以使得一个面内的物体连续性最大化，可以提高译码效率。

[0090] 在视频级的360度视频属性信令

[0091] 不同的投影几何具有不同的特征。例如，只有一个面用于等距柱状投影和等面投影。没有面边界的问题，虽然图片被拉伸。立方体贴图具有六个面，并且在帧打包图片中具有许多面边界。每个图片可以以不同的投影几何进行译码或使用相同的几何进行译码，但是具有不同的面布局、尺寸或质量。为此，如表1所示，在一些实施方式中可以引入新的参数集，用于360视频。

[0092] 表1视频参数集RBSP

[0093]

video_parameter_set_rbsp(){	描述符
…	
vps_360_extension_flag	u(1)
if(vps_360_extension_flag)	
vps_360_extension()	
…	
}	

[0094] 在示例性实施方式中，标志vps_360_extension_flag可以具有以下语义。

[0095] vps_360_extension_flag:指定视频是否为360度视频，在这种情况下，可以使用用于360视频的有效表示和压缩的特定参数和工具。当不存在时，vps_360_extension_flag的值可以推断为等于0。

[0096] 在视频级别，根据表2，在一些实施方式中可以用信号通知不同序列和/或层中使用的投影格式的总数。

[0097] 表2.视频参数集360扩展语法

[0098]

vps_360_extension(){	描述符
vps_num_360_formats_minus1	ue(v)
for(i＝0;i<＝vps_num_360_formats_minus1;i++)	
360_format()	

if(vps_num_360_formats_minus1>0)	
360_format_idx_present_flag	u(1)
if(360_format_idx_present_flag)	
for(i＝0;i<＝MaxLayersMinus1;i++)	
vps_360_format_idx[i]	ue(v)
}	

[0099] 在示例性实施方式中,表2的参数可以具有以下语义。

[0100] vps_num_360_formats_minus1:指定在不同序列和/或层中使用的投影格式的数量(减1)。当不存在时,vps_num_360_formats_minus1的值可以推断为等于0,表示仅使用一种投影格式。

[0101] 360_format_idx_present_flag:指定语法元素vps_360_format_idx[i]是否存在。当不存在时,可以推断360_format_idx_present_flag的值等于0。

[0102] vps_360_format_idx[i]:指定VPS中360_format()语法结构列表中的适用于nuh_layer_id等于layer_id_in_nuh[i]的层的360_format()语法结构的索引。当不存在时,vps_rep_format_idx[i]的值可以推断为等于Min(i,vps_num_rep_formats_minus1)。

[0103] 对于这种提出的语法结构,在多层视频流中,每层的投影格式可能不同。例如,可以使用速率失真优化在编码器处确定每层的投影格式。编码器可以使用所有可用的投影格式对当前层进行编码,并随后测量速率失真成本。如果当前层是增强层,则可以在同一层中不仅使用帧内和帧间预测进行编码,而且可以使用与相同或不同的投影格式的另一层(例如,参考层)的层间预测。当来自参考层的投影格式与当前层的投影格式不同时,层间预测处理还可以包括投影格式转换。最后,可以选择导致最小速率失真成本的投影格式进行最终译码。

[0104] 在一些实施方式中,可以根据表3来用信号通知每个投影格式和相关参数的属性。

[0105] 表3. 360度表示格式语法

[0106]

360_format( ) {	描述符
projection_geometry	ue(v)
geometry_rotation_param_present_flag	u(1)
if(geometry_rotation_param_present_flag) {	
geometry_rotation_yaw	ue(v)
geometry_rotation_pitch	ue(v)
geometry_rotation_roll	ue(v)
}	
compact_representation_enabled_flag	u(1)
loop_filter_across_faces_enabled_flag	u(1)
num_face_rows	ue(v)
num_face_columns	ue(v)
equal_face_size_flag	u(1)
if(num_face_rows * num_face_columns > 1)	
face_qp_offset_enabled_flag	u(1)
for(i = 0; i < num_face_rows; i++) {	
for(j = 0; j < num_face_columns; j++) {	
face_idx[i][j]	ue(v)
if(!equal_face_size_flag) {	
face_width_in_luma_samples[i][j]	ue(v)
face_height_in_luma_samples[i][j]	ue(v)
}	
face_rotation_idc[i][j]	ue(v)
if(face_rotation_idc[i][j] == 4)	
face_rotation[i][j]	ue(v)
face_vertical_flip_flag[i][j]	u(1)
if(face_qp_offset_enabled_flag)	

[0107]

face_qp_offset[i][j]	se(v)
}	
}	
}	

[0108] 在示例性实施方式中,表3的参数可以具有以下语义。

[0109] projection_geometry:指定使用的投影几何的表4中的映射索引。

[0110] geometry_rotation_param_present_flag:指定语法元素geometry_rotation_yaw、geometry_rotation_pitch和geometry_rotation_roll是否存在。当不存在时,geometry_rotation_param_present_flag的值可以推断为等于0。

[0111] geometry_rotation_yaw:指定相对于绝对坐标系围绕几何坐标系的Y轴(参见图2A)的旋转。当不存在时,geometry_rotation_yaw的值可以推断为等于0。

[0112] geometry_rotation_pitch:指定相对于绝对坐标系围绕几何坐标系的Z轴(参见图2A)的旋转。当不存在时,geometry_rotation_pitch的值可以推断为等于0。

[0113] geometry_rotation_roll:指定相对于绝对坐标系围绕几何坐标系的X轴(参见图2A)旋转。当不存在时,geometry_rotation_roll的值可以推断为等于0。

[0114] compact_representation_enabled_flag:指定用于将帧打包图像填充到矩形图片中的样本或块是否被编码器跳过。当不存在时,compact_representation_enabled_flag的值可以推断为等于0。

[0115] loop_filter_across_faces_enabled_flag:指定是否可以跨面边界执行环路滤波操作。当不存在时,可以推断出loop_filter_across_faces_enabled_flag的值等于1。

[0116] num_face_rows:指定帧打包图片中的面行数。当不存在时,num_face_rows的值可以推断为等于1。

[0117] num_face_columns:指定帧打包图片中的面列数。当不存在时,num_face_columns的值可以推断为等于1。

[0118] 注意,可以用信号通知num_face_rows_minus1和num_face_columns_minus1而不是用信号通知num_face_rows和num_face_columns,来减少译码这些语法元素所需的位数。

[0119] equal_face_size_flag:指定是否所有面共享相同的大小(宽度和高度相同)。当不存在时,equal_face_size_flag的值可以推断为等于0。当equal_face_size_flag设置为1时,可以基于投影几何推断帧打包图片中所有面的宽度和高度。例如,对于立方体贴图投影,可以推断帧打包图片中所有面的亮度样本中的宽度等于pic_width_in_luma_samples/num_face_columns,而帧打包图片中所有面的亮度样本的高度可被推断为等于pic_height_in_luma_samples/num_face_rows。注意,帧打包图片中所有面的亮度样本的宽度和高度不应等于0,并且应为MinCbSizeY的整数倍。

[0120] face_qp_offset_enabled_flag指定不同的面是否使用不同的QP。当不存在时,可以推断face_qp_offset_enabled_flag的值等于0。

[0121] face_idx[i][j]:指定位于帧打包图片中的第i行和第j列的面的索引。对于只有一个单一的面的简单的几何,如等距柱状或等面积,唯一的面是面#0。对于其他几何形状,

可以使用面的默认编号和定位,如针对立方体贴图和八面体几何而在表5所示的。

[0122] face_width_in_luma_samples[i][j]:指定位于帧打包图片中第i行和第j列的面的亮度样本的宽度。可以采用技术来防止关于帧打包图片宽度的歧义。例如,可以强制地设定沿着每行的不同面宽的总和等于帧打包的图片宽度。face_width_in_luma_samples[i][j]不得等于0,并且应为MinCbSizeY的整数倍。

[0123] face_height_in_luma_samples[i][j]:指定位于帧打包图片中第i行和第j列的面的亮度样本的高度。可以采用技术来防止关于帧打包图片高度的歧义。例如,可以强制的设定沿着每列的不同面的高度的总和等于帧打包图片高度。face_height_in_luma_samples[i][j]不得等于0,并且应为MinCbSizeY的整数倍。

[0124] face_rotation_idc[i][j]:指定位于帧打包图片中第i行和第j列的面的图片坐标系和面坐标系之间的旋转的表6中的映射索引。当不存在时,face_rotation_idc[i][j]的值可以推断为等于0。

[0125] face_rotation[i][j]:指定位于帧打包图片中第i行和第j列的面的图片坐标系和面坐标系之间的旋转度。

[0126] face_vertical_flip_flag[i][j]:指定位于帧打包图片中位于第i行和第j列的面是否在旋转后被垂直翻转。当不存在时,face_vertical_flip_flag[i][j]的值可以推断为等于0。

[0127] face_qp_offset[i][j]:当确定位于帧打包图片中的第i行和第j列的面的QP值时,指定要添加到序列级QP的差异。

[0128] 表4.投影几何索引

[0129]

索引	投影几何
0	等距柱状
1	立方体贴图
2	等面积
3	八面体
4	二十面体
5	圆柱体
6	用户指定的多边形

[0130] 表5.默认面定义

[0131]

几何	**面定义**
立方体贴图	0: PX, 1: NX, 2: PY, 3: NY, 4: PZ, 5: NZ (见图 8A)
八面体	0: PXPY, 1: PXNY, 2: NXPY, 3: NXNY, 4: PYPZ, 5: NYPZ, 6: PYNZ, 7: NYNZ (见图 8B)

[0132] 表6.旋转索引

[0133]

索引	以逆时针方向旋转
0	0

1	90
2	180
3	270
4	用户指定

[0134] 将帧打包图片作考虑为面网格，这些参数可用于几何格式的非常灵活而强大的信令。对于导致单面(例如等距柱状、等面积或圆柱体)的投影几何，可以从几何和图片大小推断参数num_face_rows、num_face_columns、face_idx，face_width_in_luma_samples、face_height_in_luma_samples和face_rotation。然而，对于诸如立方体贴图，八面体或二十面体之类的其他几何，最好指定这些参数，因为面可以以不同的方式布置或具有不同的大小。例如，如图9A-9C所示，相同的立方体贴图投影可以以不同的方式进行打包，例如(a)3×4网格(图9A)或(b)2x3网格(图9B)。在3x4网格的情况下，可以将face_idx设置为高于实际面数的值，这可以从几何推断出来，以指示网格中不包含实际面的位置。例如，我们可以设置参数如下：

```
projection_geometry = 1     // 立方体贴图
face_idx[0][0] = 2          // 面 #2
face_idx[0][1] = 6          // 无效面
face_idx[0][2] = 7          // 无效面
face_idx[0][3] = 8          // 无效面
face_idx[1][0] = 1          // 面 #1
face_idx[1][1] = 4          // 面 #4
face_idx[1][2] = 0          // 面 #0
face_idx[1][3] = 5          // 面 #5
face_idx[2][0] = 3          // 面 #3
face_idx[2][1] = 9          //无效面
face_idx[2][2] = 10         //无效面
face_idx[2][3] = 11         //无效面
```

[0135]

[0136] 为了在某些方向提供更好的细节，某些面可以用更高的分辨率进行编码。这是因为与其它区域相比观看者更有可能观看某些区域，特别是靠近前面方向的区域。以这种方式，可以更有效地编码360度视频。为此，可以使用face_width_in_luma_samples和face_height_in_luma_samples参数为不同的面指定不同的大小。例如，在立方体贴图格式中，前

面可以以比其他面更高的分辨率进行译码，如图9C所示，我们可以设置参数如下：

[0137]

projection_geometry = 1

face_idx[0][0] = 0	face_width_in_luma_samples[0][0] =2W	face_height_in_luma_samples[0][0] =2H
face_idx[0][1] = 0	face_width_in_luma_samples[0][1] =2W	face_height_in_luma_samples[0][1] =2H
face_idx[0][2] = 4	face_width_in_luma_samples[0][2] =W	face_height_in_luma_samples[0][2] = H
face_idx[1][0] = 0	face_width_in_luma_samples[1][0] =2W	face_height_in_luma_samples[1][0] =2H
face_idx[1][1] = 0	face_width_in_luma_samples[1][1] =2W	face_height_in_luma_samples[1][1] =2H
face_idx[1][2] = 5	face_width_in_luma_samples[1][2] = W	face_height_in_luma_samples[1][2] = H
face_idx[2][0] = 2	face_width_in_luma_samples[2][0] = W	face_height_in_luma_samples[2][0] = H
face_idx[2][1] = 1	face_width_in_luma_samples[2][1] = W	face_height_in_luma_samples[2][1] = H
face_idx[2][2] = 3	face_width_in_luma_samples[2][2] = W	face_height_in_luma_samples[2][2] = H

[0138] 其中W是亮度样本中除面0(前面)之外的所有其他面的面宽，H是亮度样本中除面0(前面)之外的所有其他面的面高度。

[0139] 从这些参数可以推断出，前面跨越4个网格位置，因为它的大小是其他面的两倍，可以正确检索信息。

[0140] 面可以布置成具有不同的方位。例如，如立方体贴图投影所示，当与图9A的3x4网格相比时，面“2”，“1”和“3”在图9B的2x3网格中逆时针旋转90度。face_rottion_idc参数可用于指定面坐标系和帧打包图片坐标系之间的旋转。

[0141] 网格系统也可用于具有非正方形面(例如三角形面)的几何，如图11和12所示分别为八面体和二十面体。因为一些三角形面被分成两部分以用于紧凑表示(参见图11B和图12B)，可以使用两个直角三角形而不是一个等腰或等边三角形来定义一个三角形面。基本的直角三角形可以被定义为如图10A所示。由于旋转不足以使用两个直角三角形构造一个等腰三角形或等边三角形，所以旋转可以与垂直翻转(或在一些实施方式中，水平翻转)组合。通过这种表示，相同的语法可以用于具有很大灵活性的紧凑和非紧凑表示。例如，为了用信号通知图11B所示的紧凑八面体，可以如下设定参数：

```
projection_geometry = 3                 // 八面体
num_face_rows = 1                       // 1 行
num_face_columns = 16                   // 16 列
face_idx[0][0] = 15     //面#15    face_rotation_idc[0][0] = 2    face_vertical_flip_flag[0][0] = 0    // 180
face_idx[0][1] = 0      // 面#0    face_rotation_idc[0][1] = 0    face_vertical_flip_flag[0][1] = 0    // 0
face_idx[0][2] = 1      //面#1     face_rotation_idc[0][2] = 2    face_vertical_flip_flag[0][2] = 1    // 180 +垂直翻转
face_idx[0][3] = 8      //面#8     face_rotation_idc[0][3] = 0    face_vertical_flip_flag[0][3] = 1    // 0 +垂直翻转
face_idx[0][4] = 9      //面#9     face_rotation_idc[0][4] = 2    face_vertical_flip_flag[0][4] = 0    // 180
face_idx[0][5] = 2      //面#2     face_rotation_idc[0][5] = 0    face_vertical_flip_flag[0][5] = 0    // 0
face_idx[0][6] = 3      //面#3     face_rotation_idc[0][6] = 2    face_vertical_flip_flag[0][6] = 1    // 180 +垂直翻转
face_idx[0][7] = 10     //面#10    face_rotation_idc[0][7] = 0    face_vertical_flip_flag[0][7] = 1    // 0 +垂直翻转
face_idx[0][8] = 11     //面#11    face_rotation_idc[0][8] = 2    face_vertical_flip_flag[0][8] = 0    // 180
face_idx[0][9] = 4      //面#4     face_rotation_idc[0][9] = 0    face_vertical_flip_flag[0][9] = 0    // 0
face_idx[0][10] = 5     //面#5     face_rotation_idc[0][10] = 2   face_vertical_flip_flag[0][10] = 1   // 180 +垂直翻转
face_idx[0][11] = 12    //面#12    face_rotation_idc[0][11] = 0   face_vertical_flip_flag[0][11] = 1   // 0 +垂直翻转
face_idx[0][12] = 13    //面#13    face_rotation_idc[0][12] = 2   face_vertical_flip_flag[0][12] = 0   // 180
face_idx[0][13] = 6     //面#6     face_rotation_idc[0][13] = 0   face_vertical_flip_flag[0][13] = 0   // 0
face_idx[0][14] = 7     //面#7     face_rotation_idc[0][14] = 2   face_vertical_flip_flag[0][14] = 1   // 180 +垂直翻转
face_idx[0][15] =14     //面#14    face_rotation_idc[0][15] = 0   face_vertical_flip_flag[0][15] = 1   // 0 +垂直翻转
```

[0142]

[0143] face_qp_delta参数可用于指定特定面是否以较高或较低质量进行编码。例如,通过在切片或译码单元级别调整质量可以获得类似的结果。然而,切片可以覆盖几个面,并且面很可能包含多个译码单元,因此直接为每个面用信号通知质量差异可能是更有效地。

[0144] 对于由沿着每列具有相同宽度(但不同列不同宽度)且沿着每行的相同高度(但不同行不同高度)的面组成的常规帧打包网格,也可以使用较少的参数来用信号通知面属性,如表7所示。

[0145] 表7．360度表示格式替换语法

[0146]

360_format( ) {	描述符
projection_geometry	ue(v)
geometry_rotation_param_present_flag	u(1)
if(geometry_rotation_param_present_flag) {	
geometry_rotation_yaw	ue(v)
geometry_rotation_pitch	ue(v)
geometry_rotation_roll	ue(v)
}	
compact_representation_enabled_flag	u(1)
loop_filter_across_faces_enabled_flag	u(1)
num_face_rows_minus1	ue(v)
num_face_columns_minus1	ue(v)
face_qp_offset_enabled_flag	u(1)
if(num_face_rows_minus1 > 0)	
for(i = 0; i < num_face_rows_minus1; i++) {	
row_height_in_luma_samples[i]	ue(v)
if(num_face_columns_minus1 > 0)	
for(j = 0; j < num_face_columns_minus1; j++) {	
column_width_in_luma_samples[j]	ue(v)
for(i = 0; i <= num_face_rows_minus1; i++) {	
for(j = 0; j <= num_face_columns_minus1; j++) {	
face_idx[i][j]	ue(v)
face_rotation_idc[i][j]	ue(v)
if (face_rotation_idc[i][j] == 4)	
face_rotation[i][j]	ue(v)
face_vertical_flip_flag[i][j]	u(1)
if(face_qp_offset_enabled_flag)	

[0147]

face_qp_offset[i][j]	se(v)
}	
}	
}	

[0148] 在示例性实施方式中,表7的参数可以具有以下语义。

[0149] num_face_rows_minus1:指定帧打包图片中的面行数(减1)。当不存在时,可以推断num_face_rows_minus1的值等于0。

[0150] num_face_columns_minus1:指定帧打包图片中的面列数(减1)。当不存在时,可以推断num_face_columns_minus1的值等于0。

[0151] row_height_in_luma_samples[i]:指定位于帧打包图片中第i行的面的亮度样本中的高度。对于最后一行,高度可以推断为等于pic_height_in_luma_samples-$\Sigma_i$row_height_in_luma_samples[i]。row_height_in_luma_samples[i]不得等于0,并且应为MinCbSizeY的整数倍。

[0152] column_width_in_luma_samples[j]:指定位于帧打包图片中第j列的面的亮度样本中的宽度。对于最后一列,宽度可以推断为等于pic_width_in_luma_samples-$\Sigma_j$column_width_in_luma_samples[j]。column_width_in_luma_samples[j]不得等于0,并且应为MinCbSizeY的整数倍。

[0153] 面属性也可以在用于不规则面形状的面索引顺序中用信号通知。表8显示了一个例子。

[0154] 表8. 360度表示格式替换语法

[0155]

360_format( ) {	描述符
projection_geometry	ue(v)
geometry_rotation_param_present_flag	u(1)
if(geometry_rotation_param_present_flag) {	
geometry_rotation_yaw	ue(v)
geometry_rotation_pitch	ue(v)
geometry_rotation_roll	ue(v)
}	
compact_representation_enabled_flag	u(1)
face_boundary_filtering_idc	ue(v)
face_qp_delta_enabled_flag	u(1)
num_faces	ue(v)
for(i = 0; i < num_faces; i++) {	
face_idx[i]	ue(v)
num_face_vertices[i]	ue(v)
for(j=0; j < num_face_vertices[i]; j++) {	
vertex_2D_pos_x[i][j]	ue(v)
vertex_2D_pos_y[i][j]	ue(v)
if(projection_geometry == 6) {	
vertex_3D_pos_x[i][j]	se(v)
vertex_3D_pos_y[i][j]	se(v)
vertex_3D_pos_z[i][j]	se(v)
}	
}	
face_rotation_idc[i][j]	ue(v)
if(face_rotation_idc[i][j] == 4)	
face_rotation[i][j]	ue(v)

[0156]

face_vertical_flip_flag[i][j]	u(1)
if(face_qp_offset_enabled_flag)	
face_qp_delta[i][j]	se(v)
}	
}	

[0157] 在示例性实施方式中,表8的参数可以具有以下语义。

[0158] num_faces:指定帧打包图片中的面数。当不存在时,可以推断num_faces的值等于1。

[0159] 注意,代替信号num_faces,可以用信号通知num_faces_minus1以减少编码该语法元素所需的比特数。

[0160] num_face_vertices[i]:指定第i个面的顶点数。当不存在时,num_face_vertices[i]的值可以被推断为等于4,因为四边形是最常见的面多边形类型。

[0161] vertex_2D_pos_x[i][j]:指定第i个面的第j个顶点的帧打包图片中的x坐标。

[0162] vertex_2D_pos_y[i][j]:指定第i个面的第j个顶点的帧打包图片中的y坐标。

[0163] vertex_3D_pos_x[i][j]:指定第i个面的第j个顶点的3D坐标系中的x坐标。

[0164] vertex_3D_pos_y[i][j]:指定第i个面的第j个顶点的3D坐标系中的y坐标。

[0165] vertex_3D_pos_z[i][j]:指定第i个面的第j个顶点的3D坐标系中的z坐标。

[0166] 可以使用vertex_3D_pos_x[i][j],vertex_3D_pos_y[i][j]和vertex_3D_pos_z[i][j]参数来定义3D空间中用户指定的基于多边形的几何。这些参数可用于将样本从帧打包图片中的位置映射到3D几何中的相应位置。该信息可能被先进的360视频编码利用,以达到更好的压缩效率。例如,编解码器可以利用在3D表示中的未被并置在帧打包图片中的相邻面之间的冗余信息。

[0167] 序列级别360度视频属性信令

[0168] 在序列级别,可以用信号通知使用的投影格式。为此,如表9所示,可以为360视频引入新的参数集。

[0169] 表9.一般序列参数集RBSP语法

[0170]

seq_parameter_set_rbsp(){	描述符
…	
sps_360_extension_flag	u(1)
if(sps_360_extension_flag)	
sps_360_extension()	
…	
}	

[0171] 在示例性实施方式中,表9的参数可以具有以下语义。

[0172] sps_360_extension_flag:指定序列是否为360-视频,在这种情况下,可以使用用于360视频的有效压缩的特定参数和工具。

[0173] 所使用的投影格式可以根据表10用信号通知。

[0174] 表10.序列参数集360扩展语法

[0175]

sps_360_extension(){	描述符
sps_num_360_formats_minus1	ue(v)
for(i＝0;i<＝sps_num_360_formats_minus1;i++)	
sps_360_format_idx[i]	ue(v)
}	

[0176] 在示例性实施方式中,表10的参数可以具有以下语义。

[0177] sps_num_360_formats_minus1:指定序列中使用的投影格式的数量(减1)。当不存在时,sps_num_360_formats_minus1的值可以推断为等于0,表示仅使用一种投影格式。

[0178] sps_360_format_idx[i]:指定序列中使用的360_format()语法结构的VPS中360_format()语法结构列表中的索引列表。sps_360_format_idx[i]的值的范围可以是0到vps_num_360_formats_minus1,包括端值。

[0179] 注意,在VPS级别定义的所有360视频相关参数可以在SPS级别改变。尽管未在表10中示出,替代使用sps_360_format_idx索引在VPS中发送的360视频格式的集合,与表3中定义的语法元素类似的语法元素(例如,projection_geometry,face dimension parameters,face QP offset等)可以作为SPS扩展的一部分直接用信号通知,以指示参考该SPS的视频序列的360视频参数。

[0180] 图片级别360度视频属性信令

[0181] 在一些实施方式中,为了提供更大的译码优化,可以使用针对不同帧的不同投影格式对序列进行编码。在这种情况下,可以通过已经在VPS或SPS级别用信号通知的投影格式的集合中的索引来在图片级别用信号通知投影格式。为此,在一些实施方式中,如表11所示,可以为360视频引入新参数集。

[0182] 表11.一般图片参数集RBSP语法

[0183]

pic_parameter_set_rbsp(){	描述符
…	
pps_360_extension_flag	u(1)
if(pps_360_extension_flag)	
pps_360_extension()	
…	
}	

[0184] 在示例性实施方式中,表11的参数可以具有以下语义。

[0185] pps_360_extension_flag:指定引用该PPS的图片是否包含与360度视频译码相关的特定参数。当不存在时,可以推断pps_360_extension_flag的值等于sps_360_extension_flag。

[0186] 在表12和13中提供了360视频的PPS扩展的示例。

[0187] 表12.图片参数集360扩展语法

[0188]

pps_360_extension(){	描述符
pps_360_format_idx	ue(v)

coding_region_table()	
}	

[0189] 表13.译码区域表语法

[0190]

coding_region_table(){	描述符
full_sphere_range_coding_flag	u(1)
if(!full_sphere_range_coding_flag){	
pos_x_in_360_packed_frame	ue(v)
pos_y_in_360_packed_frame	ue(v)
}	
}	

[0191] 在示例性实施方式中,表12和13的参数可以具有以下语义。

[0192] pps_360_format_idx:指定由该PPS引用的在SPS定义的投影几何的集合中的索引。pps_360_format_idx的值应在0到sps_num_360_formats_minus1的范围内,包括端点值。当不存在时,pps_360_format_idx的值可以推断为等于0。

[0193] pps_360_format_idx参数用于指定在序列级别列出的可用投影格式中当前图片的投影格式。例如,如果在SPS的sps_360_format_idx列表中只有等距柱状和等面积可用,我们使用索引“0”表示等面积,“1”表示等距柱状,则该参数可以设置如下:

[0194] pps_360_format_idx=0//涉及此PPS的所有图片将以等面积格式译码

[0195] pps_360_format_idx=1//涉及此PPS的所有图片都以等距柱状格式译码。

[0196] 在相同的视频序列中,如果允许不同的图片具有不同的投影几何格式,则使用具有2个运动参数(分别为水平和垂直位移参数)的平移运动模型或使用具有4个或6个运动参数的基于仿射的运动模型进行时间运动补偿预测,可能无法再非常有效地工作。相反,如果当前图片的投影几何与其时间参考图片的投影几何不同,则可以在施加现有时间运动补偿预测之前,执行几何转换以对齐当前图片与其时间参考之间的投影几何。这可以增加时间预测效率,尽管以更高的计算复杂度为代价。当在运动补偿预测(例如,双预测)中使用多于一个的时间参考图片时,则可以在执行运动补偿预测之前,可以在当前图片与其所有参考图片之间对齐投影几何。

[0197] 在示例性实施方式中,coding_region_table()语法结构的语义可以如下:

[0198] full_sphere_range_coding_flag:指定整个球体范围是否被译码,或者是否仅其中一部分被译码。当不存在时,可以推断full_sphere_range_coding_flag的值等于1。

[0199] pos_x_in_360_packed_frame:指定帧打包图片中译码图片左上角的x坐标。

[0200] pos_y_in_360_packed_frame:指定帧打包图片中译码图片左上角的y坐标。

[0201] 由于不同的限制,例如带宽或存储器限制或解码能力,整个球体只有一部分可以被译码。可以使用full_sphere_range_coding_flag和关联的pos_x_in_360_packed_frame和pos_y_in_360_packed_frame参数来用信号通知该信息。当full_sphere_range_coding_flag设置为0时,仅对整个帧打包图片的矩形部分进行译码。然后,使用关联的pos_x_in_360_packed_frame和pos_y_in_360_packed_frame参数来用信号通知帧打包图片内的译码图片的左上角。

[0202] 图13A-B示出了用于立方体贴图(图13A)和等距柱状(图13B)投影的有限球体范围

译码的使用。在这些示例中，仅前面区域被译码。请注意，当使用有限的球体范围译码时，应禁用连接面宽/高度和译码图片宽度/高度的约束。如图3A所示，整个图片表示帧打包图片，并且矩形1305界定译码区域。如图13B所示，整个图片表示帧打包图片，矩形1310界定译码区域。

[0203] 也可以针对每种投影格式在VPS和/或PPS级别对signal_region_table()进行信号发送。

[0204] 注意，在SPS和/或VPS级别定义的一些参数可以替代地或附加地在PPS级别用信号通知。例如，特别有利的是，在PPS级别而不是在VPS或SPS级别处将面QP偏移参数用信号通知，因为它允许更多的灵活性来调整图片级别上每个单独的面的面质量。例如，它允许根据分层B预测结构中的当前帧打包图片的时间级别来灵活地调整每个单独面的面质量。例如，在较高的时间级别，对于非前面，可以将面QP偏移设置为较大值，而对于前面，可以将面QP偏移设置为较小的值(例如，0)。这可以确保不管当前图片的时间级别如何，前面总是以相对高的恒定质量进行译码，而较高时间级别图片的其他面可以被更多量化以节省比特。

[0205] 类似地，几何旋转参数(例如geometry_rotation_yaw，geometry_rotation_pitch和geometry_rotation_roll)可以在PPS级别而不是在VPS或SPS级别被定义和用信号通知，因为它允许更多的灵活性来调整在图片级别的几何旋转。在一些实施方式中，对于正被编码的内容(例如由视频内容的引导选择)，推荐的观看方向被选择，其中推荐的观看方向可以在视频进程中改变。在这样的实施方式中，几何旋转参数可以根据推荐的观察方向设置并与面QP偏移参数耦合，使得感兴趣的对象或区域投影到以最高质量译码的面。

[0206] 图14A和14B示出了帧打包图片中面的示例性替代布置。图14A和14B各自示出六个面的布置，诸如可以与立方体贴图投影结合使用。图14A和14B中的面的布置可以被用作使用本文公开的实施方式的用户指定的几何。

[0207] 使用一个或多个有线和/或无线网络节点(诸如无线发射/接收单元(WTRU)或其他网络实体)来实现本文公开的示例性实施方式。

[0208] 图15为示例性WTRU 1502的系统图，所述WTRU 1502可以被用作此处描述的实施方式中的编码器或解码器。如图15中所示，WTRU 1502可以包括处理器1518、包括收发信机1520的通信接口1519、发射/接收元件1522、扬声器/麦克风1524、键盘1526、显示屏/触摸板1528、不可移除存储器1530、可移除存储器1532、电源1534、全球定位系统(GPS)芯片组1536和传感器1538。需要理解的是，在保持与以上实施方式一致的同时，WTRU 1502可以包括上述元件的任何子集。

[0209] 处理器1518可以是通用目的处理器、专用目的处理器、常规处理器、数字信号处理器(DSP)、多个微处理器、与DSP核相关联的一个或多个微处理器、控制器、微控制器、专用集成电路(ASIC)、现场可编程门阵列(FPGA)电路、其他任何类型的集成电路(IC)、状态机等。处理器1518可以执行信号译码、数据处理、功率控制、输入/输出处理和/或使得WTRU 1502能够操作在无线环境中的其他任何功能。处理器1518可以耦合到收发信机1520，该收发信机1520可以耦合到发射/接收元件1522。尽管图15中将处理器1518和收发信机1520描述为独立的组件，但是可以理解的是处理器1518和收发信机1520可以被一起集成到电子封装或者芯片中。

[0210] 发射/接收元件1522可以被配置成通过空中接口1516将信号发送到基站，或者从

基站接收信号。例如，在一种实施方式中，发射/接收元件1522可以是被配置成发送和/或接收RF信号的天线。在另一实施方式中，发射/接收元件1522可以是被配置成发送和/或接收例如IR、UV或者可见光信号的发射器/检测器。在又一实施方式中，发射/接收元件1522可以被配置成发送和接收RF信号和光信号两者。需要理解的是发射/接收元件1522可以被配置成发送和/或接收无线信号的任意组合。

[0211] 尽管发射/接收元件1522在图15中被描述为单个元件，但是WTRU 1502可以包括任何数量的发射/接收元件1522。更特别地，客户端设备1502可以使用MIMO技术。由此，在一种实施方式中，WTRU 1502可以包括两个或更多个发射/接收元件1522(例如多个天线)以用于通过空中接口1516发射和接收无线信号。

[0212] 收发信机1520可以被配置成对将由发射/接收元件1522发送的信号进行调制，并且被配置成对由发射/接收元件1522接收的信号进行解调。如上所述，WTRU 1502可以具有多模式能力。由此，收发信机1520可以包括多个收发信机以用于使得WTRU 1502能够经由多种RAT进行通信，例如UTRA和IEEE 802.11。

[0213] WTRU 1502的处理器1518可以被耦合到扬声器/麦克风1524、键盘1526和/或显示屏/触摸板1528(例如，液晶显示(LCD)单元或者有机发光二极管(OLED)显示单元)，并且可以从上述装置接收用户输入数据。处理器1518还可以向扬声器/麦克风1524、键盘1526和/或显示屏/触摸板1528输出用户数据。此外，处理器1518可以访问来自任何类型的合适的存储器中的信息，以及向任何类型的合适的存储器中存储数据，所述存储器例如可以是不可移除存储器1530和/或可移除存储器1532。不可移除存储器1530可以包括随机接入存储器(RAM)、可读存储器(ROM)、硬盘或者任何其他类型的存储器存储装置。可移除存储器1532可以包括用户标识模块(SIM)卡、记忆棒、安全数字(SD)存储卡等类似装置。在其他实施方式中，处理器1518可以访问来自物理上未位于WTRU 1502上而诸如位于服务器或者家用计算机(未示出)上的存储器的数据，以及向上述存储器中存储数据。

[0214] 处理器1518可以从电源1534接收电能，并且可以被配置成将电能分配给WTRU 1502中的其他组件和/或对至WTRU 1502中的其他组件的电能进行控制。电源1534可以是任何适用于给WTRU 1502供电的装置。例如，电源1534可以包括一个或多个干电池(镍镉(NiCd)、镍锌(NiZn)、镍氢(NiMH)、锂离子(Li-ion)等)、太阳能电池、燃料电池等。

[0215] 处理器1518还可以耦合到GPS芯片组1536，该GPS芯片组1536可以被配置成提供关于WTRU 1502的当前位置的位置信息(例如经度和纬度)。作为来自GPS芯片组1536的信息的补充或者替代，WTRU 1502可以通过空中接口1516从基站接收位置信息，和/或基于从两个或更多个相邻基站接收到的信号的定时来确定其位置。需要理解的是，在与实施方式一致的同时，WTRU 1502可以通过任何合适的位置确定方法来获取位置信息。

[0216] 处理器1518还可以耦合到其他外围设备1538，该外围设备1538可以包括提供附加特征、功能性和/或无线或有线连接的一个或多个软件和/或硬件模块。例如，外围设备1538可以包括诸如加速度计的传感器、电子指南针(e-compass)、卫星收发信机、数码相机(用于照片和/或视频)、通用串行总线(USB)端口、震动装置、电视收发信机、免持耳机、蓝牙®模块、射频(FM)无线电单元、数字音乐播放器、媒体播放器、视频游戏播放器模块、因特网浏览器等等。

[0217] 图16描绘了可以在本公开的实施方式中使用的示例性网络实体1590，例如作为编

码器或解码器。如图16所示，网络实体1590包括通信接口1592，处理器1594和非暂时数据存储器1596，所有这些都通过总线、网络或其它通信路径1598通信地链接。

[0218] 通信接口1592可以包括一个或多个有线通信接口和/或一个或多个无线通信接口。关于有线通信，作为示例，通信接口1592可以包括一个或多个接口，例如以太网接口。关于无线通信，通信接口1592可以包括组件，诸如一个或多个天线、为一种或多种类型的无线(例如LTE)通信设计和配置的一个或多个收发信机/芯片组、和/或由相关领域的技术人员视为合适的任何其他组件。此外，关于无线通信，通信接口1592可以以适合于在无线通信(例如，LTE通信、Wi Fi通信等)的网络侧(与客户端侧相对)执行的规模和配置来配置。因此，通信接口1592可以包括用于为覆盖区域中的多个移动站、UE或其他接入终端服务的适当的设备和电路(可能包括多个收发信机)。

[0219] 处理器1594可以包括由相关领域的技术人员认为合适的任何类型的一个或多个处理器，一些示例包括通用微处理器和专用DSP。

[0220] 数据存储器1596可以采取任何非暂时的计算机可读介质或这种介质的组合的形式，一些示例包括闪速存储器、只读存储器(ROM)和随机存取存储器(RAM)等等，但是可以使用由相关领域的技术人员认为适合的任何一种或多种类型的非暂时数据存储。如图16所示，数据存储器1596包含由处理器1594执行的用于执行本文所述的各种网络实体功能的各种组合的程序指令1597。

[0221] 注意，所描述的实施方式中的一个或多个的各种硬件元件被称为与各个模块连接实施(即，执行、实行等)与本文所描述的各种功能的“模块”。如本文所使用的，针对给定实施模块包括由相关领域的技术人员认为适合的硬件(例如，一个或多个处理器、一个或多个微处理器、一个或多个微控制器、一个或多个微芯片、一个或多个专用集成电路(ASIC)、一个或多个现场可编程门阵列(FPGA)、一个或多个存储器设备)。每个所描述的模块还可以包括可执行用于执行由相应模块执行的一个或多个功能的指令，并且注意，这些指令可以采取或包括硬件(即硬连线)指令、固件指令、软件指令和/或类似的，并且可以存储在任何合适的非暂时计算机可读媒介或介质中，诸如通常称为RAM、ROM等。

[0222] 虽然本发明的特征和元素以特定的结合在以上进行了描述，但本领域普通技术人员可以理解的是，每个特征或元素可以在没有其它特征和元素的情况下单独使用，或在与本发明的任何其它特征和元素结合的各种情况下使用。此外，本发明描述的方法可以在由计算机或处理器执行的计算机程序、软件或固件中实施，其中所述计算机程序、软件或固件被包含在计算机可读存储介质中。关于计算机可读存储介质的示例包括但不局限于只读存储器(ROM)、随机存取存储器(RAM)、寄存器、缓冲存储器、半导体存储设备、磁介质(例如，内部硬盘或可移动磁盘)、磁光介质以及CD-ROM光盘和数字多功能光盘(DVD)之类的光介质。与软件有关的处理器可以被用于实施在WTRU、UE、终端、基站、RNC或者任何主计算机中使用的射频收发信机。

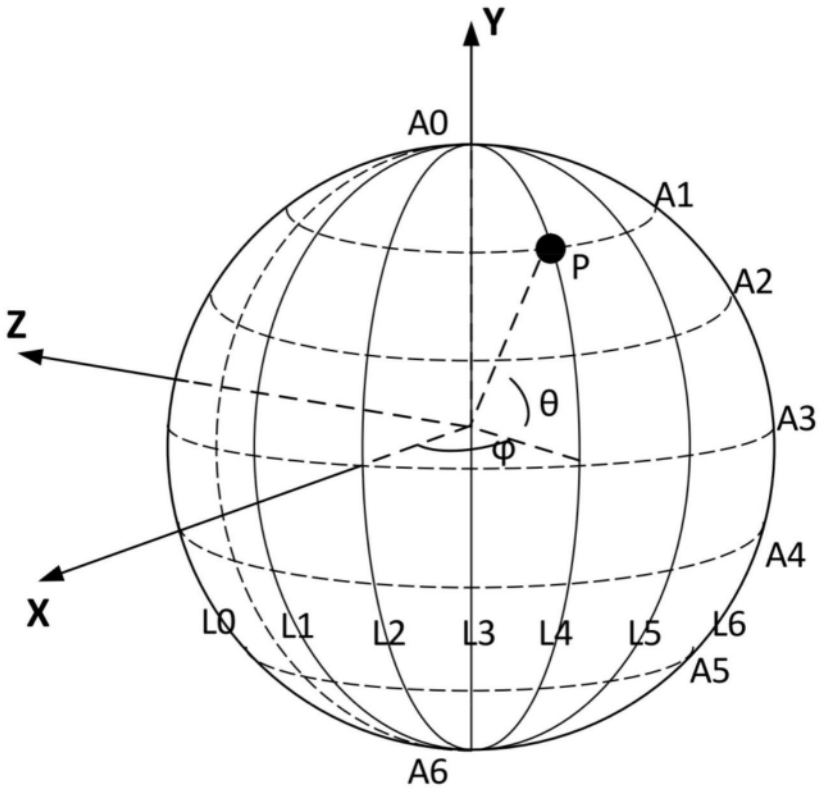


图1A

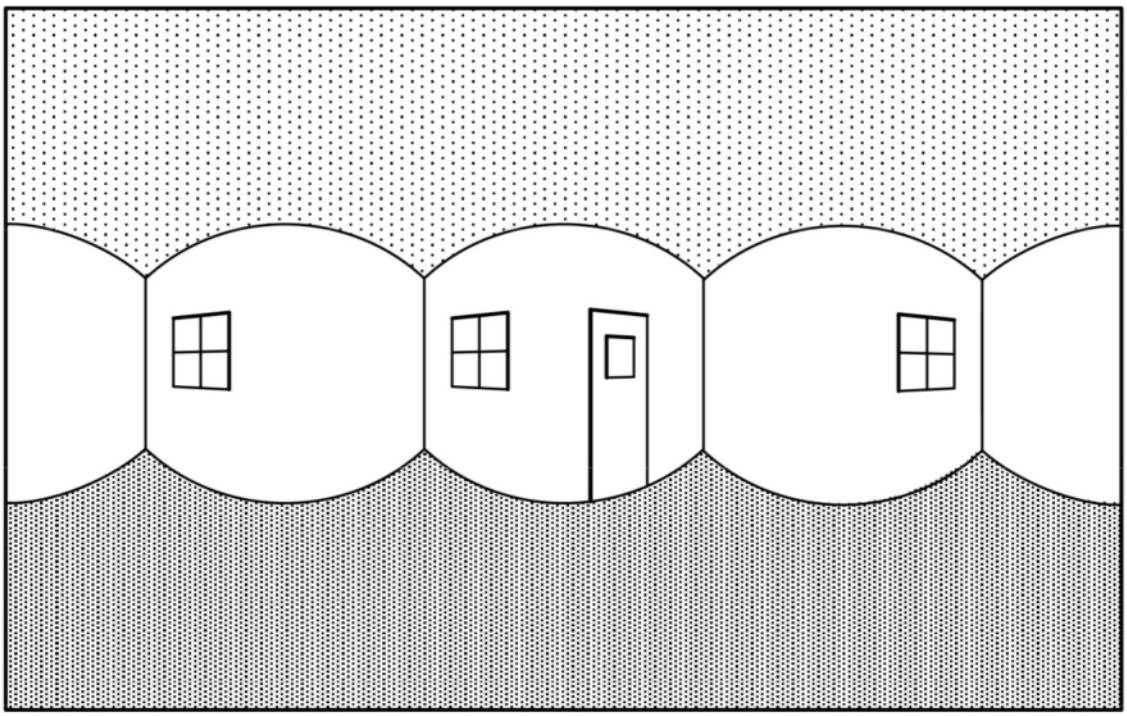


图1C

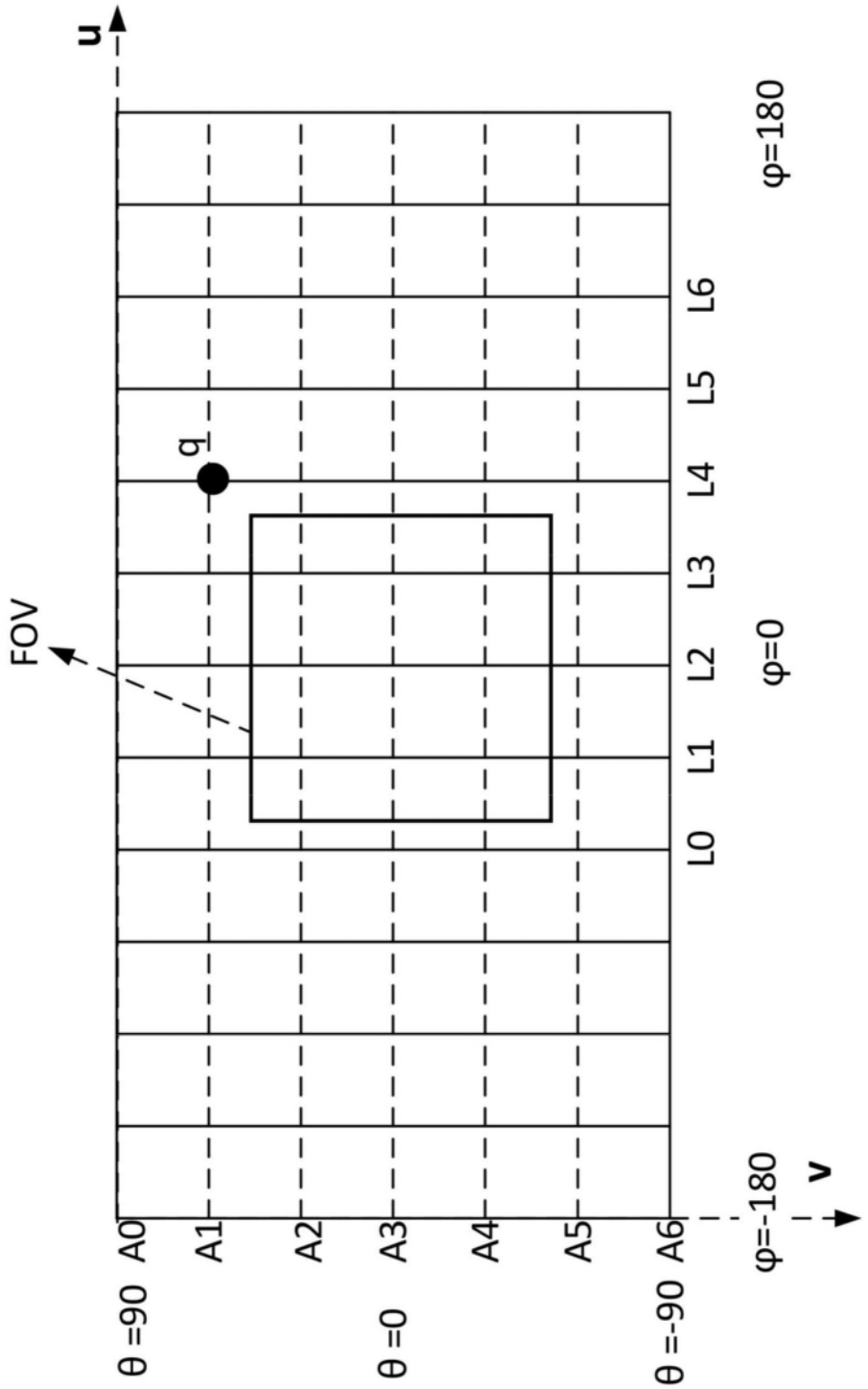

u
v
FOV
q
φ=180
φ=0
φ=-180
L0
L1
L2
L3
L4
L5
L6
A0
A1
A2
A3
A4
A5
A6
θ =90
θ =0
θ =-90

图1B

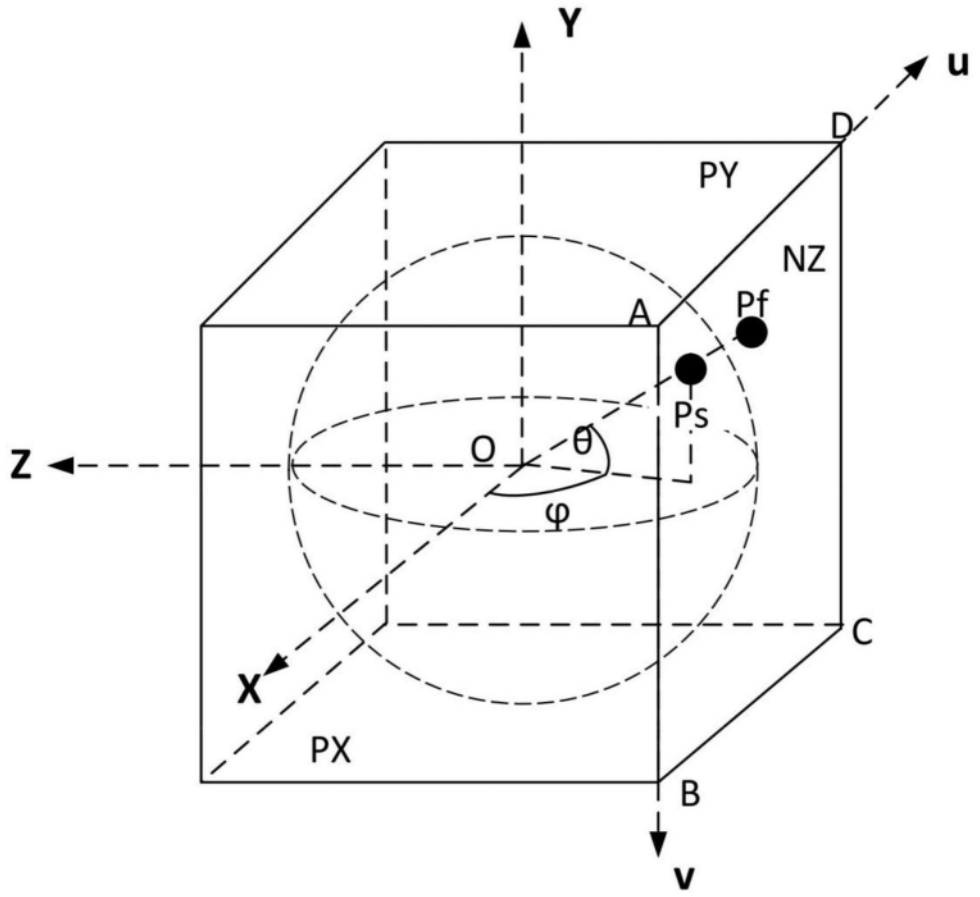


图2A

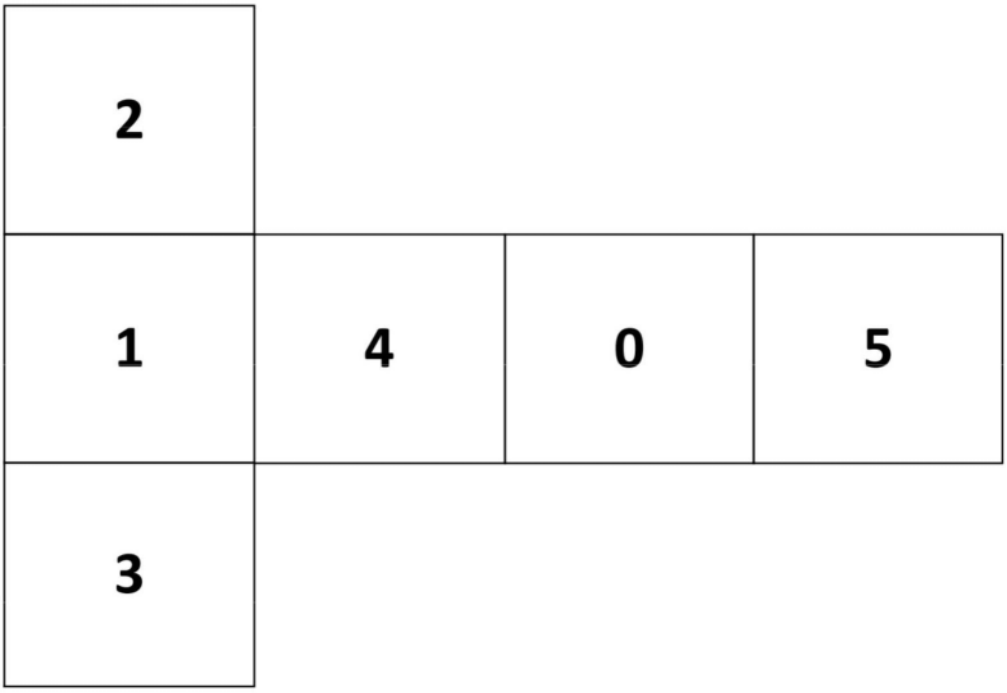


图2B

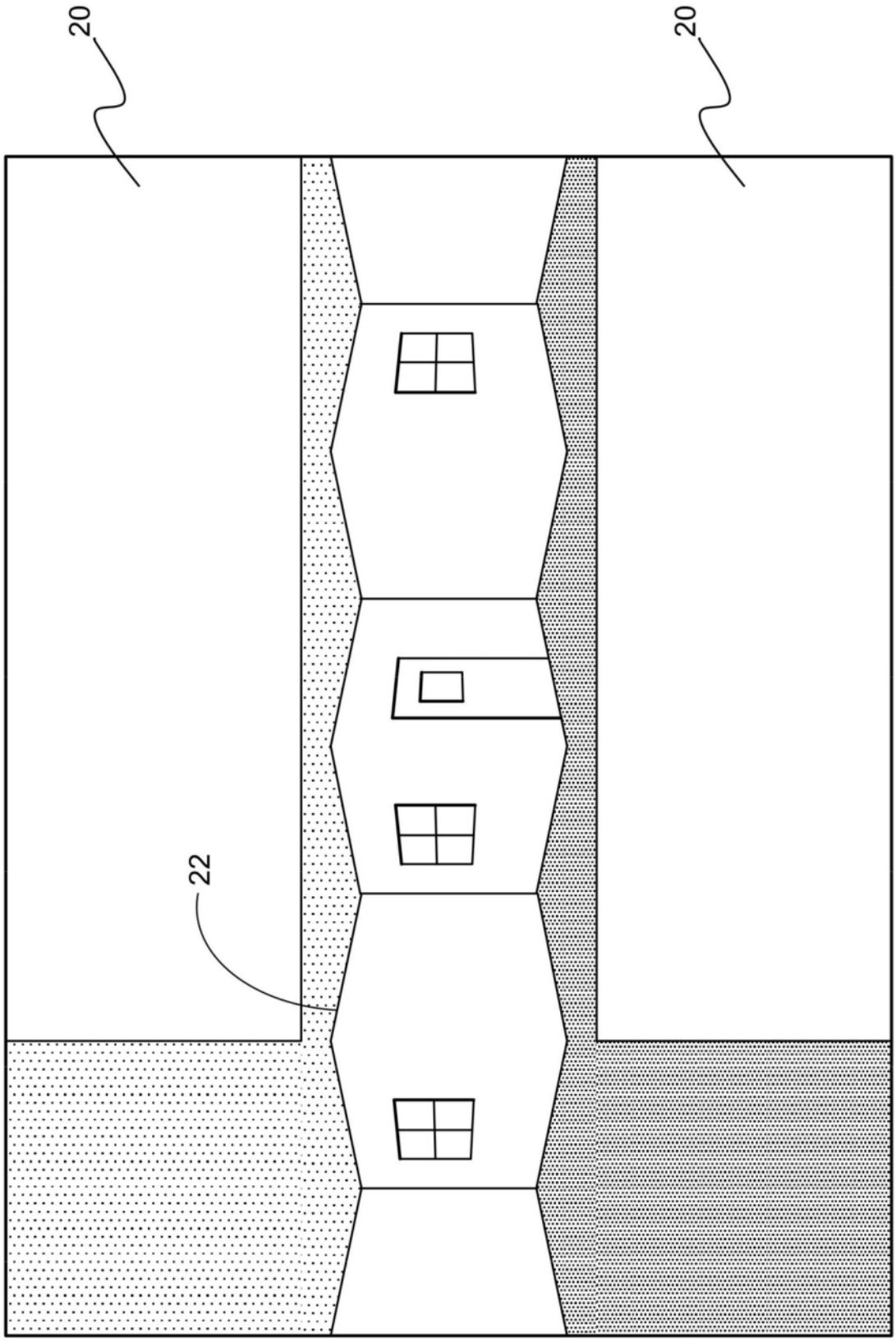


图2C

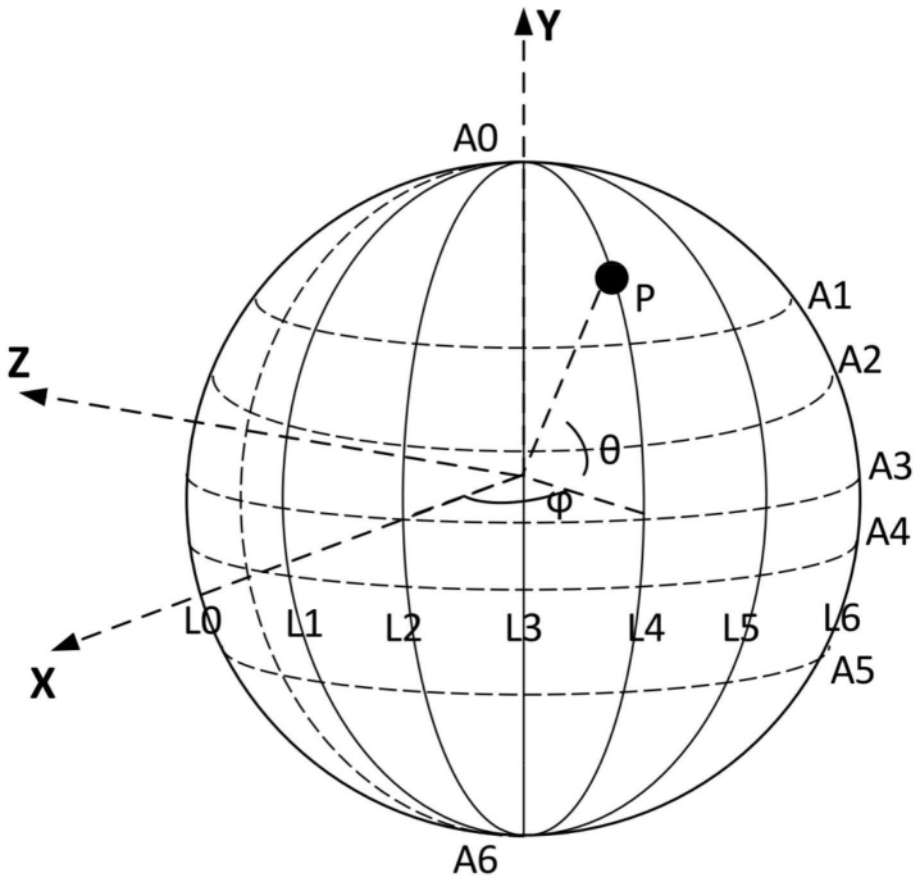


图3A

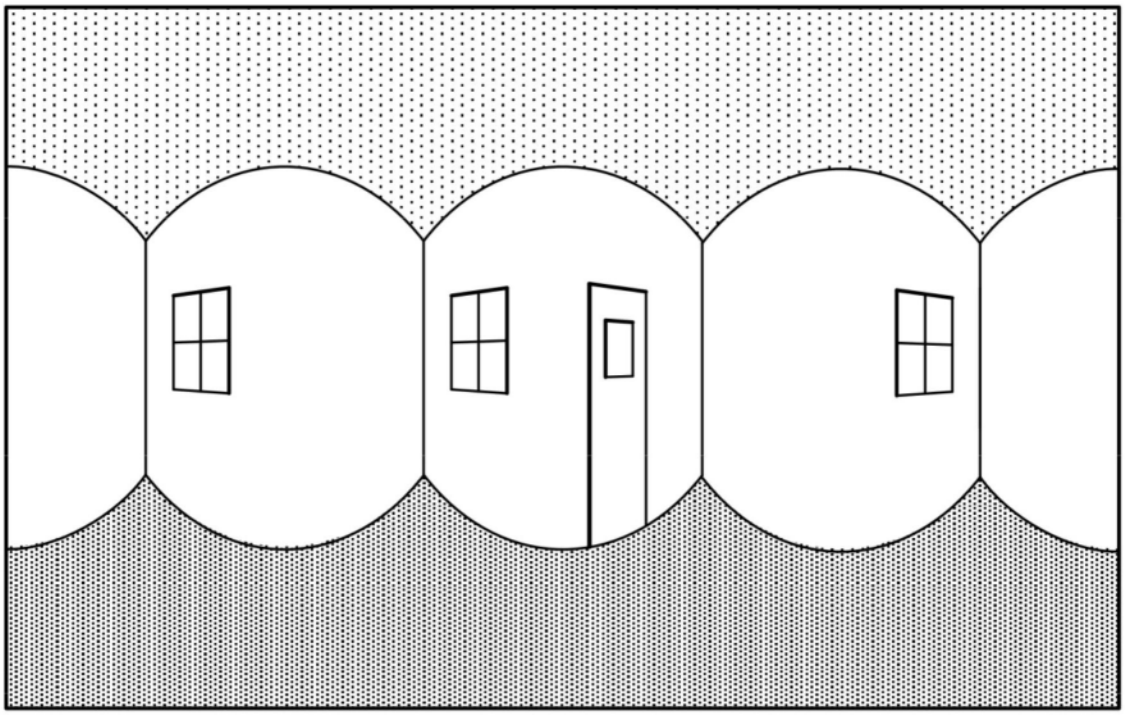


图3C

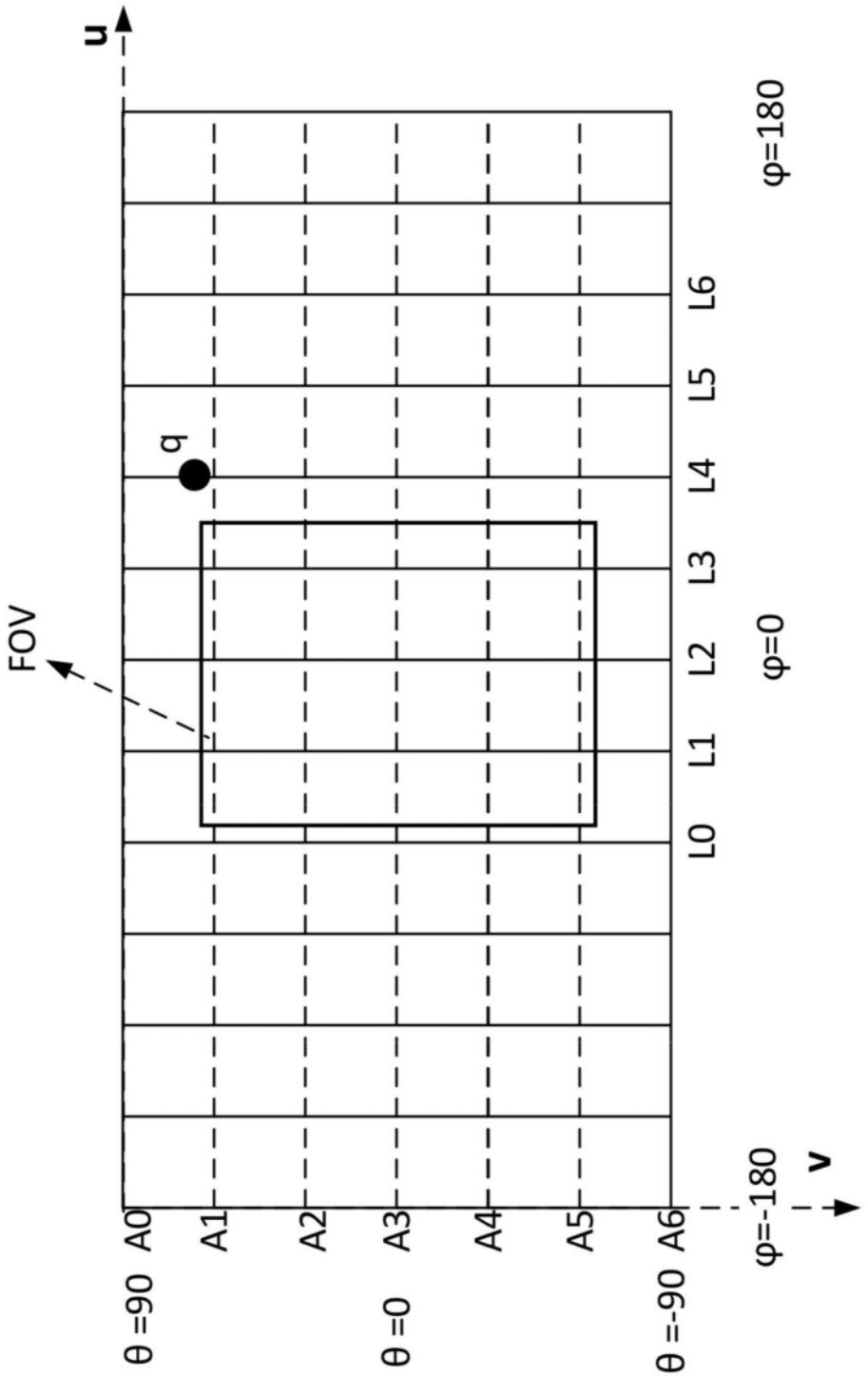

u
v
φ=180
φ=0
φ=-180
L0
L1
L2
L3
L4
L5
L6
A0
A1
A2
A3
A4
A5
A6
θ =90
θ =0
θ -90
q
FOV

图3B

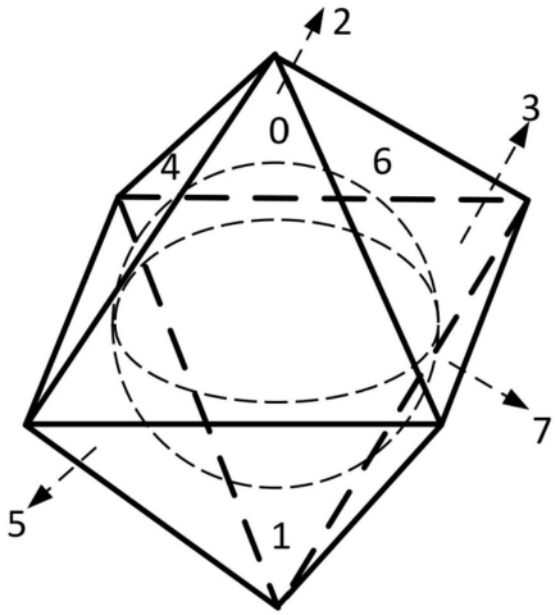


图4A

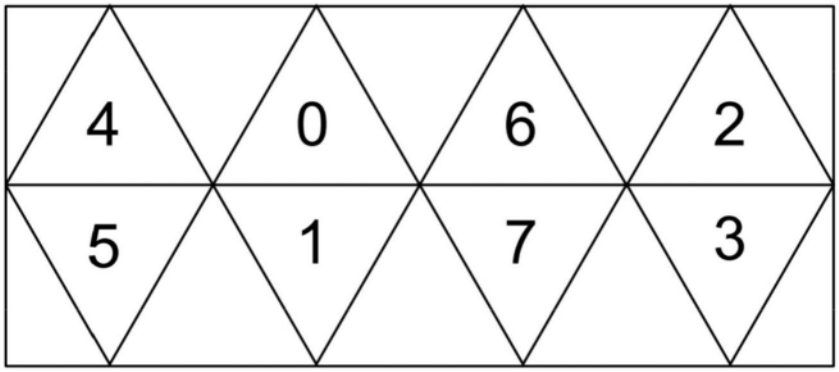


图4B

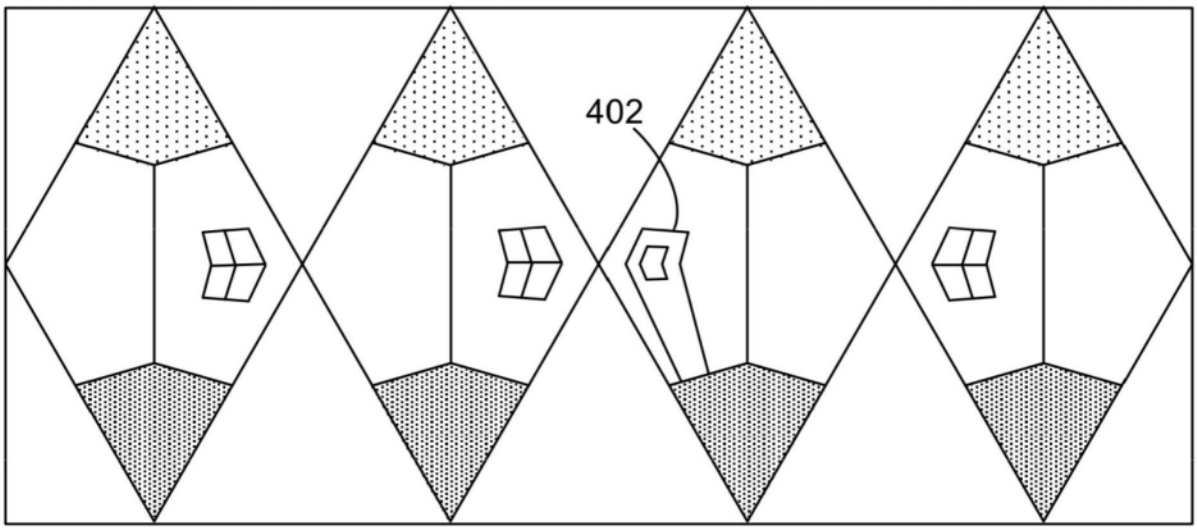


图4C

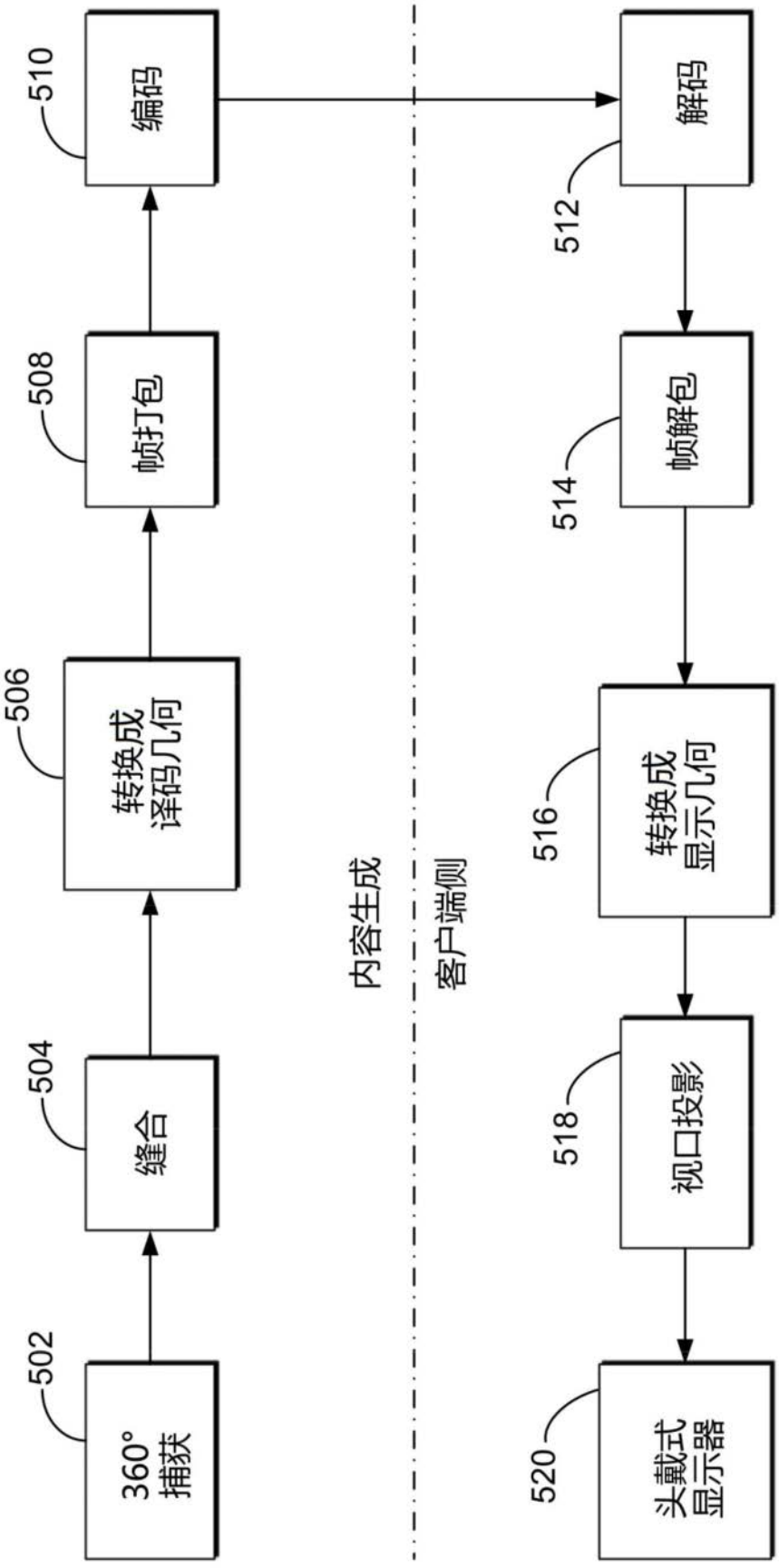

502
360°
捕获
504
缝合
506
转换成
译码几何
508
帧打包
510
编码
512
解码
514
帧解包
516
转换成
显示几何
518
视口投影
520
头戴式
显示器
内容生成
客户端侧

图5

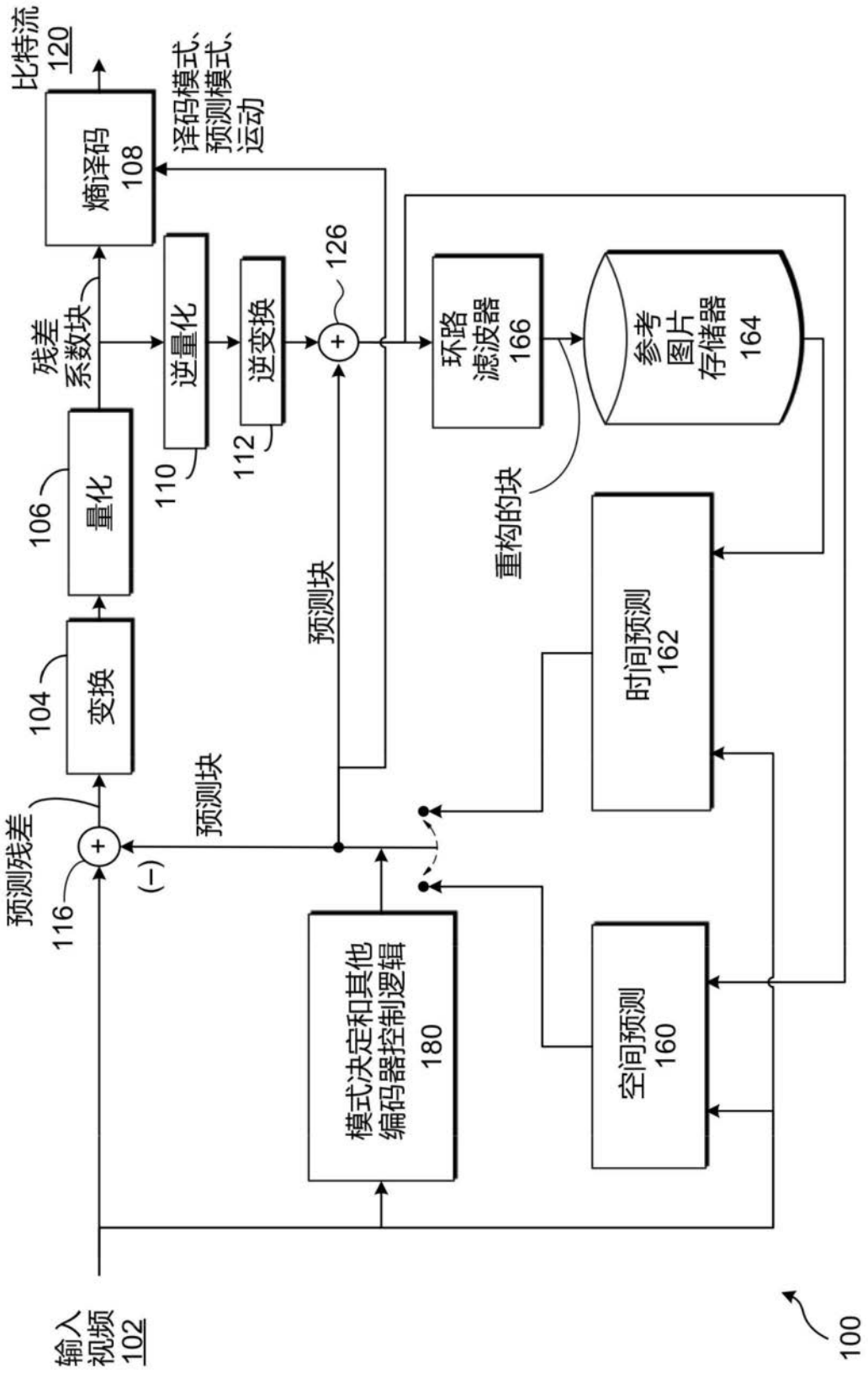

比特流 120
熵译码 108
译码模式、预测模式、运动
残差系数块
逆量化 110
逆变换 112
126
环路滤波器 166
参考图片存储器 164
量化 106
变换 104
预测块
重构的块
时间预测 162
预测残差 116
(−)
预测块
模式决定和其他编码器控制逻辑 180
空间预测 160
输入视频 102
100

图6

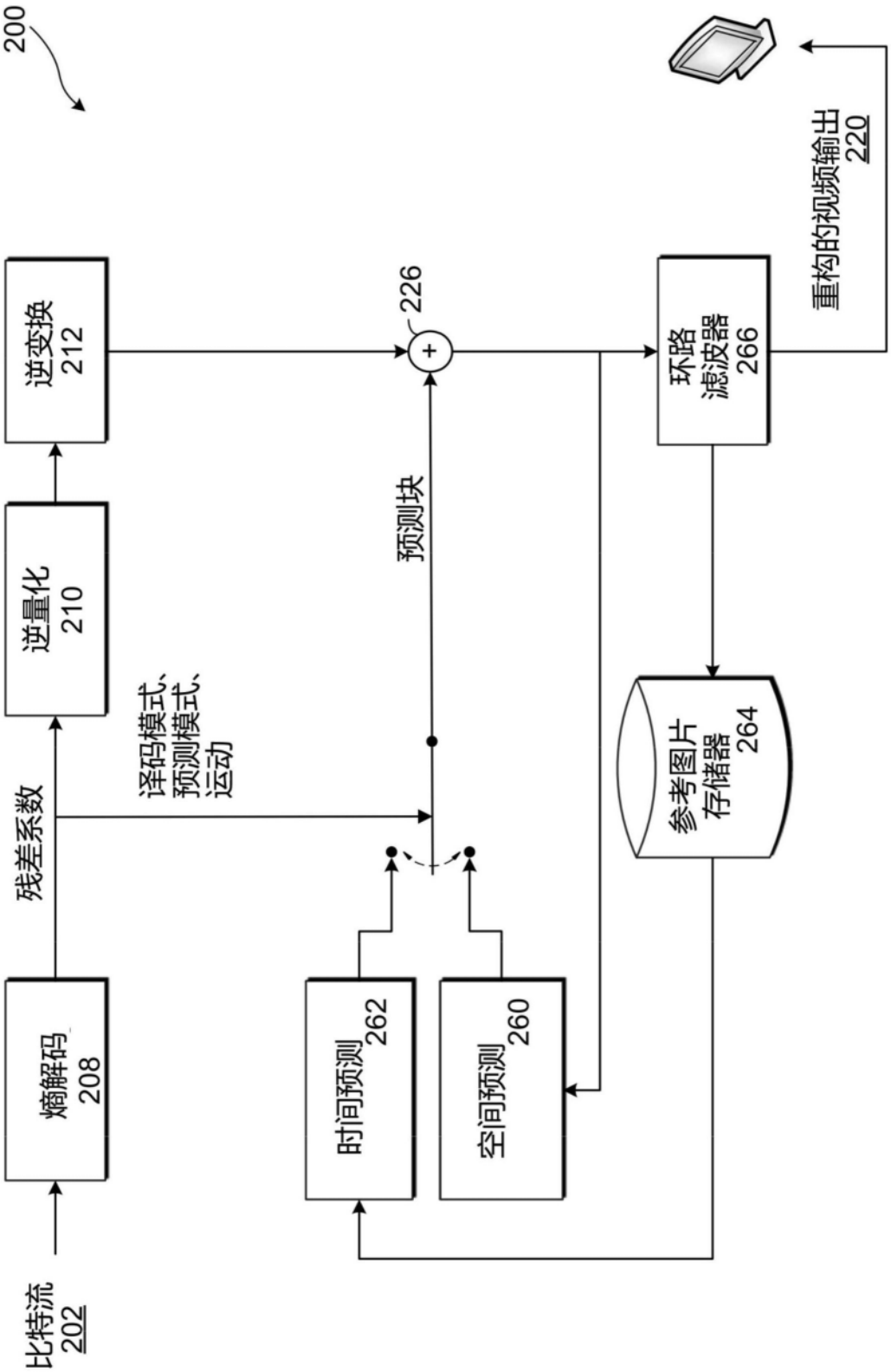

200
逆变换
212
逆量化
210
熵解码
208
比特流
202
残差系数
译码模式、
预测模式、
运动
226
预测块
时间预测
262
空间预测
260
环路
滤波器
266
参考图片
存储器
264
重构的视频输出
220

图7

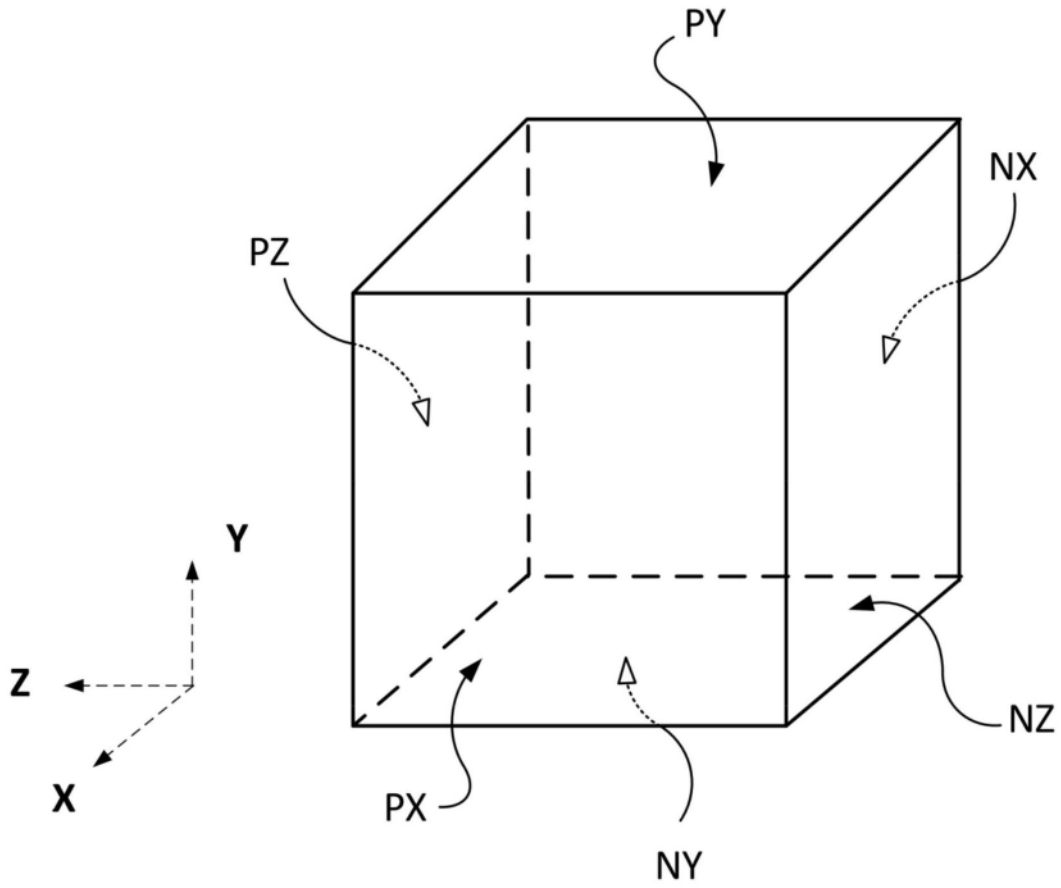


图8A

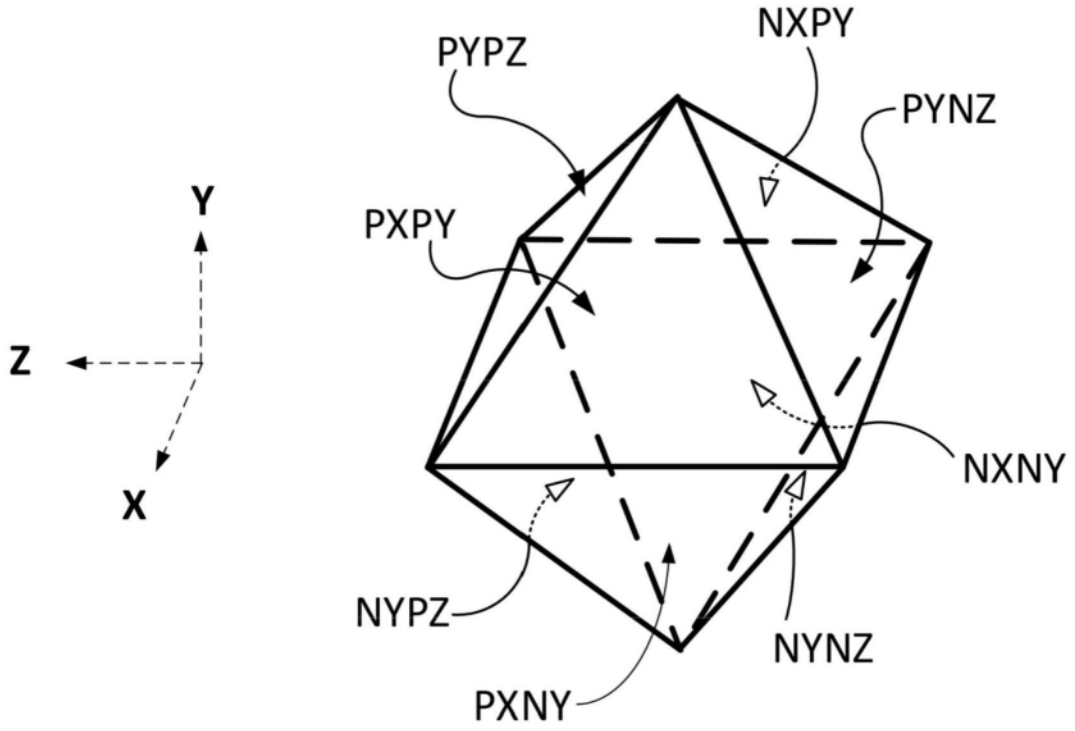


图8B

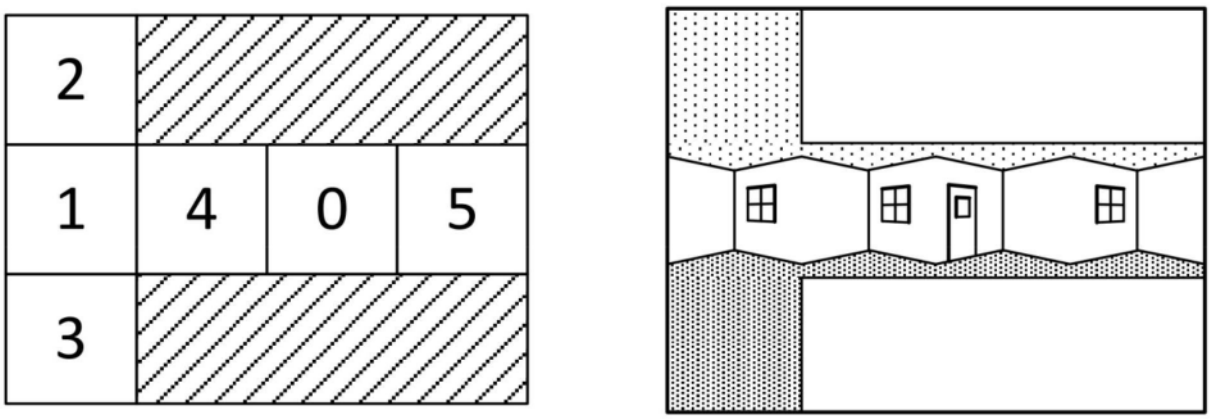


图9A

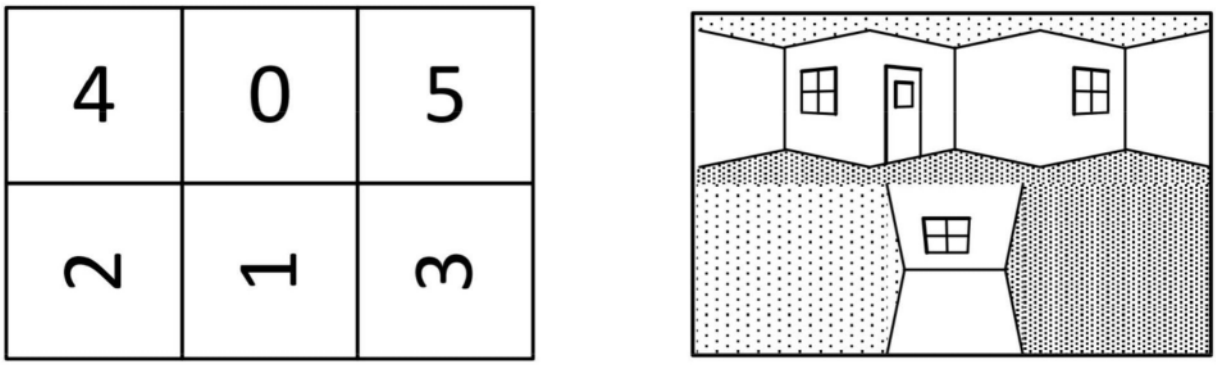


图9B

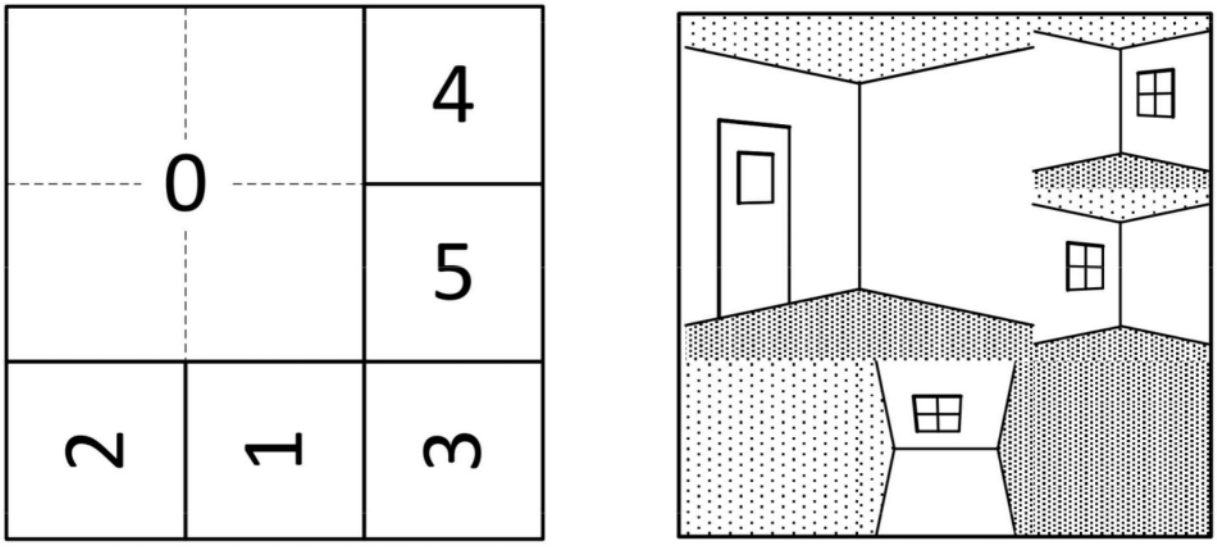


图9C



图10A

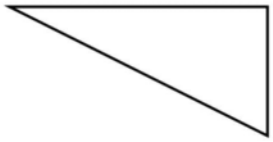


图10B



图10C

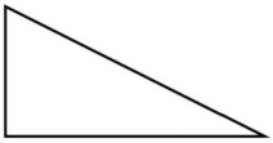


图10D



图10E

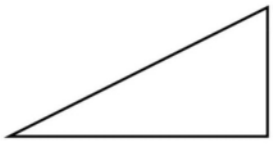


图10F



图10G

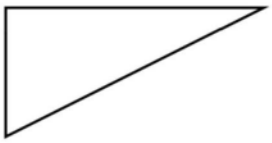


图10H

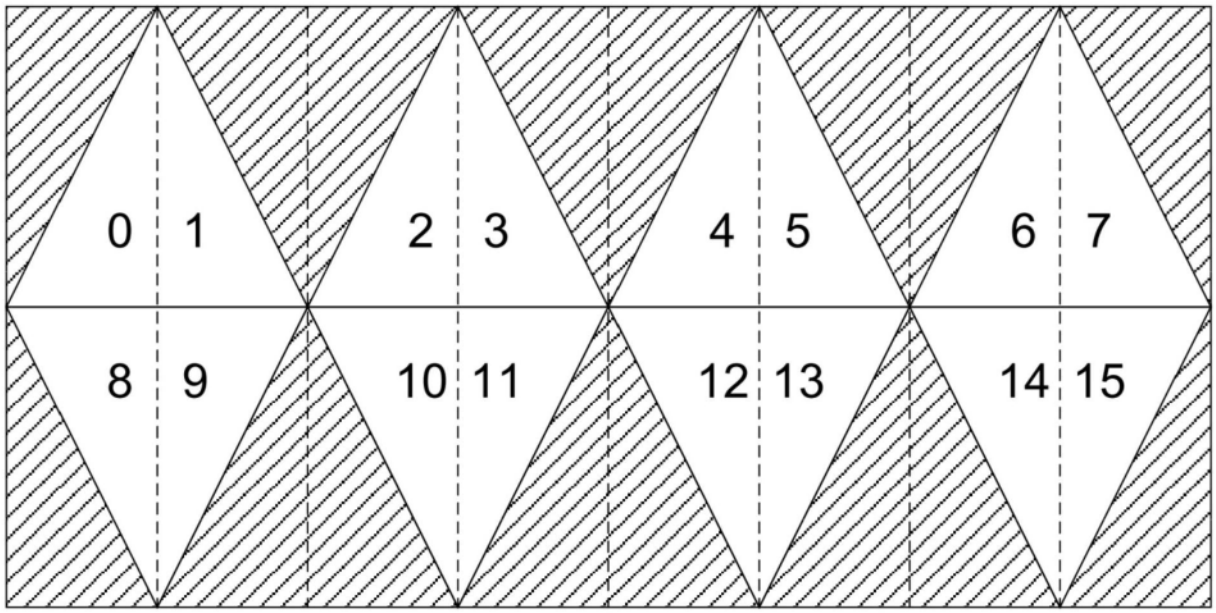


图11A

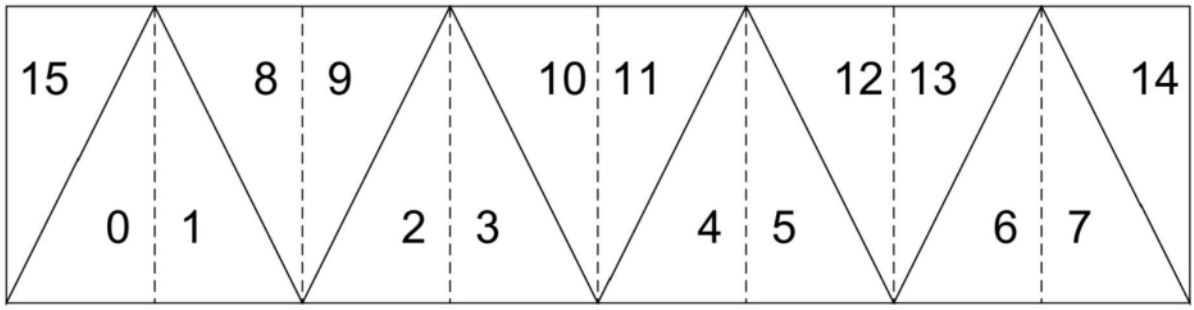


图11B

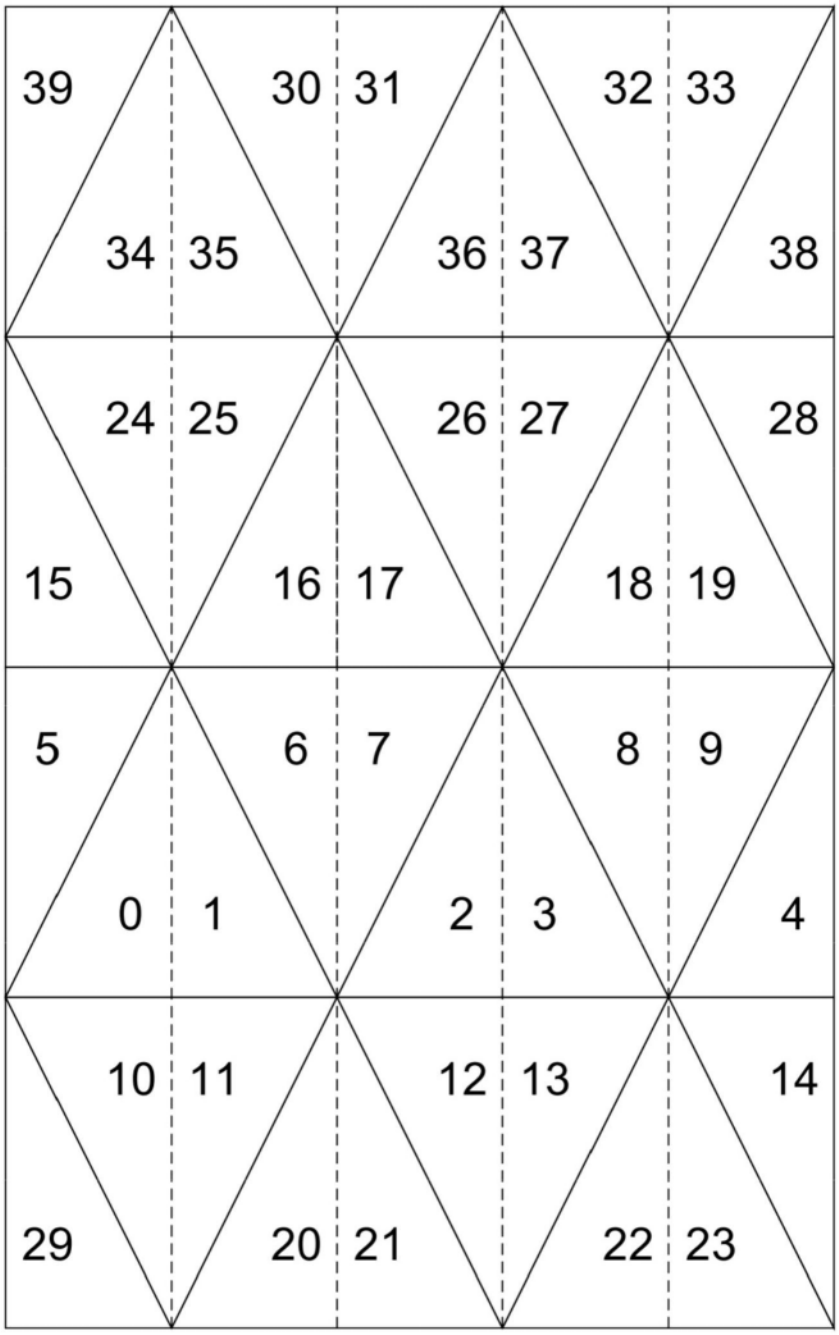

39
30
31
32
33
34
35
36
37
38
24
25
26
27
28
15
16
17
18
19
5
6
7
8
9
0
1
2
3
4
10
11
12
13
14
29
20
21
22
23

图12B

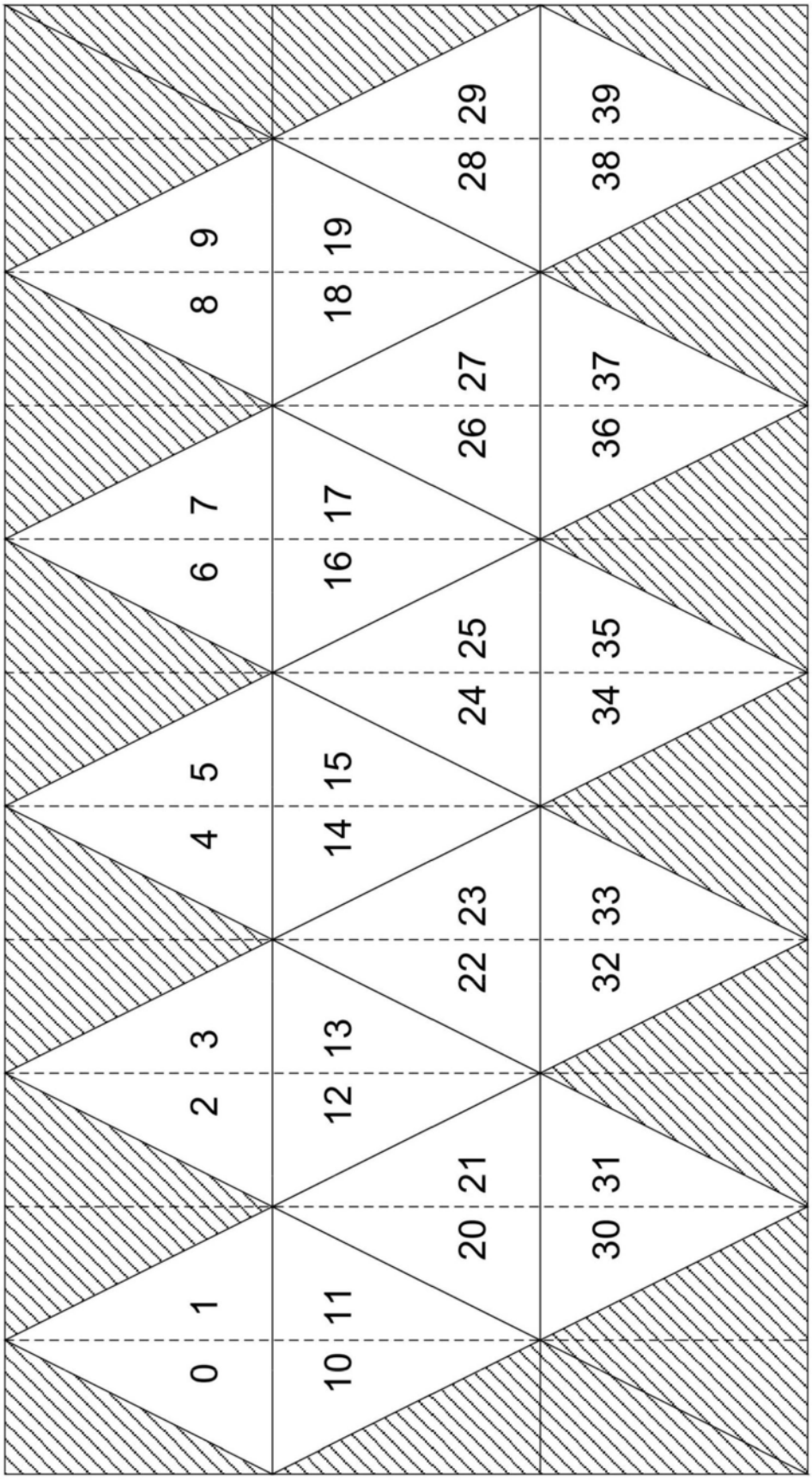

0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

图12A

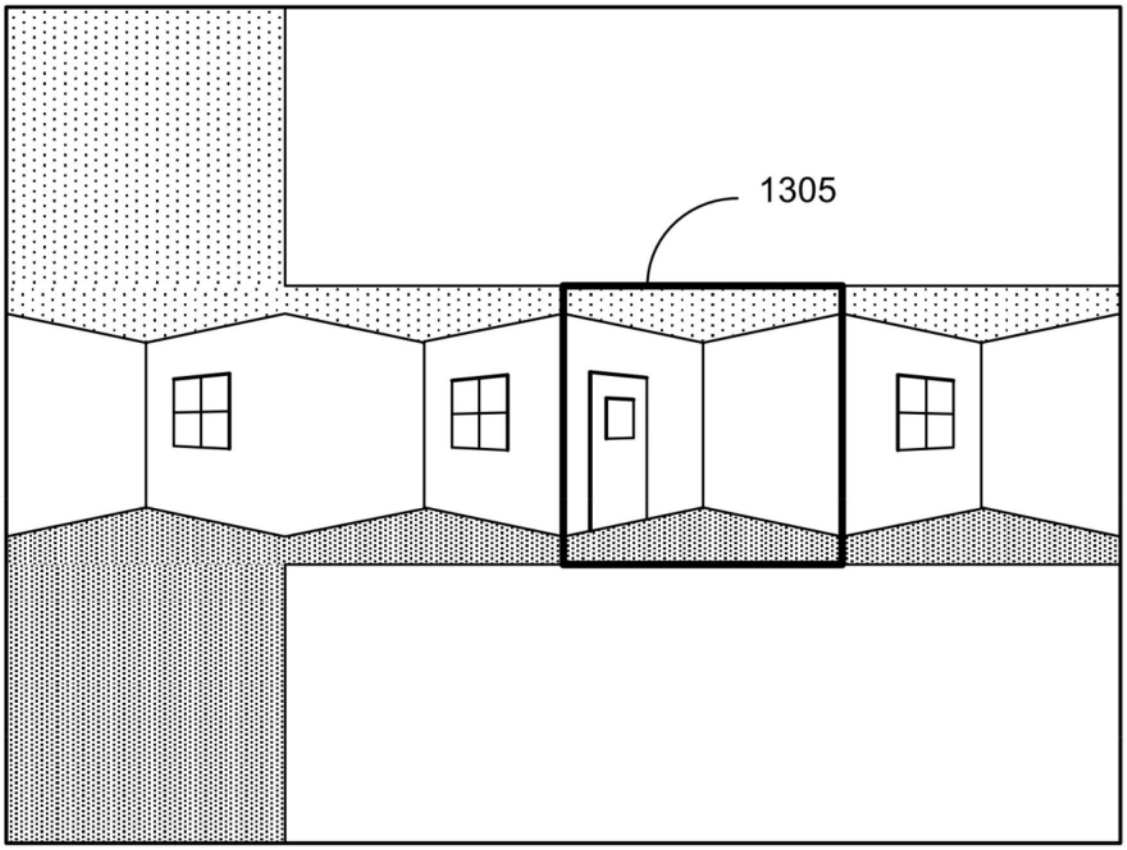


图13A

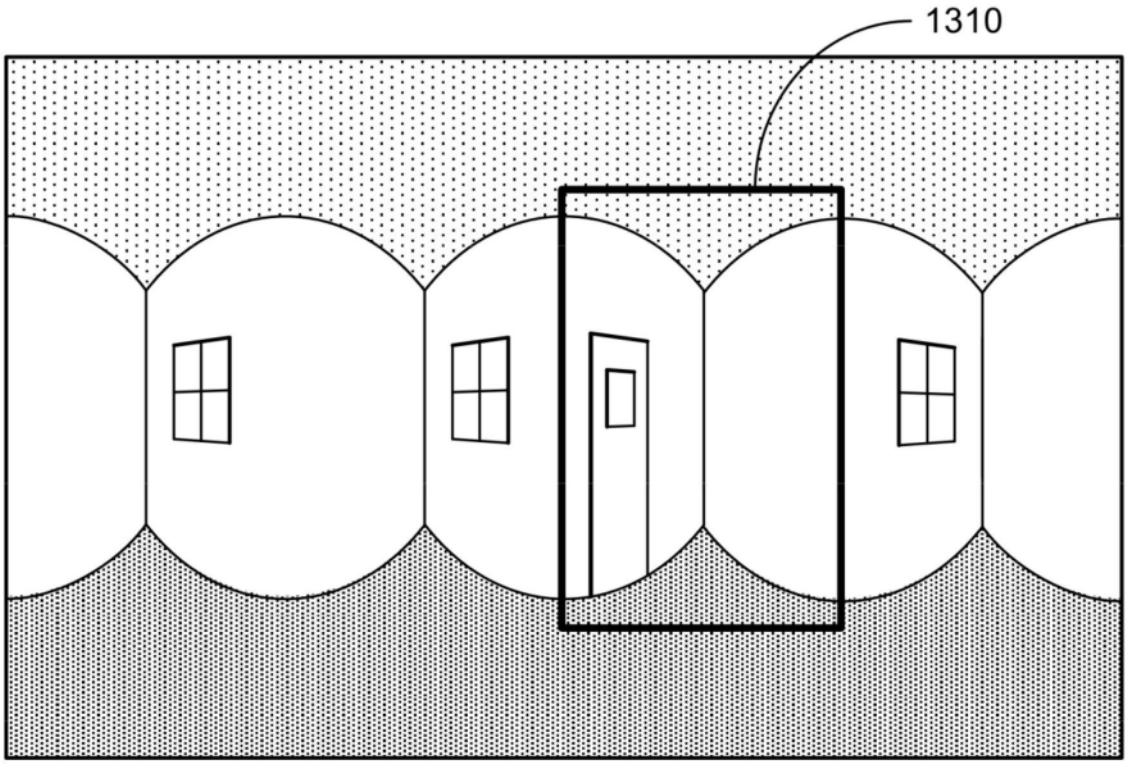


图13B

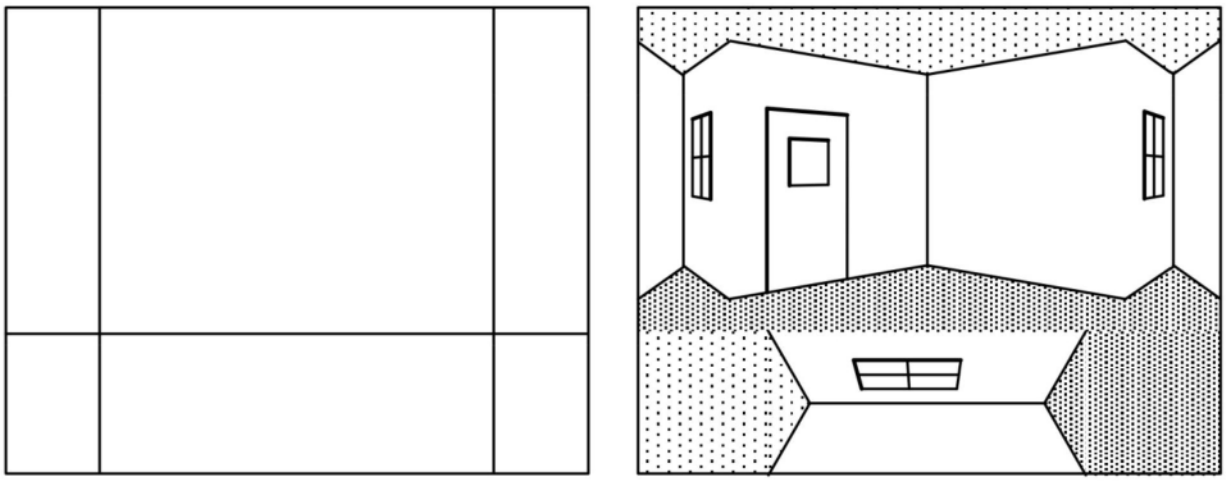


图14A

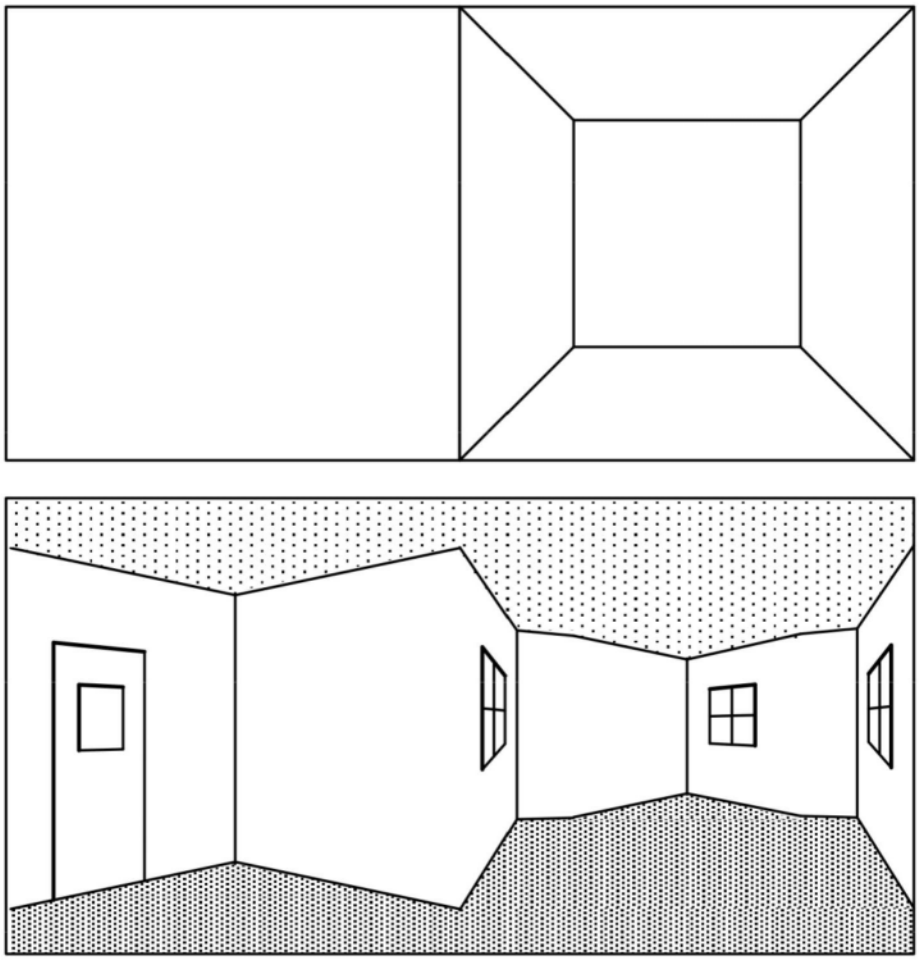


图14B

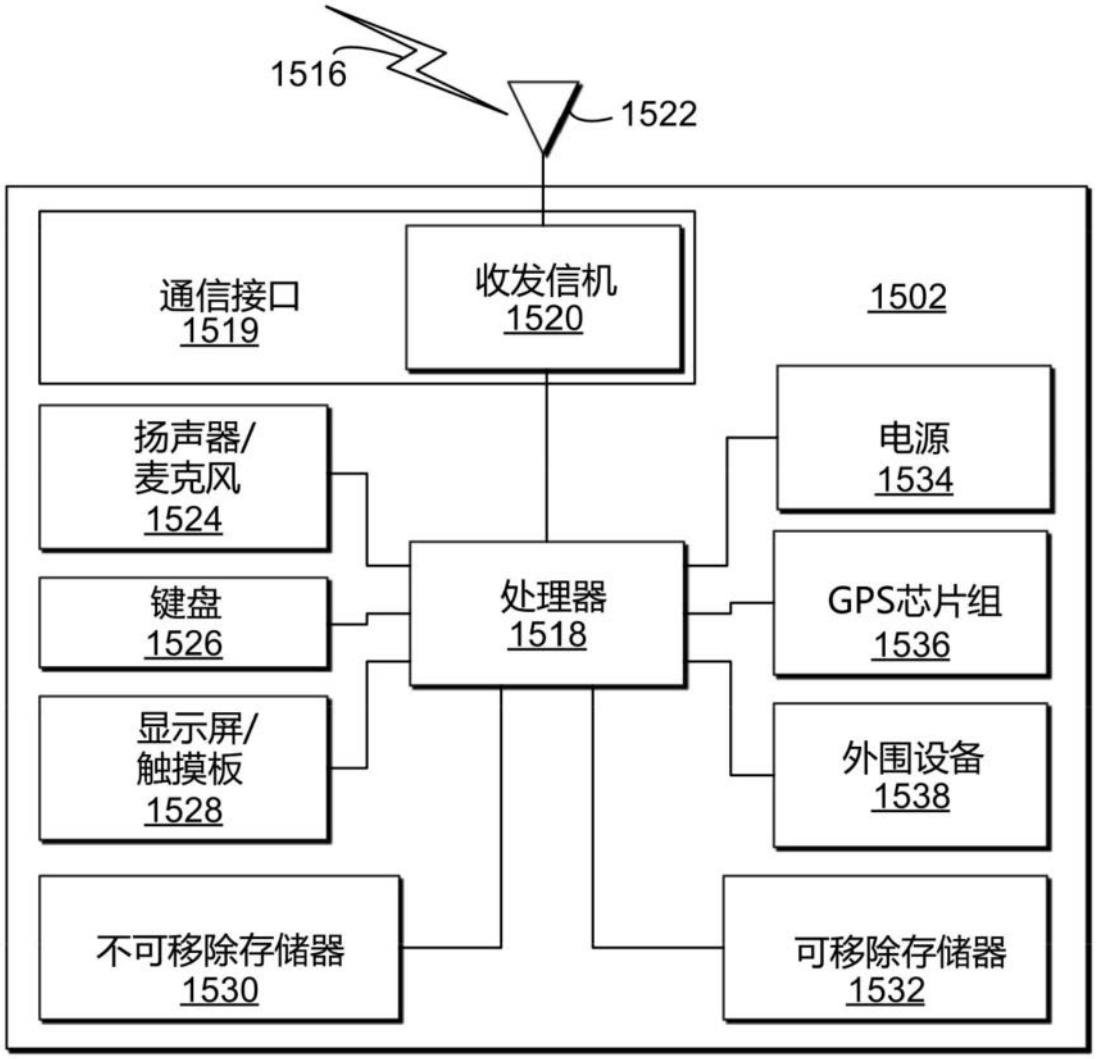

1516
1522
通信接口
1519
收发信机
1520
1502
扬声器/
麦克风
1524
电源
1534
键盘
1526
处理器
1518
GPS芯片组
1536
显示屏/
触摸板
1528
外围设备
1538
不可移除存储器
1530
可移除存储器
1532

图15

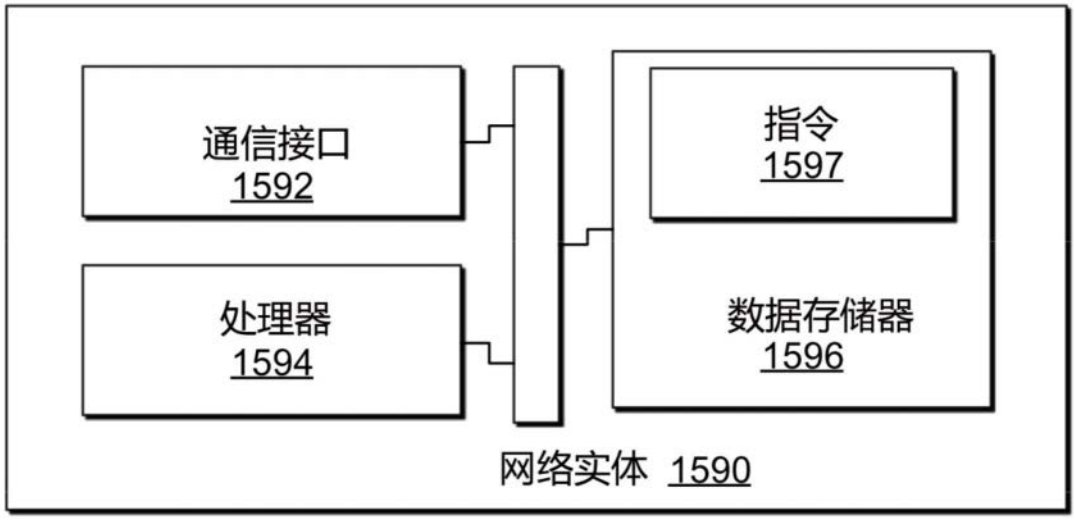

通信接口
1592
指令
1597
处理器
1594
数据存储器
1596
网络实体 1590

图16