



(12) 发明专利

(10) 授权公告号 CN 113170155 B

(45) 授权公告日 2024.02.20

(21) 申请号 202080006706.2
 (22) 申请日 2020.01.06
 (65) 同一申请的已公布的文献号
 申请公布号 CN 113170155 A
 (43) 申请公布日 2021.07.23
 (30) 优先权数据
 62/790,445 2019.01.09 US
 16/731,483 2019.12.31 US
 (85) PCT国际申请进入国家阶段日
 2021.06.03
 (86) PCT国际申请的申请数据
 PCT/US2020/012300 2020.01.06
 (87) PCT国际申请的公布数据
 W02020/146223 EN 2020.07.16
 (73) 专利权人 腾讯美国有限责任公司
 地址 美国加利福尼亚州帕洛阿尔托公园大
 道2747号
 (72) 发明人 沃索基阿拉什 芮世薰 刘杉
 (74) 专利代理机构 北京德琦知识产权代理有限
 公司 11018
 专利代理人 马媛媛 王琦

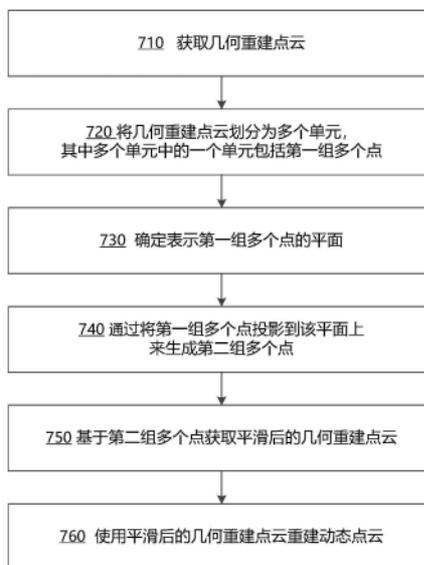
(51) Int.Cl.
 H04N 19/23 (2006.01)
 H04N 19/192 (2006.01)
 H04N 19/176 (2006.01)
 (56) 对比文件
 US 2010008424 A1,2010.01.14
 US 2010119166 A1,2010.05.13
 US 2011304619 A1,2011.12.15
 US 2013321393 A1,2013.12.05
 US 2018225866 A1,2018.08.09
 Rufael Mekuria等.An improved
 enhancement layer for octree based point
 cloud compression with plane projection
 approximation.《PROCEEDINGS OF SPIE》.2016,
 全文.
 Mouna Attia等.Efficient Initial Guess
 Determination Based on 3D Point Cloud
 Projection for ICP Algorithms.《2017
 International Conference on High
 Performance Computing & Simulation》.2017,
 全文. (续)

审查员 蒋碧波

权利要求书3页 说明书9页 附图8页

(54) 发明名称
 点云编解码方法和装置

(57) 摘要
 对使用视频点云编解码机制进行编码的视
 频流进行解码的方法和装置,包括:获取几何重
 建点云;将所述几何重建点云划分为多个单元,
 其中,所述多个单元中的一个单元包括第一组多
 个点;确定表示所述第一组多个点的平面;通过
 将所述第一组多个点投影到所述平面上来生成
 第二组多个点;基于所述第二组多个点获取平滑
 后的几何重建点云;以及使用所述平滑后的几何
 重建点云来重建动态点云。



CN 113170155 B

[接上页]

(56) 对比文件

Gustavo Sandri等.Compression of

Plenoptic Point Clouds.《IEEE TRANSACTIONS
ON IMAGE PROCESSING》.2018,全文.

1. 一种对点云视频流进行解码的方法,其特征在于,所述方法由至少一个处理器执行,包括:

获取几何重建点云;

将所述几何重建点云划分为多个单元,其中,所述多个单元中的一个单元包括第一组多个点;

确定表示所述第一组多个点的平面,其中,所述平面穿过所述第一组多个点的平均位置 \bar{p} ,并且所述平面的法向量具有与 v_3 相同的方向,所述 v_3 表示所述第一组多个点的协方差矩阵的最小特征值 λ_3 的特征向量;

通过将所述第一组多个点投影到所述平面上来生成第二组多个点,包括:将所述第一组多个点与所述第一组多个点中的点在所述平面上的投影的加权平均值,作为所述第二组多个点,其中在计算所述加权平均值时所述第一组多个点的权重为 α ,所述第一组多个点中的点在所述平面上的投影的权重为 $1-\alpha$;

基于所述第二组多个点获取平滑后的几何重建点云;以及
使用所述平滑后的几何重建点云来重建动态点云。

2. 根据权利要求1所述的方法,其特征在于,所述平均位置 \bar{p} 按如下方式确定:

$$\bar{p} = \frac{\sum_{i=1}^N x_i}{N}$$

其中, x_i 表示所述第一组多个点中索引为 i 的点, N 表示所述第一组多个点的数目。

3. 根据权利要求2所述的方法,其特征在于,所述第一组多个点的协方差矩阵 C 以如下方式确定:

$$C = \frac{\sum_{i=1}^N (x_i - \bar{p})(x_i - \bar{p})^T}{N}$$

其中, T 表示转置操作,

其中, C 的特征值包括: λ_3 、 λ_2 和 λ_1 ,其中, $\lambda_3 < \lambda_2 < \lambda_1$,

其中, λ_3 、 λ_2 、 λ_1 的特征向量由 v_3 、 v_2 、 v_1 表示,以及

其中,基于所述平均位置 \bar{p} 和 v_3 确定所述平面。

4. 根据权利要求3所述的方法,其特征在于,所述将所述第一组多个点投影到所述平面以如下方式确定:

$$X_{proj} = [(X - \bar{p}) \cdot v_1]v_1 + [(X - \bar{p}) \cdot v_2]v_2 + \bar{p}$$

其中, X 表示所述第一组多个点中的点,并且 X_{proj} 表示第一组多个点中的点在所述平面上的投影。

5. 根据权利要求4所述的方法,其特征在于,所述第二组多个点中的点以如下方式确定:

$$X \rightarrow \alpha X_{proj} + (1-\alpha)X$$

其中, X 表示所述第一组多个点中的点, X_{proj} 表示所述第一组多个点中的所述点在所述平面上的投影, $\alpha X_{proj} + (1-\alpha)X$ 表示所述第二组多个点中的点。

6. 根据权利要求1-5中任一权利要求所述的方法,其特征在于,所述平面包括表示所述

第一组多个点的平面流形。

7. 根据权利要求1-5中任一权利要求所述的方法,其特征在于,所述视频流的元数据指示以下中的至少一个:是否使用局部投影平滑的指示、所述单元的大小的指示,以及在所述投影中使用的投影权重。

8. 一种用于对点云视频流进行解码的装置,其特征在于,所述装置包括:

至少一个存储器,用于存储程序代码;以及

至少一个处理器,用于读取所述程序代码并按照所述程序代码的指示进行操作,所述程序代码包括:

第一获取代码,用于使所述至少一个处理器获取几何重建点云;

划分代码,用于使所述至少一个处理器将所述几何重建点云划分为多个单元,其中,所述多个单元中的一个单元包括第一组多个点;

确定代码,用于使所述至少一个处理器确定表示所述第一组多个点的平面,其中,所述平面穿过所述第一组多个点的平均位置 \bar{p} ,并且所述平面的法向量具有与 v_3 相同的方向,所述 v_3 表示所述第一组多个点的协方差矩阵的最小特征值 λ_3 的特征向量;

生成代码,用于使所述至少一个处理器通过将所述第一组多个点投影到所述平面上来生成第二组多个点,其中生成所述第二组多个点包括:将所述第一组多个点与所述第一组多个点中的点在所述平面上的投影的加权平均值,作为所述第二组多个点,其中在计算所述加权平均值时所述第一组多个点的权重为 α ,所述第一组多个点中的点在所述平面上的投影的权重为 $1-\alpha$;

第二获取代码,用于使所述至少一个处理器基于所述第二组多个点获取平滑后的几何重建点云;以及

重建代码,用于使所述至少一个处理器使用所述平滑后的几何重建点云来重建动态点云。

9. 根据权利要求8所述的装置,其特征在于,所述平均位置 \bar{p} 按如下方式确定:

$$\bar{p} = \frac{\sum_{i=1}^N x_i}{N}$$

其中, x_i 表示所述第一组多个点中索引为 i 的点, N 表示所述第一组多个点的数目。

10. 根据权利要求9所述的装置,其特征在于,所述第一组多个点的协方差矩阵 C 以如下方式确定:

$$C = \frac{\sum_{i=1}^N (x_i - \bar{p})(x_i - \bar{p})^T}{N}$$

其中, T 表示转置操作,

其中, C 的特征值包括: λ_3 、 λ_2 和 λ_1 ,其中, $\lambda_3 < \lambda_2 < \lambda_1$,

其中, λ_3 、 λ_2 、 λ_1 的特征向量由 v_3 、 v_2 、 v_1 表示,以及

其中,基于所述平均位置 \bar{p} 和 v_3 确定所述平面。

11. 根据权利要求10所述的装置,其特征在于,所述第一组多个点中的点在所述平面上的投影以如下方式确定:

$$X_{proj} = [(X - \bar{p}) \cdot v_1]v_1 + [(X - \bar{p}) \cdot v_2]v_2 + \bar{p}$$

其中, X 表示所述第一组多个点中的点, 并且 X_{proj} 表示所述第一组多个点中的点在所述平面上的投影。

12. 根据权利要求11所述的装置, 其特征在于, 所述第二组多个点中的点以如下方式确定:

$$X \rightarrow \alpha X_{proj} + (1 - \alpha) X$$

其中, X 表示所述第一组多个点中的点, X_{proj} 表示所述第一组多个点中的所述点在所述平面上的投影, $\alpha X_{proj} + (1 - \alpha) X$ 表示所述第二组多个点中的点。

13. 根据权利要求8-12中任一权利要求所述的装置, 其特征在于, 所述平面包括表示所述第一组多个点的平面流形。

14. 根据权利要求8-12中任一权利要求所述的装置, 其特征在于, 所述视频流的元数据指示以下中的至少一个: 是否使用局部投影平滑的指示、所述单元的大小的指示, 以及在所述投影中使用的投影权重。

15. 一种非易失性计算机可读介质, 存储有用于对点云视频流进行解码的计算机指令, 其特征在于, 所述计算机指令在由至少一个处理器执行时使所述至少一个处理器执行根据权利要求1-7中任一权利要求所述的方法。

点云编解码方法和装置

[0001] 相关申请

[0002] 本申请要求2019年1月9日提交的第62/790,445号美国临时申请的优先权、以及2019年12月31日提交的第16/731,483号美国申请的优先权,所述申请以全文引用方式并入本申请。

技术领域

[0003] 本申请涉及一组高级视频编解码技术,具体地,涉及基于视频的点云压缩。

背景技术

[0004] 世界的高级三维(3D)表示使更加沉浸式的交互和通信成为可能,并且还使机器能够理解、解释并遨游世界。3D点云(point cloud)已经成为实现此类信息的代表。目前已经发现了与点云数据相关联的多个用例,并且已经开发了对点云表示和压缩的相应要求。

[0005] 点云可以是3D空间中的一组点,每个点具有相关联的属性,例如颜色、材料属性等。点云可用于将对象或场景重建为这种点的集合。可以在各种环境中使用多个照相机和深度传感器来获得点云,点云可以由数千到数十亿个点组成,能够真实地表示重建的场景。

[0006] 需要使用压缩技术来减少用于表示点云的数据量。因此,需要针对实时通信和六自由度(degrees of freedom, DoF)虚拟现实中的点云进行有损压缩的技术。此外,在用于自动驾驶和文化遗产应用等的动态绘图的场景中,需要对点云进行无损压缩的技术。运动图像专家组(Moving Picture Experts Group, MPEG)已经开始研究一种处理几何和属性的压缩标准,属性可包括例如颜色和反射率、可缩放/渐进编解码、跨越一定时间获得的点云序列的编解码,以及对点云的子集的随机访问。

发明内容

[0007] 一个实施例中,一种对使用视频点云编解码机制进行编码的视频流进行解码的方法,由至少一个处理器执行,包括:获取几何重建点云;将所述几何重建点云划分为多个单元,其中,所述多个单元中的一个单元包括第一组多个点;确定表示所述第一组多个点的平面;通过将所述第一组多个点投影到所述平面上来生成第二组多个点;基于所述第二组多个点获取平滑后的几何重建点云;以及使用所述平滑后的几何重建点云来重建动态点云。

[0008] 一个实施例中,一种用于对使用视频点云编解码机制进行编码的视频流进行解码的装置,包括:至少一个存储器,用于存储程序代码;以及至少一个处理器,用于读取所述程序代码并按照所述程序代码的指示进行操作,所述程序代码包括:第一获取代码,用于使所述至少一个处理器获取几何重建点云;划分代码,用于使所述至少一个处理器将所述几何重建点云划分为多个单元,其中,所述多个单元中的一个单元包括第一组多个点;确定代码,用于使所述至少一个处理器确定表示所述第一组多个点的平面;生成代码,用于使所述至少一个处理器通过将所述第一组多个点投影到所述平面上来生成第二组多个点;第二获取代码,用于使所述至少一个处理器基于所述第二组多个点获取平滑后的几何重建点云;

以及重建代码,用于使所述至少一个处理器使用所述平滑后的几何重建点云来重建动态点云。

[0009] 一个实施例中,一种非易失性计算机可读介质,存储有对使用视频点云编解码机制进行编码的视频流进行解码的计算机指令,所述计算机指令由至少一个处理器执行时,使该至少一个处理器:获取几何重建点云;将所述几何重建点云划分为多个单元,其中,所述多个单元中的一个单元包括第一组多个点;确定表示所述第一组多个点的平面;通过将所述第一组多个点投影到所述平面上来生成第二组多个点;基于所述第二组多个点获取平滑后的几何重建点云;以及使用所述平滑后的几何重建点云来重建动态点云。

附图说明

[0010] 根据以下详细描述和附图,本申请主题的其他特征、性质和各种优点将变得更加明显,在附图中:

[0011] 图1是一个实施例的通信系统的简化框图的示意图。

[0012] 图2是一个实施例的流媒体系统的简化框图的示意图。

[0013] 图3是一个实施例的视频编码器的简化框图的示意图。

[0014] 图4是一个实施例的视频解码器的简化框图的示意图。

[0015] 图5是一个实施例的几何形状图像示例的示意图。

[0016] 图6是一个实施例的纹理图像示例的示意图。

[0017] 图7是一个实施例所执行的方法的流程图。

[0018] 图8是一个实施例的装置的示意图。

[0019] 图9是适于实现各实施例的计算机系统的示意图。

具体实施方式

[0020] 基于视频的点云压缩(video-based point cloud compression,V-PCC)背后的考虑因素是利用现有的视频编解码器将动态点云的几何形状、占用率和纹理压缩为三个单独的视频序列。用于解释三个视频序列的额外元数据可以单独压缩。整个比特流的一小部分是元数据,可以使用软件方式对其进行有效的编码/解码。大部分信息可以由视频编解码器处理。

[0021] 参考图1至图4,其中描述了用于实现本申请的编码结构和解码结构的本申请实施例。本申请的编码结构和解码结构可以实现上述V-PCC的一些方面。

[0022] 图1为本申请实施例的通信系统100的简化框图。系统100可包括通过网络150互联的至少两个终端110、120。对于单向数据传输,第一终端110可在本地位置编码视频数据,用于经网络150传输至另一终端120。第二终端120可从网络150接收另一终端的编码视频数据,解码该编码数据并显示恢复出的视频数据。单向数据传输常用于媒体服务应用及其它类似应用。

[0023] 图1示出第二对终端130、140,其可以支持已编码视频的双向传输,例如,在视频会议期间发生的双向传输。对于双向的数据传输,每个终端130、140可对在本地位置拍摄的视频数据进行编码,以便通过网络150传输至另一终端。每个终端130、140还可接收由另一终端传输的已编码视频数据,可解码该已编码数据,并可在本地显示器设备显示恢复出的视

频数据。

[0024] 在图1中,终端110-140可以是,例如服务器、个人计算机、智能电话和/或其它任何类型的终端。例如,终端(110-140)可以是笔记本电脑、平板电脑、媒体播放器、和/或专用的视频会议装置。网络150表示任何数量的网络,可以在终端110-140之间传输已编码视频数据,可以包括,例如,有线和/或无线通信网络。通信网络150可在电路交换和/或包交换的信道中交换数据。代表性的网络包括电信网络、局域网、广域网、和/或互联网。为本文讨论的目的,除非有明确说明,网络150的体系结构和拓扑与本申请公开的操作无关。

[0025] 作为本申请公开主题的应用示例,图2示出一种在流媒体环境下视频编码器和解码器的部署方式。本申请公开的主题可与其它支持视频的应用共同使用,包括,例如视频会议、数字电视、将压缩后的视频存储到数字媒体,包括CD、DVD、记忆棒等,的应用等。

[0026] 如图2所示,流媒体系统200可包括拍摄子系统213,其包括视频源201和编码器203。流媒体系统200还可以包括至少一个流媒体服务器205和/或至少一个流媒体客户端206。

[0027] 视频源201可以创建,例如,包括3D视频对应的3D点云的流202。视频源201可以包括,例如3D传感器(如深度传感器)或3D成像技术(如数码摄像机),以及用于利用从3D传感器或3D成像技术接收到的数据生成3D点云的计算设备。样本流202(比已编码的视频码流的数据量大)可由耦合于视频源201的编码器203处理。编码器203可包括硬件、软件或软硬件组合,从而能够支持或实现如下文详细描述的本申请公开主题的各方面。编码器203还可生成已编码视频码流204。已编码视频码流204(比未压缩视频样本流202的数据量更小)可存储于流媒体服务器205以备后续使用。一个或多个流媒体客户端206可访问流媒体服务器205,以获取视频码流209,视频码流209可以是已编码视频码流204的副本。

[0028] 流媒体客户端206可包括视频解码器210和显示器212。视频解码器210可以,例如,解码视频码流209(即,接收到的编码视频码流204的副本),并创建输出视频样本流211,其可绘制在显示器212或另一绘制设备(未示出)上。在一些流媒体系统中,视频码流204、209可根据一些视频编码/压缩标准进行编码。该标准的示例包括,但不限于,ITU-T建议H.265、多用途视频编码(Versatile Video Coding, VVC),及MPEG/V-PCC。

[0029] 参考图3至图4,下面描述可以由本申请的实施例执行的V-PCC的一些方面。

[0030] 图3为本申请实施例的视频编码器203的示例功能框图。

[0031] 如图3所示,视频编码器203可以接收一个或多个点云帧350,并基于点云帧350生成几何图像352、纹理图像356和占用图334。视频编码器203可以将几何图像352压缩成压缩后的几何图像362,将纹理图像356压缩成压缩后的纹理图像364,以及将占用图334压缩成压缩后的占用图372。视频编码器203的多路复用器328可以形成压缩后的码流374,压缩的码流374包括压缩后的几何图像362、压缩后的纹理图像364和压缩后的占用图372。

[0032] 更具体地,一些实施例中,视频编码器203可以包括将点云帧350分割成图像块(patch)的图像块生成模块302。图像块是V-PCC的有用实体。图像块生成过程包括将点云帧350分解成具有平滑边界的最小数目的图像块,同时还使重建误差最小化。本申请的编码器可实施各种方法来实现此分解。

[0033] 视频编码器203可以包括执行打包过程的图像块打包模块304。打包过程包括将提取的图像块映射到2D网格上,同时使未使用的空间最小化并保证网格的每个 $M \times M$ (例如,

16x16)块与唯一的图像块相关联。高效的图像块打包通过使未使用的空间最小化或确保时间一致性来直接影响压缩效率。图像块打包模块304可以生成占用图334。

[0034] 视频编码器203可以包括几何图像生成模块306和纹理图像生成模块308。为了更好地处理多个点投影到相同样本的情况,每个图像块可以投影到两个图像(也称为层)上。例如,几何图像生成模块306和纹理图像生成模块308可以利用图像块打包模块304在打包过程中计算出的3D到2D映射来将点云的几何形状和纹理存储为图像(也称为层)。可以根据作为参数提供的配置信息,将生成的图像/层存储为一个或多个视频帧,并使用视频编解码器(例如HM视频编解码器)对其进行压缩。

[0035] 一些实施例中,基于输入点云帧350和占用图334,几何图像生成模块306生成几何图像352,纹理图像生成模块308生成纹理图像356。图5为几何图像352的示例,图6为纹理图像356的示例。一个实施例中,几何图像352可以由YUV420-8位格式的WxH的单色帧表示。一个实施例中,占用图334图像由二进制图组成,该二进制图指示网格的每个单元属于空白空间还是属于点云。为了生成纹理图像356,纹理图像生成模块308可以利用重建后的/平滑后的几何形状358,以便计算要与重新采样的点相关联的颜色。

[0036] 视频编码器203还可以包括分别用于填充几何图像352的图像填充模块314和用于填充纹理图像356的图像填充模块316,以形成填充后的几何图像354和填充后的纹理图像360。图像填充(也称为“背景填充”)仅使用冗余信息填充图像的未使用空间。良好的背景填充是一种最低限度地增加比特率而不在图像块边界周围引入显著编码失真的操作。图像填充模块314和图像填充模块316可以使用占用图334以分别形成填充后的几何图像354和填充后的纹理图像360。一个实施例中,视频编码器203可以包括群组扩张模块320以形成填充后的纹理图像360。

[0037] 视频编码器203可以包括视频压缩模块322和视频压缩模块324,用于分别将填充后的几何图像354压缩为压缩后的几何图像362,以及将填充后的纹理图像360压缩为压缩后的纹理图像364。

[0038] 视频编码器203可以包括:熵压缩模块318,用于对占用图334进行无损编码366,以及视频压缩模块326,用于对占用图334进行有损编码368。

[0039] 各实施例中,视频编码器203可以包括平滑模块310,用于通过使用图像块信息332和由视频压缩模块322提供的重建后的几何图像365来生成平滑后的几何形状358。平滑模块310的平滑过程可以旨在减轻由于压缩伪像而在图像块边界处出现的可能的不连续性。平滑后的几何形状358可用于纹理图像生成模块308以生成纹理图像356。

[0040] 视频编码器203还可以包括辅助图像块信息压缩模块312,用于形成由多路复用器328在压缩后的码流374中提供的压缩后的辅助图像块信息370。

[0041] 图4为本申请实施例的视频解码器210的示例功能框图。

[0042] 如图4所示,视频解码器210可以从视频编码器203接收已编码的码流374,以获取压缩后的纹理图像362、压缩后的几何图像364、压缩后的占用图372和压缩后的辅助图像块信息370。视频解码器210可以对压缩后的纹理图像362、压缩后的几何图像364、压缩后的占用图372和压缩后的辅助图像块信息370进行解码,以分别获取解压缩的纹理图像460、解压缩的几何图像462、解压缩的占用图464和解压缩的辅助图像块信息466。接下来,视频解码器210可以基于解压缩的纹理图像460、解压缩的几何图像462、解压缩的占用图464和解压

缩的辅助图像块信息466生成重建的点云474。

[0043] 一个实施例中,视频解码器210可以包括解复用器402,用于从所接收的压缩后的码流374中分离出压缩后的纹理图像362、压缩后的几何图像364、压缩后的占用图372和压缩后的辅助图像块信息370。

[0044] 视频解码器210可以包括视频解压缩模块404、视频解压缩模块406、占用图解压缩模块408和辅助图像块信息解压缩模块410,分别用于对压缩后的纹理图像362、压缩后的几何图像364、压缩后的占用图372和压缩后的辅助图像块信息370进行解码。

[0045] 视频解码器210可以包括几何形状重建模块412,用于基于解压缩的几何图像462、解压缩的占用图464和解压缩的辅助图像块信息466获取重建的(三维)几何形状468。

[0046] 视频解码器210可以包括平滑模块414,用于对重建后的几何形状468进行平滑以获取平滑后的几何形状470。平滑过程可以旨在减轻由于压缩伪像而可能在图像块边界处出现的可能的不连续性。

[0047] 视频解码器210可以包括纹理重建模块416,用于基于解压缩的纹理图像460和平滑后的几何形状470获取重建后的纹理472。

[0048] 视频解码器210可以包括颜色平滑模块418,用于对重建后的纹理472的颜色进行平滑以获取重建的点云474。3D空间中的非相邻图像块在打包后通常在2D视频中彼此相邻。这意味着来自非相邻图像块的像素值可能被基于块的视频编解码器(block-based video codec)混淆。颜色平滑模块418的颜色平滑可以旨在减少图像块边界处出现的可见伪像。

[0049] 在对几何视频进行压缩后,可以使用压缩后的几何视频和相应的占用图来重建点云,也可以称为几何重建云。然后将颜色转移到得到的几何重建云中。在颜色转移之前,可以对几何重建云进行几何平滑。

[0050] 几何平滑的目的可以是恢复由于几何压缩以及高分辨率占用图到较低分辨率占用图的转换而使几何形状在补丁边界处发生的失真。更好的平滑策略可产生更好的几何重建云,进而产生更高质量的压缩后的纹理视频。

[0051] 本申请实施例可以涉及在进行平滑前在重建云上创建网格,以及在每个网格单元中执行局部平滑。在下文中,我们用 x_1, x_2, \dots, x_N 表示 $G \times G$ (G 在本申请中可由grid_cell_size表示)网格单元的点。这些点的平均值可由 \bar{p} 表示,如以下等式1所示:

$$[0052] \quad \bar{p} = \frac{\sum_{i=1}^N x_i}{N} \quad (\text{等式 1})$$

[0053] 其中, x_i 表示多个点中索引为 i 的点, N 表示该多个点的数目。

[0054] 可以按照以下等式2计算协方差矩阵:

$$[0055] \quad C = \frac{\sum_{i=1}^N (x_i - \bar{p})(x_i - \bar{p})^T}{N} \quad (\text{等式 2})$$

[0056] 其中, T 表示转置运算。 C 的特征值可由以下等式3表示:

$$[0057] \quad \lambda_3 < \lambda_2 < \lambda_1 \quad (\text{等式 3})$$

[0058] 它们对应的单元特征向量可以由 v_3, v_2, v_1 表示。

[0059] 一个实施例中,用于局部平滑的算法可以包括将平面拟合到网格单元的多个点。最能代表一个单元的多个点的平面可以是(1)穿过多个点 \bar{p} 的质心,以及(2)法向向量方向与 v_3 方向相同的平面。

[0060] 该平面可以是特征向量 v_1 和 v_2 跨越的子空间,并且被移位到该点 \bar{p} 。根据一个实施例,平滑方法可以将所有网格单元点 x_1, x_2, \dots, x_N 投影到该平面上。这样,可以抑制噪声,并且输出的可以是网格单元点的平滑后的版本。

[0061] 一个实施例中,可以使用以下等式4将给定点 $X \in \{x_1, x_2, \dots, x_N\}$ 投影到该平面上:

[0062]
$$X_{proj} = [(X - \bar{p}) \cdot v_1]v_1 + [(X - \bar{p}) \cdot v_2]v_2 + \bar{p} \quad (\text{等式 4})$$

[0063] 一个实施例中,点 X 可以移动到其加权后的投影,或者移动到连接该点与其加权投影的射线上的点,如以下等式5所示:

[0064]
$$X \rightarrow \alpha X_{proj} + (1 - \alpha) X \quad (\text{等式 5})$$

[0065] α 可以是投影权重(可以表示为projection_weight)。我们得出的该平面可以是表示当前网格单元的点的最简单的局部(平面)流形。一些实施例中,可以得出更复杂的非平面局部流形,以更精确地表示这些点。应注意,拟合的流形不能太灵活,否则它可能会对某些点过拟合,并降低平滑质量。

[0066] 用于用信号表示上述实施例的元数据语法的示例如下表1所示:

[0067] 表1

	local_projection_smoothing_metadata() {	描述符
	if(local_projection_smoothing_enabled_flag) {	
	local_projection_smoothing_present_flag	u(1)
	if(local_projection_smoothing_present_flag) {	
[0068]	grid_cell_size	u(8)
	projection_weight	f(32)
	}	
	}	
	}	

[0069] 在上表1中,local_projection_smoothing_present_flag可以指示是否使用局部投影平滑。

[0070] grid_cell_size可以指示网格单元的大小。grid_cell_size的值可以在[2,1024]的范围内。

[0071] projection_weight可以指示当前点移动到其在投影平面上的投影点的尺度。projection_weight的值可以在[0,1]的范围内。

[0072] 一个示例的解码过程可以将grid_cell_size和projection_weight作为输入。解码器,例如解码器210,可以首先在重建的未平滑的云上创建网格。然后,解码器可以逐个处理每个网格单元。对于每个单元,解码器可以拟合出最能代表该单元的各点的一个平面。这样,解码器可以计算该单元的各点的协方差矩阵的特征向量。然后,解码器可以将该单元的所有点投影到拟合出的平面上。平滑强度可以由projection_weight来控制。较大的projection_weight值可产生更平滑的结果。

[0073] 图7是一些实施例的对视频流进行解码的方法700的流程图,该视频流使用视频点云编解码进行编码,该视频点云编解码可包括通过局部几何投影进行几何平滑。在一些实施方式中,图7的至少一个过程框可以由解码器210执行。在一些实施方式中,图7的至少一个过程框可以由与编码器203分离或包括编码器203的另一设备或一组设备(例如编码器203)执行。

[0074] 如图7所示,在操作710中,方法700可以包括获取几何重建点云。

[0075] 在操作720中,方法700可以包括将几何重建点云划分为多个单元,其中多个单元中的一个单元包括第一组多个点。

[0076] 在操作730中,方法700可以包括确定表示第一组多个点的平面。

[0077] 在操作740中,方法700可以包括通过将第一组多个点投影到该平面上来生成第二组多个点。

[0078] 在操作750中,方法700可以包括基于第二组多个点获取平滑后的几何重建点云。

[0079] 在操作760中,方法700可以包括使用平滑后的几何重建点云重建动态点云。

[0080] 虽然图7示出了方法700的示例框,在一些实施方式中,相比图7中描绘的那些框,方法700可以包括更多的框、更少的框、不同的框或不同排列的框。附加地或可选地,方法700的两个或更多个框可以并行执行。

[0081] 图8是一些实施例的对视频流进行解码的装置800的示意图,该视频流使用视频点云编解码进行编码,该视频点云编解码可包括通过局部几何投影进行几何平滑。如图8所示,装置800包括第一获取代码810、划分代码820、确定代码830、生成代码840、第二获取代码850和重建代码860。

[0082] 第一获取代码810可用于,使至少一个处理器获取几何重建点云。

[0083] 划分代码820可用于,使至少一个处理器将几何重建点云划分为多个单元,其中多个单元中的一个单元包括第一组多个点。

[0084] 确定代码830可用于,使至少一个处理器确定表示第一组多个点的平面。

[0085] 生成代码840可用于,使至少一个处理器通过将第一组多个点投影到该平面上来生成第二组多个点。

[0086] 第二获取代码850可用于,使至少一个处理器基于第二组多个点获取平滑后的几何重建点云。

[0087] 重建代码860可用于使至少一个处理器使用平滑后的几何重建点云来重建动态点云。

[0088] 上述技术可以使用计算机可读指令实现为计算机软件,并且物理地存储在一个或多个计算机可读介质中。例如,图9示出适于实现本申请一些实施例的计算机系统900。

[0089] 该计算机软件可利用任何合适的机器代码或计算机语言来编码,可采用汇编、编译、链接或类似机制生成指令代码。这些指令代码可由计算机中央处理单元(central processing unit,CPU)、图形处理单元(Graphics Processing Unit,GPU)等直接执行或通过代码解释、微代码执行等操作来执行。

[0090] 这些指令可在多种类型的计算机或计算机组件中执行,包括,例如,个人计算机、平板电脑、服务器、智能电话、游戏设备、物联网设备等。

[0091] 图9所示的计算机系统900的组件仅仅是例子,而非旨在对实现本申请实施例的计

计算机软件的使用或功能范围做任何限制。也不应将组件的配置方式解释为对计算机系统900的非限制性实施例中的任一部件或其组合具有任何的依赖性要求。

[0092] 计算机系统900可以包括某些人机界面输入设备。这样的人机界面输入设备可以响应于一个或多个人类用户通过例如触觉输入(诸如键击、挥动、数据手套移动)、音频输入(诸如语音、拍击)、视觉输入(诸如姿势)、嗅觉输入(未示出)的输入。人机界面设备还可用于捕捉不必直接与人类有意识输入相关的某些介质,例如音频(诸如语音、音乐、环境声音)、图像(诸如扫描的图像、从静止图像相机获得的摄影图像)、视频(诸如二维视频,包括立体视频的三维视频)。

[0093] 人机界面输入设备可包括以下项中的一种或多种(每一种仅描绘一个):键盘901、鼠标902、触控板903、触摸屏910、数据手套、操纵杆905、麦克风906、扫描仪907、照相机908。

[0094] 计算机系统900还可以包括某些人机界面输出设备。这样的人机界面输出设备可以通过例如触觉输出、声音、光和气味/味道来刺激一个或多个人类用户的感受。这种人机界面输出设备可以包括触觉输出设备(例如通过触摸屏910、数据手套或操纵杆905的触觉反馈,但是也可以有不用作输入设备的触觉反馈设备)。例如,这种设备可以是音频输出设备(诸如扬声器909、耳机(未示出))、可视输出设备以及打印机(未示出),其中可视输出设备诸如屏幕910、虚拟现实眼镜(未示出)、全息显示器和烟雾箱(未示出),屏幕910包括阴极射线管(CRT)屏幕、液晶显示器(LCD)屏幕、等离子屏幕、有机发光二极管(OLED)屏幕,每一种都具有或不具有触摸屏输入能力,每一种都具有或不具有触觉反馈能力,这些屏幕中的一些能够通过手段(诸如立体图像输出)输出二维可视输出或多于三维的输出。

[0095] 计算机系统900还可以包括人类可访问的存储设备及其相关联的介质,诸如光学介质(包括具有CD/DVD的CD/DVDROM/RW1220)或类似介质921、拇指驱动器922、可移动硬盘驱动器或固态驱动器923、传统磁介质(诸如磁带和软盘(未示出))、基于专用ROM/ASIC/PLD的设备(诸如安全加密狗(security dongle)(未示出)),等等。

[0096] 本领域技术人员还应当理解,结合当前公开的主题使用的术语“计算机可读介质”不包括传输介质、载波或其它瞬时信号。

[0097] 计算机系统900还可以包括连接一个或多个通信网络的接口。网络可以是,例如,无线网络、有线网络、光网络。网络还可以是本地网、广域网、城域网、车联网的和工业网络、实时网络、延迟容忍网络等等。网络的示例包括局域网(诸如以太网、无线LAN)、蜂窝网络(包括全球移动通信系统(GSM)、第三代移动通信系统(3G)、第四代移动通信系统(4G)、第五代移动通信系统(5G)、长期演进(LTE)等)、电视有线或无线广域数字网络(包括有线电视、卫星电视和地面广播电视)、车辆和工业网络(包括CANBus),等等。某些网络通常需要外部网络接口适配器,该外部网络接口适配器连接到某些通用数据端口或外围总线949(诸如计算机系统900的通用串行总线(USB)端口);其他的通常通过如下所述连接到系统总线而集成到计算机系统900的核心中(例如,进入个人计算机系统的以太网接口或进入智能手机计算机系统的蜂窝网络接口)。通过使用这些网络中的任何一个,计算机系统900可以与其它实体通信。这种通信可以是使用局域或广域数字网络的到其它计算机系统的单向的、仅接收的(例如广播TV)、单向仅发送的(例如到某些CAN总线设备的CAN总线)或双向的通信。这种通信可以包括与云计算环境955的通信。可以在如上所述的那些网络和网络接口中的每一个上使用某些协议和协议栈。

[0098] 上述人机界面设备、人类可访问存储设备和网络接口954可以连接到计算机系统900的内核940。

[0099] 内核940可以包括一个或多个中央处理单元(CPU)941、图形处理单元(GPU)942、以现场可编程门阵列(FPGA)943形式存在的专用可编程处理单元、用于特定任务的硬件加速器944等。这些设备,以及只读存储器(ROM)945,随机存取存储器946,内部大容量存储器(如内部非用户可访问硬盘驱动器,SSD)947等,可以通过系统总线948相互连接。在一些计算机系统中,系统总线948可以以一个或多个物理插头的形式访问,从而通过附加的CPU,GPU等实现扩展。外围设备可以直接,或者通过外围总线948,连接到内核的系统总线949。外围总线的架构包括PCI,USB等。图形适配器950可能包括在内核940中。

[0100] CPU 941、GPU 942、FPGA 943和加速器944可以执行某些指令,这些指令组合起来可以构成前述的计算机代码。该计算机代码可以存储在ROM 945或RAM 946中。中间数据也可以存储在RAM 946中,而永久数据可以存储在,例如,内部大容量存储器947中。可以通过使用高速缓冲存储器来实现到任何存储器设备的快速存储和读取,高速缓存存储器可以与一个或多个CPU 941、GPU 942、大容量存储器947、ROM 945、RAM 946等紧密关联。

[0101] 计算机可读介质上可以具有计算机代码,在计算机代码上执行各种计算机执行的操作。介质和计算机代码可以是为本申请的目的而特别设计和构造的,也可以是计算机软件领域的技术人员所熟知和可用的介质和代码。

[0102] 作为示例而非限制,具有体系结构900的计算机系统,特别是内核940,可以提供处理器(包括CPU、GPU、FPGA、加速器等)执行在一个或多个有形的计算机可读介质中的软件而实现的功能。这样的计算机可读介质可以是与如上所述的用户可访问大容量存储器相关联的介质,以及非暂时性的内核940的某些存储,诸如内核内部大容量存储器947或ROM945。实现本申请一些实施例的软件可以存储在这样的设备中并由内核940执行。根据特定需要,计算机可读介质可包括一个或多个存储器设备或芯片。该软件可以使内核940,特别是其中的处理器(包括CPU,GPU,FPGA等),执行本文描述的特定过程或特定过程的特定部分,包括定义存储在RAM946中的数据结构,以及根据软件定义的过程修改这些数据结构。作为补充或作为替代,计算机系统可提供与电路(例如加速器944中的逻辑硬连线或其它组件相同的功能,可代替软件或与软件一起操作以执行本文所述的特定过程或特定过程的特定部分。在适当的情况下,对软件的引用可以包括逻辑,反之亦然。在适当的情况下,对计算机可读介质的引用可包括存储执行软件的电路(如集成电路(IC)),包括执行逻辑的电路,或两者兼备。本申请包括硬件和软件的任何适当组合。

[0103] 虽然本申请已对多个非限制性实施例进行了描述,但实施例的各种变更、置换和各种替代属于本申请的范围之内。因此应理解,本领域技术人员能够设计多种系统和方法,所述系统和方法虽然未在本文中明确展示或描述,但其体现了本申请的原则,因此属于本申请的精神和范围之内。

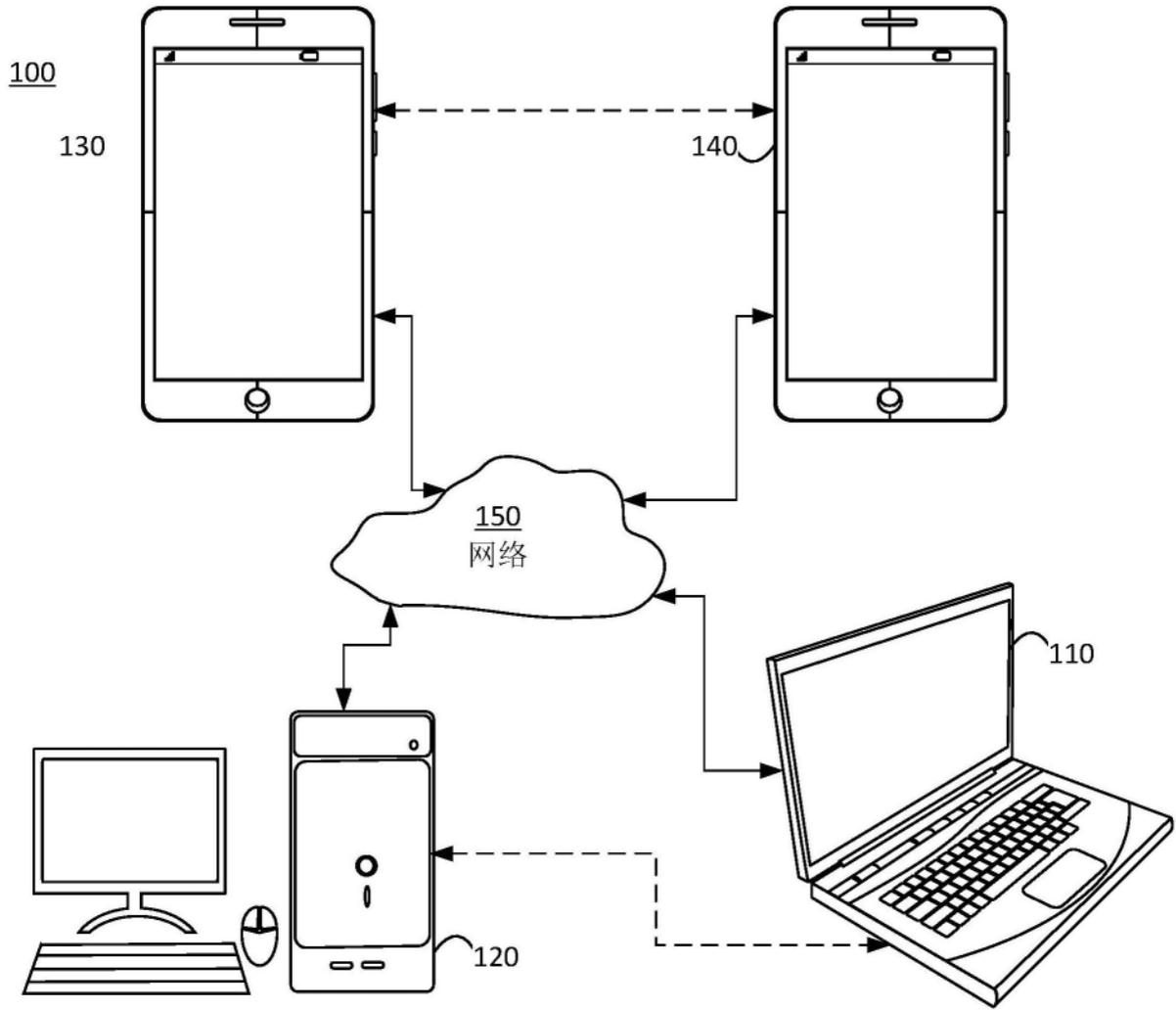


图1

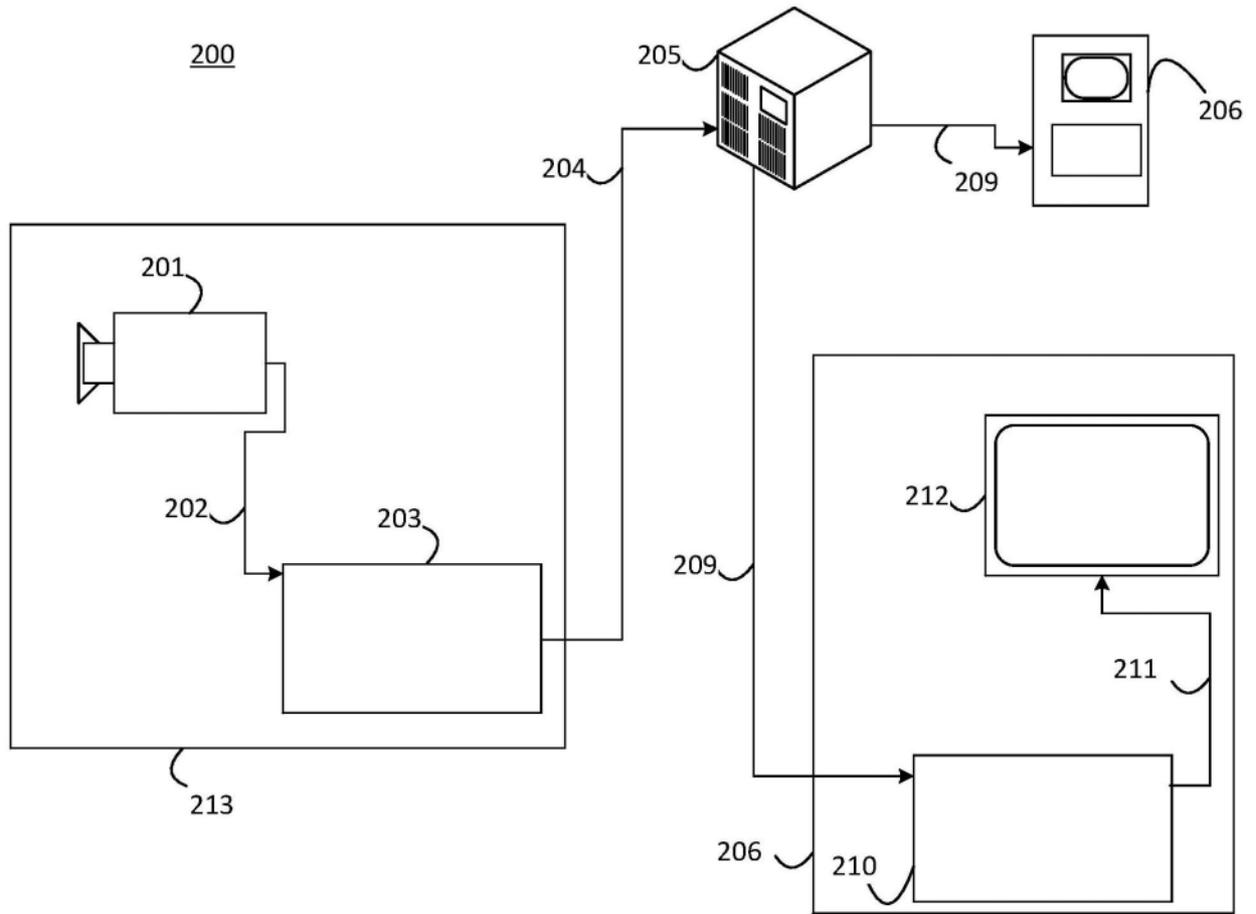


图2

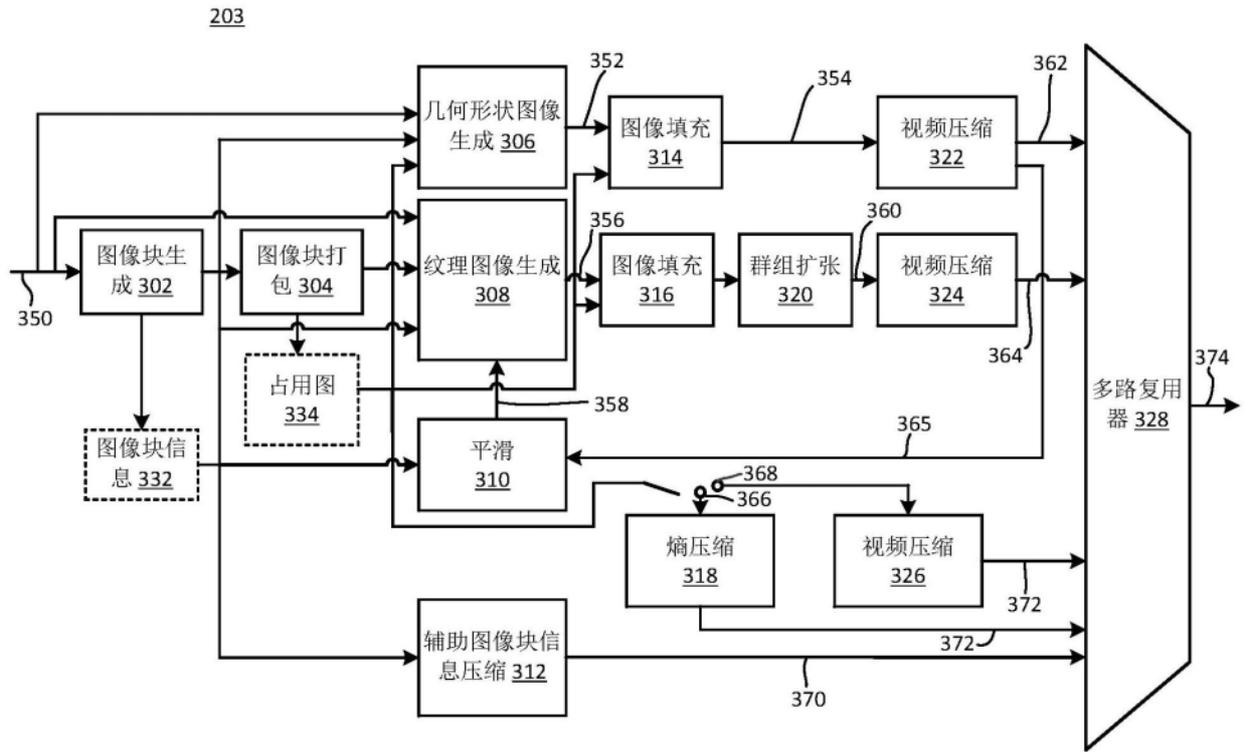


图3

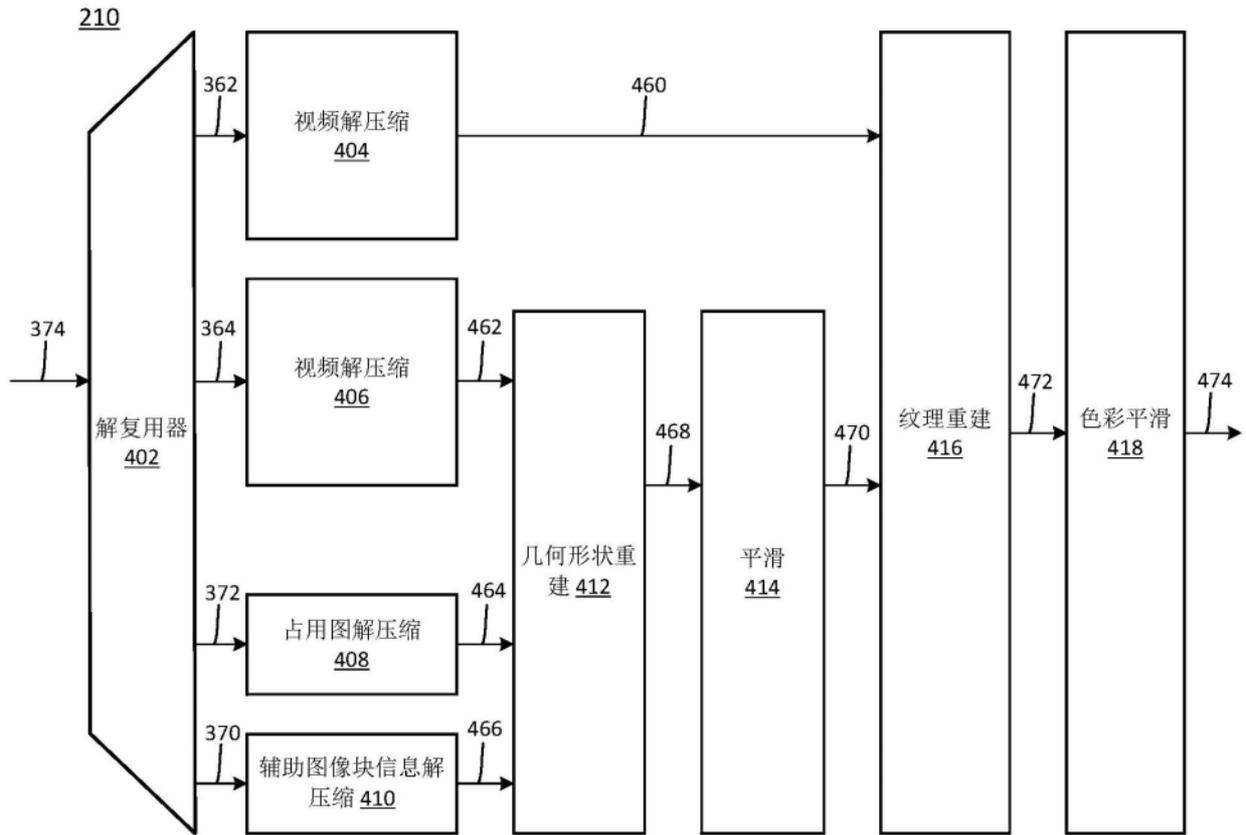


图4

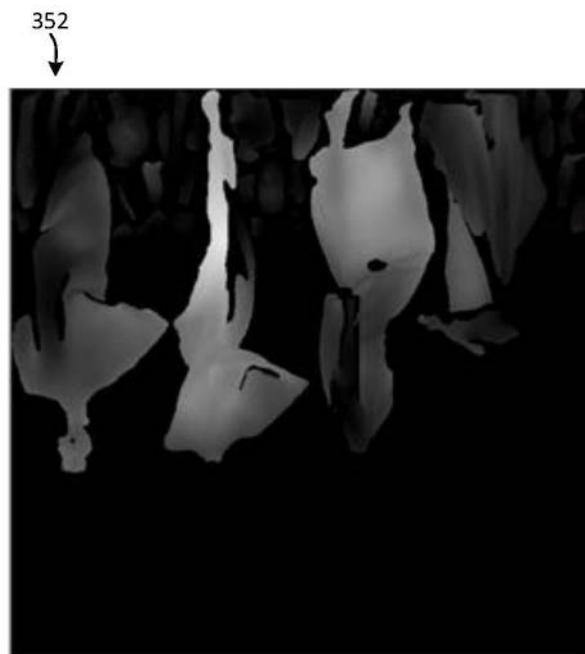


图5

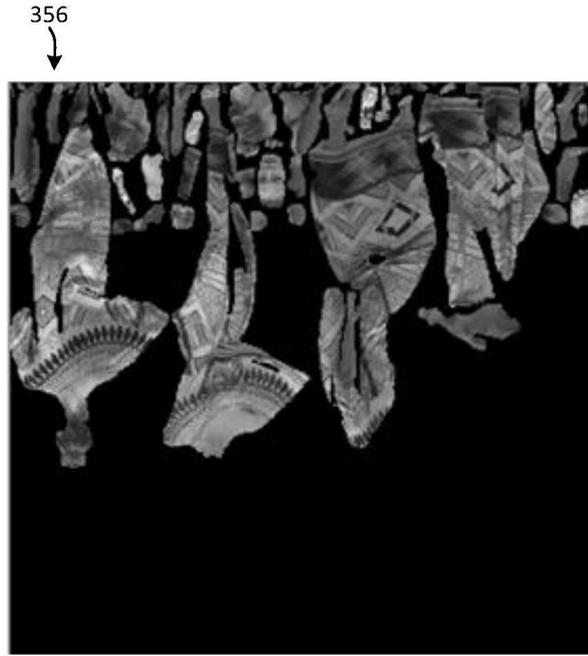


图6

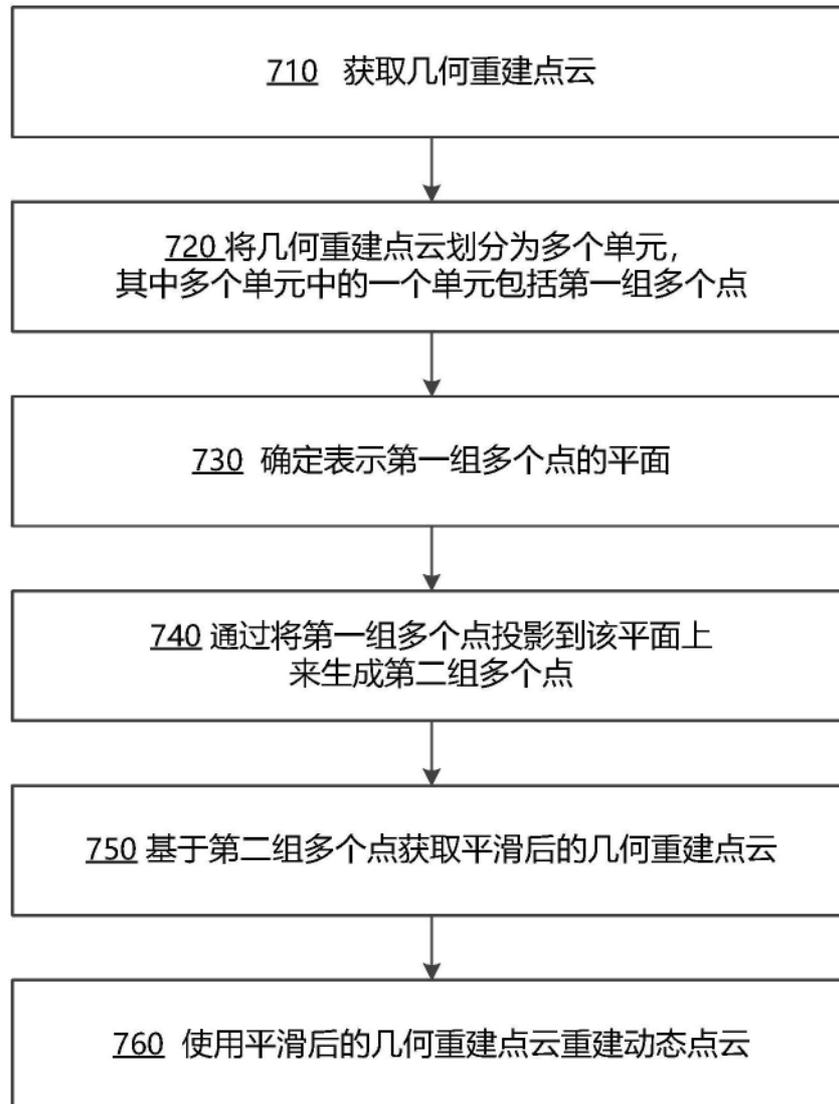


图7



图8

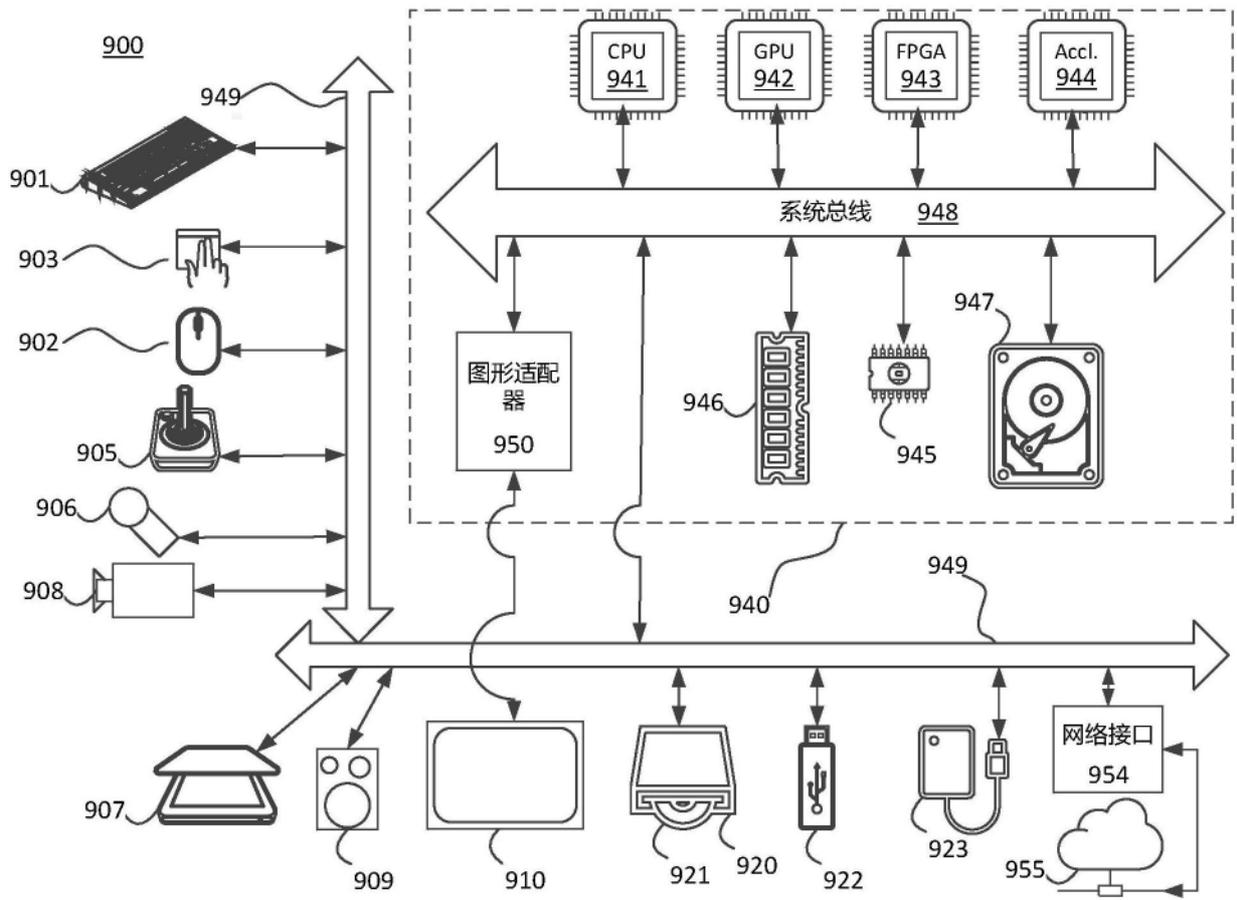


图9