



(12) 发明专利

(10) 授权公告号 CN 111937041 B

(45) 授权公告日 2025. 04. 08

(21) 申请号 201980024478.9	(73) 专利权人 谷歌有限责任公司
(22) 申请日 2019.09.17	地址 美国加利福尼亚州
(65) 同一申请的已公布的文献号 申请公布号 CN 111937041 A	(72) 发明人 迈克尔·和梅尔 阿米什·马卡迪亚
(43) 申请公布日 2020.11.13	(74) 专利代理机构 中原信达知识产权代理有限 责任公司 11219
(30) 优先权数据 16/143,165 2018.09.26 US	专利代理师 周亚荣 邓聪惠
(85) PCT国际申请进入国家阶段日 2020.09.30	(51) Int.Cl. G06T 9/00 (2006.01)
(86) PCT国际申请的申请数据 PCT/US2019/051566 2019.09.17	(56) 对比文件 W0 2015177162 A1, 2015.11.26
(87) PCT国际申请的公布数据 W02020/068492 EN 2020.04.02	审查员 陈晓海

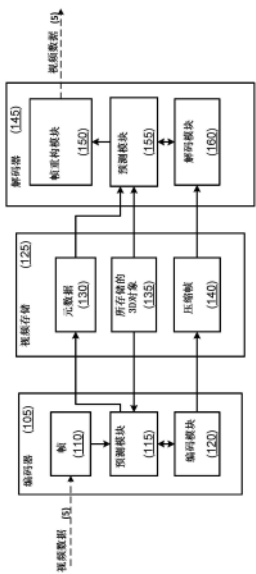
权利要求书5页 说明书37页 附图17页

(54) 发明名称

通过提供几何代理进行视频编码

(57) 摘要

压缩视频帧包括接收视频帧、在该帧中标识三维 (3D) 对象、将3D对象与所存储的3D对象进行匹配、基于3D对象和所存储的3D对象使用颜色预测方案压缩视频帧、以及存储具有元数据的压缩帧,该元数据标识3D对象,指示3D对象在视频帧中的位置,并指示3D对象在视频帧中的定向。



1. 一种用于视频编码的方法,包括:
接收视频的帧;
在所述帧中标识三维3D对象;
将所述3D对象与所存储的3D对象进行匹配;
基于所存储的3D对象来生成第一3D对象代理;
基于在所述帧中标识的所述3D对象来变换所述第一3D对象代理;
基于所变换的第一3D对象代理和基于所述3D对象生成的经变换的第二3D对象代理来使用颜色预测方案压缩所述视频的帧;以及

存储具有元数据的压缩帧,所述元数据标识所述3D对象,指示所述3D对象在所述视频的帧中的位置并指示所述3D对象在所述视频的帧中的定向。

2. 根据权利要求1所述的方法,其中基于所变换的第一3D对象代理和所述经变换的第二3D对象代理来使用所述颜色预测方案压缩所述视频的帧包括:

基于所存储的3D对象生成第二3D对象代理;
在所述视频的关键帧中标识所述3D对象;
基于在所述关键帧中标识的所述3D对象变换所述第二3D对象代理;
将颜色属性从所述3D对象映射到经变换的第一3D对象代理;
将颜色属性从在所述关键帧中标识的所述3D对象映射到经变换的第二3D对象代理;以及

基于所述经变换的第一3D对象代理的颜色属性以及所述经变换的第二3D对象代理的颜色属性生成所述3D对象的残差。

3. 根据权利要求1所述的方法,其中,基于所变换的第一3D对象代理和所述经变换的第二3D对象代理来使用所述颜色预测方案压缩所述视频的帧包括:

基于所存储的3D对象生成第二3D对象代理;
在所述视频的关键帧中标识所述3D对象;
基于在所述关键帧中标识的所述3D对象变换所述第二3D对象代理;
将颜色属性从所述3D对象映射到经变换的第一3D对象代理;以及
基于所述经变换的第一3D对象代理的颜色属性和所述经变换的第二3D对象代理的默认颜色属性生成所述3D对象的残差。

4. 根据权利要求1所述的方法,其中基于所变换的第一3D对象代理和所述经变换的第二3D对象代理来使用所述颜色预测方案压缩所述视频的帧包括:

使用自动编码器对所述第一3D对象代理进行编码,其中,变换所述第一3D对象代理包括:基于在所述帧中标识的所述3D对象变换所编码的第一3D对象代理;
基于所存储的3D对象生成第二3D对象代理;
使用自动编码器对所述第二3D对象代理进行编码;
在所述视频的关键帧中标识所述3D对象;
基于在所述关键帧中标识的所述3D对象变换所编码的第二3D对象代理;
将颜色属性从所述3D对象映射到经变换的第一3D对象代理;
将颜色属性从在所述关键帧中标识的所述3D对象映射到经变换的第二3D对象代理;以及

基于所述经变换的第一3D对象代理的颜色属性以及所述经变换的第二3D对象代理的颜色属性生成所述3D对象的残差。

5. 根据权利要求1所述的方法, 其中基于所变换的第一3D对象代理和所述经变换的第二3D对象代理来使用所述颜色预测方案压缩所述视频的帧包括:

使用自动编码器对所述第一3D对象代理进行编码, 其中, 变换所述第一3D对象代理包括: 基于在所述帧中标识的所述3D对象变换所编码的第一3D对象代理;

基于所存储的3D对象生成第二3D对象代理;

使用自动编码器对所述第二3D对象代理进行编码;

在所述视频的关键帧中标识所述3D对象;

基于在所述关键帧中标识的所述3D对象变换所编码的第二3D对象代理;

将颜色属性从所述3D对象映射到经变换的第一3D对象代理; 以及

基于所述经变换的第一3D对象代理的颜色属性和所述经变换的第二3D对象代理的默认颜色属性生成所述3D对象的残差。

6. 根据权利要求1所述的方法, 进一步包括:

在存储所述3D对象之前:

标识与所述视频相关联的至少一个感兴趣的3D对象;

确定与所述感兴趣的3D对象相关联的多个网格属性;

确定与所述感兴趣的3D对象相关联的位置;

确定与所述感兴趣的3D对象相关联的定向;

确定与所述感兴趣的3D对象相关联的多个颜色属性; 以及

使用自动编码器减少与所述感兴趣的3D对象的网格属性相关联的变量的数量。

7. 根据权利要求1所述的方法, 其中压缩所述视频的帧包括确定所述3D对象相对于关键帧中的背景3D对象的原点坐标的位置坐标。

8. 根据权利要求1所述的方法, 其中:

所存储的3D对象包括默认颜色属性, 并且

所述颜色预测方案使用所述默认颜色属性。

9. 根据权利要求1所述的方法, 进一步包括:

标识与所述视频相关联的至少一个感兴趣的3D对象;

基于所述至少一个感兴趣的3D对象, 生成至少一个所存储的3D对象, 所述至少一个所存储的3D对象中的每一个由包括通过面连接的点的集合的网格定义, 每个点存储至少一个属性, 所述至少一个属性包括相应点的位置坐标; 以及

与所述视频相关联地存储所述至少一个所存储的3D对象。

10. 一种用于视频解码的方法, 包括:

接收视频的帧;

在所述帧中标识三维3D对象;

将所述3D对象与所存储的3D对象进行匹配;

基于所存储的3D对象来生成第一3D对象代理;

基于在所述帧中标识的所述3D对象来变换所述第一3D对象代理;

基于所变换的第一3D对象代理和基于所存储的3D对象生成的经变换的第二3D对象代

理来使用颜色预测方案解压缩所述视频的帧；以及
渲染所述视频的帧。

11. 根据权利要求10所述的方法，其中基于所变换的第一3D对象代理和所述经变换的第二3D对象代理来使用所述颜色预测方案解压缩所述视频的帧包括：

基于所存储的3D对象生成第二3D对象代理；
在所述视频的关键帧中标识所述3D对象；
基于在所述关键帧中标识的所述3D对象变换所述第二3D对象代理；
将颜色属性从所述3D对象映射到经变换的第一3D对象代理；
将颜色属性从在所述关键帧中标识的所述3D对象映射到经变换的第二3D对象代理；以及

基于所述经变换的第一3D对象代理的颜色属性和所述经变换的第二3D对象代理的颜色属性生成所述3D对象的颜色属性。

12. 根据权利要求10所述的方法，其中，基于所变换的第一3D对象代理和所述经变换的第二3D对象代理来使用所述颜色预测方案解压缩所述视频的帧包括：

基于所存储的3D对象生成第二3D对象代理；
在所述视频的关键帧中标识所述3D对象；
基于在所述关键帧中标识的所述3D对象变换所述第二3D对象代理；
将颜色属性从所述3D对象映射到经变换的第一3D对象代理；以及
基于所述经变换的第一3D对象代理的颜色属性和所述经变换的第二3D对象代理的默认颜色属性生成所述3D对象的颜色属性。

13. 根据权利要求10所述的方法，其中基于所变换的第一3D对象代理和所述经变换的第二3D对象代理来使用所述颜色预测方案解压缩所述视频的帧包括：

使用自动编码器对所述第一3D对象代理进行解码；
基于与所述3D对象相关联的元数据变换已解码的第一3D对象代理；
基于所存储的3D对象生成第二3D对象代理；
使用自动编码器对所述第二3D对象代理进行解码；
在所述视频的关键帧中标识所述3D对象；
基于与在所述关键帧中标识的所述3D对象相关联的元数据变换已解码的第二3D对象代理；
将颜色属性从所述3D对象映射到所述经变换的第一3D对象代理；
将颜色属性从在所述关键帧中标识的所述3D对象映射到所述经变换的第二3D对象代理；以及

基于所述经变换的第一3D对象代理的颜色属性和所述经变换的第二3D对象代理的颜色属性生成所述3D对象的颜色属性。

14. 根据权利要求10所述的方法，其中基于所变换的第一3D对象代理和所述经变换的第二3D对象代理来使用所述颜色预测方案解压缩所述视频的帧包括：

使用自动编码器对所述第一3D对象代理进行解码；
基于与所述3D对象相关联的元数据对已解码的第一3D对象代理进行变换；
基于所存储的3D对象生成第二3D对象代理；

使用自动编码器对所述第二3D对象代理进行解码；
在所述视频的关键帧中标识所述3D对象；
基于与在所述关键帧中标识的所述3D对象相关联的元数据变换已解码的第二3D对象代理；
将颜色属性从所述3D对象映射到经变换的第一3D对象代理；以及
基于所述经变换的第一3D对象代理的颜色属性和所述经变换的第二3D对象代理的默认属性生成所述3D对象的颜色属性。

15. 根据权利要求10所述的方法,进一步包括:

接收3D形状的至少一个潜在表示；
使用机器训练的生成建模技术来：
确定与所述3D形状相关联的多个网格属性；
确定与所述3D形状相关联的位置；
确定与所述3D形状相关联的定向；以及
确定与所述3D形状相关联的多个颜色属性,以及
将所述3D形状存储为所存储的3D对象。

16. 根据权利要求10所述的方法,其中,渲染所述视频的帧包括:

接收所述3D对象相对于关键帧中的背景3D对象的原点坐标的位置坐标;以及
使用所述位置坐标将所述3D对象放置在所述帧中。

17. 根据权利要求10所述的方法,进一步包括:

接收由自动编码器的编码器使用的神经网络以减少与至少一个感兴趣的3D对象的网格属性、位置、定向和颜色属性相关联的变量的数量;

在所述自动编码器的解码器中使用所述神经网络再生与至少一个感兴趣的3D对象的网格相关联的点,再生所述点包括再生位置属性、定向属性和颜色属性;以及
将所述至少一个感兴趣的3D对象存储为所存储的3D对象。

18. 一种用于使用代理预测颜色变化的方法:

基于所存储的3D对象生成第一3D对象代理;
基于所存储的3D对象生成第二3D对象代理;
基于在视频的帧中标识的3D对象变换所述第一3D对象代理;
基于在所述视频的关键帧中标识的所述3D对象变换所述第二3D对象代理;
将颜色属性从在所述视频的帧中标识的所述3D对象映射到经变换的第一3D对象代理;
将颜色属性从在所述关键帧中标识的所述3D对象映射到经变换的第二3D对象代理;以
及

基于所述经变换的第一3D对象代理的颜色属性和所述经变换的第二3D对象代理的颜色属性生成所述3D对象的颜色数据。

19. 根据权利要求18所述的方法,进一步包括:

在变换所述第一3D对象代理之前,使用自动编码器对所述第一3D对象代理进行编码;
以及
在变换所述第二3D对象代理之前,使用所述自动编码器对所述第二3D对象代理进行编码。

20. 根据权利要求18所述的方法,进一步包括:

在变换所述第一3D对象代理之后,使用自动编码器对所述第一3D对象代理进行解码;
以及

在变换所述第二3D对象代理之后,使用所述自动编码器对所述第二3D对象代理进行解码。

21. 根据权利要求18所述的方法,其中生成所述3D对象的颜色数据包括从所述经变换的第二3D对象代理的颜色属性中减去所述经变换的第一3D对象代理的颜色属性。

22. 根据权利要求18所述的方法,其中生成所述3D对象的颜色数据包括将所述经变换的第一3D对象代理的颜色属性添加到所述经变换的第二3D对象代理的颜色属性。

通过提供几何代理进行视频编码

[0001] 相关申请

[0002] 本申请是2018年9月26日提交的美国非临时申请No.16/143,165的继续并且要求其权益,其全部内容通过引用合并于此。

技术领域

[0003] 实施例涉及对三维(3D)视频数据进行压缩和解压缩。

背景技术

[0004] 用于视频压缩的技术都通过通用方法相关。通常,视频的帧通过将该帧的块定义为残差(例如,依据从先前帧或将来帧的位移)来临时压缩。对于帧内具有可以通过平面内刚性变换表征的残差的对象(例如,随时间推移在图像平面中移位和旋转的对象),此压缩技术通常是可接受的(例如,在解压缩时具有最小的伪像或误差)。尽管此模型捕获许多视频动态源(例如,相机或场景平移),但是存在常见场景,针对其这不是最佳模型(解压缩时效率低下或包含过多的伪像或误差)。

[0005] 换句话说,典型的预测方案可以可靠地预测先前和/或未来帧(例如,关键帧)中的像素/块/补片,以用于当对象大部分线性移动和/或具有从帧到帧的可预测运动。然而,当对象具有从帧到帧的动态非线性运动时,典型的预测方案可能无法可靠地预测先前帧和/或未来帧(例如,关键帧)中的像素/块/补片以用于计算残差。因此,当对象具有从帧到帧的动态非线性运动时,使用位移预测模型将可能会导致很少的压缩。

发明内容

[0006] 示例实施方式描述通过几何代理使用颜色预测来压缩视频的帧的系统和方法。

[0007] 在总体方面,一种方法和其上存储有计算机可执行程序代码的非暂时性计算机可读存储介质,所述计算机可执行程序代码在计算机系统上执行时,使计算机系统执行步骤。所述步骤包括,接收视频的帧,在该帧中标识三维(3D)对象,将3D对象与所存储的3D对象进行匹配,基于3D对象和所存储的3D对象使用颜色预测方案压缩视频帧,以及存储具有元数据的压缩帧,该元数据标识3D对象,指示3D对象在视频帧中的位置并指示3D对象在视频帧中的定向。

[0008] 实施方式可以包括以下一个或多个特征。例如,基于3D对象和所存储的3D对象使用颜色预测方案对视频的帧进行压缩可以包括,基于所存储的3D对象生成第一3D对象代理,基于在帧中标识的3D对象变换第一3D对象代理,基于所存储的3D对象生成第二3D对象代理,在视频的关键帧中标识3D对象,基于在关键帧中标识的3D对象变换第二3D对象代理,将颜色属性从3D对象映射到经变换的第一3D对象代理,将颜色属性从在关键帧中标识的3D对象映射到经变换的第二3D对象代理,以及基于经变换的第一对象的颜色属性以及经变换的第二3D对象代理的颜色属性生成3D对象的残差。

[0009] 例如,基于3D对象和所存储的3D对象使用颜色预测方案对视频的帧进行压缩可以

包括,基于所存储的3D对象生成第一3D对象代理,基于帧中标识的3D对象变换第一3D对象代理,基于所存储的3D对象生成第二3D对象代理,在视频的关键帧中标识3D对象,基于在关键帧中标识的3D对象变换第二3D对象代理,将颜色属性从3D对象映射到经变换的第一3D对象代理,以及基于经变换的第一3D对象代理的颜色属性和经变换的第二3D对象代理的默认颜色属性生成3D对象的残差。

[0010] 例如,基于3D对象和所存储的3D对象使用颜色预测方案对视频的帧进行压缩可以包括,基于所存储的3D对象生成第一3D对象代理,使用自动编码器对第一3D对象代理进行编码,基于帧中标识的3D对象变换编码的第一3D对象代理,基于存储的3D对象生成第二3D对象代理,使用自动编码器对第二3D对象代理进行编码,标识在视频的关键帧中的3D对象,基于在关键帧中标识的3D对象变换编码的第二3D对象代理,将来自于3D对象的颜色属性映射到经变换的第一3D对象代理,将来自于在关键帧中标识的3D对象的颜色属性映射到经变换的第二3D对象代理,以及基于用于经变换的第一3D对象代理的颜色属性以及用于经变换的第二3D对象代理的颜色属性为3D对象生成残差。

[0011] 例如,基于3D对象和所存储的3D对象使用颜色预测方案对视频的帧进行压缩可以包括基于所存储的3D对象生成第一3D对象代理,使用自动编码器对第一3D对象代理进行编码,基于在帧中标识的3D对象变换所编码的第一3D对象代理,基于所存储的3D对象生成第二3D对象代理,使用自动编码器对第二3D对象代理进行编码,在视频的关键帧中标识3D对象,基于在关键帧中标识的3D对象对所编码的第二3D对象代理进行变换,将颜色属性从3D对象映射到经变换的第一3D对象代理,以及基于经变换的第一3D对象代理的颜色属性和经变换的第二3D对象代理的默认颜色属性生成该3D对象的残差。

[0012] 例如,在存储3D对象之前,这些步骤可以进一步包括:标识与视频相关联的至少一个感兴趣的3D对象,确定与感兴趣的3D对象相关联的多个网格属性,确定与感兴趣的3D对象相关联的位置,确定与感兴趣的3D对象相关联的定向,确定与感兴趣的3D对象相关联的多个颜色属性以及使用自动编码器减少与感兴趣的3D对象的网格属性相关联的变量的数量。压缩视频的帧可以包括确定3D对象相对于关键帧中的背景3D对象的原点坐标的位置坐标。所存储的3D对象可以包括默认颜色属性,并且颜色预测方案可以使用默认颜色属性。这些步骤可以进一步包括:标识与视频相关联的至少一个感兴趣的3D对象;基于至少一个感兴趣的3D对象,生成至少一个所存储的3D对象,至少一个所存储的3D对象中的每一个由包括通过面连接的点的集合的网格定义,每个点存储至少一个属性,该至少一个属性包括相应点的位置坐标,以及与视频相关联地存储至少一个所存储的3D对象。

[0013] 在另一个总体方面,一种方法和其上存储有计算机可执行程序代码的非暂时性计算机可读存储介质,所述计算机可执行程序代码在计算机系统上执行时,使计算机系统执行步骤。这些步骤包括接收视频的帧,在帧中标识三维(3D)对象,将3D对象与所存储的3D对象进行匹配,基于3D对象和所存储的3D对象使用颜色预测方案对视频的帧进行解压缩,以及渲染视频的帧。

[0014] 实施方式可以包括以下一个或多个特征。例如,基于3D对象和所存储的3D对象使用颜色预测方案对视频的帧进行解压缩可以包括,基于所存储的3D对象生成第一3D对象代理;基于在帧中标识的3D对象变换第一3D对象代理,在视频的关键帧中标识3D对象,基于在关键帧中标识的3D对象,变换第二3D对象代理,将颜色属性从3D对象映射到经变换的第一

3D对象代理,将颜色属性从在关键帧中标识的3D对象映射到经变换的第二3D对象代理,以及基于经变换的第一3D对象代理的颜色属性和经变换的第二3D对象代理的颜色属性生成3D对象的颜色属性。

[0015] 例如,基于3D对象和所存储的3D对象使用颜色预测方案对视频的帧进行解压缩可以包括基于所存储的3D对象生成第一3D对象代理,基于帧中标识的3D对象变换第一3D对象代理,在视频的关键帧中标识3D对象,基于在关键帧中标识的3D对象变换第二3D对象代理,将颜色属性从3D对象映射到经变换的第一3D对象代理,基于经变换的第一3D对象代理的颜色属性和经变换的第二3D对象代理的默认颜色属性生成3D对象的颜色属性。

[0016] 例如,基于3D对象和所存储的3D对象使用颜色预测方案对视频的帧进行解压缩可以包括基于所存储的3D对象生成第一3D对象代理,使用自动编码器对第一3D对象代理进行解码,基于与3D对象相关联的元数据变换已解码的第一3D对象代理,基于所存储的3D对象生成第二3D对象代理,使用自动编码器对第二3D对象代理进行解码,在视频的关键帧中标识3D对象,基于与在关键帧中标识的3D对象相关联的元数据对已解码的第二3D对象代理进行变换,将颜色属性从3D对象映射到经变换的第一3D对象代理,将颜色属性从在关键帧中标识的3D对象映射到经变换的第二3D对象代理,以及基于经变换的第一3D对象代理的颜色属性和经变换的第二3D对象代理的颜色属性生成3D对象的颜色属性。

[0017] 例如,基于3D对象和所存储的3D对象使用颜色预测方案对视频的帧进行解压缩可以包括基于所存储的3D对象生成第一3D对象代理,使用自动编码器对第一3D对象代理进行解码,基于与3D对象相关联的元数据对已解码的第一3D对象代理进行变换,基于所存储的3D对象生成第二3D对象代理,使用自动编码器对第二3D对象代理进行解码,在视频的关键帧中标识3D对象,基于与在关键帧中标识的3D对象相关联的元数据对已解码的第二3D对象代理进行变换,将颜色属性从3D对象映射到经变换的第一3D对象代理,以及基于经变换的第一3D对象代理的颜色属性和经变换的第二3D对象代理的默认属性生成3D对象的颜色属性。

[0018] 例如,步骤可以进一步包括使用机器训练的生成建模技术来接收3D形状的至少一个潜在表示以:确定与3D形状相关联的多个网格属性;确定与3D形状相关联的位置;确定与3D形状相关联的定向;以及确定与3D形状相关联的多个颜色属性,并将3D形状存储为所存储的3D对象。渲染视频的帧可以包括:接收3D对象相对于关键帧中的背景3D对象的原点坐标的位置坐标,以及使用位置坐标将3D对象放置在帧中。这些步骤可以进一步包括:接收由自动编码器的编码器使用的神经网络以减少与至少一个感兴趣的3D对象的网格属性、位置、定向和颜色属性相关联的变量的数量,在自动编码器的解码器中使用神经网络再生与至少一个感兴趣的3D对象的网格相关联的点,再生点包括再生位置属性、定向属性和颜色属性,以及将至少一个感兴趣的3D对象存储为所存储的3D对象。

[0019] 在又一总体方面中,一种方法和其上存储有计算机可执行程序代码的非暂时性计算机可读存储介质,所述计算机可执行程序代码在计算机系统上执行时使计算机系统执行用于使用代理预测颜色变化(color variance)的步骤。这些步骤包括:基于所存储的3D对象生成第一3D对象代理;基于所存储的3D对象生成第二3D对象代理;基于在视频的帧中标识的3D对象变换第一3D对象代理;基于在视频的关键帧中标识的3D对象变换第二3D对象代理,将颜色属性从在视频帧中标识的3D对象映射到经变换的第一3D对象代理,将颜色属性

从在关键帧中标识的3D对象映射到经变换的第二3D对象代理,以及基于经变换的第一3D对象代理的颜色属性和经变换的第二3D对象代理的颜色属性生成3D对象的颜色数据。

[0020] 实施方式可以包括以下一个或多个特征。例如,这些步骤进一步可以包括:在变换第一3D对象代理之前,使用自动编码器对第一3D对象代理进行编码,以及在变换第二3D对象代理之前,使用自动编码器对第二3D对象代理进行编码。步骤可以进一步包括在变换第一3D对象代理之后,使用自动编码器对第一3D对象代理进行解码,以及在变换第二3D对象代理之后,使用自动编码器对第二3D对象代理进行解码。生成3D对象的颜色数据可以包括从经变换的第二3D对象代理的颜色属性中减去经变换的第一3D对象代理的颜色属性。生成3D对象的颜色数据可以包括将经变换的第一3D对象代理的颜色属性添加到经变换的第二3D对象代理的颜色属性。

附图说明

[0021] 通过下面在此给出的详细描述和附图,示例实施方式将变得更加充分地理解,其中,相似的元件由相似的附图标记表示,其仅以图示的方式给出,并且因此不限制示例实施方式并且其中:

[0022] 图1图示根据示例实施方式的用于压缩视频的信号流的框图。

[0023] 图2图示根据示例实施方式的用于存储压缩视频的信号流的框图。

[0024] 图3A图示根据示例实施方式的编码器预测模块的框图。

[0025] 图3B图示根据示例实施方式的另一编码器预测模块的框图。

[0026] 图4A图示根据示例实施方式的解码器预测模块的框图。

[0027] 图4B图示根据示例实施方式的另一解码器预测模块的框图。

[0028] 图5A图示根据示例实施方式的用于对3D对象进行编码的信号流的框图。

[0029] 图5B图示根据示例实施方式的用于解码3D对象的信号流的框图。

[0030] 图6A图示用于流传输视频并在客户端设备上渲染视频的信号流的框图。

[0031] 图6B图示用于流传输视频并在客户端设备上渲染视频的另一信号流的框图。

[0032] 图7图示根据至少一个示例实施例的用于压缩视频的帧的方法的框图。

[0033] 图8图示根据至少一个示例实施例的用于压缩视频的帧的另一种方法的框图。

[0034] 图9图示根据至少一个示例实施例的用于解压缩和渲染视频的帧的方法的框图。

[0035] 图10图示根据至少一个示例实施例的用于压缩3D对象的方法的框图。

[0036] 图11图示根据至少一个示例实施例的用于解压缩3D对象的方法的框图。

[0037] 图12图示根据至少一个示例实施例的视频编码器系统。

[0038] 图13图示根据至少一个示例实施例的视频解码器系统。

[0039] 图14图示根据至少一个示例实施例的计算机设备和移动计算机设备的示例。

[0040] 应该注意的是,这些图旨在图示在某些示例实施方式中利用的方法和/或结构的一般特性,并且旨在补充在下面提供的撰写描述。然而,这些附图不是按比例的和可能不精确地反映任何给定实施方式的精确结构或性能特性,并且不应该被解释为限定或者限制由示例实施方式所包含的值或属性的范围。在各个附图中使用类似或相同的附图标记旨在指示存在类似或相同的元件或特征。

具体实施方式

[0041] 尽管示例实施例可以包括各种修改和替代形式,但是其实施例在附图中通过示例示出,并且将在本文中进行详细描述。然而,应理解,无意将示例实施例限制为所公开的特定形式,而是相反,示例实施例将覆盖落入权利要求的范围内的所有修改、等效物和替代形式。在整个附图的描述中,相同的标号指代相同的元件。

[0042] 在3D动态环境中,经历复杂3D变换的3D对象可能不允许基于先前帧和/或未来帧、先前压缩的图像或帧等进行简单的刚性变换。例如,经历由刚性、清晰度和变形分量组成的变换的人类演员将导致像素空间中的高度非线性变换。随后,附近的关键帧中可能没有合适的对应块。这些3D对象在本文中称为动态3D对象。

[0043] 非动态的可以包括由于相机或场景平移而看起来是从帧到帧移动的3D对象,可以通过几何代理使用颜色预测来进行编码/解码。在该示例中,因为捕获场景的相机正在(例如,以可预测的方式和/或方向)移动,所以静止的3D对象可以看起来是正在从帧到帧移动。在该示例中,3D对象(例如,车辆,即,火车、汽车或飞机)可以以可预测的方式(例如,以恒定的速度和/或方向)从帧到帧移动。这些对象有时在本文中称为平移3D对象。非动态3D对象的另一个示例是看起来不在场景内移动的3D对象。在此示例中,静止或固定的3D对象(例如,场景的背景、场景内的固定位置的家具或远处的缓慢移动的对象)可能看起来从帧到帧是静止的(例如,没有任何相机或场景平移)。这些对象有时在本文中称为固定或背景3D对象。

[0044] 示例实施方式将所存储的3D对象用作出现在视频中的对象的几何代理。所存储的3D对象可以是可以在预测方案中使用的可变形3D形状模型(例如,具有必要时可以进行操纵的给定属性的网格)。例如,可以基于所存储的3D对象来预测与动态3D对象相关联的像素在视频的帧中的位置。类似地,预测方案可用于压缩视频的帧中的背景图像、在视频的帧的图像平面内移动(例如,可预测的移动)的3D对象、3D对象的一部分、包括3D对象的视频帧的层(例如,作为Z阶图像的一部分)、包括3D对象的容器(例如,标准网格对象)等。

[0045] 一种或多种实施方式可以包括使用所存储的3D对象在关键帧中定位像素、块和/或补片。一种或多种实现方式可以包括将视频的帧中的像素、块和/或补片与所存储的3D对象的像素、块和/或补片匹配。一种或多种实施方式可以包括在存储所存储的3D对象之前压缩所存储的3D对象,以及在预测过程期间解压缩所存储的3D对象。一种或多种实现方式可以包括,在存储所存储的3D对象之前压缩所存储的3D对象;压缩与视频的帧相关联的3D对象;以及在预测过程中使用压缩3D对象。

[0046] 图1图示根据示例实施方式的用于压缩视频的信号流的框图。如图1中所示,编码器105包括帧110、预测模块115和编码模块120。此外,视频存储125包括元数据130、所存储的3D对象135和压缩帧140。

[0047] 视频数据5被输入到编码器105,其中从包括在视频数据5中的多个帧中选择帧110。视频数据5可以对应于3D视频(例如,单眼视图)、2D视频、视频的一部分(例如,少于视频的所有帧)等。因此,帧110可以包括与视频的帧相对应的数据。编码器105可以被配置成基于使用3D对象作为几何代理来使用颜色预测方案。编码器105可以基于将3D对象用作几何代理来使用颜色预测方案来压缩帧110。压缩帧110可以减少用于存储和/或传达帧110的数据量。压缩帧110可以包括预测步骤、量化步骤、变换步骤和熵编码步骤。

[0048] 帧间预测可以通过计算依据一个或多个邻近帧所表达的增量值来利用空间冗余(例如,帧之间的像素之间的相关性)。增量代码化可以包括定位关键帧(例如,先前的邻近关键帧、即将到来的邻近关键帧)中的相关像素/块/补片,以及然后计算正被编码的帧中的像素的增量值。增量值可以称为残差。因此,帧间预测可以生产帧中的像素/块/补片(例如,3D对象)的残差。在示例实施方式中,可以依据显式纹理(颜色)、默认纹理、预定义纹理、所标识的3D对象的纹理等来表达增量值。

[0049] 预测模块115可以被配置成在帧110和/或关键帧中定位3D对象。例如,机器视觉、计算机视觉和/或计算机图像识别技术可以用于标识和定位3D对象。一旦已经标识出3D对象,就可以使用坐标系(例如,2D笛卡尔、3D笛卡尔、极坐标等)作为与所标识的3D对象的网格相关联的点的属性来定位3D对象。

[0050] 在示例实施方式中,使用多个已知图像基于训练(机器学习)卷积神经网络的计算机图像识别技术可以用于标识和定位3D对象。例如,从所选择的帧中选择和/或标识一个块、多个块和/或补片。经训练的卷积神经网络可以对所选择的块、多个块和/或补片进行操作。可以测试结果(例如,错误测试、损耗测试、发散测试等)。如果该测试产生低于阈值的值(或可替代地,取决于测试类型而产生高于阈值的值),则可以将所选择的块、多个块和/或补片标识为3D对象。

[0051] 在示例实施方式中,视频的帧可以包括指示先前标识的感兴趣3D对象被包括在帧中的标签。标签可以包括3D对象的身份和定位。例如,可以使用计算机生成图像(CGI)工具(例如,计算机动画电影)来生成视频。可以在每帧中标识和标记计算机生成的字符。此外,可以将用于每个所标识的感兴趣的3D对象(例如,所标识的字符)的模型存储为所存储的3D对象135。

[0052] 在示例实施方式中,可以由三角形网格定义所存储的3D对象135和3D对象。三角形网格可以是由三角形面连接的点的集合。每个点可以存储各种属性。例如,属性可以包括每个点的位置、颜色、纹理坐标等。属性可以包括和/或指示(例如,多个属性可以指示)在视频的帧(例如,帧110)和/或图像中的相应3D对象的定向和/或相应3D对象的位置。

[0053] 因此,在示例实时方式中,3D对象的网格属性可以足以标识和定位3D对象。可以将包括用于多个感兴趣3D对象的网格属性的模型存储为所存储的3D对象135。该模型可以被标准化。例如,可以将用于男人、女人、青少年、儿童或更一般地人类或人类的一部分(例如,身体、头部、手等)的模型存储为3D对象135。例如,用于狗、猫、鹿或更一般的四足动物或四足动物的一部分(例如,身体、头部、腿等)的模型可以被存储为所存储的3D对象135。然后可以使用模型的属性以在帧中搜索具有相似属性的3D对象。

[0054] 位于帧110中的3D对象(以下称为3D对象)可以与所存储的3D对象135之一匹配。在示例实施方式中,由计算机图像识别技术生成的3D对象的身份可以被用于搜索所存储的3D对象135。在示例实施方式中,可以将帧中找到的3D对象的标签与所存储的3D对象135中的一个的标签进行匹配。在示例实施方式中,所存储的3D对象135之一与3D对象具有相似属性的模型可以被标识为匹配。

[0055] 然后,可以在针对帧110的颜色预测方案中使用所存储的3D对象135中的所匹配的3D对象(以下称为所存储的3D对象)。例如,与所存储的3D对象相对应的网格可以被平移并且被定向以与3D对象的定向对齐。然后,可以将与经平移和定向的所存储的3D对象相对应

的点与位于附近(临时)、先前编码的关键帧中的对应(例如,相同)的3D对象的点进行匹配。然后,预测模块115可以使用对应的3D对象的匹配点来选择或预测关键帧中的像素/块/补片,以用于计算3D对象在帧110中的残差(例如,相对于关键帧的颜色位移)。

[0056] 另外,预测模块115可以生成与帧110相关联的元数据130。元数据130可以包括与位于已经使用所存储的3D对象135之一预测的帧110中的至少一个3D对象相关联的数据。元数据130可以包括与3D对象在帧110中的定位和/或定向相关联的属性(例如,网格点属性)。元数据130是相对于(对应于)视频数据5的压缩帧140之一被存储。

[0057] 编码模块120可以被配置成对残差执行一系列编码过程。例如,对应于残差的数据可以被变换、量化和熵编码。

[0058] 变换残差可以包括将来自空间域的数据(例如,像素值)转换为变换域中的变换系数。变换系数可以对应于最初与帧110中的原始块和/或补片大小相同的系数的二维矩阵。换句话说,在帧110中的原始块和/或补片中变换系数可以与数据点(例如,像素)一样多。然而,由于变换,变换系数的一部分可以具有等于零的值。通常,变换包括Karhunen-Loève变换(KLT)、离散余弦变换(DCT)、奇异值分解变换(SVD)和非对称离散正弦变换(ADST)。

[0059] 矢量坐标(例如,表示变换系数)通常由浮点或双倍精确给出。但是,这样的表示通常比实际需要的更为精确。例如,视频数据5可以源自具有一些测量误差的视频捕获设备(例如,相机、扫描仪、计算机生成的图像程序)。因此,相对大量的较低位可能是噪声。量化将给定的浮点数转换为b位长的整数表示。因此,量化可以减少每个变换系数中的数据。量化可能涉及将相对较大范围内的值映射到相对较小范围内的值,从而减少表示量化的变换系数所需的数据量。量化可以将变换系数转换为离散的量子值,其被称为量化的变换系数或量化等级。例如,量化可以将零添加到与变换系数相关联的数据。例如,编码标准可以在标量量化过程中定义128个量化等级。

[0060] 包含在量化的变换系数中的信息的期望值(例如,离散量子值的数据集)。数据集包含的唯一值数量越多,熵就越高。重复的值减少熵并获得更好的压缩效果。因此,可以对量化的变换系数进行熵编码。使用用于基于数据熵压缩数据集的一组技术中的一种来执行熵代码化。例如,熵代码化技术可以包括霍夫曼代码化、算术代码化或使用不对称数字系统。

[0061] 然后,将(针对帧110中的所有像素/块/补片的)熵编码的系数以及解码帧110所需的信息(例如,使用的预测类型、运动矢量和量化值)一起输出作为压缩帧140之一,并与视频数据5的其他压缩帧一起存储和/或与视频数据5的其他压缩帧关联存储。

[0062] 通常,编码器105和视频存储125是云或万维网中的元素。例如,编码器105可以是被实现为在云计算设备中实现的计算机硬件和/或计算机软件的多个编码器之一,其被配置成使用由一个或多个标准(例如,H.264、H.265、HEVC、VP8、VP9、VP10、AV1等)定义的压缩方案来压缩视频数据(例如,视频数据5)。例如,视频存储125可以被实现为位于云计算设备(例如,流传输服务器)中的至少一种非易失性存储器、非暂时性计算机可读介质等。在至少一个实施方式中,编码器105将视频数据5压缩为视频,并且存储压缩的视频数据以用于未来(例如,在时间上稍后)在从云计算设备流传输到客户端设备时回放视频。

[0063] 在此实施方式中,客户端设备包括解码器(例如,解码器145)。在一些实施方式中,元数据130和/或所存储的3D对象135可以被传达到客户端设备。元数据130和/或存储的3D

对象135可以按需和/或作为初始化过程被传达到客户端设备。元数据130和/或所存储的3D对象135可以在被存储到视频存储125中时被压缩和/或在被传达到客户端设备之前按需压缩。在示例实施方式中,可以使用机器训练的生成建模技术来压缩所存储的3D对象135,以生成与针对所存储的3D对象的网格的属性和位置相关联的减少数量的变量(在本文中称为潜在表示或减少潜在表示)(如下文更详细描述)。

[0064] 因此,如图1中所示,解码器145包括帧重构模块150、预测模块155和解码模块160。解码模块160可以被配置成执行编码模块120的逆运算。解码模块160可以接收压缩帧140(例如,表示由包括解码器145的客户端设备的用户为了流传输或下载而选择的3D电影)。压缩帧140可以一次被接收一个,一次被接收多个帧或者帧的块,或者被接收作为完整的3D电影。解码器145可以选择压缩帧140之一,并且解码模块160可以对所选择的压缩帧进行熵解码、去量化和逆变换。

[0065] 在示例实施方式中,解码模块160使用选择的压缩帧内的数据元素,并通过熵解码(例如,使用霍夫曼代码化、算术代码化或非对称数字系统代码化的逆变换)对数据进行解压缩以产生量化的变换系数集。对量化的变换系数进行去量化,并且对去量化的变换系数进行逆变换(例如,使用KLT、DCT、SVD或ADST的逆变换)以生成可以与通过预测模块115生成的残差相同(或近似相同)的导数残差。

[0066] 预测模块155可以被配置成确定所选择的压缩帧是否包括所存储的3D对象135之一。在示例实施方式中,元数据130可以用于确定所选择的压缩帧是否包括所存储的3D对象之一。例如,预测模块155可以在所选择的压缩帧中查询元数据130,并且如果返回元数据,则确定所选择的压缩帧包括所存储的3D对象135之一。

[0067] 预测模块155可以被配置成基于返回的元数据来为所选择的压缩帧标识3D对象、3D对象的位置(例如,定位)和3D对象的定向。预测模块155可以被配置成从所存储的3D对象135中选择所存储的3D对象(以下称为所存储的3D对象),并使用所存储的3D对象来解压缩所选择的帧。

[0068] 在示例实施方式中,可以基于所标识的3D对象的位置(例如,定位)和所标识的3D对象的定向来平移和定向与所存储的3D对象相对应的网格。然后,可以将对应于经平移和定向的所存储的3D对象的点与位于附近(临时)、先前解码的关键帧中的对应(例如,相同)3D对象的点进行匹配。然后,预测模块155可以使用对应的3D对象的匹配点来选择或预测关键帧中的像素/块/补片,以用于基于所选择的帧中的标识位置处的残差来再生(例如,计算)经平移的和定向的所存储的3D对象的颜色值和/或颜色属性。

[0069] 预测模块155可以进一步被配置成基于所选择的帧的残差或剩余部分以及关键帧的对应的像素/块/补片来为所选择的帧的剩余部分再生(例如,计算)颜色值和/或颜色属性。帧重构模块150可以被配置成基于针对经平移和定向的所存储的3D对象的再生的颜色值和/或颜色属性以及针对所选择的帧的剩余部分的再生的颜色值和/或颜色属性来重构所选择的帧。在示例实施方式中,帧重构模块150可以被配置成基于所标识的3D对象的位置将针对经平移和定向的所存储的3D对象的再生的颜色值和/或颜色属性拼接到针对所选择的帧的剩余部分的再生的颜色值和/或颜色属性。

[0070] 示例实施方式可以包括标识两个或更多个3D对象,再生两个或更多个3D对象中的每个的颜色值和/或颜色属性,以及使用两个或更多个3D对象中的每个重构所选择的帧。帧

重构模块150可以被配置成基于多个重构帧来再生视频数据5。可以渲染视频数据5(例如,纹理数据和颜色数据),并校正颜色以显示在客户端设备的显示器上。

[0071] 使用被配置成基于使用3D对象作为几何代理而使用颜色预测方案的编码器(例如,编码器105)来压缩视频(或视频帧)可能不会导致对视频、视频帧或多个视频帧的最高的压缩率(例如,最小的数据大小)。因此,示例实施方式可以包括使用两个或更多个编码器来压缩视频、视频帧或多个视频帧,每个编码器能够使用标识的颜色预测方案来压缩视频数据。

[0072] 图2图示根据示例实施方式的用于存储压缩视频(和/或视频的压缩帧)的信号流的框图。如图2中所示,第一编码器205是基于将3D对象用作几何代理而使用颜色预测方案的编码器(例如,编码器105)。此外,至少一个第二编码器210-1是使用颜色预测方案1的编码器,至少一个第二编码器210-2是使用颜色预测方案2的编码器,并且至少一个第二编码器210-i是使用颜色预测方案i的编码器。第一编码器205和至少一个第二编码器210-1、210-2、...、210-i中的每一个可以被配置成生成总共具有x位的n个帧并将其传达给压缩大小比较器215。颜色预测方案1、2、...、i可以是编码标准的默认预测方案、编码标准的可配置预测方案、基于时间位移的自定义预测方案、基于将3D对象用作几何代理的替代预测方案等。

[0073] 压缩大小比较器215可以被配置成选择第一编码器205、至少一个第二编码器210-1、210-2、...或210-i的输出中的一个以保存在视频存储125中(例如,用于稍后流传输到客户端设备)。在示例实施方式中,可以基于压缩效率来选择编码器输出。例如,可以保存具有最少位数(例如,x的最小值)的压缩视频(和/或视频的压缩帧)。

[0074] 在另一示例实施方式中,颜色预测方案可以是优选预测方案。在该实施方式中,除非存在一些条件,否则保存优选颜色预测方案的压缩视频(和/或视频的压缩帧)。例如,该条件可以基于压缩效率。除非至少一个第二编码器210-1、210-2、...、210-i之一的输出至少为比第一编码器205的输出更有效的阈值百分比(例如,10%、15%、20%等),否则可以选择第一编码器205的输出。如果至少一个第二编码器210-1、210-2、...、210-i中的一个以上至少是比第一编码器205的输出更有效的阈值百分百(例如,10%、15%、20%等),可以保存至少一个第二编码器210-1、210-2、...、210-i的最高效率(例如,x的最小值)。

[0075] 在示例实施方式中,图2中图示的用于存储压缩视频(和/或视频的压缩帧)的信号流可以在逐帧视频和/或在帧集上执行。例如,视频的每个帧可以如上所述被编码。然后,压缩大小比较器215可以基于效率或基于有条件的优选颜色预测方案来选择编码器输出。在示例实施方式中,例如,关键帧可以用于在逐帧或帧集压缩之间进行选择。

[0076] 在示例实施方式中,可以始终保存编码器输出之一。例如,可以保存默认的颜色预测方案以确保向后兼容。在图2中,该实施方式被图示为来自至少一个第二编码器210-i的虚线。

[0077] 可以利用不同的代码化标准多次执行图2中图示的用于存储压缩视频(和/或视频的压缩帧)的信号流。例如,可以使用H.264、H.265、HEVC、VP8、VP9、VP10、AV1等代码化标准中的两个或更多个来执行图2中图示的信号流。因此,可以在两个或更多个代码化标准中并且在同一视频(例如,视频数据5)上实现基于将3D对象用作几何代理的颜色预测方案(例如,编码器105)。因此,视频存储125可以存储视频的多个实例(和/或视频的压缩帧),每个

实例已经使用不同的标准和/或标准的不同配置进行压缩。结果,流传输服务器可以基于请求客户端设备和/或要通过其传达视频的网路的能力来服务于视频。

[0078] 图3A图示根据示例实施方式的编码器预测模块的框图。如图3A中所示,预测模块115包括帧110、3D对象定位器模块305、3D对象匹配模块310、所存储的3D对象135、所存储的3D对象平移模块315、关键帧320、块匹配模块325和残差生成模块330。

[0079] 在示例实施方式中,所存储的3D对象135和帧110中的3D对象可以由三角形网格定义。三角形网格可以是由三角形面连接的点的集合。每个点可以存储各种属性。例如,属性可以包括每个点的位置、颜色、纹理坐标等。属性可以包括和/或指示(例如,多个属性可以指示)对应的3D对象的定向和/或对应的3D对象在视频的帧110和/或图像中的位置。

[0080] 3D对象定位器模块305可以被配置成在帧110和关键帧320中标识和定位3D对象(在下文中称为3D对象)。例如,机器视觉技术、计算机视觉技术、计算机图像识别技术等等可用于标识和定位帧110和关键帧320中的3D对象。一旦已经标识出3D对象,可以使用坐标系(例如,2D笛卡尔、3D笛卡尔、极坐标等)作为与所标识的3D对象的网格相关联的点的属性来定位3D对象。

[0081] 在示例实施方式中,使用多个已知图像基于训练(机器学习)卷积神经网络的计算机图像识别技术可以用于标识3D对象。例如,从所选择的帧中选择和/或标识一个块、多个块和/或补片。经训练的卷积神经网络可以对所选择的块、多个块和/或补片进行操作。可以测试结果(例如,错误测试、损耗测试、分散测试等)。如果测试产生低于阈值的值(或可替代的,取决于测试类型而高于阈值的值)阈值,则可以将所选择的块、多个块和/或补片标识为3D对象。

[0082] 在示例实施方式中,包括先前标识的感兴趣3D对象的视频帧还可以(例如,在头部、帧元数据等中)包括指示先前标识的感兴趣3D对象被包括在帧中的标签。标签可以包括3D对象的身份和定位(例如,与网格相关联的点的坐标属性)。例如,可以使用计算机生成图像(CGI)工具(例如,计算机动画电影)来生成视频。可以在每帧中标识和标记计算机生成的字符。此外,可以将针对每个所标识的感兴趣3D对象(例如,动画电影中的角色)的(例如,由三角形网格定义的)模型存储为所存储的3D对象135。

[0083] 在示例实施方式中,3D对象的网格属性可以足以标识和定位3D对象。可以将包括用于多个通用感兴趣3D对象的网格属性的模型存储在3D对象定位器模块中(或与之相关联)。可以标准化该模型。例如,可以存储用于男人、女人、青少年、儿童或更一般地人类或人类的一部分(例如,身体、头部、手等)的模型。例如,可以存储用于狗、猫、鹿或更一般的四足动物或四足动物的一部分(例如,身体、头部、腿等)的模型。然后,可以使用模型的属性来在该帧中搜索具有相似属性的3D对象。

[0084] 3D对象匹配模块310可以被配置为将位于帧110中的3D对象(在下文中,称为3D对象)与所存储的3D对象135之一进行匹配。在示例实施方式中,计算机图像识别技术由3D对象定位器模块305使用来标识3D对象。标识3D对象还可以包括将来自ID数据库的唯一ID分配给3D对象。唯一ID可以被用来在所存储的3D对象135中搜索3D对象。如果在所存储的3D对象135中找到唯一ID,则所存储的3D对象中的相应一个与该3D对象相匹配。

[0085] 在示例实施方式中,在帧中找到的3D对象的标签对3D对象可以是唯一的(例如,在标签和3D对象之间具有一对一关系)。标签可以被用来在所存储的3D对象135中搜索3D对

象。如果在所存储的3D对象135中找到标签,则所存储的3D对象中的相应一个与该3D对象相匹配。

[0086] 在示例实施方式中,所存储的3D对象135之一与3D对象具有相似属性的模型可以被标识为部分匹配。然后可以基于部分匹配来过滤所存储的3D对象135。然后,可以使用3D对象的一个或多个属性或属性的组合来在所存储的3D对象135中搜索该3D对象。如果在所存储的3D对象135中找到一个或多个属性或属性的组合,则所存储的3D对象中的相应一个是该3D对象的匹配。一个或多个属性或属性的组合可以以相对高的确定性,相对于所存储的3D对象135,唯一地标识3D对象。例如,身体部位的形状或相对位置(例如,面部形状、鼻子形状、眼睛,鼻子和嘴巴的相对位置等)、卡通人物的头发或皮肤的颜色、角色所穿戴的对象(例如珠宝)的形状和相对定位、车辆类型(例如汽车、拖拉机等)等可以相对于所存储的3D对象135,唯一地标识3D对象(例如,如果所存储的3D对象135对应于视频的感兴趣3D对象集)。

[0087] 如果3D对象匹配模块310在所存储的3D对象中未找到3D对象的匹配,则预测模块115可以使用由代码化标准定义的标准预测技术。换句话说,基于将3D对象用作几何代理的颜色预测方案可以不被用于该3D对象。

[0088] 如果3D对象匹配模块310未在所存储的3D对象135中找到3D对象的匹配,则3D对象匹配模块310可以被配置为将3D对象添加到所存储的3D对象135中。例如,3D对象匹配模块310可以被配置为向3D对象分配唯一ID或唯一标签,并且为该3D对象定义模型(例如,包括点和对应属性的三角形网格)。然后将模型存储为所存储的3D对象135之一,并且由唯一ID或唯一标签标识。

[0089] 在示例实施方式中,所存储的3D对象135可以具有预定网格表示和/或数据结构。因此,模型可以具有基于所存储的3D对象135的设计而预定义的大小(例如,点的数量、面的数量、顶点的数量等)、属性的数量、属性的类型等。此外,使该3D对象与所存储的3D对象135中的3D对象匹配可以包括在所存储的3D对象135中搜寻匹配之前,由3D对象匹配模块310基于所存储的3D对象135的设计来重新定义3D对象。

[0090] 所存储的3D对象平移模块315可以被配置为平移(或变换)所存储的3D对象135中的所匹配的一个(在下文中,称为所存储的3D对象)。例如,对应于所存储的3D对象的网格可以被平移和定向以与3D对象的定向对准。平移和定向一起可以被称为变换。所存储的3D对象平移模块315可以被配置为变换与帧110和关键帧320相关联的所存储的3D对象135中的所匹配的一个。因此,所存储的3D对象平移模块315可以被配置为基于与帧110相关联的针对3D对象的所存储的3D对象生成并变换第一3D对象代理。此外,所存储的3D对象平移模块315可以被配置为基于与关键帧320相关联的针对3D对象的所存储的3D对象生成和变换第二3D对象代理。

[0091] 变换3D代理可以包括通过铰接网格点(例如,像骨架),使用自动编码器生成潜在表示(如下所述)和定位与潜在表示相关联的值,发送用于Bezier,或NURBS曲面或细分曲面的控制点和值。每个网格可以具有预定连通性。预定连通性可以允许处于不同姿势的两个网格的点之间的直接对应。因此,我们可以对两个网格使用相同的纹理参数化。

[0092] 所存储的3D对象平移模块315可以被配置为生成元数据20。元数据20包括标识所存储的3D对象的信息以及与所存储的3D对象的平移和定向有关的信息。与所存储的3D对象

的平移和定向有关的信息可以用于在将来(例如,由解码器145)执行所存储的3D对象的相同平移和定向。在示例实施方式中,针对第一3D对象代理和第二3D对象代理两者生成和存储元数据。

[0093] 块匹配模块325可以被配置为匹配与关键帧320的经平移和定向的所存储的3D对象相对应的帧110的经平移和定向的所存储的3D对象的点。在示例实施方式中,块匹配模块325可以被配置为用颜色和/或纹理属性包装表示经平移和定向的所存储的3D对象的网格(例如3D网格)。例如,块匹配模块325可以被配置为将颜色属性从在帧110中标识的3D对象映射到经变换的第一3D对象代理。此外,块匹配模块325可以被配置为将颜色属性从在关键帧320中标识的3D对象映射到经变换的第二3D对象代理。将颜色属性从在帧110中标识的3D对象映射到经变换的第一3D对象代理可以包括将3D对象从3D空间(例如XYZ空间)转换到2D空间(例如UV空间)和/或将在帧110中标识的3D对象从2D空间(例如,UV空间)转换到3D空间(例如,XYZ空间)。

[0094] 每个网格可以具有预定连通性。预定连通性可以允许处于不同姿势的两个网格的点之间的直接对应。因此,我们可以对两个网格使用相同的纹理参数化。因此,映射颜色属性可以包括基于帧中的像素坐标来标识帧(例如,帧110和/或关键帧320)中的像素,然后将具有相同坐标的经变换的3D对象代理的网格表示中的点的颜色属性设置为所标识的像素的相同颜色值。在示例实施方式中,可以使用一个以上的帧来生成像素属性。此外,来自再生帧(例如,来自编码器中的再生循环)的像素属性可以被用来生成像素属性。

[0095] 在示例实施方式中,可以混合像素属性。混合像素属性(或纹理)可以使用来自那些纹理的相应纹素(texel)的每个纹素平均值来完成,这些纹理具有有关该纹素的信息,来自不同纹理的纹素可能具有置信度值(例如,当前帧与由其生成该纹理的帧之间的时间距离),还可以预测特定纹理中未观察到的纹素(降低它们的置信度),对于当前帧,使用网格的新姿势及其纹理来获取对当前帧(的一部分)的预测等。

[0096] 在一些情况下,帧110和/或关键帧320可能不具有与经变换的3D对象代理的网格表示上的点相对应的像素。因此,经变换的3D对象代理的网格表示可以保留这些点的默认颜色属性。因此,任何残差颜色计算都可以基于默认颜色属性。

[0097] 然后,块匹配模块325可以使用相应的3D对象的匹配点来选择或预测关键帧320中的像素/块/补片,以用于生成帧110中的3D对象的残差15。在示例实施方式中,在平移和匹配之前,可以使用自动编码器对所存储的3D对象和/或3D对象代理进行编码(在下文更详细地描述)。编码所存储的3D对象或3D对象代理将3D对象、所存储的3D对象和/或3D对象代理转换为潜在表示。潜在表示包括比表示3D对象、所存储的3D对象和/或3D对象代理的网格更少的值(例如,点)。因此,将3D对象、所存储的3D对象和/或3D对象代理平移为潜在表示包括操纵更少的点。此外,映射颜色属性包括当将3D对象、所存储的3D对象和/或3D对象代理编码为潜在表示时,映射较少的点。

[0098] 残差生成模块330可以被配置为生成(或计算)帧110中的3D对象的残差15(例如,相对于关键帧320的颜色位移)。例如,残差生成模块330可以通过从用于关键帧320中的预测像素/块/补片的每个匹配点的像素属性值中减去用于3D对象的三角形网格中的每个点的像素属性值来生成残差15。在示例实施方式中,可以从第二3D对象代理的网格表示中的相应点(例如,具有相同的点标识或处于网格序列中的相同位置)的颜色属性中减去第一3D

对象代理的网格表示中的点的颜色属性。

[0099] 在示例实施方式中,编码器105包括重构路径(未示出)。重构路径包括使用上述编码过程的至少一个逆过程而一起解码帧的若干组件或软件实现。例如,重构路径可以至少包括逆量化过程和逆变换过程。当编码下一连续帧时,可以使用在重构路径中生成的重构帧来代替关键帧320。编码原始帧会造成一些损失。因此,与原始帧相比,在重构路径中生成的重构帧包括一些伪像(例如,误差)。在预测模块115中可以校正也可以不校正这些伪像。在一些实施方式中,可以通过下文描述的颜色校正模块635来校正使用重构帧生成的伪像。

[0100] 图3B示出了根据示例实施方式的另一编码器预测模块的框图。如图3B所示,预测模块115包括帧110、3D对象定位器模块305、3D对象匹配模块310、所存储的3D对象135、所存储的3D对象平移模块315、块匹配模块325和残差生成模块330。

[0101] 图3B所示的预测模块115的实现基本上与图3A所示的预测模块115相同。但是,图3B中所示的预测模块115不基于关键帧生成残差15。相反,图3B中所示的预测模块115的残差生成模块330使用3D对象以及经平移和定向的所存储的3D对象来生成残差15。

[0102] 在示例实施方式中,经平移和定向的所存储的3D对象与3D对象具有直接的点对点关系。换句话说,在定义经平移和定向的所存储的3D对象的三角形网格中的点(例如,位置属性为 x_1, y_1, z_1)在定义3D对象的三角形网格中具有相应的点(例如,位置属性为 x_1, y_1, z_1)。因此,颜色属性可以用于确定3D对象和所存储的3D对象之间的相对颜色位移。因此,残差生成模块330可以被配置为基于所存储的对象来生成(或计算)帧110中的3D对象的残差15(例如,颜色位移)。例如,残差生成模块330可以通过从经平移和定向的所存储的3D对象的三角形网格中的相应点的像素属性值中减去用于3D对象的三角形网格中的每个点的像素属性值来生成残差15。

[0103] 在示例实施方式中,可以使用经平移的3D对象代理(例如上述第一3D对象代理)和具有默认颜色属性的所存储的3D对象来生成残差,该经平移的3D对象代理基于与帧110相关联的针对3D对象的所存储的3D对象,该所存储的3D对象已经映射有帧110的颜色。因此,可以从所存储的3D对象的网格表示中的相应点(例如,具有相同的点标识或处于网格序列中的相同位置)的颜色属性中减去第一3D对象代理的网格表示中的点的颜色属性。

[0104] 图4A示出了根据示例实施方式的解码器预测模块的框图。如图4A所示,预测模块155包括帧405、元数据模块410、3D对象匹配模块415、所存储的3D对象135、所存储的3D对象平移模块420、关键帧320、块匹配模块325和像素再生模块425。

[0105] 帧405是选自压缩帧140中的作为要解压缩的帧的压缩帧。关键帧320是与帧405相关联的解压缩关键帧(附近、时间上更早和/或时间上更晚)。帧405已经由解码模块160进行了熵解码、逆量化和逆变换。因此,帧405可以包括导数残差,该导数残差可以与预测模块115生成的残差相同(或近似相同)。

[0106] 元数据模块410可以被配置为确定帧405是否包括元数据和/或具有相关联的元数据(例如,元数据130)。元数据模块410可以被配置为从与帧405相关联的报头中读取元数据。元数据可以包括元数据20。元数据可以包括与位于帧405中的至少一个3D对象相关联的数据,该至少一个3D对象已经使用所存储的3D对象135中的一个进行了预测。元数据可以包括与帧405中的3D对象的定位和/或定向相关联的属性(例如,网格点属性)。元数据可以包括标识所存储的3D对象的信息和与所存储的3D对象的平移和定向有关的信息。与所存储的

3D对象的平移和定向有关的信息可以用于执行与由所存储的3D对象平移模块315执行的所存储的3D对象的相同平移和定向。

[0107] 3D对象匹配模块415可以被配置为将位于帧405中的3D对象(在下文中,称为3D对象)与所存储的3D对象135中的一个进行匹配。在示例实施方式中,元数据模块410输出可以被用来在所存储的3D对象135中搜索3D对象的唯一ID或标记。如果在所存储的3D对象135中找到唯一ID或标签,则所存储的3D对象135中的相应一个(在下文中,被称为所存储的3D对象)是3D对象的匹配。

[0108] 所存储的3D对象平移模块420可以被配置为平移所存储的3D对象。例如,在压缩帧405之前,当在帧405中定向3D对象时,可以平移和定向与所存储的3D对象相对应的网格以与3D对象的定向对齐。换句话说,当在帧110中定向3D对象时,可以平移和定向与所存储的3D对象相对应的网格以与3D对象的定向对齐。所存储的3D对象平移模块420可以被配置为基于与元数据中包括的所存储的3D对象的平移和定向有关的信息来平移所存储的3D对象。

[0109] 所存储的3D对象平移模块420可以被配置为平移与帧405和关键帧320相关联的所存储的3D对象135中的所匹配的一个。因此,所存储的3D对象平移模块315可以被配置为基于与帧405相关联的针对3D对象的所存储的3D对象而生成第一3D对象代理。此外,所存储的3D对象平移模块315可以被配置成基于与关键帧320相关联的针对3D对象的所存储的3D对象而生成第二3D对象代理。在示例实施方式中,第二3D对象代理可以每个关键帧320生成一次,然后用于每个相关联的帧405。

[0110] 块匹配模块325可以被配置为匹配与关键帧320的经平移和定向的所存储的3D对象相对应的帧405的经平移和定向的所存储的3D对象的点。在示例实施方式中,块匹配模块325可以被配置为将颜色属性从帧110中标识的3D对象映射到经变换的第一3D对象代理。此外,块匹配模块325可以被配置为将颜色属性从关键帧320中标识的3D对象映射到经变换的第二3D对象代理。将颜色属性从帧110中标识的3D对象映射到经变换的第一3D对象代理可以包括将3D对象从3D空间(例如XYZ空间)转换到2D空间(例如UV空间)和/或将在帧110中标识的3D对象从2D空间(例如,UV空间)转换到3D空间(例如,XYZ空间)。

[0111] 在示例实施方式中,映射颜色属性可以包括基于帧中的像素坐标来标识帧(例如,帧405和/或关键帧320)中的像素,然后将经变换的3D对象代理的网格表示中的具有相同坐标的点的颜色属性设置为所标识的像素的相同颜色值。在一些情况下,帧405和/或关键帧320可以不具有与经变换的3D对象代理的网格表示上的点相对应的像素。因此,经变换的3D对象代理的网格表示可以保留这些点的默认颜色属性。因此,任何残差颜色计算都可以基于默认颜色。

[0112] 然后,块匹配模块325可以使用相应的3D对象的匹配点来选择或预测关键帧320中的像素/块/补片,以用于再生帧405中的3D对象的像素。像素再生模块425可以被配置为生成(或计算)帧405中的3D对象25的像素值。例如,像素再生模块425可以基于所选择的帧中的标识位置处的残差,通过将来自用于关键帧320中的预测像素/块/补片的每个匹配点的像素属性值中的用于3D对象的三角形网格中的每个点的像素属性值与用于经平移和定向的所存储的3D对象的颜色值和/或颜色属性相加来生成像素值。

[0113] 在示例实施方式中,已经使用自动编码器对3D对象、所存储的3D对象和/或3D对象代理进行了编码。因此,在平移和匹配之前,可以使用自动编码器对3D对象、所存储的3D对

象和/或3D对象代理进行解码(在下文更详细地描述)。解码所存储的3D对象或3D对象代理将3D对象、所存储的3D对象和/或3D对象代理的潜在表示转换为再生的网格表示。

[0114] 像素再生模块425可以进一步被配置为基于所选择的帧的残差或剩余部分以及关键帧的相应像素/块/补片,再生(例如,计算)所选择的帧的剩余部分的颜色值和/或颜色属性。

[0115] 图4B示出了根据示例实施方式的另一解码器预测模块的框图。如图4B所示,预测模块155包括帧405、元数据模块410、3D对象匹配模块415、所存储的3D对象135、所存储的3D对象平移模块420、关键帧320、块匹配模块325和像素再生模块425。

[0116] 图4B所示的预测模块155的实现基本上与图4A所示的预测模块155相同。但是,图4B中所示的预测模块155不会基于关键帧再生3D对象的颜色值。相反,图4B中所示的预测模块155的像素再生模块425使用3D对象以及经平移和定向的所存储的3D对象来再生3D对象的颜色值。

[0117] 在示例实施方式中,经平移和定向的所存储的3D对象与3D对象具有直接的点对点关系。换句话说,在定义经平移和定向的所存储的3D对象的三角形网格中的点(例如,位置属性为 x_1, y_1, z_1)在定义3D对象的三角形网格中具有相应点(例如,位置属性为 x_1, y_1, z_1)。因此,经平移和定向的所存储的3D对象的颜色属性可以用于确定3D对象的颜色值。

[0118] 因此,像素再生模块425可以被配置为基于所存储的对象来生成(或计算)用于帧405中的3D对象的颜色值。例如,像素再生模块425可以通过将从帧405读取的3D对象的三角形网格中的每个点的像素值与用于经平移和定向的所存储的3D对象的三角形网格中的相应点的颜色属性值相加来再生颜色值。然后将具有所计算的颜色属性值的经平移和定向的所存储的3D对象输出为作为再生的3D对象的3D对象25。如上参考图4A所述,再生帧30的剩余部分。

[0119] 如上所述,每个所存储的3D对象135可以由三角形网格定义。三角形网格可以是由三角形面连接的点的集合。每个点可以存储各种属性。例如,属性可以包括每个点的位置、颜色、纹理坐标等。尽管上文引用了三角形网格结构,但是也可以使用其他多边形来定义所存储的3D对象135。此外,每个点的属性可以包括在本公开的上下文中可能有用或可能没有用的其他属性。

[0120] 在示例实施方式中,所存储的3D对象135可以包括用于CGI电影的感兴趣3D对象。因此,所存储的3D对象135可能包括3D模型或网格数据以及用于大量3D角色的相应属性值。通常,3D角色可以包括CGI演员(例如,主角、配角和其他角色)、CGI宠物、CGI生物、CGI怪物等。如上所述,这些3D角色中的每一个可以是存储为3D模型的感兴趣3D对象或存储在所存储的3D对象135中的网格数据。每个3D角色可以包括唯一ID或标签。

[0121] 在一些实施方式中,感兴趣3D对象的一部分可以被存储为3D模型或存储在所存储的3D对象135中的网格数据。3D角色的部分中的每一个可以包括唯一ID或标签。例如,至少一个3D角色(例如,主角和配角)可以具有表示至少一个3D角色头部或面部的相关联的3D模型或网格数据,以及表示至少一个3D角色身体的另一个相关联的3D模型或网格数据(注意头部、面部或身体也可以分成更小的部分)。通过将角色划分为多个部分,与相对较少动态的角色部分(例如,躯干和/或肩膀)相比,可以为更多动态的角色的部分(例如,头部、手臂和/或腿)存储相对较高的细节水平。

[0122] 此外,可以将网格数据配置为适合标准3D模型。例如,第一立方模型可以是旨在适合头部模型的标准尺寸,而第二立方模型可以是旨在适合手的模型的标准尺寸。其他标准形状的3D模型可以包括球形、直角棱柱、长方形棱柱、棱锥、圆柱体等。例如,手臂或腿可以使用直角棱柱3D模型或圆柱3D模型。标准3D模型还可以对用来定义3D对象的多个点(例如,顶点)设置限制(最大和/或最小)。

[0123] 划分角色的各部分并将每个部分存储为所存储的3D对象135中的感兴趣唯一3D对象可以使得所存储的3D对象135的大小呈指数增长。但是,表示角色的一部分的3D模型或网格数据可以用于许多角色。例如,表示角色A的躯干的躯干模型也可以用于角色B。所存储的3D对象135还可以包括标识表示角色的标准部分的3D模型或网格数据的元数据以及对应于用于该角色的标准部分的一些变形信息。因此,3D模型或网格数据表示标准部分。例如,可以将表示躯干的一个3D模型用作将角色A的躯干表示为高个子瘦男人的3D模型,以及用作将角色B的躯干表示为矮胖的男人的3D模型。

[0124] 随着所存储的3D对象135的数量增加(例如,对于使用大量3D标准模型的具有许多角色的全长CGI动画电影),存储所存储的3D对象135所需的资源量(例如,内存)也随之增加。例如,流传输服务器可以存储10万个、100万个及数百万个视频。随着对包括3D视频的视频需求的增加(例如,在虚拟现实、增强现实、3D CGI电影等中),在流传输服务器处存储的视频的百分比(通常存储为左眼和右眼2D视频)肯定也会增加。因此,当实现本文所述的技术时,用于与增长的视频数量一起使用。所存储的3D对象135的数量必定增加,因此存储所存储的3D对象135所需的资源量(例如,内存)也将增加。此外,在流传输活动期间,将所存储的3D对象135从流传输服务器传送到客户端设备可能需要大量带宽。因此,对于流传输操作而言,可能变得期望有效地编码和解码所存储的3D对象135。

[0125] 图5A示出了根据示例实施方式的用于编码3D对象的信号流的框图。如图5A所示,编码3D对象可以至少包括所存储的3D对象135、神经网络编码器505和3D对象的潜在表示510。

[0126] 如上所述,每个所存储的3D对象135可以由三角形网格定义。三角形网格可以是由三角形面连接的点或顶点的集合。每个点可以存储各种属性。在示例实施方式中,点或顶点的数量可以限制为一定值。

[0127] 神经网络编码器505可以使用生成建模技术来压缩3D对象。生成建模技术的示例实施方式可以包括变分和/或卷积自动编码器(VAE)、生成对抗网络(GAN)和/或组合VAE-GAN。尽管提到和/或讨论了这些生成建模技术,但是示例实施方式不限于此。此外,尽管将神经网络型编码器讨论为用于压缩3D对象的实施方式,但是编码3D对象(包括三角形网格)的其他实施方式也在本公开的范围之内。

[0128] 神经网络编码器505可以将由具有固定数量的顶点和固定连通性的网格定义的3D对象压缩为有时被称为潜在空间的相对较小数量的变量。例如,VAE可以被配置为学习用于3D形状(例如,已经在其上训练了VAE的3D形状)的紧凑的潜在空间。神经网络编码器505的神经网络和神经网络解码器515的神经网络都可以是机器训练神经网络和/或机器训练卷积神经网络。

[0129] 神经网络编码器505的神经网络可以包括与被应用于网格以便编码该网格的一个或多个卷积和/或滤波器相关联的系数。每个卷积可以具有C个滤波器、 $K \times K$ 掩码(K表示卷

积)和步长因子。系数可以对应于C个滤波器、KxK掩码和步长因子中的一个或多个。例如, KxK掩码可以是3x3掩码。3x3掩码包含用于在网格上执行卷积的九(9)个变量。换句话说, 3x3掩码具有九(9)个块,每个块都包含变量。每个变量可以分别是系数之一。

[0130] 神经网络编码器505的神经网络可以包括具有不同数量的神经元的层。KxK空间范围可以包括K列和K(或L)行。KxK空间范围可以是2x2、3x3、4x4、5x5、(KxL) 2x4等。卷积包括将KxK空间范围以网格点为中心,并且在空间范围内对所有网格点进行卷积,以及基于空间范围中的所有网格点的所有卷积(例如,总和),生成该网格点的新值。然后基于步长,将空间范围移动到新网格点,并且对新网格点重复卷积。步长可以为例如一(1)或二(2),其中,步长为1会移动到下一个网格点,而步长为2会跳过网格点。

[0131] VAE可以将用于网格的每个点的位置坐标(例如,对于任何感兴趣的大多数3D形状或具有视觉上相当大量的细节的相对大量的数据)用作神经网络编码器505的输入,并且通过将神经网络与网格的每个点的位置坐标进行卷积,生成数量减少(最好相对少量)的变量(例如从64到128个变量)。用于神经网络的C个滤波器、KxK掩码和步长因子的配置可以确定用于3D对象510的潜在表示的潜在空间中的变量数量。

[0132] 图5B示出了根据示例实施方式的用于解码3D对象的信号流的框图。如图5B所示,解码3D对象可以至少包括用于3D对象的潜在表示510、神经网络解码器515和所存储的3D对象135。3D对象的潜在表示510可以通过使用神经网络编码器505来编码所存储的3D对象135所生成的3D对象的潜在表示510。

[0133] 神经网络解码器515的神经网络可以包括与被应用于网格以便编码该网格的一个或多个卷积和/或滤波器相关联的系数。每个卷积可以具有C个滤波器、K×K掩码(K表示卷积)和步长因子。系数可以对应于C个滤波器、KxK掩码和步长因子中的一个或多个。例如, KxK掩码可以是3x3掩码。3x3掩码包含用于在网格上执行卷积的九(9)个变量。换句话说, 3x3掩码具有九(9)个块,每个块都包含变量。每个变量可以分别是系数之一。

[0134] 神经网络解码器515的神经网络可以包括具有不同数量的神经元的层。KxK空间范围可以包括K列和K(或L)行。KxK空间范围可以是2x2、3x3、4x4、5x5、(KxL) 2x4等。卷积包括将KxK空间范围以网格点为中心,并且在空间范围内对所有网格点进行卷积,以及基于空间范围中的所有网格点的所有卷积(例如,总和),生成该网格点的新值。然后基于上采样,将空间范围移动到新网格点,并且对新网格点重复卷积。上采样因子可以为例如一(1)或二(2),其中,步长因子为1会移动到下一个网格点,而步长因子为2会跳过网格点。

[0135] 卷积可以包括一个或多个补零的卷积运算和系数重组。在示例实施方式中,补零包括在非零像素之间填充零,并且系数重组可以包括具有以零填充像素为中心旋转180度的KxK掩码的卷积。

[0136] VAE可以将用于3D对象的潜在表示510的潜在空间中的变量用作神经网络解码器515的输入,并且对网格的每个点再生位置坐标。神经网络的C个滤波器、KxK掩码和步长因子的配置可以确定再生的网格点的数量。

[0137] 因此,神经网络解码器515可以被配置为使用用于3D对象的潜在表示510中的相应一个的变量来再现在由神经网络编码器505压缩3D对象之前的3D对象的形状和/或模型的近似。神经网络解码器515可以被配置为生成对应于所存储的3D对象135之一的3D模型的近似,包括生成在具有与由神经网络编码器505压缩之前的所存储的3D对象135之一相同数量

的顶点和连通性的网格中的点的位置坐标。

[0138] 根据示例实施方式,可以由具有相同数量的点的网格定义每个所存储的3D对象135,每个点具有位置坐标。此外,VAE可以包括神经网络编码器505和神经网络解码器515,均包括具有用于该神经网络的相同配置的C个滤波器、KxK掩码和步长因子的神经网络。结果,用于由神经网络编码器505生成的、用于3D对象的潜在表示510中的每一个的潜在空间中的变量数量是相同的。此外,由神经网络解码器515再生的网格中的点数相同。

[0139] 在上述示例实施方式中,所存储的3D对象135可以包括CGI电影的感兴趣3D对象。因此,所存储的3D对象135很可能包括3D模型或网格数据以及用于大量3D角色的相应属性值。如上所述,3D模型可以表示整个感兴趣3D对象和/或感兴趣3D对象的一部分。可以使用标准3D模型来表示整个感兴趣3D对象和/或感兴趣3D对象的一部分。

[0140] 在该示例实施方式中,所存储的3D对象135对应于对CGI电影感兴趣的3D对象(例如,作为视频数据5),并且可以不包括CGI电影中未包含的任何其他3D对象。通常,监督机器学习方法机器学习一个或多个规则或功能,以根据操作员所预定的、在示例输入和期望输出之间进行映射。监督机器学习方法和半监督学习方法可以使用标记的数据集。因此,机器训练或训练神经网络编码器505和神经网络解码器515可以使用将所存储的3D对象135(例如,每个3D对象用唯一ID和/或标签标记)用作输入数据和比较数据的监督机器学习方法或半监督学习方法完成。

[0141] 在训练过程期间使用的许多训练迭代可以在重构网格的精度上产生近似对数增益。因此,在半监督学习方法中,可以使用基于时间、误差(例如,损失)和/或迭代次数的阈值来停止进一步训练。例如,可以基于重构网格的点数和位置坐标,将阈值设置为重构误差或损失。在示例实施方式中,重构误差可以是基于形状重构损失(Lr)和正则化先验损失(Lp)的损失。

[0142] 形状重构损失(Lr)可以基于在神经网络编码器505和神经网络解码器515之间获得接近恒同变换。形状重构损失(Lr)可以基于在图5B中生成的所存储的3D对象135和图5A的所存储的3D对象135之间的逐点距离来计算。正则化先验损失(Lp)可以在潜在变量上使用基于以图5A的所存储的3D对象135为基础的矢量的变化或先前迭代分布的散度算法来计算,并且散度算法进一步基于潜在变量的分布。VAE损失或总损失变为 $L=L_r+\lambda L_p$,其中, $\lambda \geq 0$,并且控制变化分布与先前变化分布的相似性。

[0143] 在示例实施方式中,训练过程包括通过依次迭代通过如图5A所示的用于编码3D对象的信号流和如图5B所示的用于解码3D对象的信号流来训练神经网络编码器505和神经网络解码器515。继续上述示例,所存储的3D对象135对应于CGI电影的感兴趣3D对象。在每次迭代(例如,编码和解码每个所存储的3D对象135)之后,将在图5B中生成的所存储的3D对象135(例如,作为重构的3D对象)与图5A的所存储的3D对象135进行比较。在示例实施方式中,重构误差可以基于以形状重构损失(Lr)和正则化先验损失(Lp)为基础的损失来计算。如果重构误差高于阈值(如上所述),则可以修正用于系数的变量,该系数对应于用于神经网络编码器505和神经网络解码器515中的至少一个的C个滤波器、KxK掩码和步长因子中的一个或多个。此外,可以修正与计算VAE损失相关的变量 λ 。在示例实施方式中,变量可以基于用于前一迭代的变量值和结果以及用于当前修正的变量值和结果来修正。在变量修正后,可以开始下一个训练迭代。

[0144] 在示例实施方式中,流传输和/或下载视频可以包括在传送或流传输视频的压缩帧之前,传送初始化数据以供请求客户端设备使用(例如,用于解压缩压缩帧)。图6A示出了用于流传输视频并在客户端设备上渲染视频的信号流的框图。如图6A所示,流传输服务器605-1包括视频存储125、主动流传输模块610、流传输初始化模块615和收发器620。视频存储125包括压缩帧140、元数据130和用于3D对象的潜在表示510。流传输服务器605-1可以经由收发器625通信地耦合到客户端设备650。如图6A所示,客户端设备650包括收发器625、解码器145、渲染器630、颜色校正模块635和显示器640。

[0145] 主动流传输模块610可以被配置为将所选择的视频的帧流式传输到请求客户端设备(例如,客户端设备650)。主动流传输模块610可以从客户端设备650接收对下一帧的请求。然后,主动流传输模块610可以从压缩帧140中选择下一帧,并且从元数据130中选择相应的元数据。在示例实施方式中,所选择的下一帧可以是多个压缩帧,并且所选择的元数据可以是多个相应的元数据元素。例如,多个压缩帧可以是由(并且可能包括)前一(例如,时间上,在帧之前的)关键帧和未来(例如,时间上,在帧之后的)关键帧绑定的多个帧。

[0146] 然后,主动流传输模块610可以将所选择的下一帧(或多个帧)和所选择的元数据(或元数据元素)传送到收发器620。然后,收发器620可以构建包括所选择的下一帧和所选择的元数据的一个或多个数据包,将客户端设备650的地址分配给数据包,并且将数据包传送到客户端设备650。

[0147] 流传输初始化模块615可以被配置为响应于客户端设备650对视频的第一请求来选择与流视频相关联的数据。例如,客户端设备650的用户可以下载该视频以供未来回放和/或开始流式回放。流传输初始化模块615可以将用于视频的所存储的3D对象135选择为用于3D对象的潜在表示510的集合并且从元数据130中选择对应于该视频的元数据。然后,流传输初始化模块615可以将所选择的所存储的3D对象135和所选择的元数据传送到收发器620。然后,收发器620可以构建包括所选择的所存储的3D对象135和所选择的元数据的数据包,将客户端设备650的地址分配给该数据包,并且将该数据包传送到客户端设备650。在示例实施方式中,收发器620可以构建一个以上的数据包。

[0148] 客户端设备650经由收发器625接收包括所存储的3D对象135和元数据的数据包。收发器625将所存储的3D对象135和元数据传送到解码器145。解码器145可以被配置为与所请求的视频相关联地存储所存储的3D对象135和元数据。然后,客户端设备650经由收发器625接收包括所选择的下一帧和所选择的元数据的数据包。所选择的下一帧可以是初始回放时的第一帧。收发器625将所选择的下一帧和所选择的元数据传送到解码器145。

[0149] 然后,解码器145使用与下一帧相关联的所存储的3D对象135和元数据来解码(如上更详细所述)下一帧。在示例实施方式中,解码器145可以被实现为包括图形处理单元(GPU)的图形卡和/或芯片(例如,计算机主板上的ASIC)或在包括图形处理单元(GPU)的图形卡和/或芯片(例如,计算机主板上的ASIC)中实现,被配置为移除来自设备650的中央处理单元(CPU)的负担。GPU可以被配置为并行地处理大块视频数据。可以将GPU配置为处理(例如,解压缩)网格数据并由网格数据生成像素数据。解码器将纹理数据和颜色数据传送到渲染器630。

[0150] 在3D渲染系统中,通常使用笛卡尔坐标系(x-(左/右),y-(上/下)和z轴(近/远))。笛卡尔坐标系提供了一种精确的数学方法来定位和表示空间中的对象。为简化起见,可以

将帧的图像视为对象将处于的空间。可以基于相机位置来定向空间。例如,可以将相机放置 在原点并直视z轴。因此,(相对于相机位置的) 平移运动为z向前/向后、y向上/向下和x向 左/向右。然后,基于对象相对于空间原点的坐标以及任何相机的重新定位,将对象投影到 空间中。注意对象和/或相机可以逐帧移动。

[0151] 渲染器630可以被配置为渲染纹理数据和颜色数据以供显示。可以在几何阶段和 渲染阶段的图形管线中实现渲染或绘图。可以将GPU配置为处理几何阶段和渲染阶段。由 CPU或GPU执行的几何阶段被配置为处理所有多边形活动并且将3D空间数据转换为像素。几 何阶段过程可以包括但不限于场景(例如背景)几何生成、基于相机运动的对象运动、对象 运动、对象变换(例如旋转、平移和/或缩放)、对象可见性(例如,遮挡和剔除)和多边形(例 如三角形)的生成。

[0152] 由GPU的3D硬件加速器执行的渲染阶段被配置为管理内存和像素活动,并且处理 要绘制到显示器640的像素。渲染阶段过程可以包括但不限于阴影、纹理、深度缓冲和显示。

[0153] 渲染纹理数据和颜色数据以供显示可以包括使用3D着色器(例如,顶点着色器、几 何着色器等)来绘制与帧相关联的网格。可以将着色器配置为生成图元。可以将着色器配置 为将网格中的每个顶点的3D位置和纹理变换为出现在显示器(例如,显示器640)上的2D坐 标(例如,图元)。渲染纹理数据和颜色数据以供显示还可以包括执行光栅化。光栅化可以包 括基于纹理数据和颜色数据,对图元使用分配像素(例如,颜色)值。

[0154] 如上所述,几何阶段包括在2D坐标系中构建3D网格,而渲染阶段包括向网格添加 颜色和纹理。因此,解码器145的输出可以包括对象位置数据和颜色/纹理数据。因此,压缩 帧140可以包括压缩对象位置数据和压缩颜色/纹理数据。压缩对象位置数据和压缩颜色/ 纹理数据可以基于与前一帧或下一帧的位移、基于相机平移运动数据、基于绝对数据(例 如,x,y,z坐标、RGB值和/等)和/或其组合或变化。因此,编码器105可以对帧的场景进行编 码以生成包括具有该压缩对象位置数据和压缩颜色/纹理数据的数据的压缩帧140。

[0155] 与所渲染的下一帧相关联的数据可以被传送到颜色校正模块635。颜色校正模块 635可以被配置为对与所渲染的下一帧相关联的数据进行颜色校正。颜色校正可以包括补 偿帧之间的色差、补偿同一场景的多个视图之间的色差、校正对象变形(扭曲)、校正对象边 界变形等。

[0156] 颜色校正模块635可以将颜色校正帧传送到显示器640。显示器640可以被配置为 将帧显示为表示所请求的视频的帧序列。在示例实施方式中,显示器640包括和/或具有相 关联的缓冲器和/或队列。缓冲器和/或队列可以是先入/先出缓冲器和/或队列。因此,颜色 校正模块635可以将颜色校正帧传送到缓冲器和/或队列。显示器640包括被配置为从缓冲 器和/或队列中选择下一帧的组件。

[0157] 在示例实施方式中,流传输和/或下载视频可以包括与传送或流传输视频的压缩 帧并行地(例如,几乎同时)传送数据供请求客户端设备使用(例如,用于解压缩压缩帧)。图 6B示出了用于流传输视频并在客户端设备上渲染视频的另一信号流的框图。如图6B所示, 流传输服务器605-2包括视频存储125、主动流传输模块610和收发器620。

[0158] 在图6B所示的示例实施方式中,主动流传输模块610被进一步配置为从用于3D对 象的潜在表示510中选择用在压缩所选择的下一帧中的感兴趣的至少一个3D对象作为所存 储的3D对象135中的至少一个。主动流传输模块610将所存储的3D对象135中的至少一个与

所选择的下一帧和所选择的元数据一起传送到收发器620。然后,收发器620可以构建包括所选择的下一帧、所选择的元数据和所存储的3D对象135中的至少之一的数据包,将客户端设备650的地址分配给该数据包,并且将该数据包传送到客户端设备650。

[0159] 然后,客户端设备650经由收发器625接收包括所选择的下一帧、所选择的元数据和所存储的3D对象135中的至少一个的数据包。收发器625将所选择的下一帧、所选择的元数据和所存储的3D对象135中的至少一个传送到解码器145。解码器145可以进一步被配置为将所存储的3D对象135中的至少一个添加到与所请求的视频相关联的所存储的3D对象135和/或使用所存储的3D对象135中的至少一个来初始化用于所请求的视频的所存储的3D对象135。

[0160] 此外,尽管当编码/解码视频中的具有从帧到帧的动态非线性和/或随机运动的3D对象(有时在本文中称为动态3D对象)时,通过本文所述的几何代理技术的颜色预测可能是最有利的,但其他3D对象或非动态3D对象也可以通过几何代理技术,使用颜色预测来有利地进行编码/解码。

[0161] 例如,可以通过几何代理,使用颜色预测来编码/解码由于相机或场景平移而看似从帧到帧移动的3D对象。在该示例中,由于捕获场景的相机正在(例如,以可预测的方式和/或方向)移动,所以静止的3D对象可能看似正在从帧到帧移动。例如,可以通过几何代理,使用颜色预测来编码/解码看似在场景内以可预测方式移动的3D对象。在该示例中,3D对象(例如,车辆,即火车、汽车或飞机)可以以可预测的方式(例如,以恒定的速度和/或方向)从帧到帧移动。这些对象有时在本文中被称为平移3D对象。作为另一示例,可以通过几何代理,使用颜色预测来编码/解码看似在场景内未移动的3D对象。在该示例中,静止或固定的3D对象(例如,场景的背景、场景内固定位置处的家具或远处缓慢移动的对象)可能从帧到帧看起来是静止的(例如,没有任何相机或场景平移)。这些对象有时在本文中被称为固定或背景3D对象。

[0162] 出于多种原因,可以通过几何代理技术,使用颜色预测来有利地编码/解码上文例示的其他类型的3D对象。例如(这不旨在是详尽的列表),可以将平移3D对象、固定3D对象和/或背景3D对象的至少一个位置(例如,帧中的位置)从流传输服务器(例如,流传输服务器605-1、605-2)传送到要在渲染器和/或渲染操作中使用的客户端设备(例如,客户端设备650);由于存在几何代理(例如所存储的3D对象135),可以增加关键帧之间的帧数;Z阶分层技术可以被用于动态3D对象和非动态3D对象;可以使用(例如在被渲染前在队列中存储的)时间上较早和时间上较晚的帧和几何代理,重新创建丢失(例如,未重传)的帧;可以使用几何代理来编码/解码关键帧之间出现和消失的3D对象和/或可以使用几何代理来编码/解码帧外背景相机或场景平移。

[0163] 因此,上述的元数据20可以用在渲染器630中以更有效地渲染平移3D对象、固定3D对象和/或背景3D对象。在示例实施方式中,第一对象可以被标识为背景3D对象。元数据20可以将第一对象标识为例如所存储的3D对象135中的一个。元数据20还可以将(例如,作为该帧的背景的)第一对象的原点坐标(例如 x_0, y_0, z_0)标识为表示第一对象的网格中的点之一的位置属性。

[0164] 此外,可以将至少一个第二对象标识为平移3D对象和/或固定3D对象。元数据20可以将至少一个第二对象标识为例如所存储的3D对象135之一。元数据20还可以将至少一个

第二对象的位置(例如, x_n, y_n, z_n)标识为表示至少一个第二对象的网格中的点之一的位置属性。至少一个第二对象的所标识的位置可以相对于另一对象。至少一个第二对象的相对位置可以从帧到帧变化(例如,作为平移3D对象)。至少一个第二对象的相对位置可以从帧到帧是固定的(例如,作为固定的3D对象)。

[0165] 例如,至少一个第二对象的所标识的位置可以相对于用于该帧的背景(例如,第一对象)的所标识的原点坐标。将相对于该帧的背景定位至少一个第二对象时,将相对位置用于至少一个第二对象可以允许六个自由度。换句话说,当相对于该帧的背景定位至少一个第二对象时,至少一个第二对象可以具有平移运动(例如,向前/向后、向上/向下和/或向左/向右)以及旋转运动(例如,俯仰、滚动和/或偏航)。

[0166] 如上所述,第一对象可以被标识为背景3D对象,并且可以是所存储的3D对象135中的一个。此外,至少一个第二对象可以被标识为平移3D对象和/或固定3D对象,其也可以是所存储的3D对象135中的一个。因此,第一对象和/或至少一个第二对象可以通过自动编码器(实现神经网络)编码为相应存储的3D对象135的潜在表示(如上更详细所述)。与神经网络的结构有关的潜在表示(例如,作为用于3D对象的一个或多个潜在表示510)和信息(例如,元数据)可以从流传输服务器605-1、605-2传送到客户端650。客户端650可以使用自动编码器的解码器(实现神经网络)来重构一个或多个相应的所存储的3D对象135(如上更详细地讨论)。

[0167] 图7至图11是根据示例实施例的方法的流程图。参考图7至图11描述的步骤可以由执行在与装置(例如,如图12和13所示)相关联的存储器(例如,至少一个存储器1210、1310)中存储的软件代码来实现或由与装置相关联的至少一个处理器(例如,至少一个处理器1205、1305)执行。

[0168] 然而,可以设想替代实施例,诸如体现为专用处理器的系统。专用处理器可以是图形处理单元(GPU)。GPU可以是图形卡的组件。图形卡还可以包括视频存储器、随机存取存储器数模转换器(RAMDAC)和驱动器软件。视频存储器可以是帧缓冲器,其存储代表图像或帧的场景的数字数据。可以将RAMDAC配置为读取视频存储器的内容,将内容转换为模拟RGB信号,然后将模拟信号发送到显示器或监视器。驱动器软件可以是存储在上述存储器(例如,至少一个存储器1210、1310)中的软件代码。可以将软件代码配置为实施下文所述的步骤(和/或上文所述的组件、模块和信号流)。

[0169] 尽管下文描述的步骤被描述为由处理器和/或专用处理器执行,但是这些步骤不一定由同一处理器执行。换句话说,至少一个处理器和/或至少一个专用处理器可以执行下文参考图7至图11所述的步骤。

[0170] 图7示出了根据至少一个示例实施例的用于压缩视频的帧的方法的框图。如图7所示,在步骤S705中,接收包括视频的多个帧的文件。例如,可以将文件保存或传送到服务器(例如,流传输服务器)。该文件可以包括视频。该视频可以是CGI 3D电影。该文件可以包括感兴趣的多个3D对象(例如,3D电影中的角色)。

[0171] 在示例实施方式中,可以通过三角形网格来定义感兴趣的多个3D对象中的每一个。三角形网格可以是由三角形面连接的点的集合。每个点可以存储各种属性。例如,属性可以包括每个点的位置、颜色、纹理坐标等。感兴趣的多个3D对象中的每一个的网格可以具有相同数量的点,每个点具有相同属性。因此,当存储在存储器中时,感兴趣的多个3D对象

中的每一个的网格可以为近似相同大小(例如,位数)。

[0172] 在步骤S710中,选择多个帧之一。例如,可以将多个帧中的每个帧作为压缩目标。可以按时间顺序压缩多个帧。因此,在初始步骤中,选择时间上的第一帧。然后可以顺序地选择下一帧。

[0173] 在步骤S715中,在所选择的帧和关键帧中标识3D对象。所标识的3D对象可以是动态3D对象、非动态3D对象、固定3D对象、背景3D对象等。例如,机器视觉、计算机视觉和/或计算机图像识别技术可以被用来标识和定位3D对象。

[0174] 在示例实施方式中,基于使用多个已知图像训练(机器学习)卷积神经网络的计算机图像识别技术可以被用来标识3D对象。例如,从所选择的帧中选择和/或标识块、多个块和/或补片。经训练的卷积神经网络可以对所选块、多个块和/或补片进行操作。可以测试结果(例如,误差测试、损失测试、散度测试等)。如果测试产生低于阈值的值(或替代地,取决于测试类型而产生高于阈值的值),则可以将所选块、多个块和/或补片标识为3D对象。

[0175] 在示例实施方式中,视频的帧可以包括指示先前标识的感兴趣的3D对象被包括在帧中的标签。标签可以包括3D对象的身份和定位。例如,可以使用计算机生成图像(CGI)工具(例如计算机动画电影)来生成视频。可以在每帧中标识和标记计算机生成的角色。此外,可以将所标识的感兴趣的3D对象中的每一个(例如,所标识的角色)的模型存储为所存储的3D对象135。

[0176] 在示例实施方式中,可以由三角形网格定义3D对象。三角形网格可以是由三角形面连接的点的集合。每个点可以存储各种属性。例如,属性可以包括每个点的位置、颜色、纹理坐标等。属性可以包括和/或指示(例如,多个属性可以指示)相应的3D对象的定向和/或相应的3D对象在所选择的帧中的位置。

[0177] 因此,在示例实施方式中,3D对象的网格属性可以足以标识和定位3D对象。可以将包括用于感兴趣的多个3D对象的网格属性的模型存储为所存储的3D对象135。该模型可以被标准化。例如,可以将用于男人、女人、青少年、儿童或更一般地人类或人类的一部分(例如,身体、头部、手等)的模型存储为所存储的3D对象135。例如,可以将用于狗、猫、鹿或更一般的四足动物或四足动物的一部分(例如,身体、头部、腿等)的模型存储为所存储的3D对象135。然后,可以使用模型的属性在帧中搜索具有相似属性的3D对象。

[0178] 在步骤S720中,确定3D对象在所选择的帧和关键帧内的位置和方向。例如,可以通过具有多个点的网格来定义3D对象,每个点具有至少一个属性。3D对象的位置可以是3D对象在帧内的位置。因此,3D对象在帧内的位置可以基于用于该帧的3D笛卡尔坐标系。例如,网格的至少一个点可以位于帧内的x,y,z位置。3D对象的定向可以基于用于定义3D对象的网格中的每个点的位置坐标属性。如果3D对象是移动中的角色,则3D对象的定向可以是角色在帧和/或关键帧中的姿势。

[0179] 在步骤S725中,将3D对象与所存储的3D对象进行匹配。例如,可以将该3D对象与感兴趣的3D对象(例如,所存储的3D对象135)之一匹配。在示例实施方式中,可以使用如由计算机图像识别技术生成的3D对象的身份来搜索所存储的3D对象135。在示例实施方式中,可以将帧中找到的3D对象的标签与所存储的3D对象135之一的标签匹配。在示例实施方式中,所存储的3D对象135之一与3D对象具有相似属性的模型可以被标识为匹配。经匹配的所存储的3D对象可以被用作几何代理。

[0180] 在步骤S730中,基于3D对象来变换所存储的3D对象。例如,可以基于3D对象使所存储的3D对象变形。可以基于3D对象调整所存储的3D对象的大小。可以基于3D对象来定向所存储的3D对象。可以基于3D对象旋转所存储的3D对象。可以基于3D对象平移所存储的3D对象。变换所存储的3D对象可以导致所存储的3D对象(如果在显示上与3D对象并排地渲染)看起来与3D对象基本相似(例如,从姿势上)。注意,所存储的3D对象可以由不同的网格定义,并且可以具有与3D对象不同的颜色属性。

[0181] 在示例实施方式中,与所存储的3D对象相对应的网格可以绕x轴、y轴和z轴并且沿着x轴、y轴和z轴旋转和平移,以与3D对象的定向对齐。可以变换与帧和关键帧相关联的所存储的3D对象135中的所匹配的一个。因此,所存储的3D对象可以用于基于与帧相关联的针对3D对象的所存储的3D对象来生成第一3D对象代理。此外,所存储的3D对象可以用于基于与关键帧相关联的针对3D对象的所存储的3D对象来生成第二3D对象代理。因此,可以基于在帧中标识的3D对象来变换第一3D对象代理,并且可以基于在关键帧中标识的3D对象来变换第二3D对象代理。

[0182] 在步骤S735中,使用经变换的所存储的3D对象,通过颜色预测方案压缩所选择的帧。举例来说,定义经变换的所存储的3D对象的网格中的点可以与位于附近(时间上)的、先前编码的关键帧中的相应(例如,同一)3D对象的点匹配。然后,预测技术可以使用相应的3D对象的匹配点来选择或预测关键帧中的像素/块/补片,以用于计算所选择的帧中的3D对象的残差(例如,相对于关键帧的颜色位移)。

[0183] 在示例实施方式中,可以与在帧中标识的3D对象相对应(用颜色和/或纹理)包装经变换的第一3D对象代理,并且可以与关键帧中标识的3D对象相对应(用颜色和/或纹理)包装经变换的第二3D对象代理。将颜色属性从在帧110中标识的3D对象映射到经变换的第一3D对象代理可以包括将3D对象从3D空间(例如XYZ空间)转换到2D空间(例如UV空间)和/或将在帧110中标识的3D对象从2D空间(例如,UV空间)转换到3D空间(例如,XYZ空间)。可以通过从第二3D对象代理的网格表示中的相应点(例如,具有相同点标识或处于网格序列中的相同位置)的颜色属性中减去第一3D对象代理的网格表示中的点的颜色属性来生成用于3D对象的残差。

[0184] 在示例实施方式中,在所选择的帧中标识两个或更多个3D对象。使用上述预测技术来预测每个所标识的3D对象。帧的剩余部分使用标准预测技术来生成残差。压缩所选择的帧可以包括对残差执行一系列编码处理。例如,可以变换、量化和熵编码残差。

[0185] 另外,预测方案可以生成与所选择的帧相关联的元数据。元数据可以包括已经使用感兴趣的3D对象之一(例如,所存储的3D对象135)预测的、与位于所选择的帧中的至少一个3D对象相关联的数据。元数据可以包括3D对象在所选择的帧和关键帧中的定位和/或定向相关联的属性(例如,网格点属性)。

[0186] 在步骤S740中,将经压缩的所选择的帧与标识3D对象以及3D对象的位置和定向的元数据一起存储。例如,经压缩的所选择的帧和元数据可以存储在与流传输服务器相关联的存储器中。可以与对应于视频(或其一部分)的多个压缩帧有关地存储经压缩的所选择的帧和元数据。

[0187] 图8示出了根据至少一个示例实施例的用于压缩视频的帧的另一种方法的框图。如图8所示,在步骤S805中,接收包括视频的多个帧的文件。例如,可以将文件保存或传送到

服务器(例如,流传输服务器)。该文件可以包括视频。该视频可以是CGI 3D电影。该文件可以包括感兴趣的多个3D对象(例如,3D电影中的角色)。

[0188] 在示例实施方式中,可以通过三角形网格来定义感兴趣的多个3D对象中的每一个。三角形网格可以是由三角形面连接的点的集合。每个点可以存储各种属性。例如,属性可以包括每个点的位置、颜色、纹理坐标等。感兴趣的多个3D对象中的每一个的网格可以具有相同数量的点,每个点具有相同属性。因此,当存储在存储器中时,用于感兴趣的多个3D对象中的每一个的网格可以具有近似相同的大小(例如,位数)。

[0189] 在示例实施方式中,可以压缩(例如,使用上文参考图5A所述的技术压缩)用于感兴趣的多个3D对象中的每一个的网格。例如,可以使用生成建模技术(例如,使用神经网络、卷积神经网络、VAE等)压缩用于感兴趣的多个3D对象中的每一个的网格。可以使用具有卷积神经网络的神经网络编码器来压缩用于感兴趣的多个3D对象中的每一个的网格,其中,卷积神经网络具有基于机器训练的生成建模技术选择和训练的元素,该机器训练的生成建模技术被配置为生成与用于每个3D对象的网格属性和位置相关联的减少数量的变量。与用于3D对象的网格属性和位置相关联的所生成的数量减少的变量有时被称为用于3D对象的紧凑潜在表示。

[0190] 在步骤S810中,选择多个帧之一。例如,可以将多个帧中的每个帧作为压缩目标。可以按时间顺序压缩多个帧。因此,在初始步骤中,选择时间上的第一帧。然后可以按顺序选择下一帧。

[0191] 在步骤S815中,在所选择的帧和关键帧中标识3D对象。所标识的3D对象可以是动态3D对象、非动态3D对象、静止3D对象、背景3D对象等。例如,机器视觉、计算机视觉和/或计算机图像识别技术可以被用来标识和定位3D对象。

[0192] 在示例实施方式中,基于使用多个已知图像训练(机器学习)卷积神经网络的计算机图像识别技术可以被用来标识3D对象。例如,从所选择的帧中选择和/或标识块、多个块和/或补片。经训练的卷积神经网络可以对所选块、多个块和/或补片进行操作。可以测试结果(例如,误差测试、损失测试、散度测试等)。如果测试产生低于阈值的值(或替代地,取决于测试类型而产生高于阈值的值),则可以将所选块、多个块和/或补片标识为3D对象。

[0193] 在示例实施方式中,视频的帧可以包括指示先前标识的感兴趣的3D对象被包括在帧中的标签。标签可以包括3D对象的身份和定位。例如,可以使用计算机生成图像(CGI)工具(例如计算机动画电影)来生成视频。可以在每帧中标识和标记计算机生成的角色。此外,可以将用于每个所标识的感兴趣的3D对象(例如,所标识的角色)的模型存储为所存储的3D对象135。

[0194] 在示例实施方式中,可以由三角形网格定义3D对象。三角形网格可以是由三角形面连接的点的集合。每个点可以存储各种属性。例如,属性可以包括每个点的位置、颜色、纹理坐标等。属性可以包括和/或指示(例如,多个属性可以指示)相应的3D对象的定向和/或相应的3D对象在所选择的帧中的位置。

[0195] 因此,在示例实施方式中,3D对象的网格属性可以足以标识和定位3D对象。可以将包括用于感兴趣的多个3D对象的网格属性的模型存储为所存储的3D对象135。该模型可以被标准化。例如,可以将用于男人、女人、青少年、儿童或更一般地人类或人类的一部分(例如,身体、头部、手等)的模型存储为所存储的3D对象135。例如,可以将用于狗、猫、鹿或更一

般的四足动物或四足动物的一部分(例如,身体、头部、腿等)的模型存储为所存储的3D对象135。然后,可以使用模型的属性在帧中搜索具有相似属性的3D对象。

[0196] 在步骤S820中,确定3D对象在所选择的帧和关键帧内的位置和定向。例如,可以通过具有多个点的网格来定义3D对象,每个点具有至少一个属性。3D对象的位置可以是3D对象在帧内的位置。因此,3D对象在帧内的位置可以基于用于该帧的3D笛卡尔坐标系。例如,网格的至少一个点可以位于帧内的x,y,z位置。3D对象的定向可以基于用于定义3D对象的网格中的每个点的位置坐标属性。如果3D对象是移动中的角色,则3D对象的定向可以是角色在帧和/或关键帧中的姿势。

[0197] 在步骤S825中,将3D对象与用于3D对象的潜在表示进行匹配。例如,该3D对象可以与感兴趣的3D对象之一(例如,所存储的3D对象135)进行匹配。然后,如上所述,可以将所匹配的3D对象编码为用于3D对象的潜在表示。例如,可以将该3D对象与感兴趣的3D对象(例如,用于所存储的3D对象的潜在表示510)之一进行匹配。在示例实施方式中,可以使用如由计算机图像识别技术生成的3D对象的身份来搜索所存储的3D对象的潜在表示510。在示例实施方式中,可以将帧中找到的3D对象的标签与所存储的3D对象的潜在表示510之一的标签进行匹配。在示例实施方式中,所存储的3D对象的潜在表示510之一与3D对象的压缩网格属性具有相似属性的模型可以被标识为匹配。

[0198] 在步骤S830中,基于3D对象变换潜在表示。例如,可以基于3D对象使所存储的3D对象变形。可以基于3D对象调整所存储的3D对象的大小。可以基于3D对象来定向所存储的3D对象。可以基于3D对象旋转所存储的3D对象。可以基于3D对象平移所存储的3D对象。变换所存储的3D对象可以导致所存储的3D对象(如果在显示上与3D对象并排地渲染)看起来与3D对象基本相似。注意,所存储的3D对象可以由不同的网格定义,并且可以具有与3D对象不同的颜色属性。

[0199] 在示例实施方式中,可以绕x轴、y轴和z轴并且沿着x轴、y轴和z轴旋转和平移3D对象的潜在表示的每一点的坐标,以与3D对象的定向对齐。可以变换与帧和关键帧相关联的所匹配的3D对象的潜在表示。因此,可以基于与该帧相关联的针对3D对象的所存储的3D对象而将3D对象的潜在表示生成第一3D对象代理。此外,可以基于与关键帧相关联的针对3D对象的所存储的3D对象而将3D对象的潜在表示生成第二3D对象代理。因此,可以基于在帧中标识的3D对象来变换第一3D对象代理,并且可以基于在关键帧中标识的3D对象来变换第二3D对象代理。

[0200] 在示例实施方式中,将所存储的3D对象存储为基于3D对象的紧凑潜在表示。因此,在变换所存储的3D对象之前,可以使用基于3D对象的紧凑潜在表示来再生(例如,使用上文参考图5B所述的技术解压缩)所存储的3D对象。例如,可以将用于所存储的3D对象的紧凑潜在空间中的变量输入到神经网络解码器,以再生定义所存储的3D对象的网格的点以及用于网格的每个点的位置坐标。

[0201] 在步骤S835中,使用具有用于3D对象的紧凑潜在表示的经变换的所存储的3D对象,用颜色预测方案来压缩所选择的帧。如上所述,具有用于3D对象的紧凑潜在表示的经变换的所存储的3D对象可以是经变换的再生的所存储的3D对象。例如,定义经变换的再生的所存储的3D对象的网格中的点可以与位于附近(时间上)、先前编码的关键帧中的相应(例如,相同)3D对象的点匹配。然后,预测技术可以使用相应的3D对象的匹配点来选择或预测

关键帧中的像素/块/补片,以用于计算所选择的帧中的3D对象的残差(例如,相对于关键帧的颜色位移)。

[0202] 在示例实施方式中,可以相应于在帧中标识的3D对象(用颜色和/或纹理)包装经变换的第一3D对象代理,并且可以相应于关键帧中标识的3D对象(用颜色和/或纹理)包装经变换的第二3D对象代理。将颜色属性从在帧110中标识的3D对象映射到经变换的第一3D对象代理可以包括将3D对象从3D空间(例如XYZ空间)转换到2D空间(例如UV空间)和/或将在帧110中标识的3D对象从2D空间(例如,UV空间)转换到3D空间(例如,XYZ空间)。在示例实施方式中,可以通过从第二3D对象代理的潜在表示中的相应点(例如,具有相同坐标或处于相同的位置)的颜色属性减去第一3D对象代理的潜在表示中的点的颜色属性来生成用于3D对象的残差。

[0203] 在示例实施方式中,可以解码用于经变换的包装的第一3D对象代理的潜在表示和经变换的包装的第二3D对象代理的潜在表示。如上所述,解码3D对象的潜在表示可以再生包括3D对象的颜色属性的网格表示。可以通过从第二3D对象代理的再生网格表示中的相应点(例如,具有相同点标识或处于网格序列中的相同位置)的颜色属性中减去第一3D对象代理的再生网格表示中的点的颜色属性来生成用于3D对象的残差。

[0204] 在示例实施方式中,在所选择的帧中标识两个或更多个3D对象。使用上述预测技术来预测每个所标识的3D对象。帧的剩余部分使用标准预测技术来生成残差。压缩所选择的帧可以包括对残差执行一系列编码处理。例如,可以变换、量化和熵编码残差。

[0205] 另外,预测方案可以生成与所选择的帧相关联的元数据。元数据可以包括已经使用感兴趣的3D对象之一(例如,所存储的3D对象135)预测的、与位于所选择的帧中的至少一个3D对象相关联的数据。元数据可以包括3D对象在所选择的帧和关键帧中的定位和/或定向相关联的属性(例如,网格点属性)。元数据可以包括与用于自动编码器的神经网络相关联的信息,该自动编码器用于生成3D对象的潜在表示(例如,编码)以及再生包括3D对象的颜色属性的网格表示(例如,解码)。

[0206] 在步骤S840中,将经压缩的所选择的帧与标识3D对象以及3D对象的位置和定向的元数据一起存储。例如,经压缩的所选择的帧和元数据可以存储在与流传输服务器相关联的存储器中。可以与对应于视频(或其一部分)的多个压缩帧有关地存储经压缩的所选择的帧和元数据。

[0207] 图9示出了根据至少一个示例实施例的用于解压缩和渲染视频的帧的方法的框图。如图9所示,在步骤S905中,接收包视频的至少一个压缩帧的数据包。例如,客户端设备可以从流传输服务器请求视频的下一帧。响应于该请求,流传输服务器可以选择下一帧(或多个下一帧),确定所选择的下一帧(或多个下一帧)是否具有相关联的元数据并选择相关联的元数据。流传输服务器可以生成包括所选择的下一帧(或多个下一帧)和所选择的相关联的元数据的数据包(或多个数据包),并且将该数据包(或多个数据包)传送给请求客户端设备。

[0208] 在步骤S910中,选择至少一帧进行解压缩。例如,可以从所接收的数据包(或多个数据包)中选择帧。在步骤S915中,确定帧是否包括元数据。响应于确定帧包括元数据,处理继续到步骤S925。否则,处理继续到步骤S920,并且通过一些其他预测方案(例如,不基于将3D对象用作几何代理的预测方案)解码所选择的帧。

[0209] 例如,当将数据包从流传输服务器传送到客户端时,与用于所选择的帧的数据包相关联的报头可以被配置为包含元数据。因此,与用于所选择的帧的数据包相关联的报头可以(例如,经由收发器625)被读取,并且将其与所选择的帧一起传送到解码器(例如,解码器145)。因此,确定所选择的帧包括元数据可以包括确定元数据已经与所选择的帧一起被传送。替代地,确定所选择的帧包括元数据可以包括确定与所选择的帧相关联地存储元数据(例如,存储在解码器中)。

[0210] 在步骤S925中,基于元数据在所选择的帧中标识3D对象。例如,元数据可以包括已经将3D对象用作几何代理预测的、与位于所选择的帧中的至少一个3D对象(例如,所存储的3D对象135中的至少一个)相关联的数据。元数据可以包括标识用作几何代理的3D对象的信息。

[0211] 在步骤S930中,基于元数据,确定3D对象在所选择的帧和关键帧内的位置和定向。例如,元数据可以包括与3D对象在所选择的帧和关键帧中的定位和/或定向相关联的属性(例如,网格点属性)。

[0212] 在步骤S935中,将3D对象与所存储的3D对象进行匹配。例如,元数据可以包括标识3D对象具有相应的所存储的3D对象的信息。在示例实施方式中,元数据包括可以被用来在所存储的3D对象135中搜索3D对象的唯一ID或标签。如果在所存储的3D对象135中找到唯一ID或标签,则所存储的3D对象135中的相应一个是3D对象的匹配。

[0213] 在步骤S940中,基于元数据变换所存储的3D对象。例如,元数据可以包括标识3D对象具有相应的所存储的3D对象的信息以及与在编码过程期间对所存储的3D对象执行的变换有关的信息。与在编码过程期间对所存储的3D对象执行的变换有关的信息可以被用来对所存储的3D对象执行与在所存储的3D对象上执行的相同变换以编码所选择帧。

[0214] 可以基于与在编码过程期间在所存储的3D对象上执行的变换有关的信息使所存储的3D对象变形。可以基于与在编码过程期间在所存储的3D对象上执行的变换有关的信息来调整所存储的3D对象的大小。可以基于与在编码过程期间在所存储的3D对象上执行的变换有关的信息来定向所存储的3D对象。可以基于与在编码过程期间在所存储的3D对象上执行的变换有关的信息来旋转所存储的3D对象。可以基于与在编码过程期间在所存储的3D对象上执行的变换有关的信息来平移所存储的3D对象。变换所存储的3D对象可以使得所存储的3D对象(如果在显示器上与3D对象并排渲染)以与3D对象基本类似地出现,因为3D对象在压缩所选择的帧之前出现在所选择的帧中。

[0215] 在示例实施方式中,可以变换与帧和关键帧相关联的所存储的3D对象中的所匹配的3D对象。因此,可以使用所存储的3D对象来基于与帧相关联的针对3D对象的所存储的3D对象来生成第一3D对象代理。此外,可以使用所存储的3D对象来基于与关键帧相关联的针对3D对象的所存储的3D对象来生成第二3D对象代理。因此,可以基于在帧中标识的用于3D对象的元数据来变换第一3D对象代理,并且可以基于在关键帧中标识的用于3D对象的元数据来变换第二3D对象代理。

[0216] 在示例实施方式中,第一3D对象代理和第二3D对象代理是3D对象的潜在表示。因此,第一3D对象代理可以被用来再生用于第一3D对象代理的网格表示,并且第二3D对象代理可以被用来再生用于第二3D对象代理的网格表示。如上所述,使用自动编码器解码用于3D对象的潜在表示可以再生包括用于3D对象的颜色属性的网格表示。自动编码器可以使用

具有从元数据读取的结构神经网络。

[0217] 在步骤S945中,使用针对3D对象的所存储的3D对象(作为几何代理),用颜色预测方案,解压缩所选择的帧。初始地,可以熵解码、逆量化和逆变换所选择的帧,以生成与当编码所选择的帧时并且在变换、量化和熵编码残差之前由编码器生成的残差相同(或近似相同)的导数残差。

[0218] 在示例实施方式中,预测方案包括将对应于经变换的(例如,经平移和定向的)所存储的3D对象的点与关键帧的点匹配。然后,使用匹配点来选择或预测关键帧中的像素/块/补片,以用于再生所选择的帧中的3D对象的像素的颜色值。

[0219] 再生像素的颜色值可以包括将来自用于关键帧中的预测像素/块/补片的每个匹配点的像素属性值中的、用于3D对象的三角形网格中的每个点的像素属性值与用于基于所选择的帧中的所标识的位置处的残差的经平移和定向的所存储的3D对象的颜色值和/或颜色属性相加。可以基于所选择的帧的残差或剩余部分以及关键帧的相应像素/块/补片来再生用于所选择的帧的剩余部分的颜色值和/或颜色属性。

[0220] 在示例实施方式中,可以对应于在帧中标识的3D对象(用颜色和/或纹理)包装经变换的第一3D对象代理。换句话说,可以用对应于在帧中标识的3D对象的残差来包装经变换的第一3D对象代理。此外,可以对应于在关键帧中标识的3D对象(用颜色和/或纹理)包装经变换的第二3D对象代理。可以通过将第一3D对象代理的网格表示中的点的颜色属性与第二3D对象代理的网格表示中的相应点(例如,具有相同的点标识或处于网格序列中的相同位置)的颜色属性相加来再生用于3D对象的像素的颜色值。将颜色属性从帧中标识的3D对象映射到经转换的第一3D对象代理可以包括将3D对象从3D空间(例如XYZ空间)转换到2D空间(例如UV空间)和/或将在帧110中标识的3D对象从2D空间(例如,UV空间)转换到3D空间(例如,XYZ空间)。

[0221] 在示例实施方式中,第一3D对象代理和第二3D对象代理是3D对象的潜在表示。因此,第一3D对象代理可以被用来再生包括用于第一3D对象代理的颜色属性的网格表示,并且第二3D对象代理可以被用来再生包括用于第二3D对象代理的颜色属性的网格表示。如上所述,使用自动编码器解码3D对象的潜在表示可以再生包括用于3D对象的颜色属性的网格表示。自动编码器可以使用具有从元数据读取的结构神经网络。

[0222] 在步骤S950中,将经变换的所存储的3D对象缝合到解压缩的帧中。可以基于经平移和定向的所存储的3D对象的再生颜色值和/或颜色属性以及所选择的帧的剩余部分的再生颜色值和/或颜色属性来重构所选择的帧。在示例实施方式中,可以通过基于(例如,如从元数据读取的)3D对象的所标识的位置,将经平移和定向的所存储的3D对象的再生颜色值和/或颜色属性缝合到所选择的帧的剩余部分的再生颜色值和/或颜色属性中来重构所选择的帧。

[0223] 在步骤S955中,渲染包括3D对象的帧。例如,可以渲染再生纹理数据和再生颜色数据以供显示。渲染纹理数据和颜色数据以供显示可以包括使用3D着色器(例如,顶点着色器、几何着色器等)来绘制与帧相关联的网格。可以将着色器配置为生成图元。可以将着色器配置为将网格中的每个顶点的3D位置和纹理变换为出现在显示器(例如,显示器640)上的2D坐标(例如,图元)。渲染纹理数据和颜色数据以供显示还可以包括执行光栅化。光栅化可以包括基于纹理数据和颜色数据,对图元使用分配像素(例如,颜色)值。

[0224] 在步骤S960中,对渲染的帧进行颜色校正。例如,颜色校正可以包括补偿帧之间的色差、补偿同一场景的多个视图之间的色差、校正对象变形(扭曲)、校正对象边界变形等。

[0225] 示例实施方式可以包括标识两个或更多个3D对象;再生两个或更多个3D对象中的每一个的颜色值和/或颜色属性;以及使用两个或更多个3D对象中的每一个重构所选择的帧。可以基于多个重构帧来再生视频数据5。可以渲染视频数据5(例如,纹理数据和颜色数据),并且颜色校正以在客户端设备的显示器上显示。

[0226] 如上所述,当实现本文所述的技术(例如,上述方法)时,用于与增加数量的视频一起使用。感兴趣的所存储的3D对象(例如,所存储的3D对象135)的数量必定增加,并且存储感兴趣的3D对象所需的资源(例如,内存)的数量也将增加。此外,在流传输活动期间,将感兴趣的3D对象从流传输服务器传送到客户端设备可能需要大量带宽。因此,对于流传输操作而言,可能变得期望有效地编码和解码感兴趣的3D对象。

[0227] 图10示出了根据至少一个示例实施例的用于压缩3D对象的方法的框图。如图10所示,在步骤S1005中,标识视频的感兴趣的至少一个3D对象。例如,感兴趣的对象可以包括CGI电影的3D角色,包括CGI演员(例如,主角、配角和临时演员)、CGI宠物、CGI生物、CGI怪物等。感兴趣的对象可以包括车辆(例如,火车、汽车或飞机)或可以以可预测的方式(例如,以恒定的速度和/或方向)从帧到帧运动的对象。感兴趣的对象可以包括静止或固定的3D对象(例如,场景的背景,场景内固定位置处的家具或远处缓慢移动的对象)在帧到帧(例如,在没有任何相机或场景平移的情况下)可能看起来是静止的。

[0228] 可以预先确定感兴趣的3D对象并且与视频相关联地存储。可以根据需要(例如,当选择关键帧时、作为初始化操作的一部分)确定感兴趣的3D对象,并且将其与视频相关联地存储或添加到先前存储的感兴趣的3D对象中。

[0229] 在步骤S1010中,确定感兴趣的3D对象的网格属性(例如,顶点和连通性)和位置。根据示例实施方式,每个感兴趣的3D对象可以由具有相同数量的点的网格定义,每个点具有位置坐标。点的其他属性可以根据需要添加。因此,用于每个感兴趣的3D对象的网格属性可以包括具有变化的连通性的相同数量的顶点。

[0230] 在步骤S1015中,使用机器训练的生成建模技术来生成与感兴趣的3D对象的网格属性和位置相关联的数量减少的变量。例如,可以使用VAE来生成数量减少的变量。这种数量减少的变量有时被称为用于3D对象的潜在表示或用于3D对象的减少的潜在表示。VAE可以包括神经网络编码器和神经网络解码器,每个神经网络编码器和神经网络解码器都包括具有用于神经网络的C个滤波器、KxK掩码和步长因子的相同配置的神经网络。因此,由神经网络编码器生成的用于感兴趣的3D对象的每个潜在表示的潜在空间中的变量数量对于每个感兴趣的3D对象都是相同的。此外,由神经网络解码器再生的网格中的点的数量对于每个感兴趣的3D对象均是相同的。

[0231] 在示例实施方式中,可以压缩(例如,使用上文参考图5A所述的技术压缩)用于感兴趣的多个3D对象的每一个的网格。例如,可以使用生成建模技术(例如,使用神经网络、卷积神经网络、VAE等)压缩用于感兴趣的多个3D对象中的每一个的网格。可以使用具有卷积神经网络的神经网络编码器来压缩用于感兴趣的多个3D对象中的每一个的网格,其中,卷积神经网络具有基于机器训练的生成建模技术选择和训练的元素,该机器训练的生成建模技术被配置为生成与用于每个3D对象的网格属性和位置相关联的减少数量的变量。与用于

3D对象的网格属性和位置相关联的所生成的数量减少的变量有时被称为用于3D对象的紧凑潜在表示。

[0232] 在步骤S1020中,将变量存储为与感兴趣的3D对象相关联并与视频相关联的感兴趣的3D对象的紧凑可解码表示(或紧凑潜在表示)。例如,变量可以被存储为用于3D对象510的潜在表示。用于3D对象510的潜在表示可以被存储在流传输服务器和/或包括至少一个编码器(例如,编码器105)的设备上。

[0233] 图11示出了根据至少一个示例实施例的用于解压缩3D对象的方法的框图。如图11所示,在步骤S1105中,接收与视频相关联的3D对象的紧凑可解码表示。例如,可以将存储为3D对象510的潜在表示的变量存储在流传输服务器和/或包括至少一个编码器(例如,编码器105)的设备上。在视频流传输操作期间,可以从流传输服务器接收存储为3D对象的潜在表示的至少一组变量。

[0234] 在步骤S1110中,可以使用机器训练的生成建模技术来生成3D对象的网格属性(例如,顶点和连通性)和位置。例如,VAE可以使用可以被用作神经网络解码器的输入的3D对象的潜在表示的潜在空间中的变量,以及VAE可以再生网格的每个点的位置坐标。用于神经网络的C个滤波器、KxK掩码和步长因子的配置可以确定再生的网格的点的数量。当再生网格的点时,神经网络解码器可以再生3D对象的网格属性和位置。

[0235] 在步骤S1115中,存储与视频相关联的作为所存储的3D对象的3D对象的网格属性、位置、定向和颜色属性。例如,当3D对象被再生为包括由面连接的点的集合的网格时,可以将3D对象存储为所存储的3D对象135。每个再生的点可以存储各种属性。例如,属性可以包括每个点的位置、颜色、纹理坐标等。因此,在基于将3D对象用作如在3D编码器和/或3D解码器中实现的几何代理的颜色预测方案中,可以使用包括在所存储的3D对象135中的再生的3D对象。

[0236] 图12图示根据至少一个示例实施例的视频编码器系统1200。如图12中所示,视频编码器系统1200包括至少一个处理器1205、至少一个存储器1210、控制器1220和视频编码器105。至少一个处理器1205、至少一个存储器1210、控制器1220、视频编码器105和视频编码器105通过总线1215通信地耦合。

[0237] 在图12的示例中,视频编码器系统1200可以是或包括至少一个计算设备,并且应该被理解为实际上代表被配置成执行本文描述的方法的任何计算设备。这样,视频编码器系统1200可以被理解为包括可以被用来实现本文描述的技术或其不同或未来版本的各种组件。举例来说,视频编码器系统1200被图示为包括至少一个处理器1205,以及至少一个存储器1210(例如,非暂时性计算机可读存储介质)。

[0238] 可以理解,可以利用至少一个处理器1205来执行存储在至少一个存储器1210上的指令,从而由此实现本文所述的各种特征和功能,或者附加的或替代的特征和功能。当然,至少一个处理器1205和至少一个存储器1210可以用于各种其他目的。特别地,可以理解的是,至少一个存储器1210可以被理解为代表各种类型的存储器以及相关的硬件和软件的示例,其可以被用于实现本文描述的模块中的任何一个。

[0239] 至少一个存储器1210可以被配置成存储与视频编码器系统1200相关联的数据和/或信息。至少一个存储器1210可以是共享资源。例如,视频编码器系统1200可以是较大系统(例如,服务器、个人计算机、移动设备等)的元件。因此,至少一个存储器1210可以被配置成

存储与较大系统内的其他元素(例如,图像/视频服务、网络浏览或有线/无线通信)相关联的数据和/或信息。

[0240] 控制器1220可以被配置成生成各种控制信号,并且将控制信号传达至视频编码器系统1200中的各个块。控制器1220可以被配置成生成控制信号以实现上述技术。根据示例实施例,控制器1220可以被配置成控制视频编码器1225以对图像、图像序列、视频帧、视频序列等进行编码。例如,控制器1220可以生成与视频质量相对应的控制信号。

[0241] 视频编码器105可以被配置成接收视频流输入5并输出压缩(例如,编码)视频位10。视频编码器105可以将视频流输入5转换为离散的视频帧。视频流输入5也可以是图像,因此,压缩(例如,编码)视频位10也可以是压缩的图像位。视频编码器105可以进一步将每个离散视频帧(或图像)转换成块矩阵(以下称为块)。例如,视频帧(或图像)可以被转换为分别具有多个像素的块的 16×16 、 16×8 、 8×8 、 4×4 或 2×2 的矩阵。尽管列出五个示例矩阵,但是示例实施例不限于此。

[0242] 压缩视频位10可以表示视频编码器系统1200的输出。例如,压缩视频位10可以表示编码视频帧(或编码图像)。例如,压缩视频位10可以准备就绪以传输到接收设备(未示出)。例如,视频位可以被传输到系统收发器(未示出)以用于传输到接收设备。

[0243] 至少一个处理器1205可以被配置成执行与控制器1220和/或视频编码器105相关联的计算机指令。至少一个处理器1205可以是共享资源。举例来说,视频编码器系统1200可为较大系统的元件(例如,移动设备、服务器、流传输服务器等)。因此,至少一个处理器1205可以被配置成执行与较大系统内的其他元件相关联的计算机指令(例如,图像/视频服务、网络浏览或有线/无线通信)。

[0244] 在示例实施方式中,视频编码器系统1200可以被实现为包括图形处理单元(GPU)的图形卡和/或芯片(例如,计算机主板上的ASIC)或在图形卡和/或芯片中实现,该图形处理单元(GPU)被配置成从中央处理器(CPU)上去除负载。至少一个处理器1205可以被实现为被配置成并行处理大视频数据块的GPU。可以将GPU配置成处理(例如,压缩)网格数据并从网格数据生成像素数据。至少一个存储器1210可以包括视频存储器和驱动器软件。视频存储器可以是帧缓冲器,其存储代表帧图像或场景的数字数据。视频存储器可以在GPU处理之前和之后存储数字数据。驱动器软件可以包括配置成解压缩视频数据的编解码器。编解码器可以包括基于将3D对象用作此处所述的几何代理来实现颜色预测方案。

[0245] 图13图示根据至少一个示例实施例的视频解码器系统1300。如图13中所示,视频解码器系统1300包括至少一个处理器1305、至少一个存储器1310、控制器1320和视频解码器145。至少一个处理器1305、至少一个存储器1310、控制器1320、视频解码器145和视频解码器145通过总线1315通信地耦合。

[0246] 在图13的示例中,视频解码器系统1300可以是至少一个计算设备,并且应该被理解为实际上代表被配置成执行本文描述的方法的任何计算设备。这样,视频解码器系统1300可以被理解为包括可以用来实现本文描述的技术或其不同或未来版本的各种组件。举例来说,视频解码器系统1300被图示为包括至少一个处理器1305,以及至少一个存储器1310(例如,计算机可读存储介质)。

[0247] 因此,可以理解的是,可以利用至少一个处理器1305来执行存储在至少一个存储器1310上的指令,从而实现本文所述的各种特征和功能,或者附加的或替代的特征以及功

能。当然,至少一个处理器1305和至少一个存储器1310可以用于各种其他目的。特别地,可以理解的是,至少一个存储器1310可以被理解为代表各种类型的存储器以及可以用于实现本文描述的模块中的任何一个的相关硬件和软件的示例。根据示例实施例,视频编码器系统1200和视频解码器系统1300可以被包括在相同的更大的系统(例如,个人计算机、移动设备等)中。

[0248] 至少一个存储器1310可以被配置成存储与视频解码器系统1300相关联的数据和/或信息。至少一个存储器1310可以是共享资源。例如,视频解码器系统1300可以是较大系统(例如,个人计算机、移动设备等)的元件。因此,至少一个存储器1310可以被配置成存储与较大系统内的其他元素(例如,网页浏览或无线通信)相关联的数据和/或信息。

[0249] 控制器1320可以被配置成生成各种控制信号并且将该控制信号传达至视频解码器系统1300中的各个块。控制器1320可以被配置成生成控制信号以便实现下面所描述的视频编码/解码技术。根据示例实施例,控制器1320可以被配置成控制视频解码器145以解码视频帧。

[0250] 视频解码器145可以被配置成接收输入并输出视频流5的压缩(例如,编码)视频位10。视频解码器145可以将压缩视频位10的离散视频帧转换为视频流5。压缩(例如,编码)视频位10也可以是压缩图像位,因此,视频流5也可以是图像。

[0251] 至少一个处理器1305可以被配置成执行与控制器1320和/或视频解码器145相关联的计算机指令。至少一个处理器1305可以是共享资源。例如,视频解码器系统1300可以是较大系统(例如,个人计算机、移动设备等)的元件。因此,至少一个处理器1305可以被配置成执行与较大系统内的其他元素(例如,网页浏览或无线通信)相关联的计算机指令。

[0252] 在示例实施方式中,视频解码器系统1300可以被实现为或包括图形处理单元(GPU)的图形卡和/或芯片(例如,计算机主板上的ASIC)上或者在其中,该图形处理单元(GPU)被配置成去除来自中央处理器(CPU)的负载。至少一个处理器1305可以被实现为被配置成并行处理大视频数据块的GPU。可以将GPU配置成处理(例如,解压缩)网格数据并从网格数据生成像素数据。至少一个存储器1310可以包括视频存储器和驱动器软件。视频存储器可以是帧缓冲器,其存储代表帧图像或场景的数字数据。视频存储器可以在GPU处理之前和之后存储数字数据。驱动器软件可以包括配置成解压缩视频数据的编解码器。编解码器可以包括基于将3D对象用作此处所述的几何代理来实现颜色预测方案。

[0253] 图14示出可以与这里描述的技术一起使用的计算机设备1400和移动计算机设备1450的示例。计算设备1400旨在表示各种形式的数字计算机,诸如膝上型电脑、台式机、工作站、个人数字助理、服务器、刀片服务器、大型机和其他适当的计算机。计算设备1450旨在表示各种形式的移动设备,诸如个人数字助理、蜂窝电话、智能电话和其他类似的计算设备。这里示出的组件、其连接和关系及其功能仅意在为示例性的,而不意在限制本文档中描述和/或要求保护的发明的实施方式。

[0254] 计算设备1400包括处理器1402、存储器1404、存储设备1406、连接到存储器1404和高速扩展端口1410的高速接口1408、以及连接到低速总线1414和存储设备1406的低速接口1412。组件1402、1404、1406、1408、1410和1412中的每一个均使用各种总线来互连,并且可以被安装在公共主板上或者酌情以其他方式安装。处理器1402可处理用于在计算设备1400内执行的指令,包括存储在存储器1404中或者在存储设备1406上以在诸如耦合到高速接口

1408的显示器1416的外部输入或者输出设备上显示用于GUI的图形信息的指令。在其他实施方式中,可以酌情连同多个存储器和多种类型的存储器一起使用多个处理器和/或多个总线。另外,可以连接多个计算设备1400,其中每个设备提供必要操作的部分(例如,作为服务器组、一组刀片服务器或多处理器系统)。

[0255] 存储器1404存储计算设备1400内的信息。在一个实施方式中,存储器1404是一个或多个易失性存储器单元。在另一实施方式中,存储器1404是一个或多个非易失性存储器单元。存储器1404还可以是另一形式的计算机可读介质,诸如磁盘或光盘。

[0256] 存储设备1406能够为计算设备1400提供大容量存储。在一个实施方式中,存储设备1406可以是或者包含计算机可读介质,诸如软盘设备、硬盘设备、光盘设备或磁带设备、闪存存储器或其他类似的固态存储设备或设备的阵列,包括存储区域网络或其他配置中的设备。计算机程序产品可被有形地具体实现在信息载体中。计算机程序产品还可以包含指令,该指令当被执行时执行一个或多个方法,诸如上面描述的那些方法。信息载体是计算机可读介质或机器可读介质,诸如存储器1404、存储设备1406、或处理器1402上的存储器。

[0257] 高速控制器1408管理计算设备1400的带宽密集运算,而低速控制器1412管理较低带宽密集运算。功能的这种分配仅是示例性的。在一个实施方式中,高速控制器1408耦合到存储器1404、显示器1416(例如,通过图形处理器或加速器),并且耦合到高速扩展端口1410,该高速扩展端口1410可以接受各种扩展卡(未示出)。在该实施方式中,低速控制器1412耦合到存储设备1406和低速扩展端口1414。可以包括各种通信端口(例如,USB、蓝牙、以太网、无线以太网)的低速扩展端口可以例如通过网络适配器耦合到一个或多个输入或者输出设备,诸如键盘、指点设备、扫描仪、或诸如交换机或路由器的联网设备。

[0258] 如图中所示,可以以许多不同的形式实现计算设备1400。例如,它可以作为标准服务器1420被实现,或者被多次实现在一组此类服务器中。它还可以作为机架服务器系统1424的一部分被实现。此外,它可以被实现在诸如膝上型计算机1422的个人计算机中。可替换地,来自计算设备1400的组件可以与诸如设备1450的移动设备(未示出)中的其他组件组合。此类设备中的每一个均可以包含计算设备1400、1450中的一个或多个,并且整个系统可以由彼此通信的多个计算设备1400、1450组成。

[0259] 计算设备1450包括处理器1452、存储器1464、诸如显示器1454的输入或者输出设备、通信接口1466、和收发器1468以及其他组件。设备1450还可以被提供有存储设备,诸如微驱动器或其他设备,以提供附加存储。组件1450、1452、1464、1454、1466和1468中的每一个均使用各种总线来互连,并且若干组件可以被安装在公共主板上或者酌情以其他方式安装。

[0260] 处理器1452可执行计算设备1450内的指令,包括存储在存储器1464中的指令。处理器可以作为芯片的芯片组被实现,该芯片包括单独的和多个模拟和数字处理器。处理器可以例如提供用于设备1450的其他组件的协调,诸如对用户界面、由设备1450运行的应用和由设备1450进行的无线通信的控制。

[0261] 处理器1452可以通过耦合到显示器1454的控制接口1458和显示接口1456来与用户进行通信。显示器1454可以是例如TFT LCD(薄膜晶体管液晶显示器)或OLED(有机发光二极管)显示器或其他适当的显示技术。显示接口1456可以包括用于驱动显示器1454以向用户呈现图形和其他信息的适当的电路。控制接口1458可以从用户接收命令并且对它们进行

转换以便提交给处理器1452。此外,可以提供与处理器1452通信的外部接口1462,以使得能实现设备1450与其他设备的近区域通信。外部接口1462可以在一些实施方式中例如提供用于有线通信,或者在其他实施方式中用于无线通信,并且还可以使用多个接口。

[0262] 存储器1464存储计算设备1450内的信息。存储器1464可作为一个或多个计算机可读介质、一个或多个易失性存储器单元、或者一个或多个非易失性存储器单元中的一种或多种而被实现。还可以提供扩展存储器674并且通过扩展接口1472将它连接到设备1450,扩展接口1472可以包括例如SIMM(单列直插存储器模块)卡接口。这种扩展存储器1474可以为设备1450提供附加的存储空间,或者还可以为设备1450存储应用或其他信息。具体地,扩展存储器1474可以包括用于执行或者补充上述过程的指令,并且还可以包括安全信息。因此,例如,扩展存储器1474可以作为用于设备1450的安全模块被提供,并且可以被编程有允许安全使用设备1450的指令。此外,可以经由SIMM卡提供安全应用以及附加信息,诸如以不可破解的方式将识别信息放置在SIMM卡上。

[0263] 存储器可以包括例如闪速存储器和/或NVRAM存储器,如在下面所讨论的。在一个实施方式中,计算机程序产品被有形地具体实现在信息载体中。计算机程序产品包含指令,该指令当被执行时执行一个或多个方法,诸如上述的那些方法。信息载体是可以例如通过收发器1468或外部接口1462接收的计算机或机器可读介质,诸如存储器1464、扩展存储器1474、或处理器1452上的存储器。

[0264] 设备1450可以通过通信接口1466以无线方式通信,该通信接口1466必要时可以包括数字信号处理电路。通信接口1466可以提供用于各种模式或协议下的通信,各种模式或协议诸如GSM语音呼叫、SMS、EMS或MMS消息传送、CDMA、TDMA、PDC、WCDMA、CDMA2000或GPRS等。这种通信可以例如通过射频收发器1468而发生。此外,可以发生短距离通信,诸如使用蓝牙、WiFi或其他这种收发器(未示出)。此外,GPS(全球定位系统)接收器模块570可以向设备1450提供附加的导航和位置相关无线数据,其可以酌情由在设备1450上运行的应用使用。

[0265] 设备1450还可以使用音频编解码器1460可听地通信,该音频编解码器1460可以从用户接收口语信息并且将它转换为可用的数字信息。音频编解码器1460可以同样地诸如通过扬声器为用户生成可听声音——例如,在设备1450的头戴式耳机中。这种声音可以包括来自语音电话呼叫的声音,可以包括记录的声音(例如,语音消息、音乐文件等)并且还可以包括由在设备1450上操作的应用所生成的声音。

[0266] 如图中所示,可以以许多不同的形式实现计算设备1450。例如,它可以作为蜂窝电话1480被实现。它还可以作为智能电话1482、个人数字助理或其他类似的移动设备的一部分被实现。

[0267] 这里描述的系统和技术各种实施方式可用数字电子电路、集成电路、专门地设计的ASIC(专用集成电路)、计算机硬件、固件、软件、和/或其组合加以实现。这些各种实施方式可包括在可编程系统上可执行和/或可解释的一个或多个计算机程序中的实施方式,该可编程系统包括至少一个可编程处理器,其可以是专用的或通用的,耦合以从存储系统、至少一个输入设备和至少一个输出设备接收数据和指令,并且以向存储系统、至少一个输入设备和至少一个输出设备发送数据和指令。这里描述的系统和技术各种实施方式可作为可组合软件和硬件方面的电路、模块、块或系统被实现和/或一般地在本文中被作为可组

合软件和硬件方面的电路、模块、块或系统。例如,模块可以包括在处理器(例如,形成在硅基底、GaAs基底等上的处理器)或某个其他可编程数据处理装置上执行的功能/行为/计算机程序指令。

[0268] 上述示例实施例中的一些被描述为作为流程图描绘的过程或方法。尽管流程图将操作描述为顺序过程,然而可以并行地、并发地或同时地执行许多操作。此外,可以重新布置操作的次序。这些过程可以在其操作完成时被终止,但是也可以具有未包括在图中的附加步骤。这些过程可以对应于方法、函数、程序、子例程、子程序等。

[0269] 上面讨论的方法——其中的一些通过流程图来图示——可以通过硬件、软件、固件、中间件、微码、硬件描述语言或其任何组合来实现。当用软件、固件、中间件或微码加以实现时,用于执行必要的任务的程序代码或代码段可以被存储在诸如存储介质的机器或计算机可读介质中。处理器可以执行必要的任务。

[0270] 本文中公开的具体结构和功能细节仅仅是表示性的以用于描述示例实施例的目的。然而,示例实施例被以许多替代形式具体实现,而不应该被解释为限于仅本文中阐述的实施例。

[0271] 应理解的是,尽管可以在本文中使用术语第一、第二等来描述各种元件,然而这些元件不应该受这些术语限制。这些术语仅用于区分一个元件和另一元件。例如,第一元件能被称为第二元件,并且类似地,第二元件能被称为第一元件,而不脱离示例实施例的范围。如本文中所使用的,术语和/或包括相关列举项目中的一个或多个的任何和所有组合。

[0272] 应理解的是,当一个元件被称为连接或者耦合到另一元件时,它可直接地连接或者耦合到另一元件或者可以存在中间元件。相比之下,当一个元件被称为直接地连接或者直接地耦合到另一元件时,不存在中间元件。应该以相似的方式解释用于描述元件之间的关系的其他单词(例如,在…之间对直接地在…之间、相邻对直接地相邻等)。

[0273] 本文中使用的术语仅用于描述特定实施例的目的而不旨在限制示例实施例。如本文中所使用的,除非上下文另外清楚地指示,否则单数形式一(a)、一个(an)和该(the)也旨在包括复数形式。应进一步理解的是,术语含、含有、包括和/或包括有当在本文中使用,指定存在陈述的特征、整数、步骤、操作、元件和/或组件,但是不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或其组。

[0274] 还应该注意的是,在一些替代实施方式中,所指出的功能/行为可以不按图中指出的次序发生。例如,取决于所涉及的功能性或者行为,相继示出的两个图实际上可以被并发地执行或者有时可以以相反的次序被执行。

[0275] 除非另外定义,否则本文中使用的术语(包括技术和科学术语)具有与由示例实施例所属的领域的普通技术人员所通常理解的相同的含义。应进一步理解的是,除非在本文中明确地如此定义,否则术语(例如,在常用词典中定义的那些术语)应该被解释为具有与其在相关领域的上下文中的含义一致的含义,而不应在理想化或过于正式的意义上进行解释。

[0276] 上述示例实施例的各部分和对应的详细描述是按软件或算法以及对计算机存储器内的数据比特的操作的符号表示来呈现的。这些描述和表示是本领域的普通技术人员用来有效地将其工作的实质传达给本领域的其他普通技术人员的描述和表示。算法(当在这里使用该术语时,并且当一般地使用它时)被认为是导致期望结果的步骤的自相一致序列。

这些步骤是要求对物理量物的理操纵的那些步骤。通常,尽管不一定,然而这些量采取能够被存储、转移、组合、比较和以其他方式操纵的光学信号、电信号或磁信号的形式。有时主要由于通用的原因,将这些信号称为比特、值、元素、符号、字符、术语、数字等已证明是方便的。

[0277] 在上述说明性实施例中,对可以作为程序模块或功能过程被实现的操作(例如,以流程图的形式)的行为和符号表示的引用包括执行特定任务或者实现特定抽象数据类型并且可以使用现有结构元件处的现有硬件来描述和/或实现的例程、程序、对象、组件、数据结构等。这种现有硬件可以包括一个或多个中央处理单元(CPU)、数字信号处理器(DSP)、专用集成电路、现场可编程门阵列(FPGA)计算机等。

[0278] 然而,应该记住的是,所有这些和类似的术语都将与适当的物理量相关联并且仅仅是应用于这些量的方便标签。除非另外具体地陈述,或者如从讨论中显而易见的,诸如处理或计算或计算出或确定显示等的术语指代计算机系统或类似的电子计算设备的动作和过程,所述计算机系统或类似的电子计算设备将被表示为计算机系统的寄存器和存储器内的物理、电子量的数据操纵并变换成被类似地表示为计算机系统存储器或寄存器或其他此类信息存储、传输或显示设备内的物理量的其他数据。

[0279] 另外注意的是,示例实施例的软件实现的方面通常被编码在某种形式的非暂时性程序存储介质上或者实现在某种类型的传输介质上。程序存储介质可以是磁的(例如,软盘或硬盘驱动器)或光学的(例如,紧致盘只读存储器或CD ROM),并且可以是只读或随机存取的。类似地,传输介质可以是双绞线对、同轴电缆、光纤或为本领域所知的某个其他适合的传输介质。示例实施例不受任何给定实施方式的这些方面限制。

[0280] 最后,还应该注意的是,虽然所附权利要求书陈列本文中描述的特征的特定组合,但是本公开的范围不限于此后要求保护的特定组合,而是替代地扩展到包含本文中公开的特征或实施例的任何组合,而不管此时是否已在所附权利要求中具体地枚举了该特定组合。

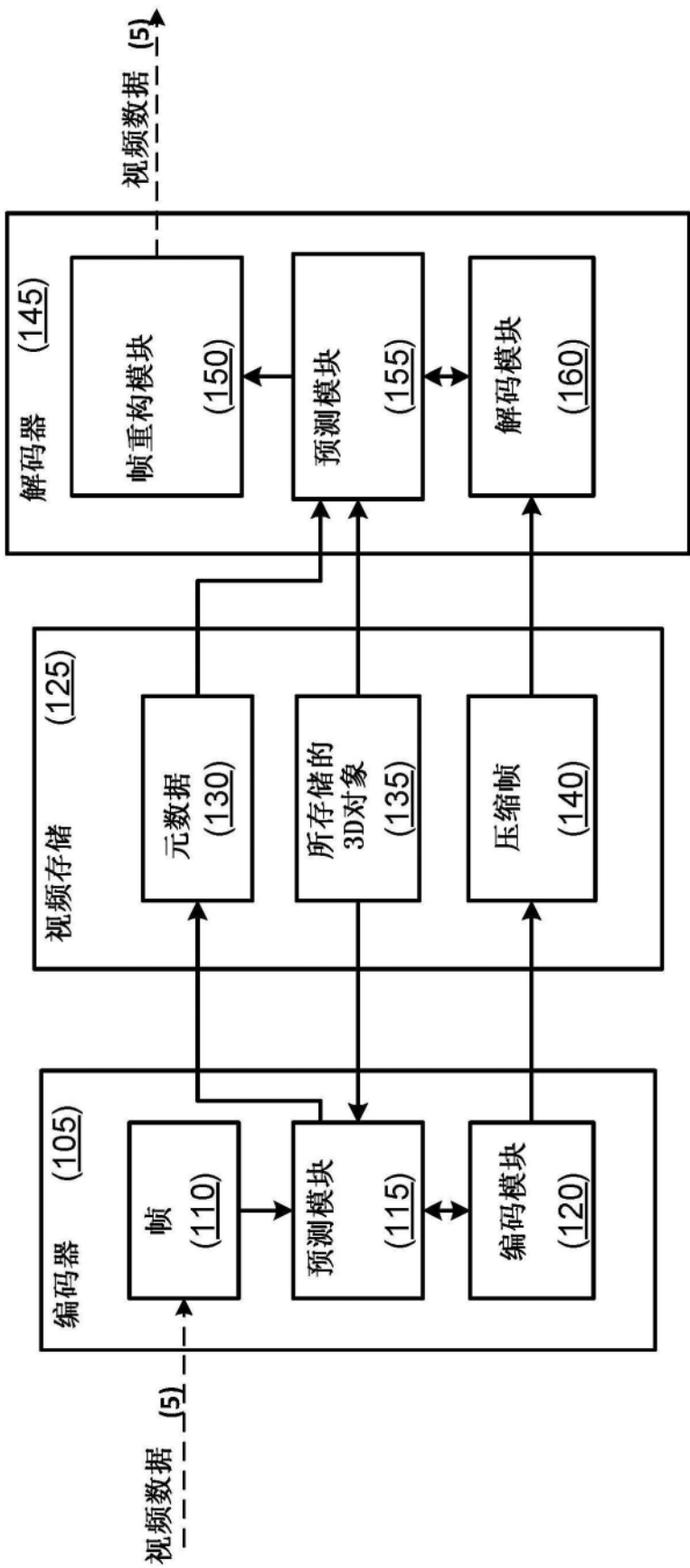


图1

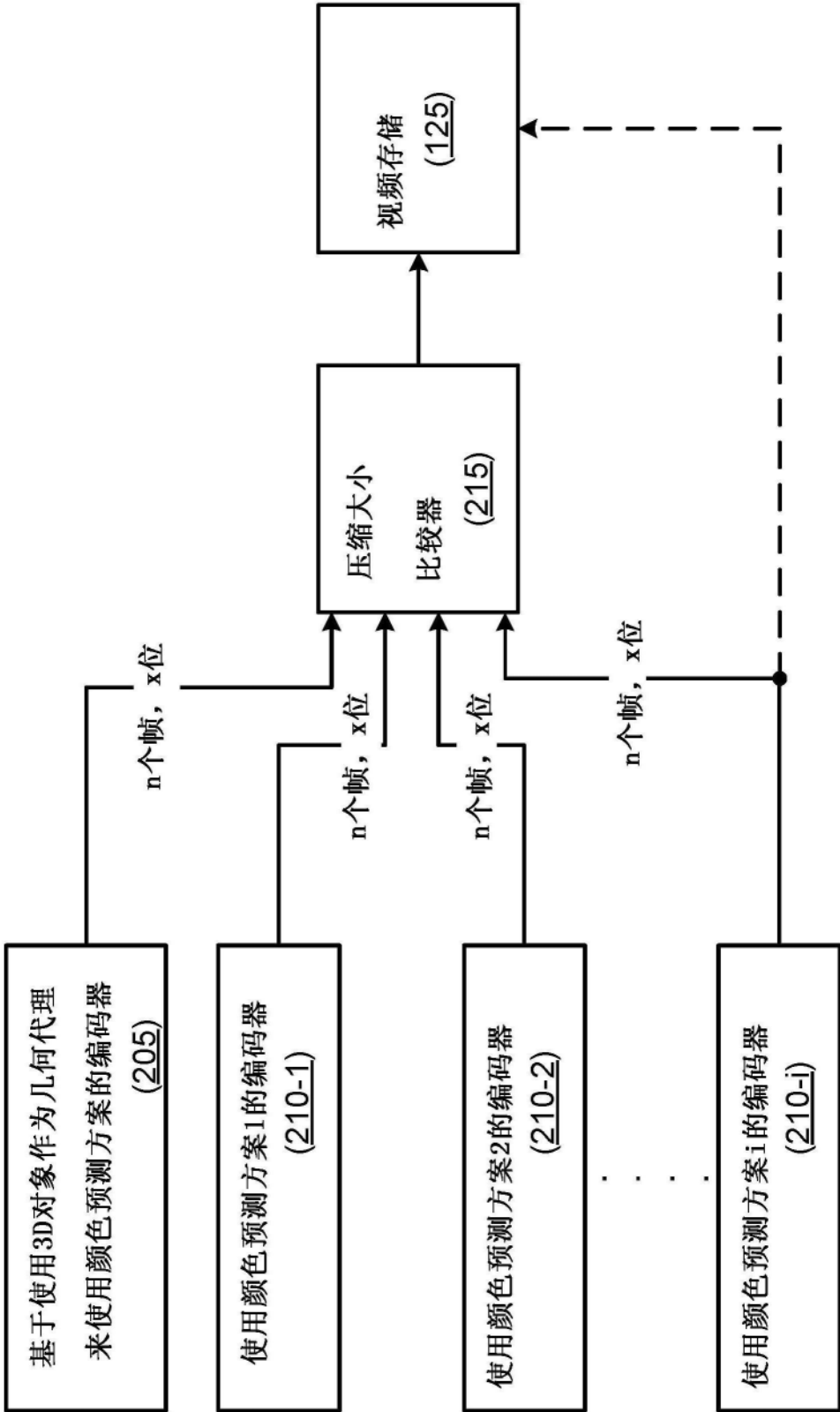


图2

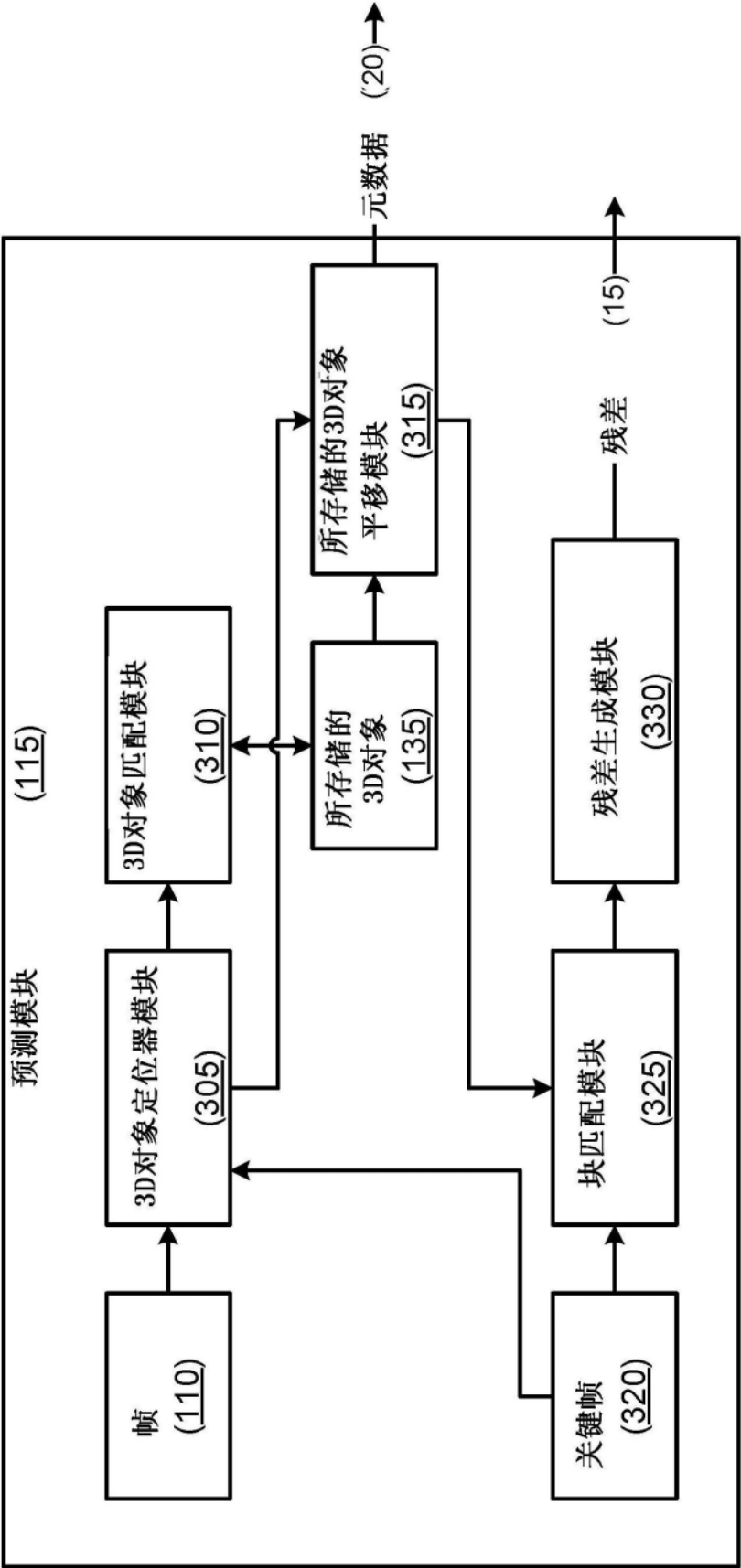


图3A

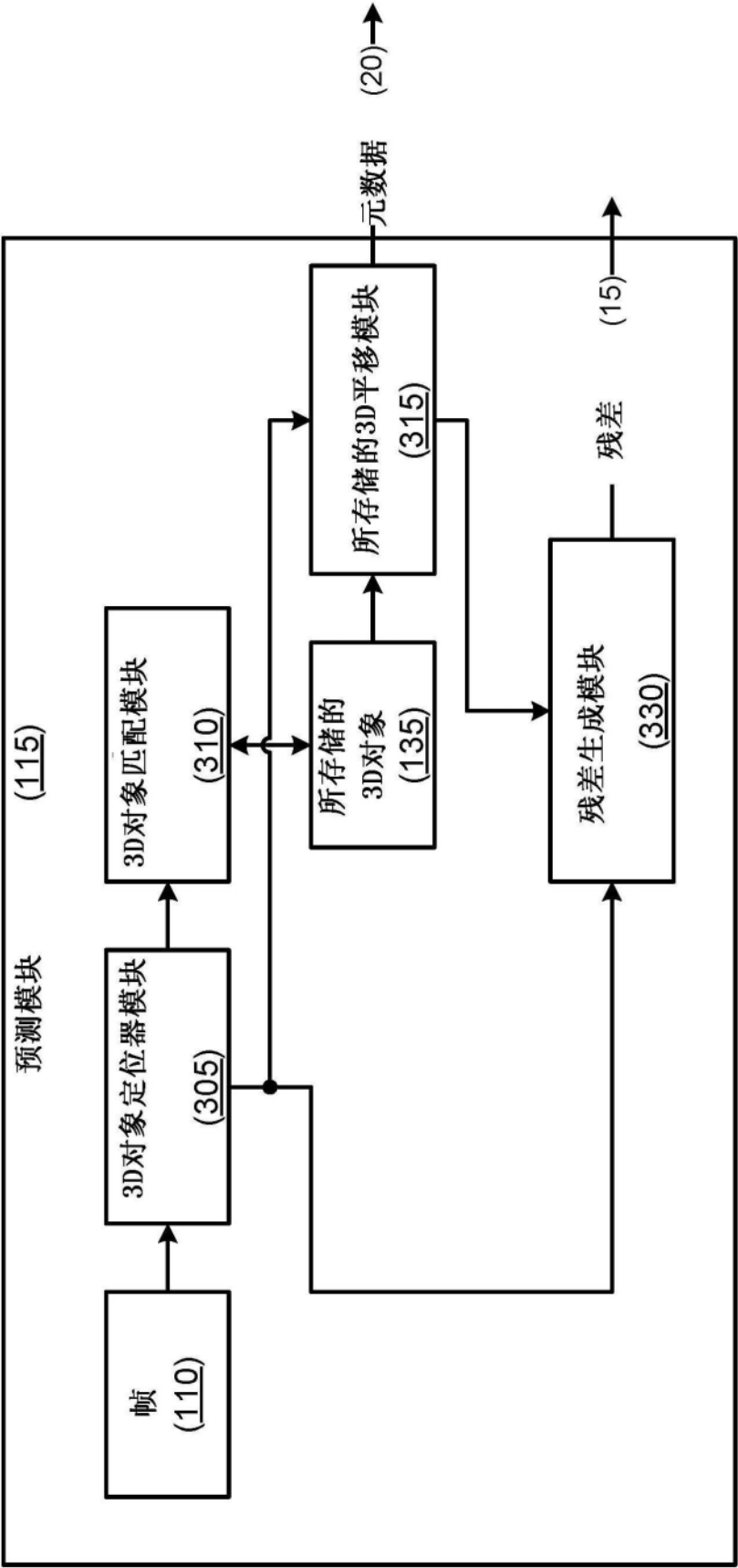


图3B

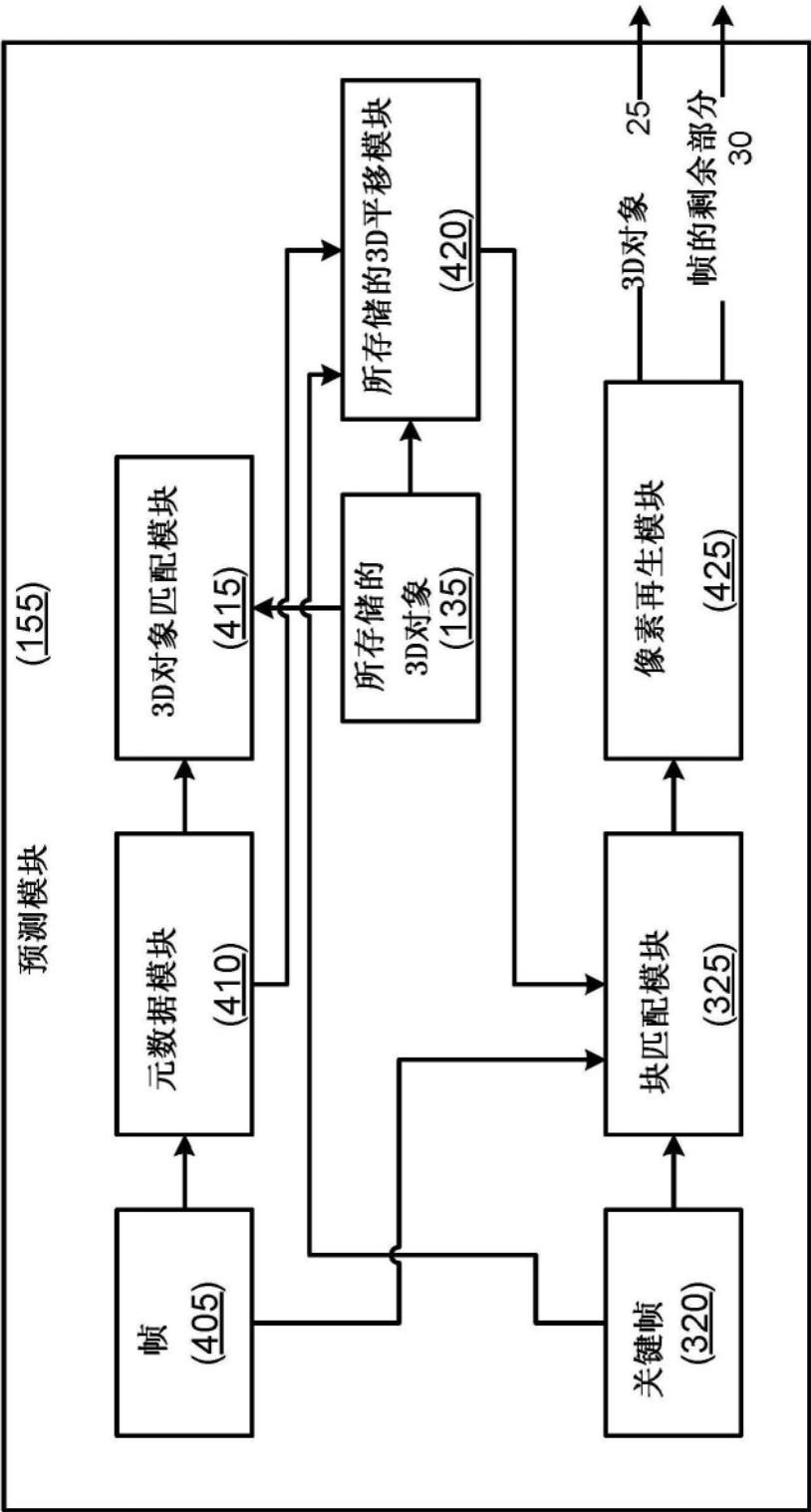


图4A

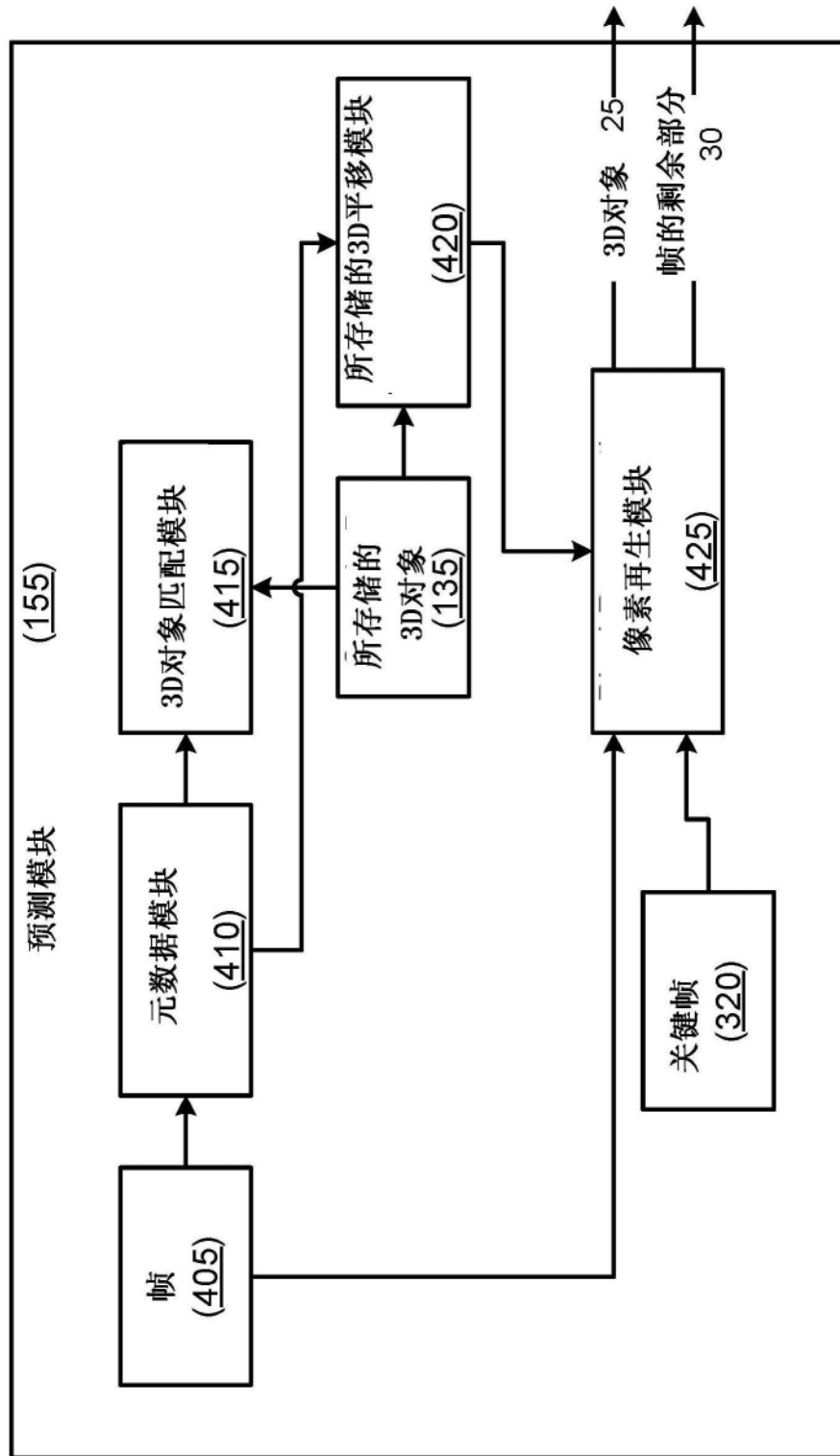


图4B

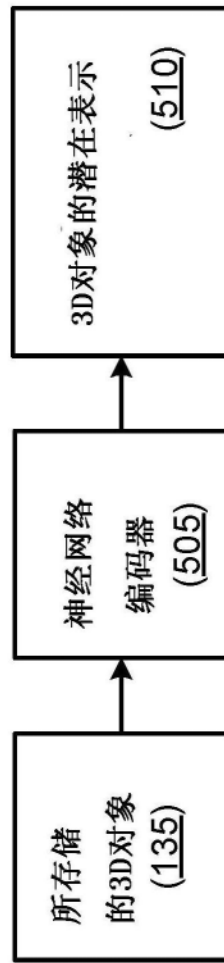


图5A

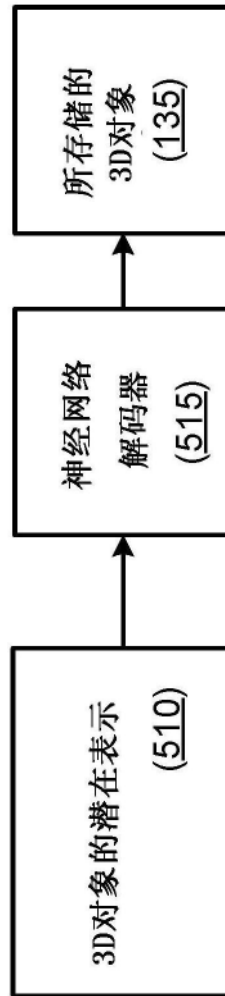


图5B

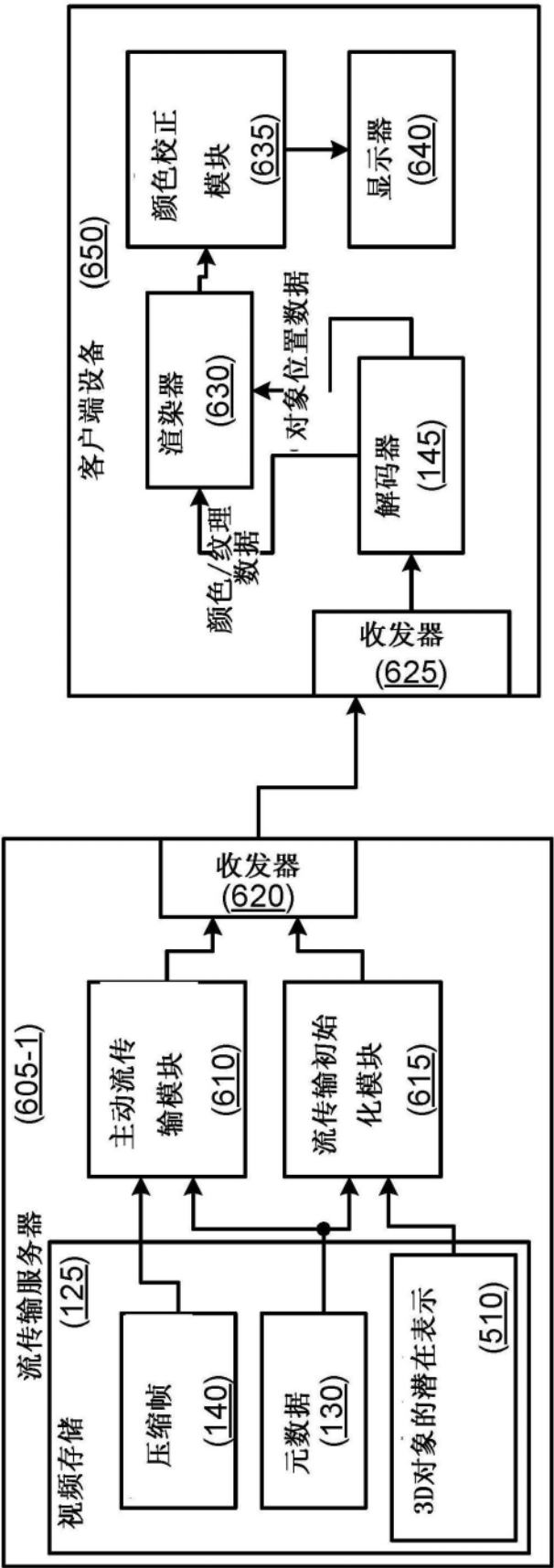


图6A

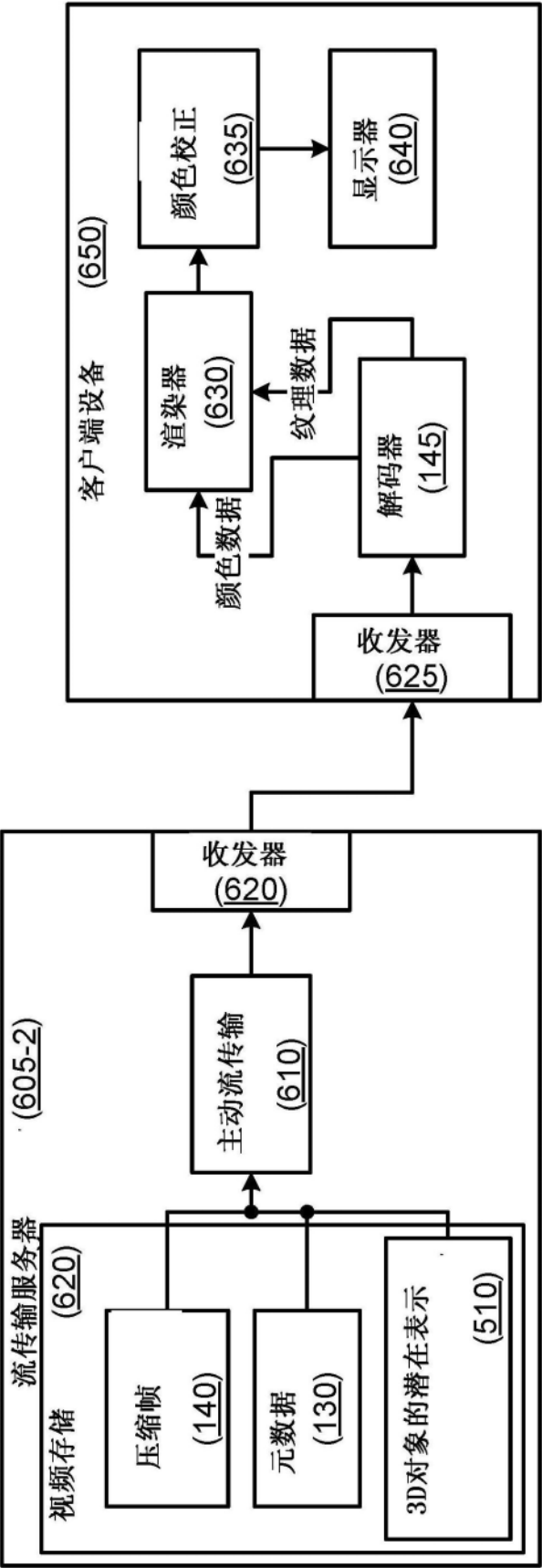


图6B

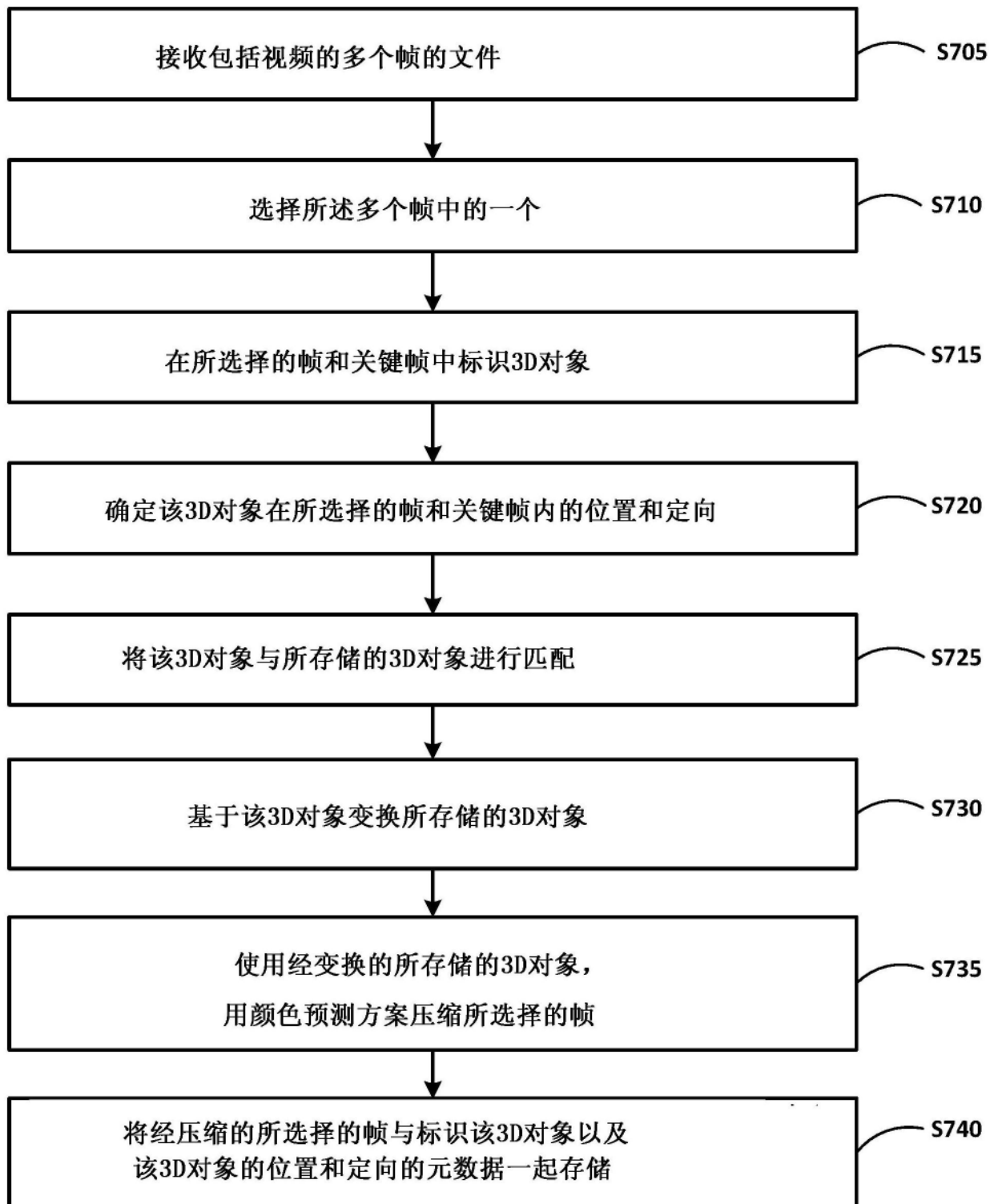


图7

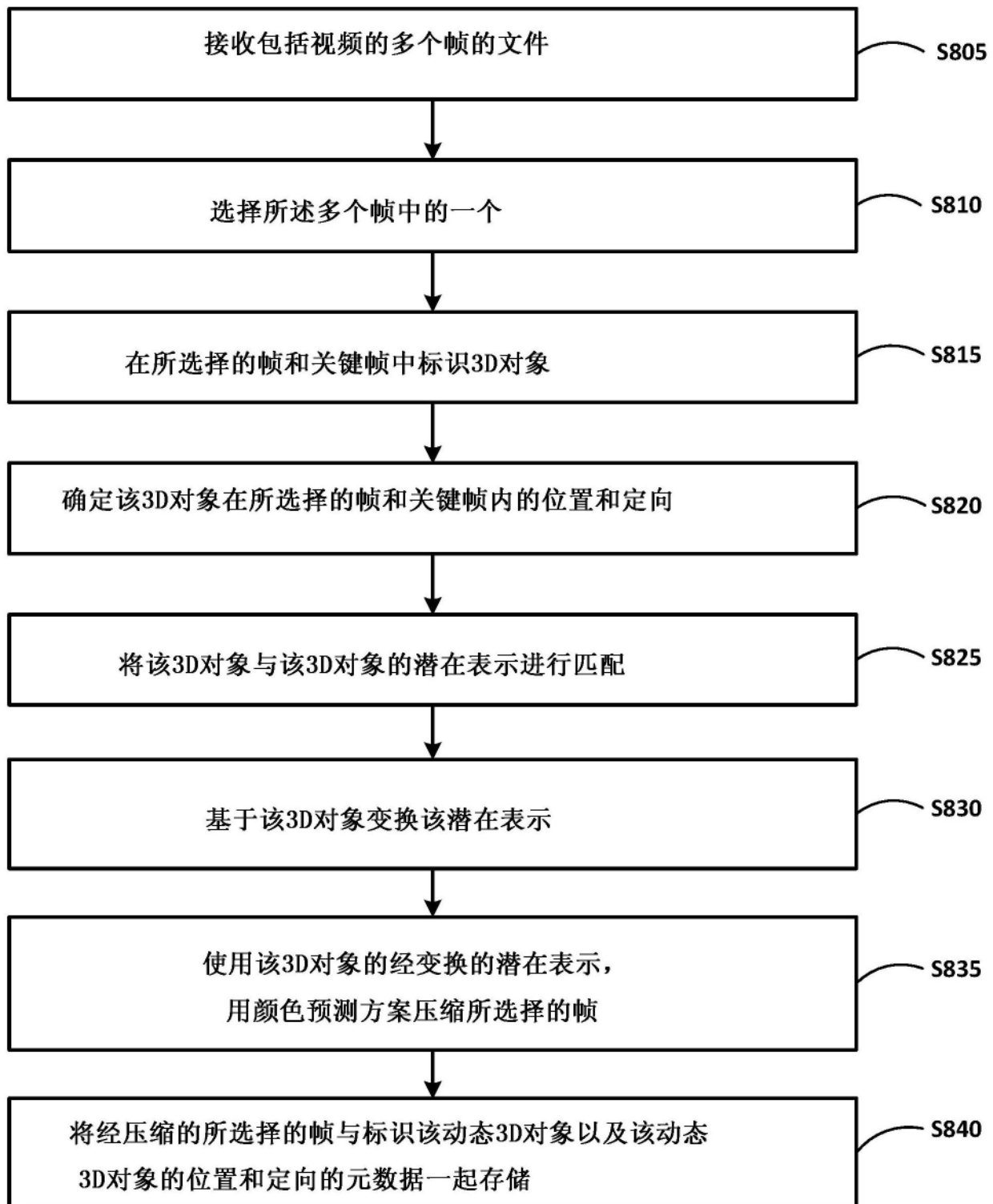


图8

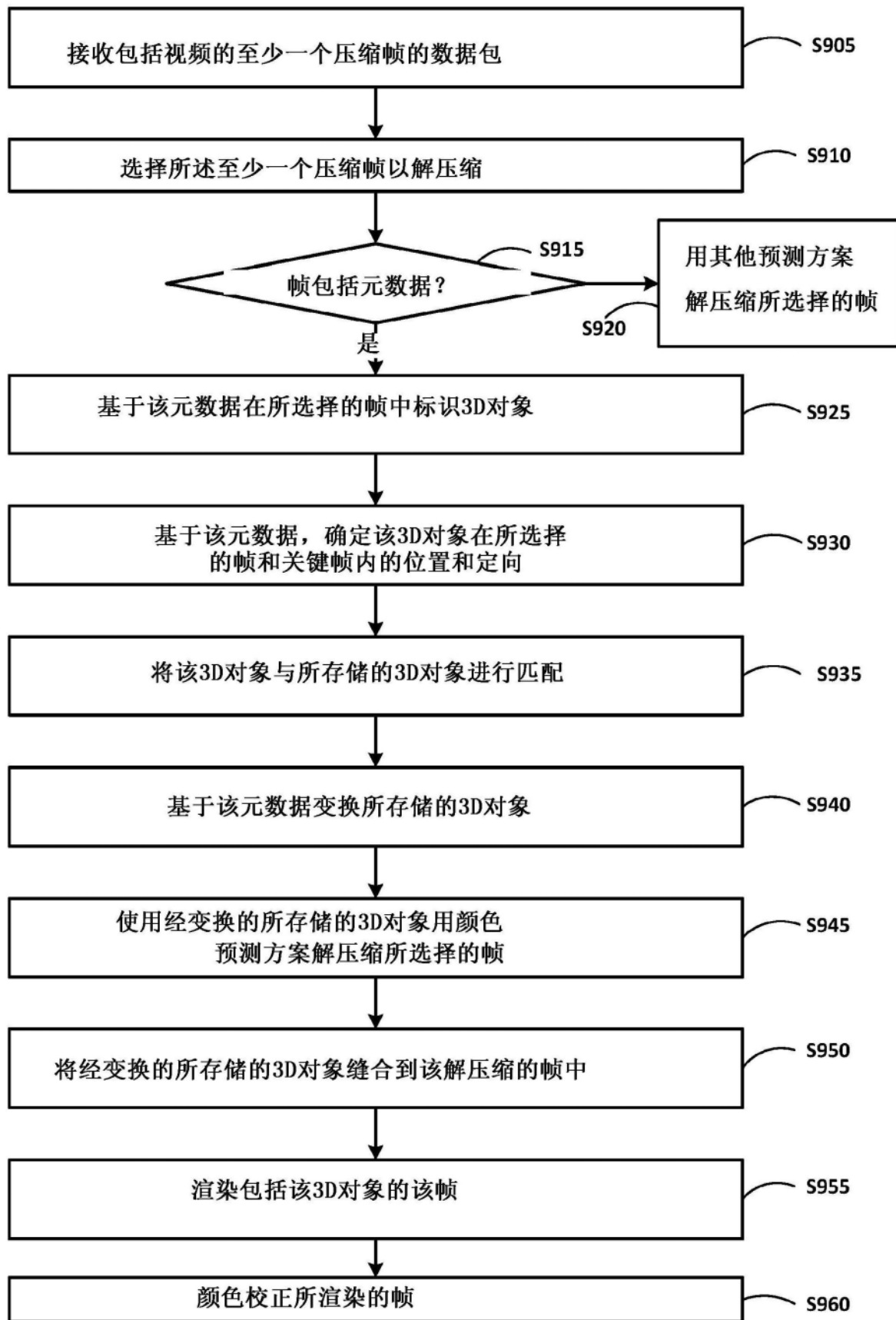


图9

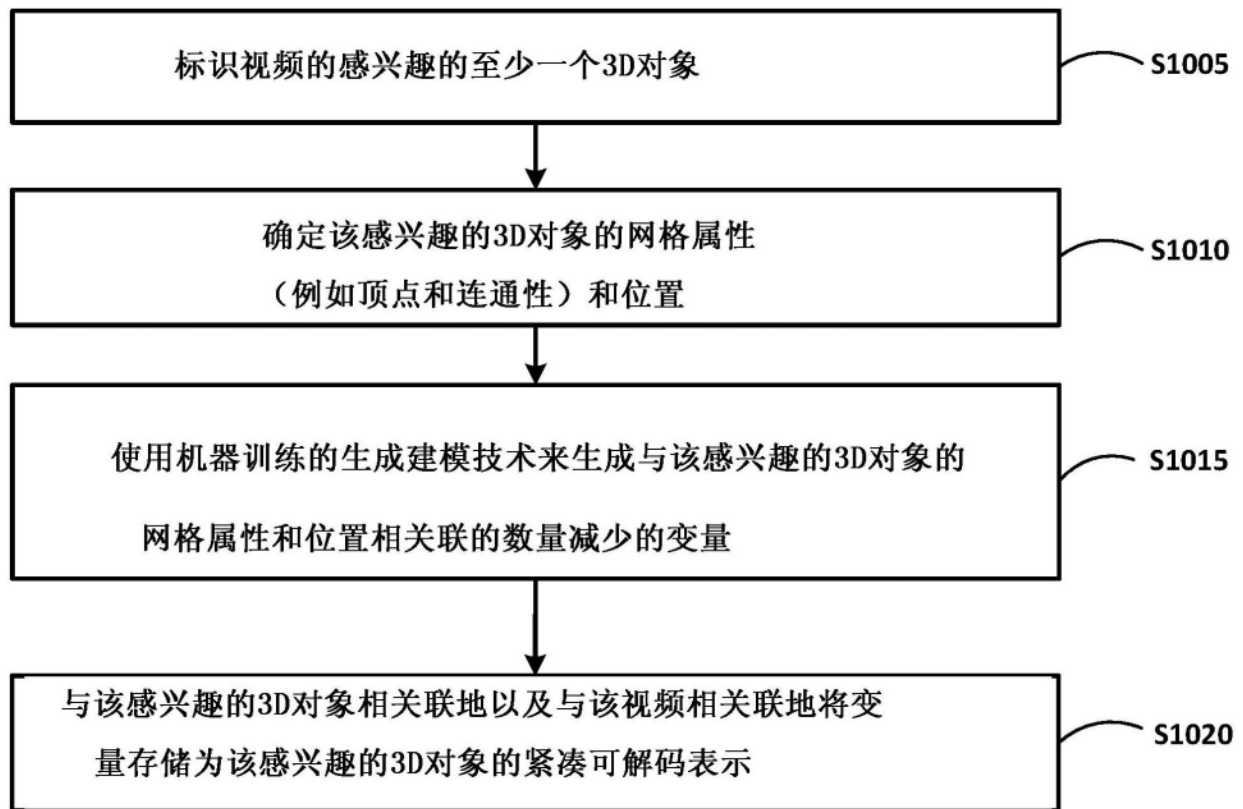


图10

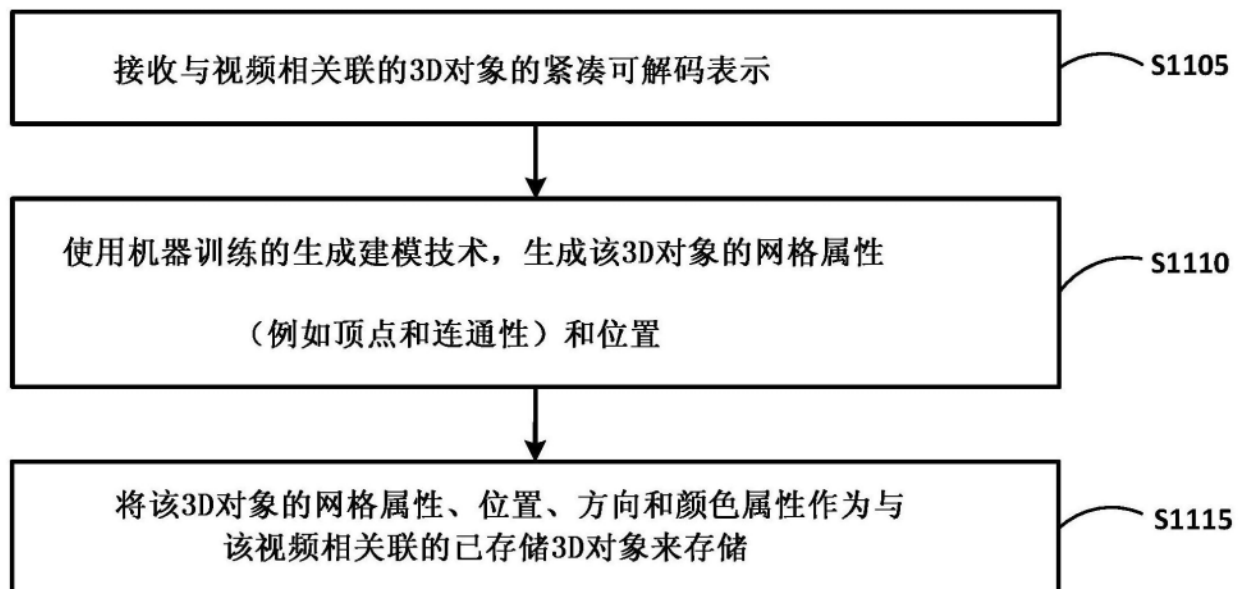


图11

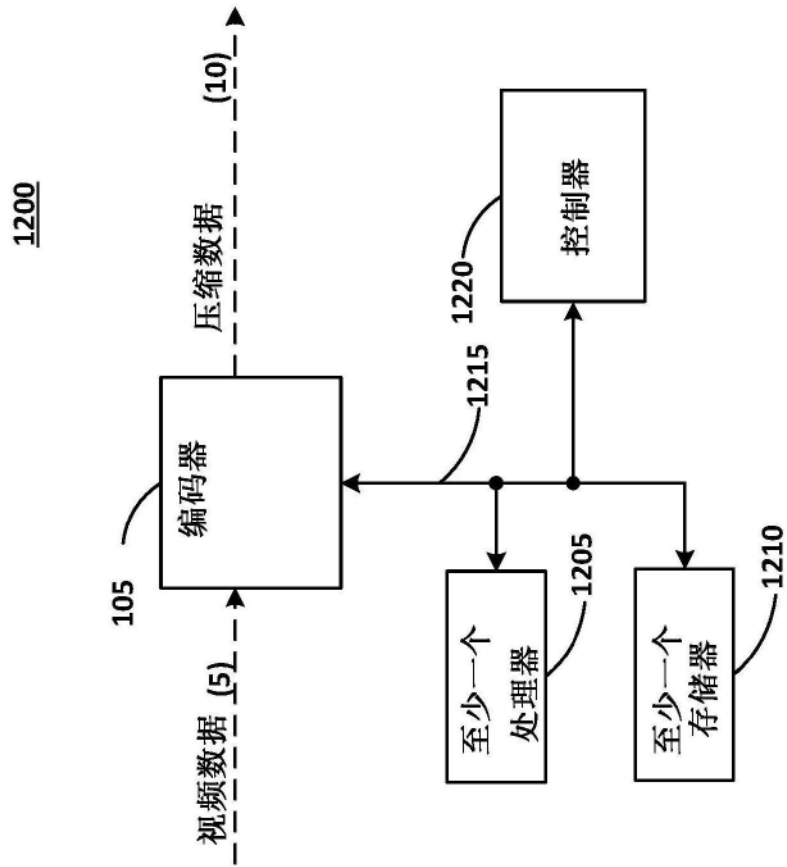


图12

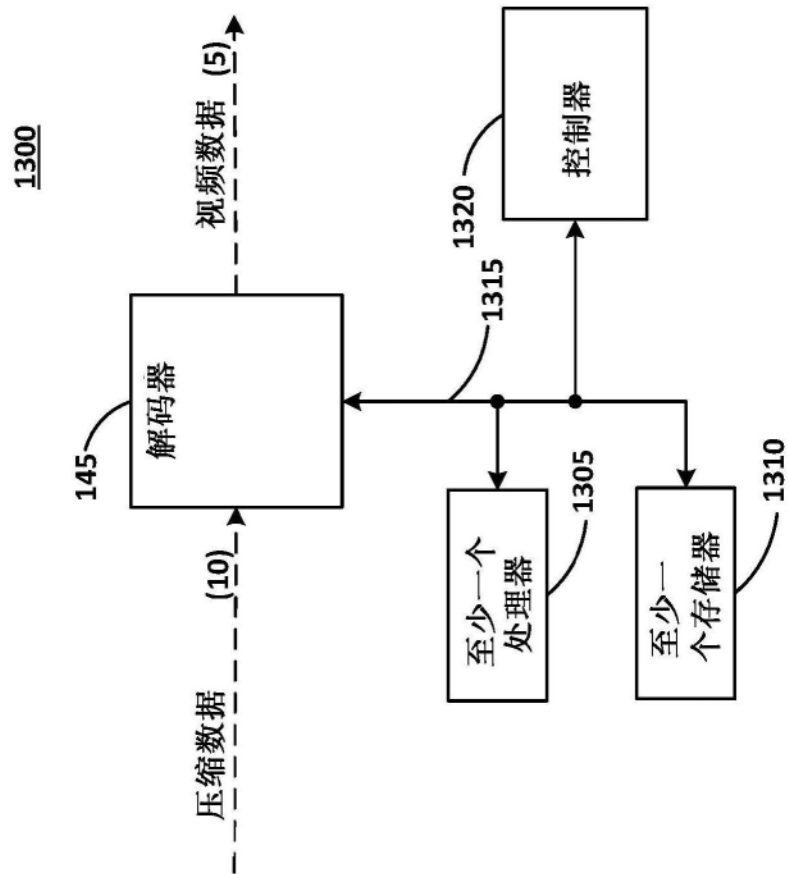


图13

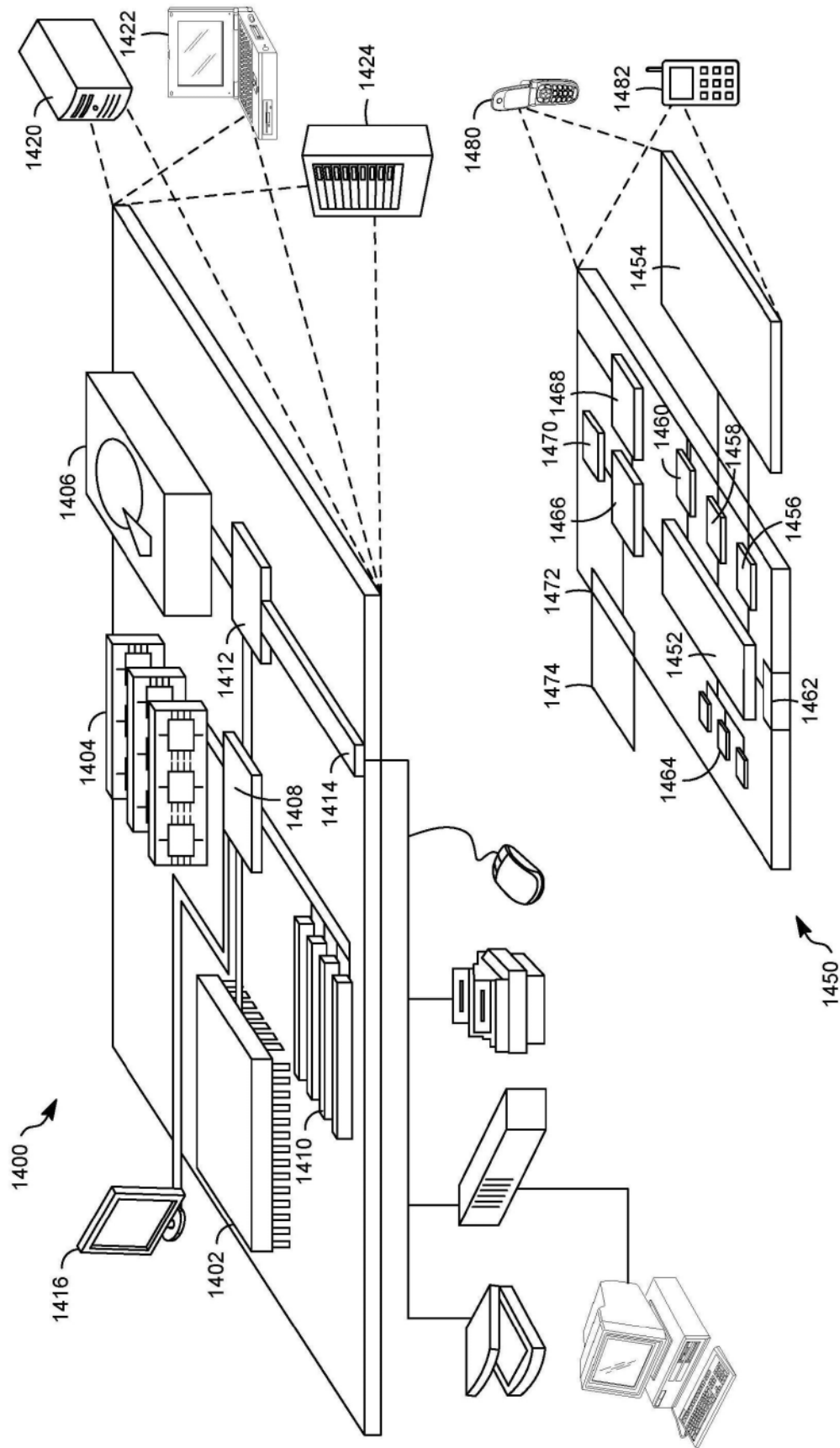


图14