



(12) 发明专利申请

(10) 申请公布号 CN 119856491 A

(43) 申请公布日 2025. 04. 18

(21) 申请号 202280099894.7

(51) Int. Cl.

(22) 申请日 2022.09.20

H04N 19/142 (2006.01)

H04N 19/177 (2006.01)

(85) PCT国际申请进入国家阶段日
2025.03.10

(86) PCT国际申请的申请数据
PCT/CN2022/119802 2022.09.20

(87) PCT国际申请的公布数据
W02024/059998 EN 2024.03.28

(71) 申请人 高通股份有限公司
地址 美国加利福尼亚

(72) 发明人 张楠

(74) 专利代理机构 永新专利商标代理有限公司
72002

专利代理师 赵腾飞

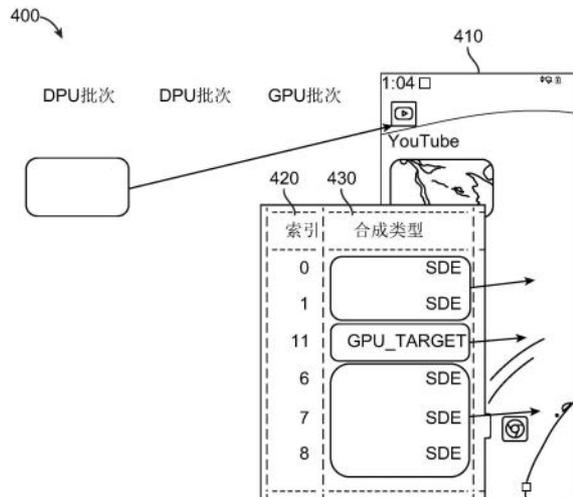
权利要求书3页 说明书37页 附图10页

(54) 发明名称

用于视频译码的可变帧内(I帧)时间间隔和
图片组(GOP)长度

(57) 摘要

提供了用于处理视频数据的系统和技术。例如,过程可包括获得与计算设备的显示器相关联的视频数据帧,其中该视频数据帧包括一个或多个层。可将与该视频数据帧中所包括的该一个或多个层相关联的层信息和与先前视频数据帧中所包括的一个或多个层相关联的层信息进行比较。基于确定与该视频数据帧相关联的帧几何形状变化,可使用该视频数据帧生成帧间预测帧。可基于与该视频数据帧中所包括的该一个或多个层相关联的该层信息来确定更新的图片组(GOP)长度。



1. 一种用于处理视频数据的装置,所述装置包括:
至少一个存储器;和
至少一个处理器,所述至少一个处理器耦合到所述至少一个存储器,所述至少一个处理器被配置为:
获得与计算设备的显示器相关联的视频数据帧,其中所述视频数据帧包括一个或多个层;
比较与所述视频数据帧中所包括的所述一个或多个层相关联的层信息和与先前视频数据帧中所包括的一个或多个层相关联的层信息;
基于确定与所述视频数据帧相关联的帧几何形状变化来使用所述视频数据帧生成帧间预测帧;以及
基于与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息来确定更新的照片组(GOP)长度。
2. 根据权利要求1所述的装置,其中所述至少一个处理器被配置为基于与所述视频数据帧中所包括的主要层相关联的层信息来确定所述更新的GOP长度。
3. 根据权利要求2所述的装置,其中所述视频数据帧中所包括的所述主要层是使用比与所述视频数据帧中所包括的一个或多个附加层相关联的相应z顺序更大的z顺序来渲染的。
4. 根据权利要求1所述的装置,其中所述层信息包括针对所述一个或多个层中所包括的每个相应层的与每个相应层相关联的层名称、与每个相应层相关联的层格式以及与每个相应层相关联的一个或多个坐标中的至少一者。
5. 根据权利要求1所述的装置,其中所述层信息包括层数量或帧层数中的至少一者。
6. 根据权利要求1所述的装置,其中所述至少一个处理器被配置为基于比较与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息和与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息来确定与所述视频数据帧相关联的所述帧几何形状变化。
7. 根据权利要求6所述的装置,其中为了确定与所述视频数据帧相关联的所述帧几何形状变化,所述至少一个处理器被配置为:
确定与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息相比于与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息已改变大于阈值量。
8. 根据权利要求1所述的装置,其中所述至少一个处理器还被配置为:
基于与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息相比于与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息改变小于阈值量,确定帧几何形状变化不与所述视频数据帧相关联。
9. 根据权利要求8所述的装置,其中所述至少一个处理器还被配置为:
基于所述视频数据帧以及预先确定的数量的先前视频数据帧不与帧几何形状变化相关联来检测到与所述视频数据帧相关联的显示器空闲状态;以及
应用显示器空闲GOP长度,其中所述显示器空闲GOP长度大于所述更新的GOP长度。
10. 根据权利要求8所述的装置,其中所述至少一个处理器还被配置为:

基于所述视频数据帧不与帧几何形状变化相关联来将所述视频数据帧编码为预测帧(P帧)或双向帧(B帧)。

11. 根据权利要求1所述的装置,其中所述视频数据帧包括在所述计算设备的显示器上显示的视频显示数据。

12. 根据权利要求1所述的装置,其中所述视频数据帧是与从所述计算设备到第二计算设备的无线显示共享相关联的所捕获的视频显示数据的帧。

13. 根据权利要求12所述的装置,其中所述视频数据帧和所述先前视频数据帧是所捕获的视频显示数据的多个帧中所包括的连续帧。

14. 根据权利要求13所述的装置,其中所述至少一个处理器还被配置为:
使用所述帧间预测帧对所捕获的视频显示数据的所述多个帧的至少一部分进行编码。

15. 根据权利要求14所述的装置,其中所述至少一个处理器还被配置为:
将编码的视频数据发送到与同所述计算设备相同的无线显示共享会话相关联的第二设备;

其中所述编码的视频数据包括所述帧间预测帧以及基于对所捕获的视频显示数据的所述多个帧的至少所述部分进行编码而生成的单向预测帧或双向预测帧中的至少一者。

16. 一种用于处理视频数据的方法,所述方法包括:

获得与计算设备的显示器相关联的视频数据帧,其中所述视频数据帧包括一个或多个层;

比较与所述视频数据帧中所包括的所述一个或多个层相关联的层信息和与先前视频数据帧中所包括的一个或多个层相关联的层信息;

基于确定与所述视频数据帧相关联的帧几何形状变化来使用所述视频数据帧生成帧间预测帧;以及

基于与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息来确定更新的照片组(GOP)长度。

17. 根据权利要求16所述的方法,其中所述更新的GOP长度是基于与所述视频数据帧中所包括的主要层相关联的层信息来确定的。

18. 根据权利要求17所述的方法,其中所述视频数据帧中所包括的所述主要层是使用比与所述视频数据帧中所包括的一个或多个附加层相关联的相应z顺序更大的z顺序来渲染的。

19. 根据权利要求16所述的方法,其中所述层信息包括针对所述一个或多个层中所包括的每个相应层的与每个相应层相关联的层名称、与每个相应层相关联的层格式以及与每个相应层相关联的一个或多个坐标中的至少一者。

20. 根据权利要求16所述的方法,其中所述层信息包括层数量或帧层数中的至少一者。

21. 根据权利要求16所述的方法,其中确定与所述视频数据帧相关联的所述帧几何形状变化基于比较与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息和与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息。

22. 根据权利要求21所述的方法,其中确定与所述视频数据帧相关联的所述帧几何形状变化包括:

确定与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息相比于与

所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息已改变大于阈值量。

23. 根据权利要求16所述的方法,所述方法还包括:

基于与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息相比于与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息改变小于阈值量,确定帧几何形状变化不与所述视频数据帧相关联。

24. 根据权利要求23所述的方法,所述方法还包括:

基于所述视频数据帧以及预先确定的数量的先前视频数据帧不与帧几何形状变化相关联来检测到与所述视频数据帧相关联的显示器空闲状态;以及

应用显示器空闲GOP长度,其中所述显示器空闲GOP长度大于所述更新的GOP长度。

25. 根据权利要求23所述的方法,所述方法还包括:

基于所述视频数据帧不与帧几何形状变化相关联来将所述视频数据帧编码为预测帧(P帧)或双向帧(B帧)。

26. 根据权利要求16所述的方法,其中所述视频数据帧包括在所述计算设备的显示器上显示的视频显示数据。

27. 根据权利要求16所述的方法,其中所述视频数据帧是与从所述计算设备到第二计算设备的无线显示共享相关联的所捕获的视频显示数据的帧。

28. 根据权利要求27所述的方法,其中所述视频数据帧和所述先前视频数据帧是所捕获的视频显示数据的多个帧中所包括的连续帧。

29. 根据权利要求28所述的方法,所述方法还包括:

使用所述帧间预测帧对所捕获的视频显示数据的所述多个帧的至少一部分进行编码。

30. 根据权利要求29所述的方法,所述方法还包括:

将编码的视频数据发送到与同所述计算设备相同的无线显示共享会话相关联的第二设备;

其中所述编码的视频数据包括所述帧间预测帧以及基于对所捕获的视频显示数据的所述多个帧的至少所述部分进行编码而生成的单向预测帧或双向预测帧中的至少一者。

用于视频译码的可变帧内(I帧)时间间隔和图片组(GOP)长度

技术领域

[0001] 本公开整体涉及视频译码(例如,包括视频数据的编码和/或解码)。例如,本公开的各方面涉及改进与可变帧内时间间隔和/或图片组(GOP)长度相关的视频译码技术。

背景技术

[0002] 数字视频能力可被并入到各种各样的设备中,包括数字电视机、数字直播系统、无线广播系统、个人数字助理(PDA)、膝上型计算机或台式计算机、平板计算机、电子书阅读器、数字相机、数字记录设备、数字媒体播放器、视频游戏设备、视频游戏控制台、蜂窝或卫星无线电话(所谓的“智能电话”)、视频电话会议设备、视频流式传输设备等。此类设备允许视频数据被处理并输出以供消费。数字视频数据包括大量数据,以满足消费者和视频提供者的需求。例如,视频数据的消费者期望最高质量的视频,具有高保真度、高分辨率、高帧速率等。结果,为了满足这些需求所需要的大量视频数据为处理并存储视频数据的通信网络和设备带来了负担。

[0003] 数字视频设备可实现视频译码技术以压缩视频数据。根据一个或多个视频译码标准或格式执行视频译码。例如,视频译码标准或格式包括通用视频译码(VVC)、高效视频译码(HEVC)、高级视频译码(AVC)、MPEG-2部分2译码(MPEG代表运动图像专家组)等等,以及专有视频编解码器/格式,诸如由开放媒体联盟开发的AOMedia视频1(AV1)。视频译码通常利用预测方法(例如,帧间预测、帧内预测等),这些预测方法利用在视频图像或序列中存在的冗余。视频译码技术的目标是将视频数据压缩为使用较低比特率的形式,同时避免或最小化视频质量的降级。随着不断发展的视频服务变得可用,需要具有更好的译码效率的译码技术。

发明内容

[0004] 在一些示例中,描述了用于使用可变帧内(I帧)时间间隔和/或可变长度图片组(GOP)长度来执行视频译码的系统和技术。例如,该系统和技术可使用基于诸如视频帧层信息、视频帧层几何形状等信息确定的可变I帧间隔和/或可变GOP长度来执行视频译码(例如,编码和/或解码)。根据至少一个例示性示例,提供了一种用于处理视频数据的方法。该方法包括:获得与计算设备的显示器相关联的视频数据帧,其中视频数据帧包括一个或多个层;比较与视频数据帧中所包括的一个或多个层相关联的层信息和与先前视频数据帧中所包括的一个或多个层相关联的层信息;基于确定与视频数据帧相关联的帧几何形状变化来使用视频数据帧生成帧间预测帧;以及基于与视频数据帧中所包括的一个或多个层相关联的层信息来确定更新的图片组(GOP)长度。

[0005] 在另一示例中,提供了一种装置,该装置包括至少一个存储器(例如,被配置为存储数据)和至少一个处理器(例如,在电路中实现),该至少一个处理器耦合到至少一个存储器。该至少一个处理器被配置为并可:获得与计算设备的显示器相关联的视频数据帧,其中视频数据帧包括一个或多个层;比较与视频数据帧中所包括的一个或多个层相关联的层信

息和与先前视频数据帧中所包括的一个或多个层相关联的层信息;基于确定与视频数据帧相关联的帧几何形状变化来使用视频数据帧生成帧间预测帧;以及基于与视频数据帧中所包括的一个或多个层相关联的层信息来确定更新的图片组 (GOP) 长度。

[0006] 在另一示例中,提供了一种非暂态计算机可读介质,该非暂态计算机可读介质上存储有指令,这些指令在由一个或多个处理器执行时使一个或多个处理器:获得与计算设备的显示器相关联的视频数据帧,其中视频数据帧包括一个或多个层;比较与视频数据帧中所包括的一个或多个层相关联的层信息和与先前视频数据帧中所包括的一个或多个层相关联的层信息;基于确定与视频数据帧相关联的帧几何形状变化来使用视频数据帧生成帧间预测帧;以及基于与视频数据帧中所包括的一个或多个层相关联的层信息来确定更新的图片组 (GOP) 长度。

[0007] 在另一示例中,提供了一种装置,该装置包括:用于获得与计算设备的显示器相关联的视频数据帧的部件,其中视频数据帧包括一个或多个层;用于比较与视频数据帧中所包括的一个或多个层相关联的层信息和与先前视频数据帧中所包括的一个或多个层相关联的层信息的部件;用于基于确定与视频数据帧相关联的帧几何形状变化来使用视频数据帧生成帧间预测帧的部件;以及用于基于与视频数据帧中所包括的一个或多个层相关联的层信息来确定更新的图片组 (GOP) 长度的部件。

[0008] 在一些方面,本文所述的装置中的一个或多个装置是以下设备、是以下设备的部分和/或包括以下设备:移动设备或无线通信设备(例如,移动电话或其他移动设备)、扩展现实(XR)设备或系统(例如,虚拟现实(VR)设备、增强现实(AR)设备或混合现实(MR)设备)、可穿戴设备(例如,网络连接的手表或其他可穿戴设备)、相机、个人计算机、膝上型计算机、交通工具或交通工具的计算设备或组件、服务器计算机或服务器设备、另一设备或它们的组合。在一些方面,该装置包括用于捕获一个或多个图像的一个相机或多个相机。在一些方面,该装置还包括用于显示一个或多个图像、通知和/或其他可显示数据的显示器。在一些方面,上述装置可包括一个或多个传感器(例如,一个或多个惯性测量单元(IMU),诸如一个或多个陀螺仪、一个或多个陀螺测试仪、一个或多个加速度计、它们的任何组合和/或其他传感器)。

[0009] 该发明内容不旨在标识所要求保护的主题的关键或必要特征,其也不旨在孤立地用于确定所要求保护的主题的范围。本主题应当参考本专利的整个说明书的合适部分、任何或所有附图以及每项权利要求来理解。

[0010] 前述内容以及其他特征和方面将在参考以下说明书、权利要求书和所附附图时变得更明显。

附图说明

[0011] 本申请的例示性方面在下文参照以下附图进行了详细的描述:

[0012] 图1是例示根据本公开的一些示例的编码设备和解码设备的示例的框图;

[0013] 图2A是例示根据一些示例的角度预测模式的示例的示图;

[0014] 图2B是例示根据一些示例的通用视频译码(VVC)中的定向帧内预测模式的示例的示图;

[0015] 图3是例示根据一些示例的图片组(GOP)长度和帧间(I帧)时间间隔的示例的示

图;

[0016] 图4A是例示根据一些示例的与所捕获的视频显示数据的帧相关联的帧层的示例的示图;

[0017] 图4B是例示根据一些示例的与所捕获的视频显示数据的帧相关联的视频帧层堆栈的示例的示图;

[0018] 图4C是例示根据一些示例的包括与所捕获的视频显示数据的给定帧中所包括的多个层相关联的层信息的示例列表的示图;

[0019] 图5是例示根据一些示例的用于使用可变I帧时间间隔和可变GOP长度来执行视频译码的过程的示例的流程图;

[0020] 图6是例示根据一些示例的用于使用可变I帧时间间隔和可变GOP长度来执行视频译码的过程的另一示例的流程图;

[0021] 图7是例示根据一些示例的示例视频编码设备的框图;并且

[0022] 图8是例示根据一些示例的示例视频解码设备的框图。

具体实施方式

[0023] 以下提供本公开的某些方面和各方面。这些方面和各方面中的一些可独立地应用,并且它们中的一些可组合应用,这对于本领域技术人员来说是显而易见的。在以下描述中,出于解释目的阐述了具体细节以提供对本申请的各方面的透彻理解。然而,显然的是,可在没有这些具体细节的情况下实施各个方面。各附图和描述不旨在是限制性的。

[0024] 以下描述仅提供了示例性方面,并且并不旨在限定本公开的范围、适用性或配置。相反,对示例性方面的以下描述将向本领域技术人员提供能够用于实现示例性方面的描述。应当理解,在不脱离如所附权利要求所阐述的本申请的实质和范围的情况下,可对元素的功能和排列做出各种改变。

[0025] 数字视频数据可包括大量数据,尤其是随着对高质量视频数据的需求持续增长。例如,视频数据的消费者通常期望越来越高质量的视频,具有高保真度、高分辨率、高帧速率等。然而,满足这样的需求所需要的大量视频数据可能给通信网络以及处理和存储视频数据的设备带来显著的负担。

[0026] 视频译码设备(例如,编码设备、解码设备或组合的编码-解码设备)实现视频压缩技术以有效地对视频数据进行译码(例如,编码和/或解码)。视频压缩技术可以包括应用不同的预测模式,包括空间预测(例如,帧内预测(intra-frame prediction或intra-prediction)、时间预测(例如,帧间预测(inter-frame prediction或inter-prediction)、层间预测(跨视频数据的不同层)、和/或用于减少或去除视频序列中固有的冗余的其他预测技术。视频编码器可以将原始的视频序列的每个图片分割为多个矩形区域,这些矩形区域被称为视频块或译码单元(下文将更详细地进行描述)。这些视频块可以是使用特定的预测模式来进行编码的。

[0027] 视频块可以以一种或多种方式被划分为一组或多组较小的块。块可以包括译码树块、预测块、变换块或其他合适的块。除非另外指定,否则通常对“块”的引用可以指代此类视频块(例如,如本领域普通技术人员将理解的,译码树块、译码块、预测块、变换块、或其他适当的块或子块)。此外,这些块中的每个块在本文中也可以被互换地称为“单元”(例如,译

码树单元 (CTU)、译码单元、预测单元 (PU)、变换单元 (TU) 等)。在一些情况下,单元可指示在比特流中编码的译码逻辑单元,而块可指示视频帧缓冲器中的过程所针对的一部分。

[0028] 对于帧间预测模式,视频编码器可以在位于另一时间位置的帧(或图片)(其被称为参考帧或参考图片)中搜索与被编码的块相似的块。视频编码器可以将搜索限制为距要被编码的块一定的空间位移。可以使用包括水平位移分量和竖直位移分量的二维(2D)运动矢量来定位最佳匹配。对于帧内预测模式,视频编码器可基于来自相同图片内的先前被编码的相邻块的数据,使用空间预测技术来形成预测块。

[0029] 视频编码器可以确定预测误差。例如,可将预测确定为正被编码的块和预测块中的像素值之间的差。预测误差也可以称为残差。视频编码器还可将变换应用于预测误差(例如,离散余弦变换(DCT)或其他合适的变换)以生成变换系数。在变换之后,视频编码器可以对变换系数进行量化。量化的变换系数和运动矢量可以使用语法元素来表示,并且与控制信息一起形成对视频序列的译码的表示。在一些实例中,视频编码器可以对所量化的变换系数和/或语法元素进行熵编码,从而进一步减少用于其表示所需要的比特数量。

[0030] 在对所接收的比特流进行熵解码和解量化之后,视频解码器可以使用以上讨论的语法元素和控制信息来构造用于解码当前帧的预测数据(例如,预测块)。例如,视频解码器可将预测块和压缩的预测误差相加。视频解码器可通过使用所量化的系数对变换基函数进行加权来确定压缩的预测误差。在重建的帧与原始帧之间的差被称为重建误差。

[0031] 可以根据特定的视频译码标准来执行视频译码。视频译码标准的示例包括但不限于ITU-T H.261、ISO/IEC MPEG-1视频、ITU-T H.262或ISO/IEC MPEG-2视频、ITU-T H.263、ISO/IEC MPEG-4视频、高级视频译码(AVC)或ITU-T H.264(包括其可缩放视频译码(SVC)和多视图视频译码(MVC)扩展)、高效视频译码(HEVC)或ITU-T H.265(包括其范围和屏幕内容译码、3D视频译码(3D-HEVC)、多视图(MV-HEVC)和可缩放(SHVC)扩展)、通用视频译码(VVC)或ITU-T H.266和其扩展、VP9、开放媒体联盟(AOMedia)视频1(AV1)、基本视频译码(EVC)等等。

[0032] 如上文所指出,视频编码器可将原始视频序列的每个图片分割为一个或多个较小块或矩形区域,该一个或多个较小块或矩形区域然后可使用例如帧内预测(intra-prediction或intra-frame prediction)进行编码以移除原始视频序列所固有的空间冗余。如果在帧内预测模式下对块进行编码,则基于先前编码且重建的块(例如,包括在相同的视频数据帧中)来形成预测块,该先前编码且重建的块在视频编码器和视频解码器两者中可用于形成预测参考。例如,可使用与相同视频数据帧相关联的先前编码的相邻块的像素值来确定当前块(例如,当前编码或当前解码)内部的像素值的空间预测。这些像素值用作参考像素。参考像素可被组织成一个或多个参考像素线和/或参考像素组。在一些示例中,帧内预测可应用于块的亮度分量和色度分量两者。

[0033] 可利用多个不同帧内预测模式提供不同空间预测技术以基于来自相同图片内的先前编码的相邻块(例如,来自参考像素)的数据来形成预测参考或预测块。帧内预测模式可包括平面和DC模式和/或定向帧内预测模式(也被称为“规则帧内预测模式”)。在一些示例中,可使用单个平面帧内预测和单个DC帧内预测模式以及多个定向帧内预测模式。帧内预测模式描述了用于基于参考像素值来计算正被译码的区域中的像素值的不同变体或方法。在一个例示性示例中,HEVC标准提供了33个定向帧内预测模式。在另一例示性示例中,

VVC和/或VVC测试模型5 (VTM5) 扩展HEVC定向帧内预测模式以提供总共93个定向帧内预测模式。

[0034] 在视频解码器处,每个块的帧内预测模式选择(例如,对应于视频编码器在生成编码块时所做出的帧内预测模式选择)可由解码器确定(例如,推导)或可向视频解码器发信号通知,诸如在比特流的语法中。例如,在一些情况下,可使相邻块之间的帧内预测模式相关(例如,如果使用帧内预测模式2来预测两个先前编码的相邻块,则用于当前块的最佳帧内预测模式可能也是帧内预测模式2)。在一些示例中,对于每个当前块,视频编码器和视频解码器可计算最可能的帧内预测模式。视频编码器还可向视频解码器发信号通知帧内预测模式(例如,使用标记、模式参数、模式选择器等)。

[0035] 在当前VVC标准中,提供了93个定向帧内预测模式,如先前所指出。每个帧内预测模式与不同的角度方向相关联,使得帧内预测模式是唯一且非重叠的。定向帧内预测模式可被分类为整数角度模式或分数(非整数角度)模式。对于给定视频数据块,整数角度帧内预测模式在整数位置处具有参考像素,例如,整数角度帧内预测模式具有穿过位于当前译码块的周边处的参考像素的位置的斜率。相比之下,分数帧内预测模式不具有整数位置处的参考像素,而是具有穿过两个相邻参考像素之间某处的点的斜率(例如,分数位置 $i+f$ (i :整数部分, f :分数部分)处的像素的斜率穿过像素 i 和像素 $i+1$)。

[0036] 在一些方面,可使用帧内预测帧(I帧)、预测帧(P帧)和/或双向帧(B帧)的组合来执行视频译码。例如,I帧可仅包括使用帧内预测的视频数据块。在一些方面,基于I帧仅包括以相同I帧内的其他视频数据块为参考的视频数据块,I帧可用作关键帧。例如,相对于顺序位于I帧之前或之后的一个或多个P帧和/或B帧,I帧可用作关键帧。在一些方面,可生成一个或多个P帧以参考先前编码的I帧(例如,充当关键帧)和/或先前编码的P帧。在一些示例中,可生成一个或多个B帧以参考先前编码和随后解码(例如,未来)的帧两者,它们均可包括I帧(例如,充当关键帧)。

[0037] 在一些方面,两个关键帧(例如,I帧)之间的距离可被称为图片组(GOP)或GOP长度。GOP长度可用两个I帧关键帧之间的帧数或两个I帧关键帧之间的时间量来测量。例如,如果对于每秒30帧的视频的每一秒,插入一个I帧并将其用作关键帧,则GOP长度是30帧或一秒。

[0038] 在一些示例中,可使用可变I帧时间间隔(例如,也被称为“可变GOP长度”)来执行视频译码。例如,可使用可变I帧时间间隔或GOP长度来提高视频译码的效率。在一些方面,不同类型的视频内容的最佳视频译码效率可与不同GOP长度相关联。例如,可使用相对较大的GOP长度来对相对静态的视频内容进行译码(例如,基于视频内容随时间相对较少地改变),而可使用相对较小的GOP长度来对具有快速移动或许多移动对象的视频内容进行译码(例如,基于视频内容随时间相对快速地改变)。

[0039] 在一些方面,可变长度GOP视频译码(例如,编码和/或解码)可基于视频内容分析和/或运动分析以确定与视频内容相关联的随时间的改变量。此类技术可实现提高的视频译码性能,但通常与高计算复杂性和功率消耗相关联(例如,与同视频译码分开地执行视频内容分析和/或运动分析相关联)。在一些情况下,可基于与给定视频译码设备相关联的计算约束、功率约束和/或译码时间约束来利用固定长度GOP视频译码(例如,固定I帧时间间隔视频译码)。

[0040] 例如,WiFi显示技术(例如,诸如Miracast[®])可用于通过捕获和编码源设备(例如,智能电话)的显示内容并将编码的视频数据发送到接收设备(例如,电视机)而在WiFi设备之间无线地共享视频内容。基于源设备(例如,智能电话)处可用的计算资源和/或功率和能量资源,WiFi显示技术(诸如Miracast[®])通常利用固定长度GOP来对源设备处显示的视频数据进行编码(并随后解码)。在一些方面,至少部分基于源设备(例如,智能电话)视频显示数据在源、类型、内容等方面变化,使用固定长度GOP可降低针对WiFi显示技术(诸如Miracast[®])执行的视频译码的效率。需要可用于执行可变长度GOP视频译码(例如,可变I帧时间间隔视频译码)而不使用计算密集型技术(诸如视频内容分析或运动分析)的系统和技术。还需要可用于以可用于实现WiFi显示共享(例如,Miracast[®])的功率有效方式执行可变长度GOP视频译码(例如,可变I帧时间间隔视频译码)的系统和技术。

[0041] 如本文更详细描述,本文描述了用于使用可变帧内(I帧)时间间隔和/或可变长度图片组(GOP)长度来执行视频译码的系统、装置、方法和计算机可读介质(统称为“系统和技术”)。例如,该系统和技术可使用基于诸如视频帧层信息、视频帧层几何形状等信息确定的可变I帧间隔和/或可变GOP长度来执行视频译码(例如,编码和/或解码)。例如,该系统和技术可获得并分析与待编码的多个帧中所包括的一个或多个层相关联的层信息。在一些方面,待编码的多个帧可由用于执行无线显示共享(例如,Miracast[®])的智能电话或其他移动计算设备捕获。

[0042] 在一些方面,一个或多个层可以是表示合成工作及与显示硬件(例如,与编码设备或其他计算设备相关联的显示器或其他显示硬件)的交互的层基元类型。在一些示例中,层也可被称为合成单元。层可以是表面和SurfaceControl的实例的组合。每个层可具有定义该层如何与其他层交互的特性集合。例如,层特性可包括(但不限于)下面描述的层特性中的一者或多者。

[0043] “位置”层特性可指示该层在其对应显示器上出现在何处(例如,显示器是与层组合起来可表示合成工作以及与显示硬件的交互的另一类型的基元)。“位置”层特性可包括诸如层边缘的位置和该层相对于其他层的z次序(例如,该层是位于其他层的前面还是后面)之类的信息。

[0044] “内容”层特性可指示应如何在层的界限内呈现层上的内容显示(例如,由位置特性给出)。“内容”层特性可包括诸如裁剪信息(例如,扩展内容的一部分以填充层的边界)和变换信息(例如,示出旋转或翻转的内容)的信息。

[0045] “合成”层特性可指示该层应当如何与其他层合成,并且可包括诸如混合模式和用于阿尔法合成的层宽阿尔法值的信息。

[0046] “优化”层特性可指示或以其他方式包括信息,该信息可不直接用于合成该层,而是可由硬件合成器(HWC)用于优化其合成性能。例如,“优化”层特性可包括诸如该层的可见区域和自从前一帧以来该层的哪个(哪些)部分已经被更新的信息。

[0047] 在一些方面,层信息可包括单独层的几何形状,其中与源设备相关联的所捕获的显示数据的每个帧包括一个或多个层。在一些情况下,层信息可包括与所捕获的显示数据的帧的单独层相关联的坐标信息、格式信息等。在一些方面,该系统和技术可使用层信息来确定自适应和/或可变长度GOP(例如,自适应和/或可变I帧时间间隔)。例如,层信息可包括可比基于像素的内容或运动分析更有效地分析的丰富场景信息。基于从层信息确定的丰富

场景信息,该系统和技术可实现场景变化分析、通知和/或I帧触发。

[0048] 例如,可分析来自源设备的所捕获的显示数据的当前帧的层信息并且将其与同来自源设备的所捕获的显示数据的一个或多个先前帧相关联的层信息进行比较。在一些方面,基于确定层信息中所表示的层中的一些(或全部)层已改变大于阈值量或百分比(例如,基于所捕获的显示数据的帧几何形状充分改变),可触发I帧译码并且可将新(例如,可变或自适应)GOP长度应用于新生成的I帧。在一些方面,本文所述的系统和技术可基于如上所述的层信息并进一步基于显示器空闲确定来执行自适应和/或可变长度GOP视频译码(例如,自适应和/或可变I帧时间间隔视频译码)。

[0049] 例如,该系统和技术可确定来自源设备的所捕获的显示数据的帧中所表示的视频内容已转变到空闲状态。在一些方面,空闲状态可与在预先确定的时间段内未检测到内容变化和/或未检测到所捕获的显示数据的帧的额外渲染或刷新相关联。在一些示例中,检测到或确定显示器空闲状态可使该系统和技术应用与显示器空闲状态相关联的新的相对较长的GOP长度。例如,显示器空闲GOP长度可以是300帧,尽管更多或更少数量的帧也可用于显示器空闲GOP长度。

[0050] 在一些示例中,可包括于来自源设备的所捕获的显示数据的给定帧中的每种类型的层可与其自身的GOP长度(例如,以I帧关键帧之间的帧数或时间间隔计)相关联。在一些方面,每种层类型可与不同GOP长度(例如,以I帧关键帧之间的帧数或时间间隔计)相关联。在一些情况下,每种层类型可与其自身的GOP长度相关联,其中层类型中的一者或多者与相同GOP长度值(例如,以I帧关键帧之间的帧数或时间间隔计)相关联。在一些示例中,可预先确定层类型GOP长度中的一者或多者(或全部)。例如,可基于与不同层类型的每个相应层相关联的视频内容的类型和/或运动的类型来预先确定不同层类型GOP长度。在一些方面,可基于与每个给定帧相关联的主要内容层或主要内容层类型来为所捕获的显示数据的多个帧中所包括的帧确定GOP长度。例如,可使用与为所捕获的显示数据的多个帧中的每个给定帧确定的主要内容层相关联的GOP长度来自适应地或可变地确定帧GOP长度。

[0051] 将相对于附图来描述关于系统和技术其他细节。

[0052] 图1是示出包括编码设备104和解码设备112的系统100的示例的框图。编码设备104可以是源设备的一部分,并且解码设备112可以是接收设备的一部分。源设备和/或接收设备可包括电子设备,诸如移动或固定电话手机(例如,智能电话、蜂窝电话等)、台式计算机、膝上型或笔记本计算机、平板计算机、机顶盒、电视机、相机、显示设备、数字媒体播放器、视频游戏控制台、视频流式传输设备、互联网协议(IP)相机或任何其他合适的电子设备。在一些示例中,源设备和接收设备可包括用于无线通信的一个或多个无线收发器。本文所描述的译码技术适用于各种多媒体应用中的视频译码,包括流式视频发送(例如,通过互联网)、电视广播或发送、对数字视频进行编码以用于存储在数据存储介质上、对存储在数据存储介质上的数字视频进行解码、或其他应用。如本文所用,术语译码可指编码和/或解码。在一些示例中,系统100可支持单向或双向视频发送,以支持诸如视频会议、视频流式发送、视频回放、视频广播、游戏和/或视频电话的应用。

[0053] 编码设备104(或编码器)可以用于使用视频译码标准、格式、编解码器或协议来对视频数据进行编码,以生成编码的视频比特流。视频译码标准和格式、编解码器的示例包括ITU-T H.261、ISO/IEC MPEG-1视频、ITU-T H.262或ISO/IEC MPEG-2视频、ITU-T H.263、

ISO/IEC MPEG-4视频、ITU-T H.264(也被称为ISO/IEC MPEG-4AVC)(包括其可缩放视频译码(SVC)和多视图视频译码(MVC)扩展)、高效视频译码(HEVC)或ITU-T H.265和通用视频译码(VVC)或ITU-T H.266。存在对HEVC的各种扩展处理多层视频译码,包括范围和屏幕内容译码扩展、3D视频译码(3D-HEVC)和多视图扩展(MV-HEVC)和可缩放扩展(SHVC)。HEVC及其扩展已由视频译码联合合作小组(JCT-VC)以及ITU-T视频译码专家组(VCEG)和ISO/IEC动画专家组(MPEG)的3D视频译码扩展开发联合合作小组(JCT-3V)开发。VP9、由开放媒体联盟(AOMedia)开发的AOMedia视频1(AV1)和基本视频译码(EVC)是可对其应用本文所描述的技术的其他视频译码标准。

[0054] VVC是由ITU-T和ISO/IEC的联合视频专家组(JVET)开发以针对广泛应用范围至少部分地实现超出HEVC的高压缩能力的最新视频译码标准。VVC规范在2020年7月完成并由ITU-T和ISO/IEC两者发布。VVC规范指定标准比特流和图片格式、高级语法(HLS)和译码单元级语法、解析过程、解码过程等。VVC还在附件中指定配置文件/层次/层级(PTL)限制、字节流格式、假设参考解码器和补充增强信息(SEI)。

[0055] 本文所描述的系统和技术可以应用于现有视频编解码器(例如,VVC、HEVC、AVC或其他合适的现有视频编解码器)中的任一者,并且/或者可以是用于正在开发的任何视频译码标准和/或未来视频译码标准的有效译码工具。例如,可使用诸如VVC、HEVC、AVC和/或它们的扩展之类的视频编解码器来执行本文所描述的示例。然而,本文所描述的技术和系统也可以适用于其他译码标准、编解码器或格式,诸如MPEG、JPEG(或用于静态图像的其他译码标准)、VP9、AV1、它们的扩展、或已经可用或尚不可用或尚未开发的其他合适的译码标准。例如,在一些示例中,编码设备104和/或解码设备112可以根据专有视频编解码器/格式(诸如AV1、AVI的扩展和/或AV1的后继版本(例如,AV2))或其他专有格式或工业标准来操作。因此,尽管可参考特定的视频译码标准来描述本文所述的技术和系统,但是本领域的普通技术人员将明白的是,不应当将描述解释为仅适用于该特定的标准。

[0056] 参考图1,视频源102可以将视频数据提供给编码设备104。视频源102可以是源设备的一部分,或者可以是除了源设备之外的设备的一部分。视频源102可以包括视频捕获设备(例如,摄像机、照相电话、视频电话等),包括被存储的视频的视频存档、提供视频数据的视频服务器或内容提供商、从视频服务器或内容提供者接收视频的视频馈送接口、用于生成计算机图形视频数据的计算机图形系统、此类源的组合、或任何其他合适的视频源。

[0057] 来自视频源102的视频数据可以包括一个或多个输入图片或帧。图片或帧是静态图像,在一些情况下其是视频的一部分。在一些示例中,来自视频源102的数据可以是不作为视频的一部分的静态图像。在HEVC、VVC和其他视频译码规范中,视频序列可以包括一系列图片。图片可以包括三个样本阵列,其被表示为SL、SCb和SCr。SL是亮度样本的二维阵列,SCb是Cb色度样本的二维阵列,并且SCr是Cr色度样本的二维阵列。色度(chrominance)样本在本文中也可被称为“chroma(色度)”样本。像素可指图片的阵列中的给定位置的所有三个分量(亮度样本和色度样本)。在其他情况下,图片可以是单色的并且可仅包括亮度样本阵列,在这种情况下,术语像素和样本可互换地使用。关于出于例示性目的而引用各个样本的本文所描述的示例技术,可将相同技术应用于像素(例如,图片的阵列中的给定位置的所有三个样本分量)。关于出于例示性目的而引用像素(例如,图片的阵列中的给定位置的所有三个样本分量)的本文所描述的示例技术,可将相同技术应用于各个样本。

[0058] 编码设备104的编码器引擎106(或编码器)对视频数据进行编码以生成编码的视频比特流。在一些示例中,编码的视频比特流(或“视频比特流”或“比特流”)是一系列的一个或多个译码的视频序列。译码的视频序列(CVS)包括一系列访问单元(AU),这一系列访问单元从在基本层中具有随机访问点图片并且具有某些属性的AU开始,直到在基本层中具有随机访问点图片并且具有某些属性的下一AU并且不包括该下一AU。例如,开始CVS的随机访问点图片的某些属性可以包括等于1的RASL标记(例如,NoRaslOutputFlag)。否则,随机访问点图片(具有等于0的RASL标记)不开始CVS。访问单元(AU)包括一个或多个译码的图片以及与共享相同的输出时间的译码的图片相对应的控制信息。图片的译码的切片在比特流级别被封装为数据单元,该数据单元被称为网络抽象层(NAL)单元。例如,HEVC视频比特流可以包括一个或多个CVS,其包括NAL单元。NAL单元中的每个NAL单元具有NAL单元报头。在一个示例中,该报头对于H.264/AVC为一个字节(多层扩展除外)并且对于HEVC为两个字节。NAL单元报头中的语法元素采用指定的比特,并且因此对于所有种类的系统 and 传输层(诸如传输流、实时传输(RTP)协议、文件格式等)都是可见的。

[0059] 在HEVC标准中存在两类NAL单元,包括视频译码层(VCL)NAL单元和非VCL NAL单元。VCL NAL单元包括译码的图片数据的一个切片或切片段(下文描述),并且非VCL NAL单元包括与一个或多个译码的图片有关的控制信息。在一些情况下,NAL单元可被称为分组。HEVC AU包括包含译码的图片数据的VCL NAL单元和与译码的图片数据相对应的非VCL NAL单元(如果有的话)。

[0060] NAL单元可以包括形成对视频数据的译码的表示的比特序列(例如,编码的视频比特流、比特流的CVS等),诸如对在视频中的图片的译码的表示。编码器引擎106通过将每个图片分割为多个切片来生成对图片的译码的表示。切片是独立于其他切片的,使得可以对该切片中的信息进行译码,而不依赖于来自相同的图片内的其他切片的数据。切片包括一个或多个切片段,其包括独立的切片段以及(如果存在的话)依赖于先前切片段的一个或多个从属切片段。将切片分割成亮度样本和色度样本的译码树块(CTB)。亮度样本的CTB和色度样本的一个或多个CTB连同用于样本的语法一起被称为译码树单元(CTU)。CTU还可被称为“树块”或“最大译码单元”(LCU)。CTU是用于HEVC编码的基本处理单元。CTU可以被拆分为大小不同的多个译码单元(CU)。CU包含被称为译码块(CB)的亮度样本阵列和色度样本阵列。

[0061] 亮度CB和色度CB可以还被拆分为预测块(PB)。PB是亮度分量或色度分量的样本块,其使用相同的运动参数来进行帧间预测或帧内块复制预测(当可用或被启用时)。亮度PB和一个或多个色度PB连同相关联的语法一起形成预测单元(PU)。对于帧间预测,运动参数集合(例如,一个或多个运动矢量、参考索引等)是在用于每个PU的比特流中用信号通知的,以及用于亮度PB和一个或多个色度PB的帧间预测。运动参数也可以被称为运动信息。CB也可以被分割为一个或多个变换块(TB)。TB表示颜色分量的样本的正方形块,其中对其应用了残差变换(例如,在一些情况下,相同的二维变换)以对预测残差信号进行译码。变换单元(TU)表示亮度样本和色度样本的TB以及相应的语法元素。

[0062] CU的大小对应于译码模式的大小并且可以是正方形形状。例如,CU的大小可以是8x8个样本、16x16个样本、32x32个样本、64x64个样本、或高达对应的CTU的大小的任何其他适当的大小。短语“NxN”在本文中用于指代视频块在垂直维度和水平维度方面的像素尺寸

(例如,8个像素x8个像素)。块中的像素可以是按行和列来排列的。在一些示例中,块在水平方向上可不具有与垂直方向上相同数量的像素。与CU相关联的语法数据可以描述例如将CU分割为一个或多个PU。分割模式可能根据CU是帧内预测模式编码的还是帧间预测模式编码的而有所不同。PU可以被分割为非正方形形状。与CU相关联的语法数据还可以例如描述根据CTU将CU分割为一个或多个TU。TU可以是正方形或非正方形形状。

[0063] 根据HEVC标准,可使用变换单元(TU)来执行变换。TU可针对不同CU而变化。可以基于给定CU内的PU的大小来设定TU的大小。TU可以具有与PU相同的大小或者小于PU。在一些示例中,可以使用被称为残差二叉树(RQT)的四叉树结构来将与CU相对应的残差样本细分为更小的单元。RQT的叶节点可以对应于TU。与TU相关联的像素差值可以被变换以产生变换系数。可由编码器引擎106对变换系数进行量化。

[0064] 一旦视频数据的图片被分割为CU,编码器引擎106就使用预测模式来预测每个PU。从原始视频数据中减去预测单元或预测块以获得残差(下文描述)。对于每个CU,可以使用语法数据在比特流内用信号通知预测模式。预测模式可以包括帧内预测(或图片内预测)或帧间预测(或图片间预测)。帧内预测利用在图片内的空间上相邻的样本之间的相关性。例如,使用帧内预测,每个PU是使用例如DC预测来从相同的图片中的相邻图像数据进行预测的以找到用于PU的平均值,使用平面预测以使平面表面适配PU,使用方向预测来从相邻数据进行推断,或者使用任何其他合适的预测类型。帧间预测使用在图片之间的时间相关性,以便推导出用于图像样本块的运动补偿预测。例如,使用帧间预测,每个PU是使用运动补偿预测从一个或多个参考图片(按照输出次序在当前图片之前或之后)中的图像数据预测的。例如,可以在CU级别处作出是使用图片间预测还是使用图片内预测来对图片区域进行译码的决策。

[0065] 编码器引擎106和解码器引擎116(下文更加详细地描述)可被配置为根据VVC进行操作。根据VVC,视频译码器(诸如编码器引擎106和/或解码器引擎116)将图片分割为多个译码树单元(CTU)(其中亮度样本的CTB和色度样本的一个或多个CTB连同用于样本的语法一起被称为CTU)。视频译码器可以根据树结构(诸如四叉树-二叉树(QTBT)结构或多类型树(MTT)结构)来分割CTU。QTBT结构去除多种分区类型的概念,诸如HEVC的CU、PU和TU之间的分隔。QTBT结构包括两个级别,包括根据四叉树分割进行分割的第一级别、以及根据二叉树分割进行分割的第二级别。QTBT结构的根节点对应于CTU。二叉树的叶节点对应于译码单元(CU)。

[0066] 在MTT分割结构中,可以使用四叉树分割、二叉树分割以及一种或多种类型的三叉树分割来对块进行分割。三叉树分割是其中块被拆分为三个子块的分割。在一些示例中,三叉树分割将块划分为三个子块,而不通过中心划分原始块。MTT中的分割类型(例如,四叉树、二叉树和三叉树)可以是对称的或不对称的。

[0067] 当根据AV1编解码器操作时,编码设备104和解码设备112可被配置为以块为单位对视频数据进行译码。在AV1中,能够处理的最大译码块被称为超级块。在AV1中,超级块可以是 128×128 亮度样本或 64×64 亮度样本。然而,在后续视频译码格式(例如,AV2)中,超级块可由不同(例如,更大)的亮度样本大小来定义。在一些示例中,超级块是块四叉树的顶层。编码设备104可进一步将超级块分割成较小的译码块。编码设备104可使用正方形或非正方形分割将超级块和其他译码块分割成较小的块。非正方形块可包括 $N/2 \times N$ 块、 $N \times N/2$

块、 $N/4 \times N$ 块和 $N \times N/4$ 块。编码设备104和解码设备112可对每个译码块执行单独的预测和变换处理。

[0068] AV1还定义了视频数据的图块。图块是可以独立于其他图块进行译码的超级块的矩形阵列。也就是说,编码设备104和解码设备112可在不使用来自其他图块的视频数据的情况下分别对图块内的译码块进行编码和解码。然而,编码设备104和解码设备112可跨越图块边界执行滤波。图块的大小可以是均匀的或不均匀的。基于图块的译码可实现编码器和解码器具体实施的并行处理和/或多线程。

[0069] 在一些示例中,编码设备104和解码设备112可以使用单个QTBT或MTT结构来表示亮度分量和色度分量中的每一者,而在其他示例中,视频译码器可以使用两个或更多个QTBT或MTT结构,诸如用于亮度分量的一个QTBT或MTT结构以及用于两个色度分量的另一QTBT或MTT结构(或者用于相应色度分量的两个QTBT和/或MTT结构)。

[0070] 编码设备104和解码设备112可被配置为使用每HEVC的四叉树分割、QTBT分割、MTT分割、或其他分割结构。

[0071] 在一些示例中,给图片的一个或多个切片指派切片类型。切片类型包括I切片、P切片和B切片。I切片(帧内,可独立解码)是仅通过帧内预测进行译码的图片的切片,并且因此可独立解码,因为I切片仅需要帧内的数据来预测切片的任何预测单元或预测块。P切片(单向预测帧)是可用帧内预测以及用单向帧间预测进行译码的图片的切片。用帧内预测或帧间预测对P切片内的每个预测单元或预测块进行译码。当应用帧间预测时,预测单元或预测块仅由一个参考图片预测,并且因此参考样本仅来自一个帧的一个参考区。B切片(双向预测性帧)是可用帧内预测和帧间预测(例如,双向预测或单向预测任一者)进行译码的图片的切片。可从两个参考图片双向地预测B切片的预测单元或预测块,其中每个图片贡献于一个参考区域并且对这两个参考区域的样本集合进行加权(例如,用相等权重或用不同权重)以产生双向预测块的预测信号。如上所述,一个图片的切片被独立地译码。在一些情况下,可将图片译码为仅一个切片。

[0072] 如上文所指出,图片内预测利用图片内的在空间上相邻的样本之间的相关性。存在多种帧内预测模式(也被称为“帧内模式”)。在一些示例中,亮度块的帧内预测包括35种模式,其包括平面模式、DC模式和33种角度模式(例如,对角线帧内预测模式和与对角线帧内预测模式相邻的角度模式)。编码设备104和/或解码设备112可以针对每个块选择使预测块与待编码的块之间的残差最小化的预测模式(例如,基于绝对误差总和(SAE)、绝对差总和(SAD)、绝对变换差总和(SATD)或类似性的其他量度)。例如,可以通过获得待编码的块中的每个像素(或样本)与用于比较的预测块中的对应像素(或样本)之间的绝对差来计算SAE。将像素(或样本)的差求和以产生块类似性的度量,诸如差图像的L1范数、两个图像块之间的曼哈顿距离或其他计算。使用SAE作为示例,用于使用帧内预测模式中的每个帧内预测模式的每个预测的SAE指示预测误差的量值。具有与实际当前块的最佳匹配的帧内预测模式由给出最小SAE的帧内预测模式给出。

[0073] 为帧内预测的35个模式编索引,如下表1所示。在其他示例中,可以定义更多的帧内模式,包括可能尚未由33种角度模式表示的预测角度。在其他示例中,与角度模式相关联的预测角度可以不同于在HEVC中使用的那些预测角度。

[0074]

帧内预测模式	相关联的名称
--------	--------

0	INTRA_PLANAR
1	INTRA_DC
2..34	INTRA_ANGULAR2..INTRA_ANGULAR34

[0075] 表1-帧内预测模式和相关联的名称的规范

[0076] 为了对 $N \times N$ 块执行平面预测,对于位于 (x, y) 处的每个样本 p_{xy} ,可通过将双线性滤波器应用于四个特定相邻重构样本(用作帧内预测的参考样本)来计算预测样本值。四个参考样本包括右上重构样本TR、左下重构样本BL以及位于当前样本的相同列 $(r_{x, -1})$ 和行 $(r_{-1, y})$ 处的两个重构样本。平面模式可用公式表示如下:

$$[0077] \quad p_{xy} = ((N-x1) * L + (N-y1) * T + x1 * R + y1 * B) / (2 * N),$$

[0078] 其中 $x1 = x + 1, y1 = y + 1, R = TR$ 并且 $B = BL$ 。

[0079] 对于DC模式,用相邻重构样本的平均值填充预测块。一般来讲,平面模式和DC模式都应用于对平滑变化和恒定的图像区域进行建模。

[0080] 对于HEVC中的角度帧内预测模式(其包括33个不同预测方向),帧内预测过程可描述如下。对于每个给定角度帧内预测模式,可相应地标识帧内预测方向;例如,帧内模式18对应于纯水平预测方向,并且帧内模式26对应于纯竖直预测方向。在图2A的示例示图200a中示出了角度预测模式。在一些编解码器中,可使用不同数量的帧内预测模式。例如,除了平面模式和DC模式之外,还可定义93个角度模式,其中模式2指示 -135° 的预测方向,模式34指示 -45° 的预测方向,并且模式66指示 45° 的预测方向。在一些编解码器(例如,VVC)中,也可定义超过 -135° (小于 -135°)和超过 45° (大于 45°)的角度;这些可被称为广角帧内模式。尽管本文的描述是关于HEVC中的帧内模式设计(例如,具有35个模式),但所公开的技术还可应用于更多帧内模式(例如,由VVC或其他编解码器定义的帧内模式)。

[0081] 预测块的每个样本的坐标 (x, y) 沿着特定帧内预测方向(例如,角度帧内预测模式中的一者)投影。例如,在给定特定帧内预测方向的情况下,预测块的样本的坐标 (x, y) 首先沿着帧内预测方向投影到相邻重构样本的行/列。在 (x, y) 投影到两个相邻重构样本L和R之间的分数位置 α 的情况下;则可使用两抽头双线性内插滤波器来计算 (x, y) 的预测值,用公式表示如下:

$$[0082] \quad p_{xy} = (1 - \alpha) \cdot L + \alpha \cdot R$$

[0083] 为了避免浮点运算,在HEVC中,可使用整数算术将以上计算估计为:

$$[0084] \quad p_{xy} = ((32 - a') \cdot L + a' \cdot R + 16) \gg 5,$$

[0085] 其中 a' 是等于 $32 * a$ 的整数。

[0086] 在一些示例中,在帧内预测之前,使用2抽头双线性或3抽头(1, 2, 1)/4滤波器来对相邻参考样本进行滤波,这可被称为帧内参考平滑或模式依赖帧内平滑(MDIS)。当执行帧内预测时,在给定帧内预测模式索引(predModeIntra)和块大小(nTbS)的情况下,决定是否执行参考平滑过程以及使用哪个平滑滤波器。帧内预测模式索引是指示帧内预测模式的索引。

[0087] 图片间预测使用在图片之间的时间相关性,以便推导出用于图像样本块的运动补偿预测。使用平移运动模型,先前解码的图片(参考图片)中的块的位置由运动矢量 $(\Delta x, \Delta y)$ 指示,其中 Δx 指定参考块相对于当前块的位置的水平位移,并且 Δy 指定参考块相对于当前块的位置的竖直位移。在一些情况下,运动矢量 $(\Delta x, \Delta y)$ 可具有整数样本准确度(也

称为整数准确度),在这种情况下,运动矢量指向参考帧的整数像素栅格(或整数像素采样栅格)。在一些情况下,运动矢量($\Delta x, \Delta y$)可具有分数样本准确度(也称为分数像素准确度或非整数准确度)以更准确地捕获下层对象的移动,而限于参考帧的整数像素栅格。运动矢量的准确度可以由运动矢量的量化水平来表示。例如,量化水平可以是整数准确度(例如,1像素)或分数像素准确度(例如,1/4像素、1/2像素或其他子像素值)。当对应的运动矢量具有分数采样准确度时,对参考图片应用插值以导出预测信号。例如,可对整数位置处可用的样本进行滤波(例如,使用一个或多个插值滤波器)以估计分数位置处的值。通过参考图片列表的参考索引(refIdx)来指示先前解码的参考图片。运动矢量和参考索引可被称为运动参数。可执行两种图片间预测,包括单向预测和双向预测。

[0088] 在使用双向预测的帧间预测(也被称为双向帧间预测)的情况下,使用两组运动参数($\Delta x_0, y_0, \text{refIdx}_0$ 和 $\Delta x_1, y_1, \text{refIdx}_1$)来生成两个运动补偿预测(从同一参考图片或者可能从不同的参考图片)。例如,在双向预测的情况下,每个预测块使用两个运动补偿预测信号,并且生成B个预测单元。将两个运动补偿预测进行组合以获得最终的运动补偿预测。例如,可以通过进行平均来组合两个运动补偿预测。在另一示例中,可以使用加权预测,在这种情况下,可以将不同的权重应用于每个运动补偿预测。可以在双向预测中使用的参考图片被存储在两个单独的列表中,分别被表示为列表0和列表1。可以在编码设备104处使用运动估计过程来推导运动参数。

[0089] 在使用单向预测的帧间预测(也被称为单向帧间预测)的情况下,使用一组运动参数($\Delta x_0, y_0, \text{refIdx}_0$)从参考图片生成运动补偿预测。例如,在单向预测的情况下,每个预测块最多使用一个运动补偿预测信号,并且生成P个预测单元。

[0090] PU可以包括与预测过程相关的数据(例如,运动参数或其他合适的数据)。例如,当PU使用帧内预测来编码时,PU可以包括描述用于PU的帧内预测模式的数据。作为另一示例,当PU使用帧间预测来编码时,PU可以包括定义用于PU的运动矢量的数据。定义用于PU的运动矢量的数据可描述例如运动矢量的水平分量(Δx)、运动矢量的垂直分量(Δy)、用于运动矢量的分辨率(例如,整数精度、四分之一像素精度或八分之一像素精度)、运动矢量所指向的参考图片、参考索引、用于运动矢量的参考图片列表(例如,列表0、列表1或列表C)、或它们的任何组合。

[0091] AV1包括用于对视频数据的译码块进行编码和解码的两种通用技术。这两种通用技术是帧内预测(例如,帧内预测或空间预测)和帧间预测(例如,帧间预测或时间预测)。在AV1的内容中,当使用帧内预测模式来预测视频数据的当前帧的块时,编码设备104和解码设备112不使用来自视频数据的其他帧的视频数据。对于大多数帧内预测模式,视频编码设备104基于当前块中的样本值与从同一帧中的参考样本生成的预测值之间的差来编码当前帧的块。视频编码设备104基于帧内预测模式确定从参考样本生成的预测值。

[0092] 在使用帧内预测和/或帧间预测执行预测之后,编码设备104可执行变换和量化。例如,在预测之后,编码器引擎106可计算与PU相对应的残差值。残差值可以包括在正被译码的当前像素块(PU)与用于预测当前块的预测块(例如,当前块的预测版本)之间的像素差值。例如,在生成预测块(例如,使用帧间预测或帧内预测)之后,编码器引擎106可以通过从当前块中减去由预测单元生成的预测块来生成残差块。残差块包括像素差值集合,其对当前块的像素值与预测块的像素值之间的差进行量化。在一些示例中,可以用二维块格式(例

如,像素值的二维矩阵或阵列)来表示残差块。在此类示例中,残差块是对像素值的二维表示。

[0093] 使用块变换来对在执行预测之后可能剩余的任何残差数据进行变换,块变换可以基于离散余弦变换、离散正弦变换、整数变换、小波变换、其他合适的变换函数、或它们的任何组合。在一些情况下,可以将一个或多个块变换(例如,大小为 32×32 、 16×16 、 8×8 、 4×4 或其他合适的大小)应用于每个CU中的残差数据。在一些示例中,可将TU用于由编码器引擎106实现的变换和量化过程。具有一个或多个PU的给定CU也可以包括一个或多个TU。如下文进一步详细描述,残差值可以使用块变换而被变换为变换系数,以及可以使用TU进行量化和扫描,以产生用于熵译码的串行化变换系数。

[0094] 在一些示例中,在使用CU的PU进行帧内预测或帧间预测译码之后,编码器引擎106可以计算用于CU的TU的残差数据。PU可以包括空间域(或像素域)中的像素数据。TU可以包括在应用块变换之后在变换域中的系数。如先前所指出,残差数据可以对应于在未编码的图像的像素与对应于PU的预测值之间的像素差值。编码器引擎106可形成包括用于CU的残差数据的TU,并且可对TU进行变换以产生用于CU的变换系数。

[0095] 编码器引擎106可以执行对变换系数的量化。量化通过对变换系数进行量化以减少用于表示系数的数据量来提供进一步的压缩。例如,量化可以减小与系数中的一些或所有系数相关联的比特深度。在一个示例中,具有n比特的值的系数可以在量化期间向下舍入为m比特的值,其中n大于m。

[0096] 一旦执行了量化,则译码的视频比特流包括量化的变换系数、预测信息(例如,预测模式、运动矢量、块矢量等)、分割信息以及任何其他合适的数据(诸如其他语法数据)。译码的视频比特流的不同元素可由编码器引擎106进行熵编码。在一些示例中,编码器引擎106可以利用预定义的扫描次序来扫描所量化的变换系数,以产生可以被熵编码的串行化矢量。在一些示例中,编码器引擎106可执行自适应扫描。在扫描所量化的变换系数以形成矢量(例如,一维矢量)之后,编码器引擎106可以对该矢量进行熵编码。例如,编码器引擎106可使用上下文自适应可变长度译码、上下文自适应二元算术译码、基于语法的上下文自适应二元算术译码、概率区间分割熵译码或另一合适的熵编码技术。

[0097] 编码设备104的输出端110可在通信链路120上将构成编码的视频比特流数据的NAL单元传送到接收设备的解码设备112。解码设备112的输入端114可以接收NAL单元。通信链路120可包括由无线网络、有线网络或有线网络和无线网络的组合提供的信道。无线网络可包括任何无线接口或无线接口的组合,并且可包括任何合适的无线网络(例如,互联网或其他广域网、基于分组的网络、WiFiTM、射频(RF)、UWB、WiFi直连、蜂窝、长期演进(LTE)、WiMaxTM等)。有线网络可以包括任何有线接口(例如,光纤、以太网、电力线以太网、基于同轴电缆的以太网、数字信号线(DSL)等)。有线网络和/或无线网络可以使用各种装备(诸如基站、路由器、接入点、网桥、网关、交换机等)来实现。编码的视频比特流数据可以根据通信标准(诸如无线通信协议)进行调制,并且被发送给接收设备。

[0098] 在一些示例中,编码设备104可将编码的视频比特流数据存储于存储装置108中。输出端110可以从编码器引擎106或从存储装置108检索编码的视频比特流数据。存储装置108可包括各种分布式或本地访问的数据存储介质中的任何一种。例如,存储装置108可以包括硬盘驱动器、存储盘、闪存存储器、易失性或非易失性存储器、或用于存储编码的视频

数据的任何其他合适的数字存储介质。存储装置108还可以包括用于存储供在帧间预测中使用的参考图片的所解码的图片缓冲器 (DPB)。在进一步的示例中,存储装置108可以对应于文件服务器或可以存储由源设备生成的编码的视频的另一中间存储设备。在此类情况下,包括解码设备112的接收设备可经由流式传输或下载来从存储设备访问所存储的视频数据。文件服务器可以是能够存储编码的视频数据并且将编码的视频数据发送给接收设备的任何类型的服务器。示例文件服务器包括web服务器(例如,用于网站)、FTP服务器、网络附加存储(NAS)设备或本地磁盘驱动器。接收设备可通过任何标准数据连接(包括互联网连接)访问编码的视频数据,并且可包括适于访问被存储在文件服务器上的编码的视频数据的无线信道(例如,Wi-Fi连接)、有线连接(例如,DSL、电缆调制解调器等)、或这两者的组合。编码的视频数据从存储装置108的发送可以是流式发送、下载发送或它们的组合。

[0099] 解码设备112的输入端114接收编码的视频比特流数据,以及可将视频比特流数据提供给解码器引擎116,或者提供给存储装置118以供解码器引擎116稍后使用。例如,存储装置118可包括用于存储在帧间预测中使用的参考图片的DPB。包括解码设备112的接收设备可以经由存储装置108来接收要被解码的编码的视频数据。编码的视频数据可以根据通信标准(诸如无线通信协议)进行调制,并且被发送给接收设备。用于发送编码的视频数据的通信介质可以包括任何无线或有线通信介质,诸如射频(RF)频谱或一条或多条物理发送线。通信介质可以形成诸如以下各项的基于分组的网络的一部分:局域网、广域网或全球网络(诸如互联网)。通信介质可以包括路由器、交换机、基站、或对于促进从源设备到接收设备的通信而言可以有用的任何其他装备。

[0100] 解码器引擎116可以通过熵解码(例如,使用熵解码器)并且提取构成编码的视频数据的一个或多个编码的视频序列的元素,从而对编码的视频比特流数据进行解码。解码器引擎116可重新缩放编码的视频比特流数据并且对其执行逆变换。残差数据被传递到解码器引擎116的预测级。解码器引擎116预测像素块(例如,PU)。在一些示例中,将预测与逆变换的输出(残差数据)相加。

[0101] 解码设备112可将解码的视频输出到视频目的地设备122,该视频目的地设备可包括用于将解码的视频数据显示给内容的消费者的显示器或其他输出设备。在一些方面,视频目的地设备122可以是包括解码设备112的接收设备的一部分。在一些方面,视频目的地设备122可以是不同于接收设备的单独的设备的一部分。

[0102] 在一些示例中,视频编码设备104和/或视频解码设备112可分别与音频编码设备和音频解码设备集成。视频编码设备104和/或视频解码设备112还可以包括用于实现上文描述的译码技术所必需的其他硬件或软件,诸如一个或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、分立逻辑、软件、硬件、固件或它们的任何组合。视频编码设备104和视频解码设备112可以被集成为相应的设备中的组合的编码器/解码器(编解码器)的一部分。下文参考图7描述了编码设备104的具体细节的示例。下文参考图8描述了解码设备112的具体细节的示例。

[0103] 在图1中示出的示例系统是在本文中使用的的一个例示性示例。用于使用本文所描述的技术来处理视频数据的技术可以由任何数字视频编码和/或解码设备来执行。尽管一般来讲本公开的技术由视频编码设备或视频解码设备执行,但这些技术也可由组合的视频编码器-解码器(通常称为“编解码器”)执行。此外,本公开的技术还可由视频预处理器

执行。源设备和接收设备仅仅是此类译码设备的示例,其中,源设备生成译码的视频数据以用于发送给接收设备。在一些示例中,源设备和接收设备可以以基本上对称的方式进行操作,使得这些设备中的每个设备包括视频编码和解码组件。因此,示例系统可以支持在视频设备之间的单向或双向视频发送,例如,以用于视频流式传输、视频回放、视频广播或视频电话。

[0104] 如先前所描述,HEVC比特流包括NAL单元组,包括VCL NAL单元和非VCL NAL单元。VCL NAL单元包括形成译码的视频比特流的译码的图片数据。例如,形成译码的视频比特流的比特序列存在于VCL NAL单元中。除了其他信息之外,非VCL NAL单元还可以包含具有与编码的视频比特流有关的高级信息的参数集。例如,参数集可以包括视频参数集(VPS)、序列参数集(SPS)和图片参数集(PPS)。参数集的目标的示例包括比特率效率、容错性以及提供系统层接口。每个切片引用单个活动PPS、SPS和VPS以访问解码设备112可用于对切片进行解码的信息。可针对每个参数集对标识符(ID)进行译码,包括VPS ID、SPS ID和PPS ID。SPS包括SPS ID和VPS ID。PPS包括PPS ID和SPS ID。每个切片报头包括PPS ID。使用ID,可针对给定切片标识有效参数集。

[0105] PPS包括应用于给定图片中的所有切片的信息。在一些示例中,图片中的所有切片引用同一PPS。不同图片中的切片也可引用同一PPS。SPS包括应用于同一译码的视频序列(CVS)或比特流中的所有图片的信息。如先前所描述,译码的视频序列是一系列访问单元(AU),这一系列访问单元从在基础层中并且具有某些属性(上文所描述)的随机访问点图片(例如,瞬时解码参考(IDR)图片或断链访问(BLA)图片,或其他适当的随机访问点图片)开始,直到在基础层中具有随机访问点图片并且具有某些属性的下一AU(或比特流的末端)并且不包括该下一AU。SPS中的信息在译码的视频序列内可不随图片而改变。译码的视频序列中的图片可使用同一SPS。VPS包括应用于译码的视频序列或比特流内的所有层的信息。VPS包括具有应用于整个译码的视频序列的语法元素的语法结构。在一些示例中,VPS、SPS或PPS可与编码的比特流一起在带内发送。在一些示例中,VPS、SPS或PPS可在与包含译码的视频数据的NAL单元分开的发送中带外发送。

[0106] 本公开通常可以涉及“发信号通知”某些信息(诸如语法元素)。术语“发信号通知”通常可以指代对用于对编码视频数据进行解码的语法元素的值和/或其他数据的传达。例如,视频编码设备104可以用信号通知比特流中语法元素的值。通常,发信号通知是指在比特流中生成值。如上文所指出,视频源102可以基本上实时地或不实时地(诸如可能在将语法元素存储到存储装置108以供视频目的地设备122稍后检索时发生)将比特流传输到视频目的地设备122。

[0107] 视频比特流还可包括补充增强信息(SEI)消息。例如,SEI NAL单元可以是视频比特流的一部分。在一些情况下,SEI消息可包含解码过程不需要的信息。例如,SEI消息中的信息对于解码器解码比特流的视频图片来说可能不是必要的,但是解码器可以使用该信息来改进图片(例如,所解码的输出)的显示或处理。SEI消息中的信息可以是嵌入的元数据。在一个示例性示例中,SEI消息中的信息可被解码器侧实体用来改进内容的可视性。在一些实例中,某些应用标准可指示此类SEI消息在比特流中的存在,使得可为符合应用标准的所有设备带来质量的改进(例如,承载用于帧兼容平面立体3DTV视频格式的帧封装SEI消息,其中针对视频的每个帧承载SEI消息;处理恢复点SEI消息;使用DVB中的拉移式扫描矩形

SEI消息;以及许多其他示例)。

[0108] 如上文所指出,编码设备104可通过使用帧内预测(intra-prediction)和/或帧内预测(intra-frame prediction)来对原始视频序列的图片的一个或多个块或矩形区域进行编码以移除空间冗余。解码设备112可通过使用与编码设备104所使用的相同帧内预测模式来对编码的块进行解码。帧内预测模式描述了用于基于参考像素值来计算正被译码的区域中的像素值的不同变体或方法。在VVC标准中,可基于帧内预测模式来选择一个或多个平滑滤波器和内插滤波器,随后将其应用于参考像素和/或当前块的帧内预测。在该方法中,将用于帧内预测的平滑滤波器与内插滤波器之间的相同选择应用于所有块大小,例如,将固定程度的平滑应用于所有可能块大小。在VVC标准中提供不同的定向帧内预测模式。

[0109] 图2B例示了VVC中的定向帧内预测模式(也被称为“角度帧内预测模式”)的示例图200b。在一些示例中,平面模式和DC模式在VVC中保持与HEVC中相同。如所例示,具有2与66之间的偶数索引的帧内预测模式可等同于33个HEVC帧内预测模式,其中图2B的剩余帧内预测模式表示VVC中新添加的帧内预测模式。作为一个例示性示例,为了更好地捕获自然视频中呈现的任意边缘方向,将VTM5(VVC测试模型5)中的定向帧内预测模式的数量从33个HEVC方向增加到总共93个方向。帧内预测模式更详细地描述于B. Bross, J. Chen, S. Liu, “Versatile Video Coding (Draft 10)”, 第19届JVET会议, 远程会议, 2020年7月, JVET-S2001中, 该文献据此全文以引用的方式并入以用于所有目的。在一些示例中,在VVC标准中引入的较密集定向帧内预测模式可应用于所有块大小并应用于亮度帧内预测和色度帧内预测两者。在一些情况下,这些定向帧内预测模式可与多个参考线(MRL)和/或与帧内子划分模式(ISP)组合使用。进一步的细节描述于J. Chen, Y. Ye, S. Kim, “Algorithm description for Versatile Video Coding and Test Model 10 (VTM10)”, 第19届JVET会议, 远程会议, 2020年7月, JVET-S2002中, 该文献据此全文以引用的方式并入以用于所有目的。

[0110] 如先前所述,可使用帧内预测帧(I帧)、预测帧(P帧)和/或双向帧(B帧)的组合来执行视频译码。在一些方面,I帧可用作用于对视频数据的多个帧进行译码(例如,编码或解码)的关键帧。例如,I帧可用作用于对所捕获的视频显示数据的多个帧进行编码或解码的关键帧,如下文将更深入地描述。

[0111] 图3是例示可用于对视频数据的多个帧进行译码(例如,编码和/或解码)的图片组(GOP)长度和帧间(I帧)时间间隔的示例的示图。例如,图3描绘了包括第一I帧310和第二I帧320的视频数据的多个帧300。在一些方面,I帧310和320可用作用于对顺序位于I帧之前或之后的一个或多个B帧和/或一个或多个P帧进行译码的关键帧。例如,I帧310和320可用作也包括在视频数据的多个帧300中的多个P帧和B帧的关键帧。在一些方面,基于I帧310和320仅包括以相同I帧内的其他视频数据块为参考的视频数据块,I帧310和320可用作关键帧。

[0112] 在一些方面,可生成P帧以参考先前编码的I帧关键帧。例如,可生成图3中所描绘的四个P帧中的一者或多者(或全部)以将第一I帧310作为关键帧来参考(例如,基于在视频数据的多个帧300中所包括的四个P帧中的任一者之前生成或编码第一I帧310)。在一些示例中,可生成图3中所描绘的十个B帧中的一者或多者(或全部)以将第一I帧310作为关键帧来参考,可生成图3中所描绘的十个B帧中的一者或多者(或全部)以将第二I帧320作为关键

帧来参考,或可生成图3中所描绘的十个B帧中的一者或多者(或全部)以将第一I帧310和第二I帧320两者作为关键帧来参考。

[0113] 两个关键帧之间(例如,两个I帧310、320之间)的距离可被称为图片组(GOP)长度和/或I帧时间间隔。例如,多个帧300包括第一I帧310、十四个中间帧(例如,每个中间帧是P帧或B帧)以及第二I帧320。在一些方面,与多个帧300相关联的开放GOP长度为15帧。例如,可将15帧的开放GOP长度确定为十四个P帧和B帧加上第一I帧310(例如,第一I帧310及十四个P帧和B帧包括长度为 $1+14=15$ 帧的图片组(GOP))。在一些方面,可将15帧的开放GOP长度确定为十四个P帧和B帧加上下一I帧(例如,第二I帧320)。例如,可使用15帧的开放GOP长度来指示第一I帧310(例如,第一关键帧)之后的下一I帧(例如,关键帧)将位于第一I帧310之后的15帧处。在一些情况下,可将封闭GOP长度确定为位于两个I帧之间的中间帧(例如,P帧和/或B帧)的数量,其中封闭GOP长度是排除与计算相关联的相应I帧所占据的帧而确定的。例如,多个帧300可与14帧的封闭GOP长度相关联。

[0114] 如先前所提及,GOP长度可测量一系列视频帧(例如,诸如多个视频帧300)中的两个关键帧(例如,I帧)之间的距离。在一个例示性示例中,GOP长度在本文中也可被称为I帧时间间隔。GOP长度(例如,或I帧间隔)可用I帧之间的帧的数目或数量来测量(例如,如上文所描述)和/或可被测量为I帧之间的时间量。例如,在图3所例示的多个视频帧300的情况下,对于每15帧的视频,可插入I帧并将其用作关键帧。如果多个视频帧300与30帧/秒(fps)的回放速度相关联,则GOP长度也可被测量为0.5秒(例如,每一秒视频两个I帧关键帧)。

[0115] 在一些示例中,可基于使用可变I帧时间间隔(例如,“可变GOP长度”)来提高视频译码的效率和/或性能。在一些方面,不同类型的视频内容的提高的视频译码效率可与不同GOP长度相关联。例如,可使用相对较大的GOP长度来对相对静态的视频内容进行译码(例如,基于视频内容随时间相对较少地改变),而可使用相对较小的GOP长度来对具有快速移动或许多移动对象的视频内容进行译码(例如,基于视频内容随时间相对快速地改变)。

[0116] 在一些示例中,可基于视频内容分析和/或运动分析来执行可变长度GOP视频译码。例如,可执行视频内容分析和/或运动分析以确定与视频内容相关联的改变量(例如,随时间的改变量)。此类技术可实现提高的视频译码性能,通常可与高计算复杂性和功率消耗相关联。例如,可基于与视频译码分开执行视频内容分析和/或运动分析而增加计算复杂性和功率消耗。

[0117] 在一些情况下,可基于与给定视频译码设备相关联的计算约束、功率约束和/或译码时间约束来利用固定长度GOP视频译码(例如,固定I帧时间间隔视频译码)。例如,无线显示共享和/或WiFi显示技术(例如,诸如Miracast[®])可用于通过捕获和编码源设备(例如,智能电话)的显示内容并将编码的视频数据发送到接收设备(例如,电视机)而在WiFi设备之间无线地共享视频内容。基于源设备(例如,智能电话)处可用的计算资源和/或功率和能量资源,WiFi显示技术(诸如Miracast[®])通常利用固定长度GOP来对源设备处显示的视频数据进行编码(并随后解码)。在一些方面,至少部分基于源设备(例如,智能电话)视频显示数据在源、类型、内容等方面变化,使用固定长度GOP可降低针对WiFi显示技术(诸如Miracast[®])执行的视频译码的效率。需要可用于执行可变长度GOP视频译码(例如,可变I帧时间间隔视频译码)而不使用计算密集型技术(诸如视频内容分析或运动分析)的系统和技术。还需要可用于以可用于实现WiFi显示共享(例如,Miracast[®])的功率有效方式执行可变长度GOP视频译码

(例如,可变I帧时间间隔视频译码)的系统和技术。

[0118] 如下文更详细描述,提供了用于使用可变帧内(I帧)时间间隔和/或可变长度图片组(GOP)长度来执行视频译码的系统和技术。例如,该系统和技术可使用基于诸如视频帧层信息、视频帧层几何形状等信息确定的可变I帧间隔和/或可变GOP长度来执行视频译码(例如,编码和/或解码)。例如,该系统和技术可获得并分析与待编码的多个帧中所包括的一个或多个层相关联的层信息。在一些方面,待编码的多个帧可由用于执行无线显示共享(例如, Miracast[®])的智能电话或其他移动计算设备捕获。

[0119] 例如,图4A是例示可与所捕获的视频显示数据的帧相关联的帧层的示例的示图400。在一个例示性示例中,所捕获的视频显示数据的帧410可由充当用于无线显示共享(例如,WiFi显示共享、Miracast[®]等)的源设备的智能电话或其他移动计算设备捕获或以其他方式从该智能电话或其他移动计算设备获得。所捕获的视频显示数据的帧410可与至少包括层索引420和层类型430的层信息相关联。例如,所捕获的视频显示数据的每个帧(例如,诸如所捕获的视频显示数据的帧410)可包括一个或多个层,其中单独层与在所捕获的视频显示数据的帧410中显示的不同图标、应用、UI元素等相关联。

[0120] 例如,第一层可与在所捕获的视频显示数据的帧410中显示或包括的背景相关联;第二层可与状态栏系统用户界面(UI)元素相关联(例如,显示诸如当前时间、通知、连接的网络、通知、电池百分比等系统信息);第三层可与在背景显示层的顶部上渲染的应用图标相关联;第四层可与在其他层的顶部上显示移动应用的画中画(PIP)或浮动窗口UI元素相关联;等等。

[0121] 在一些示例中,该系统和技术可利用所捕获的视频显示数据的帧,其是 Miracast[®]帧或与无线显示共享相关联的其他帧。在一些方面, Miracast[®]帧可包括给定帧中所包括的每个层的层信息或与该层信息相关联。在一些方面,所捕获的视频显示数据的每个帧(例如,诸如所捕获的视频显示数据的帧410)可包括一个或多个层。例如,帧410可包括1至15层,这些层组成帧410的最终渲染视图,该最终渲染视图作为 Miracast[®]或其他无线显示共享帧来捕获并共享。

[0122] 例如,所捕获的视频显示数据的帧可由执行无线显示共享的源设备(例如,智能电话)生成或以其他方式获得。每个相应帧的层信息可在无线显示共享过程期间由源设备确定(例如,由源设备填充)。例如,源设备可生成用于无线显示共享的所捕获的视频显示数据的 Miracast[®]帧或其他帧,它们包括每个相应帧中所包括的单独层的几何形状、每个层的坐标、每个层的格式或层类型等。

[0123] 例如,如图4A所例示,所捕获的视频显示数据的帧410可包括具有索引0、1、6、7和8以及层类型(例如,合成类型430)“SDE”(例如,骁龙显示引擎或其他硬件显示处理单元)的层。例如,层类型或合成类型“SDE”可指示对应的层将由SDE或其他硬件显示处理单元构成。帧410可另外包括具有索引11和层类型(例如,合成类型430)“GPU_TARGET”的层。在一些方面,可由与所捕获的视频显示数据的帧410相关联的源设备提供诸如层索引420和层合成类型430的层信息。

[0124] 图4B是例示根据一些示例的与所捕获的视频显示数据的帧相关联的示例层堆栈450的示图。例如,层堆栈450可以是应用层堆栈,其中在智能电话或其他移动计算设备上运行(例如,当前在其显示器上渲染)的应用各自与层堆栈450中的一个或多个层相关联。

[0125] 层堆栈450可包括以z顺序排序(例如,“堆叠”)的多个不同层,其中在堆栈450的顶部上的层是在层堆栈450在智能电话或其他计算设备的显示器上渲染时可见的最顶层。例如,与“应用4”相关联的层可以是在与示例层堆栈450相关联的所捕获的视频显示数据的帧中可见的最顶层,而与“应用0”相关联的层可以是在层堆栈450的其余层之下渲染的(例如,层堆栈450中所包括的层中的)最底层。

[0126] 如先前所提及,所捕获的视频显示数据的每个帧可包括多个层(例如,也被称为“子层”)。在一些方面,所捕获的视频显示数据的连续帧之间的层的数量和/或类型可改变。每个层可与层信息和/或层参数相关联。例如,每个层可与层名称、z顺序、格式、层类别、一个或多个变换、一个或多个标记、元数据、帧区域坐标等相关联。在一个例示性示例中,本文所述的系统和技术所利用的层信息可与源设备(例如,参与无线显示共享的源设备)用来确定帧组成、几何形状渲染、色调映射和/或各种其他显示后处理操作的信息相同或类似。

[0127] 在一些方面,层信息可确定所得帧的几何形状。例如,可基于与五个层中的每个相应层相关联的层信息来确定包括五个层的帧的几何形状。基于确定所得帧几何形状的层信息,层信息还可确定在基于层信息生成的最终帧中渲染(或以其他方式可见)的最终内容像素。

[0128] 层名称可由应用(例如,在源设备上运行)和/或操作系统(OS)框架(例如,源设备的OS框架)指定。在一些方面,层名称也可被称为层可读ID或层标识符。基于至少部分地在源设备上运行的应用指定的层名称参数,与不同应用相关联的层可具有不同的名称和命名约定。在一个例示性示例中,与给定层相关联的层名称可用于确定与给定层相关联的使用或用例、与给定层相关联的一个或多个内容分类和/或与给定层相关联的应用信息。例如,该系统和技术可使用层名称参数来确定给定层是游戏层还是非游戏层(例如,因为游戏层可与快速变化的内容相关联并且可从比非游戏层更短的GOP长度获益)。

[0129] 层信息还可包括与所捕获的视频显示数据的给定帧相关联的层中的一些(或全部)层的z顺序。层的z顺序可指示层沿z轴的相对定位(例如,相对于相同帧的其他层)。例如,层的z轴排序可用于确定哪些层在其他层的顶部上,哪些层在其他层的下面,等等。

[0130] 与给定帧相关联的层信息可包括帧层数或帧层数量参数,其指示所捕获的视频显示数据的给定帧中所包括的层的数量。例如,帧层数信息可指示所捕获的视频显示数据的给定帧包括总共八层。在一些方面,帧层数信息可另外标识与特定应用或源相关联的层的子集。例如,帧层数信息可指示所捕获的视频显示数据的给定帧包括总共八层,其中四个层与第一应用相关联,两个层与第二应用相关联,并且两个层与源设备OS相关联。

[0131] 在一些方面,层信息可包括变换信息。例如,变换信息可指示应用于层内容的一个或多个变换。在一些方面,层变换信息可指示给定层如何旋转或翻转(例如,如果有的话)。例如,层变换信息可指示层不经历变换、顺时针旋转90度、逆时针旋转90度,等等。

[0132] 在一些方面,层信息可包括显示帧信息。例如,显示帧信息可指示给定层位于其相关联的帧(例如,包括给定层的所捕获的视频显示数据的帧)内的何处。在一个例示性示例中,显示帧信息可包括二维(2D)坐标,其指示给定层在其相关联的帧内的定位以及给定层在其相关联的帧内的范围(例如,大小)。例如,显示帧信息可包括指示帧中所包括的给定层的一个或多个角部的位置的(x,y)坐标。

[0133] 在一些方面,层信息可包括元数据信息。例如,元数据信息可以是与渲染给定层的

一些(或全部)视觉内容或像素相关联的信息。在一个例示性示例中,元数据信息可包括层颜色、层亮度、高动态范围(HDR)信息、电光传递函数(EOTF)信息和/或光电传递函数(OETF)信息等中的一者或多者。

[0134] 在一些方面,层信息可包括合成类型(composition type)(例如,也被称为“comp type(合成类型)”)。合成类型可指示给定层将由数据处理单元(DPU)、图形处理单元(GPU)还是其他硬件处理器构成。

[0135] 在一些方面,层信息可包括指示给定层所需的特定处理的一个或多个标记。例如,层可与指示该层需要安全处理、仅外部显示处理等的标记相关联。

[0136] 在一些方面,层信息可包括层格式。层格式可包括RGB或YUV、RGB1010102或FP16等。

[0137] 在一些示例中,可为所捕获的视频显示数据的给定帧中所包括的每个层获得层信息。在一个例示性示例中,可在层信息的组合列表中获得给定帧中所包括的层的层信息。例如,图4C描绘了包括可为所捕获的视频显示数据的帧获得的层信息的示例列表的示例表400c。

[0138] 在一些方面,常规视频流(例如,由相机或者在智能电话或其他移动设备中所包括的相机捕获)可不包括任何层信息,因为此类视频流是作为单层记录捕获的。在一个例示性示例中,本文所述的系统和技术可使用所捕获的视频显示数据的帧中所包括或以其他方式与所捕获的视频显示数据的帧相关联的层信息来确定用于设定自适应或可变长度GOP(例如,自适应或可变I帧时间间隔)的场景信息。在一些情况下,基于层信息用于最初在参与无线显示共享的源设备的显示器上渲染视频显示数据,层信息可与所捕获的视频显示数据的帧相关联。在一些示例中,与所捕获的视频显示数据的帧相关联的层信息可包括可比基于像素的内容或运动分析更有效地分析的丰富场景信息。基于从层信息确定的丰富场景信息,该系统和技术可实现场景变化分析、通知和/或I帧触发,如下文将更深入地描述。

[0139] 图5是例示可由本文所述的系统和技术实现的过程500的示例的流程图,以使用基于与视频显示数据的一个或多个帧相关联的层信息来自适应地确定的可变长度GOP(例如,可变I帧时间间隔)执行视频译码。

[0140] 在一个例示性示例中,视频显示数据的一个或多个帧可以是参与无线显示共享(例如,诸如Miracast®)的源设备获得的所捕获的视频显示数据的帧。例如,所捕获的视频显示数据的多个帧510可包括当前帧504(例如,与时间t相关联)、先前帧502(例如,与时间t-1相关联)以及n个附加帧,直到与时间t+n相关联的最终帧。

[0141] 基于监测并分析多个帧510中所包括的相应帧中的一些(或全部)帧的层信息,可用可变GOP长度和可变I帧时间间隔来编码所捕获的视频显示数据的多个帧510。例如,在框550处,该系统和技术可监测并分析为当前帧504(例如,与当前时间t相关联的帧)获得的或以其他方式与该当前帧相关联的层信息。在一些方面,在框550处监测并分析当前帧504的层信息可包括在执行该分析之前获得当前帧504的层信息。

[0142] 在框552处,该系统和技术可将当前帧504与一个或多个先前帧(例如,诸如与时间t-1相关联的先前帧502)进行比较以确定是否已发生场景变化和/或确定是否已发生帧几何形状变化。例如,可将与先前帧502相关联的层信息与同当前帧504相关联的层信息(例如,在框550处确定并分析的层信息)进行比较。在一些方面,与先前帧502相关联的层信息

可存储在先前时间步骤中,在该先前时间步骤中帧502是提供给框550以供分析的当前帧。

[0143] 在一个例示性示例中,可在框552处基于先前帧502的层信息中所包括和/或当前帧504的层信息中所包括的一个或多个参数改变大于预先确定的阈值的量来检测场景变化或帧几何形状变化。例如,可在框552处基于层数改变(例如,从帧502中的四个层增加到帧504中的六个层)和/或基于一个或多个层大小改变(例如,层大小从帧502到帧504增加至少25%)来检测场景或帧几何形状变化。在一些情况下,可在框552处基于一个或多个层位置改变(例如,向上/向下、向左/向右移动大于层的对应尺寸的50%的量)和/或基于一个或多个层格式改变(例如,从帧502中的HDR YUV 10比特改变为帧504中的SDR YUV 8比特)来检测场景或帧几何形状变化。

[0144] 在一些方面,可附加地或另选地在框552处基于标识主要层从先前帧502到当前帧504的变化来检测场景变化或帧几何形状变化。例如,所捕获的视频显示数据的给定帧510的主要层可以是大小足够大的最高定位层(例如,以图4B所例示的层堆栈450的z顺序)。在一些情况下,给定帧的主要层可以是超过预先确定的大小阈值的最高定位层(例如,在给定帧的顶部处或顶部附近)。例如,预先确定的大小阈值可以是给定帧的像素面积的30%。

[0145] 基于在框552处检测到场景变化或帧几何形状变化(例如,图5中所例示的“是”选项),该系统和技术可在框554处自动触发I帧编码。例如,可基于当前帧504来对新的I帧进行编码,并将新的I帧用作为所捕获的视频显示数据的多个帧510生成的编码的视频数据的关键帧。

[0146] 在框558处可另外确定并应用新的GOP长度(例如,结合响应于在框552处检测到显著场景变化或帧几何形状变化而在框554处触发的新的I帧编码)。例如,新的GOP长度可指示在另一个I帧关键帧要被编码之前将被编码为B帧或P帧的帧的数量(例如,所捕获的视频显示数据的多个帧510的数量)。在一个例示性示例中,在框558处确定并应用的新GOP长度可基于针对为当前编码的帧504标识的主要层确定的GOP长度。如先前所提及,所捕获的视频显示数据的给定帧510的主要层可以是大小足够大的最高定位层(例如,以图4B所例示的层堆栈450的z顺序)。在一些方面,GOP长度可基于主要层的类型和/或可基于与主要层相关联的层信息。

[0147] 例如,可基于与所捕获的视频显示数据的当前编码的帧(例如,当前帧504)的所标识的主要层相关联的层名称、层格式、层元数据和/或层大小中的一者或多者来确定GOP长度。在一些方面,可预先确定GOP长度值到层信息组合的映射。在一些示例中,可训练机器学习(ML)网络或分类器以基于与主要层相关联的层信息来生成GOP长度值。例如,可将神经网络或深度神经网络(DNN)作为分类器在诸如层名称、层格式、层元数据、层大小等的层信息上训练。在一些示例中,可训练神经网络分类器以基于接收主要层的层信息作为输入来输出GOP长度值。在一些示例中,可训练神经网络分类器以基于接收主要层的层信息作为输入来输出指示主要层类型的语义分类(例如,并且随后可将语义分类映射到相应语义分类的预先确定的GOP长度值)。

[0148] 至少部分地基于主要层的内容类型和/或主要层的内容类型的帧几何形状变化的预期频率,可为不同类型的主要层确定或利用不同的GOP长度。例如,作为游戏层的主要层可与相对较短的GOP长度(例如,60帧)相关联,而作为电子邮件客户端/应用层的主要层可与相对较长的GOP长度(例如,120帧)相关联。在一些示例中,视频回放层和/或YUV颜色格式

层可与比游戏层GOP长度更短的GOP长度相关联(例如,视频回放层和/或YUV颜色格式层可与30帧的GOP长度相关联)。在一些方面,包括相对静态内容(诸如来自基于文本的聊天或消息收发应用的内容)的主要层可具有相对较长的GOP长度(例如,300帧)。类似地,包括相对静态内容(诸如来自幻灯片演示应用的内容)的主要层也可具有相对较长的GOP长度(例如,300帧)。

[0149] 在框554处对当前编码的帧504触发新的I帧(例如,关键帧)编码并且基于当前编码的帧504的主要层的层信息或层类型来应用新的GOP长度(例如,在框558处)之后,该系统和技术可在框566处前进到下一帧。例如,在编码当前帧504(例如,与时间t相关联)之后,该系统和技术可从框566前进到通过返回到框550来分析下一帧(例如,与时间t+1相关联)的层信息。

[0150] 返回到对当前帧504进行编码的讨论,如果在框552处未检测到场景变化或帧几何形状变化(例如,“否”选项),则该系统和技术可前进到框556,该框确定是否已发生或正发生显示器空闲状态。例如,在框556处,该系统和技术可确定来自源设备的所捕获的显示数据的帧中所表示的视频内容已转变到空闲状态。在一些方面,空闲状态可与在预先确定的时间段内未检测到内容变化和/或未检测到所捕获的显示数据的帧的额外渲染或刷新相关联。例如,可基于在框552处未在预先确定的连续帧数内检测到场景或帧几何形状变化来在框556处检测到或触发显示器空闲状态。

[0151] 在一些示例中,检测到或确定显示器空闲状态可使该系统和技术应用与显示器空闲状态相关联的新的相对较长的GOP长度(例如,也被称为“显示器空闲GOP长度”)。例如,显示器空闲GOP长度可以是300帧,尽管更多或更少数量的帧也可用于显示器空闲GOP长度。例如,如果在框556处检测到显示器空闲状态(例如,“是”选项),则该系统和技术可在框560处应用显示器空闲GOP长度。如果在框556处未检测到显示器空闲状态(例如,“否”选项),则该系统和技术可在框562处应用(例如,保持)当前GOP长度。在一个例示性示例中,可在框562处应用或保持的当前GOP长度(例如,当未检测到显示器空闲时)可与同最近I帧相关联的最近应用的新GOP长度(例如,从上次到达框558时起最近应用的新GOP长度)相同。

[0152] 在框560处应用显示器空闲GOP长度或在框562处保持当前GOP长度之后,该系统和技术可将当前编码的帧504编码为P帧或B帧(例如,非I帧,因为只有在基于在框552处检测到场景或帧几何形状变化而到达框554时,才会将当前帧504编码为I帧)。在框564处将当前编码的帧504编码为P帧或B帧之后,该系统和技术可基于框566将过程500返回到框550来前进到下一帧(例如,与时间t+1相关联的帧)。

[0153] 在一个例示性示例中,如果先前(例如,与时间t-1相关联)帧502包括四个层并且当前(例如,与时间t相关联)帧504包括六个层,则可在框552处检测到场景变化和/或帧几何形状变化,可在框554处触发新的I帧编码,并且可在框558处应用新的GOP长度(例如,其中新的GOP长度基于与当前编码的帧504的主要层相关联的GOP长度)。

[0154] 在另一例示性示例中,如果先前帧502包括具有由0 0 1080 2400给出的层大小信息的层并且当前帧504包括具有现在由220 180 540 1200给出的层大小信息的相同层,则可在框552处检测到帧几何形状变化,可在框554处触发新的I帧编码,并且可在框558处应用新的GOP长度。

[0155] 在另一例示性示例中,如果先前帧502包括具有指示该层是HDR YUV 10比特格式

的元数据信息和/或格式信息的层,并且当前帧504包括具有现在指示该层是SDR YUV 8比特层的元数据和/或格式信息的相同层,则可在框552处检测到场景变化和/或帧几何形状变化,可在框554处触发新的I帧编码,并且可在框558处应用新的GOP长度。

[0156] 在另一示例中,如果层名称在先前帧502与当前帧504之间改变,则可在框552处检测到场景变化。例如,可基于确定来自先前帧502的层名称不被包括在当前帧504中来检测到场景变化。在一些方面,存在于先前帧502中但不存在于当前帧504中的层名称可指示一个或多个新层被包括在当前帧504中。在一些情况下,存在于先前帧502中但不存在于当前帧504中的层名称可指示新的主要层对于当前帧504将为活动的或被包括在该当前帧中。在任一情况下,可在框552处检测到场景变化,可在框554处触发新的I帧编码,并且可在框558处应用新的GOP长度。

[0157] 在一些方面,所捕获的视频显示数据的多个帧510可以是被捕获或获得并随后实时编码的连续帧。在一些方面,所捕获的视频显示数据的多个帧510可包括与无线显示共享(例如,Miracast)会话相关联的多个帧,其中源设备(例如,智能电话或其他计算设备)将其显示内容镜射到目标设备(例如,电视机)。例如,帧502可以是无线显示共享会话捕获的第一帧,并且帧508可表示为相同无线显示共享会话捕获的最后一帧。

[0158] 如先前所提及,用于对所捕获的视频显示数据的多个帧510进行译码的GOP长度和I帧间隔可随时间变化。例如,如果帧502包括八个层并且帧504包括四个层,则在框552处检测到场景变化,从而在框554处触发新的I帧(例如,关键帧)编码并在框558处应用新的GOP长度(例如,特定于帧504的主要层)。如果当前编码的帧504的主要层是具有4K分辨率的YUV格式层,则在框558处应用的新的GOP长度可以是与YUV 4K层相关联的目标GOP长度。例如,如果YUV 4K层与30帧的目标GOP长度相关联,则在框558处应用的新的目标GOP长度可以是30帧。

[0159] 如果随后在一分钟内未检测到超过预先确定的阈值的场景变化或帧几何形状变化(例如,对于在帧504之后的一分钟量的帧510,框552处的确定为“否”),则目标GOP长度将保持在30帧并且每30帧触发新的I帧。如果帧510与30fps的回放速度相关联,则每秒触发一次新的I帧,在一分钟时间段内触发总共60个I帧,其中在框552处未检测到场景变化或帧几何形状变化。

[0160] 如果在一分钟标记处当前编码的帧层数从四增加到六,则在框552处检测到场景变化,在框554处立即触发新的I帧(例如,与是否已经满足30帧的目标GOP长度无关),并且在框558处应用新的目标GOP长度。

[0161] 如果在框552处在接下来的20分钟内未检测到进一步的场景变化或帧几何形状变化,则I帧触发将基于当前目标GOP长度继续进行20分钟时间段。最初,目标GOP长度可以在帧层计数从四增加到六时在框558处确定的GOP长度。在框552处分析预先确定的数量的连续帧而未检测到场景变化或帧几何形状变化之后,接着将目标GOP长度更新为显示器空闲GOP长度(例如,在框560处)。例如,显示器空闲GOP长度可以是相对较长的GOP长度(例如,诸如对于30fps视频而言的300帧或10秒)。

[0162] 如果在20分钟标记处,帧层数保持不变(例如,保持在六层)但层大小和/或层坐标中的一者或多者改变,则将在框552处检测到场景变化并且将在框554处立即触发新的I帧编码(例如,即使尚未达到显示器空闲GOP长度)。随后可基于与层大小变化和/或层坐标变

化相关联的帧的主要层来确定新的GOP长度,并且可在框558处应用该新的GOP长度以替换显示器空闲GOP长度。

[0163] 在一个示例性示例中,本文所述的系统和技术可用于执行实时或“在线”视频译码。例如,在从源设备到目标设备的无线显示共享(例如,Miracast)的上下文中,在源设备处获得所捕获的视频显示数据的帧,对这些帧进行编码,并且发送这些帧以供在目标设备上基本上实时地镜像显示。基于在不执行基于分析像素级信息的内容分析或运动分析的情况下检测到场景变化和/或帧几何形状变化,该系统和技术可利用可变GOP长度和/或I帧间隔来更有效地提供实时视频编码和解码。

[0164] 图6是示例根据本文所述的各方面的用于对视频数据进行编码或解码(译码)的过程600的示例的流程图。在框602处,过程600包括获得与计算设备的显示器相关联的视频数据帧,其中视频数据帧包括一个或多个层。例如,可从图1所例示的与编码设备104的显示器相关联的视频源102获得视频数据帧。在一些示例中,视频数据帧可与图4A所例示的帧410(其包括图4B所例示的一个或多个层450)相同或类似。在一些情况下,视频数据帧可与图5所例示的帧510中的一者或多者(例如,帧502、504、508等中的一者或多者)相同或类似。

[0165] 在一些情况下,视频数据帧包括在计算设备的显示器上显示的视频显示数据。在一些示例中,视频数据帧可以是与从计算设备到第二计算设备的无线显示共享(例如,Miracast)相关联的所捕获的视频显示数据的帧。例如,视频数据帧可以是与从图1所例示的编码设备104到图1所例示的解码设备112的无线显示共享相关联的所捕获的视频显示数据的帧。

[0166] 在框604处,过程600包括比较与视频数据帧中所包括的一个或多个层相关联的层信息和与先前视频数据帧中所包括的一个或多个层相关联的层信息。例如,层信息可至少包括图4A所例示的层索引值420和合成类型430信息。在一些情况下,层信息可与视频数据帧中所包括的一个或多个层(诸如图4B所例示的应用层堆栈450中所包括的一个或多个层)相关联。在一些情况下,层信息可包括图4C中描绘的表400c中所例示的层信息中的一些(或全部)。

[0167] 在一些情况下,层信息可包括对于一个或多个层中所包括的每个相应层,与每个相应层相关联的层名称、与每个相应层相关联的层格式以及与每个相应层相关联的一个或多个坐标中的至少一者。在一些示例中,层信息可包括层数量或帧层数中的至少一者(例如,诸如图4A所例示的层索引值420)。在一些情况下,视频数据帧和先前视频数据帧可以是所捕获的视频显示数据的连续帧。例如,视频数据帧可与图5所例示的所捕获的显示数据的帧504相同或类似,并且先前视频数据帧可与图5所例示的所捕获的显示数据的帧502相同或类似。处于

[0168] 在一些示例中,视频数据帧和先前视频数据帧可以是所捕获的视频显示数据的多个帧(例如,诸如图5所例示的多个帧510)中所包括的连续帧。在一些示例中,可使用帧间预测帧对所捕获的视频显示数据的多个帧的至少一部分进行编码,如下文更深入地描述。在一些情况下,可将编码的视频数据发送到与同计算设备相同的无线显示共享会话相关联的第二设备。编码的视频可包括帧间预测帧以及基于对所捕获的视频显示数据的多个帧的至少该部分进行编码而生成的单向预测帧或双向预测帧中的至少一者。

[0169] 在框606处,过程600包括基于确定与视频数据帧相关联的帧几何形状变化来使用

视频数据帧生成帧间预测帧。例如,确定与视频数据帧相关联的帧几何形状变化可基于比较与视频数据帧中所包括的一个或多个层相关联的层信息和与先前视频数据帧中所包括的一个或多个层相关联的层信息。在一些情况下,可基于确定与视频数据帧中所包括的一个或多个层相关联的层信息相比于与先前视频数据帧中所包括的一个或多个层相关联的层信息已改变大于阈值量来确定帧几何形状变化。

[0170] 在框608处,过程600包括基于与视频数据帧中所包括的一个或多个层相关联的层信息来确定更新的图片组(GOP)长度。例如,可基于与视频数据帧中所包括的主要层相关联的层信息来确定更新的GOP长度。在一些情况下,可使用比与视频数据帧中所包括的一个或多个附加层相关联的相应z顺序更大的z顺序来渲染视频数据帧中所包括的主要层。

[0171] 在一些示例中,过程600可包括基于与视频数据帧中所包括的一个或多个层相关联的层信息相比于与先前视频数据帧中所包括的一个或多个层相关联的层信息改变小于阈值量,确定帧几何形状变化不与视频数据帧相关联。在一些情况下,可基于视频数据帧以及预先确定的数量的先前视频数据帧不与帧几何形状变化相关联来检测到与视频数据帧相关联的显示器空闲状态。在一些示例中,可应用显示器空闲GOP长度,其中显示器空闲GOP长度大于更新的GOP长度。在一些情况下,可基于视频数据帧不与帧几何形状变化相关联来将视频数据帧编码为预测帧(P帧)或双向帧(B帧)。

[0172] 在一些情况下,过程600可由解码设备(例如,图1和图8的解码设备112)执行。在一些情况下,过程600可由编码设备(例如,图1和图7的编码设备104)执行。例如,过程600可包括生成包括与视频数据块相关联的信息的编码的视频比特流。在一些示例中,过程600可包括存储编码的视频比特流(例如,存储在装置的至少一个存储器中)。在一些示例中,过程600可包括发送编码的视频比特流(例如,使用装置的发送器)。

[0173] 在一些具体实施中,本文所描述的过程(或方法)可由计算设备或装置(诸如图1所示的系统100)执行。例如,可由图1和图7所示的编码设备104、由另一视频源侧设备或视频发送设备、由图1和图8所示的解码设备112和/或由另一客户端侧设备(诸如播放器设备、显示器或任何其他客户端侧设备)来执行这些过程。在一些情况下,计算设备或装置可以包括处理器、微处理器、微型计算机或被配置为进行本文所描述的过程的步骤的设备的其他组件。在一些示例中,计算设备或装置可包括被配置为捕获包括视频帧的视频数据(例如,视频序列)的相机。在一些示例中,捕获视频数据的相机或其他捕获设备与计算设备分离,在这种情况下,计算设备接收或获得所捕获的视频数据。计算设备还可以包括被配置为传达视频数据的网络接口。网络接口可被配置为传达基于互联网协议(IP)的数据或其他类型的数据。在一些示例中,计算设备或装置可包括用于显示输出视频内容(诸如视频比特流的图片的样本)的显示器。

[0174] 这些过程可相对于逻辑流程图进行描述,这些逻辑流程图的操作表示能够以硬件、计算机指令或它们的组合来实现的操作序列。在计算机指令的上下文中,各操作表示存储在一个或多个计算机可读存储介质上的计算机可执行指令,该指令在由一个或多个处理器执行时执行所叙述的操作。一般来讲,计算机可执行指令包括执行特定功能或实现特定数据类型的例程、程序、对象、组件和数据结构等。描述操作的次序不旨在被解释为限制,并且任何数目的所描述的操作可按照任何次序和/或并行地组合以实现过程。

[0175] 此外,这些过程可以在配置有可执行指令的一个或多个计算机系统的控制下执

行,并且可以通过硬件实现为在一个或多个处理器上共同执行的代码(例如,可执行指令、一个或多个计算机程序或一个或多个应用),或其组合。如上文所指出,代码可存储在计算机可读或机器可读存储介质上,例如,以包括能够由一个或多个处理器执行的多个指令的计算机程序的形式。计算机可读或机器可读存储介质可以是非暂态的。

[0176] 本文所讨论的译码技术可以在示例视频编码和解码系统(例如,系统100)中实现。在一些示例中,一种系统包括源设备,该源设备提供稍后将由目的地设备解码的编码的视频数据。具体地,源设备经由计算机可读介质向目的地设备提供视频数据。源设备和目的地设备可包括各种各样的设备中的任一者,包括台式计算机、笔记本(例如,膝上型)计算机、平板计算机、机顶盒、电话手持设备(诸如所谓的“智能”电话、所谓的“智能”平板)、电视机、相机、显示设备、数字媒体播放器、视频游戏控制台、视频流式传输设备等。在一些情况下,源设备和目的地设备可被配备用于无线通信。

[0177] 目的地设备可经由计算机可读介质接收待解码的编码的视频数据。计算机可读介质可包括能够将编码的视频数据从源设备移动到目的地设备的任何类型的介质或设备。在一个示例中,计算机可读介质可包括使得源设备能够将编码的视频数据直接实时发送到目的地设备的通信介质。编码的视频数据可以根据通信标准(诸如无线通信协议)进行调制,并且被发送给目的地设备。通信介质可包括任何无线或有线通信介质,诸如射频(RF)频谱或一条或多条物理发送线。通信介质可以形成诸如以下各项的基于分组的网络的一部分:局域网、广域网或全球网络(诸如互联网)。通信介质可以包括路由器、交换机、基站、或对于促进从源设备到目的地设备的通信而言可以有用的任何其他装备。

[0178] 在一些示例中,可将编码的数据从输出接口输出到存储设备。相似地,可通过输入接口从存储设备访问编码的数据。存储设备可包括各种分布式或本地访问的数据存储介质中的任何数据存储介质,诸如硬盘驱动器、蓝光光盘、DVD、CD-ROM、闪存存储器、易失性或非易失性存储器、或用于存储编码的视频数据的任何其他合适的数字存储介质。在另一个示例中,存储设备可对应于文件服务器或可存储由源设备生成的编码的视频的另一中间存储设备。目的地设备可以经由流式传输或下载来从存储设备访问所存储的视频数据。文件服务器可以是能够存储编码的视频数据并且将编码的视频数据发送给目的地设备的任何类型的服务器。示例文件服务器包括web服务器(例如,用于网站)、FTP服务器、网络附加存储(NAS)设备或本地磁盘驱动器。目的地设备可通过任何标准数据连接(包括互联网连接)来访问编码的视频数据。这可以包括适于访问被存储在文件服务器上的编码的视频数据的无线信道(例如,Wi-Fi连接)、有线连接(例如,DSL、电缆调制解调器等)、或这两者的组合。编码的视频数据从存储设备的发送可以是流式发送、下载发送或它们的组合。

[0179] 本公开的技术不一定限于无线应用或设置。这些技术可以应用于视频译码,以支持各种多媒体应用中的任何多媒体应用,诸如空中电视广播、有线电视发送、卫星电视发送、互联网流式视频发送(诸如基于HTTP的动态自适应流式传输(DASH))、被编码到数据存储介质上的数字视频、对存储在数据存储介质上的数字视频的解码、或其他应用。在一些示例中,系统可被配置为支持单向或双向视频发送,以支持诸如视频流式传输、视频回放、视频广播和/或视频电话之类的应用。

[0180] 在一个示例中,源设备包括视频源、视频编码器和输出接口。目的地设备可包括输入接口、视频解码器和显示设备。源设备的视频编码器可被配置为应用本文所公开的技术。

在其他示例中,源设备和目的地设备可以包括其他组件或布置方式。例如,源设备可以从诸如外部相机之类的外部视频源接收视频数据。同样,目的地设备可以与外部显示设备对接,而不是包括集成的显示设备。

[0181] 上述示例系统仅仅是一个示例。用于同时处理视频数据的技术可以由任何数字视频编码和/或解码设备来执行。尽管一般来讲本公开的技术由视频编码设备执行,但这些技术也可由视频编码器/解码器(通常称为“编解码器”)执行。此外,本公开的技术还可通过视频预处理器执行。源设备和目的地设备仅仅是此类译码设备的示例,其中源设备生成译码的视频数据以发送给目的地设备。在一些示例中,源设备和目的地设备可以以基本上对称的方式进行操作,使得这些设备中的每个设备包括视频编码和解码组件。因此,示例系统可以支持在视频设备之间的单向或双向视频发送,例如,以用于视频流式传输、视频回放、视频广播或视频电话。

[0182] 视频源可包括视频捕获设备,诸如摄像机、包括先前捕获的视频的视频存档单元、和/或用于从视频内容提供者接收视频的视频馈送接口。作为另外的替代方案,视频源可以生成基于计算机图形的数据作为源视频,或者生成实时视频、存档视频和计算机生成的视频的组合。在一些情况下,如果视频源为摄像机,则源设备和目的地设备可形成所谓的相机电话或视频电话。然而,如上文所提及,本公开中所描述的技术通常可适用于视频译码,并且可应用于无线和/或有线应用。在每种情况下,所捕获的、预先捕获的或计算机生成的视频可由视频编码器编码。然后编码的视频信息可由输出接口输出到计算机可读介质上。

[0183] 如上文所指出,计算机可读介质可包括暂态介质(诸如无线广播或有线网络发送)或存储介质(即,非暂态存储介质),诸如硬盘、闪存驱动器、压缩光盘、数字视频光盘、蓝光光盘或其他计算机可读介质。在一些示例中,网络服务器(未示出)可从源设备接收编码的视频数据并且例如经由网络发送将编码的视频数据提供到目的地设备。相似地,介质生产设施(诸如光盘冲压设施)的计算设备可从源设备接收编码的视频数据并且生产包含编码的视频数据的光盘。因此,在各种示例中,计算机可读介质可被理解为包括各种形式的一种或多种计算机可读介质。

[0184] 目的地设备的输入接口从计算机可读介质接收信息。计算机可读介质的信息可包括由视频编码器定义的语法信息,其也由视频解码器使用,该语法信息包括描述块和其他译码的单元(例如,图片组(GOP))的特性和/或处理的语法元素。显示设备向用户显示解码的视频数据,并且可包括各种显示设备中的任一者,诸如阴极射线管(CRT)、液晶显示器(LCD)、等离子显示器、有机发光二极管(OLED)显示器、或另一类型的显示设备。已经描述了本申请的各种示例。

[0185] 编码设备104和解码设备112的具体细节分别在图7和图8中示出。图7是例示可实现本公开所述的技术中的一种或多种技术的示例编码设备104的框图。例如,编码设备104可生成本文所描述的语法结构(例如,VPS、SPS、PPS或其他语法元素的语法结构)。编码设备104可执行视频切片内的视频块的帧内预测译码和帧间预测译码。如先前所述,帧内译码至少部分地依赖于空间预测来减少或移除给定视频帧或图片内的空间冗余。帧间译码至少部分地依赖于时间预测来减少或移除视频序列的邻近或周围帧内的时间冗余。帧内模式(I模式)可指代若干基于空间的压缩模式中的任一者。帧间模式(诸如单向预测(P模式)或双向预测(B模式))可指代若干基于时间的压缩模式中的任一者。

[0186] 编码设备104包括分割单元35、预测处理单元41、滤波器单元63、图片存储器64、求和器50、变换处理单元52、量化单元54和熵编码单元56。预测处理单元41包括运动估计单元42、运动补偿单元44和帧内预测处理单元46。对于视频块重建,编码设备104还包括逆量化单元58、逆变换处理单元60和求和器62。滤波器单元63旨在表示一个或多个环路滤波器,诸如去块滤波器、自适应环路滤波器(ALF)和样本自适应偏移(SAO)滤波器。尽管滤波器单元63在图8中示出为环路内滤波器,但是在其他配置中,滤波器单元63可被实现为环路后滤波器。后处理设备57可对由编码设备104生成的编码的视频数据执行附加处理。在一些示例中,本公开的技术可由编码设备104实现。然而,在其他情况下,本公开的技术中的一种或多种技术可由后处理设备57实现。

[0187] 如图7所示,编码设备104接收视频数据,并且分割单元35将该数据分割为视频块。分割还可包括分割成切片、切片、图块或其他较大单元以及视频块分割(例如,根据LCU和CU的四叉树结构)。编码设备104大致示出了对待编码的视频切片内的视频块进行编码的组件。切片可被划分成多个视频块(并且可能划分成被称为图块的视频块集合)。预测处理单元41可基于误差结果(例如,译码速率和失真程度等)针对当前视频块选择多种可能译码模式中的一种,诸如多种帧内预测译码模式中的一种或多种帧间预测译码模式中的一种。预测处理单元41可将所得的帧内译码的块或帧间译码的块提供到求和器50以生成残差块数据并且提供到求和器62以重建编码的块以用作参考图片。

[0188] 预测处理单元41内的帧内预测处理单元46可相对于与待译码的当前块相同的帧或切片中的一个或多个相邻块执行当前视频块的帧内预测译码以提供空间压缩。预测处理单元41内的运动估计单元42和运动补偿单元44相对于一个或多个参考图片中的一个或多个预测性块执行当前视频块的帧间预测性译码以提供时间压缩。

[0189] 运动估计单元42可被配置为根据视频序列的预定模式来确定视频切片的帧间预测模式。预定模式可将序列中的视频切片指定为P切片、B切片或GPB切片。运动估计单元42和运动补偿单元44可高度集成,但出于概念目的而单独地示出。由运动估计单元42执行的运动估计是生成运动矢量的过程,其估计视频块的运动。例如,运动矢量可指示当前视频帧或图片内的视频块的预测单元(PU)相对于参考图片内的预测性块的位移。

[0190] 预测性块是被发现在像素差方面与待译码的视频块的PU紧密匹配的块,该像素差可通过绝对差总和(SAD)、平方差总和(SSD)或其他差量度来确定。在一些示例中,编码设备104可计算存储在图片存储器64中的参考图片的子整数像素位置的值。例如,编码设备104可对参考图片的四分之一像素位置、八分之一像素位置或其他分数像素位置的值进行插值。因此,运动估计单元42可执行相对于全像素位置和分数像素位置的运动搜索并且输出具有分数像素精度的运动矢量。

[0191] 运动估计单元42通过将帧间译码的切片中的视频块的PU的位置与参考图片的预测性块的位置进行比较来计算PU的运动矢量。参考图片可选自第一参考图片列表(列表0)或第二参考图片列表(列表1),其中每一者标识存储在图片存储器64中的一个或多个参考图片。运动估计单元42将所计算的运动矢量传送到熵编码单元56和运动补偿单元44。

[0192] 由运动补偿单元44执行的运动补偿可涉及基于由运动估计确定的运动矢量来提取或生成预测性块,从而可能执行对子像素精度的插值。在接收到当前视频块的PU的运动矢量后,运动补偿单元44可在参考图片列表中定位运动矢量所指向的预测性块。编码设备

104通过从正被译码的当前视频块的像素值减去预测性块的像素值从而形成像素差值来形成残差视频块。像素差值形成块的残差数据,并且可包括亮度差分量和色度差分量两者。求和器50表示执行该减法运算的一个或多个组件。运动补偿单元44还可生成与视频块和视频切片相关联的语法元素以供解码设备112在对视频切片的视频块进行解码时使用。

[0193] 帧内预测处理单元46可对当前块进行帧内预测,作为由运动估计单元42和运动补偿单元44执行的帧间预测的替代方案,如上文所描述。具体地,帧内预测处理单元46可确定用于对当前块进行编码的帧内预测模式。在一些示例中,帧内预测处理单元46可例如在单独的编码过程期间使用各种帧内预测模式来对当前块进行编码,并且帧内预测处理单元46可从测试的模式中选择适当的帧内预测模式来使用。例如,帧内预测处理单元46可使用针对各种测试的帧内预测模式的率失真分析来计算率失真值,并且可在测试的模式当中选择具有最佳率失真特性的帧内预测模式。率失真分析通常确定编码的块与被编码以产生编码的块的原始未编码块之间的失真(或误差)量,以及用于产生编码的块的比特率(即,比特数量)。帧内预测处理单元46可根据各种编码的块的失真和速率来计算比率以确定哪种帧内预测模式表现出块的最佳率失真值。

[0194] 在任何情况下,在选择用于块的帧内预测模式之后,帧内预测处理单元46可将指示用于块的所选择的帧内预测模式的信息提供到熵编码单元56。熵编码单元56可对指示所选择的帧内预测模式的信息进行编码。编码设备104可在所发送的比特流配置数据中包括用于各种块的编码上下文的定义以及对用于上下文中的每一者的最可能的帧内预测模式、帧内预测模式索引表和修改的帧内预测模式索引表的指示。比特流配置数据可包括多个帧内预测模式索引表和多个修改的帧内预测模式索引表(也被称为码字映射表)。

[0195] 在预测处理单元41经由帧间预测或帧内预测生成用于当前视频块的预测性块之后,编码设备104通过从当前视频块减去预测性块来形成残差视频块。残差块中的残差视频数据可包括在一个或多个TU中并且应用于变换处理单元52。变换处理单元52使用变换(诸如离散余弦变换(DCT)或概念上类似的变换)将残差视频数据变换为残差变换系数。变换处理单元52可将残差视频数据从像素域转换到变换域,诸如频域。

[0196] 变换处理单元52可将所得变换系数传送到量化单元54。量化单元54对变换系数进行量化以进一步减小比特率。量化过程可减小与系数中的一些或所有系数相关联的比特深度。可通过调整量化参数来修改量化程度。在一些示例中,然后量化单元54可执行对包括所量化的变换系数的矩阵的扫描。另选地,熵编码单元56可执行扫描。

[0197] 在量化之后,熵编码单元56对所量化的变换系数进行熵编码。例如,熵编码单元56可执行上下文自适应可变长度译码(CAVLC)、上下文自适应二元算术译码(CABAC)、基于语法的上下文自适应二元算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码技术。在由熵编码单元56进行熵编码之后,可将编码的比特流发送到解码设备112,或存档以供稍后发送或由解码设备112检索。熵编码单元56还可对正被译码的当前视频切片的运动矢量和其他语法元素进行熵编码。

[0198] 逆量化单元58和逆变换处理单元60分别应用逆量化和逆变换以在像素域中重建残差块以供稍后用作参考图片的参考块。运动补偿单元44可通过将残差块添加到参考图片列表内的参考图片中的一个参考图片的预测性块来计算参考块。运动补偿单元44还可将一个或多个插值滤波器应用于重建的残差块以计算子整数像素值以供在运动估计中使用。求

和器62将重建的残差块添加到由运动补偿单元44产生的运动补偿的预测块以产生参考块以供存储在图片存储器64中。参考块可由运动估计单元42和运动补偿单元44用作参考块以对后续视频帧或图片中的块进行帧间预测。

[0199] 以此方式,图7的编码设备104表示被配置为执行本文所述的技术的视频编码器的示例。例如,编码设备104可以执行本文所描述的技术中的任一者,包括本文所描述的过程。在一些情况下,本公开的技术中的一些技术还可由后处理设备57实现。

[0200] 图8是例示示例解码设备112的框图。解码设备112包括熵解码单元80、预测处理单元81、逆量化单元86、逆变换处理单元88、求和器90、滤波器单元91和图片存储器92。预测处理单元81包括运动补偿单元82和帧内预测处理单元84。在一些示例中,解码设备112可执行大致与相对于来自图7的编码设备104所描述的编码遍次互逆的解码遍次。

[0201] 在解码过程期间,解码设备112接收表示由编码设备104传送的编码的视频切片的视频块和相关联的语法元素的编码的视频比特流。在一些示例中,解码设备112可从编码设备104接收编码的视频比特流。在一些示例中,解码设备112可从网络实体79诸如服务器、媒体感知网络元件(MANE)、视频编辑器/拼接器或被配置为实现上述技术中的一种或多种技术的其他此类设备接收编码的视频比特流。网络实体79可以包括或不包括编码设备104。本公开中所描述的技术中的一些技术可在网络实体79将编码的视频比特流发送到解码设备112之前由网络实体79实现。在一些视频解码系统中,网络实体79和解码设备112可以是单独设备的部分,而在其他实例中,关于网络实体79所描述的功能性可由包括解码设备112的同一设备执行。

[0202] 解码设备112的熵解码单元80对比特流进行熵解码以生成所量化的系数、运动矢量和其他语法元素。熵解码单元80将运动矢量和其他语法元素转发到预测处理单元81。解码设备112可在视频切片级别和/或视频块级别接收语法元素。熵解码单元80可处理并解析一个或多个参数集(诸如VPS、SPS和PPS)中的固定长度语法元素和可变长度语法元素两者。

[0203] 当视频切片被译码为帧内译码(I)的切片时,预测处理单元81的帧内预测处理单元84可基于用信号通知的帧内预测模式和来自当前帧或图片的先前被解码的块的数据来生成用于当前视频切片的视频块的预测数据。当视频帧被译码为帧间译码的(即,B、P或GPB)切片时,预测处理单元81的运动补偿单元82基于从熵解码单元80接收的运动矢量和和其他语法元素来产生用于当前视频切片的视频块的预测性块。可从参考图片列表内的参考图片中的一个参考图片产生预测性块。解码设备112可基于存储在图片存储器92中的参考图片使用默认构建技术来构建参考帧列表(列表0和列表1)。

[0204] 运动补偿单元82通过解析运动矢量和其他语法元素来确定用于当前视频切片的视频块的预测信息,并且使用该预测信息来产生用于正被解码的当前视频块的预测性块。例如,运动补偿单元82可使用参数集中的一个或多个语法元素来确定用于对视频切片的视频块进行译码的预测模式(例如,帧内预测或帧间预测)、帧间预测切片类型(例如,B切片、P切片或GPB切片)、切片的一个或多个参考图片列表的构建信息、切片的每个帧间编码的视频块的运动矢量、切片的每个帧间译码的视频块的帧间预测状态,以及用于对当前视频切片中的视频块进行解码的其他信息。

[0205] 运动补偿单元82还可基于插值滤波器来执行插值。运动补偿单元82可使用如由编码设备104在视频块的编码期间使用的插值滤波器来计算参考块的子整数像素的经插值的

值。在这种情况下,运动补偿单元82可从所接收的语法元素确定由编码设备104使用的插值滤波器,并且可使用该插值滤波器来产生预测块。

[0206] 逆量化单元86对比特流中所提供且由熵解码单元80解码的所量化的变换系数进行逆量化或解量化。逆量化过程可包括使用由编码设备104针对视频切片中的每个视频块计算的量化参数来确定量化程度,并且同样地确定应当应用的逆量化程度。逆变换处理单元88将逆变换(例如,逆DCT或其他合适的逆变换)、逆整数变换或概念上类似的逆变换过程应用于变换系数以便产生像素域中的残差块。

[0207] 在运动补偿单元82基于运动矢量和其他语法元素来生成用于当前视频块的预测性块之后,解码设备112通过将来自逆变换处理单元88的残差块与由运动补偿单元82生成的对应预测性块相加来形成解码的视频块。求和器90表示执行该求和运算的一个或多个组件。如果需要,还可使用环路滤波器(在译码环路中或在译码环路之后)来平滑像素转变,或以其他方式改进视频质量。滤波器单元91旨在表示一个或多个环路滤波器,诸如去块滤波器、自适应环路滤波器(ALF)和样本自适应偏移(SAO)滤波器。尽管滤波器单元91在图8中示出为环路内滤波器,但是在其他配置中,滤波器单元91可被实现为环路后滤波器。然后将给定帧或图片中的所解码的视频块存储在图片存储器92中,该图片存储器存储用于后续运动补偿的参考图片。图片存储器92还存储解码的视频以供稍后在显示设备(诸如图1所示的视频目的地设备122)上呈现。

[0208] 以此方式,图8的解码设备112表示被配置为执行本文所述的技术的视频解码器的示例。例如,解码设备112可以执行本文所描述的技术中的任一者,包括本文所描述的过程。

[0209] 如本文所用,术语“计算机可读介质”包括但不限于便携式或非便携式存储设备、光学存储设备以及能够存储、包含或携带指令和/或数据的各种其他介质。计算机可读介质可包括非暂态介质,该非暂态介质中可存储数据并且不包括无线地或在有线连接上传播的载波和/或暂态电子信号。非暂态介质的示例可包括但不限于磁盘或磁带、光学存储介质(诸如压缩盘(CD)或数字多功能盘(DVD))、闪存存储器、存储器或存储器设备。计算机可读介质可在其上存储有代码和/或机器可执行指令,该代码和/或机器可执行指令可表示过程、函数、子程序、程序、例程、子例程、模块、软件包、类别,或者指令、数据结构或程序语句的任何组合。通过传递和/或接收信息、数据、自变量、参数或存储器内容,代码段可耦合到另一代码段或硬件电路。信息、自变量、参数、数据等可经由任何合适的手段来传递、转发或发送,这些手段包括存储器共享、消息传递、令牌传递、网络发送等。

[0210] 在一些示例中,计算机可读存储设备、介质和存储器可包括包含比特流等的电缆或无线信号。然而,在被提及时,非暂态计算机可读存储介质明确排除诸如功耗、载波信号、电磁波以及信号本身等介质。

[0211] 在以上描述中提供了具体细节以提供对本文所提供的方面和示例的详尽理解。然而,本领域普通技术人员将理解的是,可以在没有这些具体细节的情况下实施示例。为了清楚说明,在一些情况下,本技术可被呈现为包括单独的功能块,包括包含设备、设备组件、以软件或硬件和软件的组合实现的方法中的步骤或例程的功能块。可使用除附图所示和/或本文所描述的那些组件之外的附加组件。例如,电路、系统、网络、过程和其他组件可用框图形式被示出为组件以避免使这些示例和方面混淆在不必要的细节中。在其他情况下,可在没有不必要的细节的情况下示出公知的电路、过程、算法、结构和技术以避免混淆各示例和

方面。

[0212] 单独示例和方面在上文可被描述为过程或方法,该过程或方法被描绘为流程图、流程示图、数据流示图、结构示图或框图。尽管流程图可将操作描述为顺序过程,但是操作中的许多操作可被并行或并发地执行。此外,可重新排列操作的次序。当过程的操作完成时过程被终止,但是过程可具有附图中未包括的附加步骤。过程可与方法、函数、过程、子例程、子程序等相对应。当过程与函数相对应时,该过程的终止可与该函数返回调用函数或主函数相对应。

[0213] 根据上述示例的过程和方法可使用被存储的计算机可执行指令或以其他方式从计算机可读介质获取的计算机可执行指令来实现。此类指令可包括例如致使或以其他方式将通用计算机、专用计算机或处理设备配置为执行某一功能或功能群组的指令和数据。可通过网络访问所使用的计算机资源的部分。计算机可执行指令可以是例如二进制、中间格式指令,诸如汇编语言、固件、源代码等。可用于存储指令、所使用的信息和/或在根据所描述的示例的方法期间创建的信息的计算机可读介质的示例包括磁盘或光盘、闪存、具有非易失性存储器的USB设备、联网存储设备等。

[0214] 实现根据这些公开内容的过程和方法的设备可包括硬件、软件、固件、中间件、微代码、硬件描述语言或它们的任何组合,并且可采取多种形状因子中的任何形状因子。当以软件、固件、中间件或微代码实现时,用于执行必要任务的程序代码或代码段(例如,计算机程序产品)可被存储在计算机可读或机器可读介质中。处理器可执行必要任务。形状因子的典型示例包括膝上型计算机、智能手机、移动手机、平板设备或其他小形状因子个人计算机、个人数字助理、机架安装设备、独立设备等。本文所述的功能性也可被实现在外围设备或内插式卡中。借由另外的示例,此类功能性还可被实现在单个设备上执行的不同芯片或不同过程当中的电路板上。

[0215] 指令、用于输送此类指令的介质、用于执行它们的计算资源以及用于支持此类计算资源的其他结构是用于提供本公开中所描述的功能的示例部件。

[0216] 在前面的描述中,参考本申请的特定示例描述了本申请的各方面,但是本领域技术人员将认识到,本申请不限于此。尽管本申请的例示性示例和方面已在本文中详细描述,但是应当理解,发明概念可以以其他方式被实现和采用,并且所附权利要求书旨在被解释为包括这些变型,除非受到现有技术的限制。上述应用的各种特征和方面可单独地或联合地使用。此外,各示例和方面可在超出本文所述的那些环境和应用的任何数量的环境和应用中利用而不背离本说明书的更宽泛的实质和范围。因此,说明书和附图应当被认为是例示性的而非限制性的。出于例示的目的,按照特定次序来描述各方法。应当明白的是,在另选的示例中,可以以与所描述的次序不同的次序来执行方法。

[0217] 本领域普通技术人员将领会,本文所使用的小于("<")和大于(">")符号或术语可以分别用小于等于("≤")和大于等于("≥")符号来代替而不背离本说明书的范围。

[0218] 在组件被描述为“被配置为”执行某些操作的情况下,可例如通过设计电子电路或其他硬件以执行操作、通过编程可编程电子电路(例如,微处理器或其他合适的电子电路)以执行操作或它们的任何组合来实现此类配置。

[0219] 短语“耦合到”是指任何组件直接或间接物理连接到另一组件,和/或任何组件直接或间接与另一组件通信(例如,通过有线或无线连接和/或其他合适的通信接口连接到另

一组件)。

[0220] 叙述集合“中的至少一者”和/或集合中的“一者或多者”的权利要求语言或其他语言指示该集合中的一个成员或该集合中的多个成员(以任何组合)满足权利要求。例如,叙述“A和B中的至少一者”的权利要求语言意指A、B或A和B。在另一示例中,叙述“A、B和C中的至少一者”的权利要求语言意指A、B、C、或A和B、或A和C、或B和C、或A和B和C。语言集合“中的至少一者”和/或集合中的“一者或多者”不将该集合限制为该集合中所列的项目。例如,叙述“A和B中的至少一者”的权利要求语言可意指A、B或A和B,并且可附加地包括A和B的集合中未列出的项目。

[0221] 结合本文所公开的各示例和方面所描述的各种例示性逻辑块、模块、电路和算法步骤可被实现为电子硬件、计算机软件、固件或它们的组合。为了清楚地例示硬件和软件的这种可互换性,已经在其功能性方面大致描述了各种例示性组件、框、模块、电路和步骤。将此类功能性实现为硬件还是软件取决于特定应用和对整个系统提出的设计约束。技术人员可针对每种具体的应用以不同方式来实现所描述的功能性,但此类具体实施决策不应被解读为使脱离本申请的范围。

[0222] 本文所描述的技术还可以电子硬件、计算机软件、固件或它们的任何组合实现。此类技术可在多种设备中的任何设备中实现,多种设备诸如通用计算机、无线通信设备手机、或具有多种用途的集成电路设备,多种用途包括在无线通信设备手机和其他设备中的应用。被描述为模块或组件的任何特征可一起在集成逻辑器件中实现或分开地实现为分立但可互操作的逻辑器件。如果以软件来实现,则这些技术可至少部分地由包括程序代码的计算机可读数据存储介质来实现,该程序代码包括指令,该指令在被执行时执行上述方法中的一者或多者。计算机可读数据存储介质可形成计算机程序产品的一部分,其可包括封装材料。计算机可读介质可包括存储器或数据存储介质,诸如随机存取存储器(RAM)(诸如同步动态随机存取存储器(SDRAM))、只读存储器(ROM)、非易失性随机存取存储器(NVRAM)、电可擦除可编程只读存储器(EEPROM)、闪存存储器、磁性或光学数据存储介质等。附加地或另选地,该技术可至少部分地由计算机可读通信介质来实现,该计算机可读通信介质携带或传达呈指令或数据结构形式的且可由计算机访问、读取和/或执行的程序代码,诸如传播的信号或波。

[0223] 程序代码可由处理器执行,该处理器可包括一个或多个处理器,诸如一个或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其他等效集成或分立逻辑电路。此类处理器可被配置为执行本公开中所描述的技术中的任何技术。通用处理器可以是微处理器,但在另选方案中,处理器可以是任何传统的处理器、控制器、微控制器或状态机。处理器还可被实现为计算设备的组合,例如DSP与微处理器的组合、多个微处理器、与DSP核心结合的一个或多个微处理器,或任何其他此类配置。因此,如本文所用的术语“处理器”可以是指前述结构中的任何结构、前述结构的任何组合或适合于实现本文所描述的技术的任何其他结构或装置。另外,在一些方面,本文所描述的功能性可提供在被配置用于编码和解码的专用软件模块或硬件模块内,或并入组合的视频编解码器-解码器(CODEC)中。

[0224] 本公开的例示性实施例包括:

[0225] 方面1:一种用于处理视频数据的装置,所述装置包括:至少一个存储器;以及至少

一个处理器,所述至少一个处理器耦合到所述至少一个存储器,所述至少一个处理器被配置为:获得与计算设备的显示器相关联的视频数据帧,其中所述视频数据帧包括一个或多个层;比较与所述视频数据帧中所包括的所述一个或多个层相关联的层信息和与先前视频数据帧中所包括的一个或多个层相关联的层信息;基于确定与所述视频数据帧相关联的帧几何形状变化来使用所述视频数据帧生成帧间预测帧;以及基于与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息来确定更新的图片组(GOP)长度。

[0226] 方面2:根据方面1所述的装置,其中所述至少一个处理器被配置为基于与所述视频数据帧中所包括的主要层相关联的层信息来确定所述更新的GOP长度。

[0227] 方面3:根据方面2所述的装置,其中所述视频数据帧中所包括的所述主要层是使用比与所述视频数据帧中所包括的一个或多个附加层相关联的相应z顺序更大的z顺序来渲染的。

[0228] 方面4:根据方面1至3中任一项所述的装置,其中所述层信息包括针对所述一个或多个层中所包括的每个相应层的与每个相应层相关联的层名称、与每个相应层相关联的层格式以及与每个相应层相关联的一个或多个坐标中的至少一者。

[0229] 方面5:根据方面1至4中任一项所述的装置,其中所述层信息包括层数量或帧层数中的至少一者。

[0230] 方面6:根据方面1至5中任一项所述的装置,其中所述至少一个处理器被配置为基于比较与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息和与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息来确定与所述视频数据帧相关联的所述帧几何形状变化。

[0231] 方面7:根据方面6所述的装置,其中为了确定与所述视频数据帧相关联的所述帧几何形状变化,所述至少一个处理器被配置为:确定与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息相比于与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息已改变大于阈值量。

[0232] 方面8:根据方面1至7中任一项所述的装置,其中所述至少一个处理器还被配置为:基于与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息相比于与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息改变小于阈值量,确定帧几何形状变化不与所述视频数据帧相关联。

[0233] 方面9:根据方面8所述的装置,其中所述至少一个处理器还被配置为:基于所述视频数据帧以及预先确定的数量的先前视频数据帧不与帧几何形状变化相关联来检测到与所述视频数据帧相关联的显示器空闲状态;以及应用显示器空闲GOP长度,其中所述显示器空闲GOP长度大于所述更新的GOP长度。

[0234] 方面10:根据方面8至9中任一项所述的装置,其中所述至少一个处理器还被配置为:基于所述视频数据帧不与帧几何形状变化相关联来将所述视频数据帧编码为预测帧(P帧)或双向帧(B帧)。

[0235] 方面11:根据方面1至10中任一项所述的装置,其中所述视频数据帧包括在所述计算设备的显示器上显示的视频显示数据。

[0236] 方面12:根据方面1至11中任一项所述的装置,其中所述视频数据帧是与从所述计算设备到第二计算设备的无线显示共享相关联的所捕获的视频显示数据的帧。

[0237] 方面13:根据方面12所述的装置,其中所述视频数据帧和所述先前视频数据帧是所捕获的视频显示数据的多个帧中所包括的连续帧。

[0238] 方面14:根据方面13所述的装置,其中所述至少一个处理器还被配置为:使用所述帧间预测帧对所捕获的视频显示数据的所述多个帧的至少一部分进行编码。

[0239] 方面15:根据方面14所述的装置,其中所述至少一个处理器还被配置为:将编码的视频数据发送到与同所述计算设备相同的无线显示共享会话相关联的第二设备;其中所述编码的视频数据包括所述帧间预测帧以及基于对所捕获的视频显示数据的所述多个帧的至少所述部分进行编码而生成的单向预测帧或双向预测帧中的至少一者。

[0240] 方面16:一种用于处理视频数据的方法,所述方法包括:获得与计算设备的显示器相关联的视频数据帧,其中所述视频数据帧包括一个或多个层;比较与所述视频数据帧中所包括的所述一个或多个层相关联的层信息和与先前视频数据帧中所包括的一个或多个层相关联的层信息;基于确定与所述视频数据帧相关联的帧几何形状变化来使用所述视频数据帧生成帧间预测帧;以及基于与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息来确定更新的图片组(GOP)长度。

[0241] 方面17:根据方面16所述的方法,其中所述更新的GOP长度是基于与所述视频数据帧中所包括的主要层相关联的层信息来确定的。

[0242] 方面18:根据方面17所述的方法,其中所述视频数据帧中所包括的所述主要层是使用比与所述视频数据帧中所包括的一个或多个附加层相关联的相应z顺序更大的z顺序来渲染的。

[0243] 方面19:根据方面16至18中任一项所述的方法,其中所述层信息包括针对所述一个或多个层中所包括的每个相应层的与每个相应层相关联的层名称、与每个相应层相关联的层格式以及与每个相应层相关联的一个或多个坐标中的至少一者。

[0244] 方面20:根据方面16至19中任一项所述的方法,其中所述层信息包括层数量或帧层数中的至少一者。

[0245] 方面21:根据方面16至20中任一项所述的方法,其中确定与所述视频数据帧相关联的所述帧几何形状变化基于比较与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息和与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息。

[0246] 方面22:根据方面21所述的方法,其中确定与所述视频数据帧相关联的所述帧几何形状变化包括:确定与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息相比于与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息已改变大于阈值量。

[0247] 方面23:根据方面16至22中任一项所述的方法,所述方法还包括:基于与所述视频数据帧中所包括的所述一个或多个层相关联的所述层信息相比于与所述先前视频数据帧中所包括的所述一个或多个层相关联的所述层信息改变小于阈值量,确定帧几何形状变化不与所述视频数据帧相关联。

[0248] 方面24:根据方面23所述的方法,所述方法还包括:基于所述视频数据帧以及预先确定的数量的先前视频数据帧不与帧几何形状变化相关联来检测到与所述视频数据帧相关联的显示器空闲状态;以及应用显示器空闲GOP长度,其中所述显示器空闲GOP长度大于

所述更新的GOP长度。

[0249] 方面25:根据方面23至24中任一项所述的方法,所述方法还包括:基于所述视频数据帧不与帧几何形状变化相关联来将所述视频数据帧编码为预测帧(P帧)或双向帧(B帧)。

[0250] 方面26:根据方面16至25中任一项所述的方法,其中所述视频数据帧包括在所述计算设备的显示器上显示的视频显示数据。

[0251] 方面27:根据方面16至26中任一项所述的方法,其中所述视频数据帧是与从所述计算设备到第二计算设备的无线显示共享相关联的所捕获的视频显示数据的帧。

[0252] 方面28:根据方面27所述的方法,其中所述视频数据帧和所述先前视频数据帧是所捕获的视频显示数据的多个帧中所包括的连续帧。

[0253] 方面29:根据方面28所述的方法,所述方法还包括:使用所述帧间预测帧对所捕获的视频显示数据的所述多个帧的至少一部分进行编码。

[0254] 方面30:根据方面29所述的方法,所述方法还包括:将编码的视频数据发送到与同所述计算设备相同的无线显示共享会话相关联的第二设备;其中所述编码的视频数据包括所述帧间预测帧以及基于对所捕获的视频显示数据的所述多个帧的至少所述部分进行编码而生成的单向预测帧或双向预测帧中的至少一者。

[0255] 方面31:一种装置,所述装置包括:至少一个存储器;以及至少一个处理器,所述至少一个处理器耦合到所述至少一个存储器,其中所述至少一个处理器被配置为执行根据方面1至30中任一项所述的操作。

[0256] 方面32:一种装置,所述装置包括用于执行根据方面1至30中任一项所述的操作的部件。

[0257] 方面33:一种其上存储有指令的非暂态计算机可读介质,所述指令在由一个或多个处理器执行时使所述一个或多个处理器执行根据方面1至30中任一项所述的操作。

100

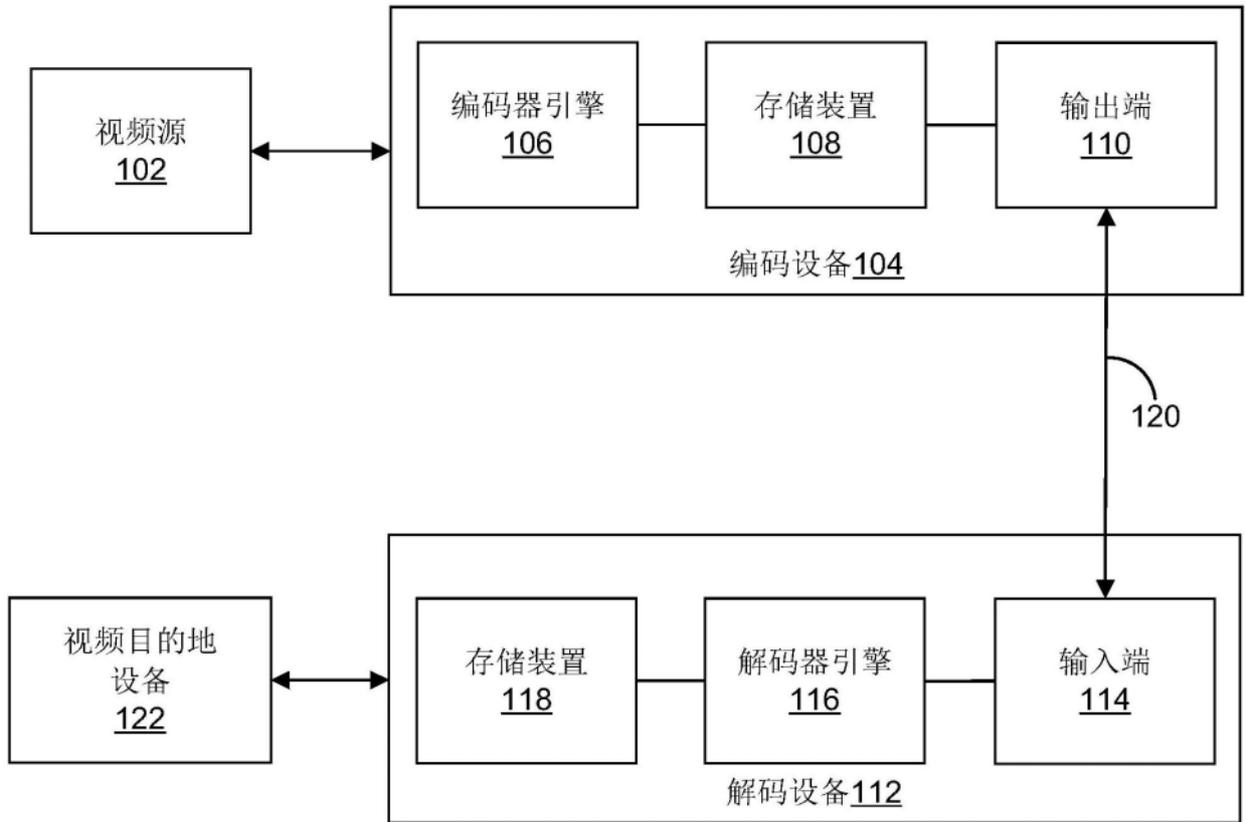


图1

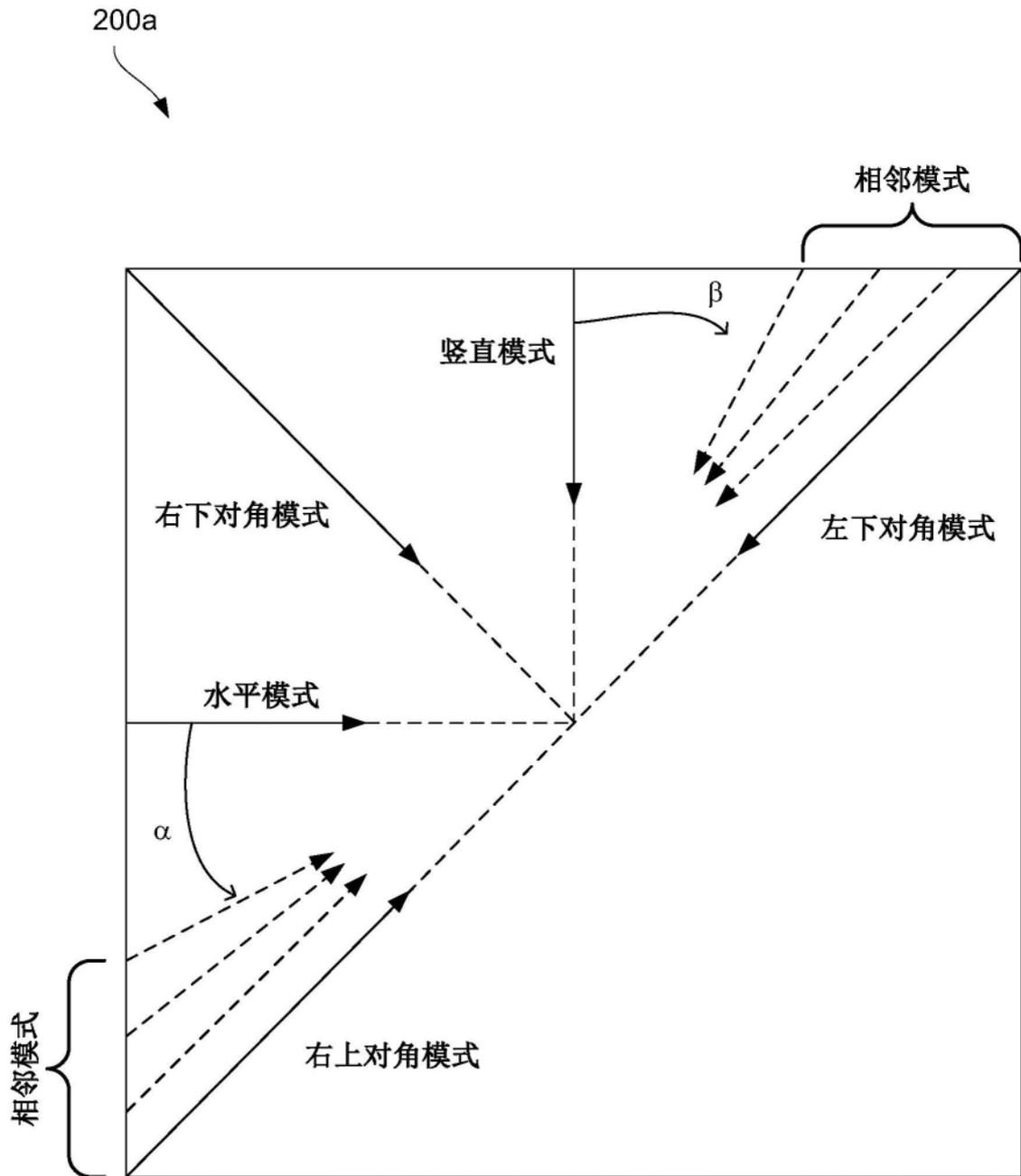


图2A

200b

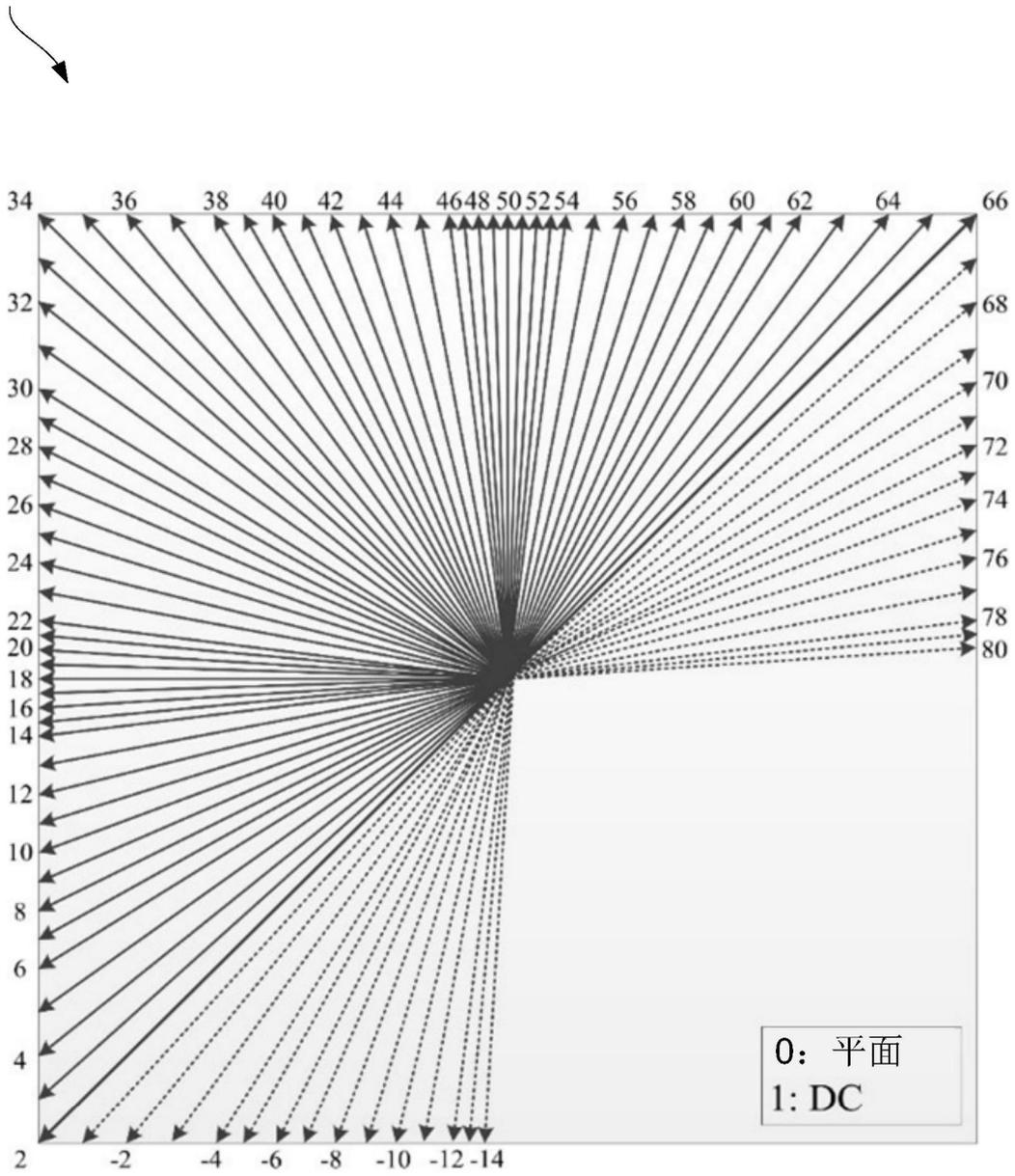


图2B

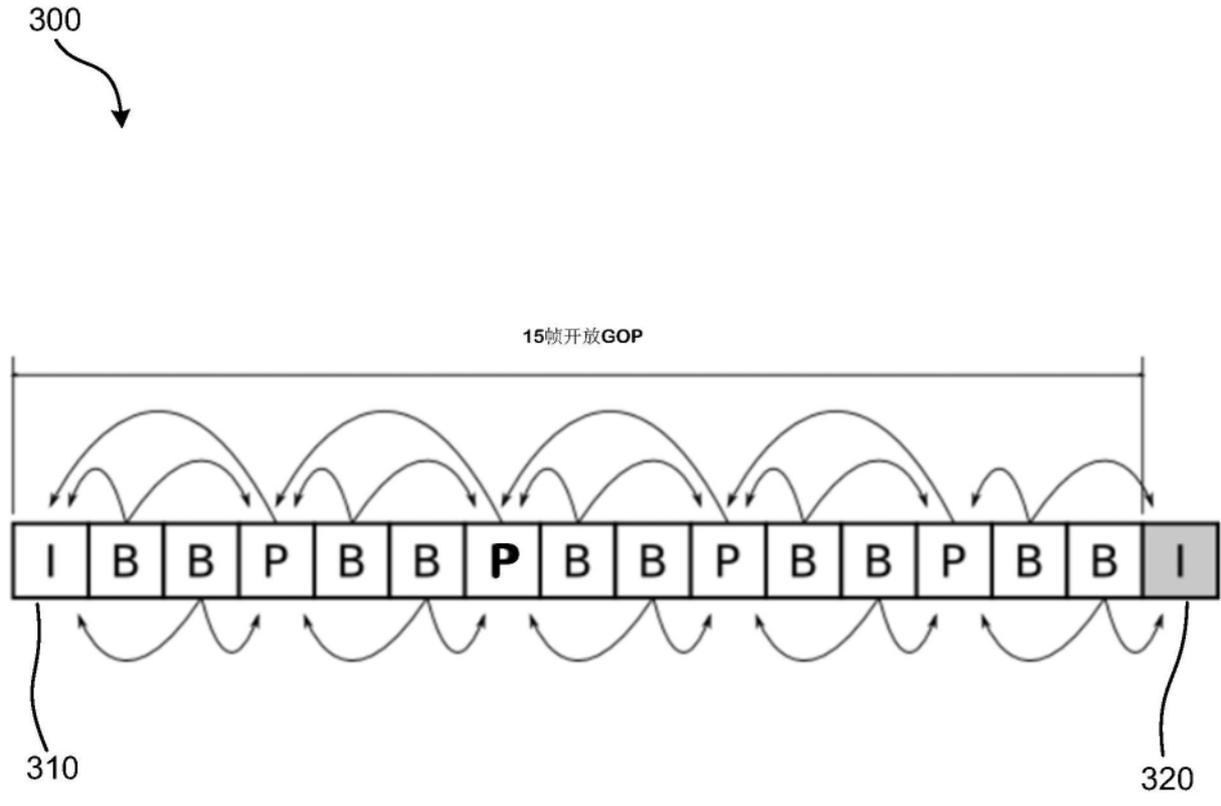


图3

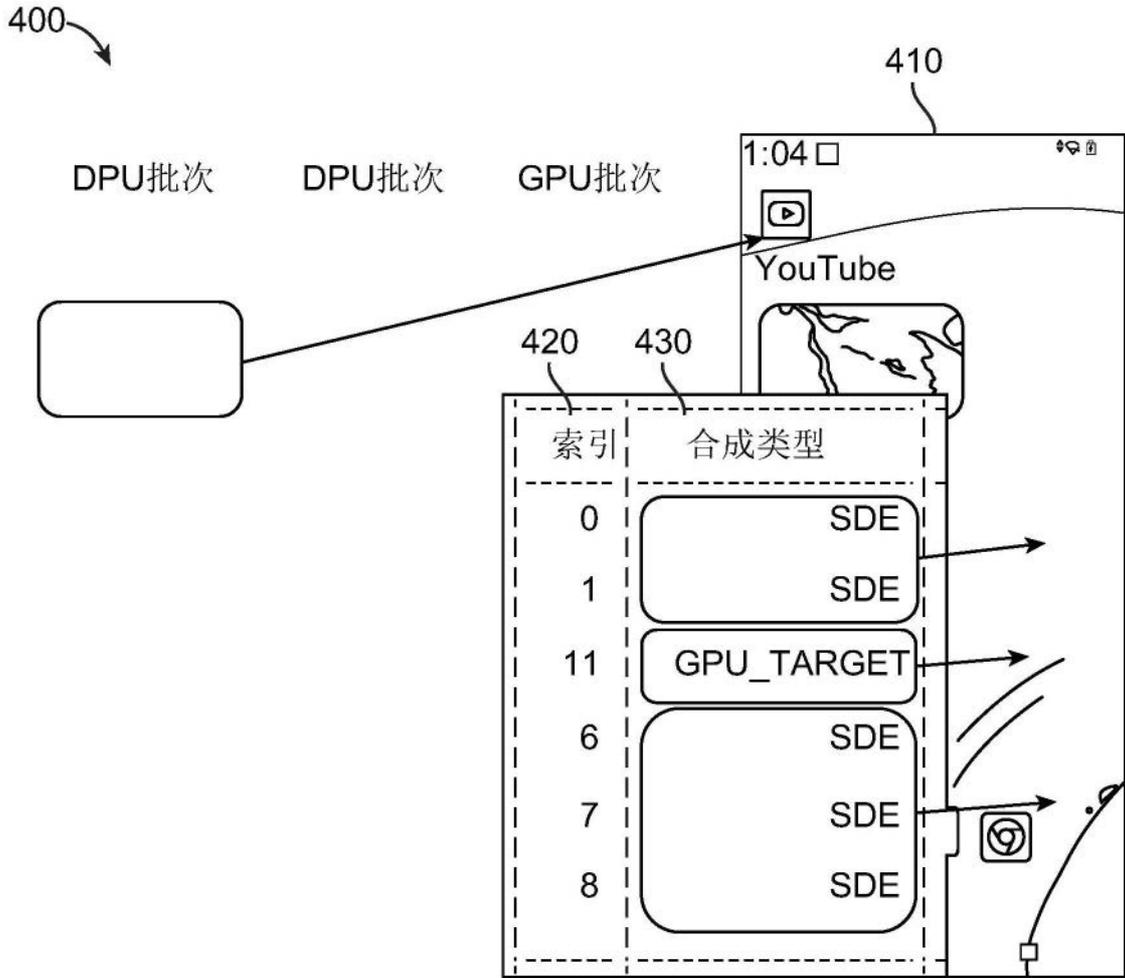


图4A

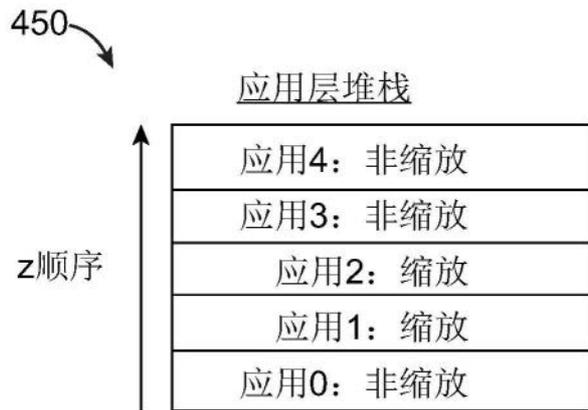


图4B

400c

层名称	Z	窗口类型	层类	合成类型	变换	显示帧 (LTRB)	源裁剪 (LTRB)	帧
帧率 (显式) (无缝性) [聚焦]								

SurfaceView[com.happyelements.Android[...].jents.hellolua.MainActivity](BLAST)#0								
rel	0	0	设备	0	0	0 1080 2400	0.0 0.0 1080.0 2400.0	60.00fps
Exact SeamedAndSeamless [*]								

com.happyelements.AndroidAnimal/com.happyelements.hellolua.MainActivity#0								
rel	0	1	设备	0	0	0 1080 2400	0.0 0.0 1080.0 2400.0	60.00fps
Exact SeamedAndSeamless [*]								

com.miui.securitycenter/.FloatingWindow#0								
rel	0	2003	设备	0	0	0 527 57 717	0.0 0.0 0.0 57.0	57.0
190.0 []								

com.miui.securitycenter/.FloatingWindow#1								
rel	0	2003	设备	0	0	0 1023 527 1080 717	0.0 0.0 0.0 57.0	57.0
190.0 []								

NavigationBar#0								
rel	0	2019	设备	0	0	0 2358 1080 2400	0.0 0.0 1080.0 42.0	
[]								

RoundCornerTop#0								
rel	0	2024	设备	0	0	0 1080 225	0.0 0.0 1080.0 225.0	
[]								

RoundCornerBottom#0								
rel	0	2024	设备	0	0	0 2175 1080 2400	0.0 0.0 1080.0 225.0	
[]								

图4C

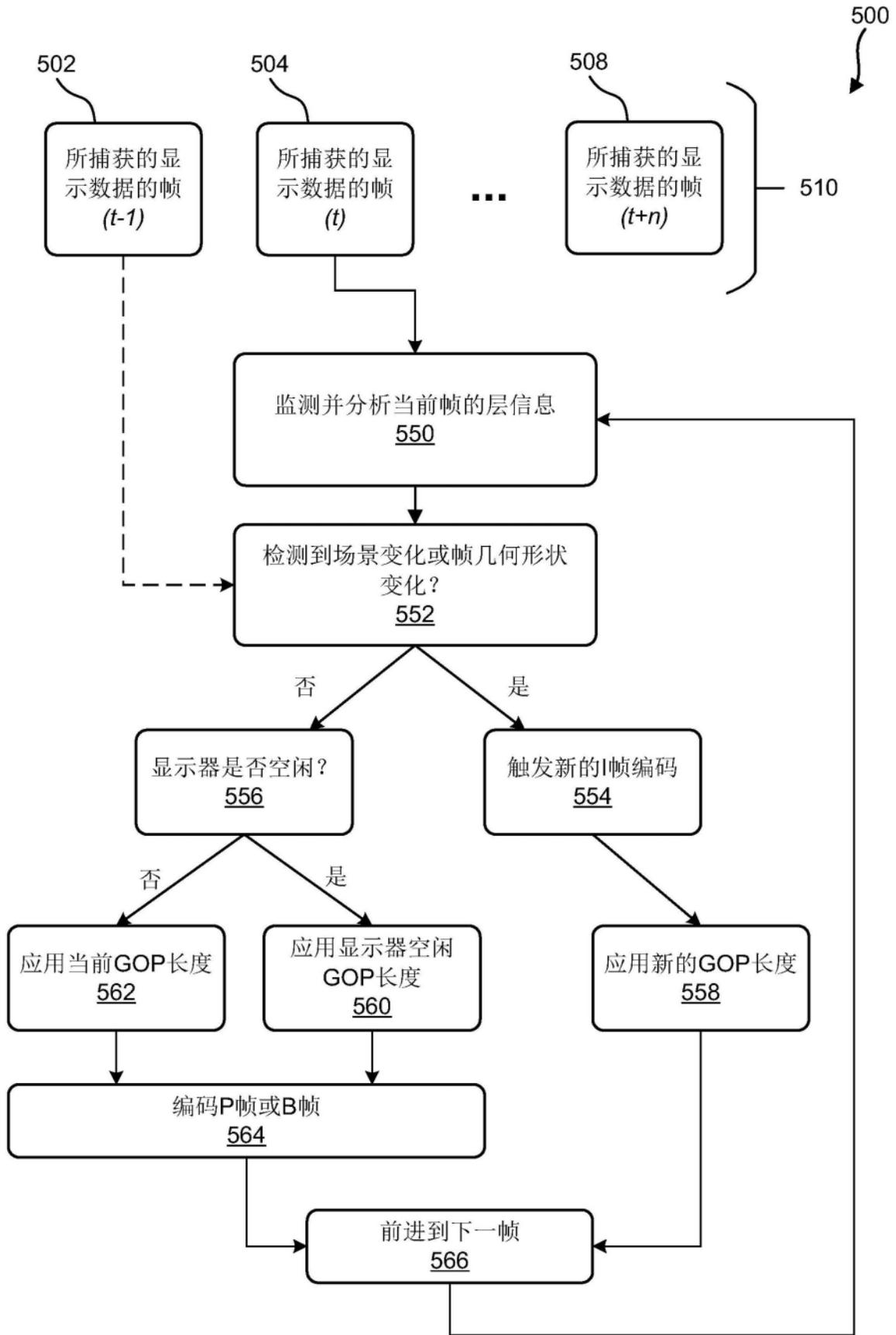


图5

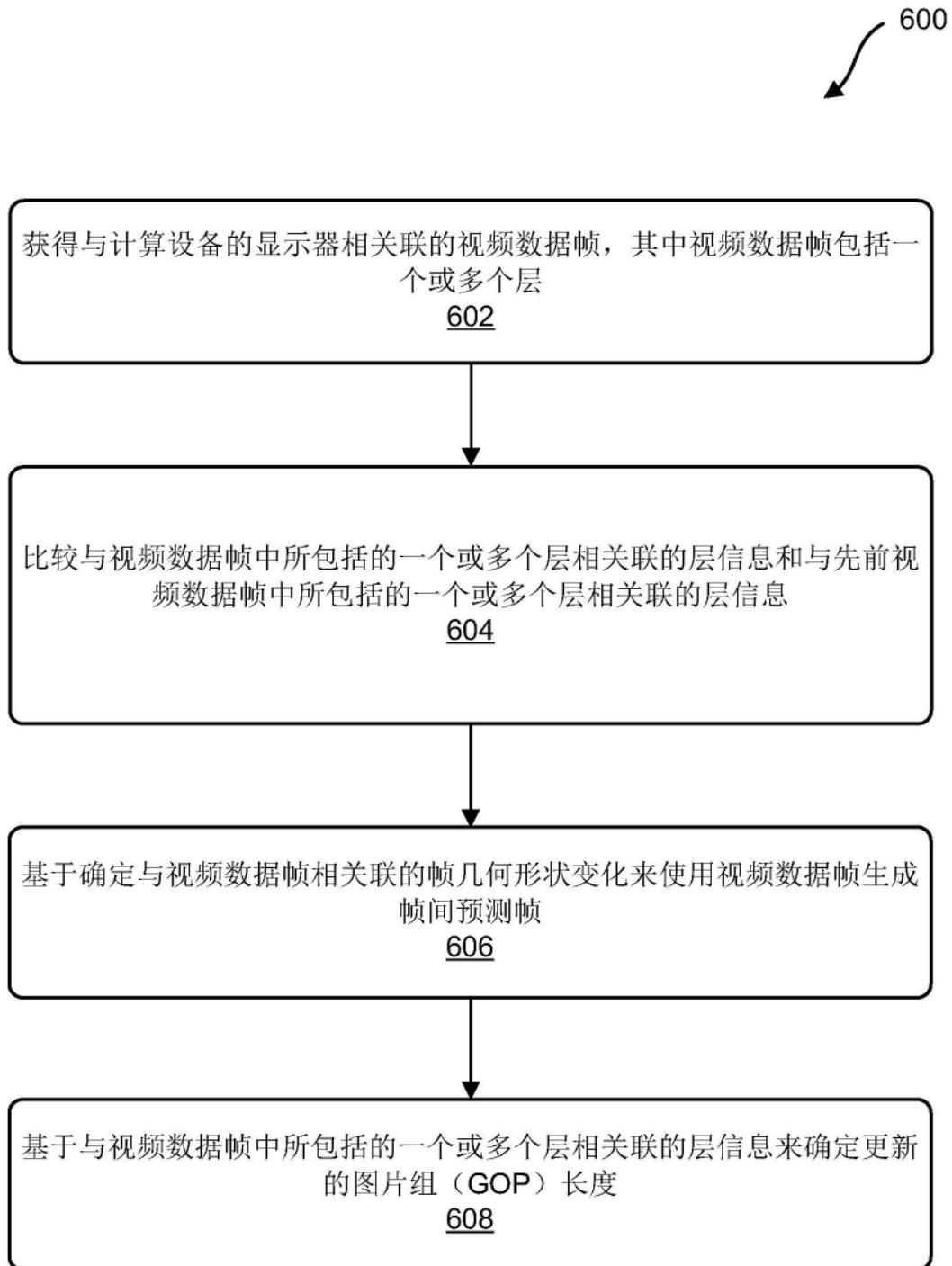


图6

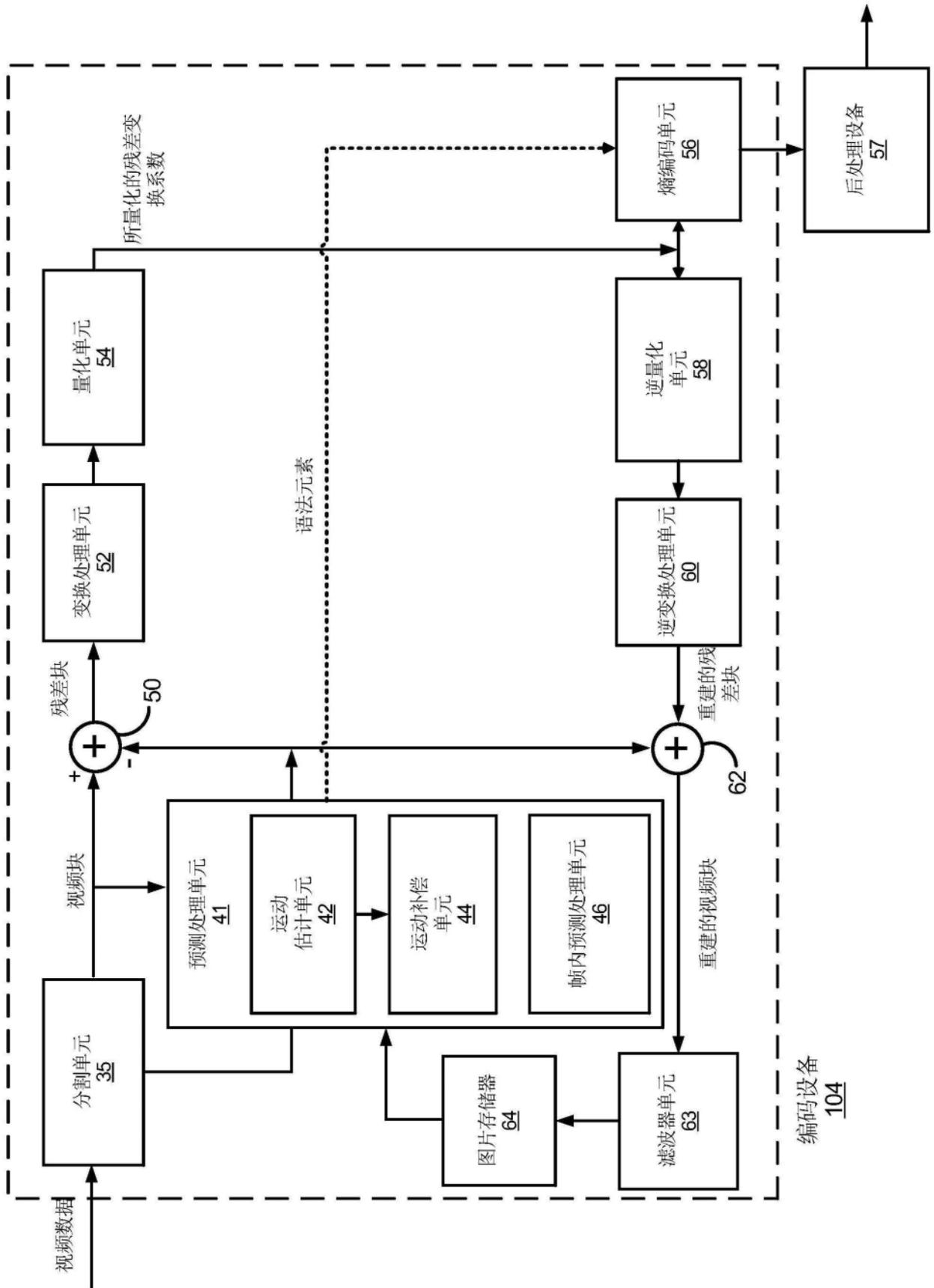


图7

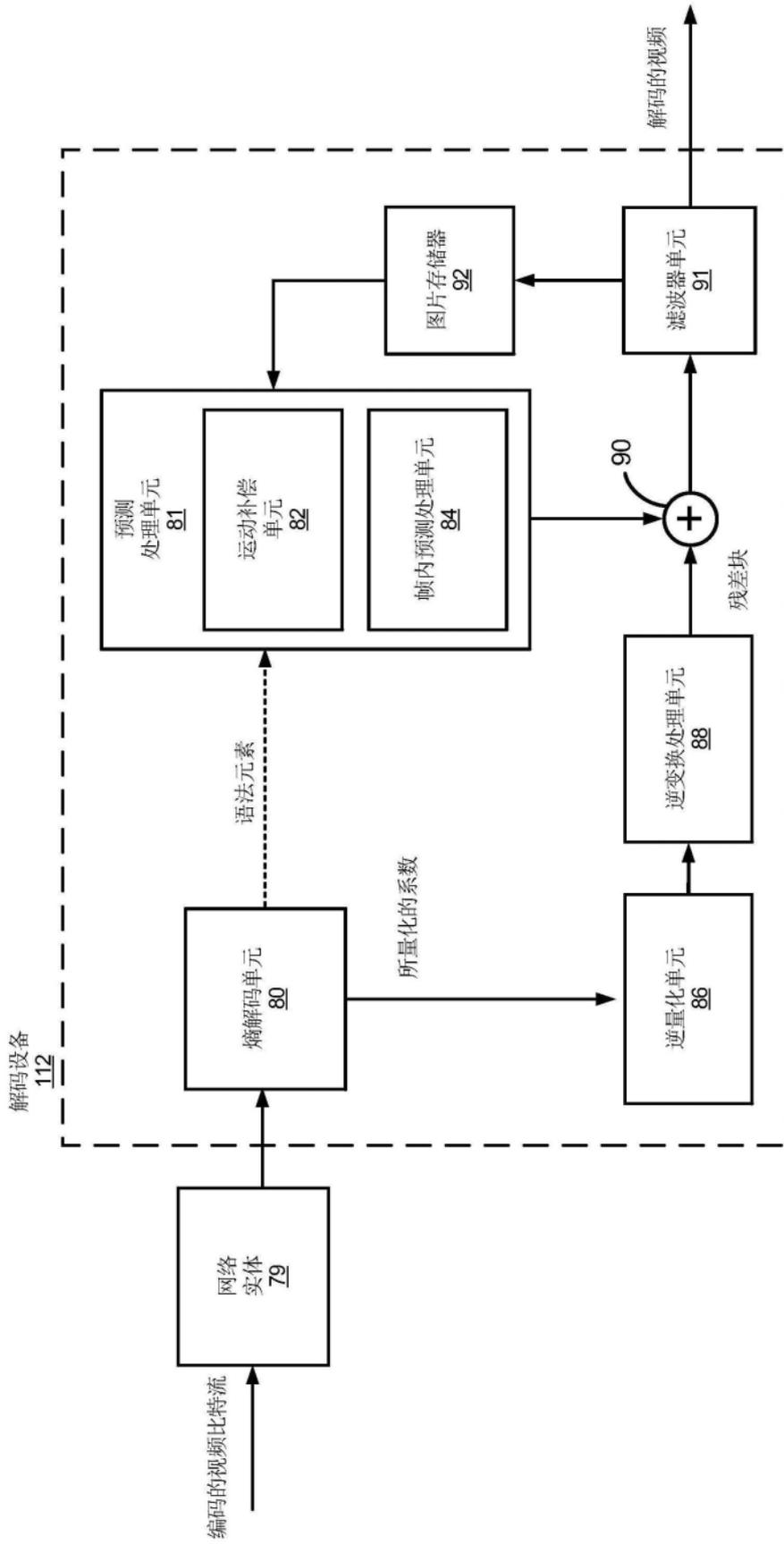


图8