(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 118827975 A
(43)申请公布日 2024.10.22

(21)申请号 202411215992.X

(22)申请日 2021.04.01

(30)优先权数据
63/005,511 2020.04.06 US
17/217,595 2021.03.30 US

(62)分案原申请数据
202180003511.7 2021.04.01

(71)申请人 腾讯美国有限责任公司
地址 美国加利福尼亚州帕洛阿尔托公园大道2747号

(72)发明人 李贵春 李翔 许晓中 刘杉

(74)专利代理机构 北京德琦知识产权代理有限公司 11018
专利代理师 刘俊英 陈舒维

(51)Int.Cl.
H04N 19/105(2014.01)
H04N 19/107(2014.01)
H04N 19/139(2014.01)
H04N 19/176(2014.01)

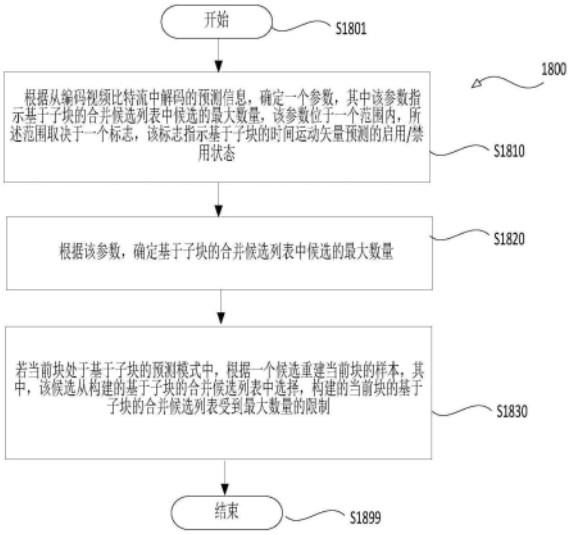
权利要求书2页 说明书25页 附图13页

(54)发明名称

视频编解码方法、装置和介质

(57)摘要

本公开提供了一种视频编解码方法、装置和介质。视频解码装置包括接收电路和处理电路。在一些示例中,所述处理电路根据从编码视频比特流解码的预测信息,确定一个参数。所述参数位于一个范围内,所述范围取决于一个标志,所述标志指示基于子块的时间运动矢量预测的启用/禁用状态。然后,所述处理电路根据所述参数,计算所述基于子块的合并候选列表中候选的最大数量,若当前块处于基于子块的预测模式中,根据一个候选重建所述当前块的样本,其中,所述候选从构建的所述当前块的基于子块的合并候选列表中选择。所述构建的当前块的基于子块的合并候选列表受到所述最大数量的限制。



CN 118827975 A

1.一种视频解码方法,其特征在于,包括:

基于编码视频比特流确定编码参数;

根据所述编码参数,确定基于子块的合并候选列表中候选的最大数量;所述编码参数的值在0到5-sps_sbtmvp_enabled_flag的范围内,其中,所述sps_sbtmvp_enabled_flag指示在所述编码视频比特流的序列参数集SPS的时间运动矢量预测的启用/禁用状态,所述sps_sbtmvp_enabled_flag等于1表示:基于子块的时间运动矢量预测值可用于解码具有非帧内编码I类型的切片的图片,所述sps_sbtmvp_enabled_flag等于0表示在编码视频中不使用基于子块的时间运动矢量预测值;

构建当前块的基于子块的合并候选列表,所述合并候选列表受到所述候选的最大数量的限制;

若当前块处于基于子块的预测模式中,从所述合并候选列表中选择一个候选,根据选择的所述候选重建所述当前块的样本。

2.根据权利要求1所述的方法,其特征在于,进一步包括:

当在所述编码视频比特流发信号通知的标志sps_affine_enabled_flag等于1时,根据默认数量和所述编码参数,确定所述基于子块的合并候选列表中候选的最大数量,其中,所述sps_affine_enabled_flag表示基于仿射模型的运动补偿是否可用于帧间预测。

3.根据权利要求2所述的方法,其特征在于,所述根据默认数量和所述编码参数,确定所述基于子块的合并候选列表中候选的最大数量包括:

通过从默认数量中减去所述编码参数,计算所述基于子块的合并候选列表中候选的最大数量。

4.根据权利要求3所述的方法,其特征在于,所述默认数量是5。

5.根据权利要求1所述的方法,其特征在于,进一步包括:

当在所述编码视频比特流发信号通知的标志sps_affine_enabled_flag等于0时,根据所述sps_sbtmvp_enabled_flag和一个标志ph_temporal_mvp_enabled_flag,确定所述基于子块的合并候选列表中候选的最大数量;其中,所述sps_affine_enabled_flag表示基于仿射模型的运动补偿是否可用于帧间预测,所述ph_temporal_mvp_enabled_flag指示在所述编码视频比特流的图片头PH级的时间运动矢量预测的启用/禁用状态。

6.根据权利要求5所述的方法,其特征在于,

所述根据所述sps_sbtmvp_enabled_flag和一个标志ph_temporal_mvp_enabled_flag,确定所述基于子块的合并候选列表中候选的最大数量,包括:

当所述sps_sbtmvp_enabled_flag和所述ph_temporal_mvp_enabled_flag中的至少一个取值为0时,确定所述基于子块的合并候选列表中候选的最大数量为0;

当所述sps_sbtmvp_enabled_flag和所述ph_temporal_mvp_enabled_flag的取值均为1时,确定所述基于子块的合并候选列表中候选的最大数量为1。

7.根据权利要求1所述的方法,其特征在于,进一步包括:

接收在所述编码视频比特流中发信号通知的所述编码参数。

8.根据权利要求1所述的方法,其特征在于,进一步包括:

若未在所述编码视频比特流中发信号通知所述编码参数,根据默认数量和所述sps_sbtmvp_enabled_flag推断所述编码参数。

9.根据权利要求1所述的方法,其特征在于,所述sps_sbtmvp_enabled_flag处于序列参数集SPS级别。

10.根据权利要求1至9任一项所述的方法,其特征在于,所述编码参数在取决于处于序列参数集SPS级别的所述sps_sbtmvp_enabled_flag和一个标志的范围内,所述标志指示在图片头PH级的时间运动矢量预测的启用/禁用状态。

11.根据权利要求10所述的方法,其特征在于,进一步包括:

若未在所述编码视频比特流中发信号通知所述编码参数,根据默认数量、所述sps_sbtmvp_enabled_flag、所述标志推断所述编码参数。

12.一种视频编码方法,用于产生编码视频比特流,其特征在于,包括:

根据视频数据,设置一个参数,所述参数用于计算基于子块的合并候选列表中候选的最大数量;所述参数的值在0到5-sps_sbtmvp_enabled_flag的范围内,其中,所述sps_sbtmvp_enabled_flag指示在所述编码视频比特流的序列参数集SPS级的时间运动矢量预测的启用/禁用状态,所述sps_sbtmvp_enabled_flag等于1表示:基于子块的时间运动矢量预测值可用于编码具有非帧内编码I类型的切片的图片,所述sps_sbtmvp_enabled_flag等于0表示在视频中不使用基于子块的时间运动矢量预测值;

构建当前块的基于子块的合并候选列表,所述合并候选列表受到所述候选的最大数量的限制;

若当前块处于基于子块的预测模式中,从所述合并候选列表中选择一个候选,根据选择的所述候选编码所述当前块的样本。

13.一种视频解码装置,其特征在于,包括处理电路,所述处理电路用于执行权利要求1至11任一项所述的方法。

14.一种视频编码装置,其特征在于,包括处理电路,所述处理电路用于执行权12所述的方法。

15.一种非易失性计算机可读介质,其特征在于,所述非易失性计算机可读介质存储指令,当所述指令由用于视频解码的计算机执行时,使得所述计算机执行权1至12任一项所述的方法。

16.一种处理视频比特流的方法,其特征在于,在非易失性计算机可读介质上存储视频比特流,所述视频比特流根据权利要求12所述的视频编码方法产生,或者基于权利要求1-11任一项所述的视频解码方法进行解码。

视频编解码方法、装置和介质

[0001] 引用并入

[0002] 本申请要求于2021年3月30日提交的美国专利申请第17/217,595号“视频编解码的方法和装置”的优先权,该美国申请要求了于2020年4月6日递交的美国临时申请第63/005,511号“设置子块合并候选数的方法”的优先权。在先申请的全部内容通过引用并入本文。

技术领域

[0003] 本公开描述了与视频编解码相关的实施例。

背景技术

[0004] 本文所提供的背景描述旨在整体呈现本申请的背景。在背景技术部分以及本说明书的各个方面中所描述的目前已署名的发明人的工作所进行的程度,并不表明其在本申请提交时作为现有技术,且从未明示或暗示其被承认为本申请的现有技术。

[0005] 通过具有运动补偿的帧间图片预测技术,可以进行视频编码和解码。未压缩的数字视频可包括一系列图片,每个图片具有例如1920×1080亮度样本及相关色度样本的空间维度。所述系列图片具有固定的或可变的图片速率(也非正式地称为帧率),例如每秒60个图片或60Hz。未压缩的视频具有非常大的比特率要求。例如,每个样本8比特的1080p60 4:2:0的视频(1920x1080亮度样本分辨率,60Hz帧率)要求接近1.5Gbit/s带宽。一小时这样的视频就需要超过600GB的存储空间。

[0006] 视频编码和解码的一个目的,是通过压缩减少输入视频信号的冗余信息。视频压缩可以帮助降低对上述带宽或存储空间的要求,在某些情况下可降低两个或更多数量级。无损和有损压缩,以及两者的组合均可采用。无损压缩是指从压缩的原始信号中重建原始信号精确副本的技术。当使用有损压缩时,重建信号可能与原始信号不完全相同,但是原始信号和重建信号之间的失真足够小,使得重建信号可用于预期应用。有损压缩广泛应用于视频。容许的失真量取决于应用。例如,相比于电视应用的用户,某些消费流媒体应用的用户可以容忍更高的失真。可实现的压缩比反映出:较高的允许/容许失真可产生较高的压缩比。

[0007] 运动补偿可以是一种有损压缩技术,且可涉及如下技术:来自先前重建的图片或重建图片一部分(参考图片)的样本数据块在空间上按运动矢量(下文称为MV)指示的方向移位后,用于新重建的图片或图片部分的预测。在某些情况下,参考图片可与当前正在重建的图片相同。MV可具有两个维度X和Y,或者三个维度,其中第三个维度表示使用中的参考图片(后者间接地可为时间维度)。

[0008] 在一些视频压缩技术中,应用于某个样本数据区域的MV可根据其它MV来预测,例如根据与正在重建的区域空间相邻的另一个样本数据区域相关的、且按解码顺序在该MV前面的那些MV。这样做可以大大减少编码MV所需的数据量,从而消除冗余信息并增加压缩量。MV预测可以有效地进行,例如,当对从相机导出的输入视频信号(称为自然视频)进行编码

时,存在一种统计上的可能性,即面积大于单个MV适用区域的区域,会朝着类似的方向移动,因此,在某些情况下,可以用邻近区域的MV导出的相似运动矢量进行预测。这导致针对给定区域发现的MV与根据周围MV预测的MV相似或相同,并且在熵编码之后,又可以用比直接编码MV时使用的比特数更少的比特数来表示。在某些情况下,MV预测可以是对从原始信号(即样本流)导出的信号(即MV)进行无损压缩的示例。在其它情况下,MV预测本身可能是有损的,例如由于根据几个周围MV计算预测值时产生的取整误差。

[0009] H.265/HEVC(ITU-T Rec.H.265,“高效视频编码”,2016年12月)描述了各种MV预测机制。在H.265所提供的多种MV预测机制中,本文描述的是一种下文称为“空间合并”的技术。

[0010] 参照图1,当前块(101)包括编码器在运动搜索过程中发现的样本,所述样本可以根据空间移动了相同大小的先前块进行预测。不直接对该MV进行编码,而是通过使用与五个周围样本中的任何一个相关联的MV,从与一个或多个参考图片相关联的元数据中导出该MV,例如从最近的(按解码顺序)参考图片中导出该MV。其中,五个周围样本分别用A0、A1和B0、B1、B2(分别为102到106)表示。在H.265中,MV预测可使用相邻块正在使用的同一参考图片的预测值。

发明内容

[0011] 本公开的多个方面提供了视频编码/解码的方法和装置。在一些示例中,视频解码装置包括接收电路和处理电路。例如,所述处理电路根据从编码视频比特流解码的预测信息,确定一个参数。所述参数位于一个范围内,所述范围取决于一个标志,所述标志指示基于子块的时间运动矢量预测的启用/禁用状态。然后,所述处理电路根据所述参数,计算所述基于子块的合并候选列表中候选的最大数量,若当前块处于基于子块的预测模式中,根据一个候选重建所述当前块的样本,其中,所述候选从构建的所述当前块的基于子块的合并候选列表中选择。所述构建的当前块的基于子块的合并候选列表受到所述基于子块的合并候选列表中候选的最大数量的限制。

[0012] 在一些示例中,所述处理电路通过从默认数量中减去所述参数,确定所述基于子块的合并候选列表中候选的最大数量。在一个示例中,所述默认数量是5。

[0013] 在一些实施例中,所述范围的上限取决于所述标志,所述标志指示所述基于子块的时间运动矢量预测的启用/禁用状态。

[0014] 在一个示例中,所述处理电路接收在所述编码视频比特流中发信号通知的所述参数。

[0015] 在另一个示例中,若未在所述编码视频比特流中发信号通知所述参数,所述处理电路根据默认数量和所述标志推断所述参数,其中,所述标志指示基于子块的时间运动矢量预测的启用/禁用状态。

[0016] 在一些示例中,所述标志指示在序列参数集SPS级的所述基于子块的时间运动矢量预测的启用/禁用状态。

[0017] 在一些实施例中,所述参数在取决于第一标志和第二标志的范围内,所述第一标志指示在序列参数集SPS级的所述基于子块的时间运动矢量预测的启用/禁用状态,所述第二标志指示在图片头PH级的时间运动矢量预测的启用/禁用状态。

[0018] 在一些示例中,若未在所述编码视频比特流中发信号通知所述参数,所述处理电路可以根据默认数量、所述第一标志、所述第二标志推断所述参数,所述第一标志指示在所述序列参数集SPS级的所述基于子块的时间运动矢量预测的启用/禁用状态,并且所述第二标志指示在所述图片头PH级的所述时间运动矢量预测的启用/禁用状态。

[0019] 本公开的多个方面还提供了一种存储指令的非易失性计算机可读介质,其中,当所述指令由用于视频解码的计算机执行时,使得所述计算机执行上述视频解码方法。

附图说明

[0020] 根据以下详细描述和附图,所公开的主题的其他特征、性质和各种优点将进一步明确,其中:

[0021] 图1是一个示例中当前块及其周围空间合并候选的示意图。

[0022] 图2是根据实施例的通信系统的简化框图的示意图。

[0023] 图3是根据实施例的通信系统的简化框图的示意图。

[0024] 图4是根据实施例的解码器的简化框图的示意图。

[0025] 图5是根据实施例的编码器的简化框图的示意图。

[0026] 图6示出了根据另一个实施例的编码器的框图。

[0027] 图7示出了根据另一个实施例的解码器的框图。

[0028] 图8A-8B示出了仿射运动模型。

[0029] 图9示出了每个子块的仿射运动矢量场的示例。

[0030] 图10示出了在仿射合并模式中确定继承控制点运动矢量的示例。

[0031] 图11示出了根据本公开一些实施例的空间邻居和时间邻居的示例。

[0032] 图12示出了在一些示例中用于仿射运动数据继承的运动矢量使用的图。

[0033] 图13示出了根据本公开一个实施例的使用光流方法进行预测修正示例的示意图。

[0034] 图14-15示出了根据本公开一些实施例的基于子块的时间运动矢量预测(subblock based temporal motion vector prediction,SbTMVP)过程的示例。

[0035] 图16示出了在一些示例中序列参数集(sequence parameter set,SPS)的语法表示例。

[0036] 图17示出了在一些示例中图片头结构的语法表示例。

[0037] 图18示出了根据本公开一些实施例的概述过程示例的流程图。

[0038] 图19是根据实施例的计算机系统的示意图。

具体实施方式

[0039] 图2是根据本公开实施例的通信系统(200)的简化框图。通信系统(200)包括可以经由,例如网络(250),彼此通信的多个终端装置。举例来说,通信系统(200)包括通过网络(250)互连的第一对终端装置(210)和终端装置(220)。在图2的示例中,第一对终端装置(210)和终端装置(220)执行单向数据传输。举例来说,终端装置(210)可对视频数据(例如由终端装置(210)采集的视频图片流)进行编码,以通过网络(250)传输到另一个终端装置(220)。已编码的视频数据以一个或多个已编码视频码流形式传输。终端装置(220)可从网络(250)接收已编码视频数据,对已编码视频数据进行解码以恢复视频数据,并根据恢复的

视频数据显示视频图片。单向数据传输在媒体服务等应用中是较常见的。

[0040] 在另一实施例中,通信系统(200)包括执行已编码视频数据的双向传输的第二对终端装置(230)和终端装置(240),所述双向传输可例如在视频会议期间发生。对于双向数据传输,在一个示例中,终端装置(230)和终端装置(240)中的每个终端装置可对视频数据(例如由终端装置采集的视频图片流)进行编码,以通过网络(250)传输到终端装置(230)和终端装置(240)中的另一终端装置。终端装置(230)和终端装置(240)中的每个终端装置还可接收由终端装置(230)和终端装置(240)中的另一终端装置传输的已编码视频数据,且可对所述已编码视频数据进行解码以恢复视频数据,且可根据恢复的视频数据在可访问的显示装置上显示视频图片。

[0041] 在图2的实施例中,终端装置(210)、终端装置(220)、终端装置(230)和终端装置(240)可为服务器、个人计算机和智能电话,但本申请公开的原理可不限于此。本申请公开的实施例适用于膝上型计算机、平板电脑、媒体播放器和/或专用视频会议设备。网络(250)表示在终端装置(210)、终端装置(220)、终端装置(230)和终端装置(240)之间传送已编码视频数据的任何数目的网络,包括例如有线(连线的)和/或无线通信网络。通信网络(250)可在电路交换和/或分组交换信道中交换数据。该网络可包括电信网络、局域网、广域网和/或互联网。出于本申请的目的,除非在下文中有所解释,否则网络(250)的架构和拓扑对于本申请公开的操作来说可能是无关紧要的。

[0042] 作为实施例,图3示出视频编码器和视频解码器在流式传输环境中的放置方式。本申请所公开主题可同等地适用于其它支持视频的应用,包括例如视频会议、数字TV、在包括CD、DVD、存储棒等的数字介质上存储压缩视频等等。

[0043] 流式传输系统可包括采集子系统(313),所述采集子系统可包括数码相机等视频源(301),所述视频源创建未压缩的视频图片流(302)。在实施例中,视频图片流(302)包括由数码相机拍摄的样本。相较于已编码的视频数据(304)(或已编码的视频码流),视频图片流(302)被描绘为粗线以强调高数据量的视频图片流,视频图片流(302)可由电子装置(320)处理,所述电子装置(320)包括耦接到视频源(301)的视频编码器(303)。视频编码器(303)可包括硬件、软件或软硬件组合以实现或实施如下文更详细地描述的所公开主题的各方面。相较于视频图片流(302),已编码的视频数据(304)(或已编码的视频码流(304))被描绘为细线以强调较低数据量的已编码的视频数据(304)(或已编码的视频码流(304)),其可存储在流式传输服务器(305)上以供将来使用。一个或多个流式传输客户端子系统,例如图3中的客户端子系统(306)和客户端子系统(308),可访问流式传输服务器(305)以检索已编码的视频数据(304)的副本(307)和副本(309)。客户端子系统(306)可包括例如电子装置(330)中的视频解码器(310)。视频解码器(310)对已编码的视频数据的传入副本(307)进行解码,且产生可在显示器(312)(例如显示屏)或另一呈现装置(未描绘)上呈现的输出视频图片流(311)。在一些流式传输系统中,可根据某些视频编码/压缩标准对已编码的视频数据(304)、视频数据(307)和视频数据(309)(例如视频码流)进行编码。该些标准的实施例包括ITU-T H.265。在实施例中,正在开发的视频编码标准非正式地称为下一代视频编码(Versatile Video Coding,VVC),本申请可用于VVC标准的上下文中。

[0044] 应注意,电子装置(320)和电子装置(330)可包括其它组件(未示出)。举例来说,电子装置(320)可包括视频解码器(未示出),且电子装置(330)还可包括视频编码器(未示

出)。

[0045] 图4是根据本申请公开的实施例的视频解码器(410)的框图。视频解码器(410)可设置在电子装置(430)中。电子装置(430)可包括接收器(431)(例如接收电路)。视频解码器(410)可用于代替图3实施例中的视频解码器(310)。

[0046] 接收器(431)可接收将由视频解码器(410)解码的一个或多个已编码视频序列;在同一实施例或另一实施例中,一次接收一个已编码视频序列,其中每个已编码视频序列的解码独立于其它已编码视频序列。可从信道(401)接收已编码视频序列,所述信道可以是通向存储已编码的视频数据的存储装置的硬件/软件链路。接收器(431)可接收已编码的视频数据以及其它数据,例如,可转发到它们各自的使用实体(未标示)的已编码音频数据和/或辅助数据流。接收器(431)可将已编码视频序列与其它数据分开。为了防止网络抖动,缓冲存储器(415)可耦接在接收器(431)与熵解码器/解析器(420)(此后称为“解析器(420)”)之间。在某些应用中,缓冲存储器(415)是视频解码器(410)的一部分。在其它情况下,所述缓冲存储器(415)可设置在视频解码器(410)外部(未标示)。而在其它情况下,视频解码器(410)的外部设置缓冲存储器(未标示)以例如防止网络抖动,且在视频解码器(410)的内部可配置另一缓冲存储器(415)以例如处理播出定时。而当接收器(431)从具有足够带宽和可控性的存储/转发装置或从等时同步网络接收数据时,也可能不需要配置缓冲存储器(415),或可以将所述缓冲存储器做得较小。当然,为了在互联网等业务分组网络上使用,也可能需要缓冲存储器(415),所述缓冲存储器可相对较大且可具有自适应性大小,且可至少部分地实施于操作系统或视频解码器(410)外部的类似元件(未标示)中。

[0047] 视频解码器(410)可包括解析器(420)以根据已编码视频序列重建符号(421)。这些符号的类别包括用于管理视频解码器(410)的操作的信息,以及用以控制显示装置(412)(例如,显示屏)等显示装置的潜在信息,所述显示装置不是电子装置(430)的组成部分,但可耦接到电子装置(430),如图4中所示。用于显示装置的控制信息可以是辅助增强信息(Supplemental Enhancement Information,SEI消息)或视频可用性信息(Video Usability Information,VUI)的参数集片段(未标示)。解析器(420)可对接收到的已编码视频序列进行解析/熵解码。已编码视频序列的编码可根据视频编码技术或标准进行,且可遵循各种原理,包括可变长度编码、霍夫曼编码(Huffman coding)、具有或不具有上下文灵敏度的算术编码等等。解析器(420)可基于对应于群组的至少一个参数,从已编码视频序列提取用于视频解码器中的像素的子群中的至少一个子群的子群参数集。子群可包括图片群组(Group of Pictures,GOP)、图片、图块、切片、宏块、编码单元(Coding Unit,CU)、块、变换单元(Transform Unit,TU)、预测单元(Prediction Unit,PU)等等。解析器(420)还可从已编码视频序列提取信息,例如变换系数、量化器参数值、运动矢量等等。

[0048] 解析器(420)可对从缓冲存储器(415)接收的视频序列执行熵解码/解析操作,从而创建符号(421)。

[0049] 取决于已编码视频图片或一部分已编码视频图片(例如:帧间图片和帧内图片、帧间块和帧内块)的类型以及其它因素,符号(421)的重建可涉及多个不同单元。涉及哪些单元以及涉及方式可由解析器(420)从已编码视频序列解析的子群控制信息控制。为了简洁起见,未描述解析器(420)与下文的多个单元之间的此类子群控制信息流。

[0050] 除已经提及的功能块以外,视频解码器(410)可在概念上细分成如下文所描述的

数个功能单元。在商业约束下运行的实际实施例中,这些单元中的许多单元彼此紧密交互并且可以彼此集成。然而,出于描述所公开主题的目的,概念上细分成下文的功能单元是适当的。

[0051] 第一单元是缩放器/逆变换单元(451)。缩放器/逆变换单元(451)从解析器(420)接收作为符号(421)的量化变换系数以及控制信息,包括使用哪种变换方式、块大小、量化因子、量化缩放矩阵等。缩放器/逆变换单元(451)可输出包括样本值的块,所述样本值可输入到聚合器(455)中。

[0052] 在一些情况下,缩放器/逆变换单元(451)的输出样本可属于帧内编码块;即:不使用来自先前重建的图片的预测性信息,但可使用来自当前图片的先前重建部分的预测性信息的块。此类预测性信息可由帧内图片预测单元(452)提供。在一些情况下,帧内图片预测单元(452)采用从当前图片缓冲器(458)提取的已重建信息生成大小和形状与正在重建的块相同的周围块。举例来说,当前图片缓冲器(458)缓冲部分重建的当前图片和/或完全重建的当前图片。在一些情况下,聚合器(455)基于每个样本,将帧内预测单元(452)生成的预测信息添加到由缩放器/逆变换单元(451)提供的输出样本信息中。

[0053] 在其它情况下,缩放器/逆变换单元(451)的输出样本可属于帧间编码和潜在运动补偿块。在此情况下,运动补偿预测单元(453)可访问参考图片存储器(457)以提取用于预测的样本。在根据符号(421)对提取的样本进行运动补偿之后,这些样本可由聚合器(455)添加到缩放器/逆变换单元(451)的输出(在这种情况下被称作残差样本或残差信号),从而生成输出样本信息。运动补偿预测单元(453)从参考图片存储器(457)内的地址获取预测样本可受到运动矢量控制,且所述运动矢量以所述符号(421)的形式而供运动补偿预测单元(453)使用,所述符号(421)例如是包括X、Y和参考图片分量。运动补偿还可包括在使用子样本精确运动矢量时,从参考图片存储器(457)提取的样本值的内插、运动矢量预测机制等等。

[0054] 聚合器(455)的输出样本可在环路滤波器单元(456)中被各种环路滤波技术采用。视频压缩技术可包括环路内滤波器技术,所述环路内滤波器技术受控于包括在已编码视频序列(也称作已编码视频码流)中的参数,且所述参数作为来自解析器(420)的符号(421)可用于环路滤波器单元(456)。然而,在其他实施例中,视频压缩技术还可响应于在解码已编码图片或已编码视频序列的先前(按解码次序)部分期间获得的元信息,以及响应于先前重建且经过环路滤波的样本值。

[0055] 环路滤波器单元(456)的输出可以是样本流,所述样本流可输出到显示装置(412)以及存储在参考图片存储器(457),以用于后续的帧间图片预测。

[0056] 一旦完全重建,某些已编码图片就可用作参考图片以用于将来预测。举例来说,一旦对应于当前图片的已编码图片被完全重建,且已编码图片(通过例如解析器(420))被识别为参考图片,则当前图片缓冲器(458)可变为参考图片存储器(457)的一部分,且可在开始重建后续已编码图片之前重新分配新的当前图片缓冲器。

[0057] 视频解码器(410)可根据例如ITU-T H.265标准中的预定视频压缩技术执行解码操作。在已编码视频序列遵循视频压缩技术或标准的语法以及视频压缩技术或标准中记录的配置文件的意义上,已编码视频序列可符合所使用的视频压缩技术或标准指定的语法。具体地说,配置文件可从视频压缩技术或标准中可用的所有工具中选择某些工具作为在所

述配置文件下可供使用的仅有工具。对于合规性,还要求已编码视频序列的复杂度处于视频压缩技术或标准的层级所限定的范围内。在一些情况下,层级限制最大图片大小、最大帧率、最大重建取样率(以例如每秒兆(mega)个样本为单位进行测量)、最大参考图片大小等。在一些情况下,由层级设定的限制可通过假想参考解码器(Hypothetical Reference Decoder,HRD)规范和在已编码视频序列中用信号表示的HRD缓冲器管理的元数据来进一步限定。

[0058] 在实施例中,接收器(431)可连同已编码视频一起接收附加(冗余)数据。所述附加数据可以是已编码视频序列的一部分。所述附加数据可由视频解码器(410)用以对数据进行适当解码和/或较准确地重建原始视频数据。附加数据可呈例如时间、空间或信噪比(signal noise ratio,SNR)增强层、冗余切片、冗余图片、前向纠错码等形式。

[0059] 图5是根据本申请公开的实施例的视频编码器(503)的框图。视频编码器(503)设置于电子装置(520)中。电子装置(520)包括传输器(540)(例如传输电路)。视频编码器(503)可用于代替图3实施例中的视频编码器(303)。

[0060] 视频编码器(503)可从视频源(501)(并非图5实施例中的电子装置(520)的一部分)接收视频样本,所述视频源可采集将由视频编码器(503)编码的视频图像。在另一实施例中,视频源(501)是电子装置(520)的一部分。

[0061] 视频源(501)可提供将由视频编码器(503)编码的呈数字视频样本流形式的源视频序列,所述数字视频样本流可具有任何合适位深度(例如:8位、10位、12位……)、任何色彩空间(例如BT.601Y CrCB、RGB……)和任何合适取样结构(例如Y CrCb 4:2:0、Y CrCb 4:4:4)。在媒体服务系统中,视频源(501)可以是存储先前已准备的视频的存储装置。在视频会议系统中,视频源(501)可以是采集本地图像信息作为视频序列的相机。可将视频数据提供为多个单独的图片,当按顺序观看时,这些图片被赋予运动。图片自身可构建为空间像素阵列,其中取决于所用的取样结构、色彩空间等,每个像素可包括一个或多个样本。所属领域的技术人员可以很容易理解像素与样本之间的关系。下文侧重于描述样本。

[0062] 根据实施例,视频编码器(503)可实时或在由应用所要求的任何其它时间约束下,将源视频序列的图片编码且压缩成已编码视频序列(543)。施行适当的编码速度是控制器(550)的一个功能。在一些实施例中,控制器(550)控制如下文所描述的其它功能单元且在功能上耦接到这些单元。为了简洁起见,图中未标示耦接。由控制器(550)设置的参数可包括速率控制相关参数(图片跳过、量化器、率失真优化技术的λ值等)、图片大小、图片群组(group of pictures,GOP)布局,最大运动矢量搜索范围等。控制器(550)可用于具有其它合适的功能,这些功能涉及针对某一系统设计优化的视频编码器(503)。

[0063] 在一些实施例中,视频编码器(503)在编码环路中进行操作。作为简单的描述,在实施例中,编码环路可包括源编码器(530)(例如,负责基于待编码的输入图片和参考图片创建符号,例如符号流)和嵌入于视频编码器(503)中的(本地)解码器(533)。解码器(533)以类似于(远程)解码器创建样本数据的方式重建符号以创建样本数据(因为在本申请所考虑的视频压缩技术中,符号与已编码视频码流之间的任何压缩是无损的)。将重建的样本流(样本数据)输入到参考图片存储器(534)。由于符号流的解码产生与解码器位置(本地或远程)无关的位精确结果,因此参考图片存储器(534)中的内容在本地编码器与远程编码器之间也是按比特位精确对应的。换句话说,编码器的预测部分“看到”的参考图片样本与解码

器将在解码期间使用预测时所“看到”的样本值完全相同。这种参考图片同步性基本原理(以及在例如因信道误差而无法维持同步性的情况下产生的漂移)也用于一些相关技术。

[0064] “本地”解码器(533)的操作可与例如已在上文结合图4详细描述视频解码器(410)的“远程”解码器相同。然而,另外简要参考图4,当符号可用且熵编码器(545)和解析器(420)能够无损地将符号编码/解码为已编码视频序列时,包括缓冲存储器(415)和解析器(420)在内的视频解码器(410)的熵解码部分,可能无法完全在本地解码器(533)中实施。

[0065] 此时可以观察到,除存在于解码器中的解析/熵解码之外的任何解码器技术,也必定以基本上相同的功能形式存在于对应的编码器中。出于此原因,本申请侧重于解码器操作。可简化编码器技术的描述,因为编码器技术与全面地描述的解码器技术互逆。仅在某些区域中需要更详细的描述,并且在下文提供。

[0066] 在操作期间,在一些实施例中,源编码器(530)可执行运动补偿预测编码。参考来自视频序列中被指定为“参考图片”的一个或多个先前已编码图片,所述运动补偿预测编码对输入图片进行预测性编码。以此方式,编码引擎(532)对输入图片的像素块与参考图片的像素块之间的差异进行编码,所述参考图片可被选作所述输入图片的预测参考。

[0067] 本地视频解码器(533)可基于源编码器(530)创建的符号,对可指定为参考图片的图片的已编码视频数据进行解码。编码引擎(532)的操作可为有损过程。当已编码视频数据可在视频解码器(图5中未示)处被解码时,重建的视频序列通常可以是带有一些误差的源视频序列的副本。本地视频解码器(533)复制解码过程,所述解码过程可由视频解码器对参考图片执行,且可使重建的参考图片存储在参考图片高速缓存(534)中。以此方式,视频编码器(503)可在本地存储重建的参考图片的副本,所述副本与将由远端视频解码器获得的重建参考图片具有共同内容(不存在传输误差)。

[0068] 预测器(535)可针对编码引擎(532)执行预测搜索。即,对于将要编码的新图片,预测器(535)可在参考图片存储器(534)中搜索可作为所述新图片的适当预测参考的样本数据(作为候选参考像素块)或某些元数据,例如参考图片运动矢量、块形状等。预测器(535)可基于样本块逐像素块操作,以找到合适的预测参考。在一些情况下,根据预测器(535)获得的搜索结果,可确定输入图片可具有从参考图片存储器(534)中存储的多个参考图片取得的预测参考。

[0069] 控制器(550)可管理源编码器(530)的编码操作,包括例如设置用于对视频数据进行编码的参数和子群参数。

[0070] 可在熵编码器(545)中对所有上述功能单元的输出进行熵编码。熵编码器(545)根据例如霍夫曼编码、可变长度编码、算术编码等技术对各种功能单元生成的符号进行无损压缩,从而将所述符号转换成已编码视频序列。

[0071] 传输器(540)可缓冲由熵编码器(545)创建的已编码视频序列,从而为通过通信信道(560)进行传输做准备,所述通信信道可以是通向将存储已编码的视频数据的存储装置的硬件/软件链路。传输器(540)可将来自视频编码器(503)的已编码视频数据与要传输的其它数据合并,所述其它数据例如是已编码音频数据和/或辅助数据流(未示出来源)。

[0072] 控制器(550)可管理视频编码器(503)的操作。在编码期间,控制器(550)可以为每个已编码图片分配某一已编码图片类型,但这可能影响可应用于相应的图片的编码技术。例如,通常可将图片分配为以下任一种图片类型:

[0073] 帧内图片(I图片),其可以是不将序列中的任何其它图片用作预测源就可被编码和解码的图片。一些视频编解码器容许不同类型的帧内图片,包括例如独立解码器刷新(Independent Decoder Refresh,“IDR”)图片。所属领域的技术人员了解I图片的变体及其相应的应用和特征。

[0074] 预测性图片(P图片),其可以是可使用帧内预测或帧间预测进行编码和解码的图片,所述帧内预测或帧间预测使用至多一个运动矢量和参考索引来预测每个块的样本值。

[0075] 双向预测性图片(B图片),其可以是可使用帧内预测或帧间预测进行编码和解码的图片,所述帧内预测或帧间预测使用至多两个运动矢量和参考索引来预测每个块的样本值。类似地,多个预测性图片可使用多于两个参考图片和相关联元数据以用于重建单个块。

[0076] 源图片通常可在空间上细分成多个样本块(例如,4×4、8×8、4×8或16×16个样本的块),且逐块进行编码。这些块可参考其它(已编码)块进行预测编码,根据应用于块的相应图片的编码分配来确定所述其它块。举例来说,I图片的块可进行非预测编码,或所述块可参考同一图片的已经编码的块来进行预测编码(空间预测或帧内预测)。P图片的像素块可参考一个先前编码的参考图片通过空间预测或通过时域预测进行预测编码。B图片的块可参考一个或两个先前编码的参考图片通过空间预测或通过时域预测进行预测编码。

[0077] 视频编码器(503)可根据例如ITU-T H.265建议书的预定视频编码技术或标准执行编码操作。在操作中,视频编码器(503)可执行各种压缩操作,包括利用输入视频序列中的时间和空间冗余的预测编码操作。因此,已编码视频数据可符合所用视频编码技术或标准指定的语法。

[0078] 在实施例中,传输器(540)可在传输已编码的视频时传输附加数据。源编码器(530)可将此类数据作为已编码视频序列的一部分。附加数据可包括时间/空间/SNR增强层、冗余图片和切片等其它形式的冗余数据、SEI消息、VUI参数集片段等。

[0079] 采集到的视频可作为呈时间序列的多个源图片(视频图片)。帧内图片预测(常常简化为帧内预测)利用给定图片中的空间相关性,而帧间图片预测则利用图片之间的(时间或其它)相关性。在实施例中,将正在编码/解码的特定图片分割成块,正在编码/解码的特定图片被称作当前图片。在当前图片中的块类似于视频中先前已编码且仍被缓冲的参考图片中的参考块时,可通过称作运动矢量的矢量对当前图片中的块进行编码。所述运动矢量指向参考图片中的参考块,且在使用多个参考图片的情况下,所述运动矢量可具有识别参考图片的第三维度。

[0080] 在一些实施例中,双向预测技术可用于帧间图片预测中。根据双向预测技术,使用两个参考图片,例如按解码次序都在视频中的当前图片之前(但按显示次序可能分别是过去和将来)第一参考图片和第二参考图片。可通过指向第一参考图片中的第一参考块的第一运动矢量和指向第二参考图片中的第二参考块的第二运动矢量对当前图片中的块进行编码。具体来说,可通过第一参考块和第二参考块的组合来预测所述块。

[0081] 此外,合并模式技术可用于帧间图片预测中以改善编码效率。

[0082] 根据本申请公开的一些实施例,帧间图片预测和帧内图片预测等预测的执行以块为单位。举例来说,根据HEVC标准,将视频图片序列中的图片分割成编码树单元(coding tree unit,CTU)以用于压缩,图片中的CTU具有相同大小,例如64×64像素、32×32像素或16×16像素。一般来说,CTU包括三个编码树块(coding tree block,CTB),所述三个编码树

块是一个亮度CTB和两个色度CTB。更进一步的,还可将每个CTU以四叉树拆分为一个或多个编码单元(coding unit,CU)。举例来说,可将64×64像素的CTU拆分为一个64×64像素的CU,或4个32×32像素的CU,或16个16×16像素的CU。在实施例中,分析每个CU以确定用于CU的预测类型,例如帧间预测类型或帧内预测类型。此外,取决于时间和/或空间可预测性,将CU拆分为一个或多个预测单元(prediction unit,PU)。通常,每个PU包括亮度预测块(prediction block,PB)和两个色度PB。在实施例中,编码(编码/解码)中的预测操作以预测块为单位来执行。以亮度预测块作为预测块为例,预测块包括像素值(例如,亮度值)的矩阵,例如8×8像素、16×16像素、8×16像素、16×8像素等等。

[0083] 图6是根据本申请公开的另一实施例的视频编码器(603)的图。视频编码器(603)用于接收视频图片序列中的当前视频图片内的样本值的处理块(例如预测块),且将所述处理块编码到作为已编码视频序列的一部分的已编码图片中。在本实施例中,视频编码器(603)用于代替图3实施例中的视频编码器(303)。

[0084] 在HEVC实施例中,视频编码器(603)接收用于处理块的样本值的矩阵,所述处理块为例如8×8样本的预测块等。视频编码器(603)使用例如率失真(rate-distortion,RD)优化来确定是否使用帧内模式、帧间模式或双向预测模式来编码所述处理块。当在帧内模式中编码处理块时,视频编码器(603)可使用帧内预测技术以将处理块编码到已编码图片中;且当在帧间模式或双向预测模式中编码处理块时,视频编码器(603)可分别使用帧间预测或双向预测技术将处理块编码到已编码图片中。在某些视频编码技术中,合并模式可以是帧间图片预测子模式,其中,在不借助预测值外部的已编码运动矢量分量的情况下,从一个或多个运动矢量预测值导出运动矢量。在某些其它视频编码技术中,可存在适用于主题块的运动矢量分量。在实施例中,视频编码器(603)包括其它组件,例如用于确定处理块模式的模式决策模块(未示出)。

[0085] 在图6的实施例中,视频编码器(603)包括如图6所示的耦接到一起的帧间编码器(630)、帧内编码器(622)、残差计算器(623)、开关(626)、残差编码器(624)、通用控制器(621)和熵编码器(625)。

[0086] 帧间编码器(630)用于接收当前块(例如处理块)的样本、比较所述块与参考图片中的一个或多个参考块(例如先前图片和后来图片中的块)、生成帧间预测信息(例如根据帧间编码技术的冗余信息描述、运动矢量、合并模式信息)、以及基于帧间预测信息使用任何合适的技术计算帧间预测结果(例如已预测块)。在一些实施例中,参考图片是基于已编码的视频信息解码的已解码参考图片。

[0087] 帧内编码器(622)用于接收当前块(例如处理块)的样本、在一些情况下比较所述块与同一图片中已编码的块、在变换之后生成量化系数、以及在一些情况下还(例如根据一个或多个帧内编码技术的帧内预测方向信息)生成帧内预测信息。在实施例中,帧内编码器(622)还基于帧内预测信息和同一图片中的参考块计算帧内预测结果(例如已预测块)。

[0088] 通用控制器(621)用于确定通用控制数据,且基于所述通用控制数据控制视频编码器(603)的其它组件。在实施例中,通用控制器(621)确定块的模式,且基于所述模式将控制信号提供到开关(626)。举例来说,当所述模式是帧内模式时,通用控制器(621)控制开关(626)以选择供残差计算器(623)使用的帧内模式结果,且控制熵编码器(625)以选择帧内预测信息且将所述帧内预测信息添加在码流中;以及当所述模式是帧间模式时,通用控制

器(621)控制开关(626)以选择供残差计算器(623)使用的帧间预测结果,且控制熵编码器(625)以选择帧间预测信息且将所述帧间预测信息添加在码流中。

[0089] 残差计算器(623)用于计算所接收的块与选自帧内编码器(622)或帧间编码器(630)的预测结果之间的差(残差数据)。残差编码器(624)用于基于残差数据操作,以对残差数据进行编码以生成变换系数。在实施例中,残差编码器(624)用于将残差数据从空间域转换到频域,且生成变换系数。变换系数接着经由量化处理以获得量化的变换系数。在各种实施例中,视频编码器(603)还包括残差解码器(628)。残差解码器(628)用于执行逆变换,且生成已解码残差数据。已解码残差数据可适当地由帧内编码器(622)和帧间编码器(630)使用。举例来说,帧间编码器(630)可基于已解码残差数据和帧间预测信息生成已解码块,且帧内编码器(622)可基于已解码残差数据和帧内预测信息生成已解码块。适当处理已解码块以生成已解码图片,且在一些实施例中,所述已解码图片可在存储器电路(未示出)中缓冲并用作参考图片。

[0090] 熵编码器(625)用于将码流格式化以产生已编码的块。熵编码器(625)根据HEVC标准等合适标准产生各种信息。在实施例中,熵编码器(625)用于获得通用控制数据、所选预测信息(例如帧内预测信息或帧间预测信息)、残差信息和码流中的其它合适的信息。应注意,根据所公开的主题,当在帧间模式或双向预测模式的合并子模式中对块进行编码时,不存在残差信息。

[0091] 图7是根据本申请公开的另一实施例的视频解码器(710)的图。视频解码器(710)用于接收作为已编码视频序列的一部分的已编码图像,且对所述已编码图像进行解码以生成重建的图片。在实施例中,视频解码器(710)用于代替图3实施例中的视频解码器(310)。

[0092] 在图7实施例中,视频解码器(710)包括如图7中所示耦接到一起的熵解码器(771)、帧间解码器(780)、残差解码器(773)、重建模块(774)和帧内解码器(772)。

[0093] 熵解码器(771)可用于根据已编码图片来重建某些符号,这些符号表示构成所述已编码图片的语法元素。此类符号可包括例如用于对所述块进行编码的模式(例如帧内模式、帧间模式、双向预测模式、后两者的合并子模式或另一子模式)、可分别识别供帧内解码器(772)或帧间解码器(780)用以进行预测的某些样本或元数据的预测信息(例如帧内预测信息或帧间预测信息)、呈例如量化的变换系数形式的残差信息等等。在实施例中,当预测模式是帧间或双向预测模式时,将帧间预测信息提供到帧间解码器(780);以及当预测类型是帧内预测类型时,将帧内预测信息提供到帧内解码器(772)。残差信息可经由逆量化并提供到残差解码器(773)。

[0094] 帧间解码器(780)用于接收帧间预测信息,且基于所述帧间预测信息生成帧间预测结果。

[0095] 帧内解码器(772)用于接收帧内预测信息,且基于所述帧内预测信息生成预测结果。

[0096] 残差解码器(773)用于执行逆量化以提取解量化的变换系数,且处理所述解量化的变换系数,以将残差从频域转换到空间域。残差解码器(773)还可能需要某些控制信息(用以获得量化器参数QP),且所述信息可由熵解码器(771)提供(未标示数据路径,因为这仅仅是低量控制信息)。

[0097] 重建模块(774)用于在空间域中组合由残差解码器(773)输出的残差与预测结果

(可由帧间预测模块或帧内预测模块输出)以形成重建的块,所述重建的块可以是重建的图片的一部分,所述重建的图片继而可以是重建的视频的一部分。应注意,可执行解块操作等其它合适的操作来改善视觉质量。

[0098] 应注意,可使用任何合适的技术来实施视频编码器(303)、视频编码器(503)和视频编码器(603)以及视频解码器(310)、视频解码器(410)和视频解码器(710)。在实施例中,可使用一个或多个集成电路来实施视频编码器(303)、视频编码器(503)和视频编码器(603)以及视频解码器(310)、视频解码器(410)和视频解码器(710)。在另一实施例中,可使用执行软件指令的一个或多个处理器来实施视频编码器(303)、视频编码器(503)和视频编码器(603)以及视频解码器(310)、视频解码器(410)和视频解码器(710)。

[0099] 本公开的多个方面提供了高级视频编解码器中的帧间预测领域的技术。该技术可用于设置候选列表中的候选个数,该候选列表可被称为子块合并候选列表。

[0100] 在各种实施例中,对于帧间预测编码单元(Coding Unit,CU),包括运动矢量、参考图片索引、参考图片列表使用索引、和/或、其它额外信息的运动参数,可用于生成帧间预测样本。帧间预测可包括单向预测、双向预测等。在单向预测中,可使用参考图片列表(例如,第一参考图片列表、或列表0(L0)、或第二参考图片列表、或列表1(L1))。在双向预测中,可以使用L0和L1。参考图片列表使用索引可以指示参考图片列表包括L0、L1,或L0和L1。

[0101] 可以以显式或隐式的方式发信号通知运动参数。当使用跳过模式对CU进行编码时,CU可与一个PU相关联,并且可以不包括明显残差系数(significant residual coefficients)(例如,残差系数为零),不包括已编码运动矢量差(motion vector difference,MVD)或不包括参考图片索引。

[0102] 可使用合并模式,其中,可从相邻CU获得当前CU的运动参数,相邻CU包括空间和时间合并候选以及可选的其它合并候选。合并模式可应用于帧间预测CU,且可用于跳过模式。或者,可以显式地传输、或发信号通知运动参数。例如,对于每个CU,可以显式地发信号通知运动矢量、每个参考图片列表的对应参考图片索引、参考图片列表使用标志以及其它信息。

[0103] 在一些示例(例如,通用视频编码,Versatile Video Coding,VVC)中,使用以下帧间预测编解码工具中的一个或多个:(1)扩展合并预测,(2)具有运动矢量差的合并模式(merge mode with motion vector difference,MMVD),(3)对称MVD(symmetric MVD,SMVD)信令,(4)仿射运动补偿预测,(5)基于子块的时间运动矢量预测(subblock-based temporal motion vector prediction,SbTMVP),(6)自适应运动矢量分辨率(Adaptive Motion Vector Resolution,AMVR),(7)运动场存储:1/16亮度样本MV存储和8×8运动场压缩,(8)具有CU级权重的双向预测(bi-prediction with CU-level weight,BCW),(9)双向光流(Bi-Directional Optical Flow,BDOF),(10)解码器侧运动矢量修正(decoder side motion vector refinement,DMVR),(11)几何分区模式(Geometric Partitioning Mode,GPM),以及(12)组合的帧间和帧内预测(Combined Inter and Intra Prediction,CIIP)。

[0104] 根据本公开的一方面,一些帧间预测编解码工具可根据基于子块的合并候选列表进行操作。在一个示例中,仿射运动补偿预测可在仿射合并模式(在一些示例中也称为基于子块的合并模式)中执行。在仿射合并模式中,可基于仿射合并候选列表执行预测,该仿射合并候选列表在一些示例中称为基于子块的合并候选列表。在另一示例中,基于子块的时间运动矢量预测(SbTMVP)还可根据基于子块的合并候选列表进行操作。

[0105] 对于仿射运动补偿预测,在一些示例(例如,HEVC)中,仅将平移运动模型(translation motion model)应用于运动补偿预测(Motion Compensation Prediction,MCP)。现实世界具有许多种运动,例如,放大/缩小、旋转、透视运动和其它不规则运动。在一些示例(例如,VVC)中,应用基于块的仿射变换运动补偿预测。

[0106] 图8A至图8B示出了仿射运动模型。图8A示出了由两个控制点CP0和CP1(4参数仿射模型)的运动信息描述的块的仿射运动场,并且图8B示出了由三个控制点CP0、CP1和CP2(6参数仿射模型)描述的块的仿射运动场。

[0107] 在一些实施例中,对于4参数仿射运动模型,块中样本位置(x,y)处的运动矢量(mv_x,mv_y)可导出为(等式1),且对于6参数仿射运动模型,块中样本位置(x,y)处的运动矢量可以导出为(等式2):

[0108]
$$\begin{cases} mv_x = \frac{mv_{1x}-mv_{0x}}{W}x + \frac{mv_{1y}-mv_{0y}}{W}y + mv_{0x} \\ mv_y = \frac{mv_{1y}-mv_{0y}}{W}x + \frac{mv_{1y}-mv_{0x}}{W}y + mv_{0y} \end{cases} \quad (等式\ 1)$$

$$\begin{cases} mv_x = \frac{mv_{1x}-mv_{0x}}{W}x + \frac{mv_{2x}-mv_{0x}}{H}y + mv_{0x} \\ mv_y = \frac{mv_{1y}-mv_{0y}}{W}x + \frac{mv_{2y}-mv_{0y}}{H}y + mv_{0y} \end{cases} \quad (等式\ 2)$$

[0109] 其中,mv_0或(mv_{0x},mv_{0y})表示左上角控制点CP0的运动矢量,mv_1或(mv_{1x},mv_{1y})是右上角控制点CP1的运动矢量,mv_2或(mv_{2x},mv_{2y})是左下角控制点CP2的运动矢量,W表示块的宽度,并且H表示块的高度。

[0110] 为了简化运动补偿预测,应用基于块的仿射变换预测。

[0111] 图9示出了每个子块的仿射MV场的示例。在一个示例中,将当前CU 910(例如,具有16×16亮度样本)划分为4×4亮度子块(每个子块可以是4×4亮度样本)。为了导出每个4×4亮度子块的运动矢量,如图9所示,根据以上等式(等式1)和(等式2),计算每个子块的中心样本的运动矢量。例如,可以将运动矢量四舍五入到1/16分数精度。然后,应用运动补偿插值滤波器生成每个子块的预测,其中,每个子块的预测具有导出的运动矢量。在一些示例中,也可以将色度分量的子块大小设置为4×4,因此4×4色度子块包括四个对应的4×4亮度子块。在一个示例中,4×4色度子块的MV计算为四个对应的4×4亮度子块的MV的平均值。

[0112] 注意,可以将子块定义为具有其它合适数量的亮度样本。还要注意,在一些示例中,子块称为子CU。

[0113] 对于平移运动帧间预测,可以使用称为仿射合并(AF_MERGE)模式和仿射高级MVP(仿射AMVP)模式的两个仿射运动帧间预测模式。

[0114] 对于仿射合并预测,在示例中,AF_MERGE模式可以应用于宽度和高度都大于或等于8的CU。在AF_MERGE模式中,基于空间相邻CU的运动信息,生成当前CU的控制点运动矢量(Control Point Motion Vectors,CPMV)。在示例中,仿射合并候选列表(也称为基于子块的合并候选列表)可包括多达五个控制点运动矢量预测器(Control Point Motion Vector Predictor,CPMVP)候选,且发信号通知一个索引,以指示将用于当前CU的一个CPMVP候选。在示例中,使用三种类型的CPMVP候选,从而形成仿射合并候选列表。第一种类型的CPMVP候选是从相邻CU的CPMV推断的继承仿射合并候选。第二种类型的CPMVP候选是使用相邻CU的平移MV导出的构建的仿射合并候选CPMVP。第三种类型的CPMVP候选使用零MV。

[0115] 在一些示例中,例如在VVC中,可使用最多两个继承仿射候选。在示例中,从相邻块的仿射运动模型导出两个继承仿射候选,一个来自左相邻CU(称为左预测器),并且一个来自上相邻CU(称为上预测器(above predictor))。使用图1所示的相邻块作为示例,对于左预测器,扫描顺序是A0->A1,并且对于上预测器,扫描顺序是B0->B1->B2。在一个示例中,仅选择每一侧可用的第一继承候选。在一些示例中,在两个继承候选之间不执行修剪检查(pruning check)。当识别相邻仿射CU时,使用相邻仿射CU的控制点运动矢量来导出当前CU的仿射合并候选列表中的CPMVP候选。

[0116] 图10示出了在仿射合并模式中确定继承控制点运动矢量的示例。如图10所示,当以仿射模式编码相邻左下子块A时,可以获得包含子块A的CU的左上角的运动矢量mv_2、右上角的运动矢量mv_3和左下角的运动矢量mv_4。当使用4参数仿射模型编码子块A时,根据mv_2和mv_3计算当前CU的两个CPMV。在使用6参数仿射模型对子块A进行编码的情况下,根据mv_2、mv_3和mv_4计算当前CU的三个CPMV。

[0117] 在一些示例中,通过组合每一控制点的邻近平移运动信息,构建所构建的仿射候选。控制点的运动信息可以从指定的空间邻居和时间邻居导出。

[0118] 图11示出了根据本公开一些实施例的空间邻居(例如,子块A0-A2和B0-B3)和时间邻居(例如,T)的示例。在示例中,$CPMV_k$(k=1、2、3、4)表示第k个控制点。对于$CPMV_1$,检查B2块->B3块->A2块(->用于表示检查顺序),并且将第一可用块的MV用作$CPMV_1$。对于$CPMV_2$,检查B1块->B0块,并将第一可用块的MV用作$CPMV_2$。对于$CPMV_3$,检查A1块->A0块,并且将第一可用块的MV用作$CPMV_3$。对于TMVP,检查T,如果块T的MV可用,则将T用作$CPMV_4$。

[0119] 在获得四个控制点$CPMV_1$-$CPMV_4$的MV之后,基于该运动信息构建仿射合并候选。按顺序构建控制点MV的以下组合:{$CPMV_1$,$CPMV_2$,$CPMV_3$},{$CPMV_1$,$CPMV_2$,$CPMV_4$},{$CPMV_1$,$CPMV_3$,$CPMV_4$},{$CPMV_2$,$CPMV_3$,$CPMV_4$},{$CPMV_1$,$CPMV_2$},{$CPMV_1$,$CPMV_3$}。

[0120] 3个CPMV的组合可以构建6参数仿射合并候选,并且2个CPMV的组合可以构建4参数仿射合并候选。在一个示例中,为了避免运动缩放过程,当控制点的参考索引不同时,可以丢弃控制点MV的相关组合。

[0121] 在一个示例中,在检查继承仿射合并候选和构建的仿射合并候选之后,如果候选列表仍然未满,则将零MV插入到列表的末尾。

[0122] 对于仿射AMVP预测,仿射AMVP模式可应用于宽度和高度均大于或等于16的CU。在一些示例中,在比特流(例如,已编码视频比特流)中发信号通知CU级的仿射标志,以指示在CU中是否使用仿射AMVP模式,然后发信号通知另一标志,以指示使用4参数仿射还是6参数仿射。在仿射AMVP模式中,可以在比特流中发信号通知当前CU的CPMV与其预测器的CPMVP的差。仿射AMVP候选列表大小为2,且仿射AMVP候选列表是通过按顺序使用以下四种类型的CPMV候选而生成:(1)从相邻CU的CPMV推断的继承仿射AMVP候选;(2)使用相邻CU的平移MV导出的构建的仿射AMVP候选的CPMVP;(3)相邻CU的平移MV;以及(4)零MV。

[0123] 在一些示例中,继承仿射AMVP候选的检查顺序与继承仿射合并候选的检查顺序相同。在一个示例中,仿射合并预测与仿射AMVP预测之间的唯一差异在于,对于AMVP候选,仅考虑与当前块具有相同参考图片的仿射CU。在一个示例中,当将继承仿射运动预测值插入到候选列表中时,不应用修剪过程。

[0124] 在一些示例中,可从图11所示的指定空间邻居导出构建的AMVP候选。在一个示例

中,使用与仿射合并预测的候选结构中相同的检查顺序。此外,还检查相邻块的参考图片索引。使用检查顺序中的第一块,其中,该第一块已进行帧间编码,并具有与当前CU相同的参考图片。当使用4参数仿射模式对当前CU进行编码,并且两个控制点的运动矢量(例如,{$CPMV_1$,$CPMV_2$})均可用时,将两个控制点的运动矢量作为一个候选添加到仿射AMVP列表中。当使用6参数仿射模式对当前CU进行编码,并且控制点CPMV的所有三个运动矢量(例如,{$CPMV_1$,$CPMV_2$,$CPMV_3$})都可用时,将控制点CPMV的所有三个运动矢量作为一个候选添加到仿射AMVP列表中。否则,将构建的AMVP候选设置为不可用。

[0125] 当对继承仿射AMVP候选和构建的AMVP候选进行检查之后,仿射AMVP列表候选的数量仍小于2时,将按顺序添加$CPMV_1$、$CPMV_2$和$CPMV_3$作为平移MV,以在可用时预测当前CU的所有控制点MV。最后,如果仿射AMVP列表仍然未满,则使用零MV来填充仿射AMVP列表。

[0126] 根据本公开的一些方面,可以将运动信息存储在合适的缓冲器中,例如,本地缓冲器、图片行缓冲器等。本地缓冲器用于存储CTU级的运动信息,例如,CTU中4×4块的运动矢量。举例来说,当基于帧间预测对CTU中的CU进行解码时,CU的相应4×4块的运动矢量可存储在本地缓冲器中,且可用于对稍后的CU进行解码。图片行缓冲器用于存储当前CTU上方的CTU的运动信息,例如,上方CTU底部的4×4块的运动矢量。当前CTU的上方CTU可以称为上方CTU行。

[0127] 在一些示例(例如,VVC)中,将仿射CU的CPMV与4×4块的运动矢量分开存储。在一个示例中,本地缓冲器包括第一部分和第二部分,其中,第一部分用于存储CTU中4×4块的运动矢量,第二部分用于存储CTU中仿射CU的CPMV。本地缓冲器的第二部分中存储的CPMV可用于,在仿射合并模式和仿射AMVP模式下,为最近编码的CU生成继承的CPMVP。从CPMV导出的子块MV用于运动补偿、平移MV的合并/AMVP列表的MV推导,以及去块(de-blocking)。

[0128] 在一些实施例中,图片行缓冲器不存储上方CTU行中仿射CU的额外CPMV。在一些示例中,上方CTU的CU的仿射运动数据继承和相同CTU行中正常相邻CU的继承的处理方式不同。当仿射运动数据继承的候选CU在上方CTU行中时,图片行缓冲器中的左下和右下子块MV(而不是CPMV)用于推导仿射MVP。因此,在一些示例中,CPMV只存储在本地缓冲器中,而不存储在图片行缓冲器中。在候选CU是6参数仿射编码的示例中,仿射模型可以降级为4参数模型。

[0129] 图12示出了在一些示例中用于仿射运动数据继承的运动矢量使用的图。在图12中,每个小正方形表示4×4的子块,并且子块的运动矢量可以是子块中心处的运动矢量。此外,当前CU位于当前CTU的顶部位置。如图12所示,在当前CU的相邻CU中,对CU-E和CU-D进行仿射编码。CU-D在与当前CU相同的CTU行中,而CU-E在当前CU的上方CTU行中。CU-D的CPMV可以存储在本地缓冲器中。例如,对于4参数仿射模型,mv_{D0}和mv_{D1}存储在本地缓冲器中,并且,可以根据mv_{D0}和mv_{D1}以及mv_{D0}和mv_{D1}的控制点的对应位置,计算当前CU的CPMV(例如,mv_0和mv_1)。

[0130] 在一个示例中,图片行缓冲器存储位于上方CTU行底部的子块的运动矢量。CU-E的CPMV,如mv_{E0}和mv_{E1}所示,不存储在图片行缓冲器中。在一个示例中,CU-E的左下子块和右下子块的运动矢量(例如,由mv_{LE0}和mv_{LE1}所示)用于当前CU的仿射继承。例如,可以根据mv_{LE0}和mv_{LE1}以及两个子块的对应中心位置,计算当前CU的CPMV(例如,mv_0和mv_1)。

[0131] 在一些实施例中,可实施具有光流的预测修正(Prediction Refinement with

Optical Flow,PROF)(也称为PROF方法),从而改进基于子块的仿射运动补偿,以在不增加运动补偿的存储器访问带宽的情况下,实现更细粒度的运动补偿。在一个实施例(例如,VVC)中,在执行基于子块的仿射运动补偿之后,可将基于光流方程导出的差(或修正值、修正、预测修正)添加到预测样本(例如,亮度预测样本(luma predicted samples)或亮度预测样本(luma prediction samples)),从而获得修正预测样本。

[0132] 图13示出了根据本公开实施例的PROF方法的示例的示意图。可将当前块(1310)划分为四个子块(1312、1314、1316和1318)。每个子块(1312、1314、1316和1318)可以具有4×4像素或样本的大小。可以根据当前块1310的CPMV,例如,可以使用子块(1312)的中心位置和仿射运动模型(例如,4参数仿射运动模型、6参数仿射运动模型),导出子块(1312)的子块MV(1320)。子块MV(1320)可以指向参考图片中的参考子块(1332)。可以根据参考子块(1332)确定初始子块预测样本。

[0133] 在一些示例中,如子块MV(1320)所描述的从参考子块(1332)到子块(1312)的平移运动可能无法以高精度预测子块(1312)。除了由子块MV(1320)描述的平移运动之外,子块(1312)还可经历非平移运动(例如,如图13中所示的旋转)。参考图13,参考图片中具有阴影样本(例如,样本(1332a))的子块(1350)对应于,并且可用于重建子块(1312)中的样本。可以通过像素MV(1340)移动阴影样本(1332a),以便高精度的重建子块(1312)中的样本(1312a)。因此,在一些示例中,当发生非平移运动时,为了提高预测的准确性,可以在仿射运动模型中应用合适的预测修正方法,如下所述。

[0134] 在一个示例中,使用以下四个步骤实现PROF方法。在步骤(1)中,可执行基于子块的仿射运动补偿,以生成当前子块(例如,子块(1312))的预测,例如,初始子块预测I(i,j),其中,i和j是对应于当前子块(1312)中的位置(i,j)处的样本(也称为样本位置(sample position)、样本位置(sample location))的坐标。

[0135] 在步骤(2)中,可以执行梯度计算,其中,可以使用例如3抽头滤波器[-1,0,1],根据如下等式3和等式4,计算每个样本位置(i,j)处的初始子块预测I(i,j)的空间梯度$g_x(i,j)$和$g_y(i,j)$:

[0136] $$g_x(i,j)=I(i+1,j)-I(i-1,j) \quad \text{(等式3)}$$

[0137] $$g_y(i,j)=I(i,j+1)-I(i,j-1) \quad \text{(等式4)}$$

[0138] 子块预测可在每一侧扩展一个像素以用于梯度计算。在一些实施例中,为了降低存储器带宽和复杂度,可以从参考图片(例如,包括子块(1332)的参考图片)中的最近整数像素位置复制扩展边界上的像素。因此,可以避免对填充区域进行额外插值。

[0139] 在步骤(3)中,可以通过如下等式5(例如,光流方程)计算预测修正ΔI(i,j)。

[0140] $$\Delta I(i,j)=g_x(i,j)\times\Delta mv_x(i,j)+g_y(i,j)\times\Delta mv_y(i,j) \quad \text{(等式5)}$$

[0141] 其中,Δmv(i,j)(例如,ΔMV(1342))是样本位置(i,j)处的像素MV或样本MV mv(i,j)(例如,像素MV(1340))与样本位置(i,j)所在的子块(例如,子块(1312))的子块MV Mv_{SB}(例如,子块MV(1320))之间的差MV。Δmv(i,j)也可称为样本位置(i,j)或样本(i,j)处的样本的MV修正(MV Refinement,MVR)。可以使用如下等式6来确定Δmv(i,j)。

[0142] $$\Delta mv(i,j)=mv(i,j)-mv_{SB} \quad \text{(等式6)}$$

[0143] $\Delta mv_x(i,j)$和$\Delta mv_y(i,j)$分别是差MV Δmv(i,j)的x分量(例如,水平分量)和y分量(例如,垂直分量)。

[0144] 由于仿射模型参数和相对于子块中心位置的像素位置从一个子块到另一个子块没有改变,因此,可以为第一子块(例如,子块(1312))计算Δmv(i,j),并将Δmv(i,j)重新用于同一当前块(1310)中的其它子块(例如,子块(1314)、子块(1316)和子块(1318))。在一些示例中,x和y表示样本位置(i,j)相对于子块(1312)的中心位置的水平移位和垂直移位,可以通过如下等式7导出Δmv(i,j)(例如,包括$\Delta mv_x(i,j)$和$\Delta mv_y(i,j)$),

[0145] $$\begin{cases}\Delta mv_x(x,y) = a \times x + b \times y \\ \Delta mv_y(x,y) = c \times x + d \times y\end{cases} \quad (等式7)$$

[0146] 其中,$\Delta mv_x(x,y)$是x分量$\Delta mv_x(i,j)$,并且$\Delta mv_y(x,y)$是y分量$\Delta mv_y(i,j)$。

[0147] 在一个示例中,对于4参数仿射运动模型,参数a-d通过(等式1)描述。对于6参数仿射运动模型,参数a-d通过上述(等式2)描述。

[0148] 在步骤(4)中,可将预测修正ΔI(i,j)(例如,亮度预测修正)添加到初始子块预测I(i,j)以生成另一预测,例如,修正的预测I′(i,j)。可以使用如下等式8,生成样本(i,j)的修正的预测I′(i,j)。

[0149] $I'(i,j) = I(i,j) + \Delta I(i,j)$. (等式8)

[0150] 在一些情况下,不将PROF应用于仿射编码的CU。在一个示例中,所有控制点MV是相同的,这指示CU仅具有平移运动,并且不应用PROF。在另一示例中,仿射运动参数大于指定限制,然后应用PROF。在第二种情况下,将基于子块的仿射运动补偿降级为基于CU的运动补偿,以避免大的存储器访问带宽需求。

[0151] 在一些实施例中,可以应用快速编码方法,从而降低具有PROF的仿射运动估计的编码复杂度。在快速编码方法中,在以下两种情况下,在仿射运动估计阶段不应用PROF。在第一种情况下,如果当前CU不是根块,并且当前CU的父块不选择仿射模式作为最佳模式,则不应用PROF,因为当前CU选择仿射模式作为最佳模式的可能性低。在第二种情况下,如果四个仿射参数(a-d)的幅度都小于预定阈值,并且当前图片不是低延迟图片,则不应用PROF,因为对于这种情况PROF引入的改进很小。这样,可以加速具有PROF的仿射运动估计。

[0152] 在一些示例(例如,VVC)中,可使用基于子块的时间运动矢量预测(SbTMVP)。类似于HEVC中的时间运动矢量预测(Temporal Motion Vector Prediction,TMVP),SbTMVP使用位置相同的图片中的运动场,从而改进当前图片中CU的运动矢量预测和合并模式。在一些示例中,由TMVP使用的同一位置的相同图片用于SbTMVP。SbTMVP在两个方面不同于TMVP。在第一方面中,TMVP预测CU级的运动,但SbTMVP预测子CU级的运动。在第二方面中,TMVP从位置相同的图片中的位置相同块获取时间运动矢量(该位置相同块是相对于当前CU的右下块或中心块),SbTMVP在从位置相同的图片获取时间运动信息之前应用运动移位(motion shift)。运动移位是从当前CU的空间相邻块之一的运动矢量获得的。

[0153] 图14至图15示出了根据本公开一些实施例的SbTMVP过程的示例。SbTMVP通过两个步骤预测当前CU内的子CU的运动矢量。在第一步骤中,检查图14所示的空间邻居A1。如果空间邻居A1具有运动矢量,其中,该运动矢量使用位置相同的图片作为参考图片,则选择该运动矢量作为即将应用的运动移位。如果没有识别出这种运动,则将运动移位设置为(0,0)。

[0154] 在第二步骤中,应用在第一步骤中识别的运动移位(即,添加到当前块的坐标),以从图15所示的位置相同的图片中获得子CU级运动信息(运动矢量和参考索引)。在图15的示例中,将A1的运动矢量设置为运动移位(1510)。然后,对于每个子CU,使用位置相同的图片

中的对应块(覆盖中心样本的最小运动网格)的运动信息来导出子CU的运动信息。在识别位置相同的子CU的运动信息之后,使用与HEVC的TMVP过程类似的方式,将位置相同的子CU的运动信息转换为当前子CU的运动矢量和参考索引。举例来说,应用时间运动缩放,从而将时间运动矢量的参考图片与当前CU的参考图片对齐。

[0155] 在一些示例中,例如在VVC中,基于子块的合并候选列表用于基于子块的合并模式的信令。基于子块的合并候选列表可以包括SbTMVP候选和仿射合并候选,并且在一些示例中称为基于组合子块的合并候选列表(combined subblock based merge candidate list)。SbTMVP模式由标志(例如,序列参数集(Sequence Parameter Set,SPS)标志)启用/禁用。当启用SbTMVP模式时,在一个示例中,添加SbTMVP预测值作为基于组合子块的合并候选列表的第一条目,且随后是仿射合并候选。在一些示例(例如,VVC)中,基于组合子块的合并候选列表的最大允许大小是5。注意,基于组合子块的合并候选列表的最大允许大小可以是其它合适的数量。

[0156] 在一个示例中,SbTMVP中使用的子CU大小固定为8×8,和仿射合并模式一样,SbTMVP模式仅适用于宽度和高度都大于或等于8的CU。

[0157] 在一些实施例中,附加SbTMVP合并候选的编码逻辑与其它合并候选的编码逻辑相同。在一个示例中,对于P或B切片中的每一个CU,执行额外的率失真检查,以决定是否使用SbTMVP候选。

[0158] 根据本公开的一些方面,可以发信号通知基于组合子块的合并候选列表中候选的最大数量。

[0159] 图16示出了一些示例中序列参数集(SPS)的语法表示例(1600)。SPS包括可应用于一系列连续编码视频图片(也称为编码视频序列)的信息。

[0160] 在语法表示例(1600)中,如(1610)所示,发信号通知标志sps_temporal_mvp_enabled_flag。等于1的标志sps_temporal_mvp_enabled_flag表示时间运动矢量预测值可用于编码视频中;并且,等于0的标志sps_temporal_mvp_enabled_flag表示在编码视频中不使用时间运动矢量预测值。在一些示例中,编码视频可称为编码层视频序列(Coded Layer Video Sequence,CLVS),CLVS为属于同一层的图片组,从随机访问点开始,然后是可能相互依赖的图片和随机访问点图片。

[0161] 在相关语法表示例(1600)中,当标志sps_temporal_mvp_enabled_flag等于1时,可以发信号通知两个标志sps_sbtmvp_enabled_flag和sps_affine_enabled_flag,如(1620)和(1630)所示。等于1的标志sps_sbtmvp_enabled_flag表示,基于子块的时间运动矢量预测值可用于解码在编码视频中具有切片的图片,其中,切片的切片类型不等于I(帧内编码)。等于0的标志sps_sbtmvp_enabled_flag表示,在编码视频中不使用基于子块的时间运动矢量预测值。在一个示例中,当未发信号通知标志sps_sbtmvp_enabled_flag时,可以推断标志sps_sbtmvp_enabled_flag等于0。

[0162] 标志sps_affine_enabled_flag表示基于仿射模型的运动补偿是否可用于帧间预测。在一些示例中,当标志sps_affine_enabled_flag等于0时,语法受到限制,因此编码视频中不使用基于仿射模型的运动补偿。否则(sps_affine_enabled_flag等于1),编码视频中可以使用基于仿射模型的运动补偿。

[0163] 在语法表示例(1600)中,当标志sps_affine_enabled_flag等于1时,可发信号通

知参数,例如,five_minus_max_num_subblock_merge_cand。参数five_minus_max_num_subblock_merge_cand表示从5减去SPS中所支持的基于子块的合并候选的最大数量。在一些示例中,five_minus_max_num_subblock_merge_cand的值在0到5的范围(包括端值)内。例如,当five_minus_max_num_subblock_merge_cand的值为2时,基于组合子块的合并候选列表中候选的最大数量为3(5减2)。

[0164] 在一些示例中,可以在图片头级启用/禁用时间运动矢量预测器。图17示出了一些示例中图片头结构的语法表示例(1700)。

[0165] 在语法表示例(1700)中,当SPS级标志sps_temporal_mvp_enabled_flag等于1时,发信号通知标志ph_temporal_mvp_enabled_flag,如(1710)所示。标志ph_temporal_mvp_enabled_flag表示,时间运动矢量预测器是否可用于与图片头相关联的切片的帧间预测。如果ph_temporal_mvp_enabled_flag等于0,则约束与图片头相关联的切片的语法元素,使得在切片的解码中不使用时间运动矢量预测器。否则(ph_temporal_mvp_enabled_flag等于1),时间运动矢量预测器可用于解码与图片头相关联的切片。当不存在时,在一个示例中,推断ph_temporal_mvp_enabled_flag的值等于0。当解码图片缓冲区中没有参考图片具有与当前图片相同的空间分辨率时,ph_temporal_mvp_enabled_flag的值应等于0。

[0166] 基于发信号通知或推断的标志和参数,可导出基于子块的合并候选的最大数量。在一个示例中,变量MaxNumSubblockMergeCand用于表示基于子块的合并候选的最大数量。在一个示例中,当sps_affine_enabled_flag等于1时,根据(等式9)导出MaxNumSubblockMergeCand,并且当sps_affine_enabled_flag等于0时,根据(等式1 0)导出MaxNumSubblockMergeCand:

[0167] MaxNumSubblockMergeCand=5-five_minus_max_num_subblock_merge_cand (等式9)

[0168] MaxNumSubblockMergeCand=sps_sbtmvp_enabled_flag&&ph_temporal_mvp_enabled_flag (等式10)

[0169] 在一些示例中,MaxNumSubblockMergeCand的值在0到5的范围(包括端值)内。

[0170] 根据本公开的一方面,当发信号通知sps_affine_enabled_flag为1时,MaxNumSubblockMergeCand是从five_minus_max_num_subblock_merge_cand导出的,如(等式9)所述。在一些示例中,允许这样的场景:发信号通知sps_affine_enabled_flag为1,并且发信号通知five_minus_max_num_subblock_merge_cand等于5。在此场景中,导出基于子块的合并候选的最大数量MaxNumSubblockMergeCand为0,这可关闭仿射合并模式以及SbTMVP,而不管SbTMVP启用标志,并且当SbTMVP启用标志指示sbTMVP启用时可能引起冲突。

[0171] 本公开的一些方面提供了技术,从而根据基于子块的合并候选的默认数量(例如,由N表示),和仿射和/或SbTMVP编解码工具的相关高级使用标志,设置基于子块的合并候选的数量的值的范围(也称为基于子块的合并候选的最大数量)。例如,当SbTMVP启用标志指示启用SbTMVP时,基于子块的合并候选的最大数量不是0。

[0172] 在一些实施例中,参数five_minus_max_num_subblock_merge_cand与基于子块的合并候选的最大数量呈负相关,并且,根据SbTMVP使能标志确定参数five_minus_max_num_subblock_merge_cand的上限。

[0173] 在一个实施例中,参数five_minus_max_num_subblock_merge_cand指定从N中减

去SPS支持的基于子块的合并运动矢量预测候选的最大数量,并且进一步的,将参数five_minus_max_num_subblock_merge_cand的值限制在0到N-sps_sbtmvp_enabled_flag的范围(包括端值)内。参数five_minus_max_num_subblock_merge_cand的上限取决于标志sps_sbtmvp_enabled_flag。

[0174] 在一些示例中,默认数字N为5,当标志sps_sbtmvp_enabled_flag是0(禁用SbTMVP)时,参数five_minus_max_num_subblock_merge_cand的值可以在0到5的范围内(包括端值)。但是,当标志sps_sbtmvp_enabled_flag为1(启用SbTMVP)时,参数five_minus_max_num_subblock_merge_cand的值可以在0到4的范围内(包括端值)。在一个示例中,在编码器侧,当标志sps_sbtmvp_enabled_flag是1,并且,参数five_minus_max_num_subblock_merge_cand的计算值是5且高于上限时,在编码视频比特流中发信号通知的参数five_minus_max_num_subblock_merge_cand的值是4,4属于0到4的范围(包括端值)内。

[0175] 在一些示例中,当参数five_minus_max_num_subblock_merge_cand的值等于范围的上限时,可以不在来自编码器侧的编码视频比特流中发信号通知参数five_minus_max_num_subblock_merge_cand。在解码器侧,当解码器检测到编码视频比特流中不存在参数five_minus_max_num_subblock_merge_cand时,解码器可以推断出参数five_minus_max_num_subblock_merge_cand的值是范围的上限。可以根据SbTMVP启用标志,确定范围的上限。例如,推断five_minus_max_num_subblock_merge_cand的值等于N-sps_sbtmvp_enabled_flag。在一个示例中,默认数字N为5,当标志sps_sbtmvp_enabled_flag是0(禁用SbTMVP)时,可以推断参数five_minus_max_num_subblock_merge_cand的值为5。但是,当标志sps_sbtmvp_enabled_flag为1(启用SbTMVP)时,可以推断参数five_minus_max_num_subblock_merge_cand的值为4。

[0176] 在另一个实施例中,参数five_minus_max_num_subblock_merge_cand的上限根据多个SbTMVP使能标志的组合确定,例如,SPS级的第一标志sps_sbtmvp_enabled_flag,以及图片头级的第二标志ph_temporal_mvp_enabled_flag。在一个示例中,将参数five_minus_max_num_subblock_merge_cand的值限制在0到N-(sps_sbtmvp_enabled_flag&&ph_temporal_mvp_enabled_flag)的范围(包括端值)内。当编码视频比特流中不存在参数five_minus_max_num_subblock_merge_cand时,推断five_minus_max_num_subblock_merge_cand的值等于N-(sps_sbtmvp_enabled_flag&&ph_temporal_mvp_enabled_flag)。

[0177] 在一些示例中,默认数字N为5,当第一标志sps_sbtmvp_enabled_flag和第二标志ph_temporal_mvp_enabled_flag中的至少一个为0(禁用SbTMVP)时,参数five_minus_max_num_subblock_merge_cand的值可以在0到5的范围(包括端值)内。但是,当第一标志sps_sbtmvp_enabled_flag和第二标志ph_temporal_mvp_enabled_flag均为1(启用SbTMVP)时,参数five_minus_max_num_subblock_merge_cand的值可以在0到4的范围(包括端值)内。在一个示例中,在编码器侧,当第一标志sps_sbtmvp_enabled_flag和第二标志ph_temporal_mvp_enabled_flag均为1时,并且,参数five_minus_max_num_subblock_merge_cand的计算值是5,高于上限,在编码视频比特流中发信号通知的参数five_minus_max_num_subblock_merge_cand的值是4,4被限制在0到4的范围(包括端值)内。

[0178] 在一些示例中,当参数five_minus_max_num_subblock_merge_cand的值等于范围的上限时,可以不在来自编码器侧的编码视频比特流中发信号通知参数five_minus_max_

num_subblock_merge_cand。在解码器侧,当解码器检测到编码视频比特流中不存在参数five_minus_max_num_subblock_merge_cand时,解码器可以推断出参数five_minus_max_num_subblock_merge_cand的值是范围的上限。例如,可以基于第一标志sps_sbtmvp_enabled_flag和第二标志ph_temporal_mvp_enabled_flag的适当组合,确定范围的上限。例如,推断five_minus_max_num_subblock_merge_cand的值等于N-(sps_sbtmvp_enabled_flag&&ph_temporal_mvp_enabled_flag)。在一个示例中,默认数字N为5,当第一标志sps_sbtmvp_enabled_flag和第二标志ph_temporal_mvp_enabled_flag的至少一个为0(禁用SbTMVP)时,可以推断参数five_minus_max_num_subblock_merge_cand的值为5。但是,当第一标志sps_sbtmvp_enabled_flag和第二标志ph_temporal_mvp_enabled_flag均为1(启用SbTMVP)时,可以推断参数five_minus_max_num_subblock_merge_cand的值为4。

[0179] 图18示出了根据本公开一个实施例的概述过程(1800)的流程图。过程(1800)可以用于块的重建,从而为正在重建的块生成一个预测块。在各种实施例中,过程(1800)由处理电路执行,例如,终端设备(210)、终端设备(220)、终端设备(230)和终端设备(240)中的处理电路,执行视频编码器(303)功能的处理电路,执行视频解码器(310)功能的处理电路,执行视频解码器(410)功能的处理电路,执行视频编码器(503)功能的处理电路等。在一些实施例中,过程(1800)在软件指令中实现。因此,当处理电路执行软件指令时,处理电路执行过程(1800)。该过程开始于(S1801),并进行到(S1810)。

[0180] 在(S1810),基于从编码视频比特流解码的预测信息确定参数(例如,five_minus_max_num_subblock_merge_cand指示基于子块的合并候选列表中候选的最大数量)。该参数位于一个范围内,所述范围取决于一个标志,其中,该标志指示基于子块的时间运动矢量预测的启用/禁用状态。在一些示例中,该范围的上限取决于该标志,其中,该标志指示基于子块的时间运动矢量预测的启用/禁用状态。在一个示例中,该标志指示SPS级的基于子块的时间运动矢量预测的启用/禁用状态。

[0181] 在一个实施例中,在编码视频比特流中发信号通知参数的值。在另一个示例中,当未在编码视频比特流中发信号通知参数的值时,可以推断参数的值为范围的上限。例如,当未在编码视频比特流中发信号通知参数时,可以基于默认数量和该标志推断该参数,其中,该标志指示基于子块的时间运动矢量预测的启用/禁用状态。

[0182] 在一些实施例中,该参数位于取决于第一标志和第二标志的范围内,其中,第一标志指示SPS级的基于子块的时间运动矢量预测的启用/禁用状态,第二标志指示图片头(picture header,PH)级的时间运动矢量预测的启用/禁用状态。在一些示例中,当未在编码视频比特流中发信号通知参数时,可以基于默认数量、第一标志和第二标志推断该参数,其中,第一标志指示SPS级的基于子块的时间运动矢量预测的启用/禁用状态,第二标志指示PH级的时间运动矢量预测的启用/禁用状态。

[0183] 在(S1820),根据该参数,计算基于子块的合并候选列表中候选的最大数量。在一些示例中,通过从默认数量中减去该参数,从而计算基于子块的合并候选列表中候选的最大数量,例如使用(等式9)。在一个示例中,默认数量是5。

[0184] 在(S1830),若当前块处于基于子块的预测模式,根据一个候选对当前块的样本进行重建,其中,所述候选从构建的当前块的基于子块的合并候选列表中选择。构建的当前块的基于子块的合并候选列表受到基于子块的合并候选列表中候选的最大数量的限制。

[0185] 然后,过程进行到(S1899)并终止。

[0186] 上述技术可以通过计算机可读指令实现为计算机软件,并且物理地存储在一个或多个计算机可读介质中。例如,图19示出了计算机系统(1900),其适于实现所公开主题的某些实施例。

[0187] 所述计算机软件可通过任何合适的机器代码或计算机语言进行编码,通过汇编、编译、链接等机制创建包括指令的代码,所述指令可由一个或多个计算机中央处理单元(CPU),图形处理单元(GPU)等直接执行或通过译码、微代码等方式执行。

[0188] 所述指令可以在各种类型的计算机或其组件上执行,包括例如个人计算机、平板电脑、服务器、智能手机、游戏设备、物联网设备等。

[0189] 图19所示的用于计算机系统(1900)的组件本质上是示例性的,并不用于对实现本申请实施例的计算机软件的使用范围或功能进行任何限制。也不应将组件的配置解释为与计算机系统(1900)的示例性实施例中所示的任一组件或其组合具有任何依赖性或要求。

[0190] 计算机系统(1900)可以包括某些人机界面输入设备。这种人机界面输入设备可以通过触觉输入(如:键盘输入、滑动、数据手套移动)、音频输入(如:声音、掌声)、视觉输入(如:手势)、嗅觉输入(未示出),对一个或多个人类用户的输入做出响应。所述人机界面设备还可用于捕获某些媒体,气与人类有意识的输入不必直接相关,如音频(例如:语音、音乐、环境声音)、图像(例如:扫描图像、从静止影像相机获得的摄影图像)、视频(例如二维视频、包括立体视频的三维视频)。

[0191] 人机界面输入设备可包括以下中的一个或多个(仅绘出其中一个):键盘(1901)、鼠标(1902)、触控板(1903)、触摸屏(1910)、数据手套(未示出)、操纵杆(1905)、麦克风(1906)、扫描仪(1907)、照相机(1908)。

[0192] 计算机系统(1900)还可以包括某些人机界面输出设备。这种人机界面输出设备可以通过例如触觉输出、声音、光和嗅觉/味觉来刺激一个或多个人类用户的感觉。这样的人机界面输出设备可包括触觉输出设备(例如通过触摸屏(1910)、数据手套(未示出)或操纵杆(1905)的触觉反馈,但也可以有不用作输入设备的触觉反馈设备)、音频输出设备(例如,扬声器(1909)、耳机(未示出))、视觉输出设备(例如,包括阴极射线管屏幕、液晶屏幕、等离子屏幕、有机发光二极管屏的屏幕(1910),其中每一个都具有或没有触摸屏输入功能、每一个都具有或没有触觉反馈功能——其中一些可通过诸如立体画面输出的手段输出二维视觉输出或三维以上的输出;虚拟现实眼镜(未示出)、全息显示器和放烟箱(未示出))以及打印机(未示出)。

[0193] 计算机系统(1900)还可以包括人可访问的存储设备及其相关介质,如包括具有CD/DVD的高密度只读/可重写式光盘(CD/DVD ROM/RW)(1920)或类似介质(1921)的光学介质、拇指驱动器(1922)、可移动硬盘驱动器或固体状态驱动器(1923),诸如磁带和软盘(未示出)的传统磁介质,诸如安全软件保护器(未示出)等的基于ROM/ASIC/PLD的专用设备,等等。

[0194] 本领域技术人员还应当理解,结合所公开的主题使用的术语“计算机可读介质”不包括传输介质、载波或其它瞬时信号。

[0195] 计算机系统(1900)还可以包括通往一个或多个通信网络的接口。例如,网络可以是无线的、有线的、光学的。网络还可为局域网、广域网、城域网、车载网络和工业网络、实时

网络、延迟容忍网络等等。网络还包括以太网、无线局域网、蜂窝网络(GSM、3G、4G、5G、LTE等)等局域网、电视有线或无线广域数字网络(包括有线电视、卫星电视、和地面广播电视)、车载和工业网络(包括CANBus)等等。某些网络通常需要外部网络接口适配器,用于连接到某些通用数据端口或外围总线(1949)(例如,计算机系统(1900)的USB端口);其它系统通常通过连接到如下所述的系统总线集成到计算机系统(1900)的核心(例如,以太网接口集成到PC计算机系统或蜂窝网络接口集成到智能电话计算机系统)。通过使用这些网络中的任何一个,计算机系统(1900)可以与其它实体进行通信。所述通信可以是单向的,仅用于接收(例如,无线电视),单向的仅用于发送(例如CAN总线到某些CAN总线设备),或双向的,例如通过局域或广域数字网络到其它计算机系统。上述的每个网络和网络接口可使用某些协议和协议栈。

[0196] 上述的人机界面设备、人可访问的存储设备以及网络接口可以连接到计算机系统(1900)的核心(1940)。

[0197] 核心(1940)可包括一个或多个中央处理单元(CPU)(1941)、图形处理单元(GPU)(1942)、以现场可编程门阵列(FPGA)(1943)形式的专用可编程处理单元、用于特定任务的硬件加速器(1944)等。这些设备以及只读存储器(ROM)(1945)、随机存取存储器(1946)、内部大容量存储器(例如内部非用户可存取硬盘驱动器、固态硬盘等)(1947)等可通过系统总线(1948)进行连接。在某些计算机系统中,可以以一个或多个物理插头的形式访问系统总线(1948),以便可通过额外的中央处理单元、图形处理单元等进行扩展。外围装置可直接附接到核心的系统总线(1948),或通过外围总线(1949)进行连接。外围总线的体系结构包括外部控制器接口PCI、通用串行总线USB等。

[0198] CPU(1941)、GPU(1942)、FPGA(1943)和加速器(1944)可以执行某些指令,这些指令组合起来可以构成上述计算机代码。该计算机代码可以存储在ROM(1945)或RAM(1946)中。过渡数据也可以存储在RAM(1946)中,而永久数据可以存储在例如内部大容量存储器(1947)中。通过使用高速缓冲存储器可实现对任何存储器设备的快速存储和检索,高速缓冲存储器可与一个或多个CPU(1941)、GPU(1942)、大容量存储器(1947)、ROM(1945)、RAM(1946)等紧密关联。

[0199] 所述计算机可读介质上可具有计算机代码,用于执行各种计算机实现的操作。介质和计算机代码可以是为本申请的目的而特别设计和构造的,也可以是计算机软件领域的技术人员所熟知和可用的介质和代码。

[0200] 作为实施例而非限制,具有体系结构(1900)的计算机系统,特别是核心(1940),可以作为处理器(包括CPU、GPU、FPGA、加速器等)提供执行包含在一个或多个有形的计算机可读介质中的软件的功能。这种计算机可读介质可以是与上述的用户可访问的大容量存储器相关联的介质,以及具有非易失性的核心(1940)的特定存储器,例如核心内部大容量存储器(1947)或ROM(1945)。实现本申请的各种实施例的软件可以存储在这种设备中并且由核心(1940)执行。根据特定需要,计算机可读介质可包括一个或一个以上存储设备或芯片。该软件可以使得核心(1940)特别是其中的处理器(包括CPU、GPU、FPGA等)执行本文所述的特定过程或特定过程的特定部分,包括定义存储在RAM(1946)中的数据结构以及根据软件定义的过程来修改这种数据结构。另外或作为替代,计算机系统可以提供逻辑硬连线或以其它方式包含在电路(例如,加速器(1944))中的功能,该电路可以代替软件或与软件一起运

行以执行本文所述的特定过程或特定过程的特定部分。在适当的情况下,对软件的引用可以包括逻辑,反之亦然。在适当的情况下,对计算机可读介质的引用可包括存储执行软件的电路(如集成电路(IC)),包含执行逻辑的电路,或两者兼备。本申请包括任何合适的硬件和软件组合。

[0201] 附录A:首字母缩略词

[0202] JEM:联合开发模式

[0203] VVC:通用视频编码

[0204] BMS:基准集合

[0205] MV:运动向量

[0206] HEVC:高效视频编码

[0207] SEI:补充增强信息

[0208] VUI:视频可用性信息

[0209] GOP:图片组

[0210] TU:变换单元

[0211] PU:预测单元

[0212] CTU:编码树单元

[0213] CTB:编码树块

[0214] PB:预测块

[0215] HRD:假设参考解码器

[0216] SNR:信噪比

[0217] CPU:中央处理单元

[0218] GPU:图形处理单元

[0219] CRT:阴极射线管

[0220] LCD:液晶显示

[0221] OLED:有机发光二极管

[0222] CD:光盘

[0223] DVD:数字化视频光盘

[0224] ROM:只读存储器

[0225] RAM:随机存取存储器

[0226] ASIC:专用集成电路

[0227] PLD:可编程逻辑设备

[0228] LAN:局域网

[0229] GSM:全球移动通信系统

[0230] LTE:长期演进

[0231] CANBus:控制器局域网络总线

[0232] USB:通用串行总线

[0233] PCI:外围设备互连

[0234] FPGA:现场可编程门阵列

[0235] SSD:固态驱动器

[0236] IC:集成电路

[0237] CU:编码单元

[0238] 虽然本申请已对多个示例性实施例进行了描述,但实施例的各种变更、排列和各种等同替换均属于本申请的范围内。因此应理解,本领域技术人员能够设计多种系统和方法,所述系统和方法虽然未在本文中明确示出或描述,但其体现了本申请的原则,因此属于本申请的精神和范围之内。

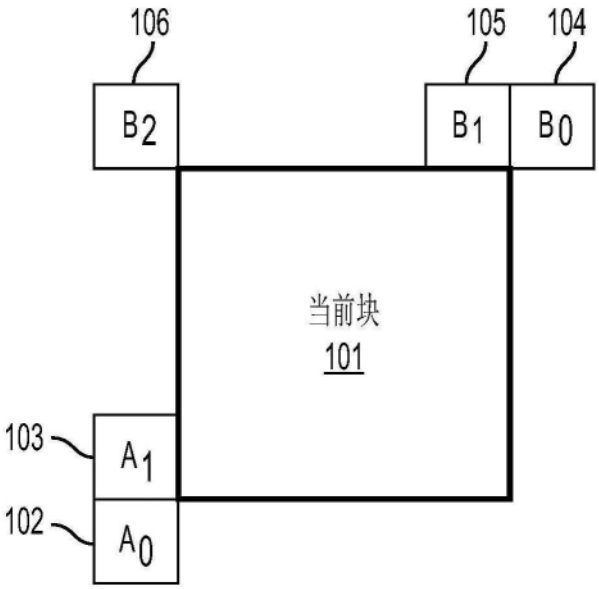

106
105
104
B2
B1
B0
当前块
101
103
A1
102
A0

图1(现有技术)

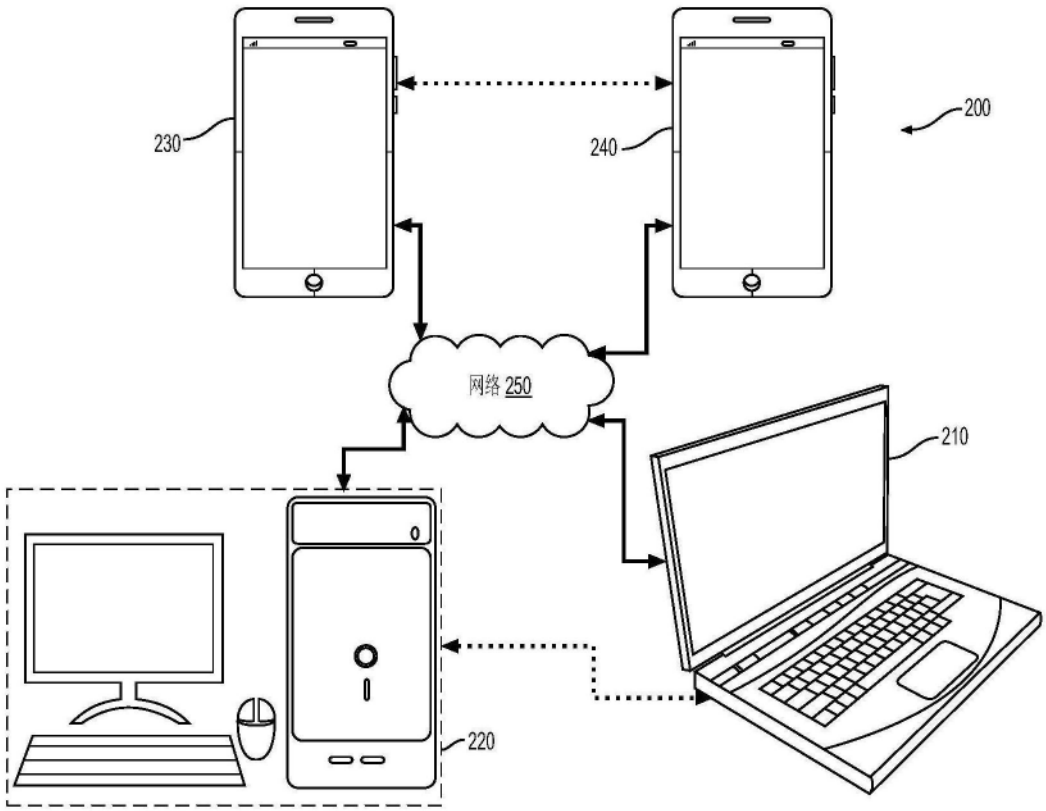

230
240
200
网络 250
210
220

图2

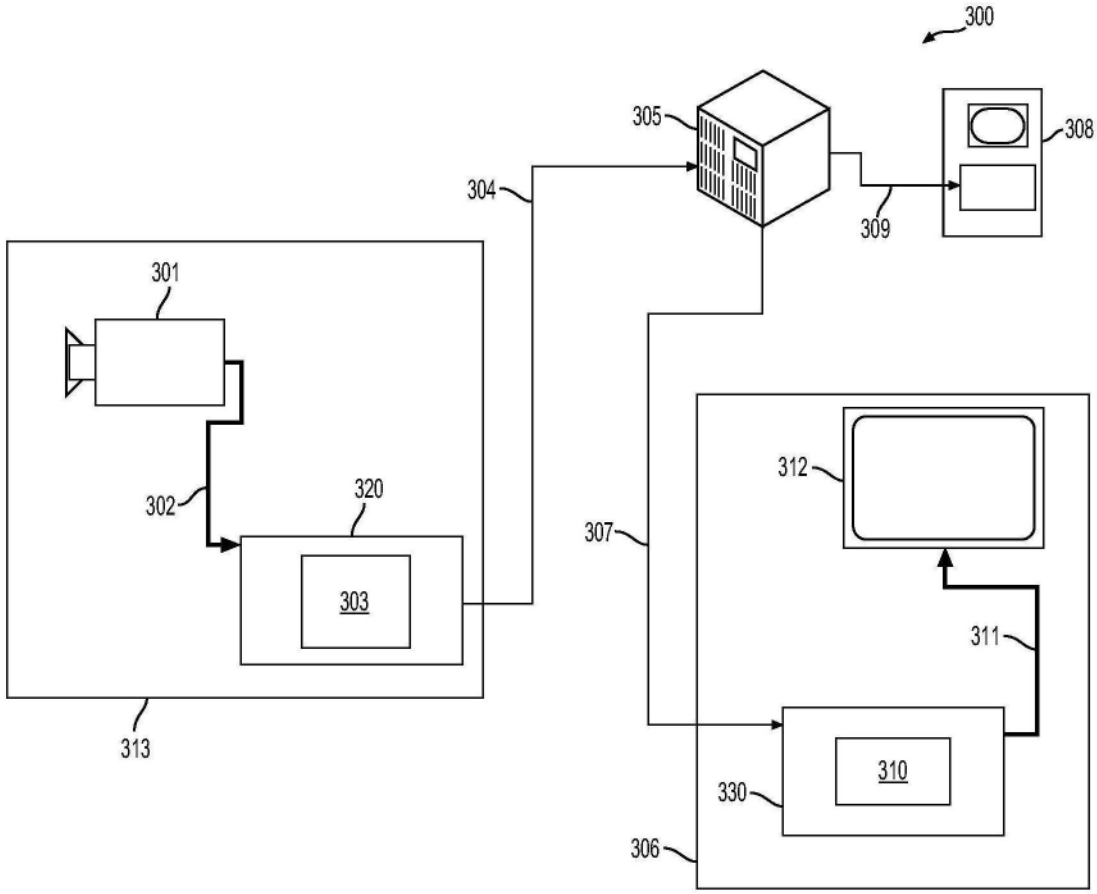


图3

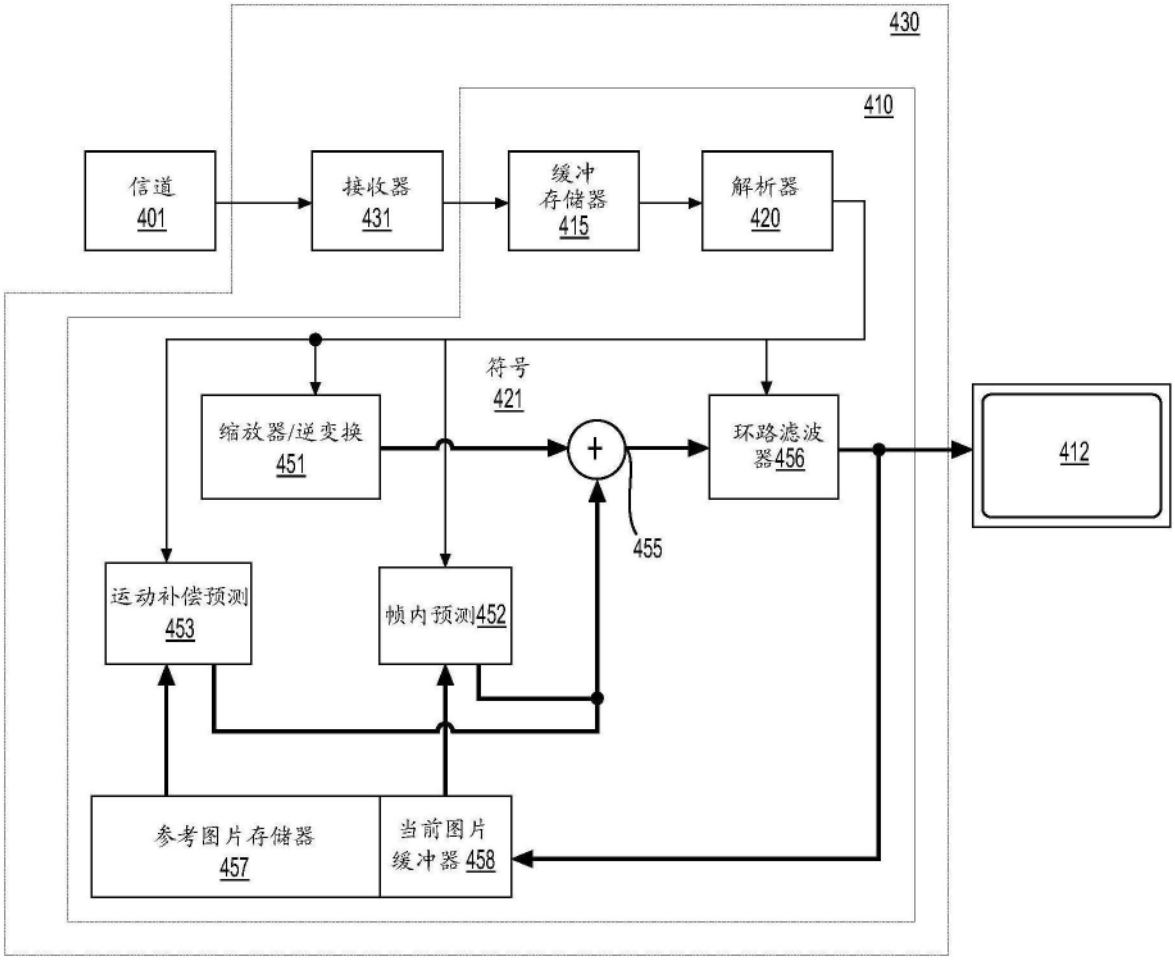

430
410
信道
401
接收器
431
缓冲
存储器
415
解析器
420
符号
421
缩放器/逆变换
451
+
环路滤波
器456
412
455
运动补偿预测
453
帧内预测452
参考图片存储器
457
当前图片
缓冲器 458

图4

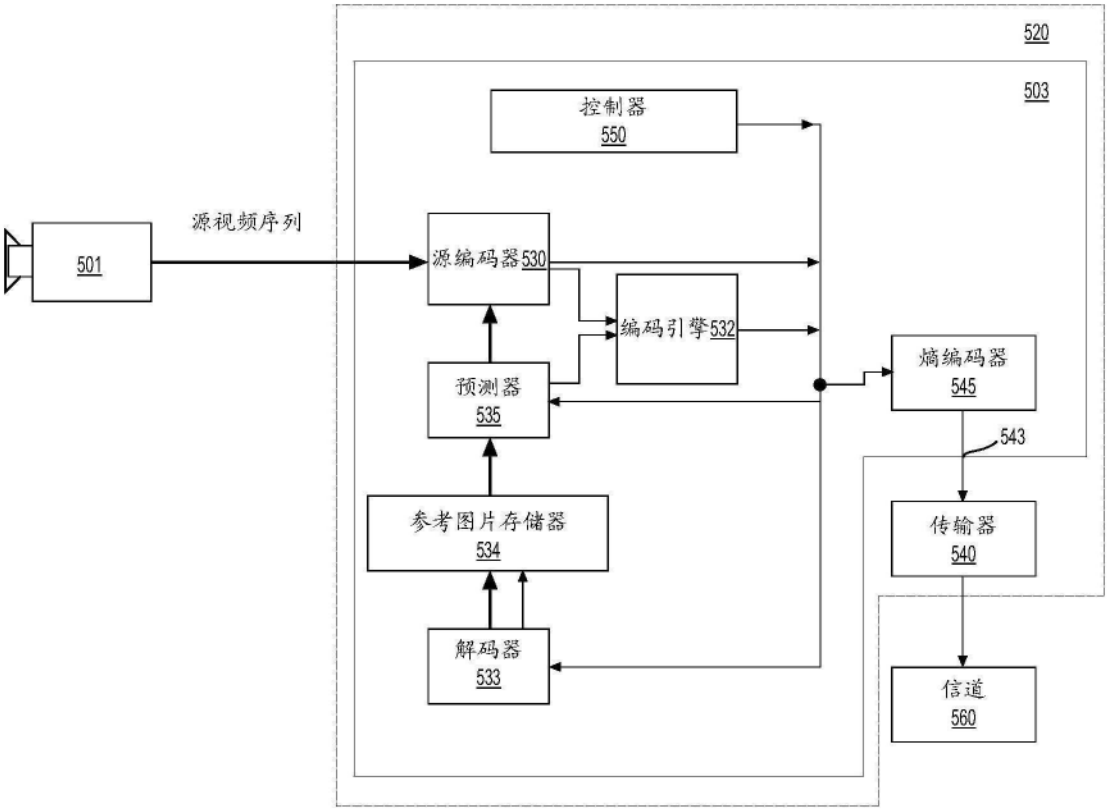

520
503
控制器
550
源视频序列
501
源编码器530
编码引擎532
预测器
535
熵编码器
545
543
参考图片存储器
534
传输器
540
解码器
533
信道
560

图5

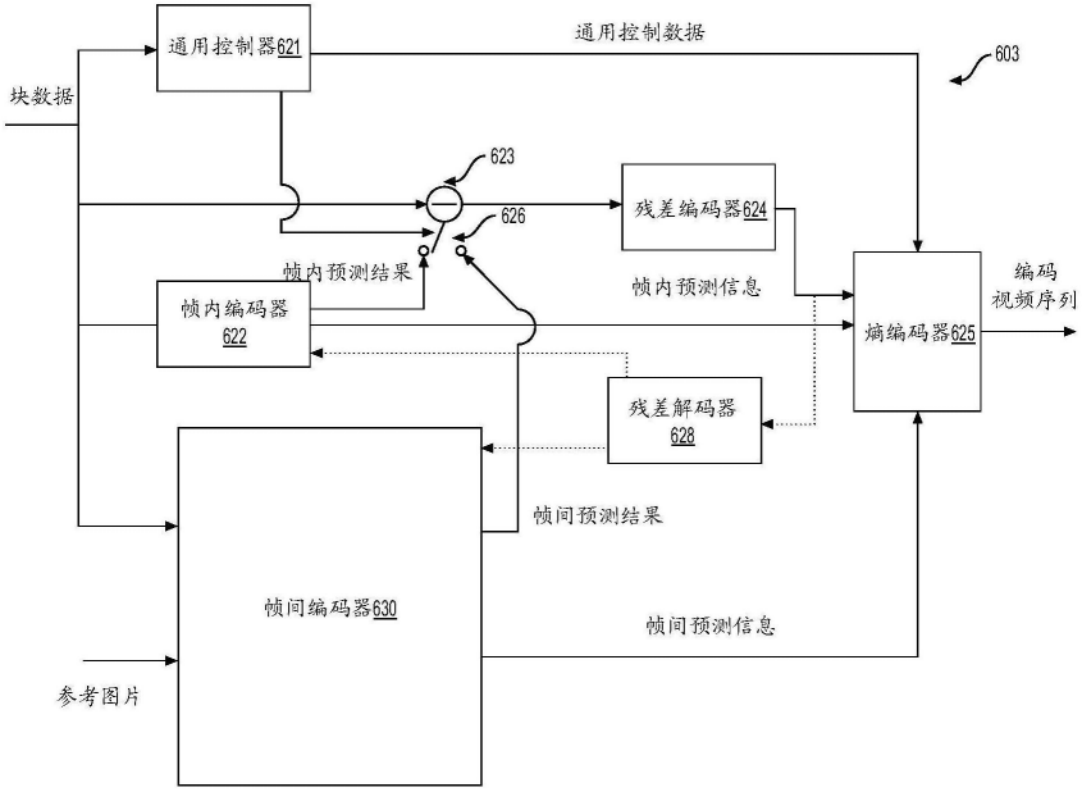

通用控制器621
通用控制数据
603
块数据
623
残差编码器624
626
帧内预测结果
帧内预测信息
编码
视频序列
帧内编码器
622
熵编码器625
残差解码器
628
帧间预测结果
帧间编码器630
帧间预测信息
参考图片

图6

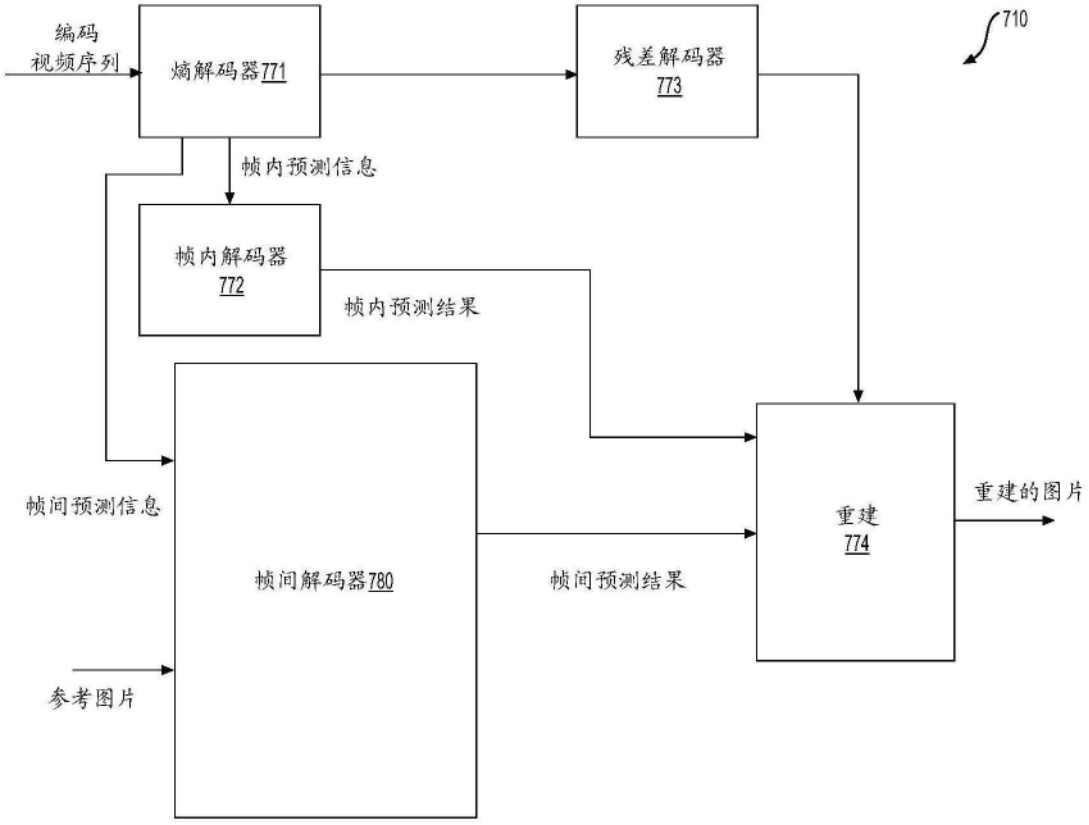

编码
视频序列
熵解码器771
残差解码器
773
710
帧内预测信息
帧内解码器
772
帧内预测结果
帧间预测信息
帧间解码器780
帧间预测结果
重建
774
重建的图片
参考图片

图7

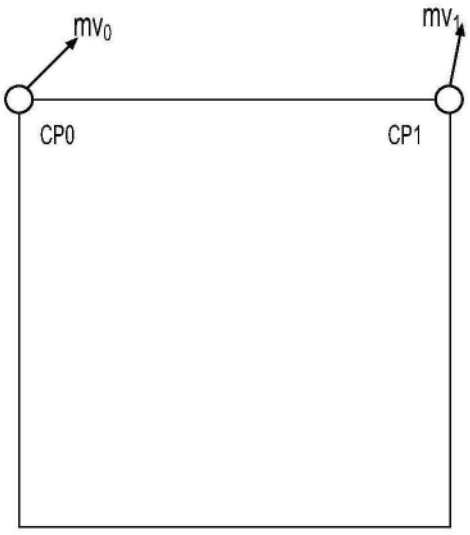

mv0
mv1
CP0
CP1

图8A

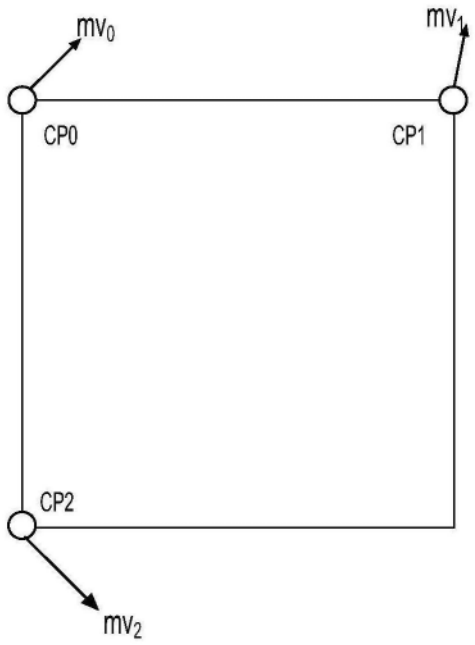


图8B

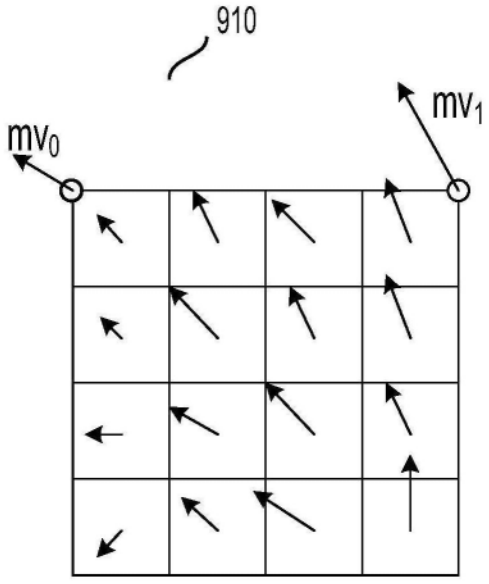


图9

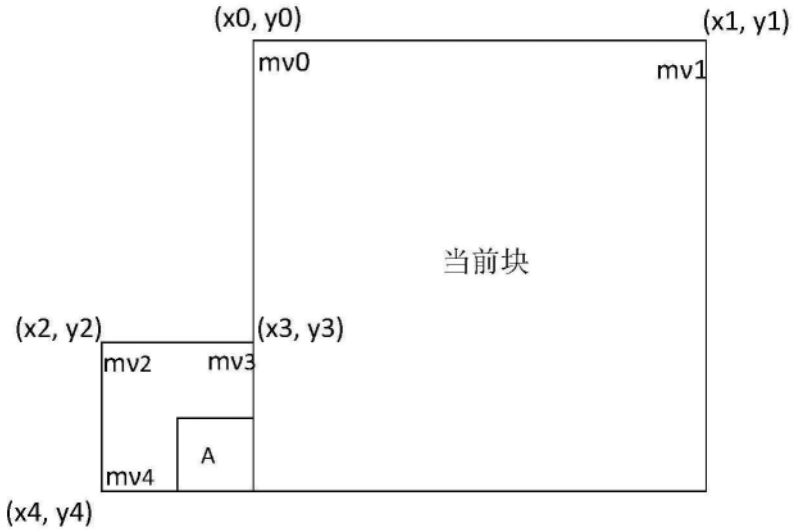

(x0, y0)
(x1, y1)
mv0
mv1
当前块
(x2, y2)
(x3, y3)
mv2
mv3
A
mv4
(x4, y4)

图10

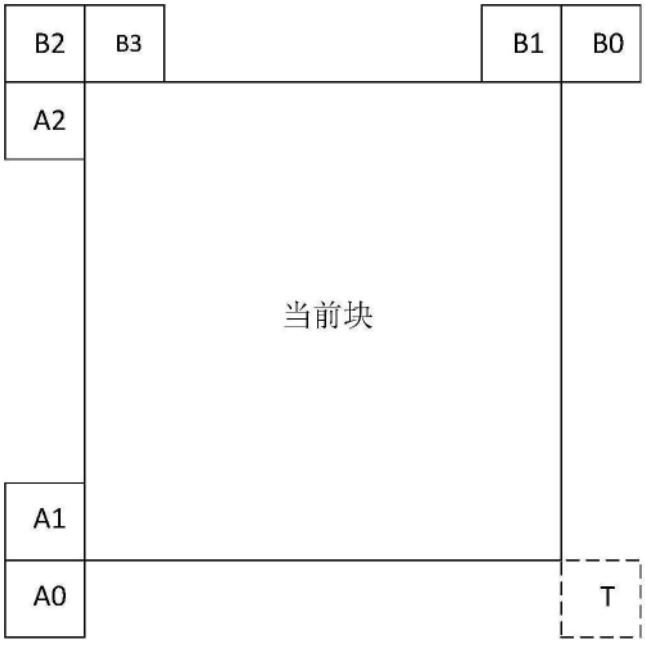

B2
B3
B1
B0
A2
当前块
A1
A0
T

图11

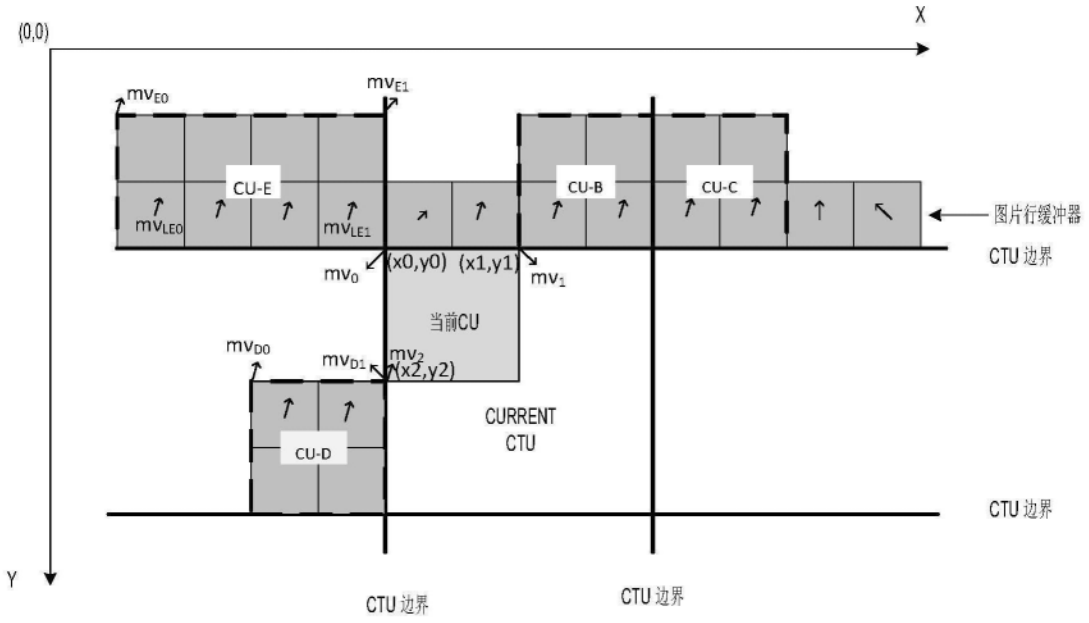

(0,0)
X
Y
mv_{E0}
mv_{E1}
CU-E
CU-B
CU-C
mv_{LE0}
mv_{LE1}
图片行缓冲器
CTU 边界
mv_0
(x0,y0)
(x1,y1)
mv_1
当前CU
mv_{D0}
mv_{D1}
mv_2
(x2,y2)
CU-D
CURRENT
CTU
CTU 边界
CTU 边界
CTU 边界

图12

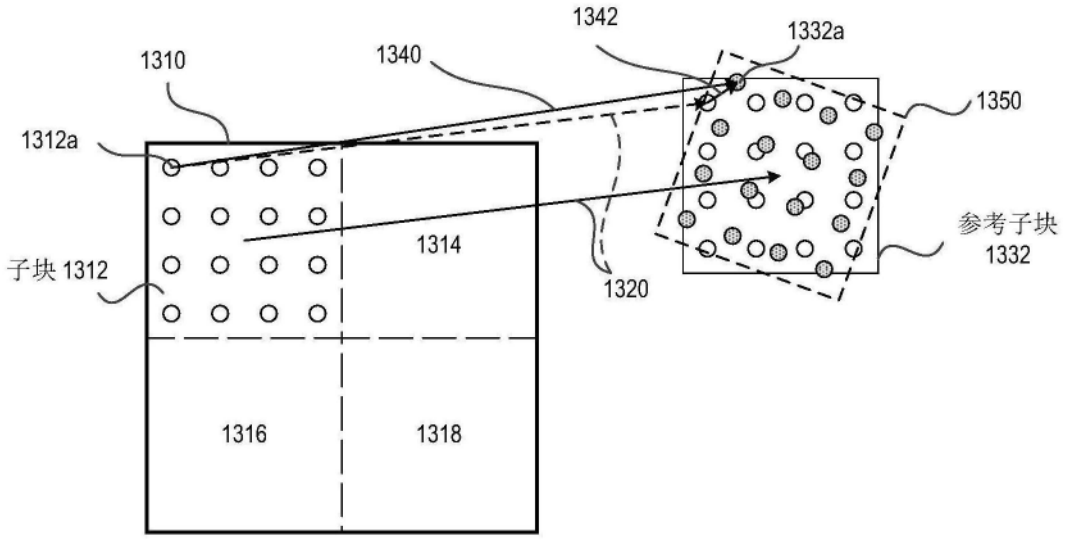

1342
1332a
1310
1340
1350
1312a
1314
参考子块
1332
子块 1312
1320
1316
1318

图13

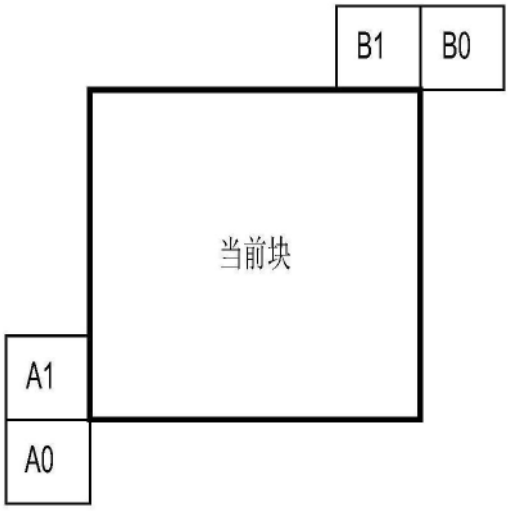


图14

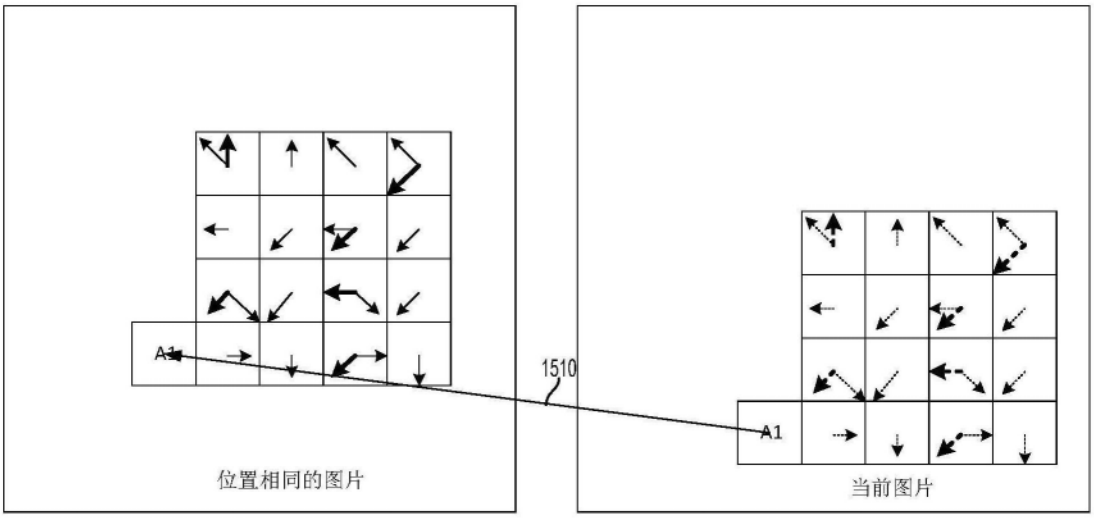


图15

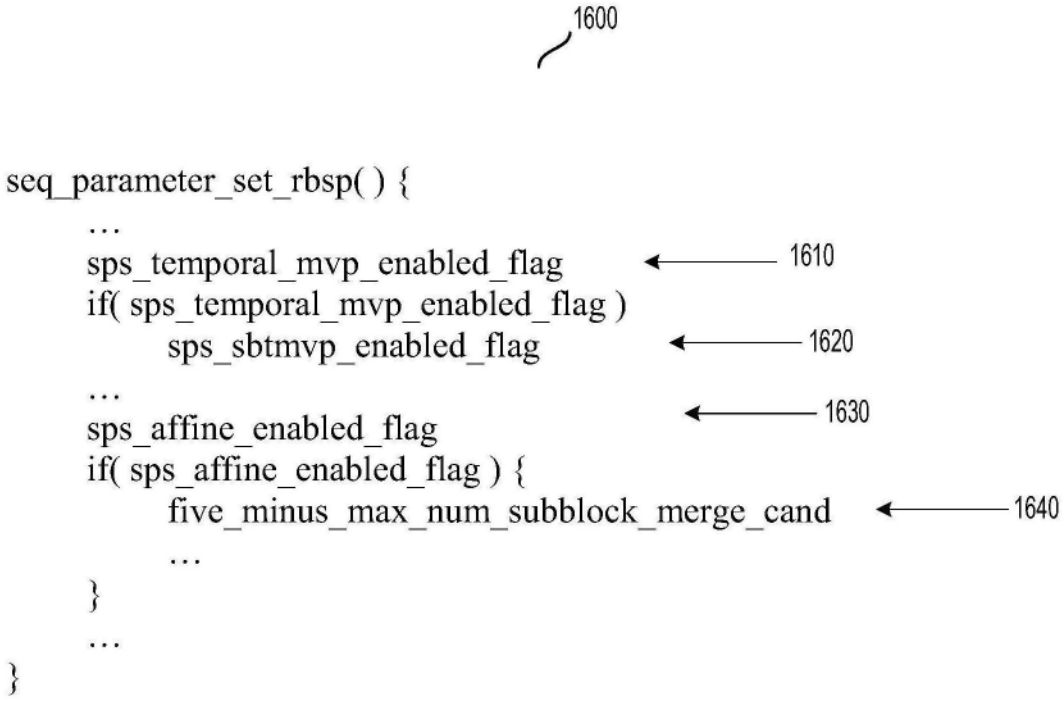

1600
seq_parameter_set_rbsp( ) {
…
sps_temporal_mvp_enabled_flag
1610
if(sps_temporal_mvp_enabled_flag)
sps_sbtmvp_enabled_flag
1620
…
1630
sps_affine_enabled_flag
if(sps_affine_enabled_flag) {
five_minus_max_num_subblock_merge_cand
1640
…
}
…
}

图16

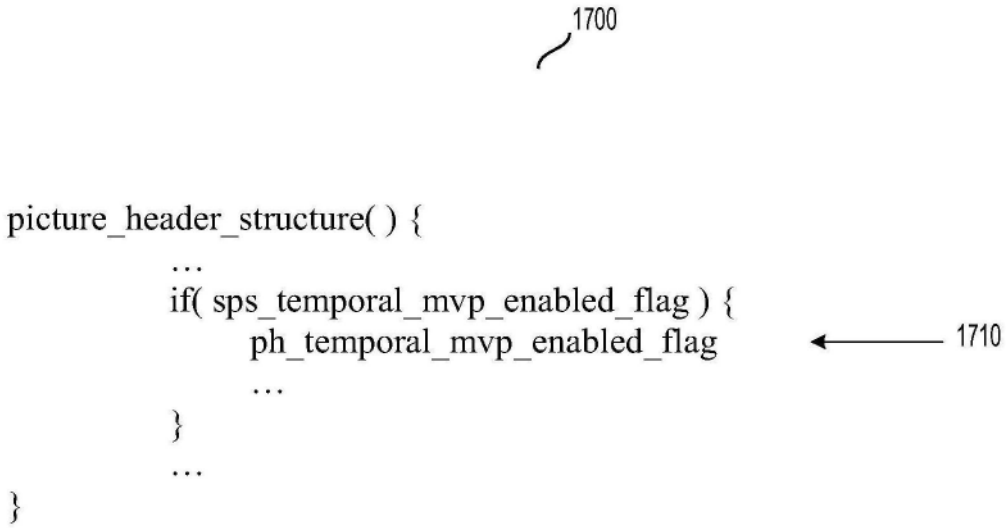

1700
picture_header_structure( ) {
…
if(sps_temporal_mvp_enabled_flag) {
ph_temporal_mvp_enabled_flag
1710
…
}
…
}

图17

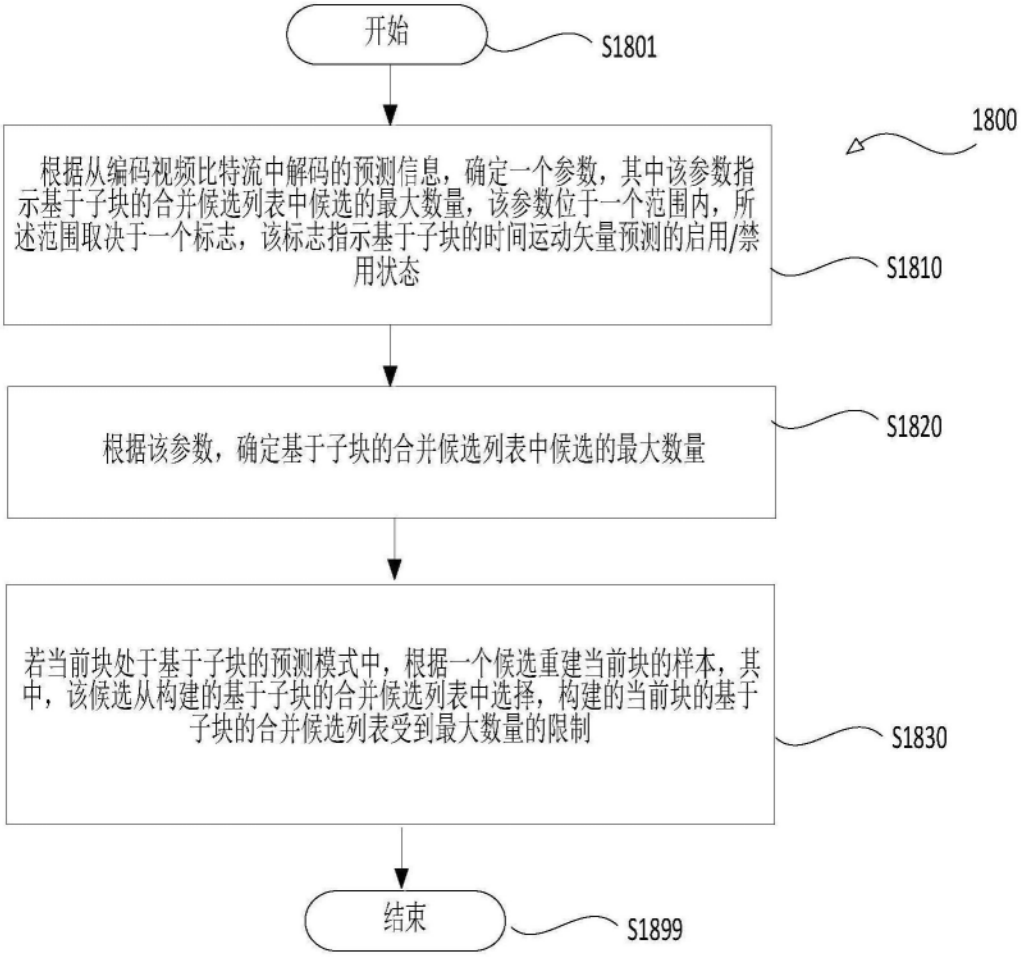

开始
S1801
1800
根据从编码视频比特流中解码的预测信息，确定一个参数，其中该参数指示基于子块的合并候选列表中候选的最大数量，该参数位于一个范围内，所述范围取决于一个标志，该标志指示基于子块的时间运动矢量预测的启用/禁用状态
S1810
根据该参数，确定基于子块的合并候选列表中候选的最大数量
S1820
若当前块处于基于子块的预测模式中，根据一个候选重建当前块的样本，其中，该候选从构建的基于子块的合并候选列表中选择，构建的当前块的基于子块的合并候选列表受到最大数量的限制
S1830
结束
S1899

图18

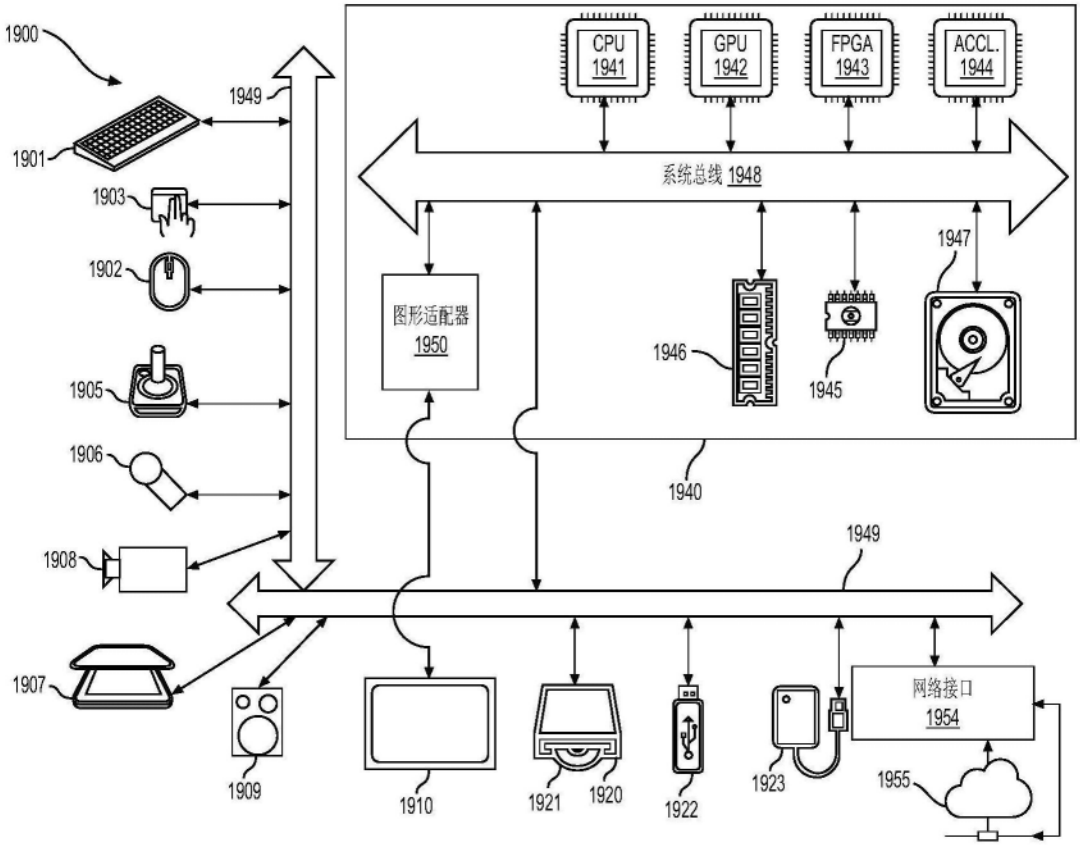

1900
1901
1903
1902
1905
1906
1908
1907
1909
1910
1921
1920
1922
1923
1955
1949
1949
CPU
1941
GPU
1942
FPGA
1943
ACCL.
1944
系统总线 1948
图形适配器
1950
1946
1945
1947
1940
网络接口
1954

图19