



(12) 发明专利申请

(10) 申请公布号 CN 118714345 A

(43) 申请公布日 2024. 09. 27

(21) 申请号 202411049428.5

(22) 申请日 2022.08.16

(30) 优先权数据

63/233,490 2021.08.16 US

17/888,193 2022.08.15 US

(62) 分案原申请数据

202280006651.4 2022.08.16

(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 李翔 陈联霏 刘杉

(74) 专利代理机构 北京康信知识产权代理有限公司 11240

专利代理师 周婷婷

(51) Int.Cl.

H04N 19/513 (2014.01)

H04N 19/109 (2014.01)

H04N 19/139 (2014.01)

H04N 19/13 (2014.01)

H04N 19/46 (2014.01)

H04N 19/159 (2014.01)

H04N 19/577 (2014.01)

H04N 19/105 (2014.01)

H04N 19/172 (2014.01)

H04N 19/157 (2014.01)

H04N 19/573 (2014.01)

H04N 19/70 (2014.01)

H04N 19/176 (2014.01)

H04N 19/174 (2014.01)

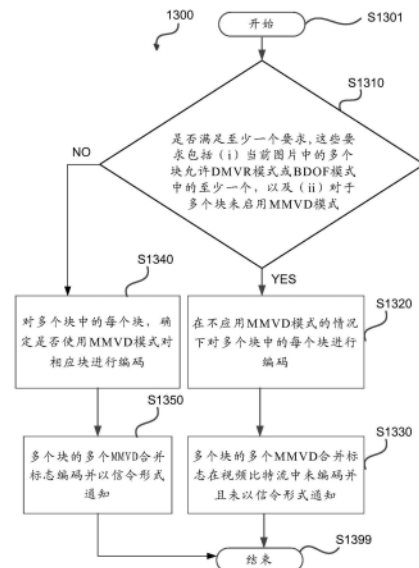
权利要求书2页 说明书32页 附图12页

(54) 发明名称

用于视频编码、解码的方法、装置和存储介质

(57) 摘要

本公开的各方面提供了一种用于视频编码、解码的方法和装置。该装置包括处理电路，该处理电路确定当前图片是否满足第一要求和第二要求中的任意一个，其中，第一要求包括：当前图片中的多个块允许运动矢量细化DMVR模式或双向光流BDOF模式中的至少一个；第二要求包括：多个块未启用运动矢量差分合并MMVD模式；响应于确定当前图片满足第一要求和第二要求中的任意一个，在不应用运动矢量差分合并MMVD模式的情况下对多个块中的每个块进行编码；其中，多个块的多个MMVD的合并标志在视频比特流中未编码并且未以信令形式通知。



1. 一种视频编码的方法,其特征在于,包括:

确定当前图片是否满足第一要求和第二要求中的任意一个,其中,所述第一要求包括:所述当前图片中的多个块允许运动矢量细化DMVR模式或双向光流BDOF模式中的至少一个;所述第二要求包括:所述多个块未启用运动矢量差分合并MMVD模式;

响应于确定所述当前图片满足所述第一要求和所述第二要求中的任意一个,在不应用运动矢量差分合并MMVD模式的情况下对所述多个块中的每个块进行编码;

其中,所述多个块的多个MMVD的合并标志在视频比特流中未编码并且未以信令形式通知。

2. 根据权利要求1所述的方法,其特征在于,在所述第一要求被满足的情况下,响应于所述当前图片中的所述多个块允许所述DMVR模式或所述BDOF模式中的至少一个的确定结果,在不应用MMVD模式的情况下对所述多个块中的每个块进行编码。

3. 根据权利要求2所述的方法,其特征在于,还包括:

对所述多个块对应的预测信息中的至少一个标志进行编码,其中,所述至少一个标志指示所述多个块允许DMVR模式或BDOF模式中的至少一个。

4. 根据权利要求2所述的方法,其特征在于,所述第一要求还包括:所述当前图片根据前向参考图片和后向参考图片确定,且所述当前图片位于所述前向参考图片和所述后向参考图片的中间。

5. 根据权利要求4所述的方法,其特征在于,所述第一要求还包括:所述前向参考图片与所述当前图片之间的第一图片阶次计数(POC)差值的绝对值,等于所述后向参考图片与所述当前图片之间的第二图片阶次计数(POC)差值的绝对值。

6. 根据权利要求5所述的方法,其特征在于,其中,第一图片阶次计数(POC)差值的绝对值和所述第二图片阶次计数(POC)差值的绝对值小于阈值。

7. 根据权利要求1所述的方法,其特征在于,在所述第二要求被满足的情况下,响应于所述当前图片中的所述多个块未启用所述MMVD模式的确定结果,在不应用所述MMVD模式的情况下对所述多个块中的每个块进行编码。

8. 根据权利要求7所述的方法,其特征在于,对所述多个块对应的预测信息中的MMVD合并标志,其中,所述MMVD合并标志指示所述多个块未启用所述MMVD模式。

9. 根据权利要求1所述的方法,其特征在于,所述多个块是所述当前图片中的切片。

10. 根据权利要求1所述的方法,其特征在于,包括:

在所述第一要求和所述第二要求均未满足的情况下,对所述多个块中的每个块进行单独确定,以确定是否应用所述MMVD模式对相应块进行编码;

在视频比特流中对所述多个块中的所述每个块的MMVD合并标志进行编码并以信令通知。

11. 一种视频解码的方法,其特征在于,包括:

从编码视频比特流接收当前图片中的多个块的预测信息;

基于所述预测信息,确定所述多个块是否允许解码器侧运动矢量细化DMVR模式或双向光流BDOF模式中的至少一个;

响应于所述多个块允许所述DMVR模式或所述BDOF模式中的至少一个的确定结果,确定多个具有运动矢量差分合并MMVD的合并标志被推断为对于所述多个块为假,多个MMVD的合

并标志分别指示MMVD模式是否被应用于所述多个块;以及

响应于所述多个MMVD的合并标志被推断为对于所述多个块为假的确定结果,在不应用所述MMVD模式的情况下重构所述多个块中的每个块。

12. 一种用于视频解码的装置,其特征在于,包括:处理电路,所述处理电路被配置为执行如权利要求11所述的方法。

13. 一种用于视频编码的装置,其特征在于,包括:处理电路,所述处理电路被配置为执行如权利要求1-10任一项所述的方法。

14. 一种非暂时性计算机可读存储介质,其特征在于,存储可由至少一个处理器执行的程序,以执行权利要求1-10或11任一项所述的方法。

## 用于视频编码、解码的方法、装置和存储介质

[0001] 援引并入

[0002] 本申请要求2022年8月15日提交的题为“运动矢量差分合并的信令改善”的美国专利申请号17/888,193的权益,该申请要求2021年8月16日提交的题为“运动矢量差分合并MMVD的信令改善”的美国临时申请号63/233,490的优先权。在先申请的公开内容在此全部通过引用结合于此。

### 技术领域

[0003] 本公开描述了总体上涉及视频编码、解码方法及相关装置的实施例。

### 背景技术

[0004] 未压缩的数字图像和/或视频可以包括一系列图片,每个图片具有例如1920×1080亮度样本和相关色度样本的空间维度。该系列图片可以具有固定或可变的图像速率(非正式地也称为帧速率),例如,每秒60幅图片或60Hz。未压缩的图像和/或视频有特定的比特率要求。例如,每样本8比特的1080p60 4:2:0视频(60Hz帧速率的1920x1080亮度样本分辨率)需要接近1.5Gbit/s的带宽。一小时这样的视频需要超过600G字节的存储空间。

[0005] 图像和/或视频编码和解码的一个目的是通过压缩来减少输入图像和/或视频信号中的冗余。压缩有助于降低上述带宽和/或存储空间需求,在某些情况下可降低两个数量级或更多。尽管本文这里描述所使用的视频编码/解码是作为说明性示例,但是在不脱离本公开的精神的情况下,相同的技术可以以类似的方式应用于图像编码/解码。可以采用无损压缩和有损压缩及其组合。无损压缩是指可以从压缩的原始信号中重构出原始信号的精确副本的技术。当使用有损压缩时,重构信号可能与原始信号不相同,但是原始信号和重构信号之间的失真足够小,使得重构信号对预期应用有用。对于视频来说,广泛采用的是有损压缩。被容许的失真量取决于应用;例如,某些消费者流应用的用户可能比电视分发应用的用户容忍更高的失真。可实现的压缩比可以反映出:更高的容许/可容忍失真可以产生更高的压缩比。

[0006] 视频编码器和解码器可以利用几大类技术,包括例如运动补偿、变换处理、量化和熵编码。

[0007] 视频编解码器技术可以包括称为帧内编码的技术。在帧内编码中,样本值是在不参考来自先前重构的参考图片的样本或其他数据的情况下表示的。在某些视频编解码器中,图片在空间上被细分为样本块。当所有样本块都以帧内模式编码时,该图片可以是帧内图片。帧内图片及其派生图片(例如,独立的解码器刷新图片)可以用于重置解码器状态,因此可以用作编码视频比特流和视频会话中的第一图片,或者用作静止图像。可帧内块的样本可以被变换,并且在熵编码之前可以对变换系数进行量化。帧内预测可以是一种在预变换域中最小化样本值的技术。在一些情况下,变换后的DC值越小,AC系数越小,在给定量化步长下表示熵编码后的块所需的比特就越少。

[0008] 例如,从MPEG-2代编码技术中使用的传统帧内编码不使用帧内预测。然而,一些较

新的视频压缩技术包括尝试基于诸如数据块的编码和/或解码期间获得的周围样本数据和/或元数据执行的预测技术。这种技术此后被称为“帧内预测”技术。注意,至少在一些情况下,帧内预测所使用的参考数据来自重构中的当前图片,而不是来自参考图片。

[0009] 可以有多种形式不同的帧内预测。当在给定的视频编码技术中可以使用多于一种这样的技术时,所使用的特定技术可以被编码为使用特定技术的特定帧内预测模式。在某些情况下,帧内预测模式可以有子模式和/或参数,其中,子模式和/或参数可以被单独编码或包含在模式码字中,其定义所使用的预测模式。对于给定的模式、子模式和/或参数组合,使用哪个码字会对通过帧内预测的编码效率增益产生影响,因此,对于将码字转换成比特流的熵编码技术也会产生影响。

[0010] H.264引入了特定的帧内预测模式,在H.265中进行了改进,并在更新的编码技术中进一步改进,例如,共同探索模型(Joint Exploration Model,简称JEM)、通用视频编码(Versatile Video Coding,简称VVC)和基准集(Benchmark Set,简称BMS)。使用已经是可用样本的相邻样本值来形成预测块。根据方向将相邻样本的样本值复制到预测块中。对使用中的方向的参考可以被编码仅比特流中,或者可以被其本身预测出。

[0011] 参考图1A,右下方描绘了从H.265中定义的33个可能的预测方向(对应于35个帧内模式的33个角度模式)中已知的9个预测方向的子集。箭头会聚的点(101)表示被预测的样本。箭头表示的是来自哪个被预测的样本的方向。例如,箭头(102)指示样本(101)是从与水平线成 $45^\circ$ 角的右上方的一个或多个样本预测到的。类似地,箭头(103)指示样本(101)是从与水平线成 $22.5^\circ$ 角的样本(101)左下方的一个或多个样本预测到的。

[0012] 仍然参考图1A,在左上方描绘了 $4 \times 4$ 样本的正方形块(104)(由粗虚线表示)。正方形块(104)包括16个样本,每个样本标有“S”,其在Y维度上的位置(例如,行索引)和其在X维度上的位置(例如,列索引)。例如,样本S21是Y维度中的第二个样本(从顶部算起)和X维度中的第一个样本(从左侧算起)。类似地,样本S44在Y和X维度上都是块(104)中的第四个样本。由于该块的大小为 $4 \times 4$ 个样本,所以S44位于右下角。还示出了遵循类似编号方案的参考样本。用R标记的参考样本、其相对于块的Y位置(例如,行索引)和X位置(列索引)来标记(104)。在H.264和H.265中,预测样本与重构中的块相邻;因此,不需要使用负值。

[0013] 帧内图片预测可以从合适的信号预测方向的相邻样本中复制参考样本值来工作。例如,假设已编码的视频比特流包含针对此块的信号,其指示与箭头(102)一致的预测方向——即,从与水平成 $45^\circ$ 角的右上方的样本来预测样本。在这种情况下,从相同的参考样本R05预测样本S41、S32、S23和S14。然后从参考样本R08预测样本S44。

[0014] 在某些情况下,多个参考样本的值可以组合,例如,通过插值,以便计算参考样本;尤其是当方向不能被 $45^\circ$ 整除时。

[0015] 随着视频编码技术的发展,可能的方向的数量已经增加。在H.264(2003年)中,可以表示九个不同方向。这在H.265(2013年)中增加到33个。目前,JEM/VVC/BMS可以支持多达65个方向。已经通过实验来识别最可能的方向,并且熵编码中的某些技术用于以少量比特来表示那些可能的方向,对于不太可能的方向接受一定的惩罚。此外,方向本身有时可以从相邻的已经解码的块中使用的相邻方向来预测。

[0016] 图1B示出了示意图(110),描绘了根据JEM的65个帧内预测方向,以说明预测方向的数量随着时间的推移而增加。

[0017] 编码视频比特流中帧内预测方向比特的映射所表示的方向可以根据视频编码技术的不同而不同。这种映射的范围可以从诸如简单直接映射到码字,到涉及最有可能模式的复杂自适应方案,以及类似的技术。然而,在大部分情况下,与某些其他方向相比,某些方向在统计上不太可能出现在视频内容中。由于视频压缩的目标是减少冗余,在较好的视频编码技术中,那些不太可能的方向将由更多的比特来表示,相对更可能的方向。

[0018] 可以使用带运动补偿的图片帧间预测进行图像和/或视频编码和解码。运动补偿可以是有损压缩技术,并且可以涉及这样的技术,其中,来自先前重构的图片或其一部分(参考图片)的样本数据块在运动矢量(此后称为MV)指示的方向上进行空间移位之后,可以用于预测新重构的图片或图片部分。在某些情况下,参考图片可以与当前正在重构的图片相同。MV可以具有两个维度X和Y,或者三个维度,第三个维度是使用中的参考图片的指示(参考图片间接可以是时间维度)。

[0019] 在一些视频压缩技术中,可应用于样本数据的某个区域的MV可以从其他MV中预测,例如,从与空间上邻近正在重构的区域的并且在解码顺序上在该MV之前的样本数据的另一个区域相关的MV中预测。这样做可以大大减少编码MV所需的数据量,从而消除冗余并提高压缩率。MV预测可以有效地工作,例如,因为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,存在统计可能性,即比单个MV可应用的区域大的区域在相似的方向上移动,因此,在某些情况下,可以使用从相邻区域的MV导出的相似运动矢量来预测。这使得给定区域发现的MV与从周围MV预测出的MV相似或相同,并且在熵编码之后,这又可以用比直接编码MV时所使用的更少的比特数来表示。在某些情况下,MV预测可以从原始信号(即:样本流)导出的信号(即:MV)的无损压缩的示例。在其他情况下,MV预测本身可能是有损耗的,例如,当从几个周围的MV计算预测值时,会出现舍入误差。

[0020] 在H.265/HEVC(ITU-T Rec.H.265,“High Efficiency Video Coding”,2016年12月)中描述各种MV预测机制。在H.265提供的许多MV预测机制中,参考图2描述了一种此后被称为“空间合并”的技术。

[0021] 参考图2,当前块(201)包括由编码器在运动搜索过程中发现的样本,以便从已经被空间移位的相同大小的先前块中预测。并非直接编码该MV,可以使用与五个周围样本(表示为A0、A1和B0、B1、B2(分别为202至206))中的任一个相关联的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(按照解码顺序)参考图片中导出。在H.265中,MV预测可以使用来自相邻块正在使用的相同参考图片的预测器。

## 发明内容

[0022] 本公开的各方面提供了用于视频编码和解码的方法和装置。

[0023] 在一些示例中,一种用于视频解码的装置包括处理电路。处理电路被配置为从编码视频比特流接收当前图片中的多个块的预测信息。处理电路基于预测信息确定多个块是否允许解码器侧运动矢量细化DMVR模式或双向光流BDOF模式中的至少一个。响应于多个块允许DMVR模式或BDOF模式中的至少一个的确定结果,处理电路确定多个具有运动矢量差分合并MMVD的合并标志被推断为对于多个块为假。多个MMVD的合并标志分别指示MMVD模式是否被应用于多个块。响应于多个MMVD的合并标志被推断为对于多个块为假的确定结果,处理电路在不应用MMVD模式的情况下重构多个块中的每个块。

[0024] 在一个实施例中,预测信息包括至少一个标志,标志指示多个块是否允许DMVR模式或BDOF模式中的至少一个。

[0025] 在一个示例中,处理电路确定 (i) 当前图片参考前向参考图片和后向参考图片,以及 (ii) 当前图片在前向参考图片和后向参考图片的中间。处理电路确定多个MMVD的合并标志并非以信令形式通知,并且基于在当前图片中不允许MMVD模式的确定结果,来推断多个MMVD的合并标志的取值。

[0026] 在一个示例中,处理电路确定前向参考图片和当前图片之间的第一图片顺序计数POC差的绝对值等于后向参考图片和当前图片之间的第二图片顺序计数POC差的绝对值。例如,第一图片顺序计数POC差的绝对值和第二图片顺序计数POC差的绝对值小于阈值。

[0027] 在一个实施例中,预测信息包括多个块的MMVD的合并启用标志,MMVD的合并启用标志指示对于多个块,未启用MMVD模式,处理电路确定多个块的多个MMVD的合并标志被推断为假。

[0028] 在一个示例中,多个块是当前图片中的切片。

[0029] 在一个实施例中,响应于多个MMVD的合并标志未被推断为对于多个块为假的确定结果,多个块中每个块相应的MMVD的合并标志的取值将在编码视频比特流中以信令形式通知。处理电路可以基于多个块是否允许DMVR模式或BDOF模式中的一个或多个,在第一上下文和第二上下文中选择上下文。处理电路可以使用具有所选上下文的上下文自适应二进制算术编码CABAC来解码多个块中每个块相应的MMVD的合并标志。

[0030] 在一个实施例中,预测信息指示是否满足选择第一上下文的条件。条件包括多个块允许DMVR模式或BDOF模式中的一种或多种,多个块的参考图片包括前向参考图片和后向参考图片,并且前向参考图片和当前图片之间的第一图片顺序计数POC差的绝对值等于后向参考图片和当前图片之间的第二图片顺序计数POC差的绝对值。处理电路可以基于所满足的条件,选择第一上下文,作为所选的上下文。处理电路可以基于多个条件中未满足的一个条件,选择第二上下文,作为所选的上下文。

[0031] 在一个实施例中,基于第一概率来初始化所述第一上下文,第一概率指示多个块中每个块相应的MMVD的合并标志为假的概率高于MMVD的合并标志为真的概率的概率。

[0032] 本公开的各方面还提供了用于视频解码器中进行视频解码的方法,其中,包括:从编码视频比特流接收当前图片中的多个块的预测信息;基于预测信息,确定多个块是否允许解码器侧运动矢量细化DMVR模式或双向光流BDOF模式中的至少一个;响应于多个块允许DMVR模式或BDOF模式中的至少一个的确定结果,确定多个具有运动矢量差分合并MMVD的合并标志被推断为对于多个块为假,多个MMVD的合并标志分别指示MMVD模式是否被应用于多个块;以及响应于多个MMVD的合并标志被推断为对于多个块为假的确定结果,在不应用MMVD模式的情况下重构多个块中的每个块。

[0033] 在一个实施例中,预测信息包括至少一个标志,标志指示多个块是否允许DMVR模式或BDOF模式中的至少一个。

[0034] 在一个示例中,该方法还包括:确定 (i) 当前图片参考前向参考图片和后向参考图片,以及 (ii) 当前图片在前向参考图片和后向参考图片的中间;确定多个MMVD的合并标志并非以信令形式通知;以及基于在当前图片中不允许MMVD模式的确定结果,来推断多个MMVD的合并标志。

[0035] 在一个示例中,当前图片在前向参考图片和后向参考图片的中间的结果,将确定出前向参考图片和当前图片之间的第一图片顺序计数POC差的绝对值等于后向参考图片和当前图片之间的第二图片顺序计数POC差的绝对值。

[0036] 在一个示例中,第一图片顺序计数POC差的绝对值和第二图片顺序计数POC差的绝对值小于阈值。

[0037] 在一个实施例中,预测信息包括多个块的MMVD的合并启用标志,MMVD的合并启用标志指示对于多个块,未启用MMVD模式,以及该方法包括确定多个块的多个MMVD的合并标志被推断为假。

[0038] 在一个实施例中,多个块是当前图片中的切片。

[0039] 在一个实施例中,响应于多个MMVD的合并标志未被推断为对于多个块为假的结果,多个块中每个块相应的MMVD的合并标志将在编码视频比特流中以信令形式通知,并且该方法还包括:基于多个块是否允许DMVR模式或BDOF模式中的一个或多个,在第一上下文和第二上下文中选择上下文,以及使用具有所选上下文的上下文自适应二进制算术编码CABAC,来解码多个块中每个块相应的MMVD的合并标志。

[0040] 在一个示例中,预测信息指示是否满足选择第一上下文的条件,条件包括:多个块允许DMVR模式或BDOF模式中的一种或多种,多个块的参考图片包括前向参考图片和后向参考图片,并且前向参考图片和当前图片之间的第一图片顺序计数POC差的绝对值等于后向参考图片和当前图片之间的第二图片顺序计数POC差的绝对值,并且选择上下文包括:基于所满足的条件,选择第一上下文,作为所选的上下文;以及基于多个条件中未满足的一个条件,选择第二上下文,作为所选的上下文。

[0041] 在一个示例中,基于第一概率来初始化第一上下文,第一概率指示多个块中每个块相应的MMVD的合并标志为假的概率高于MMVD的合并标志为真的概率的概率。

[0042] 本公开的各方面还提供了用于视频编码的方法,其中,包括:确定当前图片是否满足至少一个要求,其中,至少一个要求包括:当前图片中的多个块允许运动矢量细化DMVR模式或双向光流BDOF模式中的至少一个;多个块未启用MMVD模式;响应于多个块允许DMVR模式或BDOF模式中的至少一个的结果,在不应用运动矢量差分合并MMVD模式的情况下对多个块中的每个块进行编码;响应于多个块未启用MMVD模式的结果,在不应用运动矢量差分合并MMVD模式的情况下对多个块中的每个块进行编码;在至少一个要求均未满足的情况下,对多个块中的每个块进行单独确定,以确定是否应用MMVD模式对相应块进行编码,并根据单独确定的结果对相应块进行编码。

[0043] 本公开的各方面还提供了用于视频编码的装置,其中,包括处理电路。处理电路被配置为确定当前图片是否满足至少一个要求,其中,至少一个要求包括:当前图片中的多个块允许运动矢量细化DMVR模式或双向光流BDOF模式中的至少一个;多个块未启用MMVD模式;响应于多个块允许DMVR模式或BDOF模式中的至少一个的结果,在不应用运动矢量差分合并MMVD模式的情况下对多个块中的每个块进行编码;响应于多个块未启用MMVD模式的结果,在不应用运动矢量差分合并MMVD模式的情况下对多个块中的每个块进行编码;在至少一个要求均未满足的情况下,对多个块中的每个块进行单独确定,以确定是否应用MMVD模式对相应块进行编码,并根据单独确定的结果对相应块进行编码。

[0044] 本公开的各方面还提供了一种存储程序的非暂时性计算机可读存储介质,该程序

可由至少一个处理器执行,以执行用于视频编码和解码的方法。

### 附图说明

[0045] 从以下详细描述和附图中,所公开主题的进一步特征、性质和各种优点将变得更加明显,其中:

[0046] 图1A是帧内预测模式的示例集的示意图;

[0047] 图1B是示例性帧内预测方向的图示;

[0048] 图2示出了当前块和周围样本的示例;

[0049] 图3是一实施例的通信系统的示例性框图的示意图;

[0050] 图4是另一实施例的通信系统的示例性框图的示意图;

[0051] 图5是解码器的示例性框图的示意图;

[0052] 图6是编码器的示例性框图的示意图;

[0053] 图7示出了示例性编码器的框图;

[0054] 图8示出了示例性解码器的框图;

[0055] 图9示出了双向光流BDOF模式的示例;

[0056] 图10示出了解码器侧运动矢量细化DMVR模式的示例;

[0057] 图11示出了具有运动矢量差的合并模式MMVD中的搜索过程的示例;

[0058] 图12示出了MMVD模式中的搜索点的示例;

[0059] 图13示出了概述根据本公开的实施例的编码过程的流程图;

[0060] 图14示出了概述根据本公开的实施例的解码过程的流程图;

[0061] 图15是根据一个实施例的计算机系统的示意图。

### 具体实施方式

[0062] 图3示出了通信系统(300)的示例性框图。通信系统(300)包括能够经由例如网络(350)相互通信的多个终端设备。例如,通信系统(300)包括经由网络(350)互连的第一对终端设备(310)和(320)。在图3的示例中,第一对终端设备(310)和(320)执行数据的单向传输。例如,终端设备(310)可以对视频数据(例如,由终端设备(310)捕捉的视频图片流)进行编码,以便经由网络(350)传输到另一个终端设备(320)。编码视频数据可以以一个或多个编码视频比特流的形式传输。终端设备(320)可以从网络(350)接收已编码视频数据,解码上述已编码视频数据,以恢复视频图片,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应用等场景中是常见的。

[0063] 在另一示例中,通信系统(300)包括第二对终端设备(330)和(340),其执行对已编码视频数据的双向传输,例如对于在视频会议期间产生的编码视频数据。对于数据的双向传输,在一个示例中,终端设备(330)和(340)的每个终端设备可以对视频数据(例如,由终端设备捕捉的视频图片流)进行编码,以便经由网络(350)传输到终端设备(330)和(340)的另一个终端设备。终端设备(330)和(340)的每个终端设备还可以接收由终端设备(330)和(340)的另一个终端设备传输的编码视频数据,并且可以解码已编码的视频数据,以恢复视频图片,并且可以根据恢复的视频数据在可访问的显示设备上显示视频图片。

[0064] 在图3的示例中,终端设备(310)、终端设备(320)、终端设备(330)和终端设备

(340) 分别被示为服务器、个人计算机和智能电话,但是本公开的原理可以不限于此。本公开的实施例适用于膝上型计算机、平板计算机、媒体播放器和/或专用视频会议设备。网络(350)表示在终端设备(310)、终端设备(320)、终端设备(330)和终端设备(340)之间传送编码视频数据的任意数量的网络,包括例如线路(有线)和/或无线通信网络。通信网络(350)可以在电路交换和/或分组交换信道中交换数据。代表性的网络包括电信网络、局域网、广域网和/或互联网。出于当前讨论的目的,网络(350)的架构和拓扑对于本公开的操作可能是不重要的,除非在下文中解释。

[0065] 作为所公开的主题的应用的示例,图4示出了在流环境中的视频编码器和视频解码器。所公开的主题同样可应用于其他支持视频的应用,包括例如视频会议、数字电视、流服务、在包括CD、DVD、记忆棒等的数字媒体上存储压缩视频等。

[0066] 流式传输系统可以包括:捕捉子系统(413),其可以包括视频源(401),例如,数码相机;创建例如未压缩的视频图片流(402)。在一个示例中,视频图片流(402)包括由数码相机拍摄的样本。当与编码视频数据(404)(或编码视频比特流)相比时,视频图片流(402)被描绘为强调高数据量的粗线,可以由包括耦合到视频源(401)的视频编码器(403)的电子设备(420)来处理。视频编码器(403)可以包括硬件、软件或其组合,以实现或实施如下面更详细描述的主题的各方面。编码视频数据(404)(或编码视频比特流)被描绘为细线,以强调当与视频图片流(402)相比时其为较低的数据量,可以存储在流服务器(405)上,以供将来使用。一个或多个流客户端子系统(例如,图4中的客户端子系统(406)和(408))可以访问流服务器(405),以检索编码视频数据(404)的副本(407)和(409)。客户端子系统(406)可以包括例如电子设备(430)中的视频解码器(410)。视频解码器(410)对编码视频数据的副本(407)进行解码,并创建可以在显示器(412)(例如,显示屏)或其他渲染设备(未示出)上渲染的输出视频图片流(411)。在一些流系统中,编码的视频数据(404)、编码视频数据的副本(407)和(409)(例如,编码视频比特流)可以根据某些视频编码/压缩标准来编码。这些标准的示例包括ITU-T建议H.265。在一个示例中,正在开发的视频编码标准被非正式地称为通用视频编码(VVC)。所公开的主题可以在VVC的环境中使用。

[0067] 需要注意的,电子设备(420)和(430)可以包括其他组件(未示出)。例如,电子设备(420)可以包括视频解码器(未示出),并且电子设备(430)也可以包括视频编码器(未示出)。

[0068] 图5示出了视频解码器(510)的示例性框图。视频解码器(510)可以包括在电子设备(530)中。电子设备(530)可以包括接收机(531)(例如,接收电路)。视频解码器(510)可以用来代替图4示例中的视频解码器(410)。

[0069] 接收机(531)可以接收将由视频解码器(510)解码的一个或多个编码视频序列。在一个实施例中,一次接收一个编码视频序列,其中,每个编码视频序列的解码独立于其他编码视频序列的解码。可以从信道(501)接收编码视频序列,该信道可以是到存储编码视频数据的存储设备的硬件/软件链接。接收机(531)可以接收编码视频数据和其他数据,例如,编码的音频数据和/或辅助数据流,这些数据可以被转发到其相应的使用实体(未示出)。接收机(531)可以将编码视频序列与其他数据分离。为了对抗网络抖动,缓冲存储器(515)可以耦合在接收机(531)和熵解码器/解析器(520)(以下称为“解析器(520)”)之间。在某些应用中,缓冲存储器(515)是视频解码器(510)的一部分。在其他情况下,可以在视频解码器

(510) 之外 (未示出)。在其他情况下,在视频解码器 (510) 之外可以有缓冲存储器 (未示出),例如,用于对抗网络抖动,此外,在视频解码器 (510) 内部可以有另一个缓冲存储器 (515),例如,用于处理播放定时。当接收机 (531) 从具有足够带宽和可控性的存储/转发设备或者从等同步网络中接收数据时,缓冲存储器 (515) 可以是不需要的,或者可以是小的。为了在诸如因特网之类的尽力而为的分组网络上使用,可能需要缓冲存储器 (515),该缓冲存储器可以相对较大,并且可以有利地具有自适应大小,并且可以至少部分地在视频解码器 (510) 外部的操作系统或类似元件 (未示出) 中实现。

[0070] 视频解码器 (510) 可以包括解析器 (520),以从编码视频序列中重构符号 (521)。这些符号的类别包括用于管理视频解码器 (510) 的操作的信息以及潜在地控制诸如渲染设备 (512) (例如,显示屏) 等渲染设备的信息,该渲染设备 (512) 并不是电子设备 (530) 的组成部分,但是可以耦合到电子设备 (530),如图5所示。用于渲染设备的控制信息可以是补充增强信息 (Supplemental Enhancement Information,简称SEI消息) 或视频可用性信息 (Video Usability Information,简称VUI) 参数集片段 (未示出) 的形式。解析器 (520) 可以对接收到的编码视频序列进行解析/熵解码。编码视频序列的编码可以根据视频编码技术或标准,并且可以遵循各种原理,包括可变长度编码、霍夫曼编码、具有或不具有上下文敏感性的算术编码等。解析器 (520) 可以基于对应于该组的至少一个参数,从编码视频序列中提取视频解码器中的至少一个像素子组的一组子组参数。子组可以包括图片组 (Groups of Pictures,简称GOP)、图片、图块、切片、宏块、编码单元 (Coding Units,简称CU)、块、变换单元 (Transform Units,简称TU)、预测单元 (Prediction Units,简称PU) 等。解析器 (520) 还可以从编码视频序列中提取信息,例如,变换系数、量化器参数值、运动矢量等。

[0071] 解析器 (520) 可以对从缓冲存储器 (515) 接收的视频序列执行熵解码/解析操作,以便创建符号 (521)。

[0072] 根据编码视频图片或其部分的类型 (例如:帧间和帧内图片、帧间和帧内块) 以及其他因素,符号 (521) 的重构可以涉及多个不同的单元。可以通过由解析器 (520) 从编码视频序列中解析的子组控制信息来控制涉及哪些单元以及如何涉及。为了清楚起见,没有描述解析器 (520) 和下面的多个单元之间的这种子组控制信息流。

[0073] 除了已经提到的功能块之外,视频解码器 (510) 可以在概念上细分成如下所述的多个功能单元。在商业限制下操作的实际实现中,许多这些单元彼此紧密交互,并且可以至少部分地彼此集成。然而,为了描述所公开的主题,在概念上细分成以下功能单元是合适的。

[0074] 第一单元是定标器/逆变换单元 (551)。定标器/逆变换单元 (551) 接收量化的变换系数以及控制信息,包括使用哪个变换、块大小、量化因子、量化缩放矩阵等,作为来自解析器 (520) 的符号 (521)。定标器/逆变换单元 (551) 可以输出包括样本值的块,这些块可以被输入到聚集器 (555) 中。

[0075] 在一些情况下,定标器/逆变换单元 (551) 的输出样本可以属于帧内编码块。帧内编码块没有使用来自先前重构图片的预测信息,但是可以使用来自当前图片的先前重构部分的预测信息的块。这种预测信息可以由帧内图片预测单元 (552) 提供。在一些情况下,帧内图片预测单元 (552) 使用从当前图片缓冲器 (558) 获取的周围已经重构的信息,来生成与重构中的块具有相同大小和形状的块。当前图片缓冲器 (558) 缓冲例如部分重构的当前图

片和/或完全重构的当前图片。在一些情况下,聚集器(555)基于每个样本将帧内预测单元(552)已经生成的预测信息添加到由定标器/逆变换单元(551)提供的输出样本信息。

[0076] 在其他情况下,定标器/逆变换单元(551)的输出样本可以属于帧间编码的并且可能是运动补偿的块。在这种情况下,运动补偿预测单元(553)可以访问参考图片存储器(557),以获取用于预测的样本。在根据与该块有关的符号(521)对获取的样本进行运动补偿之后,这些样本可以由聚集器(555)添加到定标器/逆变换单元(551)的输出(在这种情况下称为残差样本或残差信号),以便生成输出样本信息。运动补偿预测单元(553)从中获取预测样本的参考图片存储器(557)内的地址可以由运动矢量来控制,对于运动补偿预测单元(553)可行的是,以符号(521)的形式获得这些地址,这些符号可以具有例如X、Y和参考图片组件。当使用子采样精确运动矢量时,运动补偿还可以包括从参考图片存储器(557)获取的样本值的插值、运动矢量预测机制等。

[0077] 聚集器(555)的输出样本可以在环路滤波器单元(556)中经受各种环路滤波技术。视频压缩技术可以包括环路滤波技术,该技术由编码视频序列(也称为编码视频比特流)中包含的参数控制,并且该参数可以作为来自解析器(520)的符号(521)用于环路滤波器单元(556)中,而且也可以对在编码图像或编码视频序列的先前(按照解码顺序)部分的解码期间获得的元信息进行响应,以及对先前的重构的和环路滤波的样本值进行响应。

[0078] 环路滤波器单元(556)的输出可以是样本流,该样本流可以输出到渲染设备(512)以及存储在参考图片存储器(557)中,以用于将来的帧间图片预测。

[0079] 一旦完全重构,某些编码图片可以用作未来预测的参考图片。例如,一旦完全重构与当前图片对应的编码图像,并且该编码图片已经被识别为参考图片(例如,通过解析器(520)),则当前图片缓冲器(558)可以成为参考图片存储器(557)的一部分,并且在开始对接下来的编码图片的重构之前,可以重新分配新的当前图片缓冲器。

[0080] 视频解码器(510)可以根据预定视频压缩技术或诸如ITU-T Rec.H.265等标准来执行解码操作。在编码视频序列符合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载的简介的意义上,编码视频序列可以符合由所使用的视频压缩技术或标准规定的语法。具体地,简介可以从视频压缩技术或标准中所有可用的工具中选择某些工具,作为在该简档下唯一可用的工具。符合标准还需要编码视频序列的复杂度在视频压缩技术或标准的级别所定义的范围之内。在某些情况下,级别限制了最大图片大小、最大帧速率、最大重构采样率(例如,以每秒兆样本为单位测量)、最大参考图片大小等。在某些情况下,由级别设置的限制可以通过假设参考解码器(Hypothetical Reference Decoder,简称HRD)规范和编码视频序列中用于HRD缓冲管理的元数据来进一步限制。

[0081] 在一个实施例中,接收机(531)可以接收具有编码视频的额外(冗余)数据。该额外数据可以作为一部分包括在编码视频序列中。视频解码器(510)可以使用额外数据来正确解码数据和/或更准确地重构原始视频数据。额外数据可以是例如时间、空间或信噪比(SNR)增强层、冗余切片、冗余图片、前向纠错码等形式。

[0082] 图6示出了视频编码器(603)的示例性框图。视频编码器(603)包括在电子设备(620)中。电子设备(620)包括发射机(640)(例如,发射电路)。视频编码器(603)可以用来代替图4示例中的视频编码器(403)。

[0083] 视频编码器(603)可以从视频源(601)(其不是图6示例中的电子设备(620)的一部

分)接收视频样本,该视频源可以捕捉要由视频编码器(603)编码的视频图片。在另一示例中,视频源(601)是电子设备(620)的一部分。

[0084] 视频源(601)可以以数字视频样本流的形式提供要由视频编码器(603)编码的源视频序列,该数字视频样本流可以具有任何合适的比特深度(例如:8比特、10比特、12比特、…)、任何颜色空间(例如,BT.601Y CrCb、RGB、…)和任何合适的采样结构(例如,Y CrCb 4:2:0、Y CrCb4:4:4)。在媒体服务系统中,视频源(601)可以是存储先前准备的视频的存储设备。在视频会议系统中,视频源(601)可以是捕捉本地图像信息作为视频序列的相机。视频数据可以以多个单独的图片的形式提供,当按顺序观看时,这些图片赋予运动。这些图片本身可以被组织为像素的空间阵列,其中,每个像素可以包括一个或多个样本,这取决于使用中的采样结构、颜色空间等。本领域技术人员可以容易地理解像素和样本之间的关系。下面的描述集中在样本上。

[0085] 根据一个实施例,视频编码器(603)可以实时地或者在所需的任何其他时间约束下,将源视频序列的图片编码和压缩成编码视频序列(643)。实施适当的编码速度是控制器(650)的一个功能。在一些实施例中,控制器(650)控制如下所述的其他功能单元,并且在功能上耦合到其他功能单元。为了清楚起见,没有描述耦合。控制器(650)设置的参数可以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 $\lambda$ 值、…)、图片大小、图片组(GOP)布局、最大运动矢量搜索范围等。控制器(650)可以被配置为具有针对特定系统设计而优化的与视频编码器(603)相关的其他合适的功能。

[0086] 在一些实施例中,视频编码器(603)被配置为在编码循环中操作。作为简化的描述,在一个示例中,编码环路可以包括源编码器(630)(例如,用于基于待编码的输入图片和参考图片来创建符号,例如,符号流)以及嵌入在视频编码器(603)中的(本地)解码器(633)。解码器(633)以类似于(远程)解码器也将创建的方式重构符号,以创建样本数据。重构的样本流(样本数据)被输入到参考图片存储器(634)。由于符号流的解码导致位精确结果独立于解码器位置(本地或远程),所以参考图片存储器(634)中的内容在本地编码器和远程编码器之间也是位精确的。换言之,当在解码期间使用预测时,编码器的预测部分作为参考图片样本“看到”与解码器“看到”的样本值完全相同的样本值。也在一些相关技术中使用参考图片同步性(以及如果不能保持同步性而产生的漂移,例如,由于信道误差)的该基本原理。

[0087] “本地”解码器(633)的操作可以与诸如视频解码器(510)之类的“远程”解码器的操作相同,上面已经结合图5对其进行了详细描述。然而,简要地参考图5,由于符号是可用的,并且熵编码器(645)和解析器(520)对编码视频序列的符号编码/解码可以是无损的,所以可以不完全在本地解码器(633)中实现视频解码器(510)的熵解码部分,包括缓冲存储器(515)和解析器(520)。

[0088] 在一个实施例中,除了解码器中存在的解析/熵解码之外,任何解码器技术以相同或基本相同的功能形式存在于相应的编码器中。因此,所公开的主题集中于解码器操作。编码器技术的描述可以简化,因为这些技术是全面描述的了解码器技术的逆。在某些领域,下面提供了更详细的描述。

[0089] 在操作期间,在一些示例中,源编码器(630)可以执行运动补偿预测编码,其参考来自视频序列的被指定为“参考图片”的一个或多个先前已编码的图片来对输入图片进行

预测性地编码。以这种方式,编码引擎(632)对输入图片的像素块和可以被选为对输入图片进行预测参考的参考图片的像素块之间的差异进行编码。

[0090] 本地视频解码器(633)可以基于由源编码器(630)创建的符号,对可以被指定为参考图片的编码视频数据进行解码。编码引擎(632)的操作可以是有利地是有损过程。当已编码的视频数据可以在视频解码器(图6中未示出)处被解码时,重构的视频序列通常可以是具有一些误差的源视频序列的副本。本地视频解码器(633)可以复制由视频解码器对参考图片执行的解码过程,并且可以使得重构的参考图片存储在参考图片存储器(634)中。以这种方式,视频编码器(603)可以本地存储重构的参考图片的副本,这些副本具有与由远端视频解码器获得的重构的参考图片相同的内容(不存在传输误差)。

[0091] 预测器(635)可以对编码引擎(632)执行预测搜索。也就是说,对于待编码的新图片,预测器(635)可以在参考图片存储器(634)中搜索样本数据(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,其可以用作新图片的适当预测参考。预测器(635)可以在逐个样本块-像素块的基础上操作,以找到合适的预测参考。在一些情况下,如由预测器(635)获得的搜索结果所确定的,输入图片可以具有从存储在参考图片存储器(634)中的多个参考图片中提取的预测参考。

[0092] 控制器(650)可以管理源编码器(630)的编码操作,包括例如用于编码视频数据的参数和子组参数的设置。

[0093] 所有前述功能单元的输出可以在熵编码器中经历熵编码(645)。熵编码器(645)通过根据诸如Huffman编码、可变长度编码、算术编码等技术将无损压缩应用于符号,将由各种功能单元生成的符号转换成编码的视频序列。

[0094] 发射机(640)可以缓冲由熵编码器(645)创建的编码视频序列,以准备经由通信信道(660)传输,通信信道可以是到将存储编码视频数据的存储设备的硬件/软件链接。发射机(640)可以将来自视频编码器(603)的编码视频数据与待传输的其他数据合并,例如,编码音频数据和/或辅助数据流(源未示出)。

[0095] 控制器(650)可以管理视频编码器(603)的操作。在编码期间,控制器(650)可以向每个编码图片分配特定的编码图片类型,这可以影响应用于相应图片的编码技术。例如,图片通常可以被分配为以下图片类型之一:

[0096] 帧内图片(I图片)可以是不使用序列中的任何其他图片作为预测源而被编码和解码的图片。一些视频编解码器允许不同类型的帧内图片,包括例如独立解码器刷新(“IDR”)图片。本领域技术人员知道I图片的那些变体以及其相应的应用和特征。

[0097] 预测图片(P图片)可以是使用帧内预测或帧间预测来编码和解码的图片,该预测使用了最多一个运动矢量和参考索引来预测每个块的样本值。

[0098] 双向预测图片(B图片)可以是使用帧内预测或帧间预测来编码和解码的图片,该预测使用了最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,多预测图片可以使用两个以上的参考图片和相关元数据来重构单个块。

[0099] 源图片通常可以在空间上被细分成多个样本块(例如,每个样本块为 $4 \times 4$ 、 $8 \times 8$ 、 $4 \times 8$ 或 $16 \times 16$ 个样本块),并且在分块的基础上编码。参考其他(已经编码的)块来对块进行预测性地编码块,这里其他块是由应用于块的相应图片的编码分配来确定的。例如,I图片的块可以被非预测性地编码,或者可以参考同一图片的已经编码的块被预测性地编码(空

间预测或帧内预测)。参考一个先前编码的参考图片,经由空间预测或经由时间预测,对P图片的像素块进行预测性地编码。参考一个或两个先前编码的参考图片,经由空间预测或经由时间预测,可以对B图片的块进行预测性地编码。

[0100] 视频编码器(603)可以根据预定的视频编码技术或标准(例如,ITU-T Rec.H.265)来执行编码操作。在其操作中,视频编码器(603)可以执行各种压缩操作,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,编码的视频数据可以符合由正在使用的视频编码技术或标准指定的语法。

[0101] 在一个实施例中,发射机(640)可以与编码视频一起传输额外数据。源编码器(630)可以包括作为编码视频序列的一部分的数据。额外数据可以包括时间/空间/SNR增强层、其他形式的冗余数据(例如,冗余图片和切片)、SEI消息、VUI参数集片段等。

[0102] 从视频中捕捉作为时间序列中的多个源图片(视频图片)。帧内图片预测(通常缩写为帧内预测)是利用给定图片中的空间相关性,而帧间图片是预测利用图片之间的(时间或其他)相关性。在一个示例中,一个在编码/解码中的特定图片,被称为当前图片,被分割成块。在当前图片中的块类似于视频中先前编码且仍被缓冲的参考图片中的参考块时,当前图片中的块可以通过称为运动矢量的矢量来编码。在使用多个参考图片的情况下,运动矢量指向参考图片中的参考块,并且可以具有识别参考图片的第三维。

[0103] 在一些实施例中,可以在帧间图片预测中使用双向预测技术。根据双向预测技术使用两个参考图片,例如,第一参考图片和第二参考图片,这两个参考图片在解码顺序上都在视频中的当前图片之前(但是在显示顺序上可以分别在过去和未来)。当前图片中的块可由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来编码。可以通过第一参考块和第二参考块的组合来预测该块。

[0104] 此外,可以在帧间图片预测中使用合并模式技术来提高编码效率。

[0105] 根据本公开的一些实施例,例如,帧间图片预测和帧内图片预测可以以块为单位执行预测。例如,根据HEVC标准,视频图片序列中的图片被分割成编码树单元(Coding Tree Units,简称CTU),用于压缩,图片中的CTU具有相同的大小,例如,64×64像素、32×32像素或16×16像素。通常,CTU包括三个编码树块(Coding Tree Blocks,简称CTB),即一个亮度CTB和两个色度CTB。每个CTU可以被递归地二叉树分割成一个或多个编码单元(CU)。例如,64×64像素的CTU可以被分割成一个64×64像素的CU,或者4个32×32像素的CU,或者16个16×16像素的CU。在一个示例中,分析每个CU,以确定CU的预测类型,例如,帧间预测类型或帧内预测类型。根据时间和/或空间可预测性,CU被分成一个或多个预测单元(PU)。通常,每个PU包括一个亮度预测块(PB)和两个色度PB。在实施例中,以预测块为单位执行编码(编码/解码)中的预测操作。使用亮度预测块作为预测块的示例,预测块包含像素的值(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。

[0106] 图7示出了视频编码器(703)的示例性示意图。视频编码器(703)被配置为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),并将该处理块编码成作为编码视频序列的一部分的编码图像。在一个示例中,视频编码器(703)用于代替图4示例中的视频编码器(403)。

[0107] 在HEVC示例中,视频编码器(703)接收处理块的样本值矩阵,例如,8×8样本的预测块等。视频编码器(703)确定使用帧内模式、帧间模式还是使用例如率失真优化的双向预

测模式,来对处理块进行最佳编码。当要以帧内模式对处理块进行编码时,视频编码器(703)可以使用帧内预测技术将处理块编码成编码图片;并且当要以帧间模式或双向预测模式对处理块进行编码时,视频编码器(703)可以分别使用帧间预测或双向预测技术来将处理块编码成编码图片。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,从一个或多个运动矢量预测器中导出运动矢量,而不受益于预测器之外的编码运动矢量组件。在某些其他视频编码技术中,可能存在适用于对象块的运动矢量组件。在一个示例中,视频编码器(703)包括其他组件,例如,用于确定处理块的模式模式判定模块(未示出)。

[0108] 在图7的示例中,视频编码器(703)包括如图7所示耦合在一起的帧间编码器(730)、帧内编码器(722)、残差计算器(723)、交换机(726)、残差编码器(724)、通用控制器(721)和熵编码器(725)。

[0109] 帧间编码器(730)被配置为接收当前块(例如,处理块)的样本,将该块与参考图片中的一个或多个参考块(例如,先前图片和后续图片中的块)进行比较,生成帧间预测信息(例如,根据帧间编码技术的冗余信息的描述、运动矢量、合并模式信息),并且基于使用了任何合适的技术得到的帧间预测信息来计算帧间预测结果(例如,预测块)。在一些示例中,参考图片是基于编码视频信息解码的解码参考图片。

[0110] 帧内编码器(722)被配置为接收当前块(例如,处理块)的样本,在一些情况下,将该块与已经在同一图片中编码的块进行比较,在变换后生成量化系数,并且在一些情况下,还生成帧内预测信息(例如,根据一种或多种帧内编码技术的帧内预测方向信息)。在一个示例中,帧内编码器(722)还基于同一图片中的帧内预测信息和参考块来计算帧内预测结果(例如,预测块)。

[0111] 通用控制器(721)被配置为确定通用控制数据并基于通用控制数据控制视频编码器(703)的其他组件。在一个示例中,通用控制器(721)确定块的模式,并基于该模式向交换机(726)提供控制信号。例如,当模式是帧内模式时,通用控制器(721)控制交换机(726)选择帧内模式结果供残差计算器(723)使用,并控制熵编码器(725)选择帧内预测信息并将帧内预测信息包括在比特流中;当模式是帧间模式时,通用控制器(721)控制交换机(726)选择帧间预测结果供残差计算器(723)使用,并控制熵编码器(725)选择帧间预测信息并将帧间预测信息包括在比特流中。

[0112] 残差计算器(723)被配置为计算接收的块和从帧内编码器(722)或帧间编码器(730)选择的预测结果之间的差(残差数据)。残差编码器(724)被配置为基于残差数据进行操作,以编码残差数据,来生成变换系数。在一个示例中,残差编码器(724)被配置为将残差数据从空间域转换到频域,并生成变换系数。然后对变换系数进行量化处理,以获得量化的变换系数。在各种实施例中,视频编码器(703)还包括残差解码器(728)。残差解码器(728)被配置为执行逆变换,并生成解码后的残差数据。解码的残差数据可以被帧内编码器(722)和帧间编码器(730)适当地使用。例如,帧间编码器(730)可以基于解码的残差数据和帧间预测信息生成解码的块,帧内编码器(722)可以基于解码的残差数据和帧内预测信息生成解码的块。适当地处理被解码的块,以生成已解码的图片,并且已解码的图片可以缓存在存储器电路(未示出)中,并且在一些示例中用作参考图片。

[0113] 熵编码器(725)被配置为格式化比特流,以包括编码块。熵编码器(725)被配置为

根据合适的标准,例如,HEVC标准,在比特流中包括各种信息。在一个示例中,熵编码器(725)被配置为包括通用控制数据、所选预测信息(例如,帧内预测信息或帧间预测信息)、残差信息和比特流中的其他合适的信息。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,没有残差信息。

[0114] 图8示出了视频解码器(810)的示例性示意图。视频解码器(810)被配置为接收作为编码视频序列的一部分的编码图片,并对编码图片进行解码,以生成重构图片。在一个示例中,视频解码器(810)用于代替图4示例中的视频解码器(410)。

[0115] 在图8的示例中,视频解码器(810)包括如图8所示耦合在一起的熵解码器(871)、帧间解码器(880)、残差解码器(873)、重构模块(874)和帧内解码器(872)。

[0116] 熵解码器(871)可以被配置为从编码图像中重构某些符号,这些符号表示组成编码图像的语法元素。这种符号可以包括例如对块进行编码的模式(例如,帧内模式、帧间模式、双向预测模式、合并子模式或另一子模式中的后两种模式)以及可以识别出分别由帧内解码器(872)或帧间解码器(880)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)。符号还可以包括例如量化变换系数形式的残差信息等。在一个示例中,当预测模式是帧间或双向预测模式时,帧间预测信息被提供给帧间解码器(880);并且当预测类型是帧内预测类型时,帧内预测信息被提供给帧内解码器(872)。残差信息可以经历逆量化,并被提供给残差解码器(873)。

[0117] 帧间解码器(880)被配置为接收帧间预测信息,并基于帧间预测信息生成帧间预测结果。

[0118] 帧内解码器(872)被配置为接收帧内预测信息,并基于帧内预测信息生成预测结果。

[0119] 残差解码器(873)被配置为执行逆量化,以提取去量化的变换系数,并处理去量化的变换系数,以将残差信息从频域转换到空间域。残差解码器(873)还可能某些控制信息(以包括量化参数(QP)),并且该信息可以由熵解码器(871)提供(数据路径未示出,因为这可能只是少量控制信息)。

[0120] 重构模块(874)被配置为在空间域中组合由残差解码器(873)输出的残差信息和预测结果(由帧间或帧内预测模块输出,视情况而定),以形成重构块,该重构块可以是重构图片的一部分,该重构图片又可以是重构视频的一部分。注意,可以执行诸如去块操作等其他合适的操作来提高视觉质量。

[0121] 注意,视频编码器(403)、视频编码器(603)和视频编码器(703)以及视频解码器(410)、视频解码器(510)和视频解码器(810)可以使用任何合适的技术来实现。在一个实施例中,视频编码器(403)、视频编码器(603)和视频编码器(703)以及视频解码器(410)、视频解码器(510)和视频解码器(810)可以使用一个或多个集成电路来实现。在另一个实施例中,视频编码器(403)、视频编码器(603)和视频编码器(603)以及视频解码器(410)、视频解码器(510)和视频解码器(810)可以使用执行软件指令的一个或多个处理器来实现。

[0122] 在VVC中可以使用各种帧间预测模式。对于帧间预测的CU,运动参数可包括MV、一个或多个参考图片索引、参考图片列表使用索引以及将被用于帧间预测样本生成的某些编码特征的额外信息。运动参数可以以显式或隐式的信令形式通知。当以跳过模式对CU进行编码时,CU可以与PU相关联,并且可以没有显著的残差系数、没有编码的运动矢量增量或MV

差(例如,运动矢量数据, (Motion Vector Data, 简称MVD) 或参考图片索引。合并模式可以被指定,其中,从包括空间和/或时间候选的相邻CU,获得当前CU的运动参数。合并模式可以应用于帧间预测的CU,而不仅仅是用于跳过模式。在一个示例中,合并模式的替代方案是运动参数的显式传输,其中,每个CU的MV、每个参考图片列表对应的参考图片索引以及参考图片列表使用标志和其他信息被以信令的形式显示地发出。

[0123] 可以在VVC中使用某些帧间预测编码工具(例如,细化帧间预测编码工具),例如,双向光流 (Bi-Directional Optical Flow, 简称BDOF) 模式、解码器侧运动矢量细化 (Decoder-side Motion Vector Refinement, 简称DMVR) 模式和具有运动矢量差分合并 (Merge with Motion Vector Difference, 简称MMVD) 模式。

[0124] 根据本公开的实施例,BDOF模式是在帧间预测中使用的运动细化技术。BDOF模式可用于细化CU的双向预测信号,例如,在 $4 \times 4$ 子块级别。当满足某些条件时,BDOF模式可以应用于CU。在一个示例中,上述条件包含:(i) 使用“真”双向预测模式对CU进行编码,例如,两个参考图片中的一个在显示顺序中的当前图片之前,两个参考图片中的另一个在显示顺序中的当前图片之后,(ii) 从两个参考图片到当前图片的距离(例如,图片顺序计数 (Picture Order Count, 简称POC) 差) 相同,(iii) 两个参考图片是短期参考图片,(iv) CU并未使用仿射模式或基于子块的时间运动矢量预测 (Subblock-based Temporal Motion Vector Prediction, 简称SbTMVP) 合并模式来编码,(v) CU具有超过64个亮度样本,(vi) CU高度和CU宽度大于或等于8个亮度样本,(vii) 具有CU权重的双向预测 (Bi-Prediction with CU Weight, 简称BCW) 的权重索引指示相等的权重,(viii) 未针对当前CU启用加权预测 (Weighted Prediction, 简称WP), 以及 (ix) 组合帧间预测 (Combined Inter-Intra Prediction, 简称CIIP) 模式并未用于当前CU。

[0125] 在一个实施例中,BDOF模式仅应用于亮度分量。BDOF模式中的运动细化基于光流概念,例如,假设物体的运动是平滑的。参看图9,对于每个子块,例如, $4 \times 4$ 子块(例如, CU (910) 中的子块 (920)), 可通过最小化L0与L1预测样本之间的差来计算运动细化(例如,MV细化或MVR) ( $\Delta mv_x, \Delta mv_y$ )。在一个示例中,L0预测样本和L1预测样本分别是基于第一参考图片 (L0) 和第二参考图片 (L1) 的预测样本。运动细化可用于调整 $4 \times 4$ 子块(例如,子块 (920)) 中的双向预测样本值。以下步骤可应用于BDOF过程。

[0126] 可以通过直接计算两个相邻样本之间的差,来计算两个预测信号的水平 and 垂直梯度  $\frac{\partial I^{(k)}}{\partial x}(i,j)$  和  $\frac{\partial I^{(k)}}{\partial y}(i,j)$  ( $k = 0,1$ ), 等式1如下:

$$\frac{\partial I^{(k)}}{\partial x}(i,j) = ((I^{(k)}(i+1,j) \gg \text{shift1}) - (I^{(k)}(i-1,j) \gg \text{shift1}))$$

[0127]

$$\frac{\partial I^{(k)}}{\partial y}(i,j) = ((I^{(k)}(i,j+1) \gg \text{shift1}) - (I^{(k)}(i,j-1) \gg \text{shift1}))$$

[0128] 其中, $I^{(k)}(i,j)$  是列表k中的预测信号的坐标 (i, j) 处的样本值, $k=0,1$ , 并且参数  $\text{shift1}$  是基于亮度比特深度  $\text{bitDepth}$  (BD) 计算的。在一个示例中,  $\text{shift1} = \max(6, \text{bitDepth} - 6)$ 。

[0129] 在一个示例中,梯度  $S_1, S_2, S_3, S_5$  和  $S_6$  的自相关和互相关计算如下。等式2如下:

$$[0130] \quad S_1 = \sum_{(i,j) \in \Omega} \text{Abs}(\psi_x(i,j)), S_3 = \sum_{(i,j) \in \Omega} \theta(i,j) \cdot \text{Sign}(\varphi_x(i,j))$$

$$[0131] \quad S_2 = \sum_{(i,j) \in \Omega} \psi_x(i,j) \cdot \text{Sign}(\psi_y(i,j))$$

$$[0132] \quad S_5 = \sum_{(i,j) \in \Omega} \text{Abs}(\psi_y(i,j)), S_6 = \sum_{(i,j) \in \Omega} \theta(i,j) \cdot \text{Sign}(\psi_y(i,j))$$

[0133] 其中,等式3如下:

$$\psi_x(i,j) = \left( \frac{\partial I^{(1)}}{\partial x}(i,j) + \frac{\partial I^{(0)}}{\partial x}(i,j) \right) \gg n_a$$

[0134]

$$\psi_y(i,j) = \left( \frac{\partial I^{(1)}}{\partial y}(i,j) + \frac{\partial I^{(0)}}{\partial y}(i,j) \right) \gg n_a$$

$$[0135] \quad \theta(i,j) = (I^{(1)}(i,j) \gg n_b) - (I^{(0)}(i,j) \gg n_b)$$

[0136] 并且 $\Omega$ 是子块(例如,子块(920))周围的窗口,例如, $6 \times 6$ 窗口(例如,窗口(940)),并且值 $n_a$ 和 $n_b$ 分别被设置为等于 $\min(1, \text{bitDepth}-11)$ 和 $\min(4, \text{bitDepth}-8)$ 。

[0137] 可以使用自相关和互相关导出如下MVR( $\Delta mv_x, \Delta mv_y$ )。等式4如下:

$$[0138] \quad \Delta mv_x = S_1 > 0 ? \text{clip3} \left( -th'_{BIO}, th'_{BIO}, -((S_3 \cdot 2^{n_b - n_a}) \gg \lfloor \log_2 S_1 \rfloor) \right) : 0$$

[0139] 等式5如下:

$$\Delta mv_y = S_5 >$$

$$[0140] \quad 0 ? \text{clip3} \left( -th'_{BIO}, th'_{BIO}, -\left( (S_6 \cdot 2^{n_b - n_a} - ((v_x S_{2,m}) \ll n_{s_2} + v_x S_{2,s}) / 2 \right) \gg \lfloor \log_2 S_5 \rfloor \right) : 0$$

[0141] 其中, $S_{2,m} = S_2 \gg n_{s_2}$ ,  $S_{2,s} = S_2 \& (2^{n_{s_2}} - 1)$ ,  $\lfloor \cdot \rfloor$ 是底函数,  $n_{s_2} = 12$ 。在一个示例中,  $th'_{BIO} = 2^{\max(5, \text{BD}-7)}$ 。

[0142] 基于运动细化和梯度,可以为子块(例如, $4 \times 4$ 子块)中的每个样本计算以下调整(或预测细化) $b(x,y)$ ,等式6如下:

$$[0143] \quad b(x,y) = \text{rnd} \left( \frac{\left( \Delta mv_x \left( \frac{\partial I^{(1)}(x,y)}{\partial x} - \frac{\partial I^{(0)}(x,y)}{\partial x} \right) + \Delta mv_y \left( \frac{\partial I^{(1)}(x,y)}{\partial y} - \frac{\partial I^{(0)}(x,y)}{\partial y} \right) + 1 \right)}{2} \right)$$

[0144] 其中,函数‘rnd’表示舍入运算。

[0145] CU的BDOF样本可以通过如下调整双向预测样本来计算,等式7如下:

$$[0146] \quad \text{pred}_{\text{BDOF}}(x,y) = (I^{(0)}(x,y) + I^{(1)}(x,y) + b(x,y) + o_{\text{offset}}) \gg \text{shift}$$

[0147] 在一个实施例中,选择这些值,使得BDOF过程中的乘数不超过15位,并且BDOF过程中的中间参数的最大比特位宽保持在32位内。

[0148] 在一个示例中,为了导出梯度值,将生成在当前CU边界之外的列表 $k$ ( $k=0,1$ )中的预测样本 $I^{(k)}(i,j)$ 。参考图9,BDOF(例如,在VVC中)可以使用CU的边界周围的扩展行和/或扩展列(910)。为了控制生成边界外预测样本(例如,对样本(930)的预测)的计算复杂度,可

以通过在附近的整数位置取参考样本(例如,对坐标使用 $\text{floor}()$ 取整运算)来直接生成扩展区域(图9中的白色位置)中的预测样本(例如,930),而无需插值。在一个示例中,使用8抽头运动补偿内插滤波器(或8抽头内插滤波器)来生成CU(910)内的预测样本(图9中的灰色位置)。在图9所示的示例中,CU(910)包括 $8 \times 8$ 样本。在一个示例中,扩展样本值仅用于梯度计算。在一个示例中,使用等式1来执行梯度计算,如上所述。

[0149] 参考图9,对于BDOF过程中的剩余步骤,如果需要使用CU边界之外的样本和梯度值,则可以通过从对应的最近邻居填充(例如,重复)来生成CU边界之外的样本和梯度值,如图9中的箭头所示。例如,可以从(921)处的对应样本和梯度值中填充(930)处的预测样本和梯度值。

[0150] 当CU的宽度和/或高度大于阈值(例如,16个亮度样本)时,CU可以被分割成宽度和/或高度等于16个亮度样本的子块。子块的边界可以被视为BDOF过程中CU的边界。BDOF过程的最大单位大小可限于一个区域,例如, $16 \times 16$ 亮度样本。对于每个子块,可以跳过BDOF过程。在一个示例中,当初始L0和L1预测样本之间的绝对差之和(Sum of Absolute Differences,简称SAD)小于阈值时,BDOF过程不被应用于子块。阈值可以设置为 $8 \times W \times (H > 1)$ ,其中,W表示子块宽度,H表示子块高度。为了避免SAD计算的额外复杂性,可以复用在DMVR模式或DMVR过程中计算的初始L0和L1预测样本之间的SAD。

[0151] 如果为当前块启用BCW模式,例如,当BCW权重索引指示不相等的权重时,可以禁用BDOF模式。如果为当前块启用WP模式,例如,当两个参考图片中的任一个的标志(例如,luma\_weight\_lx\_flag)为1时,BDOF模式将被禁用。当CU采用对称MVD模式或CIIP模式编码时,BDOF将被禁用。

[0152] 根据本公开的一个方面,例如,在VVC中使用的DMVR模式可用于提高和/或细化合并模式中的MV的准确性。

[0153] 图10示出了根据本公开的实施例的基于双边匹配(Bilateral-Matching,简称BM)或双边模板匹配的DMVR模式的示例。当前块(1011)在当前图片(1010)中。初始MV(例如,MV0和MV1)分别指参考图片列表L0中的参考图片(Ref0)(1020)中的参考块(1022)和参考图片列表L1中的参考图片(Ref1)(1030)中的参考块(1032)。

[0154] 在一些示例中,例如,在VVC中,可以应用基于BM的DMVR来提高合并模式的运动矢量的准确性。在双向预测操作中,可分别在参考图片列表L0和参考图片列表L1中的初始MV(例如,MV0和MV1)周围搜索细化的MV。可以确定细化的参考块(或候选块)(1021)和(1031)。双边匹配方法可以基于参考图片列表L0和参考图片列表L1中的两个候选块(1021)和(1031)来计算失真。

[0155] 在一个示例中,计算由初始MV(例如,MV0和MV1)周围的相应MV候选(例如,MV0'和MV1')参考的候选块(1021)和(1031)之间的SAD。具有最低SAD的MV候选(例如,MV0'和MV1')可以是细化的MV,并且可以用于生成双向预测信号。

[0156] 可以初始MV和MV候选之间的关系描述如下。

$$[0157] \quad MV0' = MV0 + MV_{diff} \quad \text{等式8}$$

$$[0158] \quad MV1' = MV1 - MV_{diff} \quad \text{等式9}$$

[0159] 其中, $MV_{diff}$ 指示候选MV(例如,MV0')和参考图片中的一个(例如,L0)中的初始MV(例如,MV0)之间的MV差(MVD)。

[0160] 在一些示例中,例如,在VVC的一些应用中,DMVR的应用被限制。例如,DMVR模式适用于以特定模式和特征来编码的CU。在一个示例中,模式和特征包括(i)以具有双向预测MV的CU级合并模式对CU进行编码;(ii)相对于当前图片,一个参考图片(例如,后向参考图片)在过去,而另一个参考图片(例如,前向参考图片)在未来;(iii)从两个参考图片到当前图片的距离(例如,POC差)是相同的,(iv)两个参考图片是短期参考图片,(v)CU具有多于64个亮度样本,(vi)CU高度和CU宽度大于或等于8个亮度样本,(vii)BCW权重索引指示相等的权重,(viii)对于当前CU未启用WP模式,以及(ix)对于当前CU不使用CIIP模式。

[0161] 由DMVR过程导出的细化MV可用于生成帧间预测样本,并用于未来图片编码的时间运动矢量预测。原始MV(例如,初始MV)可用于去块过程中,并用于未来CU编码的空间运动矢量预测中。

[0162] 在一个实施例中,使用MMVD模式,例如,在VVC,其中,隐式导出的运动信息可用于预测CU(例如,当前CU)的样本。例如,在以信令形式通知跳过标志或合并标志之后,MMVD的合并标志可以以信令形式通知,以指定MMVD模式是否用于CU。

[0163] 在MMVD模式中,在选择了合并候选(也称为MV基础或MV起始点)之后,可以通过额外信息(例如,以信令形式通知MVD信息)来细化合并候选。额外信息可以包括用于指定运动幅度的索引(例如,距离索引,例如,mmvd\_distance\_idx[x0][y0])和用于指示运动方向的索引(例如,方向索引,例如,mmvd\_direction\_idx[x0][y0])。在MMVD模式下,可以选择合并列表中前两个候选中的一个,作为MV基础。例如,合并候选标志(例如,mmvd\_cand\_flag[x0][y0])指示合并列表中前两个候选中的一个。合并候选标志可以以信令形式通知,以指示(例如,指定)选择了前两个候选中的哪一个。额外信息可以向MV基础指示MVD(或运动偏移)。例如,运动幅度指示MVD的幅度,运动方向指示MVD的方向。

[0164] 在一个示例中,从合并候选列表中选择合并候选用于在参考图片处提供起始点或MV起始点。当前块的运动矢量可以用起始点和包括相对于起始点的运动幅度和运动方向的运动偏移(或MVD)来表示。在编码器侧,合并候选的选择和运动偏移的确定可以基于搜索过程(评估过程),如图11所示。在解码器侧,可以基于来自编码器侧的信号来确定所选择的合并候选和运动偏移。

[0165] 图11示出了MMVD模式下的搜索过程(1100)的示例。图12示出了MMVD模式中的搜索点的示例。在一些示例中,在图11的搜索过程(1100)中使用图12中的搜索点的子集或整个集合。通过执行搜索过程(1100),例如,在编码器侧,可以为当前图片(或当前帧)中的当前块(1101)确定包括合并候选标志(例如,mmvd\_cand\_flag[x0][y0])、距离索引(例如,mmvd\_distance\_idx[x0][y0])和方向索引(例如,mmvd\_direction\_idx[x0][y0])的额外信息。

[0166] 属于第一合并候选的第一运动矢量(1111)和第二运动矢量(1121)已被示出。第一运动矢量(1111)和第二运动矢量(1121)是在搜索过程(1100)中使用的MV起始点。第一合并候选可以是当前块(1101)构建的合并候选列表上的合并候选。第一运动矢量(1111)和第二运动矢量(1121)可以分别与参考图片列表L0和参考图片列表L1中的两个参考图片(1102)和(1103)相关联。参考图11-12,第一运动矢量(1111)和第二运动矢量(1121)可以分别指向参考图片(1102)和(1103)中的两个起始点(1211)和(1221),如图12所示。

[0167] 参考图12,可在参考图片(1102)和(1103)处确定图12中的两个起始点(1211)和(1221)。在一个示例中,基于起始点(1211)和(1221),可以评估在参考图片(1102)和(1103)

中在垂直方向(由+Y或-Y表示)或水平方向(由+X和-X表示)上从起始点(1211)和(1221)延伸的多个预定义点。在一个示例中,相对于相应的起始点(1211)或(1221)彼此反射的一对点(例如,一对点(1214)和(1224)(例如,由图11中的1S的移位指示)或者一对点(1215)和(1225)(例如,由图11中的2S的移位指示))可以用于确定一对运动矢量(例如,图11中的MV(1113)和(1115))。可以评估基于起始点(1211)或(1221)周围的预定义点确定的运动矢量预测值候选(例如,图11中的MV(1113)和(1123))。

[0168] 距离索引(例如,mmvd\_distance\_idx[x0][y0])可以指定运动幅度信息,并指示从由合并候选标志指示的起始点开始的预定义偏移(例如,图11中的1S或2S)。

[0169] 参考图11,可将偏移(例如,MVD(1112)或MVD(1122))应用(例如,添加)至起始MV(例如,MV(1111)或(1121))的水平分量或垂直分量。在表1中指定了距离索引(Distance Index,简称IDX)和预定义偏移的示例性关系。当全像素MMVD关闭时,例如,全像素MMVD标志(例如,slice\_fpel\_mmvd\_enabled\_flag)等于0,MMVD预定义偏移的范围可以从1/4个亮度样本到32个亮度样本。当全像素MMVD关闭时,预定义偏移可具有非整数值,例如,亮度样本的分数(例如,1/4像素或1/2像素)。当全像素MMVD开启时,例如,全像素MMVD标志(例如,slice\_fpel\_mmvd\_enabled\_flag)等于1,MMVD预定义偏移的范围可以从1个亮度样本到128个亮度样本。在一个示例中,当全像素MMVD开启时,预定义偏移仅具有整数值,例如,一个或多个亮度样本。

[0170] 表1:距离索引和偏移(例如,预定义的偏移)的示例性关系

距离 IDX	0	1	2	3	4	5	6	7
偏移(以亮度样本为单位)								
全像素 MMVD 关闭	1/4	1/2	1	2	4	8	16	32
偏移(以亮度样本为单位)								
全像素 MMVD 开启	1	2	4	8	16	32	64	128

[0172] 方向索引可以表示MVD相对于起始点的方向(或运动方向)。在一个示例中,方向索引表示表2所示的四个方向之一。表2中的MVD符号的含义可以根据起始MV的信息而变化。在一个示例中,当起始MV是单向预测MV或起始MV是双向预测MV且两个参考列表均指向当前图片的同一侧时(例如,两个参考图片的POC均大于当前图片的POC,或两个参考图片的POC均小于当前图片的POC),表2中的MVD符号指定添加到起始MV的MV偏移(或MVD)的符号。

[0173] 当起始MV为双向预测MV,且两个MV指向当前图片的不同侧时(例如,一个参考图片的POC大于当前图片的POC,另一个参考图片的POC小于当前图片的POC),表2中的MVD符号是指定添加到起始MV的list0 MV分量的MV偏移(或MVD)的符号,且list1 MV的MVD符号具有相反的值。参考图11,起始MV(1111)和(1121)是双向预测MV,两个MV(1111)和(1121)指向当前图片的不同侧。L1参考图片(1103)的POC大于当前图片的POC,并且L0参考图片(1102)的POC小于当前图片的POC。由表2中的方向索引(例如,00)指示的MVD符号(例如,x轴的符号“+”)指定添加到起始MV(例如,(1111))的list0 MV分量的MVD(例如,MVD(1112))的符号(例如,x

轴的符号“+”),起始MV(例如,(1121))的list1 MV分量的MVD(1122)的MVD符号具有相反的值,例如,与MVD(1112)的符号“+”相反的符号“-”。

[0174] 参考表2,方向索引00表示x轴的正方向,方向索引01表示x轴的负方向,方向索引10表示y轴的正方向,方向索引11表示y轴的负方向。

[0175] 表2: MV偏移符号和方向索引之间的示例性关系

方向IDX	00	01	10	11
x轴	+	-	N/A	N/A
y轴	N/A	N/A	+	-

[0177] 语法元素mmvd\_merge\_flag[x0][y0]可用于表示当前CU的MMVD的合并标志的取值。在一个示例中,MMVD的合并标志的取值(例如,mmvd\_merge\_flag[x0][y0])等于1,将指定MMVD模式用于生成当前CU的帧间预测参数。MMVD是合并标志的取值(例如,mmvd\_merge\_flag[x0][y0])等于0,将指定MMVD模式不用于生成帧间预测参数。阵列索引x0和y0可以指定所考虑的编码块(例如,当前CB)的左上亮度样本相对于图片(例如,当前图片)的左上亮度样本的位置(x0,y0)。

[0178] 当MMVD的合并标志的取值(例如,mmvd\_merge\_flag[x0][y0])对于当前CU不存在时,MMVD的合并标志的取值(例如,mmvd\_merge\_flag[x0][y0])可以被推断为对于当前CU,其值等于0。

[0179] 在一些示例中,例如,在VVC规范中,使用单个上下文以信令形式通知MMVD的合并标志(例如,mmvd\_merge\_flag)。例如,单个上下文用于在上下文自适应二进制算术编码(Context-Adaptive Binary Arithmetic Coding,简称CABAC)中编码(例如,编码和/或解码)MMVD的合并标志。

[0180] 语法元素mmvd\_cand\_flag[x0][y0]可以表示合并候选标志。在一个示例中,合并候选标志(例如,mmvd\_cand\_flag[x0][y0])指定合并候选列表中的第一(0)候选还是第二(1)候选,是否与从距离索引(例如,mmvd\_distance\_idx[x0][y0])和方向索引(例如,mmvd\_direction\_idx[x0][y0])导出的MVD一起使用。阵列索引x0和y0可以指定所考虑的编码块(例如,当前CB)的左上亮度样本相对于图片(例如,当前图片)的左上亮度样本的位置(x0,y0)。

[0181] 当合并候选标志(例如,mmvd\_cand\_flag[x0][y0])不存在时,合并候选标志(例如,mmvd\_cand\_flag[x0][y0])可以被推断为,其值等于0。

[0182] 语法元素mmvd\_distance\_idx[x0][y0]可以表示距离索引。在一个示例中,距离索引(例如,mmvd\_distance\_idx[x0][y0])指定用于导出MmvdDistance[x0][y0]的索引,如表3中所指定的。阵列索引x0和y0可以指定所考虑的编码块(例如,当前CB)的左上亮度样本相对于图片(例如,当前图片)的左上亮度样本的位置(x0,y0)。

[0183] 表3: MmvdDistance[x0][y0]和mmvd\_distance\_idx[x0][y0]之间的示例性关系

mmvd_distance_idx[x0][y0]	MmvdDistance[x0][y0]	
	slice_fpel_mmvd_enabled_flag == 0	slice_fpel_mmvd_enabled_flag == 1
0	1	4
1	2	8
[0184] 2	4	16
3	8	32
4	16	64
5	32	128
6	64	256
7	128	512

[0185] 表3中的第一列指示距离索引(例如,mmvd\_distance\_idx[x0][y0])。表3中的第二列指示当全像素MMVD关闭时的运动幅度(例如,MmvdDistance[x0][y0]),例如,全像素MMVD标志的取值(例如,slice\_fpel\_mmvd\_enabled\_flag)等于0。表3中的第三列指示当全像素MMVD开启时的运动幅度(例如,MmvdDistance[x0][y0]),例如,全像素MMVD标志的取值(例如,slice\_fpel\_mmvd\_enabled\_flag)等于1。

[0186] 在一个示例中,表3中第二列和第三列的单位为1/4个亮度样本。参考表3的第一行,当距离索引(例如,mmvd\_distance\_idx[x0][y0])为0时,当全像素mmvd关闭时(例如,slice\_fpel\_mmvd\_enabled\_flag为0),运动幅度(例如,MmvdDistance[x0][y0])为1。运动幅度(例如,MmvdDistance[x0][y0])为1×1/4个亮度样本或1/4个亮度样本。当距离索引(例如,mmvd\_distance\_idx[x0][y0])为0时,当全像素mmvd开启时(例如,slice\_fpel\_mmvd\_enabled\_flag为1),运动幅度(例如,MmvdDistance[x0][y0])为4。运动幅度(例如,MmvdDistance[x0][y0])为4×1/4个亮度样本或1个亮度样本。

[0187] 在一个示例中,表3中的第二列(以1/4个亮度样本为单位)对应于表1中的第二行(以亮度样本为单位),表3中的第三列(以1/4个亮度样本为单位)对应于表1中的第三行(以亮度样本为单位)。

[0188] 语法元素mmvd\_direction\_idx[x0][y0]可以表示方向索引。在一个示例中,方向索引(例如,mmvd\_direction\_idx[x0][y0])指定用于导出运动方向(例如,MmvdSign[x0][y0])的索引,如表4中所指定。阵列索引x0和y0指定所考虑的编码块(例如,当前CB)的左上亮度样本相对于图片(例如,当前图片)的左上亮度样本的位置(x0,y0)。表4中的第一列指示方向索引(例如,mmvd\_distance\_idx[x0][y0])。表4中的第二列指示MVD的第一分量(例如,MVD<sub>x</sub>或MmvdOffset[x0][y0][0])的第一符号(例如,MmvdSign[x0][y0][0])。表4中的第三列指示MVD的第二分量(例如,MVD<sub>y</sub>或MmvdOffset[x0][y0][1])的第二符号(例如,

MmvdSign[x0][y0][1])。

[0189] 表4:MmvdSign[x0][y0]和mmvd\_direction\_idx[x0][y0]之间的示例性关系

mmvd_direction_idx[x0][y0]	MmvdSign[x0][y0][0]	MmvdSign[x0][y0][1]
0	+1	0
1	-1	0
2	0	+1
3	0	-1

[0191] MVD的第一分量(例如,MmvdOffset[x0][y0][0])和第二分量(例如,MmvdOffset[x0][y0][1])或偏移MmvdOffset[x0][y0],可以推导如下:

[0192]  $MmvdOffset[x0][y0][0] = (MmvdDistance[x0][y0] \ll 2) \times MmvdSign[x0][y0][0]$

等式10

[0193]  $MmvdOffset[x0][y0][1] = (MmvdDistance[x0][y0] \ll 2) \times MmvdSign[x0][y0][1]$

等式11

[0194] 在一个示例中,距离索引(例如,mmvd\_distance\_idx[x0][y0])为3,方向索引(例如,mmvd\_direction\_idx[x0][y0])为2。基于表4和方向索引(例如,mmvd\_direction\_idx[x0][y0])为2,MVD的第一分量(例如,MVD<sub>x</sub>或MmvdOffset[x0][y0][0])的第一符号(例如,MmvdSign[x0][y0][0])为0,第二分量(例如,MVD<sub>y</sub>或MmvdOffset[x0][y0][1])的第二符号(例如,MmvdSign[x0][y0][1])为“+1”。在本示例中,MVD沿着正垂直方向(+y),没有水平分量。

[0195] 当全像素MMVD标志的取值(例如,slice\_fpel\_mmvd\_enabled\_flag)等于0且全像素MMVD关闭时,基于表3和距离索引(例如,mmvd\_distance\_idx[x0][y0])为3,MmvdDistance[x0][y0]指示的运动幅度为8。基于等式10-11,MVD的第一分量(例如,MmvdOffset[x0][y0][0])是 $(8 \ll 2) \times 0 = 0$ ,MVD的第二分量(例如,MmvdOffset[x0][y0][1])是 $(8 \ll 2) \times (+1) = 2$ (亮度样本)。

[0196] 当全像素MMVD标志的取值(例如,slice\_fpel\_mmvd\_enabled\_flag)等于1且全像素MMVD开启时,基于表3和距离索引(例如,mmvd\_distance\_idx[x0][y0])为3,MmvdDistance[x0][y0]指示的运动幅度为32。基于等式10-11,MVD的第一分量(例如,MmvdOffset[x0][y0][0])是 $(32 \ll 2) \times 0 = 0$ ,MVD的第二分量(例如,MmvdOffset[x0][y0][1])是 $(32 \ll 2) \times (+1) = 8$ (亮度样本)。

[0197] 本公开包括与MMVD模式下的信令传输改进相关的实施例。

[0198] 在一个实施例中,当BDOF模式和/或DMVR模式应用于具有前向参考图片和后向参考图片的双向预测图片(例如,B图片)中的编码块时,由MMVD模式实现的对双向预测图片中的编码块的编码效率提高可能不显著。因此,在一些示例中,当BDOF模式和/或DMVR模式被应用于编码块时,MMVD模式不被应用于编码块。另一方面,在各种示例中,不管MMVD模式是否应用于编码块,编码块的MMVD的合并标志的取值都以信令形式通知。以信令形式通知编码块的MMVD的合并标志的信令成本可能相对较大,例如,因为当MMVD模式不可能应用于编码块时,对编码块以信令形式通知MMVD的合并标志。

[0199] 根据本公开的一个实施例,当前图片中的每个块可以不以信令形式通知MMVD的合并标志(例如,由语法元素mmvd\_merge\_flag指示)。基于是否满足一个或多个要求,当前图片中的当前块可以有条件地以信令形式通知MMVD的合并标志。这些要求可以包括第一要求

和第二要求。第一要求可以包括,其他帧间预测模式(例如,BDOF模式和/或DMVR模式),其被包括当前块的多个块允许。第二要求可以包括,多个块的高电平标志(例如,MMVD合并启用标志)指示对于多个块未启用MMVD模式。在一些实施例中,第一要求还包括,多个块的参考图片包括前向参考图片和后向参考图片。在一些实施例中,第一要求还包括,多个块的参考图片相对于当前图片对称。

[0200] 在一个示例中,基于DMVR模式和/或BDOF模式的使用,可以有条件地以信令形式通知MMVD的合并标志(例如,mmvd\_merge\_flag)。

[0201] 在解码器侧,可以从编码视频比特流中解码当前图片中多个块的预测信息。可以基于预测信息来确定是否满足至少一个要求。在一些实施例中,要求包括(i)第一要求(例如,多个块允许DMVR模式或BDOF模式中的至少一个)和(ii)第二要求(例如,MMVD的合并启用标志指示对于多个块未启用MMVD模式)。可以基于是否满足至少一个要求来确定多个MMVD的合并标志是否被推断为对于多个块为假。多个MMVD的合并标志中的每一个可以针对多个块中的相应一个指示MMVD模式是否被应用于该相应的块。响应于多个MMVD的合并标志被推断为对于多个块为假的确定结果,可以在不应用MMVD模式的情况下重构多个块中的每个块。由于多个块的多个MMVD的合并标志并没有在编码视频比特流中以信令形式通知(例如,多个MMVD的合并标志被推断为假),所以提高编码效率。

[0202] 在一个示例中,多个块是当前图片中的切片。在一个示例中,多个块是当前图片。

[0203] 在一个实施例中,预测信息包括至少一个标志,指示多个块是否允许DMVR模式或BDOF模式中的至少一个。例如,预测信息包括第一切片级标志,指示对于当前图片的切片(例如,当前切片)中的多个块是否允许DMVR模式。例如,预测信息包括第二切片级标志,指示切片中的多个块是否允许BDOF模式。

[0204] 在一个实施例中,合并第一要求,要求多个块允许DMVR模式或BDOF模式中的至少一个,多个块的参考图片包括前向参考图片和后向参考图片,并且前向参考图片和当前图片之间的第一图片顺序计数(Picture Order Count,简称POC)差的绝对值等于后向参考图片和当前图片之间的第二图片顺序计数POC差的绝对值。例如,基于锁满足的第一要求,可以确定多个块的多个MMVD的合并标志被推断为假。在一些示例中,第一要求还包括前向参考图片和后向参考图片是短期参考图片。例如,第一要求还包括第一POC差的绝对值和第二POC差的绝对值小于阈值(例如,2)。在一个示例中,阈值是2。

[0205] 在一个实施例中,如果当前切片(例如,多个块)允许DMVR模式和/或BDOF模式,当前切片具有前向参考图片和后向参考图片,并且当前图片在时间上位于两个参考图片(例如,前向参考图片和后向参考图片)的中间,则多个块不允许MMVD模式,且多个块中的每个块的MMVD的合并标志(例如,mmvd\_merge\_flag)不以信令形式通知。多个块中的每个块的MMVD的合并标志(例如,mmvd\_merge\_flag)可以被推断为假。

[0206] 使用MMVD模式的概率可取决于两个参考图片和当前图片之间的绝对POC差(例如,第一POC差的绝对值和第二POC差的绝对值)。在一个实施例中,当绝对POC差减小时,使用MMVD模式的概率也将减小。在一个示例中,可以进一步应用对当前图片和两个参考图片之间的时间距离(例如,绝对POC差)的进一步约束。例如,绝对POC差小于1或2。

[0207] 根据本公开的一个实施例,预测信息包括与多个块中的所有块对应的单个高级MMVD的合并启用标志。单个高级MMVD的合并启用标志可以指示对于多个块是否启用MMVD模

式。可以基于MMVD的合并启用标志来推断与多个块中的每一个对应的多个MMVD的合并标志。MMVD合并启用标志可以是高级标志,例如,切片报头(Slice Header,简称SH)处的切片级标志(例如,sh\_mmvd\_merge\_enable\_flag)。MMVD的合并启用标志可以是高级标志,例如,图片报头(Picture Header,简称PH)处的图片级标志(例如,ph\_mmvd\_merge\_enable\_flag)。

[0208] 在一个实施例中,MMVD的合并启用标志可以指示对于多个块未启用MMVD模式,这满足第二要求。在这种情况下,与多个块中的每一个对应的所有多个MMVD的合并标志可以被推断为假。

[0209] 在一个实施例中,MMVD的合并启用标志可以指示对于多个块启用MMVD模式,这不满足第二要求。在这种情况下,多个MMVD的合并标志不会被推断为假。相反,与多个块中的每一个对应的多个MMVD的合并标志可以在编码视频比特流中以信令形式通知,并在解码器侧接收。

[0210] 在一个示例中,预测信息可以指示:(i) 多个块允许DMVR模式或BDOF模式中的至少一个,(ii) 多个块的参考图片包括前向参考图片和后向参考图片,以及(iii) 前向参考图片和当前图片之间的第一POC差的绝对值等于后向参考图片和当前图片之间的第二POC差的绝对值。当MMVD的合并启用标志指示对于多个块启用MMVD模式时,多个MMVD的合并标志不被推断为假。

[0211] 切片报头或图片报头处的MMVD的合并启用标志可以指示MMVD合并模式是否可用于与切片报头对应的编码切片(例如,当前切片)或与图片报头对应的编码图片(例如,当前图片)。

[0212] 在一个实施例中,如果对应于切片的切片报头处的MMVD的合并启用标志(例如,sh\_mmvd\_merge\_enable\_flag)为假,则MMVD的合并标志(例如,mmvd\_merge\_flag)被推断为对于切片中的所有块为假。

[0213] 在一个实施例中,如果对应于图片的图片报头处的MMVD的合并启用标志(例如,ph\_mmvd\_merge\_enable\_flag)为假,则MMVD的合并标志(例如,mmvd\_merge\_flag)被推断对于图片中的所有块为假。

[0214] 在一个实施例中,响应于多个MMVD的合并标志未被推断为对于多个块(例如,切片中的块或图片中的块)为假的确定结果,则可以在编码视频比特流中以信令形式通知多个块中的每个块的单独的相应MMVD的合并标志的取值。可以使用CABAC对多个MMVD的合并标志进行编码。

[0215] 在相关视频编码技术中,在CABAC期间使用单个上下文(例如,称为旧上下文)对多个MMVD的合并标志中的至少一个MMVD的合并标志进行编码。根据本公开的实施例,可以使用多个上下文来CABAC编码(例如,编码和/或解码)多个MMVD的合并标志中的至少一个MMVD的合并标志。在一个示例中,在CABAC中使用多个上下文来编码(例如,编码和/或解码)多个MMVD的合并标志中的每个MMVD的合并标志。在一个示例中,多个上下文包括第一上下文和第二上下文。在一个示例中,第二上下文是在相关视频编码技术中使用的旧上下文,而第一上下文是在相关视频编码技术中未使用的新上下文。可以基于多个块是否允许DMVR模式或BDOF模式中的一个或多个,来选择多个上下文(例如,第一上下文和第二上下文)中的一个上下文。可使用具有所选上下文的CABAC来编码(例如,编码和/或解码)多个块中的每个块

相应的MMVD的合并标志。

[0216] 在一个实施例中,预测信息指示选择第一上下文(例如,新上下文)的条件。这些条件可以与第一要求相同或不同。在一个示例中,这些条件与第一要求相同。例如,这些条件包括:多个块允许DMVR模式或BDOF模式中的至少一个,多个块的参考图片包括前向参考图片和后向参考图片,并且前向参考图片和当前图片之间的第一POC差的绝对值等于后向参考图片和当前图片之间的第二POC差的绝对值。当满足条件时,可以选择第一上下文,作为所选的上下文。基于多个条件中未满足的一个条件,可以选择第二上下文,作为所选的上下文。

[0217] 在一个示例中,当第一上下文(例如,新上下文)被选择为CABAC上下文时,利用指示每个MMVD的合并标志为假的概率高于MMVD的合并标志为真的概率的概率,来初始化第一上下文。

[0218] 在一个实施例中,第一上下文(例如,新上下文)和第二上下文(例如,旧上下文)是MMVD的合并标志(例如,mmvd\_merge\_flag)的CABAC编码中使用的上下文。第一上下文(例如,新上下文)和第二上下文(例如,旧上下文)之间的选择可以基于DMVR模式和/或BDOF模式的使用。

[0219] 在一个实施例中,在第一上下文和第二上下文之间进行选择的条件与第一要求相同。如果满足条件(例如,为真),则选择第一上下文(例如,新上下文)。否则,使用第二上下文(例如,旧上下文)。

[0220] 在一个实施例中,选择第一上下文(例如,新上下文)时,用指示MMVD的合并标志(例如,mmvd\_merge\_flag)更可能为假(例如,该概率高于MMVD的合并标志为真的概率)的概率,来初始化第一上下文(例如,新上下文)。

[0221] 图13示出了概述根据本公开的实施例的编码过程(1300)的流程图。在各种实施例中,过程(1300)由处理电路执行,例如,终端设备(310)、终端设备(320)、终端设备(330)和终端设备(340)中的处理电路、执行视频编码器(例如,视频编码器(403)、视频编码器(603)、视频编码器(703))的功能的处理电路等。在一些实施例中,过程(1300)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(1300)。该过程从(S1301)开始,并进行到(S1310)。

[0222] 在(S1310)中,可以确定是否满足至少一个要求。这些要求可以包括第一要求和第二要求。第一要求可以包括当前图片中的多个块允许DMVR模式或BDOF模式中的至少一种。第一要求可以还包括额外要求。例如,第一要求包括:(i)多个块的参考图片包括前向参考图片和后向参考图片,以及(ii)前向参考图片和当前图片之间的第一图片顺序计数(Picture Order Count,简称POC)差的绝对值等于后向参考图片和当前图片之间的第二图片顺序计数POC差的绝对值。第二要求可以包括,对于多个块未启用MMVD模式。

[0223] 当满足至少一个要求时,过程(1300)进行到(S1320)。当没有满足任何要求时,过程(1300)进行到(S1340)。

[0224] 在(S1320)中,当满足至少一个要求时,可以在不应用MMVD模式的情况下对多个块中的每个块进行编码。

[0225] 在一个示例中,第一要求还包括,多个块的参考图片包括前向参考图片和后向参考图片,并且第一POC差的绝对值等于第二POC差的绝对值。满足第一要求。可以在不应用

MMVD模式的情况下对多个块中的每个块进行编码。

[0226] 在一个示例中,满足第二要求(例如,对于多个块,未启用MMVD模式),可以在不应用MMVD模式的情况下对多个块中的每个块进行编码。

[0227] 作为包括(S1320)的分支的一部分,过程(1300)进行到(S1330)。

[0228] 在(S1330),多个块的多个MMVD的合并标志在视频比特流中未编码并且未以信令形式通知。在一个示例中,对于多个块,未启用MMVD模式,因此诸如高级标志(例如,MMVD的合并启用标志)等标志被编码并在比特流中以信令形式通知,以指示对于多个块,未启用MMVD模式。

[0229] 在一个示例中,满足第一要求,并且没有以信令形式通知高级标志(例如,MMVD的合并启用标志)。解码器可以基于满足的第一要求来推断多个块的多个MMVD的合并标志为假。

[0230] 过程(1300)然后进行到(S1399),并终止。

[0231] 在(S1340),对多个块中的每个块进行单独确定,以确定是否使用MMVD模式对相应块进行编码,并根据该确定结果对相应块进行编码。

[0232] 在一个示例中,确定将MMVD模式应用于相应的块,并通过应用MMVD模式对相应的块进行编码。

[0233] 在另一示例中,确定MMVD模式不应用于相应的块,并且在不应用MMVD模式的情况下对相应的块进行编码。

[0234] 作为包括(S1340)的分支的一部分,过程(1300)进行到(S1350)。

[0235] 在(S1350),多个块的多个MMVD的合并标志在视频比特流中编码和以信令形式通知,其中,多个被编码和以信令形式通知的MMVD的合并标志中的每一个合并标志,将指示多个块中的相应块被确定为使用MMVD模式进行编码。

[0236] 在一个示例中,对于多个块启用MMVD模式,因此MMVD的合并启用标志在视频比特流中编码和以信令形式通知,以指示对于多个块启用MMVD模式。

[0237] 在步骤(S1350)之后,过程(1300)进行到(S1399),并终止。

[0238] 过程(1300)可适用于各种情况,过程(1300)中的步骤可相应调整。可以调整、省略、重复和/或组合过程(1300)中的一个或多个步骤。可以使用任何合适的顺序来实现该过程(1300)。可以添加额外的步骤。

[0239] 在一个实施例中,满足第一要求,不满足第二要求。在图13所示的示例中,过程(1300)进行到包括(S1320)和(S1330)的分支,其中,多个块中的每个块可以在不应用MMVD模式的情况下被编码(S1320),并且多个MMVD的合并标志未编码并且未以信令形式通知(S1330)。

[0240] 可选地(图13中未示出),当满足第一要求并且不满足第二要求时,其中,对于多个块启用MMVD模式,过程(1300)可进行到包括(S1340)和(S1350)的分支。例如,可以为多个块中的每个块确定是否使用MMVD模式对相应的块进行编码。多个MMVD的合并标志可以在比特流中被编码和以信令形式通知。此外,MMVD的合并启用标志可以被编码和以信令形式通知。

[0241] 图14示出了概述根据本公开的实施例的解码过程(1400)的流程图。在各种实施例中,过程(1400)由处理电路执行,例如,终端设备(310)、终端设备(320)、终端设备(330)和终端设备(340)中的处理电路、执行视频编码器(403)功能的处理电路、执行视频解码器

(410) 功能的处理电路、执行视频解码器 (510) 功能的处理电路、执行视频编码器 (603) 功能的处理电路等。在一些实施例中,过程 (1400) 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 (1400)。该过程从 (S1401) 开始,并进行到 (S1410)。

[0242] 在 (S1410),可从编码视频比特流中解码当前图片中的多个块的预测信息。多个块可以是当前图片中的切片。多个块可以是当前图片。

[0243] 在 (S1420),可以基于预测信息确定是否满足至少一个要求。这些要求可以包括 (i) 第一要求 (例如,对于多个块允许解码器侧运动矢量细化 (DMVR) 模式或双向光流 (BDOF) 模式中的至少一个) 以及 (ii) 第二要求 (例如,MMVD的合并启用标志指示对于多个块,未启用具有运动矢量差分合并 (MMVD) 模式)。

[0244] 在一个实施例中,预测信息包括一个或多个标志,指示多个块是否允许DMVR模式或BDOF模式中的至少一个。一个或多个标志可以包括高级标志。在一个示例中,一个或多个标志包括 (i) 指示切片是否允许DMVR模式的切片级标志和/或 (ii) 指示切片是否允许BDOF模式的切片级标志。

[0245] 在 (S1430),可以基于是否满足至少一个要求,来确定多个MMVD的合并标志是否被推断为对于多个块为假。多个MMVD的合并标志中的每一个合并标志可以分别指示MMVD模式是否被应用于多个块中的相应块。

[0246] 在一个实施例中,第一要求包括: (i) 多个块允许DMVR模式或BDOF模式中的至少一个, (ii) 多个块的参考图片包括前向参考图片和后向参考图片,以及 (iii) 前向参考图片和当前图片之间的第一图片顺序计数 (POC) 差的绝对值等于后向参考图片和当前图片之间的第二图片顺序计数POC差的绝对值。例如,当满足第一要求时,可以确定多个块的多个MMVD的合并标志被推断为假。在一个示例中,当满足第一要求时,多个块的多个MMVD的合并标志将不以信令形式通知。

[0247] 在一个示例中,第一要求还包括第一POC差的绝对值和第二POC差的绝对值小于阈值。阈值可以是2。

[0248] 在一个实施例中,预测信息包括多个块的MMVD的合并启用标志。MMVD的合并启用标志可以是高级标志 (例如,高于块),例如,切片报头处的切片级标志、图片报头处的图片级标志等。

[0249] 在一个示例中,MMVD的合并启用标志指示,对于多个块未启用MMVD模式,这满足第二要求。例如,当满足第二要求时 (例如,MMVD的合并启用标志指示,对于多个块未启用MMVD模式),可以确定多个块的多个MMVD的合并标志被推断为假。

[0250] 在 (S1440),响应于多个MMVD的合并标志被推断为对于多个块为假的确定结果,可以在不应用MMVD模式的情况下重构多个块中的每个块。在一个示例中,DMVR模式和/或BDOF模式被应用于多个块中的一个或多个块,并且使用DMVR模式和/或BDOF模式来重构多个块中的一个或多个块。由于多个MMVD的合并标志被推断为假,所以在编码视频比特流中没有以信令形式通知MMVD的合并标志的取值。

[0251] 在一个实施例中,响应于多个MMVD的合并标志未被推断为假的确定结果,执行 (S1440) 的可选步骤 (未示出)。在该可选步骤中,在编码视频比特流中以信令形式通知多个MMVD的合并标志的取值。如果在编码视频比特流中以信令形式通知多个MMVD的合并标志中的第一MMVD的合并标志指示,MMVD模式被应用于多个块中的第一块,则可以通过应用MMVD

模式来重构多个块中的第一块。如果在编码视频比特流中以信令形式通知多个MMVD的合并标志中的第二MMVD的合并标志指示,MMVD模式没有被应用于多个块中的第二块,则可以在不应用MMVD模式的情况下重构多个块中的第二块。

[0252] 在(S1440)可选的步骤之后,过程(1400)进行到(S1499),并终止。

[0253] 过程(1400)可适用于各种情况,过程(1400)中的步骤可相应调整。可以调整、省略、重复和/或组合过程(1400)中的一个或多个步骤。可以使用任何合适的顺序来实现该过程(1400)。可以添加额外的步骤。

[0254] 在一个实施例中,响应于多个MMVD的合并标志未被推断为对于多个块为假的确定结果,可在编码视频比特流中以信令形式通知多个块中的每个块的相应MMVD的合并标志的取值。可以基于多个块是否允许DMVR模式或BDOF模式中的一个或多个,在第一上下文和第二上下文中选择上下文。多个块中每个块相应的MMVD的合并标志可以使用具有所选上下文的上下文自适应二进制算术编码(Context-Adaptive Binary Arithmetic Coding,简称CABAC)来解码。

[0255] 在一个示例中,预测信息指示是否满足选择第一上下文的条件。这些条件包括:(i)多个块允许DMVR模式或BDOF模式中的一个或多个,(ii)多个块的参考图片包括前向参考图片和后向参考图片,以及(iii)前向参考图片和当前图片之间的第一POC差的绝对值等于后向参考图片和当前图片之间的第二POC差的绝对值。可以基于所满足的条件,选择第一上下文,作为所选的上下文。基于多个条件中未满足的一个条件,可以选择第二上下文,作为所选的上下文。

[0256] 在一个示例中,基于第一概率来初始化第一上下文,第一概率指示多个块中每个块相应的MMVD的合并标志为假的概率高于MMVD的合并标志为真的概率的概率。

[0257] 在一个示例中,预测信息包括多个块的MMVD的合并启用标志,MMVD的合并启用标志指示对于多个块,启用MMVD模式,这不满足第二要求。当不满足第二要求时(例如,MMVD的合并启用标志指示对于多个块,启用MMVD模式),确定多个MMVD的合并标志未被推断为假,并且多个块的多个MMVD的合并标志在编码视频比特流中以信令形式通知。

[0258] 在图14未示出的示例中,无论是否满足第一要求(例如,预测信息指示:(i)多个块允许DMVR模式或BDOF模式中的至少一个,(ii)多个块的参考图片包括前向参考图片和后向参考图片,以及(iii)前向参考图片和当前图片之间的第一POC差的绝对值等于后向参考图片和当前图片之间的第二POC差的绝对值),例如,当不满足第二要求时(例如,MMVD的合并启用标志指示对于多个块,启用MMVD模式),多个MMVD的合并标志不被推断为假,并且在编码视频比特流中以信令形式通知。

[0259] 在一些实施例中,从编码视频比特流中接收当前图片中的多个块的预测信息。可以基于预测信息确定多个块是否允许DMVR模式或BDOF模式中的至少一个。响应于多个块允许DMVR模式或BDOF模式中的至少一个的确定结果,多个MMVD的合并标志可以被推断为对于多个块为假,其中,多个MMVD合并标志分别指示,MMVD模式是否被应用于多个块。响应于多个MMVD的合并标志被推断为对于多个块为假的确定结果,可以在不应用MMVD模式的情况下重构多个块中的每个块。

[0260] 在一个实施例中,(i)当前图片参考前向参考图片和后向参考图片,以及(ii)当前图片位于前向参考图片和后向参考图片的中间。基于条件(i)和(ii),多个MMVD的合并标志

被确定为未以信令形式通知。因此,可以基于确定在当前图片中不允许MMVD模式,来推断多个MMVD的合并标志。

[0261] 在一个实施例中,预测信息包括多个块的MMVD的合并启用标志。MMVD的合并启用标志指示对于多个块,未启用MMVD模式,并且多个块的多个MMVD的合并标志被推断为假。

[0262] 本公开的实施例可单独使用或以任何顺序组合使用。此外,方法(或实施例)、编码器和解码器中的每一个可以由处理电路(例如,一个或多个处理器或一个或多个集成电路)来实现。在一个示例中,一个或多个处理器执行存储在非暂时性计算机可读介质中的程序。

[0263] 上述技术可以被实现为使用计算机可读指令的计算机软件,并且物理地存储在一个或多个计算机可读介质中。例如,图15示出了适于实现所公开主题的某些实施例的计算机系统(1500)。

[0264] 计算机软件可以使用任何合适的机器代码或计算机语言来编码,其可以经受汇编、编译、链接或类似机制来创建包括指令的代码,这些指令可以由一个或多个计算机中央处理单元(Central Processing Units,简称CPU)、图形处理单元(Graphics Processing Units,简称GPU)等直接执行,或者通过解释、微代码执行等来执行。

[0265] 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、平板计算机、服务器、智能手机、游戏设备、物联网设备等。

[0266] 图15中所示的计算机系统(1500)的组件本质上是示例性的,并且不旨在对实现本公开的实施例的计算机软件的使用范围或功能提出任何限制。组件的配置也不应被解释为对计算机系统(1500)的示例性实施例中所示的任何一个组件或组件组合有任何依赖性要求。

[0267] 计算机系统(1500)可以包括某些人机接口输入设备。这种人机接口输入设备可以响应一个或多个人类用户的输入,例如触觉输入(例如:击键、滑动、数据手套移动)、音频输入(例如:语音、鼓掌)、视觉输入(例如:手势)、嗅觉输入(未示出)进行的输入。人机接口设备还可以用于捕捉不一定与人的有意识输入直接相关的某些媒体,例如,音频(例如:语音、音乐、环境声音)、图像(例如:扫描图像、从静止图像相机获得的照片图像)、视频(例如,二维视频、包括立体视频的三维视频)。

[0268] 输入人机接口设备可以包括以下一个或多个(每个仅描绘了一个):键盘(1501)、鼠标(1502)、轨迹板(1503)、触摸屏(1510)、数据手套(未示出)、操纵杆(1505)、麦克风(1506)、扫描仪(1507)、相机(1508)。

[0269] 计算机系统(1500)还可以包括某些人机接口输出设备。这种人机接口输出设备可以刺激一个或多个人类用户的感受,例如通过触觉输出、声音、光和气味/味道。这种人机接口输出设备可以包括触觉输出设备(例如,通过触摸屏(1510)、数据手套(未示出)或操纵杆(1505)的触觉反馈,但是也可以有不用作输入设备的触觉反馈设备)、音频输出设备(例如:扬声器(1509)、耳机(未示出))、视觉输出设备(例如,屏幕(1510),包括CRT屏幕、LCD屏幕、等离子屏幕、OLED屏幕,每个都具有或不具有触摸屏输入能力,每个都具有或不具有触觉反馈能力——其中一些能够输出二维视觉输出或多于三维的输出,通过诸如立体输出之类的方式;虚拟现实眼镜(未示出)、全息显示器和烟雾箱(未示出))以及打印机(未示出)。

[0270] 计算机系统(1500)还可以包括人类可访问的存储设备及其相关联的介质,例如,包括具有CD/DVD或类似介质(1520)的CD/DVD ROM/RW(1520)的光学介质、拇指驱动器

(1522)、可移动硬盘驱动器或固态驱动器(1523)、诸如磁带和软盘(未示出)之类的传统磁介质、诸如安全加密狗(未示出)之类的专用ROM/ASIC/PLD设备等。

[0271] 本领域技术人员还应该理解,结合当前公开的主题使用的术语“计算机可读介质”不包括传输介质、载波或其他瞬时信号。

[0272] 计算机系统(1500)还可以包括到一个或多个通信网络(1555)的接口(1554)。网络例如可以是无线的、有线的、光学的。网络还可以是局域的、广域的、大都市的、车辆的和工业的、实时的、延迟容忍的等。网络的示例包括诸如以太网、无线LAN之类的局域网、包括GSM、3G、4G、5G、LTE等的蜂窝网络、包括有线电视、卫星电视和地面广播电视的电视有线或无线广域数字网络、包括CANBus的车辆和工业网络等。某些网络通常需要连接到某些通用数据端口或外围总线(1549)(例如,计算机系统(1500)的USB端口)的外部网络接口适配器;其他的通常通过连接到如下所述的系统总线而集成到计算机系统(1500)的核心中(例如,PC计算机系统以太网接口或智能电话计算机系统蜂窝网络接口)。使用这些网络中的任何一个,计算机系统(1500)可以与其他实体通信。这种通信可以是单向的、只接收的(例如,广播电视)、单向的、只发送的(例如,到某些CANbus设备的CANbus),或者是双向的,例如,到使用局域或广域数字网络的其他计算机系统。如上所述,某些协议和协议栈可以用在这些网络和网络接口的每一个上。

[0273] 前述人机接口设备、人类可访问的存储设备和网络接口可以附接到计算机系统(1500)的核心(1540)。

[0274] 核心(1540)可以包括一个或多个中央处理单元(Central Processing Units,简称CPU)(1541)、图形处理单元(Graphics Processing Units,简称GPU)(1542)、现场可编程门区域(Field Programmable Gate Areas,简称FPGA)(1543)形式的专用可编程处理单元、用于某些任务的硬件加速器(1544)、图形适配器(1550)等。这些设备连同只读存储器(Read-only memory,简称ROM)(1545)、随机存取存储器(1546)、诸如内部非用户可访问硬盘驱动器、SSD之类的内部大容量存储器(1547)可以通过系统总线(1548)连接。在一些计算机系统中,系统总线(1548)可以以一个或多个物理插头的形式访问,以允许额外CPU、GPU等的扩展。外围设备可以直接或者通过外围总线(1549)连接到核心的系统总线(1548)。在一个示例中,屏幕(1510)可以连接到图形适配器(1550)。外围总线的架构包括PCI、USB等。

[0275] CPU(1541)、GPU(1542)、FPGA(1543)和加速器(1544)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(1545)或RAM(1546)中。过渡数据可以存储在RAM(1546)中,而永久数据可以存储在例如内部大容量存储器(1547)中。可以通过使用高速缓冲存储器来实现对任何存储设备的快速存储和检索,高速缓冲存储器可以与一个或多个CPU(1541)、GPU(1542)、大容量存储器(1547)、ROM(1545)、RAM(1546)等紧密关联。

[0276] 计算机可读介质上可以具有用于执行各种计算机实现的操作的计算机代码。介质和计算机代码可以是为了本公开的目的而专门设计和构造的,或者可以是计算机软件领域的技术人员公知和可获得的类型。

[0277] 作为示例而非限制,具有架构的计算机系统(1500),特别是核心(1540)可以提供作为处理器(包括CPU、GPU、FPGA、加速器等)的结果的功能,其执行嵌入在一个或多个有形的计算机可读介质中的软件。这种计算机可读介质可以是与如上所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暂时性的核心(1540)的某些存储器,例如,核心内部大容量存储器(1547)或ROM(1545)。实现本公开的各种实施例的软件可以存储在这样的设备中并由核心(1540)执行。根据特定需要,计算机可读介质可以包括一个或多个存储设备或芯片。该软件可以使核心(1540)并且特别是其中的处理器(包括CPU、GPU、FPGA等)执行本文描述的特定过程或特定过程的特定部分,包括定义存储在RAM(1546)中的数据结构并且根据软件定义的过程修改这样的数据结构。此外或作为替代,计算机系统可以提供作为硬连线或以其他方式包含在电路(例如:加速器(1544))中的逻辑的结果的功能,其可以代替软件或与软件一起操作来执行本文描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以包括存储有所要执行的软件的电路(例如,集成电路(Integrated Circuit,简称IC))、包含用于执行的逻辑的电路或者这两者。本公开包含硬件和软件的任何合适的组合。

- [0278] 附录A:缩略语
- [0279] JEM:共同探索模式(joint exploration model)
- [0280] VVC:多功能视频编码(versatile video coding)
- [0281] BMS:基准集(benchmark set)
- [0282] MV:运动矢量(Motion Vector)
- [0283] HEVC:高效视频编码(High Efficiency Video Coding)
- [0284] SEI:补充增强信息(Supplementary Enhancement Information)
- [0285] VUI:视频可用性信息(Video Usability Information)
- [0286] GOP:图片组(Groups of Pictures)
- [0287] TU:变换单元(Transform Units)
- [0288] PU:预测单元(Prediction Units)
- [0289] CTU:编码树单元(Coding Tree Units)
- [0290] CTB:编码树块(Coding Tree Blocks)
- [0291] PB:预测块(Prediction Blocks)
- [0292] HRD:假设参考解码器(Hypothetical Reference Decoder)
- [0293] SNR:信噪比(Signal Noise Ratio)
- [0294] CPU:中央处理器(Central Processing Units)
- [0295] GPU:图形处理单元(Graphics Processing Units)
- [0296] CRT:阴极射线管(Cathode Ray Tube)
- [0297] LCD:液晶显示器(Liquid-Crystal Display)
- [0298] OLED:有机发光二极管(Organic Light-Emitting Diode)
- [0299] CD:光盘(Compact Disc)
- [0300] DVD:数字视频光盘(Digital Video Disc)
- [0301] ROM:只读存储器(Read-Only Memory)
- [0302] RAM:随机存取存储器(Random Access Memory)
- [0303] ASIC:专用集成电路(Application-Specific Integrated Circuit)
- [0304] PLD:可编程逻辑器件(Programmable Logic Device)
- [0305] LAN:局域网(Local Area Network)

- [0306] GSM:全球移动通信系统(Global System for Mobile communications)
- [0307] LTE:长期演进(Long-Term Evolution)
- [0308] CANBus:控制器局域网总线(Controller Area Network Bus)
- [0309] USB:通用串行总线(Universal Serial Bus)
- [0310] PCI:外围组件互连(Peripheral Component Interconnect)
- [0311] FPGA:现场可编程门区域(Field Programmable Gate Areas)
- [0312] SSD:固态硬盘(solid-state drive)
- [0313] IC:集成电路(Integrated Circuit)
- [0314] CU:编码单元(Coding Unit)
- [0315] NIC:神经图像压缩(Neural Image Compression)
- [0316] R-D:率失真(Rate-Distortion)
- [0317] 虽然本公开已经描述了几个示例性实施例,但是存在落入本公开范围内的变更、置换和各种替代等同物。因此,应当理解,本领域技术人员将能够设计出许多系统和方法,尽管在此没有明确示出或描述,但是这些系统和方法体现了本公开的原理,并且因此在本公开的精神和范围内。

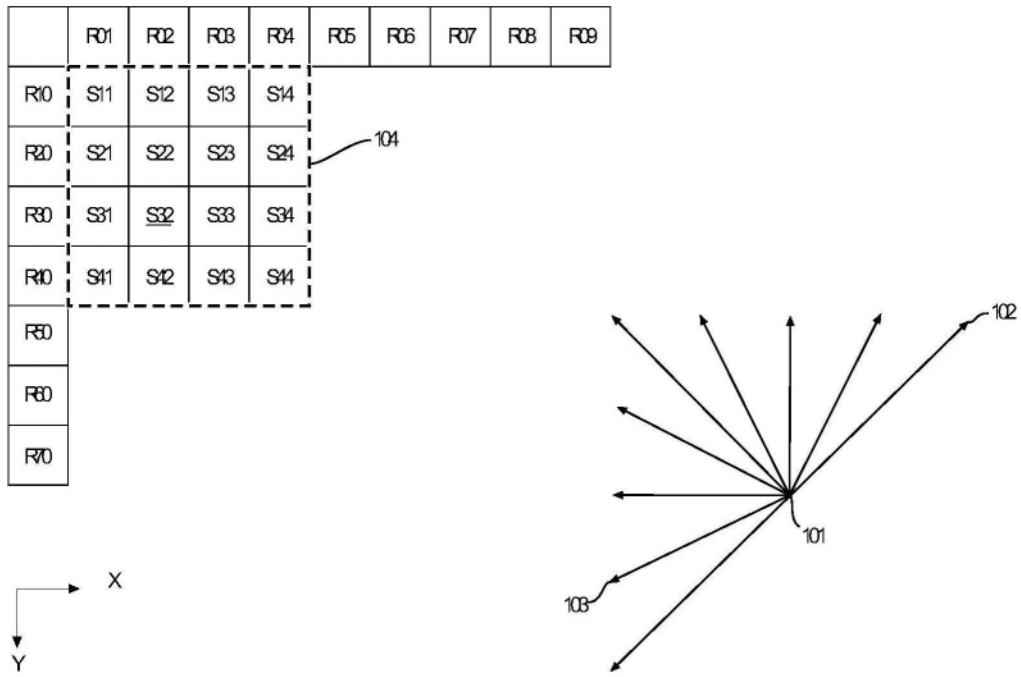


图1A

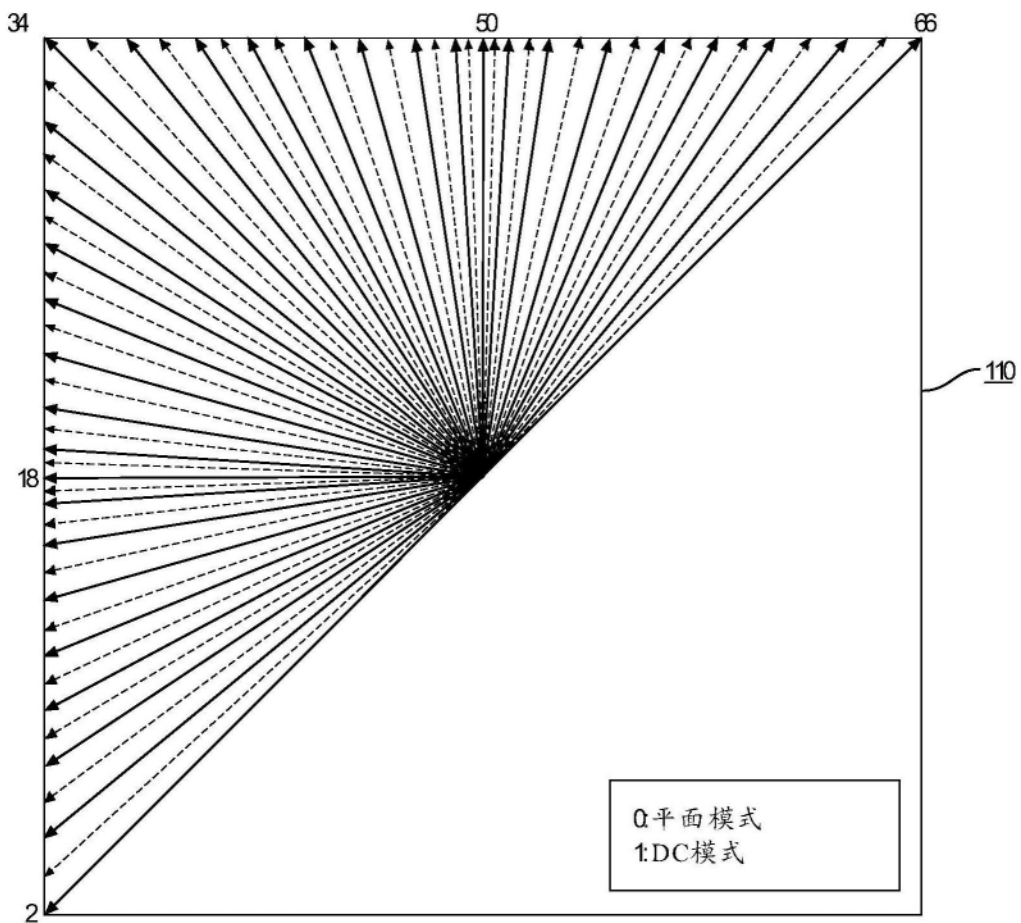


图1B

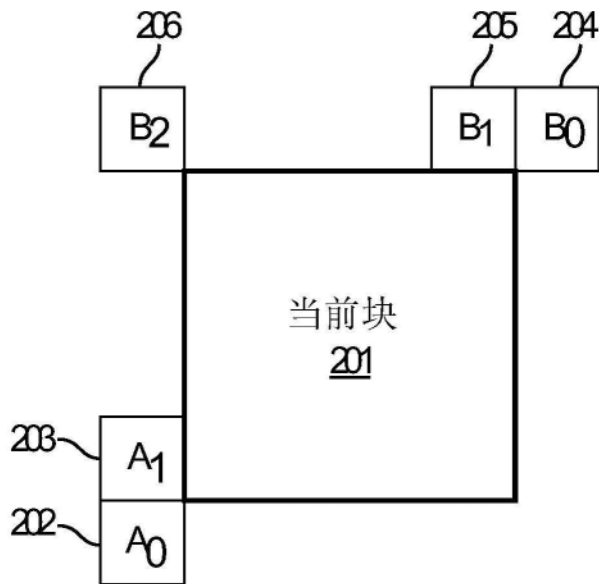


图2 (相关技术)

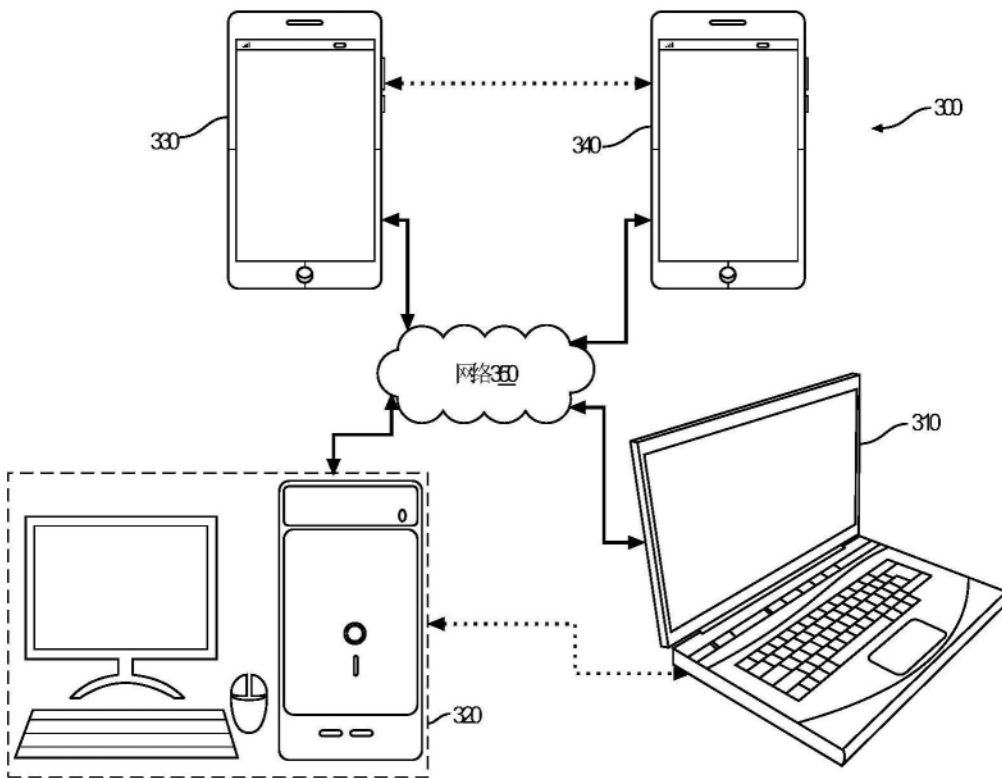


图3

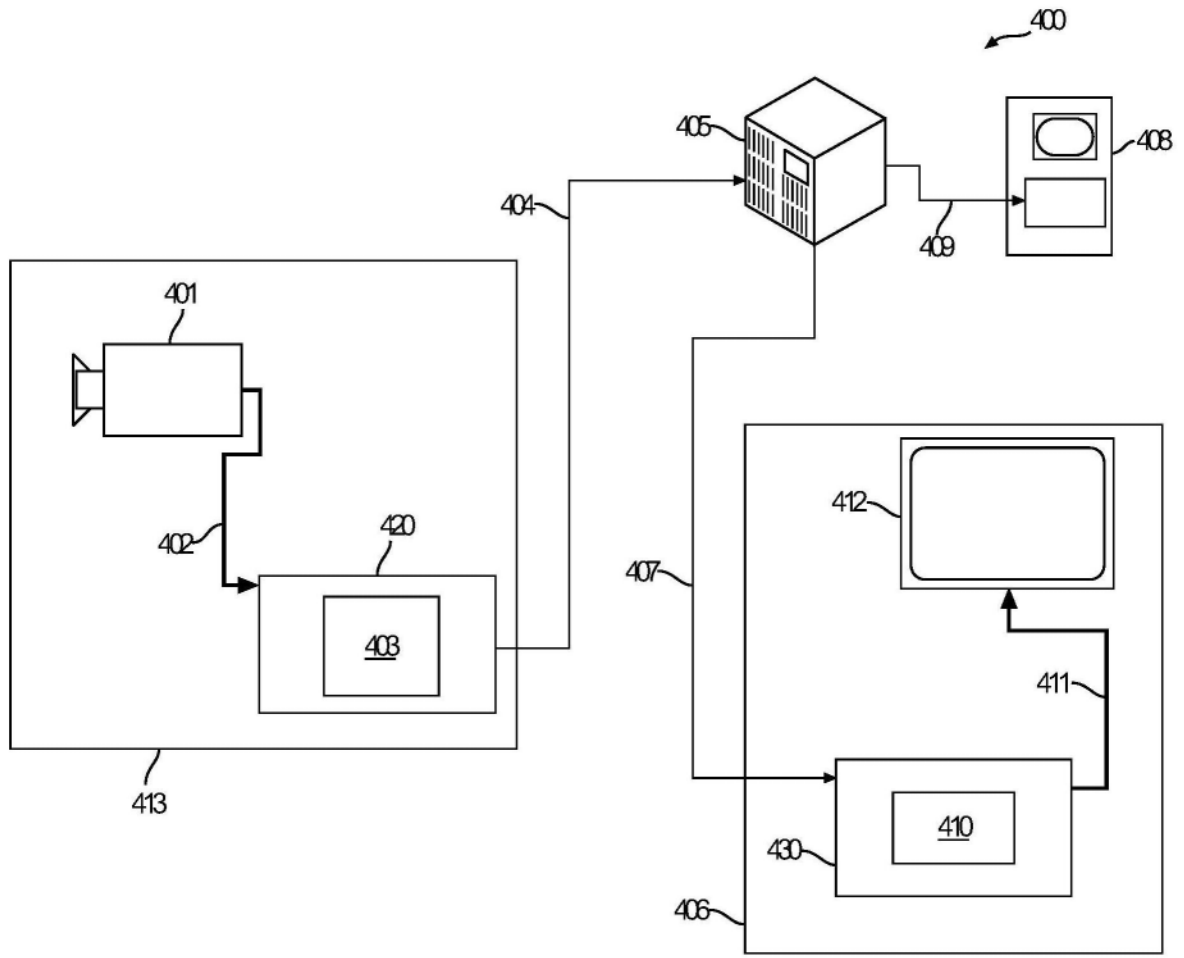


图4

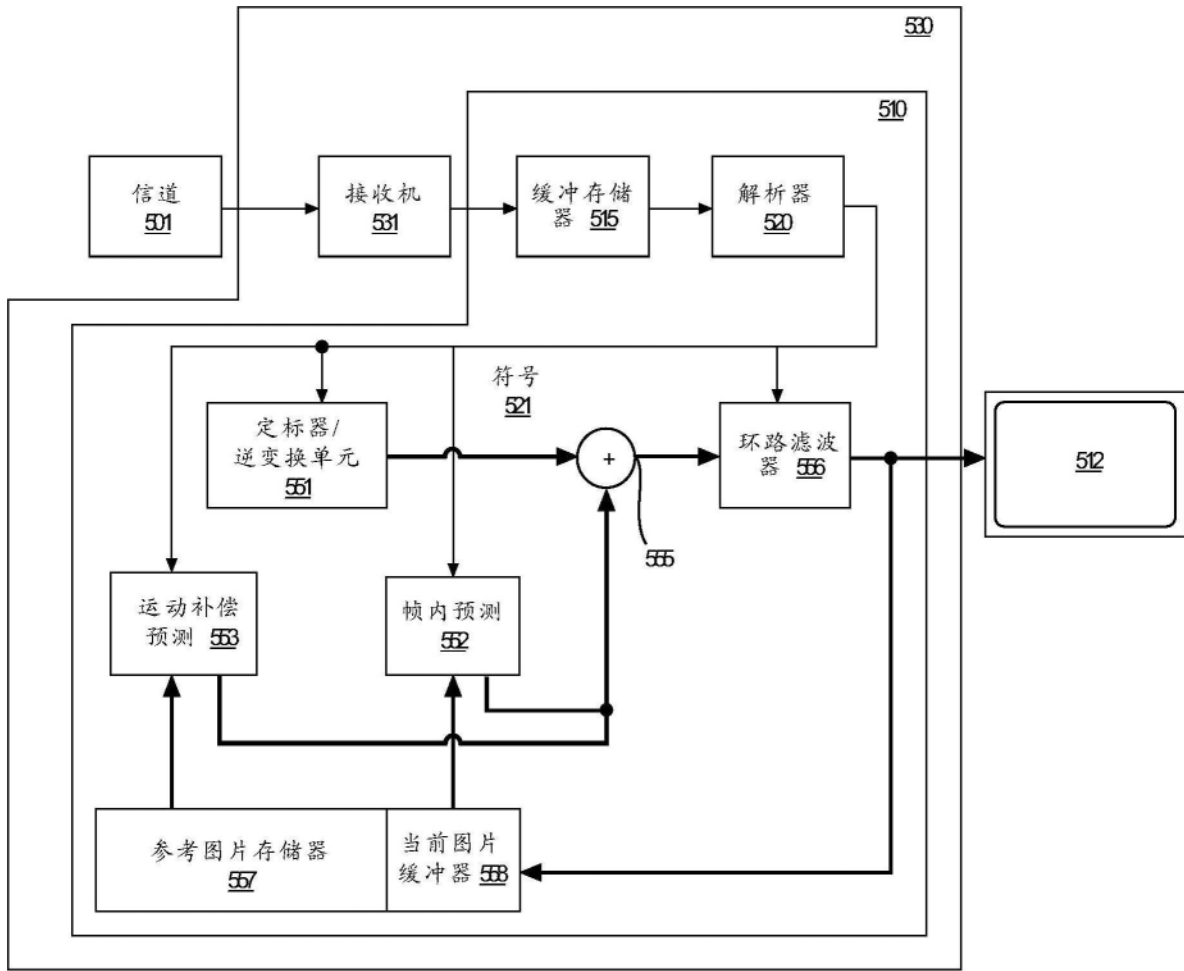


图5

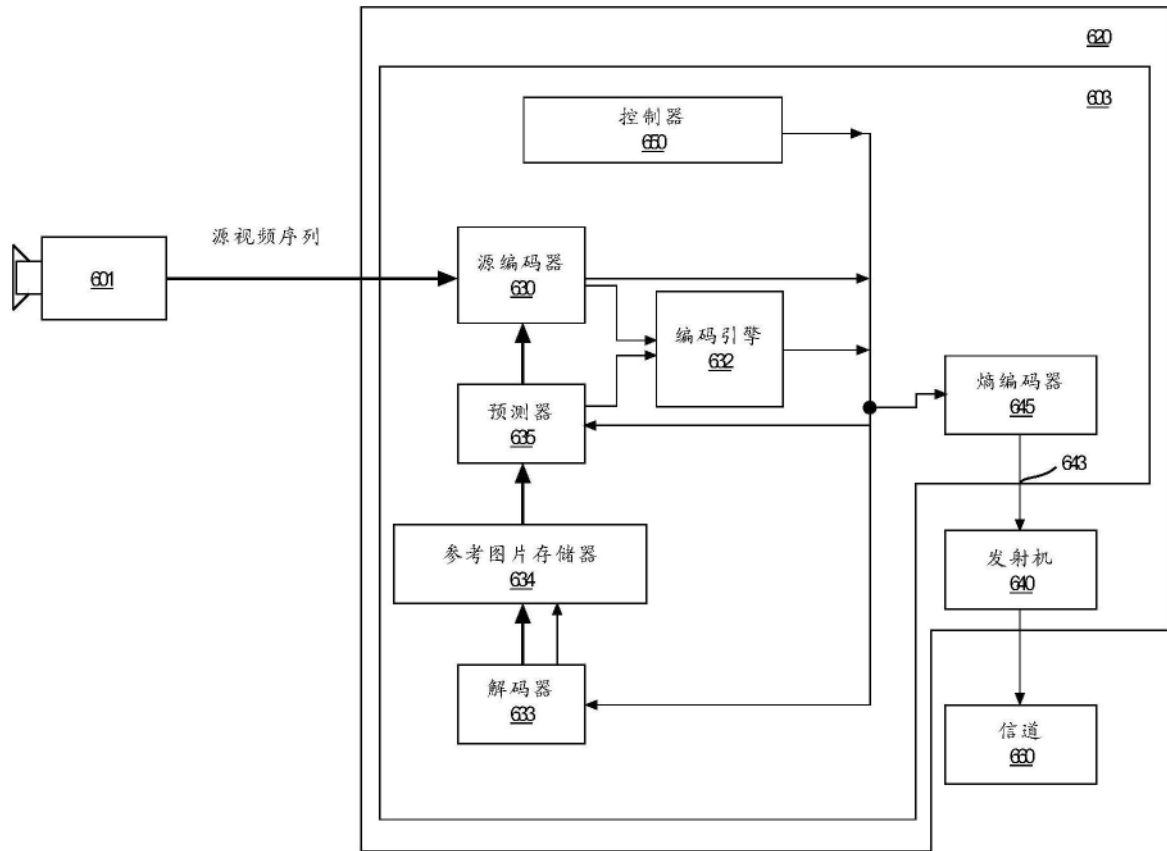


图6

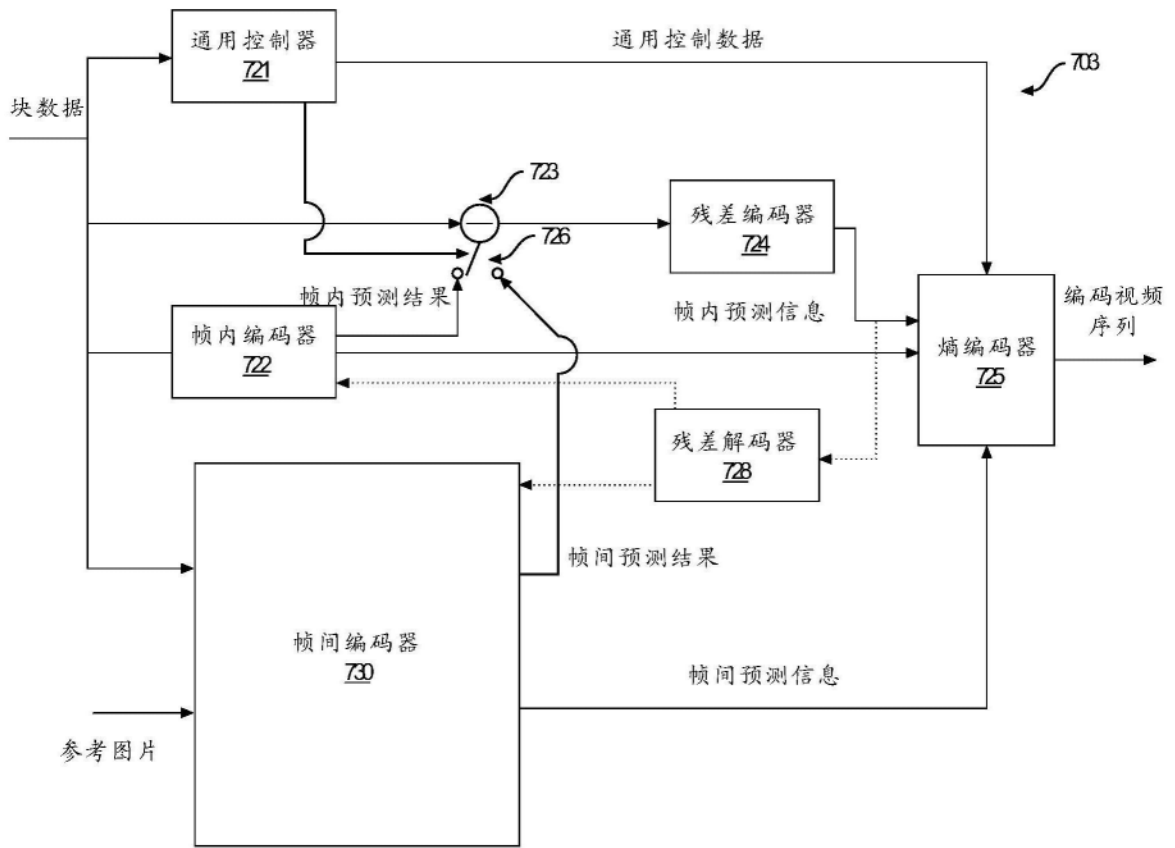


图7

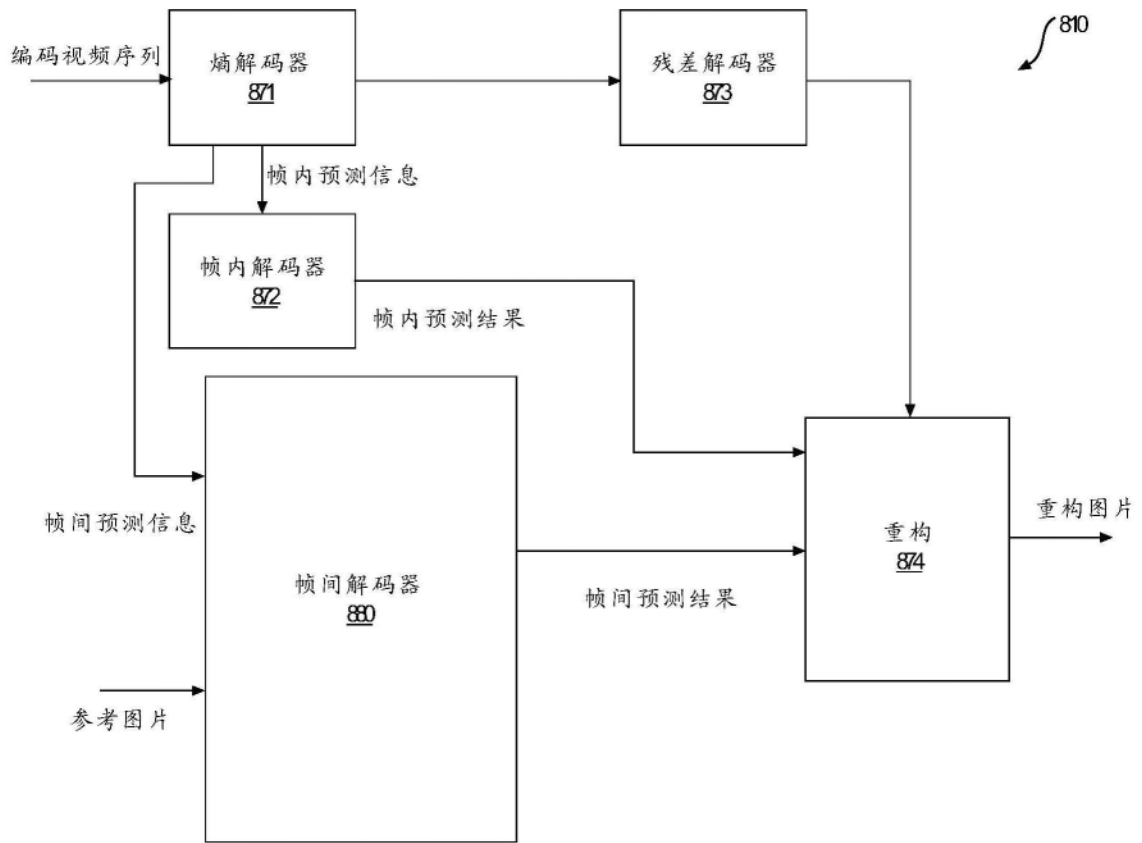


图8

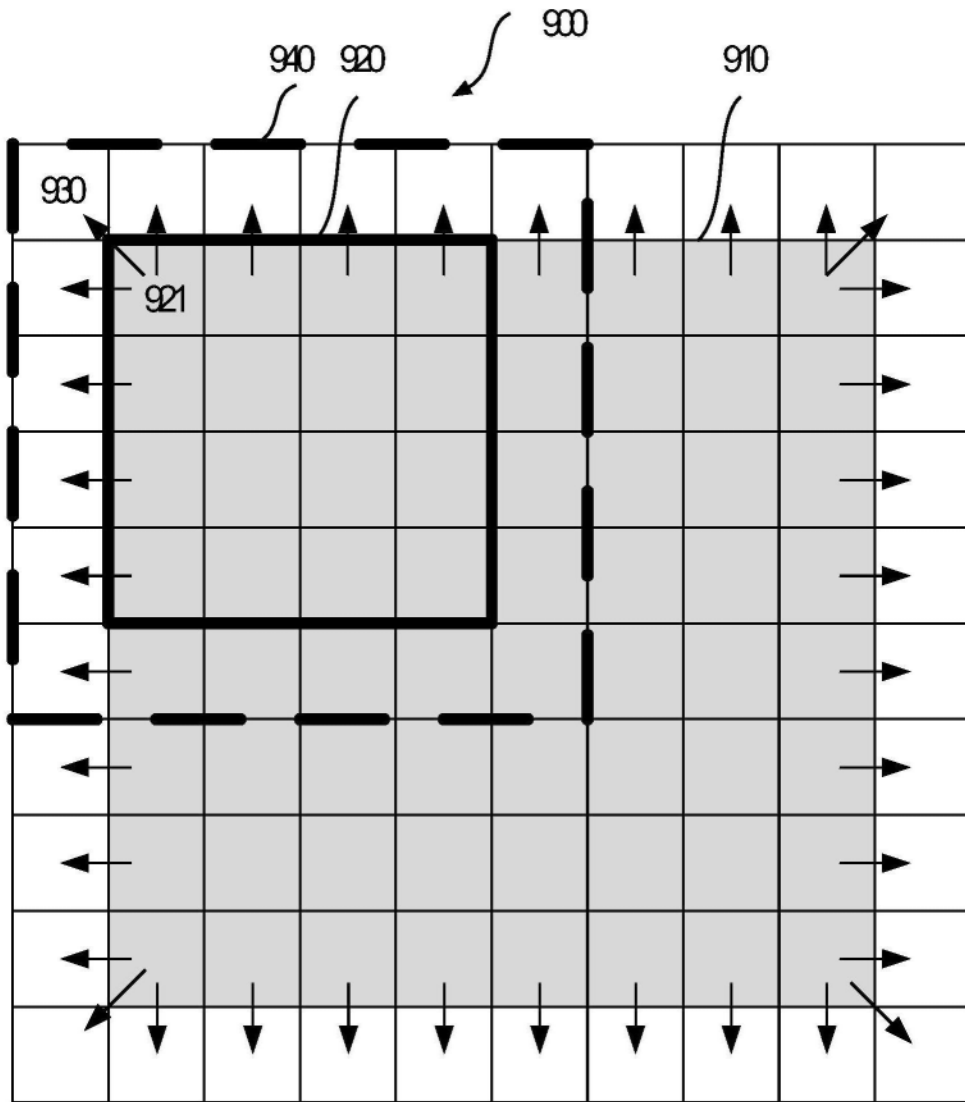


图9

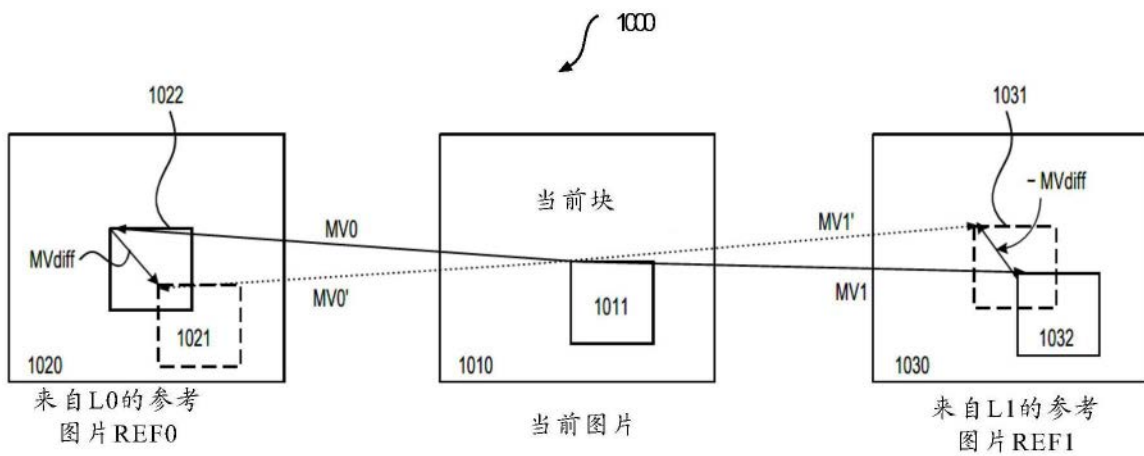


图10

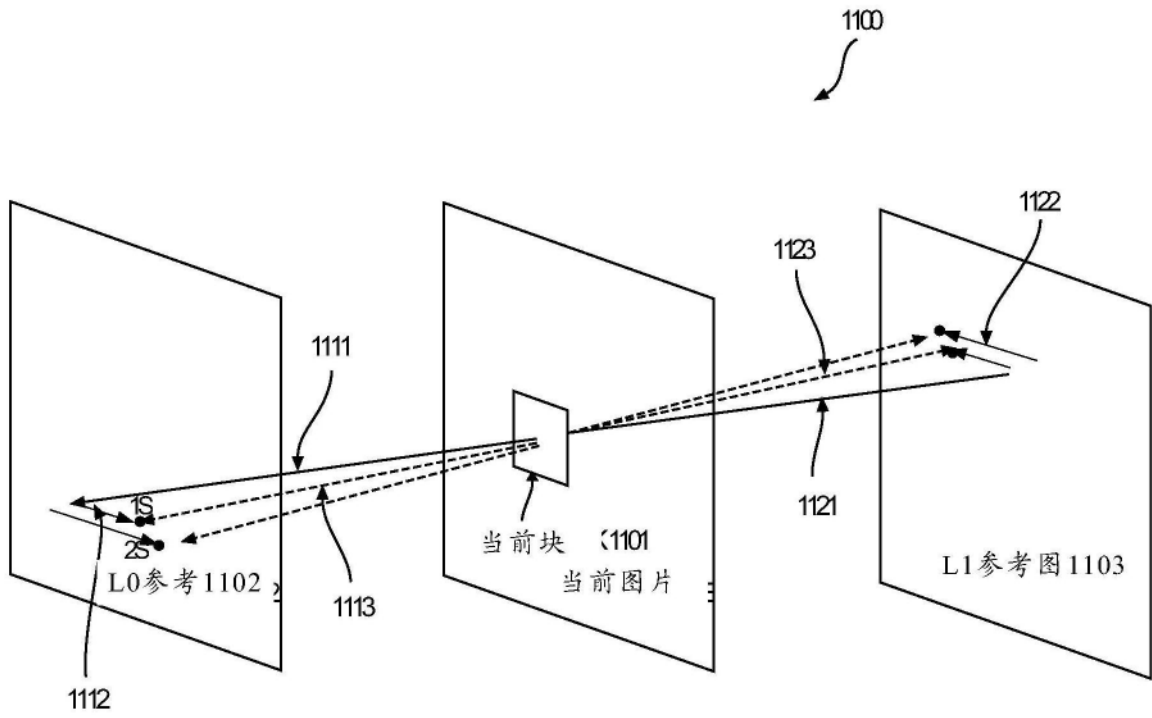


图11

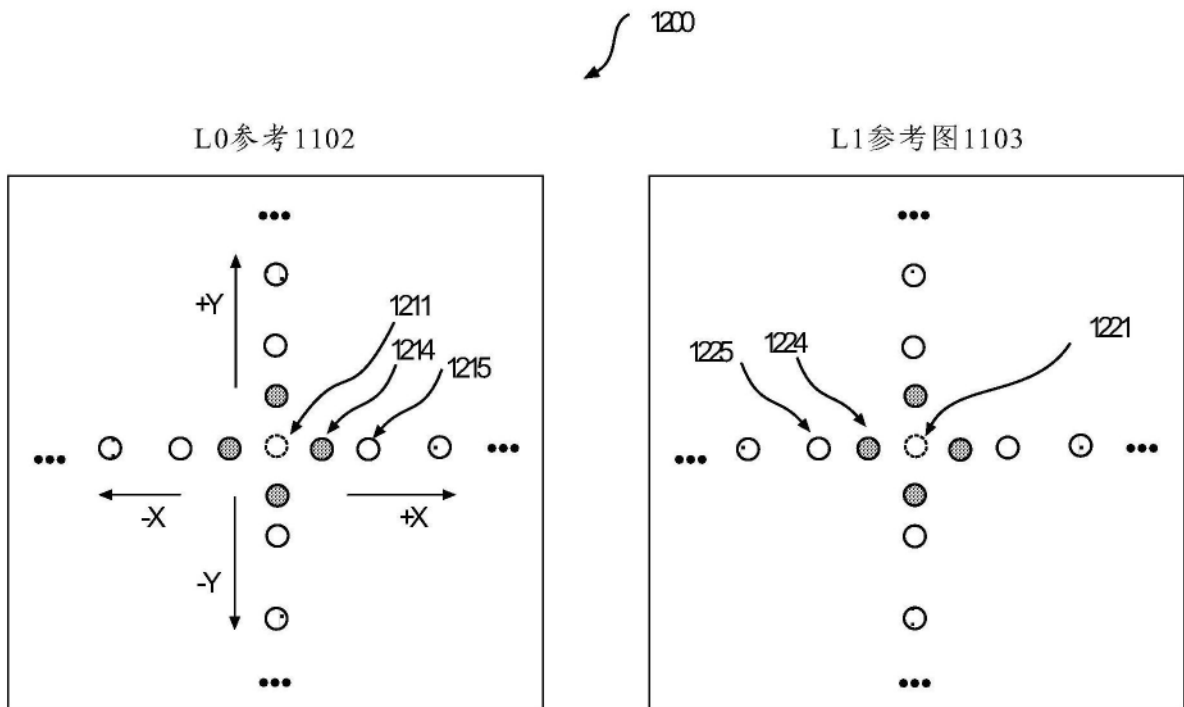


图12

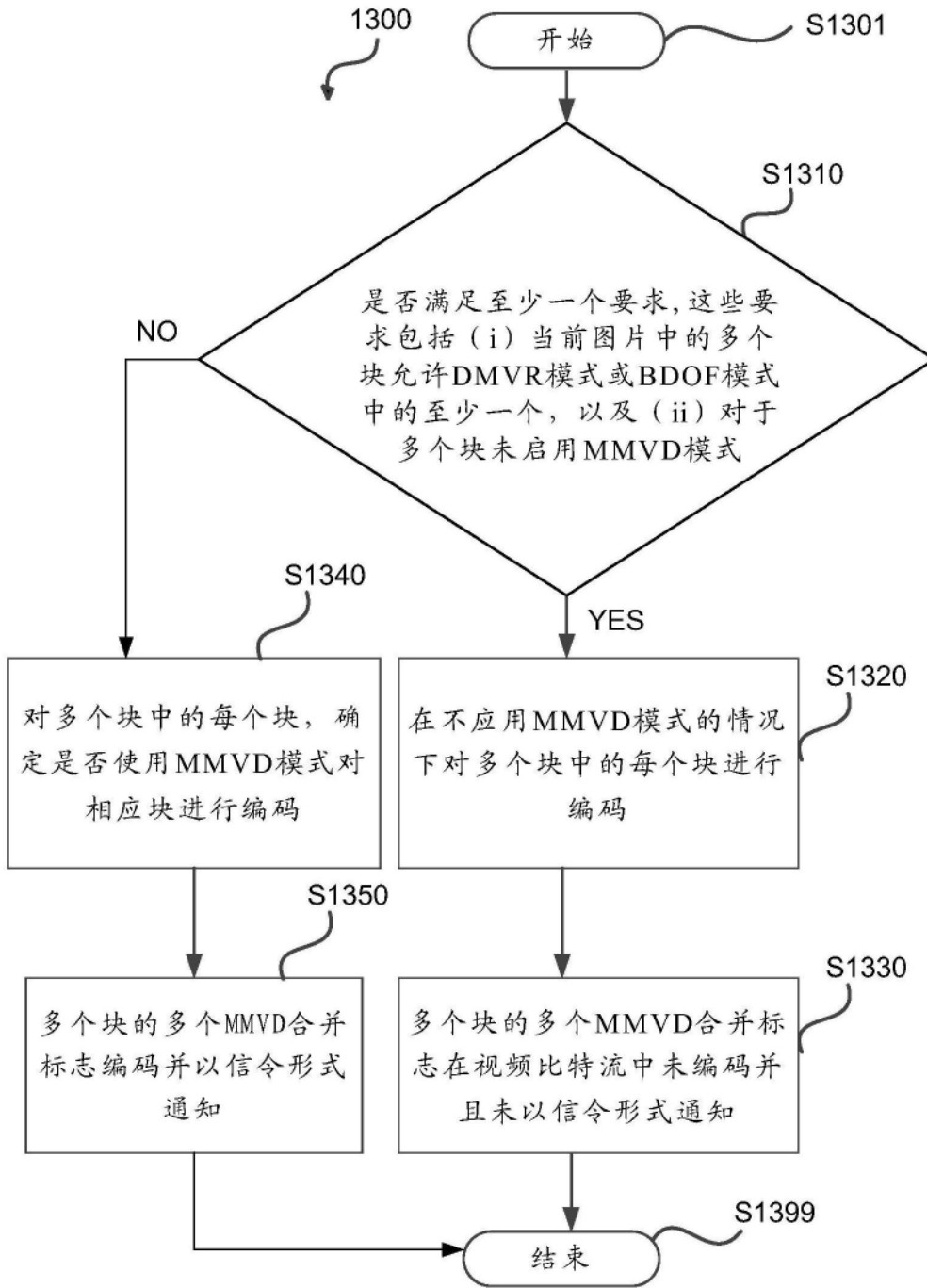


图13

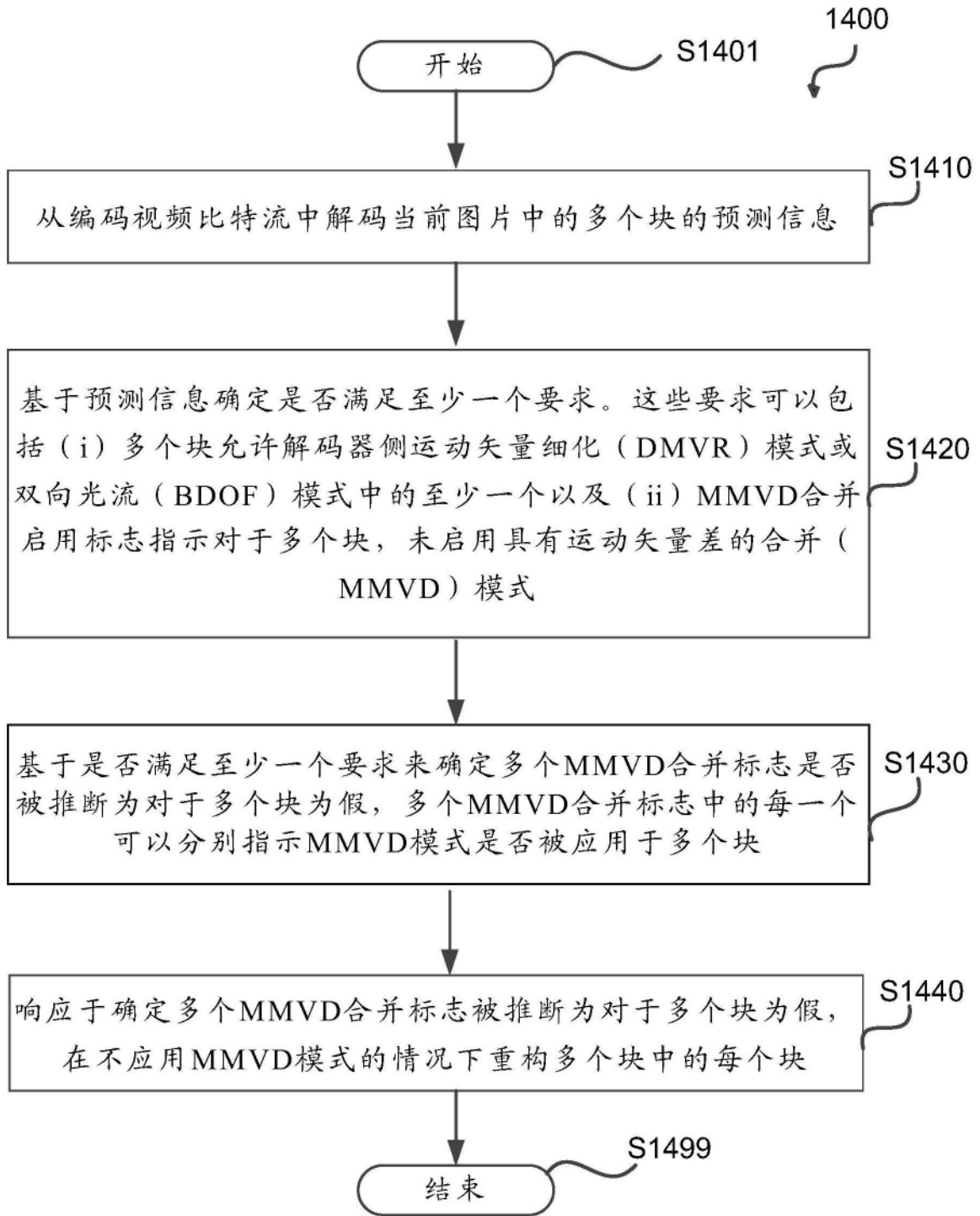


图14

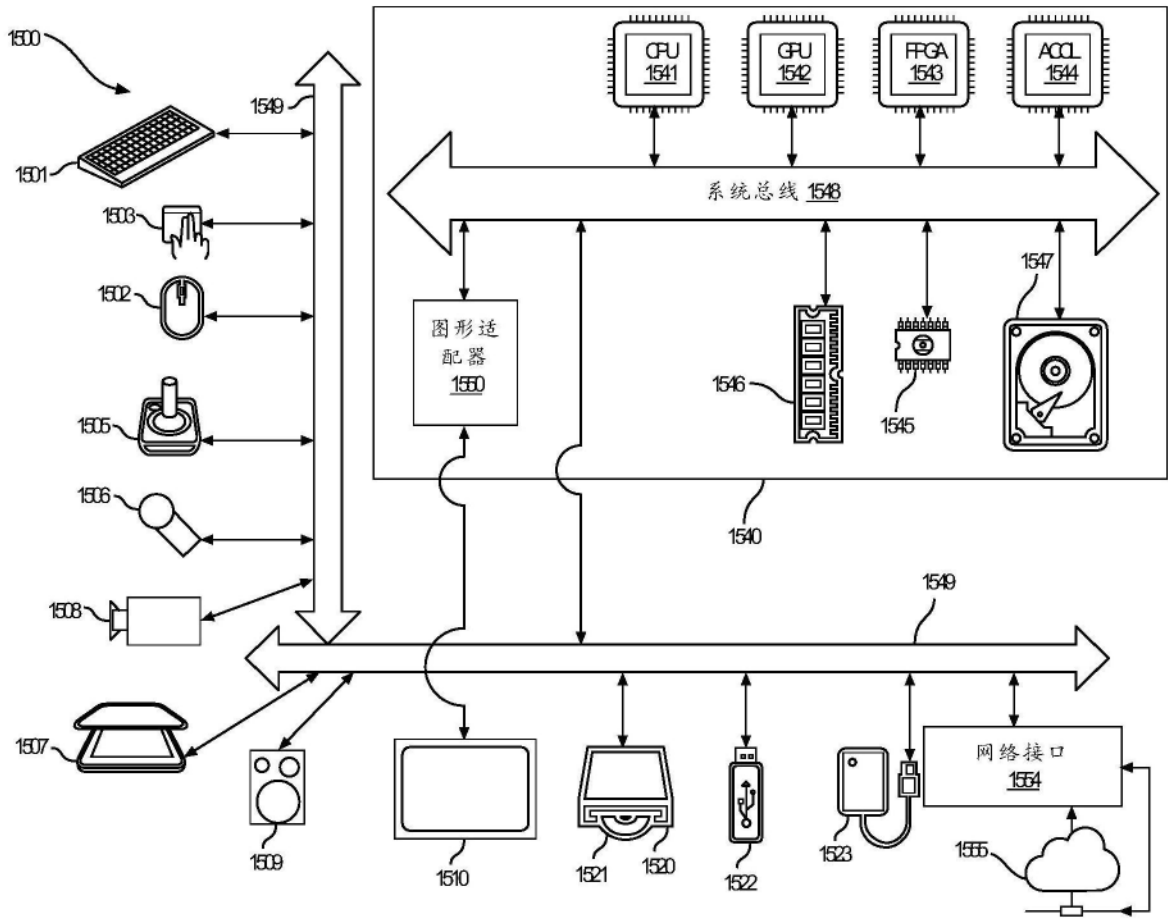


图15