



(12) 发明专利申请

(10) 申请公布号 CN 116830572 A

(43) 申请公布日 2023. 09. 29

(21) 申请号 202280012272.6

(22) 申请日 2022.06.03

(30) 优先权数据

63/302,518 2022.01.24 US

17/824,193 2022.05.25 US

(85) PCT国际申请进入国家阶段日

2023.07.28

(86) PCT国际申请的申请数据

PCT/US2022/032077 2022.06.03

(87) PCT国际申请的公布数据

W02023/140884 EN 2023.07.27

(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 赵亮 赵欣 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 李华 王琦

(51) Int.Cl.

H04N 19/10 (2006.01)

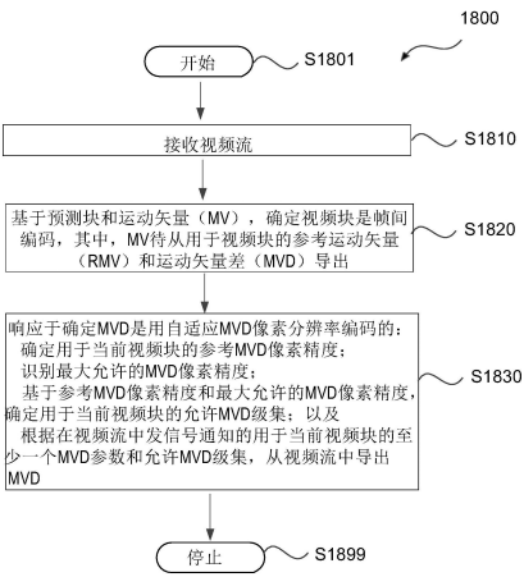
权利要求书2页 说明书42页 附图17页

(54) 发明名称

调整运动矢量差的自适应分辨率的方案

(57) 摘要

本公开总体上涉及视频编解码,更具体地涉及在针对运动矢量差实施自适应分辨率时提供用于设置允许运动矢量差值的方案的方法和系统。公开了一种处理视频流的当前视频块的示例方法。该方法可以包括接收视频流;基于预测块和运动矢量(MV),确定当前视频块是帧间编码,其中,MV待从用于当前视频块的参考运动矢量(RMV)和运动矢量差(MVD)导出。该方法进一步包括响应于确定MVD是用自适应MVD像素分辨率编码的:确定用于当前视频块的参考MVD像素精度;识别最大允许的MVD像素精度;基于参考MVD像素精度和最大允许的MVD像素精度,确定用于当前视频块的允许MVD级集;以及根据在视频流中发信号通知的用于当前视频块的至少一个MVD参数和允许MVD级集,从视频流中导出MVD。



1. 一种处理视频流的当前视频块的方法,其特征在于,包括:
接收所述视频流;
基于预测块和运动矢量(MV),确定所述当前视频块是帧间编码,其中,所述MV待从用于所述当前视频块的参考运动矢量(RMV)和运动矢量差(MVD)导出;以及
响应于确定所述MVD是用自适应MVD像素分辨率编码的:
确定用于所述当前视频块的参考MVD像素精度;
识别最大允许的MVD像素精度;
基于所述参考MVD像素精度和所述最大允许的MVD像素精度,确定用于所述当前视频块的允许MVD级集;以及
根据在所述视频流中发信号通知的用于所述当前视频块的至少一个MVD参数和所述允许MVD级集,从所述视频流中导出所述MVD。
2. 根据权利要求1所述的方法,其特征在于,所述用于所述当前视频块的参考MVD像素精度是在序列级、图片级、帧级、超块级或编码块级指定/发信号通知/导出。
3. 根据权利要求2所述的方法,其特征在于,用于所述当前视频块的所述参考MVD像素精度取决于与所述当前视频块的所述MVD相关联的MVD类别。
4. 根据权利要求2所述的方法,其特征在于,用于所述当前视频块的所述参考MVD像素精度取决于所述当前视频块的所述MVD的MVD量值。
5. 根据权利要求2所述的方法,其特征在于,所述最大允许的MVD像素精度是预定义的。
6. 根据权利要求1至5中任一项所述的方法,其特征在于,进一步包括:从预定义MVD类别集中确定当前MVD类别,其中,基于所述参考MVD像素精度和所述最大允许的MVD像素精度,确定用于所述MVD的所述允许MVD级集,包括:
从基于所述参考MVD像素精度和所述当前MVD类别确定的参考MVD级集中,排除与等于或高于所述最大允许的MVD像素精度的MVD像素精度相关联的MVD级,以确定用于所述当前视频块的所述允许MVD级集。
7. 根据权利要求6所述的方法,其特征在于,所述最大允许的MVD像素精度是1/4像素。
8. 根据权利要求1至5中任一项所述的方法,其特征在于,将与1/8像素或更高精度相关联的MVD级,排除在用于所述当前视频块的所述允许MVD级集之外。
9. 根据权利要求1至5中任一项所述的方法,其特征在于,进一步包括:从预定义MVD类别集中确定当前MVD类别,其中:
当所述当前MVD类别等于或低于阈值MVD类别时,无论所述参考MVD精度如何,与分数MVD精度相关联的MVD级是包括在所述允许MVD级集中。
10. 根据权利要求9所述的方法,其特征在于,所述阈值MVD类别是所述预定义MVD类别集中的最低MVD类别。
11. 根据权利要求1至5中任一项所述的方法,其特征在于,进一步包括:确定所述MVD的量值,其中,仅当所述MVD的量值等于或低于阈值MVD量值时,才允许在所述允许MVD级集中使用与高于阈值MVD精度的MVD精度相关联的MVD级。
12. 根据权利要求11所述的方法,其特征在于,所述阈值MVD量值为2像素或更小。
13. 根据权利要求12所述的方法,其特征在于,所述阈值MVD精度为1像素。
14. 根据权利要求11所述的方法,其特征在于,仅当所述MVD的量值等于或低于1/2像素

时,才允许使用与1/4像素或更高的MVD精度相关联的MVD级。

15.根据权利要求1至5中任一项所述的方法,其特征在于,所述最大允许的MVD像素精度不大于所述参考MVD像素精度。

16.一种处理视频流的当前视频块的方法,其特征在于,包括:

接收所述视频流;

确定所述当前视频块是帧间编码且与多个参考帧相关联;以及

基于所述视频流中的信令,确定自适应运动矢量差(MVD)像素分辨率是否应用于所述多个参考帧中的至少一个。

17.根据权利要求16所述的方法,其特征在于,所述信令包括单个比特标志,以指示自适应MVD像素分辨率是应用于所有所述多个参考帧,还是不应用于所述多个参考帧中的任何一个。

18.根据权利要求16所述的方法,其特征在于,所述信令包括分开的标志,每个标志对应于所述多个参考帧中的一个参考帧,以指示是否应用自适应MVD像素分辨率。

19.根据权利要求16所述的方法,其特征在于,对于所述多个参考帧中的每一个参考帧,所述信令包括:

隐式指示,当对应于所述多个参考帧中的每一个的MVD为零时,用于指示不应用自适应MVD像素分辨率;以及

单个比特标志,当对应于所述多个参考帧中的每一个的MVD为非零时,用于指示是否应用自适应MVD像素分辨率。

20.一种处理视频流的当前视频块的方法,其特征在于,包括:

接收所述视频流;

基于预测块和运动矢量(MV),确定所述当前视频块是帧间编码,其中,所述MV待从用于所述当前视频块的参考运动矢量(RMV)和运动矢量差(MVD)导出;

从预定义MVD类别集中确定所述MVD的当前MVD类别;

基于所述当前MVD类别,导出用于对所述视频流中的至少一个显式信令进行熵解码的至少一个上下文,所述至少一个显式信令包括在所述视频流中,以指定用于所述MVD的至少一个分量的MVD像素分辨率;以及

使用所述至少一个上下文,对所述视频流中的所述至少一个显式信令进行熵解码,以确定所述用于所述MVD的至少一个分量的MVD像素分辨率。

21.根据权利要求20所述的方法,其特征在于,所述MVD的至少一个分量包括所述MVD的水平分量和垂直分量,所述至少一个上下文包括两个单独的上下文,每个上下文与所述MVD的水平分量和垂直分量中的一者相关联,所述水平分量和所述垂直分量与单独的MVD像素分辨率相关联。

22.一种视频处理设备,其特征在于,包括处理器和用于存储计算机指令的存储器,当执行所述计算机指令时,所述处理器被配置为执行权利要求1至5和16至21中任一项所述的方法。

调整运动矢量差的自适应分辨率的方案

[0001] 引用并入

[0002] 本申请基于并要求于2022年5月25日提交的申请号为17/824,193、发明名称为“调整运动矢量差的自适应分辨率的方案(Schemes for Adjusting Adaptive Resolution for Motion Vector Difference)”的美国非临时专利申请的优先权,该美国非临时专利申请基于并要求于2022年1月24日提交的申请号为63/302,518、发明名称为“自适应MVD分辨率的进一步改进(Further Improvement for Adaptive MVD Resolution)”的美国临时专利申请优先权。这些在先申请通过引用整体并入本申请中。

技术领域

[0003] 本公开总体上涉及视频编解码,更具体地涉及在针对运动矢量差实施自适应分辨率时提供用于设置允许运动矢量差值的方案的方法和系统。

背景技术

[0004] 本文中提供的背景技术描述是为了大体上呈现本申请的上下文。在此背景技术部分描述的程度,当前署名的发明人的工作,以及在本申请提交时可能不具有作为现有技术的资格的描述的各方面,既不明确认为也不隐含认为是本申请的现有技术。

[0005] 视频编码和解码可以使用具有运动补偿的图片间预测,未压缩的数字视频可包括一系列的图片,每个图片具有一定的空间维度,例如,具有 1920×1080 的亮度样本和相关的全色度样本或子采样的色度样本。所述一系列的图片可以具有固定的或可变的图片速率(可替换地称为帧率),例如,每秒60个图片或每秒60个帧。未压缩的视频对于流式传输或数据处理有着特定的比特率要求。例如,具有 1920×1080 的像素分辨率、每秒60帧的帧率以及每个颜色通道每个像素8比特的4:2:0的色度下采样的视频需要接近1.5Gbit/s的带宽。这样的视频一小时需要超过600GB的存储空间。

[0006] 视频编码和解码的一个目的可以是通过压缩来减少未压缩的输入视频信号中的冗余。压缩可有助于降低上述带宽和/或存储空间的要求,在一些情况下,可降低两个或更多的数量级。无损压缩和有损压缩以及其组合均可以用于视频编码和解码。无损压缩是指可以由压缩的原始信号通过解码过程重建原始信号的精确副本的技术。有损压缩是指原始视频信号在编码过程中未完全保持、在解码过程中未完全恢复的编码/解码过程。当使用有损压缩时,重建的信号可能与原始信号不完全一致,但是使原始信号与重建的信号之间的失真小得足以致使重建的信号可以用于预期应用,虽然有一些信息损失。对于视频,有损压缩广泛应用于许多应用中。有损压缩可容许的失真量取决于应用。例如,与电影或电视广播应用的用户相比,某些视频流式传输应用的消费者用户可以容忍较高的失真。特定的编码算法可实现的压缩比可以选择或调节,以反应各种失真容差:可容许的失真越高,通常允许使用可产生较高损失和较高压缩比的编码算法。

[0007] 视频编码器和解码器可以使用几大类技术和步骤,包括例如运动补偿、傅里叶变换、量化和熵编码。

[0008] 视频编解码技术可以包括帧内编码技术。在帧内编码技术中,样本值的表示不参考先前已重建的参考图片中的样本或其它数据。在一些视频编解码技术中,将图片从空间上划分为样本块。当所有的样本块都是通过帧内模式进行编码时,该图片可以称作帧内图片。帧内图片以及他们的衍生图片,例如,独立解码器刷新的图片,可以用于对解码器的状态进行重置,因此可以用作已编码视频码流和视频会话中的第一幅图片,或者作为一副静止图片。然后可以将经过帧内预测的块的样本变换到频域,并且可以在进行熵编码之前,对如此生成的变换系数进行量化。帧内预测表示一种将样本值在预变换域最小化的技术。在一些情况下,变换后的DC值越小以及AC系数越小,给定量化步长大小的情况下,表示熵编码后的块所需的比特数目越少。

[0009] 传统的帧内编码技术,例如,已知的MPEG-2编码技术并不使用帧内预测。然而,一些较新的视频压缩技术包括尝试基于例如邻近的样本数据和/或元数据对块进行编码/解码的技术,所述邻近的样本数据和/或元数据是在对与正在帧内编码或解码中的数据块空间上相邻且解码顺序上在先的数据块进行编码和/或解码的过程中获得的。因此,这种技术称为“帧内预测”技术。请注意,至少在一些情况下,帧内预测仅使用正在重建的当前图片中的参考数据,而不使用其他参考图片中的参考数据。

[0010] 帧内预测可以有很多种不同的形式。当给定的视频编码技术中有一个以上的这种技术可用时,该使用中的技术可以称作帧内预测模式。一个或多个帧内预测模式可以提供在特定的编解码中。在某些情况下,一些模式具有子模式和/或与各种参数相关联,视频块的模式/子模式信息和帧内编码参数可以单独编码或者可以集体包含在模式码字中。某个给定模式/子模式和/或参数的组合使用何种码字,会通过帧内预测影响到编码效率增益,将码字转译为码流所用的熵编码技术同样对其也会产生影响。

[0011] H.264标准引入了某个模式的帧内预测,H.265标准对其进行了改进,在较新的编码技术中,例如,联合探索模型(JEM)、通用视频编码(VVC)、基准集合(BMS)等等,对其进一步进行了改进。通常,对于帧内预测,预测子(predictor)块可以使用已经可用的相邻的样本值形成。例如,可以将沿着特定方向和/或行的特定集合的相邻样本的可用值复制到预测子块中。所使用的方向的参考可以编码到码流中,或者其本身可以预测。

[0012] 参考图1A,在其右下方描绘的是H.265标准的33个可能的帧内预测子方向(predictor direction)(对应于H.265标准中规定的35个帧内模式的33个角度模式)中已知的一个具有9个预测子方向的子集。其中,各箭头的汇聚点(101)表示正在预测中的样本。箭头表示使用相邻样本对101处的样本进行预测的方向。例如,箭头(102)表示根据与水平轴成45度角度的右上角的一个或多个相邻样本,对样本(101)进行预测。类似地,箭头(103)表示根据与水平方向成22.5度角度的左下角的一个或多个相邻样本,对样本(101)进行预测。

[0013] 仍参考图1A所示,图1A的左上方描绘的是一个具有 4×4 个样本的正方形块(104)(用加粗的虚线表示)。正方形块(104)包括16个样本,每个样本标记有“S”、其在Y维度上的位置(例如,行索引)以及其在X维度上的位置(例如,列索引)。例如,样本S21是在Y维度上的第二个(从上往下数)、X维度上的第一个(从左往右数)样本。类似地,样本S44是在块(104)中X维度和Y维度上都是第四的样本。因为该块的大小是 4×4 个样本,所以S44是在其右下角。图1A进一步示出了示例参考样本,参考样本遵循类似的编号方法。参考样本标记有R、其

相对于所述块(104)的Y位置(例如,行索引)和X位置(例如,列索引)。在H.264标准和H.265标准中,使用与重建中的块紧邻的预测样本。

[0014] 块104的帧内图片预测可以开始于根据信号表示的预测方向,通过复制相邻样本的参考样本值。例如,假设已编码视频码流中包含有信令,对于所述块104,该信令表示了箭头(102)的预测方向,即,根据与水平方向成45度角度的右上角的一个或多个参考样本对所述块中的样本进行预测。在这种情况下,样本S41、S32、S23、S14是根据同一个参考样本R05进行预测的。样本S44是根据参考样本R08进行预测的。

[0015] 在某些情况下,可以例如通过插值,组合多个参考样本的值,以便计算一个参考样本;特别是当方向不能被45度整除时。

[0016] 随着视频编码技术的不断发展,可能的方向的数目也在增加。在H.264标准(2003年)中,例如,9个不同的方向可用于帧内预测。在H.265标准(2013年)中,增加到33个方向。到本申请发明时,JEM/VVC/BMS可以支持多达65个方向。目前已经进行了一些实验研究帮助识别最合适的帧内预测方向,一些熵编码技术用很少的比特数对这些最合适的方向进行编码,对于方向,接受一定的比特代价。此外,有时候这些方向本身是可以根据相邻的已解码的块在帧内预测时所使用的相邻方向进行预测的。

[0017] 图1B示出了一个描绘根据JEM的65个帧内预测方向的示意图(180),用于说明各种编码技术中预测方向的数目随着时间的变化而增加。

[0018] 在已编码视频码流中将表示帧内预测方向的比特映射到预测方向随视频编码技术的不同而不同;例如,变化范围可以是简单直接地将帧内预测模式的预测方向映射到码字,到涉及最有可能的模式和类似技术的复杂的自适应方案。然而,在所有这些情况下,依据统计,相对于其它方向而言,某些用于帧内预测的方向比较不可能出现在视频内容中。由于视频压缩的目的是减少冗余,因此,在性能较好的视频编码技术中,与比较可能的方向相比,这些比较不可能出现的方向会用较多的比特来表示。

[0019] 帧内预测或帧间预测可以基于运动补偿。在运动补偿中,来自先前已重建图片或其一部分(参考图片)的样本数据块,在由运动矢量(此后称为MV)指示的方向上空间移位之后,可以用于预测新重建的图片或图片部分(例如,块)。在一些情况下,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,或三个维度,第三维度指示使用中的参考图片(即时间维度)。

[0020] 在一些视频压缩技术中,适用于某一样本数据区域的当前MV,可根据其它MV进行预测,例如,根据与正在重建的区域空间相邻、且解码顺序在所述MV之前的另一样本数据区域相关的其它MV进行预测。这样做,通过消除相关联MV冗余,可充分减少对所述MV编码所需的总数据量,从而提高压缩效率。举例来说,MV预测可有效地运作,因为当对源自相机的输入视频信号(称为天然视频)进行编码时,存在如下的统计可能性:比单个MV适用的区域大的区域,在视频序列中类似方向上移动,且因此,可在一些情况下使用从相邻区域的MV导出的类似运动矢量进行预测。这使得给定区域中的实际MV与根据周围MV预测的MV类似或相同。这样的MV,在熵编码之后,与在直接对MV编码而不是根据相邻MV预测的情况下所使用的比特数目相比,可以用较少的比特数表示。在一些情况下,MV预测可以从原始信号(即:样本流)导出的信号(即:MV)的无损压缩的实例。在其它情况下,MV预测本身可能是有损的,例如是由于在根据周围若干MV计算预测子时的取整误差导致的。

[0021] H.265/HEVC (ITU-T H.265建议书,“高效视频编解码 (High Efficiency Video Coding)”,2016年12月)中描述了各种MV预测机制。在H.265规定的多种MV预测机制中,下文描述的是称为“空间合并”的技术。

[0022] 具体地,参考图2所示,当前块(201)包括编码器在运动搜索过程中找到的、可与当前块大小相同的前一个块(已经空间移位)预测的样本。所述MV不是直接进行编码,而是可以使用与五个周围样本(记作A0、A1、B0、B1、B2(分别是202至206))中任何一个周围样本相关联的MV,从与一个或多个参考图片(例如,最近的(按解码顺序)参考图片)相关联的元数据推导出。在H.265中,MV预测可以使用与其相邻块使用的参考图片相同的参考图像的预测子。

发明内容

[0023] 本公开总体上涉及视频编码,特别是涉及基于是否应用了帧间预测中的运动矢量差的取决于量值的自适应分辨率,信令各种运动矢量或运动矢量差相关的语法的方法和系统。

[0024] 在示例实施方式中,公开了一种处理视频流的当前视频块的方法。所述方法包括:接收视频流;基于预测块和运动矢量(MV),确定当前视频块是帧间编码,其中,所述MV待从用于当前视频块的参考运动矢量(RMV)和运动矢量差(MVD)导出。所述方法进一步包括:响应于确定MVD是用自适应MVD像素分辨率编码的:确定用于当前视频块的参考MVD像素精度;识别最大允许的MVD像素精度;基于参考MVD像素精度和最大允许的MVD像素精度,确定用于当前视频块的允许MVD级集;以及根据在所述视频流中发信号通知的用于当前视频块的至少一个MVD参数和允许MVD级集,从视频流中导出MVD。

[0025] 在上述实施方式中,用于当前视频块的参考MVD像素精度是在序列级、图片级、帧级、超块级或编码块级指定/发信号通知/导出。

[0026] 在上述任一实施方式中,用于当前视频块的参考MVD像素精度取决于与当前视频块的MVD相关联的MVD类别。

[0027] 在上述任一实施方式中,用于当前视频块的参考MVD像素精度取决于当前视频块的MVD的MVD量值。在上述任一实施方式中,最大允许的MVD像素精度是预定义的。

[0028] 在上述任一实施方式中,所述方法可以进一步包括:从预定义MVD类别集中确定当前MVD类别。基于参考MVD像素精度和最大允许的MVD像素精度,确定用于MVD的允许MVD级集,可以包括:从基于参考MVD像素精度和当前MVD类别确定的参考MVD级集中,排除与等于或高于最大允许的MVD像素精度的MVD像素精度相关联的MVD级,以确定用于当前视频块的允许MVD级集。

[0029] 在上述任一实施方式中,最大允许的MVD像素精度是1/4像素。

[0030] 在上述任一实施方式中,将与1/8像素或更高精度相关联的MVD级,排除在用于当前视频块的允许MVD级集之外。

[0031] 在上述任一实施方式中,所述方法可以进一步包括:从预定义MVD类别集中确定当前MVD类别。当当前MVD类别等于或低于阈值MVD类别时,无论参考MVD精度如何,与分数MVD精度相关联的MVD级可以是包括在允许MVD级集中。

[0032] 在上述任一实施方式中,阈值MVD类别可以是预定义MVD类别集中的最低MVD类别。

[0033] 在上述任一实施方式中,所述方法可以进一步包括:确定MVD的量值,其中,仅当MVD的量值等于或低于阈值MVD量值时,才允许在允许MVD级集中使用与高于阈值MVD精度的MVD精度相关联的MVD级。

[0034] 在上述任一实施方式中,阈值MVD量值为2像素或更小。

[0035] 在上述任一实施方式中,阈值MVD精度为1像素。

[0036] 在上述任一实施方式中,仅当MVD的量值等于或低于1/2像素时,才允许使用与1/4像素或更高的MVD精度相关联的MVD级。在上述任一实施方式中,最大允许的MVD像素精度不大于参考MVD像素精度。

[0037] 在另一实施方式中,提供了一种用于处理视频流的当前视频块的方法。所述方法包括:接收视频流;确定当前视频块是帧间编码且与多个参考帧相关联;以及基于视频流中的信令,确定自适应运动矢量差(MVD)像素分辨率是否应用于多个参考帧中的至少一个。

[0038] 在上述实施方式中,所述信令可以包括单个比特标志,以指示自适应MVD像素分辨率是应用于所有多个参考帧,还是不应用于多个参考帧中的任何一个。

[0039] 在上述任一实施方式中,所述信令包括分开的标志,每个标志对应于多个参考帧中的一个参考帧,以指示是否应用自适应MVD像素分辨率。

[0040] 在上述任一实施方式中,对于多个参考帧中的每一个,所述信令包括:隐式指示,当对应于多个参考帧中的每一个的MVD为零时,用于指示不应用自适应MVD像素分辨率;以及单个比特标志,当对应于多个参考帧中的每一个的MVD为非零时,用于指示是否应用自适应MVD像素分辨率。

[0041] 在另一实施方式中,提供了一种用于处理视频流的当前视频块的方法。所述方法包括:接收视频流;基于预测块和运动矢量(MV),确定当前视频块是帧间编码,其中,所述MV待从用于当前视频块的参考运动矢量(RMV)和运动矢量差(MVD)导出;从预定义MVD类别集中确定MVD的当前MVD类别;基于当前MVD类别,导出用于对视频流中的至少一个显式信令进行熵解码的至少一个上下文,至少一个显式信令包括在视频流中,以指定用于MVD的至少一个分量的MVD像素分辨率;以及使用至少一个上下文,对所述视频流中的所述至少一个显式信令进行熵解码,以确定用于MVD的至少一个分量的MVD像素分辨率。

[0042] 在上述实施方式中,MVD的至少一个分量可以包括MVD的水平分量和垂直分量,至少一个上下文可以包括两个单独的上下文,每个上下文与MVD的水平分量和垂直分量中的一者相关联,水平分量和垂直分量与单独的MVD像素分辨率相关联。

[0043] 本公开的各个方面还提供一种视频编码或解码设备或装置,包括被配置为执行上述任一方法实施方式的电路。

[0044] 本公开的各个方面还提供了存储指令的非易失性计算机可读介质,所述指令当由用于视频解码和/或编码的计算机执行时,使计算机执行视频解码和/或编码的方法。

附图说明

[0045] 从以下详细描述和附图中,所公开的主题的其它特征、性质和各种优点将变得更加明显,其中:

[0046] 图1A示出了帧内预测方向模式的示例性子集的示意图;

[0047] 图1B示出了示例性帧内预测方向的图示;

- [0048] 图2示出了一个示例中用于运动矢量预测的当前块及其周围空间合并候选的示意图；
- [0049] 图3示出了根据示例实施例的通信系统(300)的简化框图的示意图；
- [0050] 图4示出了根据示例实施例的通信系统(400)的简化框图的示意图；
- [0051] 图5示出了根据示例实施例的视频解码器的简化框图的示意图；
- [0052] 图6示出了根据示例实施例的视频编码器的简化框图的示意图；
- [0053] 图7示出了根据另一示例实施例的视频编码器的框图；
- [0054] 图8示出了根据另一示例实施例的视频解码器的框图；
- [0055] 图9示出了根据本公开的示例实施例的编码块分区的方案；
- [0056] 图10示出了根据本公开的示例实施例的编码块分区的另一方案；
- [0057] 图11示出了根据本公开的示例实施例的编码块分区的另一方案；
- [0058] 图12示出了根据示例分区方案将基本块分区为编码块的示例；
- [0059] 图13示出了示例三叉(ternary)分区方案；
- [0060] 图14示出了示例四叉树二叉树编码块分区方案；
- [0061] 图15示出了根据本公开的示例实施例的用于将编码块分区为多个变换块和变换块的编码顺序的一方案；
- [0062] 图16示出了根据本公开的示例实施例的用于将编码块分区为多个变换块和变换块的编码顺序的另一方案；
- [0063] 图17示出了根据本公开的示例实施例的用于将编码块分区为多个变换块的另一方案；
- [0064] 图18示出了根据本公开的示例实施例的方法的流程图；
- [0065] 图19示出了根据本公开的示例实施例的方法的另一流程图；
- [0066] 图20示出了根据本公开的示例实施例的方法的另一流程图；
- [0067] 图21示出了根据本公开的示例实施例的计算机系统的示意图示。

具体实施方式

[0068] 在整个说明书和权利要求书中，术语可以具有在明确说明的含义之外的上下文暗示的或暗含的含义。本申请中使用的短语“在一个实施例中”或“在一些实施例中”不是必须指相同的实施例，本申请中使用的短语“在另一个实施例中”或“在其它实施例中”不是必须指不同的实施例。类似地，本申请中使用的短语“在一个实施方式中”或“在一些实施方式中”不是必须指同一个实施方式，本申请中使用的短语“在另一个实施方式中”或“在其它实施方式中”不是必须指不同的实施方式。例如，旨在所要求保护的主体包括示例实施例/实施方式全部或部分的组合。

[0069] 总体而言，术语至少可以部分根据其在上下文的用法中理解。例如，本申请中使用的诸如“和”、“或”或者“和/或”等术语可以包括各种含义，这些含义可能至少部分取决于使用这些术语的上下文。通常，“或”如果用于关联列表，诸如A、B或C，旨在表示A、B和C(此处用于包容性意义)以及A、B或C(此处用于独占性意义)。另外，本申请中使用的术语“一个或多个”或者“至少一个”，至少取决于上下文，可用于描述单数意义上的任何特征、结构或特性，或者可以用于描述复数意义上的特征、结构或特性的组合。类似地，诸如“一”(“a”或“an”)

或“该”之类的术语,可以理解为传达单数用法或者用于传达复数用法,至少部分取决于上下文。另外,术语“基于”或“由……确定”可以理解为不是必须旨在表示一组排它性因素,相反,可以允许存在不一定明确描述的其它因素,同样,至少部分取决于上下文。图3图示了根据本申请公开的实施例的通信系统(300)的简化框图。通信系统(300)包括多个终端装置,所述终端装置可通过例如网络(350)彼此通信。举例来说,通信系统(300)包括通过网络(350)互连的第一对终端装置(310)和(320)。在图3的示例中,第一对终端装置(310)和(320)可以执行单向数据传输。例如,终端装置(310)可对(例如,由终端装置(310)采集的视频图片流的)视频数据进行编码,以通过网络(350)传输到另一终端装置(320)。已编码视频数据可以以一个或多个已编码视频码流的形式传输。终端装置(320)可从网络(350)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图片,并根据恢复的视频数据显示视频图片。单向数据传输可以在媒体服务等应用中实施。

[0070] 在另一示例中,通信系统(300)包括执行已编码视频数据的双向传输的第二对终端装置(330)和(340),所述双向传输可例如在视频会议期间实施。对于双向数据传输,在一示例中,终端装置(330)和终端装置(340)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的)进行编码,以通过网络(350)传输到终端装置(330)和终端装置(340)中的另一终端装置。终端装置(330)和终端装置(340)中的每个终端装置还可接收由终端装置(330)和终端装置(340)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图片,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。

[0071] 在图3的示例中,终端装置(310)、终端装置(320)、终端装置(330)和终端装置(340)可能实施为服务器、个人计算机和智能电话,但本申请公开的潜在原理的适用性可不限于此。本申请公开的实施例可以实施于笔记本电脑、膝上型计算机、平板电脑、媒体播放器、可穿戴计算机、专用视频会议设备等等之类。网络(350)表示在终端装置(310)、终端装置(320)、终端装置(330)和终端装置(340)之间传送已编码视频数据的任何数目或类型的网络,包括例如有线(连线的)和/或无线通信网络。通信网络(350)可在电路交换、分组交换信道和/或其它类型的信道中交换数据。代表性的网络包括电信网络、局域网、广域网和/或互联网。出于本论述的目的,除非在下文中明确解释,否则网络(350)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。

[0072] 作为本申请所公开主题应用的示例,图4示出了视频编码器和视频解码器在视频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视频应用,包括例如视频会议、数字TV、广播、游戏、虚拟现实、在包括CD、DVD、存储棒等的数字介质上存储压缩视频等等。

[0073] 视频流式传输系统可包括视频采集子系统(413),所述采集子系统可包括数码相机等视频源(401),所述视频源用于创建未压缩的视频图片或图像流(402)。在一个示例中,视频图片流(402)包括由视频源401的数码相机录制的样本。视频图片流(402)描绘为粗线,以强调与已编码视频数据(404)(或已编码视频码流)相比,其具有较高的数据量,视频图片流(402)可由电子装置(420)处理,所述电子装置(420)包括耦接到视频源(401)的视频编码器(403)。视频编码器(403)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。已编码视频数据(404)(或已编码的视频码流(404))描绘为细

线,以强调与未压缩的视频图片流(402)相比,其具有较低的数据量,其可存储在流式传输服务器(405)上以供将来使用或直接存储到下游视频设备(未示出)。一个或多个流式传输客户端子系统,例如,图4中的客户端子系统(406)和客户端子系统(408),可访问流式传输服务器(405),以检索已编码的视频数据(404)的副本(407)和副本(409)。客户端子系统(406)可包括例如电子装置(430)中的视频解码器(410)。视频解码器(410)对已编码的视频数据的传入副本(407)进行解码,且产生可在显示器(412)(例如,显示屏)或另一呈现装置(未描绘)上呈现的、未压缩的输出视频图片流(411)。视频解码器410可以配置来执行本公开所描述的各种功能中的一些功能或所有功能。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(404)、视频数据(407)和视频数据(409)(例如,视频码流)进行编码。该些标准的示例包括ITU-T H.265建议书。在一个示例中,正在开发的视频编码标准非正式地称为多功能视频编码(Versatile Video Coding,VVC),本申请所公开主题可用于VVC标准和其它视频编码标准的上下文中。

[0074] 应注意,电子装置(420)和电子装置(430)可包括其它组件(未示出)。举例来说,电子装置(420)可包括视频解码器(未示出),且电子装置(430)还可包括视频编码器(未示出)。

[0075] 图5是下文根据本申请公开的任一实施例的视频解码器(510)的框图。视频解码器(510)可设置在电子装置(530)中。电子装置(530)可包括接收器(531)(例如,接收电路)。视频解码器(510)可用于代替图4示例中的视频解码器(410)。

[0076] 接收器(531)可接收将由视频解码器(510)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,一次解码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。每个视频序列可以与多个视频帧或图像相关联。已编码视频序列可从信道(501)接收,所述信道可以是通向存储已编码的视频数据的存储装置或传输已编码视频数据的流式传输源的硬件/软件链路。接收器(531)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处理电路(未描绘)的已编码音频数据和/或辅助数据流。接收器(531)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(515)可设置在接收器(531)与熵解码器/解析器(520)(此后称为“解析器(520)”)之间。在某些应用中,缓冲存储器(515)可以实施为视频解码器(510)的一部分。在其它应用中,所述缓冲存储器(515)可设置在视频解码器(510)外部且与视频解码器(510)分开(未描绘)。而在其它应用中,视频解码器(510)的外部设置缓冲存储器(未描绘),以例如用于防止网络抖动,且在视频解码器(510)的内部可配置另一附加的缓冲存储器(515),以例如处理播放定时。而当接收器(531)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(515),或可以将所述缓冲存储器做得较小。为了在互联网等尽力而为业务分组网络上使用,也可能需要具有足够大小的缓冲存储器(515),所述缓冲存储器的尺寸可相对较大。这种缓冲存储器可实施为具有自适应性大小,且可至少部分地实施于操作系统或视频解码器(510)外部的类似元件(未描绘)中。

[0077] 视频解码器(510)可包括解析器(520),以根据已编码视频序列重建符号(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,以及用以控制显示器(512)(例如,显示屏)等显示装置的潜在信息,所述呈现装置可以是或不是电子装置(530)的组成部分,但可耦接到电子装置(530),如图5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 (未描绘)。解析器 (520) 可对其接收到的已编码视频序列进行解析/熵解码。已编码视频序列的熵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 (520) 可基于对应于子群的至少一个参数, 从已编码视频序列中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 (tile)、条带 (slice)、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 (520) 还可从已编码视频序列提取信息, 例如, 变换系数 (例如, 傅里叶变换)、量化器参数值、运动矢量等等。

[0078] 解析器 (520) 可对从缓冲存储器 (515) 接收的视频序列执行熵解码/解析操作, 从而创建符号 (521)。

[0079] 取决于已编码视频图片或已编码视频图片部分的类型 (例如: 帧间图片和帧内图片、帧间块和帧内块) 以及其它因素, 符号 (521) 的重建可涉及多个不同的处理单元或功能单元。涉及哪些单元以及涉及方式可由解析器 (520) 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 (520) 与下文的多个处理单元或功能单元之间的此类子群控制信息流。

[0080] 除已经提及的功能块以外, 视频解码器 (510) 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, 这些功能单元中的许多功能单元彼此紧密交互并且至少可以部分彼此集成。然而, 出于清楚地描述所公开主题的各种功能的目的, 本公开在下文采用了概念上细分功能单元。

[0081] 第一单元可以包括缩放器/逆变换单元 (551)。缩放器/逆变换单元 (551) 可以从解析器 (520) 接收作为符号 (521) 的量化变换系数以及控制信息, 包括指示使用哪种类型的逆变换、块大小、量化因子/参数、量化缩放矩阵等。缩放器/逆变换单元 (551) 可输出包括样本值的块, 所述样本值可输入到聚合器 (555) 中。

[0082] 在一些情况下, 缩放器/逆变换单元 (551) 的输出样本可属于帧内编码块; 即: 不使用来自先前已重建图片的预测性信息、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 (552) 提供。在一些情况下, 帧内图片预测单元 (552) 采用已经重建并且存储在当前图片缓冲器 (558) 中的周围块信息, 可以生成大小和形状与正在重建的块相同的块。举例来说, 当前图片缓冲器 (558) 缓冲部分重建的当前图片和/或完全重建的当前图片。在一些实施方式中, 聚合器 (555) 可以基于每个样本, 将帧内预测单元 (552) 生成的预测信息添加到由缩放器/逆变换单元 (551) 提供的输出样本信息中。

[0083] 在其它情况下, 缩放器/逆变换单元 (551) 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 (553) 可访问参考图片存储器 (557) 以提取用于帧间图片预测的样本。在根据属于所述块的符号 (521) 对提取的样本进行运动补偿之后, 这些样本可由聚合器 (555) 添加到缩放器/逆变换单元 (551) 的输出 (单元 551 的输出可以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 (553) 从参考图片存储器 (557) 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 (521) 的形

式而供运动补偿预测单元(553)使用,所述符号(521)例如是包括X、Y分量(位移)和参考图片分量(时间)。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(557)提取的样本值的插值,还可以与运动矢量预测机制相关联等等。

[0084] 聚合器(555)的输出样本可在环路滤波器单元(556)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(520)的符号(521)可用于环路滤波器单元(556),但是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。可以以各种顺序将几种类型的环路滤波器包括为环路滤波单元556的一部分,这在下文会详细介绍。

[0085] 环路滤波器单元(556)的输出可以是样本流,所述样本流可输出到显示装置(512)以及存储在参考图片存储器(557)中,以用于后续的帧间图片预测。

[0086] 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的图片间预测。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(520))被识别为参考图片,则当前图片缓冲器(558)可变为参考图片存储器(557)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。

[0087] 视频解码器(510)可根据例如ITU-T H.265建议书等标准中采用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下,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。为了符合标准,还要求已编码视频序列的复杂度可以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。

[0088] 在一些示例实施例中,接收器(531)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(510)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、冗余条带、冗余图片、前向纠错码等形式。

[0089] 图6是根据本申请公开的示例实施例的视频编码器(603)的框图。视频编码器(603)可以设置于电子装置(620)中。电子装置(620)可以进一步包括传输器(640)(例如,传输电路)。视频编码器(603)可用于代替图4示例中的视频编码器(403)。

[0090] 视频编码器(603)可从视频源(601)(并非图6示例中的电子装置(620)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(603)编码的视频图像。在另一实施例中,视频源(601)可以实施为电子装置(620)的一部分。

[0091] 视频源(601)可提供将由视频编码器(603)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.601YCrCb、RGB、XYZ……)和任何合适取样结构(例如,YCrCb 4:2:0、YCrCb

4:4:4)。在媒体服务系统中,视频源(601)可以是能够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(601)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或图像,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。所属领域的普通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。

[0092] 根据一些示例实施例,视频编码器(603)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(643)。施行适当的编码速度构成控制器(650)的一个功能。在一些实施例中,控制器(650)可以在功能上耦接到并控制如下文所描述的其它功能单元。为了简洁起见,图中未标示耦接。由控制器(650)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(650)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(603)。

[0093] 在一些示例实施例中,视频编码器(603)可以在编码环路中运作。作为简单的描述,在一个示例中,编码环路可包括源编码器(630)(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(603)中的(本地)解码器(633)。即使嵌入的解码器633处理源编码器630的未熵编码的已编码视频流,解码器(633)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(因为在本申请所公开的主题所考虑的视频压缩技术中,熵编码中符号与已编码视频码流之间的任何压缩可以是无损的)。将重建的样本流(样本数据)输入到参考图片存储器(634)。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(634)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。使用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)以提高编码质量。

[0094] “本地”解码器(633)的操作可与例如已在上文结合图5详细描述视频解码器(510)的“远程”解码器相同。然而,另外简要参考图5,当符号可用且熵编码器(645)和解析器(520)能够无损地将符号编码/解码到已编解码视频序列时,包括缓冲存储器(515)和解析器(520)在内的视频解码器(510)的熵解码部分,可能不完全在编码器中的本地解码器(633)中实施。

[0095] 此时可以观察到,除可以仅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,本申请有时侧重于解码器操作,这类似于编码器的解码部分。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。在下文中仅在某些区域或方面提供了编码器更详细的描述。

[0096] 在操作期间,在一些示例实施方式中,源编码器(630)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,对输入图片进行预测性编码。以此方式,编码引擎(632)对输入图片的像素块与参考图片的像素块之间的颜色信道中的差值(或残差)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。术

语“残差”和其形容词“残差的”可以互换使用。

[0097] 本地视频解码器(633)可基于源编码器(630)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(632)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6中未示)处解码时,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(633)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(634)中。以此方式,视频编码器(603)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(远程)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)。

[0098] 预测器(635)可针对编码引擎(632)执行预测搜索。即,对于将要编码的新图片,预测器(635)可在参考图片存储器(634)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(635)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(635)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(634)中存储的多个参考图片取得的预测参考。

[0099] 控制器(650)可管理源编码器(630)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。

[0100] 可在熵编码器(645)中对所有上述功能单元的输出进行熵编码。熵编码器(645)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。

[0101] 传输器(640)可缓冲由熵编码器(645)创建的已编码视频序列,从而为通过通信信道(660)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(640)可将来自视频编码器(603)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)。

[0102] 控制器(650)可管理视频编码器(603)的操作。在编码期间,控制器(650)可以为每个已编码图片分配某一已编码图片类型,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:

[0103] 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。所属领域的普通技术人员了解I图片的变体及其相应的应用和特征。

[0104] 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。

[0105] 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。

[0106] 源图片通常可在空间上细分成多个样本块(例如,每个具有 8×4 、 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。源图片或中间处理的图片可以细分为其它类型的块,用于其它目的。编码块和其它类型的块的划分可以遵循或不遵循相同的方式,在下文有进一步描述。

[0107] 视频编码器(603)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(603)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。相应地,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。

[0108] 在一些示例实施例中,传输器(640)可传输附加数据和已编码的视频。源编码器(630)可以包括此类数据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、冗余图片和条带等其它形式的冗余数据、SEI消息、VUI参数集片段等。

[0109] 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。例如,可以将正在编码/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。

[0110] 在一些示例实施例中,双向预测技术可用于帧间图片预测中。根据这种双向预测技术,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或将来)的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。可通过第一参考块和第二参考块的组合来联合预测所述块。

[0111] 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。

[0112] 根据本申请公开的一些示例实施例,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,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可以包括三个平行的编码树块(coding tree block, CTB):一个亮度CTB和两个色度CTB。更进一步的,还可将每个CTU以二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。32×32块中的一个或多个中的每一个可以进一步拆分为4个具有16×16像素的CU。在一些示例实施中,可以在编码过程中分析每个CU以在各种预测类型中确定用于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,可以将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。将CU拆分为PU(或具有不同的颜色通道的PB)可以以各种空间模式进行。亮度或色度PB,例如,可以包括样本的矩阵值(例如,亮度值),例如,8×8像素、16×16像素、8×16像素、16×8像素等等。

[0113] 图7示出了根据本申请公开的另一示例实施例的视频编码器(703)的图。视频编码器(703)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。示例的视频编码器(703)可以用于代替图4示例中的视频编码器(403)。

[0114] 例如,视频编码器(703)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(703)使用例如率失真优化(RDO)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。当确定以帧内模式编码处理块时,视频编码器(703)可使用帧内预测技术以将处理块编码到已编码图片中;且当确定以帧间模式或双向预测模式编码处理块时,视频编码器(703)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。在一些示例实施例中,合并模式可以用作帧间图片预测的子模式,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。在某些其它示例实施例中,可存在适用于主题块的运动矢量分量。相应地,视频编码器(703)可以包括图7中未明确示出的组件,例如,用于确定处理块的预测模式的模式决策模块。

[0115] 在图7的示例中,视频编码器(703)包括如图7所示的耦接到一起的帧间编码器(730)、帧内编码器(722)、残差计算器(723)、开关(726)、残差编码器(724)、通用控制器(721)和熵编码器(725)。

[0116] 帧间编码器(730)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,按显示顺序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)。在一些示例中,参考图片是使用嵌入在图6的示例编码器620中的解码单元633(例如,图7的残差解码器728,如下文进一步详细描述)基于已编码的视频信息解码的已解码参考图片。

[0117] 帧内编码器(722)用于接收当前块(例如,处理块)的样本、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。帧内编码器(722)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)。

[0118] 通用控制器(721)可以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(703)的其它组件。在实施例中,通用控制器(721)确定块的预测模式,且基于所述预测模式将控制信号提供到开关(726)。举例来说,当所述预测模式是帧内模式时,通用控制器(721)控制开关(726)以选择供残差计算器(723)使用的帧内模式结果,且控制熵编码器(725)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;以及当所述用于块的预测模式是帧间模式时,通用控制器(721)控制开关(726)以选择供残差计算器(723)使用的帧间预测结果,且控制熵编码器(725)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。

[0119] 残差计算器(723)可以用于计算所接收的块与选自帧内编码器(722)或帧间编码器(730)的块的预测结果之间的差(残差数据)。残差编码器(724)可以用于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(724)用于将残差数据从时域转换到频域,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。例如,视频编码器(703)还可以包括残差解码器(728)。残差解码器(728)用于执行逆变换,以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(722)和帧间编码器(730)使用。举例来说,帧

间编码器(730)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(722)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,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。

[0120] 熵编码器(725)可以用于将码流格式化以产生已编码的块并进行熵编码。熵编码器(725)用于在码流中包括各种信息。例如,熵编码器(725)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。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。

[0121] 图8示出了根据本申请公开的另一实施例的示例视频解码器(810)的图。视频解码器(810)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。在一个示例中,视频解码器(810)可以用于代替图4示例中的视频解码器(410)。

[0122] 在图8的示例中,视频解码器(810)包括如图8中的示例布置所示耦接到一起的熵解码器(871)、帧间解码器(880)、残差解码器(873)、重建模块(874)和帧内解码器(872)。

[0123] 熵解码器(871)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所述块编码的模式(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)、可分别识别供帧内解码器(872)或帧间解码器(880)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。在一个示例中,当预测模式是帧间预测模式或双向预测模式时,将帧间预测信息提供到帧间解码器(880);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(872)。残差信息可经由逆量化并提供到残差解码器(873)。

[0124] 帧间解码器(880)可以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。

[0125] 帧内解码器(872)可以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。

[0126] 残差解码器(873)可以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(873)还可能利用某些控制信息(用以获得量化器参数QP),所述信息可由熵解码器(871)提供(未标示数据路径,因为这可能仅仅是低量控制信息)。

[0127] 重建模块(874)可以用于在空间域中组合由残差解码器(873)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,所述重建的块形成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。

[0128] 应注意,可使用任何合适的技术来实施视频编码器(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)。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(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)。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(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)。

[0129] 转向用于编码和解码的块分区,一般分区可以从基本块开始并且可以遵循预定义的规则集、特定模式、分区树或任何分区结构或方案。分区可以是分层次和递归的。在遵循任何示例分区流程或以下描述的其它流程或它们的组合对基本块进行划分或分区之后,可以获得分区或编码块的最终集。这些分区中的每一个可以处于分区层次中的各种分区级别中的一个,并且可以具有各种形状。分区中的每一个可以称作编码块(CB)。对于下面进一步描述的各种示例分区实施方式,所得的每一个CB可以具有任何允许的大小和分区级别。这种分区称作编码块,因为它们可以形成一些单元,可以对这些单元做出一些基本的编码/解码决策,并且这些单元的编码/解码参数可以在已编码视频码流中被优化、确定和发信号通知。最终分区中的最高或最深级别表示树的编码块分区结构的深度。编码块可以是亮度编码块或色度编码块。每种颜色的CB树结构可以称作编码块树(CBT)。

[0130] 所有颜色信道的编码块可以统称为编码单元(CU)。用于所有颜色信道的层次结构可以统称为编码树单元(CTU)。CTU中用于各种颜色信道的分区模式或结构可以相同或不相同。

[0131] 在一些实施方式中,用于亮度信道和色度信道的分区树方案或结构可以不必相同。换句话说,亮度信道和色度信道可以具有单独的编码树结构或模式。进一步地,亮度信道和色度信道是否使用相同或不同的编码分区树结构以及要使用的实际编码分区树结构可以取决于正被编码的条带是P条带、B条带还是I条带。例如,对于I条带,色度信道和亮度信道可以具有单独的编码分区树结构或编码分区树结构模式,而对于P条带或B条带,亮度信道和色度信道可以共享相同的编码分区树方案。当应用单独的编码分区树结构或模式时,亮度信道可以通过一个编码分区树结构分区成CB,色度信道可以通过另一个编码分区树结构分区成色度CB。

[0132] 在一些示例实施方式中,预定分区模式可以应用于基本块。如图9所示,示例的4路分区树可以从第一预定义的级别(例如, 64×64 块级别或其它大小,作为基本块大小)开始,基本块可以分层次地向下分区到预定义的最低级别(例如, 4×4 级别)。例如,基本块可以经过由902、904、906和908指示的四个预定义的分区选项或模式,其中指定为R的分区允许用于递归的分区,因为可以以较低的比例重复如图9中指示的相同分区选项,直到最低级别(例如, 4×4 级别)。在一些实施方式中,可以对图9的分区方案应用附加限制。在图9的实施方式中,可以允许使用矩形分区(例如, $1:2/2:1$ 矩形分区),但是可以不允许它们是递归的,而允许正方形分区是递归的。如果需要,遵循图9的递归分区生成编码块的最终集。可以进一步定义编码树深度,以指示从根节点或根块的分割深度。例如,根节点或根块(例如, 64×64 块)的编码树深度可以设置为0,根块在遵循图9被进一步分割一次之后,编码树深度增加1。对于上述方案,从 64×64 基本块到 4×4 的最小分区的最大级别或最深级别将是4(从级别0开始)。这种分区方案可以应用于一个或多个颜色信道。可以遵循图9的方案独立地对每一个颜色信道进行分区(例如,可以针对每一个层次级别处的每一个颜色信道,独立地确定预定义模式中的分区模式或选项)。可选地,两个或更多个颜色信道可以共享图9的相同分层次的模式树(例如,可以为每一个层次级别的两个或更多个颜色信道选择预定义的模式中的相同分区模式或选项)。

[0133] 图10示出了允许使用递归的分区以形成分区树的另一示例预定义的分区模式。如图10中所示,可以预定义示例的10路分区结构或模式。根块可以开始于预定义的级别(例

如,从 128×128 级别或 64×64 级别的基本块)开始。图10的示例分区结构包括各种 $2:1/1:2$ 和 $4:1/1:4$ 矩形分区。图10的第二行中指示为1002、1004、1006和1008的具有3个子分区的分区类型可以称作“T型”分区。“T型”分区1002、1004、1006和1008可以称作左T型、顶部T型、右T型和底部T型。在一些示例实施方式中,图10的矩形分区中没有一个被允许进一步细分。可以进一步定义编码树深度,以指示从根节点或根块的分割深度。例如,根节点或根块(例如, 128×128 块)的编码树深度可以被设置为0,根块在遵循图10被进一步分割一次之后,编码树深度增加1。在一些实施方式中,可以仅允许1010中的所有正方形分区遵循图10的模式递归分区到分区树的下一个级别。换句话说,对于T型模式内的正方形分区1002、1004、1006和1008,可以不允许递归的分区。如果需要,遵循图10的递归分区过程生成编码块的最终集。这样的方案可以应用于一个或多个颜色信道。在一些实施方式中,可以为低于 8×8 级别的分区的使用增加更多的灵活性。例如,在某些情况下可以使用 2×2 色度帧间预测。

[0134] 在用于编码块分区的一些其它示例实施例中,四叉树结构可以用于将基本块或中间块分割成四叉树分区。这种四叉树分割可以是分层次和递归地应用于任何正方形分区。无论是基本块、中间块或分区进行进一步的四叉树分割,都可以适于基本块或中间块/分区的各种局部特性。图片边界的四叉树分区可以进一步调整。例如,可以在图片边界执行隐式四叉树分割,使得一个块将不断进行四叉树分割,直到大小适合图片边界。

[0135] 在一些其它示例实施方式中,可以使用从基本块进行的分层次二叉分区。对于这种方案,可以将基本块或中间级别块分区成两个分区。二叉分区可以是水平的或垂直的。例如,水平二叉分区可以将基本块或中间块分割成相等的左右分区。同样地,垂直二叉分区可以将基本块或中间块分割成相等的上下分区。这种二叉分区可以是分层次和递归的。可以在基本块或中间块的每一者之处做出二叉分区方案是否应该继续的决策,如果方案进一步继续,则做出是否应该使用水平二叉分区还是垂直二叉分区的决策。在一些实施方式中,进一步分区可以停止在预定义的最低分区大小(在一个维度或两个维度)。可选地,一旦达到从基本块开始的预定义的分区级别或深度,进一步分区便可以停止。在一些实施方式中,可以限制分区的纵横比。例如,分区的纵横比可以不小于 $1:4$ (或大于 $4:1$)。这样,具有 $4:1$ 的垂直与水平纵横比的垂直条分区可以仅进一步被垂直地二叉分区为上分区和下分区,各自具有 $2:1$ 的垂直与水平纵横比。

[0136] 在又一些其它示例中,如图13所示,三叉分区方案可以用于分区基本块或任何中间块。三叉模式可以如图13中的1302所示垂直地实施,或者如图13中的1304所示水平地实施。虽然图13中垂直或水平的示例分割比为 $1:2:1$,但是可以预定义其它比率。在一些实施方式中,可以预定义两个或更多个不同的比率。这种三叉分区方案可以用于补偿四叉树或二叉分区结构,因为这种三叉树分区能够在一个连续的分区中采集位于块中心的对象,而四叉树和二叉树总是沿着块中心进行分割,因此会将对象分割成单独的分区。在一些实施方式中,示例三叉树的分区的宽度和高度总是2的幂,以避免有附加的变换。

[0137] 以上分区方案可以在不同的分区级别以任何方式组合。作为一个示例,上述四叉树和二叉分区方案可以进行组合,以将基本块分区成四叉树—二叉树(QTBT)结构。在这种方案中,基本块或中间块/分区可以是四叉树分割或二叉分割,如果有指定预定义的条件集,则服从该预定义的条件集。在图14中图示了特定示例。在图14的示例中,基本块首先被四叉树分割成四个分区,如1402、1404、1406和1408所示。此后,所得到的每一个分区要么被

四叉树分区成四个另外的分区(诸如1408),要么在下一个级别被二叉分割成两个另外的分区(水平或垂直,诸如1402或1406,例如两者都是对称的),要么不被分割(诸如1404)。对于正方形分区,可以递归地允许二叉或四叉树分割,如1410的总体示例分区模式和1420中对应的树结构/表示所示,其中实线表示四叉树分割,虚线表示二叉分割。标志可以用于每一个二叉分割节点(非叶二叉分区),以指示二叉分割是水平的还是垂直的。例如,如1420所示,与1410的分区结构一致,标志“0”可以表示水平二叉分割,标志“1”可以表示垂直二叉分割。对于四叉树分割分区,不需要指示分割类型,因为四叉树分割总是水平地和垂直地分割块或分区以产生具有相等大小的4个子块/分区。在一些实施方式中,标志“1”可以表示水平二叉分割,标志“0”可以表示垂直二叉分割。

[0138] 在QTBT的一些示例实施方式中,四叉树和二叉分割规则集可以由以下预定义的参数和与它们相关联的对应函数来表示:

[0139] -CTU大小:四叉树的根节点大小(基本块的大小)

[0140] -MinQTSIZE:允许的最小四叉树叶节点大小

[0141] -MaxBTSIZE:允许的最大二叉树根节点大小

[0142] -MaxBTDepth:允许的最大二叉树深度

[0143] -MinBTSIZE:允许的最小二叉树叶节点大小

[0144] 在QTBT分区结构的一些示例实施方式中,CTU大小可以被设置为具有两个对应的 64×64 色度样本块的 128×128 亮度样本(当考虑并使用示例色度子采样时),MinQTSIZE可以被设置为 16×16 ,MaxBTSIZE可以被设置为 64×64 ,MinBTSIZE(对于宽度和高度两者)可以被设置为 4×4 ,MaxBTDepth可以被设置为4。四叉树分区可以首先应用于CTU以生成四叉树叶节点。四叉树叶节点的大小可以从其允许的最小大小 16×16 (即,MinQTSIZE)到 128×128 (即,CTU大小)。如果节点是 128×128 ,则其将不会首先被二叉树分割,因为该尺寸超过了MaxBTSIZE(即 64×64)。否则,不超过MaxBTSIZE的节点可以由二叉树分区。在图14的示例中,基本块是 128×128 。根据预定义的规则集,基本块只能进行四叉树分割。基本块的分区深度为0。所得到的四个分区中的每一个分区是 64×64 ,不超过MaxBTSIZE,可以进一步在级别1处进行四叉树或二叉树分割。该过程继续。当二叉树深度达到MaxBTDepth(即,4)时,可以不考虑进一步的分割。当二叉树节点的宽度等于MinBTSIZE(即,4)时,可以不考虑进一步的水平分割。类似地,当二叉树节点的高度等于MinBTSIZE时,不考虑进一步的垂直分割。

[0145] 在一些示例实施方式中,以上QTBT方案可以被配置为支持亮度和色度具有相同QTBT结构或单独的QTBT结构的灵活性。例如,对于P条带和B条带,一个CTU中的亮度CTB和色度CTB可以共享相同的QTBT结构。然而,对于I条带,亮度CTB可以通过QTBT结构被分区成CB,色度CTB可以通过另一QTBT结构被分区成色度CB。这意味着CU可以用于指代I条带中的不同颜色信道,例如,I条带可以由亮度分量的编码块或两个色度分量的编码块组成,且P条带或B条带中的CU可以由所有三个颜色分量的编码块组成。

[0146] 在一些其它实施方式中,QTBT方案可以用上述三叉方案来补充。这种实施方式可以称作多-类型-树(MTT)结构。例如,除了节点的二叉分割之外,可以选择图13的三叉分区模式中的一个。在一些实施方式中,只有正方形节点可以进行三叉分割。可以使用附加标志来指示三叉分区是水平还是垂直。

[0147] 诸如QTBT实施方式和由三叉分割补充的QTBT实施方式的两级或多级树的设计可

以主要由降低复杂度来推动。理论上,遍历树的复杂度是 T^D ,其中,T表示分割类型的数量,D是树的深度。可以通过使用多种类型(T)同时减小深度(D)进行折衷。

[0148] 在一些实施方式中,CB可以被进一步分区。例如,出于在编码和解码过程期间的帧内预测或帧间预测的目的,CB可以被进一步分区成多个预测块(PB)。换句话说,CB可以被进一步划分为不同的子分区,这些子分区中可以进行单独的预测决策/配置。与此同时,出于描绘执行视频数据的变换或逆变换的级别的目的,CB可以被进一步分区成多个变换块(TB)。CB到PB和TB的分区方案可以是相同的或不同的。例如,可以基于例如视频数据的各种特性使用其自己的流程来执行每一个分区方案。在一些示例实施方式中,PB和TB分区方案可以是独立的。在一些其它示例实施方式中,PB和TB分区方案和边界可以是相关的。在一些实施方式中,例如,TB可以在PB分区之后被分区,具体地,每一个PB(在编码块的分区之后被确定之后)然后可以被进一步分区成一个或多个TB。例如,在一些实施方式中,PB可以被分割成一个、两个、四个或其它数量的TB。

[0149] 在一些实施方式中,为了将基本块分区成编码块并且进一步分区成预测块和/或变换块,可以以不同的方式处理亮度信道和色度信道。例如,在一些实施方式中,对于亮度信道可以允许将编码块分区成预测块和/或变换块,而对于一个或多个色度信道可以不允许将编码块分区成预测块和/或变换块。在这种实施方式中,亮度块的变换和/或预测因此可以仅在编码块级执行。再举个例子,用于亮度信道和一个或多个色度信道的最小变换块大小可以不同,例如,可以允许用于亮度信道的编码块被分区成比色度信道小的变换块和/或预测块。再举一个例子,将编码块分区成变换块和/或预测块的最大深度在亮度信道与色度信道之间可以不同,例如,可以允许用于亮度信道的编码块被分区成比一个或多个色度信道更深的变换块和/或预测块。举一个具体例子,亮度编码块可以被分区成具有多个大小的变换块,这些变换块可以由向下进行多达2级的递归的分区来表示,并且可以允许诸如正方形、2:1/1:2和4:1/1:4的变换块形状和从 4×4 到 64×64 的变换块大小。然而,对于色度块,仅可以允许为亮度块指定的最大可能的变换块。

[0150] 在用于将编码块分区成PB的一些示例实施方式中,PB分区的深度、形状和/或其它特性可以取决于PB是帧内编码的还是帧间编码的。

[0151] 可以在各种示例方案中实施将编码块(或预测块)分区成变换块,示例方案包括但不限于递归地或非递归地二叉树分割和预定义的模式分割,并且附加考虑编码块或预测块的边界处的变换块。通常,所得到的变换块可以处于不同的分割级别,可以不具有相同的大小,并且可以不需要形状为正方形(例如,它们可以是具有一些允许的大小和纵横比的矩形)。以下关于图15、图16和图17更详细地描述了其它示例。

[0152] 然而,在一些其它实施方式中,经由以上任何分区方案获得的CB可以被用作用于预测和/或变换的基本块或最小编码块。换句话说,出于执行帧间预测/帧内预测的目的和/或出于变换的目的,不执行进一步的分割。例如,从以上QTBT方案获得的CB可以直接用作执行预测的单元。具体地,这种QTBT结构去除了多个分区类型的概念,即,去除了CU、PU和TU的区分,为如上所述的CU/CB分区形状提供了更大灵活性。在这种QTBT块结构中,CU/CB可以具有正方形或矩形形状。这种QTBT的叶节点被用作预测和变换处理的单元而无需任何进一步的分区。这意味着,CU、PU和TU在这种示例QTBT编码块结构中具有相同的块大小。

[0153] 以上各种CB分区方案和将CB进一步分区为PB和/或TB(不包括PB/TB分区)可以以

任何方式组合。提供了以下特定实施方式作为非限制性示例。

[0154] 以下描述编码块和变换块分区的具体示例实施方式。在这种示例实施方式中,可以使用递归的四叉树分割或上文描述的预定义的分割模式(例如图9和图10中的模式),将基本块分割成编码块。在每一个级别处,特定分区的进一步四叉树分割是否应该继续可以由本地视频数据特性来确定。所得到的CB可以处于各种四叉树分割级别,并且具有各种大小。可以在CB级(或CU级,对于所有三色信道)做出关于是否使用帧间图片(时间)预测还是帧内图片(空间)预测对图片区域进行编码的决策。可以根据预定义的PB分割类型将每一个CB进一步分割成一个、两个、四个或其它数量的PB。在一个PB内,可以应用相同的预测过程,并且可以在PB的基础上将相关信息传输到解码器。在通过应用基于PB分割类型的预测过程获得残差块之后,可以根据与用于CB的编码树类似的另一四叉树结构将CB分区成TB。在该特定实施方式中,CB或TB可以但不必限于正方形形状。进一步地,在该特定示例中,PB对于帧间预测可以是正方形或矩形形状并且对于帧内预测可以仅是正方形。编码块可以被分割成例如四个正方形TB。每一个TB可以被进一步递归地分割(使用四叉树分割)成更小的TB,称作残差四叉树(RQT)。

[0155] 下面进一步描述用于将基本块分区成CB、PB和/或TB的另一示例实施方式。例如,可以使用使用二叉和三叉分割分段结构(例如,如上所述的具有三叉分割的QTBT或QTBT)的具有嵌套多类型树的四叉树,而不是使用诸如图9或图10所示的多分区单元类型。可以放弃CB、PB和TB的区分(即,将CB分区为PB和/或TB,以及将PB分区为TB),除非需要具有对于最大变换长度而言太大的大小的CB,其中这种CB可能需要进一步分割。该示例分区方案可以被设计为,为CB分区形状提供更多灵活性,使得预测和变换都可以在CB级上执行而无需进一步分区。在这种编码树结构中,CB可以具有正方形或矩形形状。具体地,编码树块(CTB)可以首先被四叉树结构分区。然后,四叉树叶节点可以由嵌套多类型树结构进一步分区。图11中示出了使用二叉或三叉分割的嵌套多类型树结构的示例。具体地,图11的示例多类型树结构包括四种分割类型,称作垂直二叉分割(SPLIT_BT_VER) (1102)、水平二叉分割(SPLIT_BT_HOR) (1104)、垂直三叉分割(SPLIT_TT_VER) (1106)和水平三叉分割(SPLIT_TT_HOR) (1108)。然后,CB对应于多类型树的叶。在该示例实施方式中,除非CB对于最大变换长度太大,否则该分段被用于预测和变换处理而无需任何进一步的分区。这意味着,在大多数情况下,CB、PB和TB在具有嵌套多类型树编码块结构的四叉树中具有相同的块大小。当最大支持的变换长度小于CB的颜色分量的宽度或高度时,会发生例外情况。在一些实施方式中,除了二叉或三叉分割之外,图11的嵌套模式可以进一步包括四叉树分割。

[0156] 图12示出了用于一个基本块的具有块分区的嵌套多类型树编码块结构(包括四叉树、二叉和三叉分割选项)的四叉树的一个具体示例。更详细地,图12示出了基本块1200被四叉树分割成四个正方形分区1202、1204、1206和1208。为每一个四叉树分割分区作出进一步使用图11的多类型树结构和用于进一步分割的四叉树的决策。在图12的示例中,分区1204未进一步被分割。分区1202和1208中的每一个采用另一四叉树分割。对于分区1202,第二级四叉树分割的左上、右上、左下和右下分区分别采用四叉树的第三级分割、图11的水平二叉分割1104、非分割和图11的水平三叉分割1108。分区1208采用另一四叉树分割,第二级四叉树分割的左上、右上、左下和右下分区分别采用图11的垂直三叉分割1106的第三级分割、非分割、非分割和图11的水平二叉分割1104。1208的第三级左上分区的两个子分区分别

根据图11的水平二叉分割1104和水平三叉分割1108被进一步分割。分区1206采用遵循图11的垂直二叉分割1102的第二级分割模式,被分割成两个分区,这两个分区在第三级根据图11的水平三叉分割1108和垂直二叉分割1102被进一步分割。根据图11的水平二叉分割1104,第四级分割被进一步应用于它们中的一个。

[0157] 对于以上具体示例,最大亮度变换大小可以是 64×64 ,最大支持的色度变换大小可以不同于亮度,例如,为 32×32 。即使以上图12中的示例CB通常不进一步分割成更小的PB和/或TB,当亮度编码块或色度编码块的宽度或高度大于最大变换宽度或高度时,亮度编码块或色度编码块可以在水平和/或垂直方向上自动分割以满足该方向上的变换大小限制。

[0158] 在以上用于将基本块分区成CB的具体示例中,如上所述,编码树方案可以支持亮度和色度具有单独的块树结构的能力。例如,对于P条带和B条带,一个CTU中的亮度CTB和色度CTB可以共享相同的编码树结构。例如,对于I条带,亮度和色度可以具有单独的编码块树结构。当应用单独的块树结构时,亮度CTB可以通过一个编码树结构被分区成亮度CB,色度CTB通过另一编码树结构被分区成色度CB。这意味着I条带中的CU可以由亮度分量的编码块或两个色度分量的编码块组成,P条带或B条带中的CU始终由所有三个颜色分量的编码块组成,除非视频是单色的。

[0159] 当编码块被进一步分区成多个变换块时,其中的变换块可以按照各种顺序或扫描方式在码流中排序。下文进一步详细描述用于将编码块或预测块分区成变换块的示例实施方式以及变换块的编码顺序。在一些示例实施方式中,如上所述,变换分区可以支持多种形状(例如,1:1(正方形)、1:2/2:1和1:4/4:1)的变换块,其中变换块大小范围从例如 4×4 到 64×64 。在一些实施方式中,如果编码块小于或等于 64×64 ,则变换块分区可以仅应用于亮度分量,使得对于色度块,变换块大小与编码块大小相同。否则,如果编码块宽度或高度大于64,则亮度和色度编码块两者可以分别被隐式地分割成 $\min(W, 64) \times \min(H, 64)$ 和 $\min(W, 32) \times \min(H, 32)$ 变换块的倍数。

[0160] 在变换块分区的一些示例实施方式中,对于帧内编码块和帧间编码块两者,编码块可以被进一步分区成具有高达预定义数量的级别(例如,2个级别)的分区深度的多个变换块。变换块分区深度和分区大小可以是相关的。对于一些示例实施方式,从当前深度的变换大小到下一深度的变换大小的映射如下表一中所示。

[0161] 表1:变换分区大小设置

当前深度的变换大小	下一深度的变换大小
TX_4X4	TX_4X4
TX_8X8	TX_4X4
TX_16X16	TX_8X8
TX_32X32	TX_16X16
TX_64X64	TX_32X32
TX_4X8	TX_4X4
TX_8X4	TX_4X4
TX_8X16	TX_8X8
[0162] TX_16X8	TX_8X8
TX_16X32	TX_16X16
TX_32X16	TX_16X16
TX_32X64	TX_32X32
TX_64X32	TX_32X32
TX_4X16	TX_4X8
TX_16X4	TX_8X4
TX_8X32	TX_8X16
TX_32X8	TX_16X8
TX_16X64	TX_16X32
TX_64X16	TX_32X16

[0163] 基于表1的示例映射,对于1:1正方形块,下一级变换分割可以创建四个1:1正方形子变换块。变换分区可以例如在 4×4 处停止。这样,当前深度的变换大小 4×4 对应于下一深度的相同大小 4×4 。在表1的示例中,对于1:2/2:1非正方形块,下一级变换分割可以创建两个1:1正方形子变换块,而对于1:4/4:1非正方形块,下一级变换分割可以创建两个1:2/2:1子变换块。

[0164] 在一些示例实施方式中,对于帧内已编码块的亮度分量,可以针对变换块分区应用附加限制。例如,对于每一个级别的变换分区,其所有子变换块可以被限制为具有相等的大小。例如,对于 32×16 编码块,级别1变换分割创建两个 16×16 子变换块,级别2变换分割创建八个 8×8 子变换块。换句话说,第二级分割必须应用于所有第一级子块,以保持变换单元大小相等。图15中示出了按照表1的用于帧内已编码正方形块的变换块分区的示例,以及由箭头图示的编码顺序。具体地,1502示出了正方形编码块。1504中示出了根据表1分割成4个大小相等的变换块的第一级分割,编码顺序由箭头指示。1506中示出了根据表1将所有第一级大小相等的块第二级分割成16个大小相等的变换块,编码顺序由箭头指示。

[0165] 在一些示例实施方式中,对于帧间已编码块的亮度分量,可以不应用以上对帧内编码的限制。例如,在第一级变换分割之后,子变换块中的任何一个可以被进一步独立地再分割一个级别。因此,所得到的变换块可以具有或不具有相同的大小。图16中示出了将帧间已编码块分割成具有它们编码顺序的变换的示例。在图16的示例中,根据表1,帧间已编码块1602被分割成两级的变换块。在第一级,帧间已编码块被分割成四个相等大小的变

换块。然后,如1604所示,四个变换块中仅一个(不是所有的)被进一步分割成四个子变换块,产生具有两种不同的大小的总共7个变换块。这7个变换块的示例编码顺序由图16的1604中的箭头示出。

[0166] 在一些示例实施方式中,对于一个或多个色度分量,可以应用一些对变换块的附加限制。例如,对于一个或多个色度分量,变换块大小可以与编码块大小一样大,但不小于预定义的大小,例如 8×8 。

[0167] 在一些其它示例实施方式中,对于宽度(W)或高度(H)大于64的编码块,亮度编码块和色度编码块两者可以分别隐含地被分割成 $\min(W, 64) \times \min(H, 64)$ 和 $\min(W, 32) \times \min(H, 32)$ 变换单元的倍数。这里,在本公开中,“ $\min(a, b)$ ”可以返回a和b之间的较小值。

[0168] 图17进一步示出了用于将编码块或预测块分区成变换块的另一可选示例方案。如图17中所示,不使用递归的变换分区,而是可以根据编码块的变换类型,将预定义的一组分区类型应用于编码块。在图17中所示的特定示例中,可以应用6种示例分区类型中的一种,将编码块分割成各种数量的变换块。这种生成变换块分区的方案可以应用于编码块或预测块。

[0169] 更详细地,图17的分区方案提供了用于任何给定变换类型(变换类型是指例如主变换的类型,诸如ADST及其它)的多达6种示例分区类型。在该方案中,可以基于(例如)率失真成本,向每一编码块或预测块指派变换分区类型。在一示例中,可以基于编码块或预测块的变换类型,确定指派给编码块或预测块的变换分区类型。如图17所示的6种变换分区类型所示,特定变换分区类型可以对应于变换块分割大小和分割模式。各种变换类型和各种变换分区类型之间的对应关系可以预定义。下面示出了一个示例,大写标记指示可以基于速率失真成本分配给编码块或预测块的变换分区类型:

[0170] • PARTITION_NONE:分配等于块大小的变换大小。

[0171] • PARTITION_SPLIT:分配变换大小,其宽度是块大小宽度的1/2并且高度是块大小高度的1/2。

[0172] • PARTITION_HORZ:分配变换大小,其宽度与块大小宽度相同并且高度是块大小高度的1/2。

[0173] • PARTITION_VERT:分配变换大小,其宽度是块大小宽度的1/2并且高度与块大小高度相同。

[0174] • PARTITION_HORZ4:分配变换大小,其宽度与块大小宽度相同并且高度是块大小高度的1/4。

[0175] • PARTITION_VERT4:分配变换大小,其宽度是块大小宽度的1/4并且高度与块大小高度相同。

[0176] 在以上示例中,如图17所示的变换分区类型都包含用于分区后的变换块的统一变换大小。这仅仅是示例而不是限制。在一些其它实施方式中,混合的变换块大小可以用于特定分区类型(或模式)的分区后的变换块。

[0177] 从以上任何分区方案获得的PB(或CB,当未被进一步分区成预测块时也称为PB),然后可以变成用于经由帧内预测或帧间预测进行编码的各个块。对于用于当前PB的帧间预测,当前块与预测块之间的残差可以被生成、被编码并被包括在已编码码流中。

[0178] 帧间预测可以例如在单个参考模式或复合参考模式下实施。在一些实施方式中,

跳过标志可以首先包括在用于当前块的码流中(或在较高级别),以指示当前块是否是帧间编码且不被跳过。如果当前块是帧间编码,则可以进一步将另一标志包括在码流中作为信号,以指示是单个参考模式还是复合参考模式用于当前块的预测。对于单个参考模式,可以使用一个参考块来生成用于当前块的预测块。对于复合参考模式,可以使用两个或更多个参考块(例如),通过加权平均生成预测块。复合参考模式可以是指多于一个参考的模式、两个参考的模式或多个参考的模式。一个或多个参考块可以使用一个参考帧索引或多个参考帧索引附加使用对应的一个运动矢量或多个运动矢量识别,所述一个运动矢量或多个运动矢量指示一个或多个参考块与当前块之间在位置(例如,在水平和垂直像素中)上的一个或多个移位。例如,在单参考模式下,用于当前块的帧间预测块可以从由参考帧中的一个运动矢量识别为预测块的单个参考块生成,而对于复合参考模式,预测块可以通过由两个参考帧索引和两个对应的运动矢量指示的两个参考帧中的两个参考块的加权平均生成。一个或多个运动矢量可以以各种方式被编码并包括在码流中。

[0179] 在一些实施方式中,编码或解码系统可以维护已解码图片缓冲器(DPB)。一些图像/图片可以保持在DPB中等待显示(在解码系统中),DPB中的一些图像/图片可以用作参考帧,以实现帧间预测(在解码系统或编码系统中)。在一些实施方式中,DPB中的参考帧可以被标记为用于正在编码或解码的当前图像的短期参考或长期参考。例如,短期参考帧可以包括用于按解码顺序对当前帧中或与当前帧最接近的预定义的数量(例如,2)的后续视频帧中的块进行帧间预测的帧。长期参考帧可以包括DPB中的帧,这些帧可以用于按解码顺序预测远离当前帧多于预定义的数量帧的帧中的图像块。关于这种用于短期参考帧和长期参考帧的标签的信息可以称作参考图片集(RPS),并且可以被添加到已编码码流中的每一帧的头部。已编码视频流中的每一帧可以由图片顺序计数(POC)来识别,该图片顺序计数根据回放序列以绝对方式来编号或者与例如从I帧开始的图片组相关。

[0180] 在一些示例实施方式中,可以基于RPS中的信息形成一个或多个参考图片列表,其中包含用于帧间预测的短期参考帧和长期参考帧的标识。例如,可以形成单个图片参考列表用于单向帧间预测,表示为L0参考(或参考列表0),而可以形成两个图片参考列表用于双向帧间预测,表示为L0(或参考列表0)和L1(或参考列表1)用于两个预测方向中的每一个。包括在L0和L1列表中的参考帧可以以各种预定方式排序。可以在视频码流中发信号通知L0和L1列表的长度。当用于在复合预测模式中通过加权平均生成预测块的多个参考在待预测块的同一侧时,单向帧间预测可以是在单个参考模式下或在复合参考模式下。双向帧间预测可以仅是复合模式,因为双向帧间预测涉及至少两个参考块。

[0181] 在一些实施方式中,可以实施用于帧间预测的合并模式(MM)。通常,对于合并模式,用于当前PB的单个参考预测中的运动矢量或复合参考预测中的一个或多个运动矢量可以从其它一个或多个运动矢量导出,而不是独立地计算和发信号通知。例如,在编码系统中,用于当前PB的一个或多个当前运动矢量可以由一个或多个当前运动矢量与其它一个或多个已编码运动矢量(称作参考运动矢量)之间的一个或多个差表示。一个或多个运动矢量中的一个或多个这种差而不是整个一个或多个当前运动矢量可以被编码并被包括在码流中,并且可以被链接到一个或多个参考运动矢量。相应地,在解码系统中,可以基于一个或多个已解码运动矢量差和与它们链接的一个或多个已解码参考运动矢量来导出对应于当前PB的一个或多个运动矢量。作为一般合并模式(MM)帧间预测的具体形式,这种基于一个

或多个运动矢量差的帧间预测可以称作具有运动矢量差的合并模式(MMVD)。因此,通常可以实施MM,或者具体地可以实施MMVD,以利用与不同PB相关联的运动矢量之间的相关性来提高编码效率。例如,相邻PB可以具有类似的运动矢量,并且因此MVD可以很小并且可以被有效地编码。对于另一示例,运动矢量可以在时间上(帧之间)与空间中位置相似/定位的块相关联。

[0182] 在一些示例实施方式中,在编码过程期间MM标志可以包括在码流中,用于指示当前PB是否合并模式。附加地,或可选地,在编码过程期间MMVD标志可以被包括在码流中并被发信号通知,以指示当前PB是否处于MMVD模式。MM和/或MMVD标志或指示器可以在PB级、CB级、CU级、CTB级、CTU级、条带级、图片级等处被提供。对于特定示例,可以针对当前CU包括MM标志和MMVD标志两者,并且可以在跳过标志和MM标志之后立即发信号通知MMVD标志以指定是否针对当前CU使用MMVD模式。

[0183] 在MMVD的一些示例实施方式中,可以为正在被预测的块形成用于运动矢量预测的参考运动矢量(RMV)或MV预测子候选的列表。RMV候选的列表可以包含预定数量(例如,2)的MV预测子候选块,它们的运动矢量可以用于预测当前运动矢量。RMV候选块可以包括从相同帧中的相邻块和/或时间块(例如,当前帧的前一帧或后一帧中相同位置的块)中选择的块。这些选项表示在相对于当前块的空间或时间位置处可能具有与当前块相似或相同的运动矢量的块。MV预测子候选的列表的大小可以是预定的。例如,列表可以包含两个或更多个候选项。为了在RMV候选的列表上,例如,候选块可能需要具有与当前块相同的参考帧(或多个帧),必须存在(例如,当当前块接近帧的边缘时,需要执行边界检查),并且必须在编码过程期间已经被编码,和/或在解码过程期间已经被解码。在一些实施方式中,合并候选的列表可以首先填充有空间上相邻的块(以特定预定义的顺序扫描)(如果可用并且满足以上条件),并且然后填充有时间块(如果在列表中空间仍然可用)。例如,可以从当前块的左侧块和顶部块中选择相邻RMV候选块。RMV预测子候选的列表可以在各种级(序列、图片、帧、条带、超块等)处动态地形成成为动态参考列表(DRL)。可以在码流中发信号通知DRL。

[0184] 在一些实施方式中,可以发信号通知被用作用于预测当前块的运动矢量的参考运动矢量的实际MV预测子候选。在RMV候选列表包含两个候选的情况下,称作合并候选标志的1-比特标志可以用于指示参考合并候选的选择。对于以复合模式预测的当前块,使用MV预测子预测的多个运动运动矢量中的每一个可以与来自合并候选列表的参考运动矢量相关联。编码器可以确定哪个RMV候选更接近地预测当前编码块,并且将该选择作为索引发信号通知给DRL。

[0185] 在MMVD的一些示例实施方式中,在选择RMV候选并将其用作用于待预测的运动矢量的基础运动矢量预测值之后,可以在编码系统中计算运动矢量差(MVD或增量MV,表示待预测的运动矢量与参考候选运动矢量之间的差)。这种MVD可以包括表示MV差的量值和MV差的方向的信息,这两者都可以在码流中被发信号通知。运动差量值和运动差方向可以以各种方式被发信号通知。

[0186] 在MMVD的一些示例实施方式中,可以使用距离索引来指定运动矢量差的量值信息,并且指示表示来自起始点(参考运动矢量)的预定义的运动矢量差的一组预定义的偏移中的一个。然后可以根据发信号通知的索引的MV偏移添加到起始(参考)运动矢量的水平分量或垂直分量。参考运动矢量的水平分量或垂直分量是否应该被偏移可以由MVD的方向

信息来确定。在表2中指定了距离索引和预定义的偏移之间的示例预定义的关系。

[0187] 表2—距离索引和预定义的MV偏移的示例关系

[0188]	距离索引	0	1	2	3	4	5	6	7
	偏移(以亮度样本为单位)	1/4	1/2	1	2	4	8	16	32

[0189] 在MMVD的一些示例实施方式中,方向索引可以被进一步发信号通知并用于表示MVD相对于参考运动矢量的方向。在一些实施方式中,该方向可以被限制为水平方向和垂直方向中的任一个。表3中示出了示例2-比特方向索引。在表3的示例中,MVD的解释可以根据起始/参考MV的信息而变化。例如,当起始/参考MV对应于单预测块或对应于双预测块且两个参考帧列表指向当前图片的同一侧时(即,两个参考图片的POC都大于当前图片的POC,或都小于当前图片的POC),表3中的符号可以指定添加到起始/参考MV的MV偏移的符号(方向)。当起始/参考MV对应于双向预测块且参考图片在当前图片的不同侧(即,一个参考图片的POC大于当前图片的POC,且另一参考图片的POC小于当前图片的POC),且图片参考列表0中的参考POC与当前帧之间的差大于图片参考列表1中的参考POC与当前帧之间的差时,表3中的符号可以指定添加到对应于图片参考列表0中的参考图片的参考MV的MV偏移的符号,且用于对应于图片参考列表1中的参考图片的MV的偏移的符号可以具有相反值(用于偏移的相反符号)。否则,如果图片参考列表1中的参考POC与当前帧之间的差大于图片参考列表0中的参考POC与当前帧之间的差,则表3中的符号可以指定添加到与图片参考列表1相关联的参考MV的MV偏移的符号,且用于与图片参考列表0相关联的参考MV的偏移的符号具有相反值。

[0190] 表3—由方向索引指定的用于MV偏移的符号的示例实施方式

[0191]	方向IDX	00	01	10	11
	X轴(水平)	+	-	不适用	不适用
	Y轴(垂直)	不适用	不适用	+	-

[0192] 在一些示例实施方式中,可以根据每一个方向上的POC的差来对MVD进行缩放。如果两个列表中的POC的差相同,则不需要进行缩放。否则,如果参考列表0中的POC的差大于参考列表1中的POC的差,则对用于参考列表1的MVD进行缩放。如果参考列表1的POC差大于参考列表0的POC差,则可以以相同的方式对用于列表0的MVD进行缩放。如果起始MV是单预测的,则将MVD添加到可用或参考MV。

[0193] 在用于双向复合预测的MVD编码和信令的一些示例实施方式中,除了或可选地单独地编码和发信号通知两个MVD,可以实施对称MVD编码,使得仅一个MVD需要信令,而另一个MVD可以从发信号通知的MVD导出。在这种实施方式中,发信号通知包括列表-0和列表-1的参考图片索引的运动信息。然而,只发信号通知与例如参考列表-0相关联的MVD,且不发信号通知而导出与参考列表-1相关联的MVD。具体地,在条带级处,标志可以被包括在码流中,称作“mvd_l1_zero_flag”,用于指示在码流中是否不发信号通知参考列表-1。如果该标志为1,指示参考列表-1等于0(并且因此未被发信号通知),那么双向预测标志,称作

“BiDirPredFlag”,可以被设置为0,意味着不存在双向预测。否则,如果mvd_l1_zero_flag为零,则如果列表-0中的最近参考图片和列表-1中的最近参考图片形成前向和后向参考图片对或后向和前向参考图片对,则BiDirPredFlag可以被设置为1,并且列表-0和列表-1参考画面都是短期参考图片。否则BiDirPredFlag被设置为0。BiDirPredFlag为1可以指示在码流中附加地发信号通知对称模式标志。当BiDirPredFlag为1时,解码器可以从码流中提取对称模式标志。例如,可以在CU级处发信号通知(如果需要)对称模式标志,并且它可以指示对称MVD编码模式是否正用于对应的CU。当对称模式标志为1时,其指示对称MVD编码模式的使用,并且只有列表-0和列表-1两者的参考图片索引(称作“mvp_l0_flag”和“mvp_l1_flag”)用与列表-0相关联的MVD(称作“MVD0”)来发信号通知,并且其它运动矢量差“MVD1”将被导出而不是被发信号通知。例如,MVD1可以被导出为-MVD0。这样,在示例对称MVD模式中仅发信号通知一个MVD。在用于MV预测的一些其它示例实施方式中,可以使用协调方案来针对单个参考模式和复合参考模式MV预测两者实施一般合并模式、MMVD和一些其它类型的MV预测。可以使用各种语法元素来发信号通知预测用于当前块的MV的方式。

[0194] 例如,对于单个参考模式,可以发信号通知以下MV预测模式:

[0195] NEARMV-在没有任何MVD的情况下,直接使用由动态参考列表(DRL)索引指示的列表中的运动矢量预测子(MVP)中的一个。

[0196] NEWMV-使用由DRL索引发信号通知的列表中的运动矢量预测子(MVP)中的一个作为参考,并将增量应用于MVP(例如,使用MVD)。

[0197] GLOBALMV-使用基于帧级全局运动参数的运动矢量。

[0198] 同样地,对于使用对应于待预测的两个MV的两个参考帧的复合参考帧间预测模式,可以发信号通知以下MV预测模式:

[0199] NEAR_NEARMV-对于待预测的两个MV中的每一个,在没有任何MVD的情况下,使用由DRL索引发信号通知的列表中的运动矢量预测子(MVP)中的一个。

[0200] NEAR_NEWMV-为了预测两个运动矢量中的第一个,使用由DRL索引发信号通知的列表中的运动矢量预测子(MVP)中的一个作为不具有MVD的参考MV;为了预测两个运动矢量中的第二个,使用由DRL索引发信号通知的列表中的运动矢量预测子(MVP)中的一个作为参考MV,并结合附加地发信号通知的增量MV(MVD)。

[0201] NEW_NEARMV-为了预测两个运动矢量中的第二个,使用由DRL索引发信号通知的列表中的运动矢量预测子(MVP)中的一个作为不具有MVD的参考MV;为了预测两个运动矢量中的第一个,使用由DRL索引发信号通知的列表中的运动矢量预测子(MVP)中的一个作为参考MV,并结合附加地发信号通知的增量MV(MVD)。

[0202] NEW_NEWMV-使用由DRL索引发信号通知的列表中的运动矢量预测子(MVP)中的一个作为参考MV,且将其与附加地发信号通知的增量MV结合使用以预测两个MV中的每一个。

[0203] GLOBAL_GLOBALMV-基于它们的帧级全局运动参数使用来自每一个参考的MV。

[0204] 因此,以上术语“NEAR”指的是使用不具有MVD的参考MV作为一般合并模式的MV预测,而术语“NEW”指的是涉及使用参考MV并如在MMVD模式中那样用发信号通知的MVD来偏移它的MV预测。对于复合帧间预测,以上参考基本运动矢量和运动矢量增量在两个参考之间通常可以是不同的或独立的,即使它们可以是相关的,并且可以利用这种相关性来减少发信号通知两个运动矢量增量所需的信息量。在这种情况下,可以在码流中实施并指示两个

MVD的联合信令。

[0205] 以上动态参考列表(DRL)可以用于保存被动态地维持并被视为候选运动矢量预测子的索引运动矢量集。

[0206] 在一些示例实施方式中,可以允许MVD的预定义的分辨率。例如,可以允许1/8像素的运动矢量精度(或准确度)。可以以各种方式构建和发信号通知上述各种MV预测模式中的MVD。在一些实施方式中,可以使用各种语法元素来发信号通知参考帧列表0或列表1中的以上一个或多个运动矢量差。

[0207] 例如,称作“mv_joint”的语法元素可以指定与其相关联的运动矢量差的哪些分量是非零的。对于MVD,这对于所有非零分量被联合地发信号通知。例如,mv_joint具有值

[0208] 0可以指示沿水平方向或垂直方向不存在非零MVD;

[0209] 1可以指示仅沿水平方向存在非零MVD;

[0210] 2可以指示仅沿垂直方向存在非零MVD;

[0211] 3可以指示沿水平方向和垂直方向两者存在非零MVD。

[0212] 当用于MVD的“mv_joint”语法元素发信号通知不存在非零MVD分量时,则不再发信号通知MVD信息。然而,如果“mv_joint”语法发信号通知存在一个或两个非零分量,那么如下文所述,可以针对非零MVD分量中的每一个进一步发信号通知附加语法元素。

[0213] 例如,可以使用称作“mv_sign”的语法元素来附加地指定对应的运动矢量差分量是正的还是负的。

[0214] 对于另一示例,可以使用称作“mv_class”的语法元素指定用于对应的非零MVD分量的预定义的类别集中的运动矢量差的类别。例如,可以使用用于运动矢量差的预定义的类别,将运动矢量差的连续量值空间划分为非重叠范围,其中每一个范围对应于MVD类别。因此,发信号通知的MVD类别指示对应的MVD分量的量值范围。在以下表4所示的示例实现中,较高的类别对应于具有较大量值范围的运动矢量差。在表4中,符号(n,m]用于表示大于n个像素且小于或等于m个像素的运动矢量差的范围。

[0215] 表4:用于运动矢量差的量值类别

	MV 类别	MVD 的量值
	MV_CLASS_0	(0, 2]
[0216]	MV_CLASS_1	(2, 4]
	MV_CLASS_2	(4, 8]
	MV_CLASS_3	(8, 16]

[0217]	MV_CLASS_4	(16, 32]
	MV_CLASS_5	(32, 64]
	MV_CLASS_6	(64, 128]
	MV_CLASS_7	(128, 256]
	MV_CLASS_8	(256, 512]
	MV_CLASS_9	(512, 1024]
	MV_CLASS_10	(1024, 2048]

[0218] 在一些其它示例中,称作“mv_bit”的语法元素可以进一步用于指定非零运动矢量差分量与对应地发信号通知的MV类别量值范围的起始量值之间的偏移的整数部分。这样,mv_bit可以指示MVD的量值或幅度。在“mv_bit”中发信号通知每一个MVD类别的全范围所需的比特数可以作为MV类别的函数而变化。例如,表4的实施方式中的MV_CLASS 0和MV_CLASS1可能仅需要单个比特来指示来自0的起始MVD的1或2的整数像素偏移;表4的示例实施方式中的每一个较高的MV_CLASS可能需要比先前的MV_CLASS逐步多一个比特用于“mv_bit”。

[0219] 在一些其它示例中,称作“mv_fr”的语法元素可以进一步用于指定用于对应的非零MVD分量的运动矢量差的前2个小数比特,而称作“mv_hp”的语法元素可以用于指定用于对应的非零MVD分量的运动矢量差的第三个小数比特(高分辨率比特)。2-比特“mv_fr”实质上提供1/4像素MVD分辨率,而“mv_hp”比特可以进一步提供1/8像素分辨率。在一些其它实施方式中,可以使用多于一个的“mv_hp”比特来提供比1/8像素更精细的MVD像素分辨率。在一些示例实施方式中,可以在各个级中的一个或多个级处发信号通知附加标志以指示是否支持1/8像素或更高的MVD分辨率。如果MVD分辨率未应用于特定编码单元,则可以不发信号通知以上用于对应的非支持的MVD分辨率的语法元素。

[0220] 在以上的一些示例实施方式中,小数分辨率可以独立于不同类别的MVD。换句话说,无论运动矢量差的量值如何,可以使用预定义的数量值的“mv_fr”和“mv_hp”比特来提供用于运动矢量分辨率的类似选项,以用于发信号通知非零MVD分量的小数MVD。

[0221] 然而,在一些其它示例实施方式中,可以区分各种MVD量值类别中的运动矢量差的分辨率。具体地,用于较高MVD类别的大MVD量值的高分辨率MVD可能不提供压缩效率的统计上显著的改进。这样,对于对应于较高MVD量值类别的较大MVD量值范围,可以以降低的分辨率(整数像素分辨率或小数像素分辨率)对MVD进行编码。同样地,通常对于较大的MVD值,可以以降低的分辨率(整数像素分辨率或小数像素分辨率)对MVD进行编码。这种依赖于MVD类别或依赖于MVD量值的MVD分辨率通常可以称作自适应MVD分辨率、依赖于幅度的自适应MVD分辨率或依赖于量值的MVD分辨率。术语“分辨率”可以进一步称作“像素分辨率”。自适应MVD分辨率可以如以下示例实施方式所描述的那样以各种方式来实现,以实现总体更好的压缩效率。特别地,由于以非自适应方式以与低量值或低类别MVD的分辨率类似的水平处理大量值或高类别MVD的MVD分辨率可能不会显著地增加具有大量值或高类别MVD的块的帧间预测残差编码效率的统计观察,通过瞄准较不精确的MVD来减少信令比特的数量可能大于由于这种较不精确的MVD而导致的编码帧间预测残差所需的附加比特。换句话说,对大量值

或高类别MVD使用较高的MVD分辨率可能不会比使用较低的MVD分辨率产生更多的编码增益。

[0222] 在一些一般的示例实施方式中,MVD的像素分辨率或精度可以随着MVD类别的增加而降低或不增加。降低MVD的像素分辨率对应于较粗的MVD(或从一个MVD级到下一个的较大步长)。在一些实施方式中,可以指定、预定义或预配置MVD像素分辨率和MVD类别之间的对应关系,并且因此可以不需要在已编码码流中发信号通知。

[0223] 在一些示例实施方式中,表3的MV类别可以各自与不同的MVD像素分辨率相关联。

[0224] 在一些示例实施方式中,每一个MVD类可以与单个允许的分辨率相关联。在一些其它实施方式中,一个或多个MVD类别可以与两个或更多个可选的MVD像素分辨率相关联。因此,具有这种MVD类别的当前MVD分量的码流中的信号可以跟随附加信令,用于指示为当前MVD分量选择的可选的像素分辨率。

[0225] 在一些示例实施方式中,自适应地允许的MVD像素分辨率可以包括但不限于1/64-pel(像素)、1/32像素、1/16像素、1/8像素、1-4像素、1/2像素、1像素、2像素、4像素……(按分辨率的降序)。这样,每一个上升的MVD类别可以以非上升的方式与这些MVD像素分辨率中的一个相关联。在一些实施方式中,MVD类别可以与以上两个或更多个分辨率相关联,并且较高的分辨率可以低于或等于先前MVD类别的较低分辨率。例如,如果表4的MV_CLASS_3与可选的1像素和2像素分辨率相关联,则表4的MV_CLASS_4所能关联的最高分辨率将是2像素。在一些其它实施方式中,MV类别的最高可允许分辨率可以高于先前(较低)MV类别的最低可允许分辨率。然而,上升MV类别的允许的分辨率的平均值可能仅是非上升的。

[0226] 在一些实施方式中,当允许小数像素分辨率高于1/8像素时,“mv_fr”和“mv_hp”信令可以对应地扩展到总共多于3个小数比特。

[0227] 在一些示例实施方式中,小数像素分辨率可以仅允许低于或等于阈值MVD类别的MVD类别。例如,小数像素分辨率可以仅允许MVD-CLASS 0而不允许表4中的所有其它MV类别。同样地,小数像素分辨率可以仅允许低于或等于表4的其它MV类别中的任一个的MVD类别。对于高于阈值MVD类别的其它MVD类别,仅允许MVD的整数像素分辨率。以这种方式,诸如“mv_fr”和/或“mv_hp”比特中的一个或多个比特的小数分辨率信令可能不需要以高于或等于阈值MVD类别的MVD类别发信号通知的MVD发信号通知。对于分辨率低于1像素的MVD类别,可以进一步减少“mv-bit”信令中的比特数。例如,对于表4中的MV_CLASS_5,MVD像素偏移的范围是(32,64],因此需要5个比特来以1像素分辨率发信号通知整个范围。然而,如果MV_CLASS_5与2像素MVD分辨率(比1像素分辨率更低的分辨率)相关联,则“mv_bit”可能需要4个比特而不是5个比特,并且“mv_fr”和“mv_hp”中没有一个需要在作为MV-CLASS_5的“mv_class”的信令之后被发信号通知。

[0228] 在一些示例实施方式中,小数像素分辨率仅可以允许整数值低于阈值整数像素值的MVD。例如,小数像素分辨率仅可以允许小于5个像素的MVD。对应于该示例,分数分辨率可以允许表4的MV_CLASS_0和MV_CLASS_1,且不允许所有其它MV类别。对于另一示例,小数像素分辨率仅可以允许小于7个像素的MVD。对应于该示例,分数分辨率可以允许表4的MV_CLASS_0和MV_CLASS_1(范围低于5个像素),且不允许MV_CLASS_3和更高的(范围高于5个像素)。对于属于MV_CLASS_2的MVD,其像素范围包含5个像素,取决于“mv-bit”值,可以或可以允许MVD的小数像素分辨率。如果发信号通知“m-bit”值为1或2(使得发信号通知的MVD的整

数部分为5或6,计算为具有由“m-bit”指示的偏移1或2的MV_CLASS_2的像素范围的起始),则可以允许小数像素分辨率。否则,如果发信号通知“mv-bit”值为3或4(使得发信号通知的MVD的整数部分为7或8),则可以不允许小数像素分辨率。

[0229] 在一些其它实施方式中,对于等于或高于阈值MV类别的MV类别,可以仅允许单个MVD值。例如,这种阈值MV类别可以是MV_CLASS2。因此,MV_CLASS_2及以上仅可以允许具有单个MVD值且没有小数像素分辨率。可以预定义这些MV类别的单个允许的MVD值。在一些示例中,允许的单个值可以是表4中的这些MV类别的相应范围的较高端值。例如,MV_CLASS_2到MV_CLASS_10可以高于或等于MV_CLASS2的阈值类别,且可以预定义这些类别的单个允许的MVD值分别为8、16、32、64、128、256、512、1024和2048。在一些其他示例中,允许的单个值可以是表4中的这些MV类别的相应范围的中间值。例如,MV_CLASS_2到MV_CLASS_10可以高于类别阈值,且可以预定义这些类别的单个允许的MVD值分别为3、6、12、24、48、96、192、384、768和1536。范围内的任何其它值也可以被定义为相应MVD类别的单个允许的分辨率。

[0230] 在以上实施方式中,当发信号通知的“mv_class”等于或高于预定义的MVD类别阈值时,只有“mv_class”信令足以确定MVD值。然后将使用“mv_class”和“mv_sign”来确定MVD的量值和方向。

[0231] 这样,当仅针对一个参考帧(来自参考帧列表0或列表1,但不是两者)发信号通知MVD,或针对两个参考帧共同发信号通知MVD时,MVD的精度(或分辨率)可以取决于表3中的运动矢量差的相关联类别和/或MVD的量值。

[0232] 在一些其它实施方式中,MVD的像素分辨率或精度可以随着MVD量值的增加而降低或不增加。例如,像素分辨率可以取决于MVD量值的整数部分。在一些实施方式中,小数像素分辨率可以仅允许小于或等于幅度阈值的MVD量值。对于解码器,首先可以从码流中提取MVD量值的整数部分。接着可以确定像素分辨率,并且随后可以做出关于任何小数MVD是否存在于码流中且需要被解析的决策(例如,如果小数像素分辨率不允许特定提取的MVD整数量值,那么在需要提取的码流中可以不包括小数MVD比特)。与依赖于MVD类别的自适应MVD像素分辨率相关的以上示例实施方式适用于依赖于MVD量值的自适应MVD像素分辨率。对于特定示例,可以允许高于或包含量值阈值的MVD类别仅具有一个预定义的值。以上各种示例实施方式适用于单参考模式。这些实施方式也适用于MMVD下的复合预测中的示例NEW_NEARMV、NEAR_NEWMV和/或NEW_NEWMV模式。这些实施方式通常适用于用于任何MVD的自适应分辨率。

[0233] 在自适应MVD像素分辨率的特定示例实施方式中,用于1以下的MVD量值的MVD像素分辨率可以是分数的,对于MV_CLASS_1及以上的MV类别,可以仅允许使用等于表4的对应的MVD量值范围的结束值的单个MVD量值。在这种示例中,在表4中指示针对1/8、1/4或1/2像素的允许的分数像素分辨率的允许的MVD值。

[0234] 表5

[0235]	MV 类别	允许的 MVD 值
--------	-------	-----------

[0236]

MV_CLASS_0	自适应 MVD 分辨率方案中分数像素分辨率为 1/8 像素时 (1/8、2/8、3/8、4/8、5/8、6/8、7/8、1、2)； 自适应 MVD 分辨率方案中分数像素分辨率为 1/4 像素时 (2/8、4/8、6/8、1、2)； 自适应 MVD 分辨率方案中分数像素分辨率为 1/2 像素时 (1/2、1、2)； 自适应 MVD 分辨率方案中像素分辨率为 1 像素时 (1、2)。
MV_CLASS_1	4
MV_CLASS_2	8
MV_CLASS_3	16
MV_CLASS_4	32
MV_CLASS_5	64
MV_CLASS_6	128
MV_CLASS_7	256
MV_CLASS_8	512
MV_CLASS_9	1024
MV_CLASS_10	2048

[0237] 对于编码块,可以显式地或隐式地发信号通知(导出)是否使用自适应MVD像素分辨率。当发信号通知未使用自适应MVD像素分辨率时,表示不同的MVD类别可以遵循表4中指示的MVD范围,并且可以定义或发信号通知非自适应MVD像素分辨率。这种非自适应分辨率可以是分数(诸如1/8、1/4或1/2像素)或非分数(诸如1、2、4……像素),并且将适用于所有MVD类别。非自适应分辨率基本上确定发信号通知上述mv_bit、mv_fr和mv_hp所需的比特数。当非自适应分辨率是分数时,其仅可以确定发信号通知用于所有MVD类别的mv_fr和mv_hp所需的比特数(与MVD的类别无关),发信号通知Mv_bit的比特数可以取决于MVD类别。

[0238] 当发信号通知使用自适应MVD像素分辨率时,可以预定义或发信号通知诸如表5中所示的自适应方式允许的MVD级或值。例如,根据自适应MVD分辨率的特定方案,可以以各种方式在码流中发信号通知它们。在表5的示例中,可以使用信令语法集指示分数分辨率(例如,1/8像素),低于该分数分辨率的量值阈值,发信号通知的分数分辨率适用(例如,1像素的MVD量值)。可以使用另一语法集(可能更复杂)发信号通知其它自适应MVD分辨率方案。自适应MVD像素分辨率方案的这种指示可以在诸如序列级、图片级、帧级、条带级、超块级或编码块级的各种编码级中的一个发信号通知。

[0239] 在一些示例实施方式中,可以在特定编码级(例如,序列级、图片级、帧级、条带级、超块级)定义或发信号通知整体自适应MVD像素分辨率方案,包括但不限于表5中所示。这种自适应MVD像素分辨率方案可以在相同编码级或另一编码级进一步修改,使得可以在相同编码级或另一编码级调整或修改用于各种MVD类别的允许的MVD像素分辨率值。如果在特定编码级没有进行调整,则不加修改地应用发信号通知的或预定义的自适应MVD像素分辨率方案。例如,可以在帧级定义或发信号通知总体自适应MVD像素分辨率方案,而可以在一个

或多个超块级或编码块级进行调整,反之亦然。

[0240] 这种调整可以实现为MVD精度的限制或MVD精度的扩展。与这种调整相关联的信息可以预定义或发信号通知。预定义的调整可以适用于所有编码块。可选地,可以通过信令在各种编码级激活预定义的调整。

[0241] 在一些实施方式中,这种调整可以体现为最大允许的MVD精度。对于特定编码块,当应用自适应MVD分辨率时,这种最大允许的MVD精度可以不同于在图片级、或超块级、或已编码块级指定/发信号通知/导出的自适应MVD像素分辨率方案的MVD像素精度,如上所述。在这种情况下,可以通过采用由自适应MVD像素分辨率方案指定或导出的允许的值和最大允许的MVD精度两者,确定用于各种MVD类别的允许的MVD分辨率值。

[0242] 例如,假设针对某一编码级,预定义/发信号通知/导出表5的自适应MVD像素分辨率方案。进一步假设最大允许的精度是1/4像素,这意味着,无论与表5相关联的自适应MVD分辨率如何,任何MVD类别都不允许使用等于或高于1/8像素的精度。然后,通过将最大允许的像素精度作为限制无差别地应用到表5,可以将用于各种MVD类别的允许的MVD像素级或值修改为:

[0243] 表6

	MV 类别	允许的 MVD 值
[0244]	MV_CLASS_0	自适应方案中分数像素分辨率为 1/8 像素时（2/8、4/8、6/8、1、2）；
		自适应方案中分数像素分辨率为 1/4 像素时（1/4、2/4、3/4、1、2）；
		自适应方案中分数像素分辨率为 1/2 像素时（1/2、1、2）；
		自适应方案中像素分辨率为 1 像素时（1、2）。
[0245]	MV_CLASS_1 至 MV_CLASS_10	与表 5 相同

[0246] 通过定义/发信号通知最大允许的MVD像素精度为1/4像素,而不允许所有MVD类别使用1/8或更高的像素精度,仅仅是一个示例。在另一示例中,可以定义/发信号通知最大允许的像素精度为1/2像素。以上用于MV_CLASS_0的对应的允许的MVD值可以变成:(1/2、1、2)用于分数像素分辨率为1/8像素、1/4像素和1/2像素的自适应分辨率方案,(1、2)用于像素分辨率为1像素的自适应分辨率方案。

[0247] 如以上关于表6的实施方式中所举例说明的,当自适应MVD分辨率的应用是在特定编码级使用定义的/发信号通知的/导出的自适应分辨率方案,并且在相同或不同的编码级使用附加的定义的/发信号通知的最大允许的MVD精度时,可能需要/限制这种最大允许的MVD精度不大于自适应分辨率方案中的MVD分辨率。换句话说,通过考虑自适应分辨率方案和最大允许的精度两者导出的实际应用的MVD精度,是由自适应分辨率方案中的MVD分辨率进行截取(即,当最大允许的精度大于由自适应分辨率方案定义的/发信号通知的/导出的分辨率时,最大允许的精度将是无效的)。

[0248] 然而,在一些其它实施方式中,可以不需要这种截取,并且定义的/发信号通知的最大允许的MVD精度可以控制用于至少一些MVD类别的实际MVD分辨率。在这些实施方式中,当应用自适应MVD分辨率时(如通过在各种编码级的定义/发信号通知/导出所指示的,如上所述),对于用于至少一些MVD类别的MVD级的调整,可以涉及增大而不是限制在诸如与表5相关联的自适应MVD像素分辨率方案中定义的/发信号通知的/导出的自适应MVD分辨率。例如,可以进行调整,以允许使用高于指定的/发信号通知的/从MVD类别的自适应分辨率方案导出的精度,所述MV类别等于或低于定义的或发信号通知的阈值MVD类别级。如上所述,可以定义/发信号通知这种较高精度为最大允许的MVD精度。无论自适应分辨率方案中指定的/发信号通知的/导出的MVD分辨率如何,这种最大允许的精度可以施加在等于或低于阈值MVD类别级。具体地,这种阈值MVD类别级可以是(但不必是)MV_CLASS_0(或MVD类别集的最低MVD类别级,诸如表5中的MVD类别集)。最大允许的像素精度可以预定义/发信号通知。最大允许的像素精度可以是分数。对于特定示例,在表5的自适应分辨率方案中阈值MVD类别为MV_CLASS_0的情况下,当用于MV_CLASS_0的MVD像素分辨率为非分数1像素,且用于调整的最大允许的分数像素精度为1/8、1/4或1/2像素时,调整允许MVD值将为:

[0249] 表7

[0250]

MV 类别	允许的 MVD 值
MV_CLASS_0	自适应分辨率方案中像素分辨率为 1 像素时: 最大允许的分数分辨率为 1/8 像素时(1/8、 2/8、 3/8、 4/8、 5/8、 6/8、 7/8、 1、 2) ; 最大允许的分数分辨率为 1/4 像素时 (1/4、 2/4、 ¾、 1、 2) ; 最大允许的分数分辨率为 1/2 像素时 (1/2、 1、 2) ; 没有定义/发信号通知最大允许的分辨率时 (1、 2) 。
MV_CLASS_1 至 MV_CLASS_10	与表 5 相同

[0251] 在替代表7的一些示例实施方式中,可以使用阈值MVD幅度代替阈值MVD类别级。在这些实施方式中,对于量值等于或低于阈值MVD量值而不是阈值MVD类别级的MVD,指定的/发信号通知的最大允许的MVD精度可以施加更高的精度。在这种实施方式中,除了mv_class信息之外,mv_bit信息可以在视频流中足够早地发信号通知,使得可以及时确定MVD的量值,以确定允许的MVD值。例如,通过用1/2像素的阈值MVD量值替换阈值MVD类别,并且仍假设用于MV_CLASS_0的自适应MVD分辨率在自适应分辨率方案中为1像素,表7将变成下表8:

[0252] 表8

	MV 类别	允许的 MVD 值
[0253]	MV_CLASS_0	自适应分辨率方案中像素分辨率为 1 像素时： 最大允许的分数分辨率为 1/8 像素时（1/8、2/8、3/8、1/2、1、2）； 最大允许的分数分辨率为 1/4 像素时（1/4、1/2、1、2）； 最大允许的分数分辨率为 1/2 像素时（1/2、1、2）； 没有定义/发信号通知最大允许的分辨率时（1、2）。
	MV_CLASS_1	与表 5 相同
[0254]	至 MV_CLASS_10	

[0255] 在一些其它示例性实施方式中，以上调整可以涉及当MVD的量值等于或低于阈值MVD量值时，仅允许使用特定精度和较低的精度（例如分数精度1/8、1/4或1/2及更低）。在这种实施方式中，同样，除了mv_class信息之外，mv_bit信息可以在视频流中足够早地发信号通知，使得可以及时确定MVD的量值，以确定允许的MVD值。

[0256] 在这种实施方式中，不会对从自适应分辨率方案（诸如表5）导出的MVD值施加附加的分辨率。相反，当MVD的幅度高于阈值MVD量值时，可以不允许使用与等于或高于定义的/发信号通知的精度级的分辨率相关联的MVD值。再次假设，在表5的示例中，并且进一步假设对于高于1/2像素的阈值MVD量值的MVD量值，不允许使用与等于或高于1/8像素精度的定义的/发信号通知的精度度的分辨率相关联的MVD值。然后表5将调整为：

[0257] 表9

	MV 类别	允许使用的 MVD 值
[0258]	MV_CLASS_0	自适应方案中分数像素分辨率为 1/8 像素时（1/8、2/8、3/8、1/2、6/8、1、2）； 自适应方案中分数像素分辨率为 1/4 像素时（1/2、1/4、3/4、1、2）； 自适应方案中分数像素分辨率为 1/2 像素时（1/2、1、2）； 自适应方案中像素分辨率为 1 像素时（1、2）。
	MV_CLASS_1 至 MV_CLASS_10	与表 5 相同

[0259] 具体地，如上所示，用于MV_CLASS_0和1/8像素的分数分辨率的允许的MVD值（1/8、2/8、3/8、1/2、5/8、6/8、7/8、1、2）调整为（1/8、2/8、3/8、1/2、6/8、1、2），其中与1/8精度相关联的MVD值仅在等于或小于1/2像素的阈值MVD量值允许使用并保持。超过1/2像素量值，不允许使用与1/8精度相关联的MVD值，诸如5/8像素值和7/8像素值。

[0260] 类似地,在表5的示例中,假设当MVD量值高于1/2像素的阈值量值时,不允许使用与等于或高于1/4像素精度的定义的/发信号通知的精度分辨率相关联的MVD值。然后表5将调整为:

[0261] 表9

MV 类别	允许的 MVD 值
[0262] MV_CLASS_0	自适应方案中分数像素分辨率为 1/8 像素时 (1/8、2/8、3/8、1/2、1、2) ; 自适应方案中分数像素分辨率为 1/4 像素时 (1/4、1/2、1、2) ; 自适应方案中分数像素分辨率为 1/2 像素时 (1/2、1、2) ; 自适应方案中像素分辨率为 1 像素时 (1、2) ;
MV_CLASS_1 至 MV_CLASS_10	与表 5 相同

[0263] 在以上的一些实施方式中,阈值MVD量值可以是2个像素或更小,诸如在以上示例中给出的1/2像素量值阈值。

[0264] 上述示例实施方式是关于特定MVD的描述,而不管帧间预测模式是单个参考模式还是复合参考模式。在复合参考模式(其中通过多个参考帧预测MV)的一些其它示例实施方式中,可以使用定义/信令集指示是否应用自适应MVD分辨率以及其应用于多个参考帧中的哪个参考帧或哪些参考帧。

[0265] 在一些示例实施方式中,当对多个参考帧发信号通知MVD时,可以发信号通知一个(或多个)标志/索引,以指示是否应用自适应MVD分辨率。

[0266] 例如,当对多个参考帧发信号通知MVD时(例如,在上述NEW_NEWMV模式或其它复合参考帧间预测模式),可以在视频流中发信号通知一个标志/索引,以指示自适应MVD分辨率是否应用于所有多个参考帧的MVD的信令。如果该标志/索引为1(或0),则其指示自适应MVD分辨率应用于所有多个参考帧的MVD的信令。否则,如果该标志/索引为0(或1),则自适应MVD编码不应用于多个参考帧中任何一个的MVD的信令。在这种实施方式中,对于多个帧间预测参考帧,自适应MVD分辨率应用于全有或全无方案。

[0267] 在一些其它示例中,当为多个参考帧发信号通知MVD时(例如,在上述用于双参考帧复合帧间预测模式的NEW_NEWMV模式或其它复合帧间预测模式下),可以单独为每一参考帧发信号通知一个标志/索引,以指示自适应MVD分辨率是否应用于每一参考帧。在这种实施方式中,可以为参考帧中的每一个单独确定是否应用自适应MVD分辨率。可以在编码器处为多个参考帧中的每一个独立地做出是否应用自适应MVD分辨率的决定,并且在视频流中单独发信号通知。

[0268] 在一些示例实施方式中,当为多个参考帧发信号通知MVD时,对于多个参考帧中的每一个,如果该参考帧的MVD为非零,则可以发信号通知一个标志/索引,以指示自适应MVD分辨率是否应用于该参考帧。否则,不需要发信号通知标志/索引。换句话说,如果发信号通知/指示用于特定参考帧的MVD为零,则不需要确定是否应用自适应MVD分辨率,并且因此视

频流中不需要任何相应的信令。。然而,在这种实施方式中,在做出是否应用自适应分辨率的决定之前,需要发信号通知MVD为零的指示。

[0269] 进一步转向MVD分辨率的信令,在一些示例实施方式中,可以发信号通知标志/索引,以显式地指示用于当前编码块的MVD分辨率,并且用于对这种标志/索引进行熵编码的上下文可以取决于与MVD相关联的MVD类别。这种标志/索引可以是用于导出诸如表5的自适应分辨率方案的任何MVD分辨率,或上述最大允许的MVD精度。

[0270] 在关于MVD分辨率的信令的一些示例实施方式中,可以单独发信号通知MVD的各个分量。MVD例如可以包括水平分量和垂直分量。可以为MVD的水平分量和垂直分量中的每一个发信号通知标志/索引,以分别指示水平分量和垂直分量的MVD分辨率。

[0271] 在一些示例实施方式中,可以在MVD类别信息之后,发信号通知MVD分辨率标志/索引。根据发信号通知的诸如MV_CLASS_0、MV_CLASS_1、MV_CLASS_2等MVD类别信息的值,可以导出上下文值,并将其用于发信号通知用于指示MVD分辨率的MVD分辨率标志/索引。换句话说,可以使用用于不同MVD类别或不同MVD类别组的不同上下文,对用于MVD分辨率的信令的一个或多个语法进行熵编码。

[0272] 图18示出了遵循以上自适应MVD分辨率的实施方式的原理的示例方法的流程图1800。示例解码方法流程从S1801处开始。在S1810中,接收视频流。在S1820中,基于预测块和运动矢量(MV),确定视频块是帧间编码,其中,MV待从用于视频块的参考运动矢量(RMV)和运动矢量差(MVD)导出。在S1830中,响应于确定MVD是用自适应MVD像素分辨率编码的:确定用于当前视频块的参考MVD像素精度;识别最大允许的MVD像素精度;基于参考MVD像素精度和最大允许的MVD像素精度,确定用于当前视频块的允许MVD级集;以及根据在视频流中发信号通知的用于当前视频块的至少一个MVD参数和允许MVD级集,从视频流中导出MVD。示例方法在S1899处停止。

[0273] 图19示出了遵循以上自适应MVD分辨率的实施方式的原理的另一示例方法的流程图1900。示例解码方法流程从S1901处开始。在S1910中,接收视频流。在S1920中,确定当前视频块是帧间编码且与多个参考帧相关联。在S1930中,进一步基于视频流中的信令,确定自适应运动矢量差(MVD)像素分辨率是否应用于多个参考帧中的至少一个。示例方法在S1999处停止。

[0274] 图20示出了遵循以上自适应MVD分辨率的实施方式的原理的示例方法的流程图2000。示例解码方法流程从S2001处开始。在S2010中,接收视频流。在S2020中,基于预测块和运动矢量(MV),确定视频块是帧间编码,其中,MV待从用于视频块的参考运动矢量(RMV)和运动矢量差(MVD)导出。在S2030中,从预定义MVD类别集中确定MVD的当前MVD类别。在S2040中,基于当前MVD类别,导出用于对视频流中的至少一个显式信令进行熵解码的至少一个上下文,至少一个显式信令包括在视频流中,以指定用于MVD的至少一个分量的MVD像素分辨率。在S2050中,使用至少一个上下文,对视频流中的至少一个显式信令进行熵解码,以确定用于MVD的至少一个分量的MVD像素分辨率。示例方法在S2099处终止。

[0275] 在本公开的实施例和实施方式中,任何步骤和/或操作可以根据需要以任何数量或顺序来组合或布置。可以并行地执行两个或更多个步骤和/或操作。本公开中的实施例和实施方式可以单独使用或以任何顺序组合使用。进一步地,方法(或实施例)、编码器和解码器中的每一个可以通过处理电路(例如,一个或多个处理器或一个或多个集成电路)来实

施。在一个示例中，一个或多个处理器执行存储在非易失性计算机可读介质中的程序。本公开中的实施例可以应用于亮度块或色度块。术语“块”可以被解释为预测块、编码块或编码单元（即CU）。术语“块”在这里也可以用来指变换块。在以下项目中，当说到块大小时，它可以指块宽度或高度、或宽度和高度的最大值、或宽度和高度的最小值、或区大小（宽度*高度）、或块的纵横比（宽度：高度、或高度：宽度）。

[0276] 上文所描述的技术可使用计算机可读指令实施为计算机软件且以物理方式存储在一个或多个计算机可读介质中。举例来说，图21示出适于实施所公开主题的某些实施例的计算机系统(2100)。

[0277] 所述计算机软件可使用任何合适的机器代码或计算机语言来编码，所述机器代码或计算机语言可经受汇编、编译、链接或类似机制以创建包括指令的代码，所述指令可直接或通过解译、微码执行等而由一个或多个计算机中央处理单元（CPU）、图形处理单元（GPU）等执行。

[0278] 所述指令可在各种类型的计算机或计算机组件上执行，所述计算机或计算机组件包括例如个人计算机、平板电脑、服务器、智能电话、游戏装置、物联网装置等。

[0279] 图21中所示的用于计算机系统(2100)的组件在本质上是示范性的，并非旨在暗示关于实施本申请实施例的计算机软件的使用或功能的范围的任何限制。也不应将组件的配置解释为对计算机系统(2100)的示范性实施例中所示的组件中的任一个组件或组件组合有任何依赖或需求。

[0280] 计算机系统(2100)可包括某些人机接口输入装置。此类人机接口输入装置可响应于一个或多个人类用户通过例如触觉输入（例如：按键、滑动、数据手套移动）、音频输入（例如：语音、拍击）、视觉输入（例如：手势）、嗅觉输入（未描绘）进行的输入。人机接口装置还可用于捕获未必与人的有意识输入直接相关的某些媒体，例如音频（例如：话语、音乐、环境声）、图像（例如：扫描图像、从静态图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）。

[0281] 输入人机接口装置可包括以下一个或多个（每种仅描绘一个）：键盘(2101)、鼠标(2102)、轨迹垫(2103)、触摸屏(2110)、数据手套（未示出）、操纵杆(2105)、麦克风(2106)、扫描仪(2107)、相机(2108)。

[0282] 计算机系统(2100)还可包括某些人机接口输出装置。此类人机接口输出装置可通过例如触觉输出、声音、光和气味/味道刺激一个或多个人类用户的感受。此类人机接口输出装置可包括触觉输出装置（例如，触摸屏(2110)、数据手套（未示出）或操纵杆(2105)的触觉反馈，但还可存在不充当输入装置的触觉反馈装置）、音频输出装置（例如：扬声器(2109)、头戴式耳机（未描绘））、视觉输出装置（例如，屏幕(2110)，包括阴极射线管(CRT)屏幕、液晶显示(LCD)屏幕、等离子体屏幕、有机发光二极管(OLED)屏幕，各自具有或不具有触摸屏输入能力，各自具有或不具有触觉反馈能力--其中的一些能够通过例如立体平画输出的方式输出二维视觉输出或大于三维的输出；虚拟现实眼镜（未描绘）、全息显示器和烟雾箱（未描绘），以及打印机（未描绘）。

[0283] 计算机系统(2100)还可包括人类可访问的存储装置和存储装置的相关联介质，例如，光学介质，包括具有CD/DVD等介质(2121)的CD/DVD ROM/RW(2120)、拇指驱动器(2122)、可移动硬盘驱动器或固态驱动器(2123)、磁带和软盘（未描绘）等旧版磁性媒体、基于ROM/

专用集成电路(ASIC)/可编程逻辑设备(PLD)的专用装置,例如,安全保护装置(未描绘),等等。

[0284] 所属领域的技术人员还应理解,结合当前公开的主题使用的术语“计算机可读介质”并未涵盖传输介质、载波或其它瞬时信号。

[0285] 计算机系统(2100)还可包括到一个或多个通信网络(2155)的接口(2154)。网络可例如是无线的、有线的、光学的。网络还可以是本地的、广域的、城域的、车载和工业的、实时的、容忍延迟的等等。网络的实例包括例如以太网、无线LAN的局域网、包括全球移动通信系统(GSM)、第三代(3G)、第四代(4G)、第五代(5G)、长期演进(LTE)等的蜂窝网络、包括有线TV、卫星TV和地面广播TV的TV有线或无线广域数字网络、包括控制器局域网总线(CAN Bus)的车载网络和工业网络等。某些网络通常需要附接到某些通用数据端口或外围总线(2149)(例如,计算机系统(2100)的通用串行总线(USB)端口)的外部网络接口适配器;其它网络通常通过附接到如下文所描述的系统总线而集成到计算机系统(2100)的核心中(例如,通过以太网接口集成到PC计算机系统中,或通过蜂窝网络接口集成到智能电话计算机系统中)。通过使用这些网络中的任一网络,计算机系统(2100)可与其它实体通信。此类通信可以是仅单向接收(例如,广播TV)、仅单向发送(例如,连到某些CANBus装置的CANBus)或是双向的,例如,使用局域数字网络或广域数字网络连接到其它计算机系统。可在如上文所描述的那些网络和网络接口中的每一个上使用某些协议和协议栈。

[0286] 上述人机接口装置、人类可访问存储装置和网络接口可附接到计算机系统(2100)的核心(2140)。

[0287] 核心(2140)可包括一个或多个中央处理单元(CPU)(2141)、图形处理单元(GPU)(2142)、现场可编程门区域(Field Programmable Gate Areas,FPGA)形式的专用可编程处理单元(2143)、用于某些任务的硬件加速器(2144)、图形适配器(2150)等等。这些装置连同只读存储器(read-only memory,ROM)(2145)、随机存取存储器(2146)、例如内部非用户可访问的硬盘驱动器、固态驱动器(SSD)等内部大容量存储装置(2147)可通过系统总线(2148)连接。在一些计算机系统中,系统总线(2148)可通过一个或多个物理插头形式访问以实现通过额外CPU、GPU等来扩展。外围装置可直接或通过外围总线(2149)附接到核心的系统总线(2148)。在一示例中,屏幕(2110)可以连接到图形适配器(2150)。用于外围总线的架构包括外围设备互连(PCI)、USB等等。

[0288] CPU(2141)、GPU(2142)、FPGA(2143)和加速器(2144)可执行某些指令,所述指令组合起来可构成上述计算机代码。计算机代码可存储在ROM(2145)或RAM(2146)中。过渡数据也可存储在RAM(2146)中,而永久性数据可例如存储在内部大容量存储装置(2147)中。可通过使用高速缓冲存储器来实现对任一存储器装置的快速存储和检索,所述高速缓冲存储器可与一个或多个CPU(2141)、GPU(2142)、大容量存储装置(2147)、ROM(2145)、RAM(2146)等紧密关联。

[0289] 计算机可读介质上可具有用于执行各种计算机实施的操作的计算机代码。所述介质和计算机代码可以是专为本申请的目的设计和构建的介质和计算机代码,或可属于计算机软件领域中的技术人员众所周知且可用的种类。

[0290] 作为非限制性示例,具有架构(2100)且尤其是核心(2140)的计算机系统可提供因处理器(包括CPU、GPU、FPGA、加速器等)执行以一个或多个有形计算机可读介质体现的软件

而产生的功能。此类计算机可读介质可以是与上文所介绍的用户可访问大容量存储装置以及核心(2140)的非暂时性质的某些存储装置(例如,核心内部大容量存储装置(2147)或ROM(2145))相关联的介质。实施本申请的各种实施例的软件可存储在此类装置中且由核心(2140)执行。根据特定需求,计算机可读介质可包括一个或多个存储器装置或芯片。软件可使核心(2140)且具体地说使其中的处理器(包括CPU、GPU、FPGA等等)执行本文中所描述的特定过程或特定过程的特定部分,包括限定存储在RAM(2146)中的数据结构以及根据由软件限定的过程修改此类数据结构。另外或作为替代方案,计算机系统可提供由硬连线的或以其它方式体现于电路(例如:加速器(2144))中的逻辑所产生的功能,所述逻辑可代替或连同软件一起操作以执行本文描述的特定过程或特定过程的特定部分。适当时,对软件的引用可涵盖逻辑,且反之亦然。适当时,对计算机可读介质的引用可涵盖存储用于执行的软件的电路(例如,集成电路(IC))、体现用于执行的逻辑的电路或这两种电路。本申请涵盖硬件与软件的任何合适的组合。

[0291] 尽管本申请描述了若干示范性实施例,但在本申请的范围内,可以有各种改动、排列组合方式以及各种替代等同物。因此,应该理解,在申请的精神和范围内,本领域技术人员能够设计出各种虽未在本文明确示出或描述、但可以体现本申请的原理的系统和方法。

[0292] 附录A:缩略词

[0293] JEM:Joint Exploration Model,联合开发模型

[0294] VVC:Versatile Video Coding,多功能视频编码

[0295] BMS:Benchmark Set,基准集合

[0296] MV:Motion Vector,运动矢量

[0297] HEVC:High Efficiency Video Coding,高效视频编码

[0298] SEI:Supplementary Enhancement Information,辅助增强信息

[0299] VUI:Video Usability Information,视频可用性信息

[0300] GOP:Groups of Pictures,图片群组

[0301] TU:Transform Unit,变换单元

[0302] PU:Prediction Unit,预测单元

[0303] CTU:Coding Tree Unit,编码树单元

[0304] CTB:Coding Tree Block,编码树块

[0305] PB:Prediction Block,预测块

[0306] HRD:Hypothetical Reference Decoder,假想参考解码器

[0307] SNR:Signal Noise Ratio,信噪比

[0308] CPU:Central Processing Unit,中央处理单元

[0309] GPU:Graphics Processing Unit,图形处理单元

[0310] CRT:Cathode Ray Tube,阴极射线管

[0311] LCD:Liquid-Crystal Display,液晶显示

[0312] OLED:Organic Light-Emitting Diode,有机发光二极管

[0313] CD:Compact Disc,光盘

[0314] DVD:Digital Video Disc,数字视频光盘

[0315] ROM:Read-Only Memory,只读存储器

- [0316] RAM:Random Access Memory,随机存取存储器
- [0317] ASIC:Application-Specific Integrated Circuit,专用集成电路
- [0318] PLD:Programmable Logic Device,可编程逻辑设备
- [0319] LAN:Local Area Network,局域网
- [0320] GSM:Global System for Mobile communications,全球移动通信系统
- [0321] LTE:Long-Term Evolution,长期演进
- [0322] CANBus:Controller Area Network Bus,控制器局域网络总线
- [0323] USB:Universal Serial Bus,通用串行总线
- [0324] PCI:Peripheral Component Interconnect,外围设备互连
- [0325] FPGA:Field Programmable Gate Array,现场可编程门阵列
- [0326] SSD:Solid-state drive,固态驱动器
- [0327] IC:Integrated Circuit,集成电路
- [0328] HDR:high dynamic range,高动态范围
- [0329] SDR:standard dynamic range,标准动态范围
- [0330] JVET:Joint Video Exploration Team,联合视频探索小组
- [0331] MPM:most probable mode,最可能模式
- [0332] WAIP:Wide-Angle Intra Prediction,广角帧内预测
- [0333] CU:Coding Unit,编码单元
- [0334] PU:Prediction Unit,预测单元
- [0335] TU:Transform Unit,变换单元
- [0336] CTU:Coding Tree Unit,编码树单元
- [0337] PDPC:Position Dependent Prediction Combination,位置决定的帧内预测组合
ISP:Intra Sub-Partitions,帧内子分区
- [0338] SPS:Sequence Parameter Setting,序列参数设置
- [0339] PPS:Picture Parameter Set,图片参数集
- [0340] APS:Adaptation Parameter Set,适配参数集
- [0341] VPS:Video Parameter Set,视频参数集
- [0342] DPS:Decoding Parameter Set,解码参数集
- [0343] ALF:Adaptive Loop Filter,自适应环路滤波器
- [0344] SAO:Sample Adaptive Offset,样本自适应偏移
- [0345] CC-ALF:Cross-Component Adaptive Loop Filter,跨分量自适应环路滤波器
CDEF:Constrained Directional Enhancement Filter,约束方向增强滤波器
CCS0:Cross-Component Sample Offset,跨分量样本偏移
- [0346] LS0:Local Sample Offset,局部样本偏移
- [0347] LR:Loop Restoration Filter,环路恢复滤波器
- [0348] AV1:AOMedia Video 1,开放媒体联盟视频1
- [0349] AV2:AOMedia Video 2,开放媒体联盟视频2
- [0350] MVD:Motion Vector difference,运动矢量差
- [0351] CfL:Chroma from Luma,根据亮度预测色度

- [0352] SDT:Semi Decoupled Tree,半解耦树
- [0353] SDP:Semi Decoupled Partitioning,半解耦分区
- [0354] SST:Semi Separate Tree,半单独树
- [0355] SB:Super Block,超块
- [0356] IBC(或IntraBC):Intra Block Copy,帧内块复制
- [0357] CDF:Cumulative Density Function,累积密度函数
- [0358] SCC:Screen Content Coding,屏幕内容编码
- [0359] GBI:Generalized Bi-prediction,广义双向预测
- [0360] BCW:Bi-prediction with CU-level Weights,具有CU级权重的双向预测
CIIP:Combined intra-inter prediction,帧内帧间联合预测
- [0361] POC:Picture Order Count,图片顺序计数
- [0362] RPS:Reference Picture Set,参考图片集
- [0363] DPB:Decoded Picture Buffer,已解码图片缓冲器
- [0364] MMVD:Merge Mode with Motion Vector Difference,具有运动矢量差的合并模式。

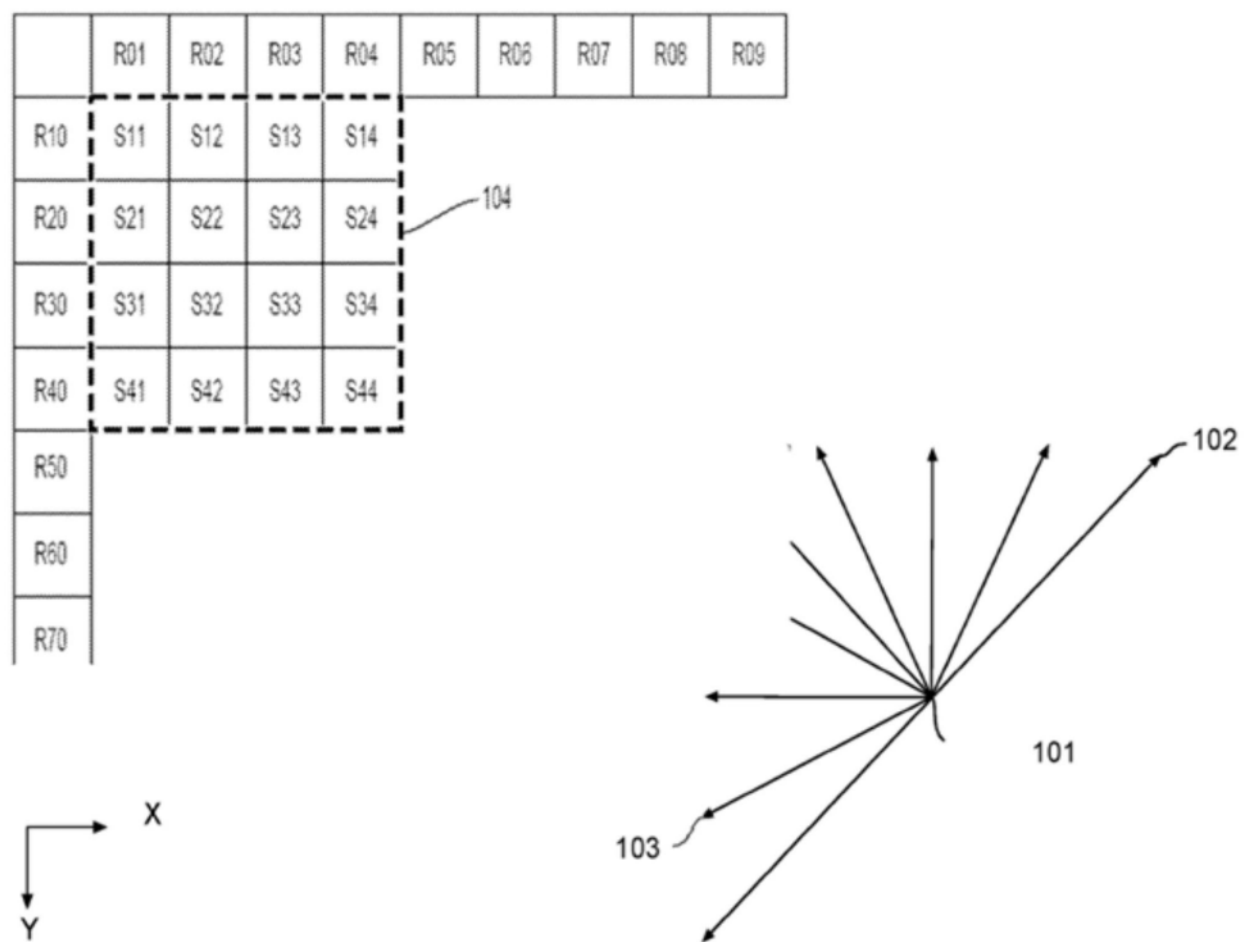


图1A

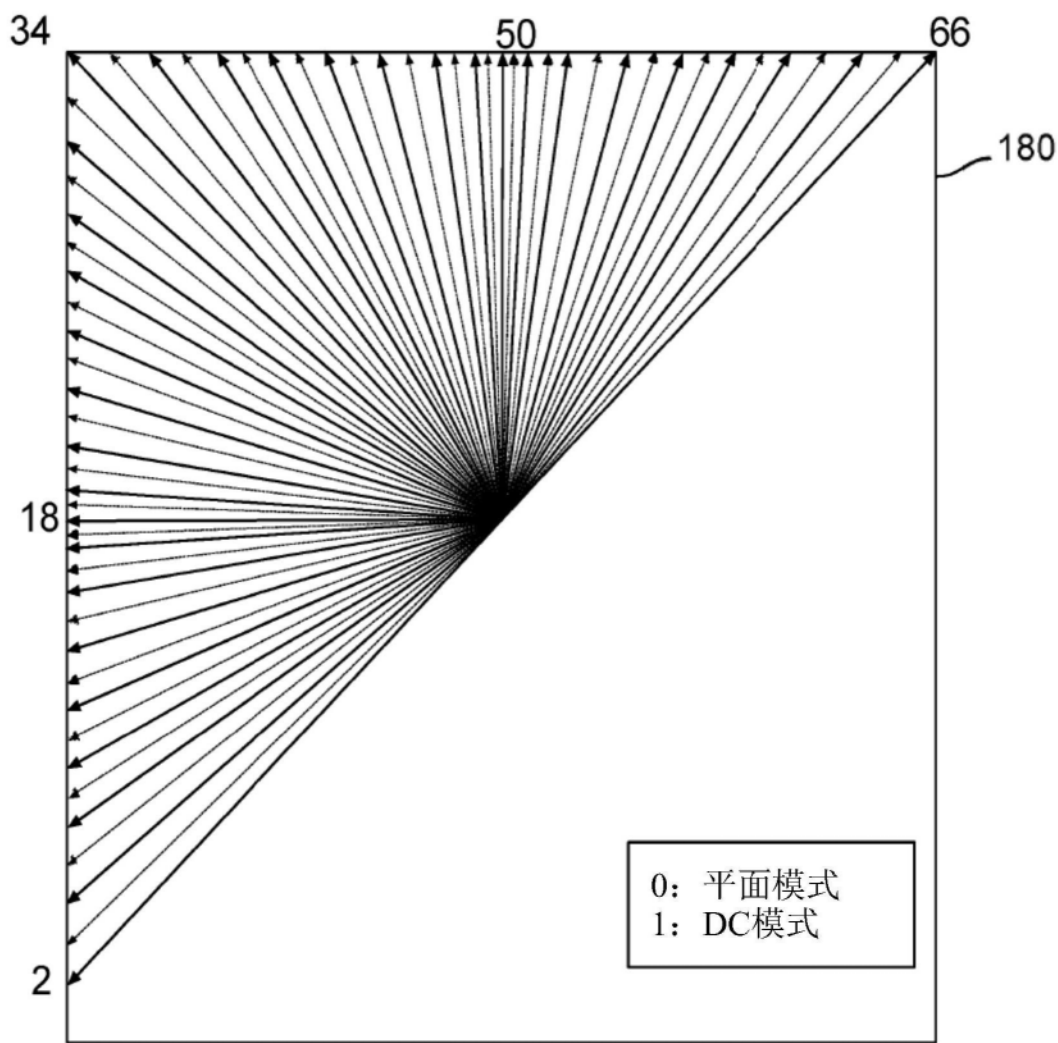


图1B

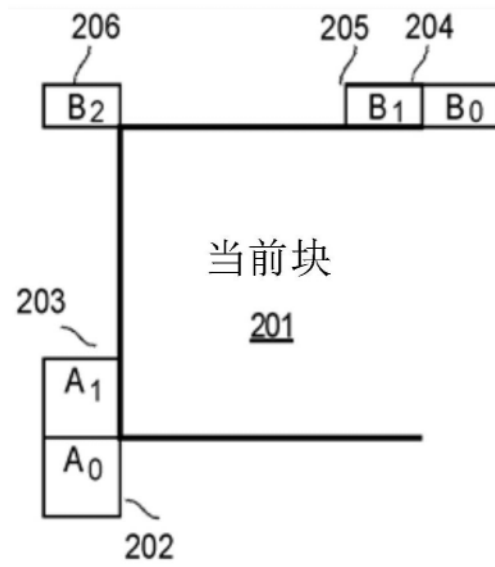


图2

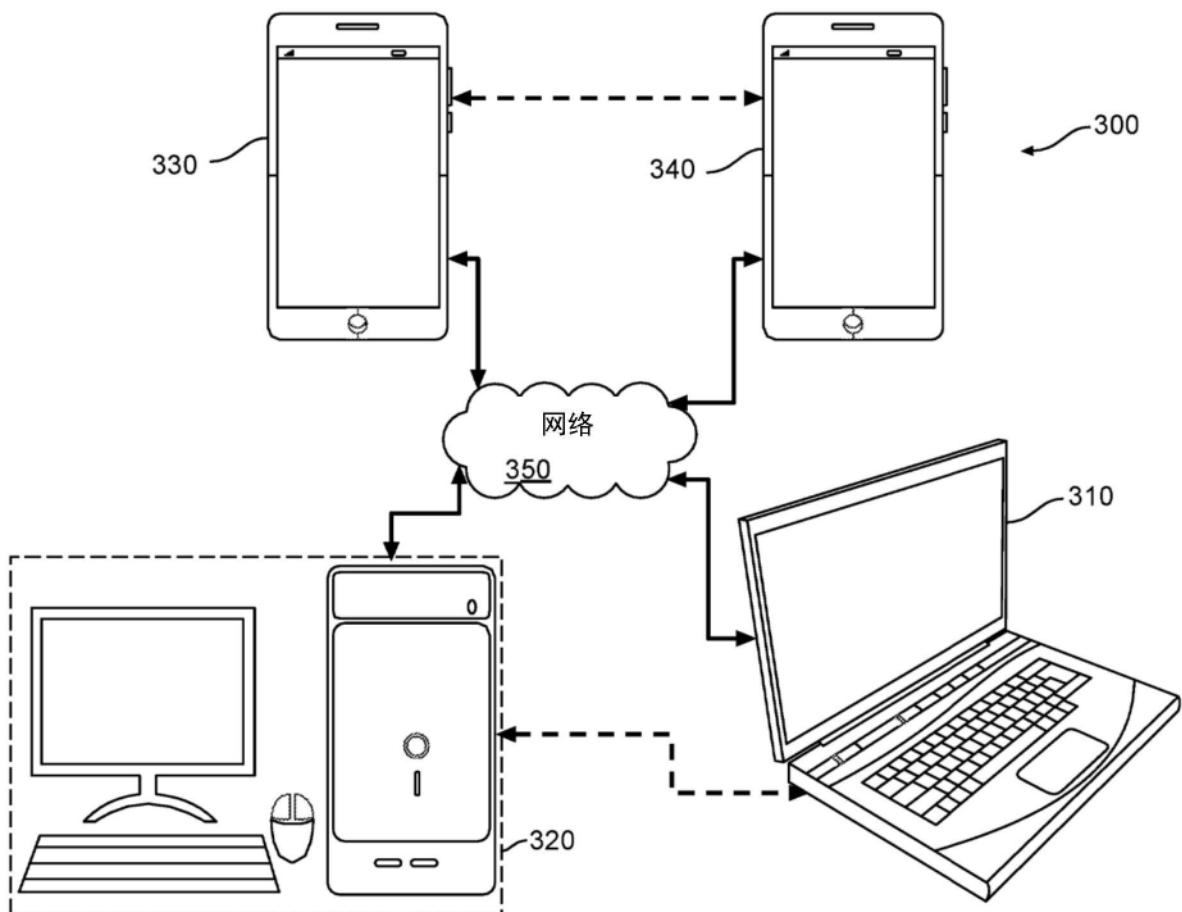


图3

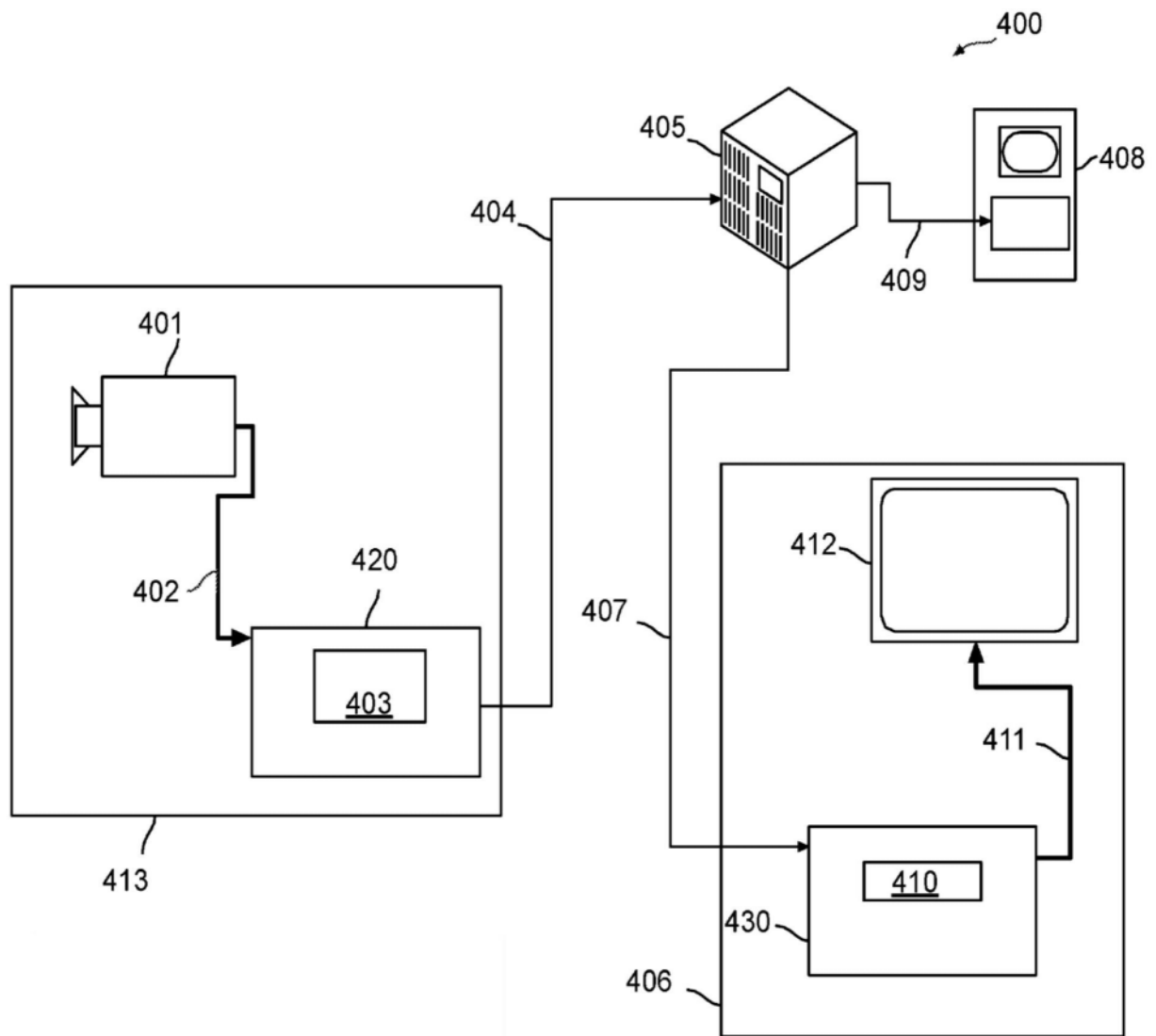


图4

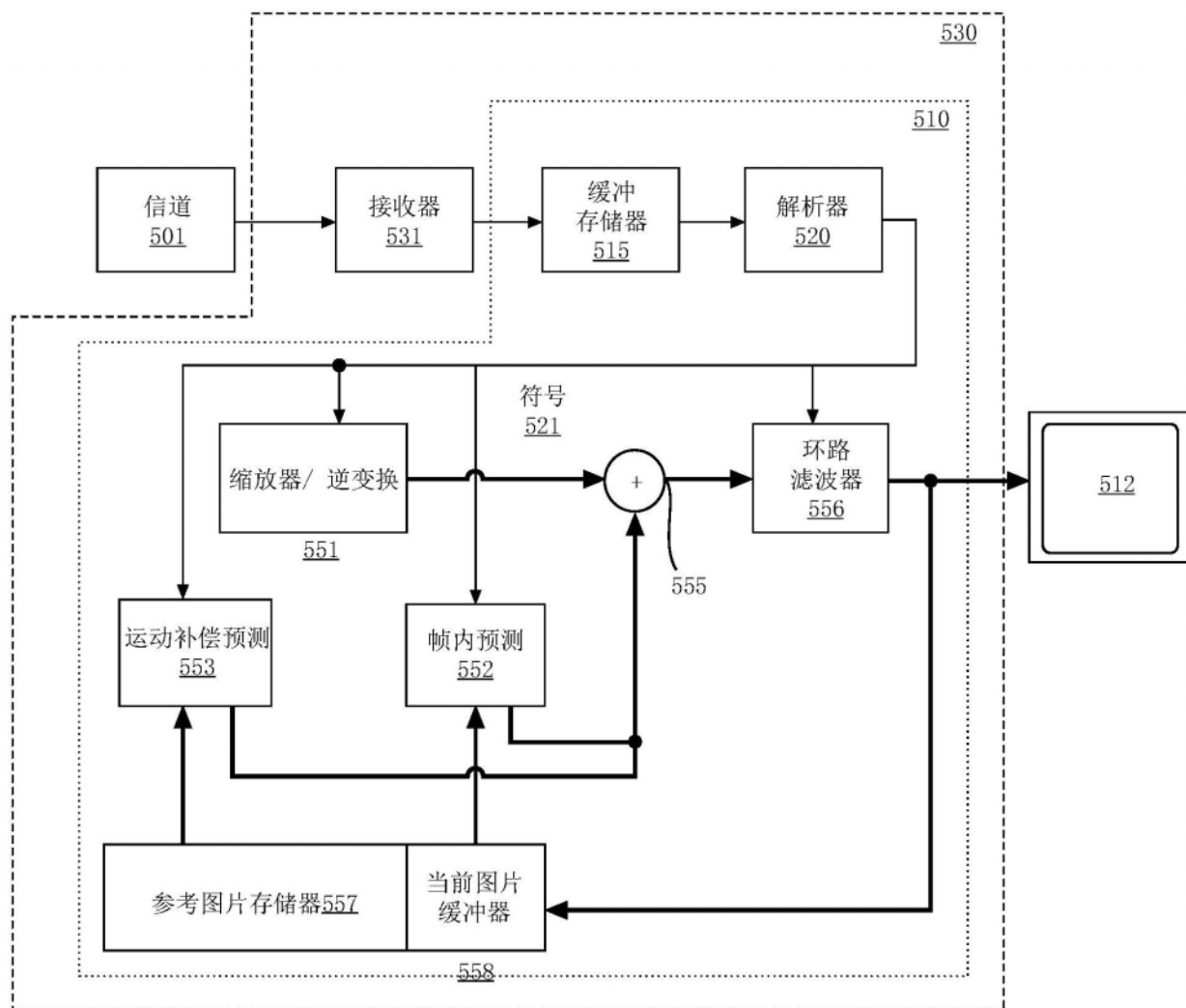


图5

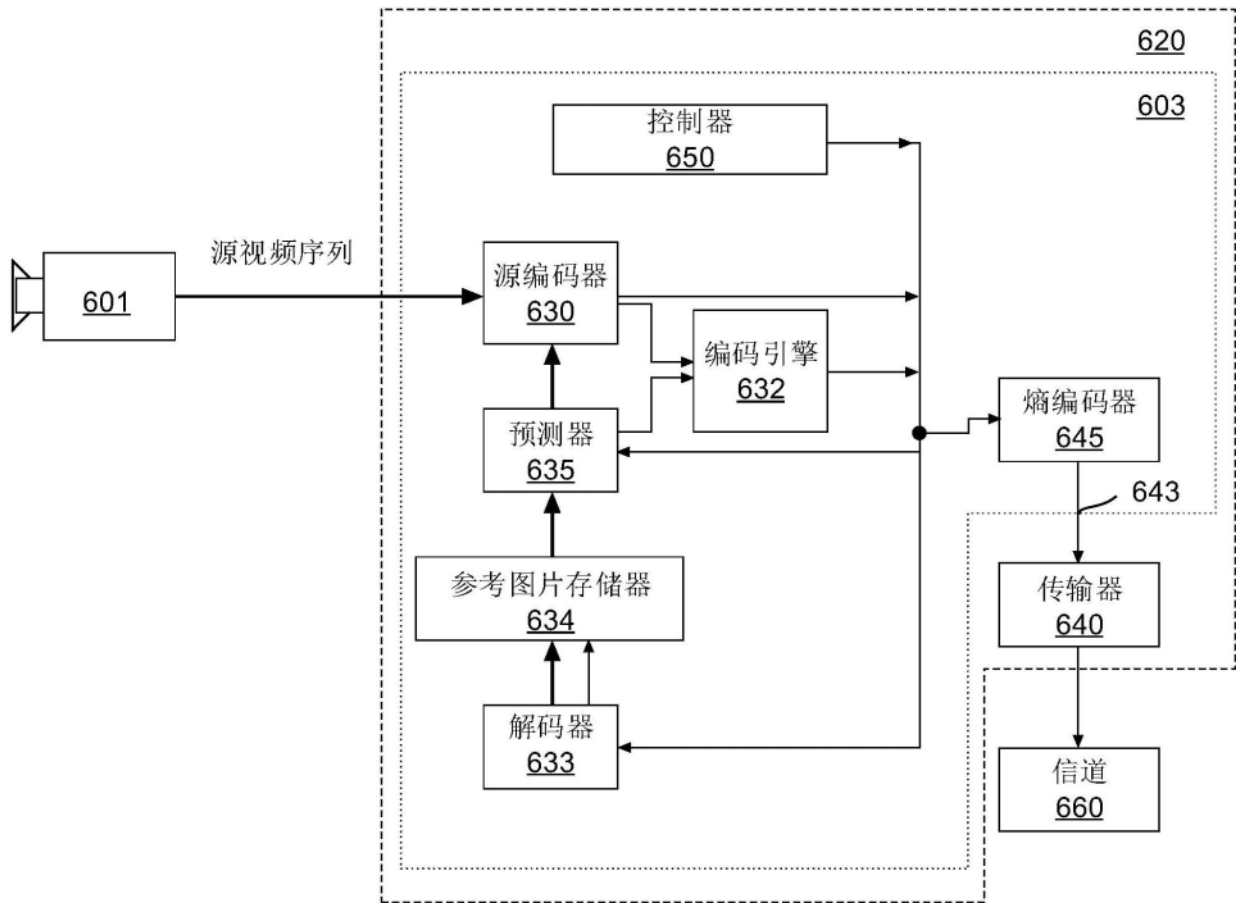


图6

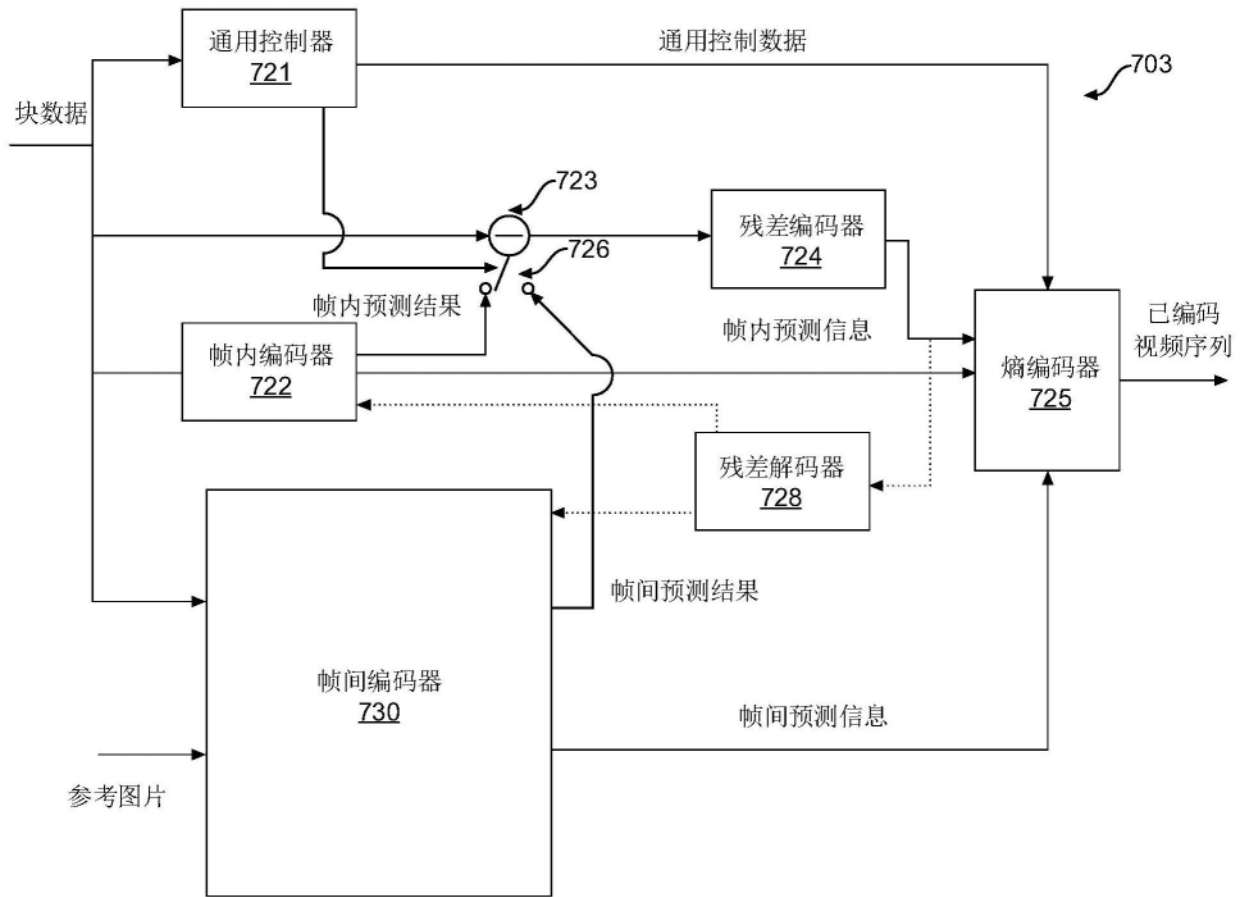


图7

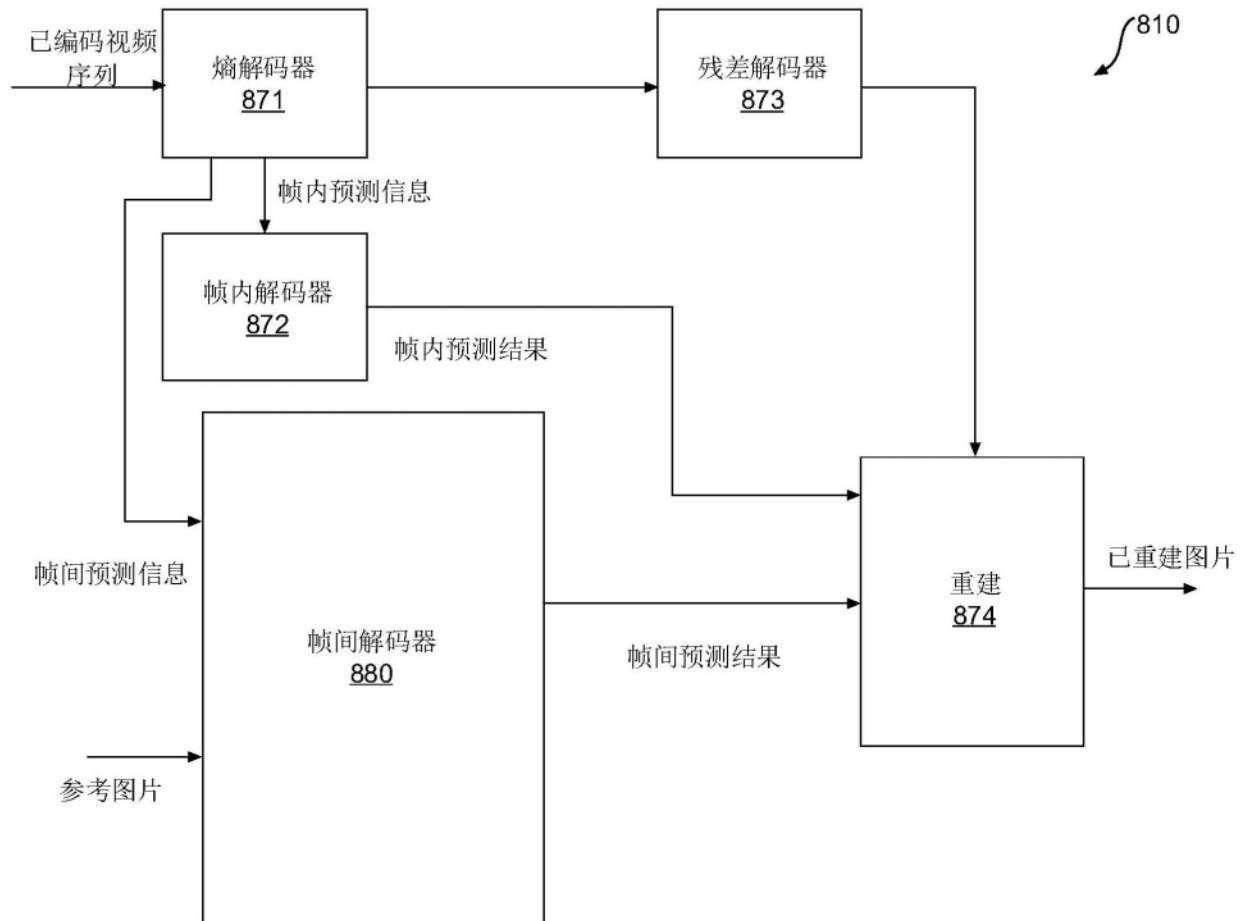


图8

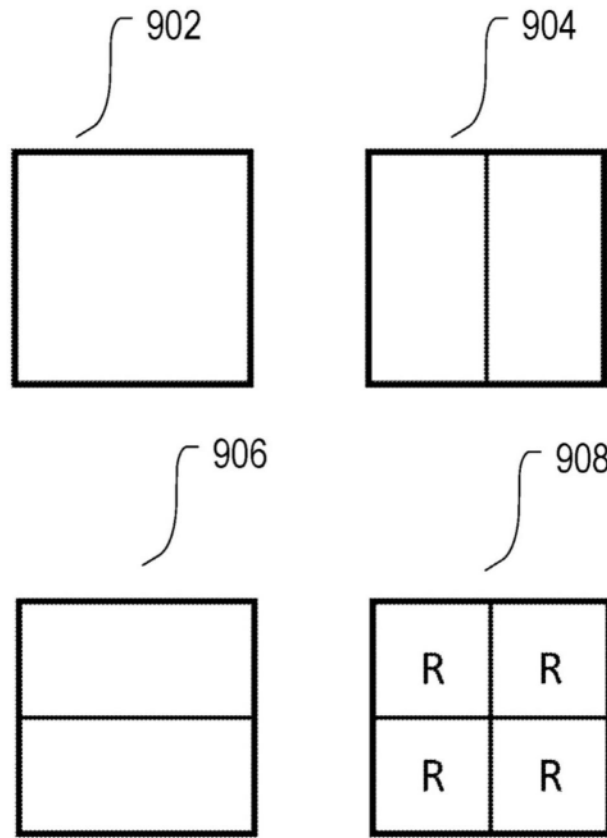


图9

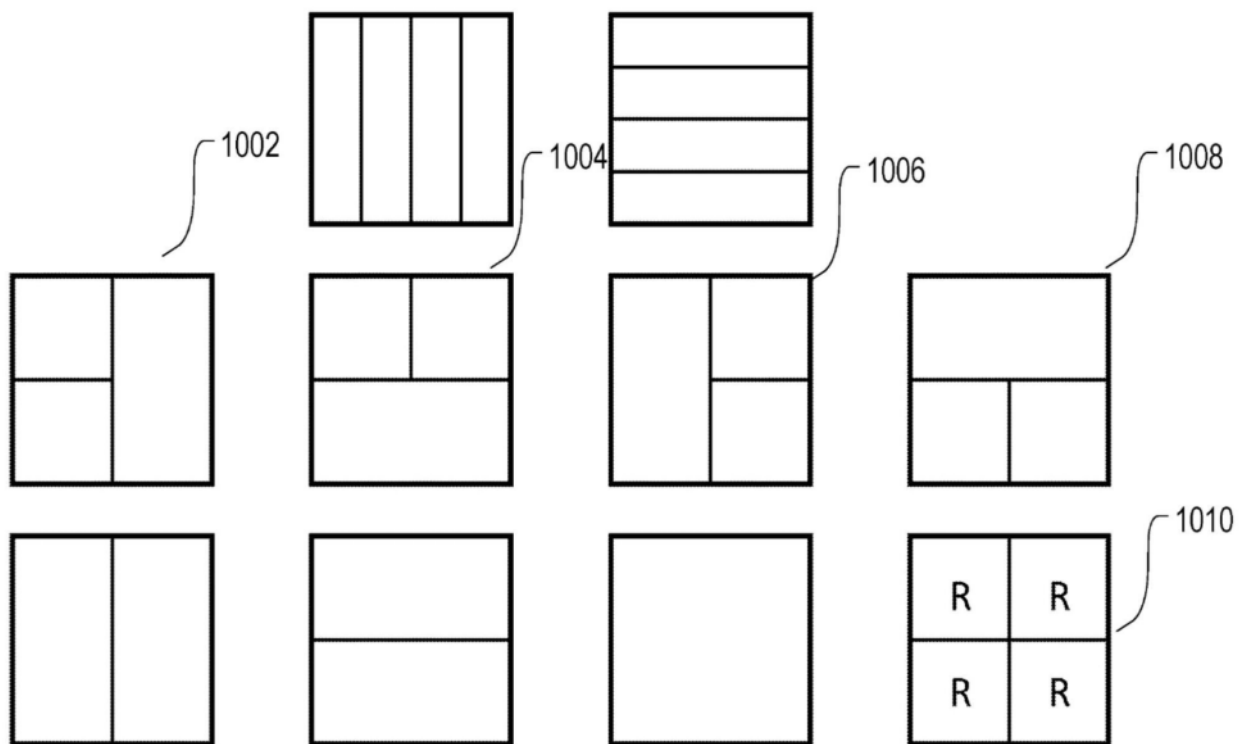


图10

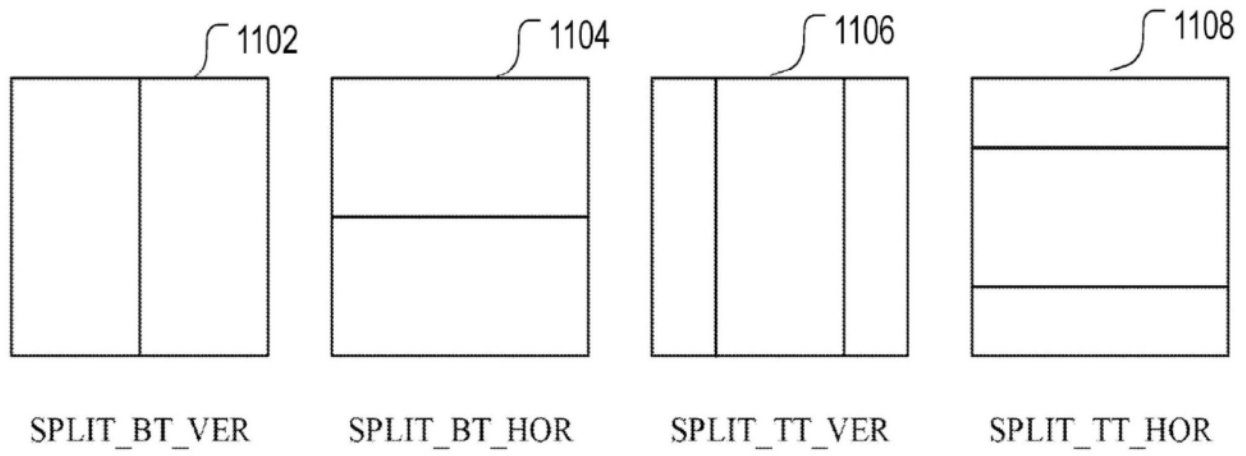


图11

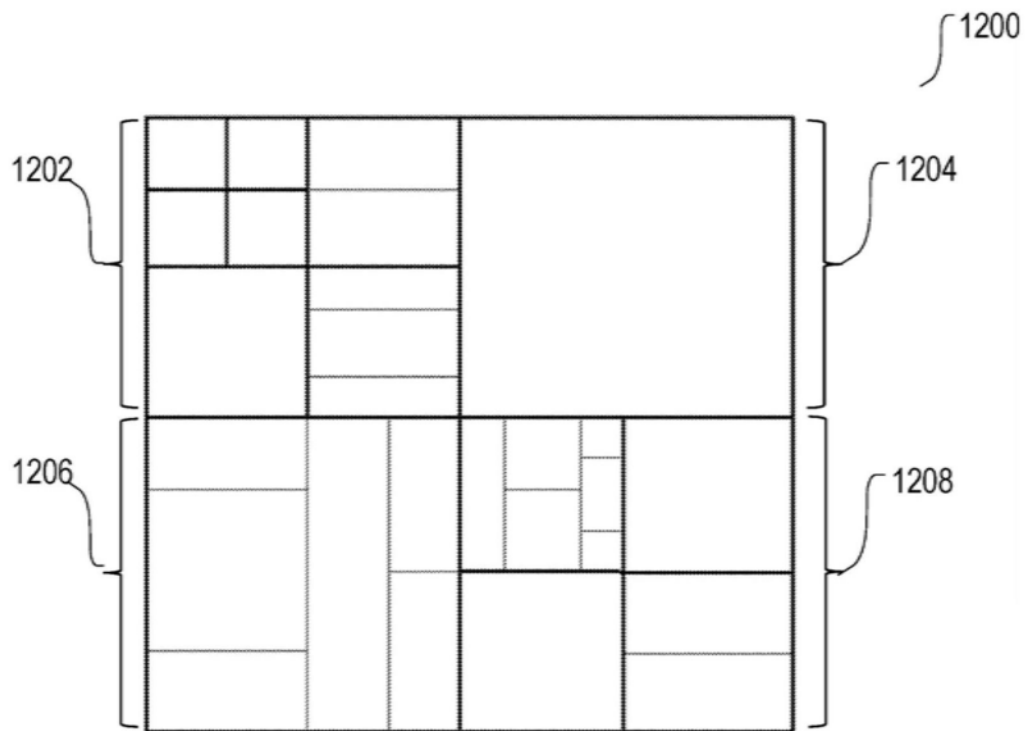


图12

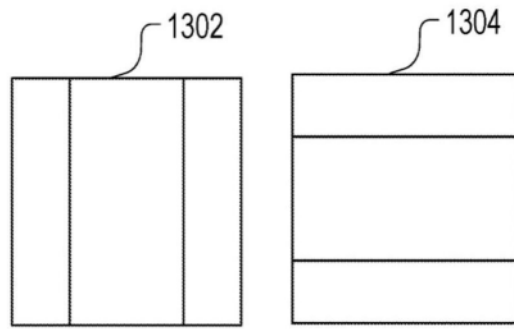


图13

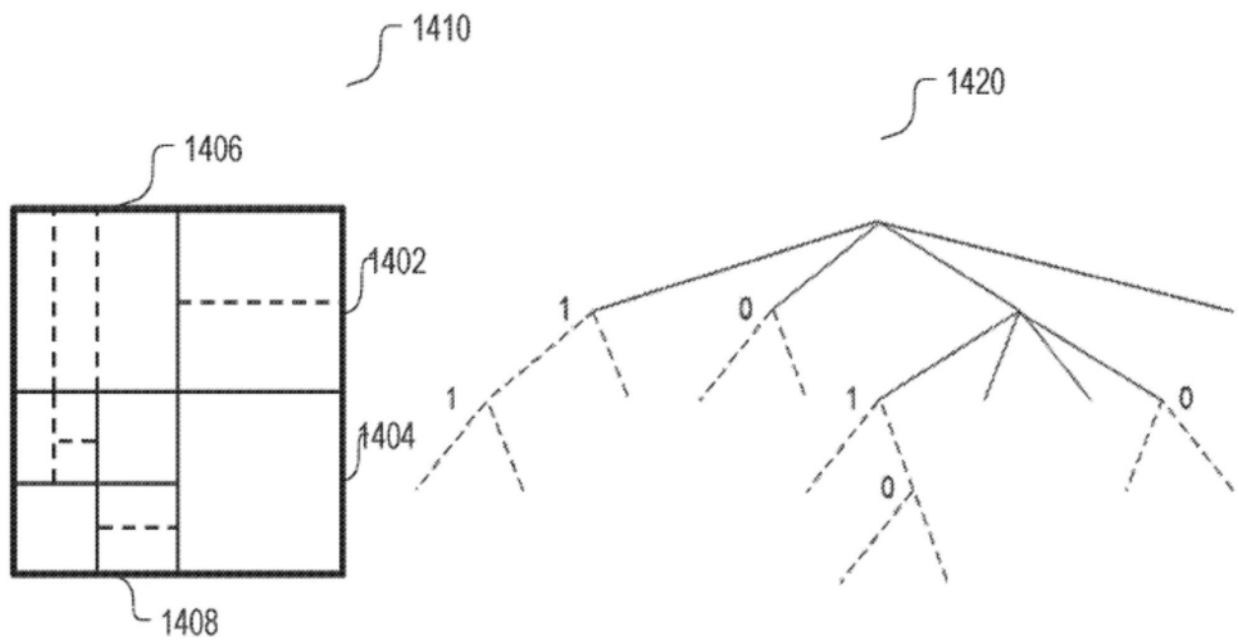


图14

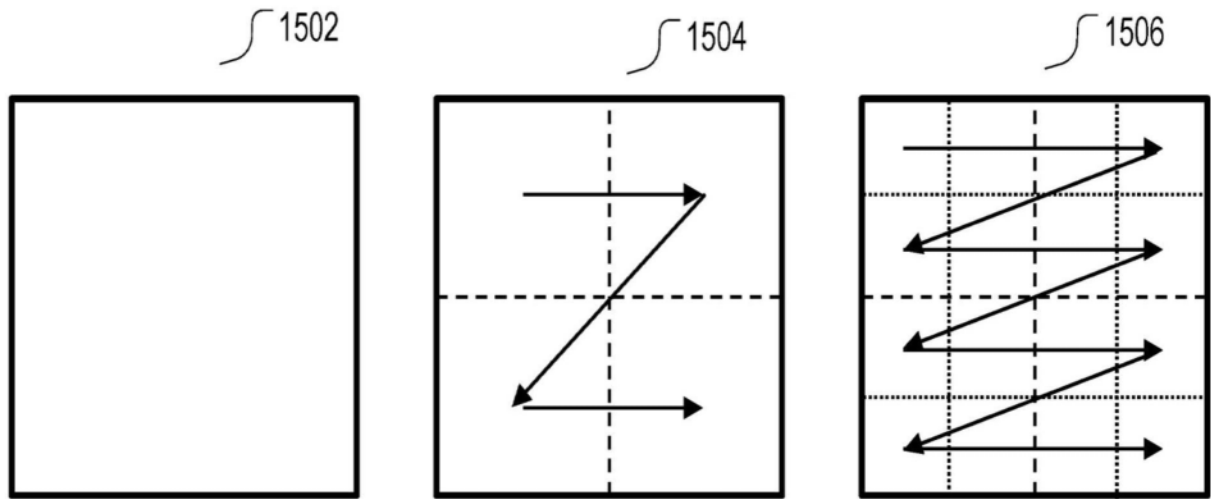


图15

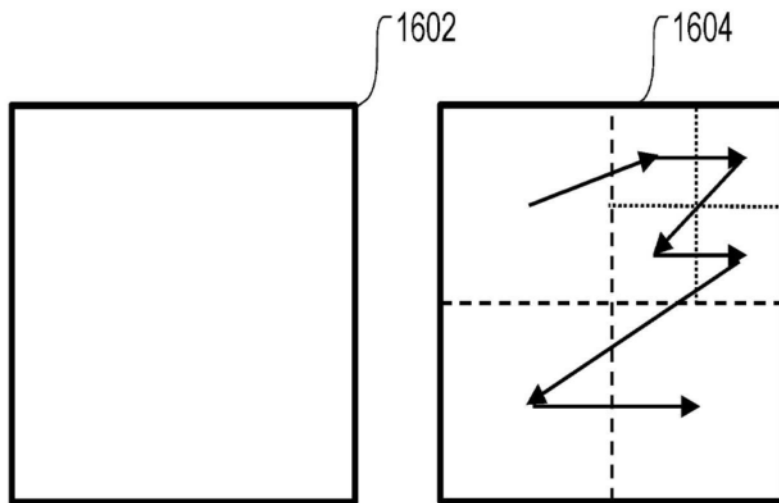


图16

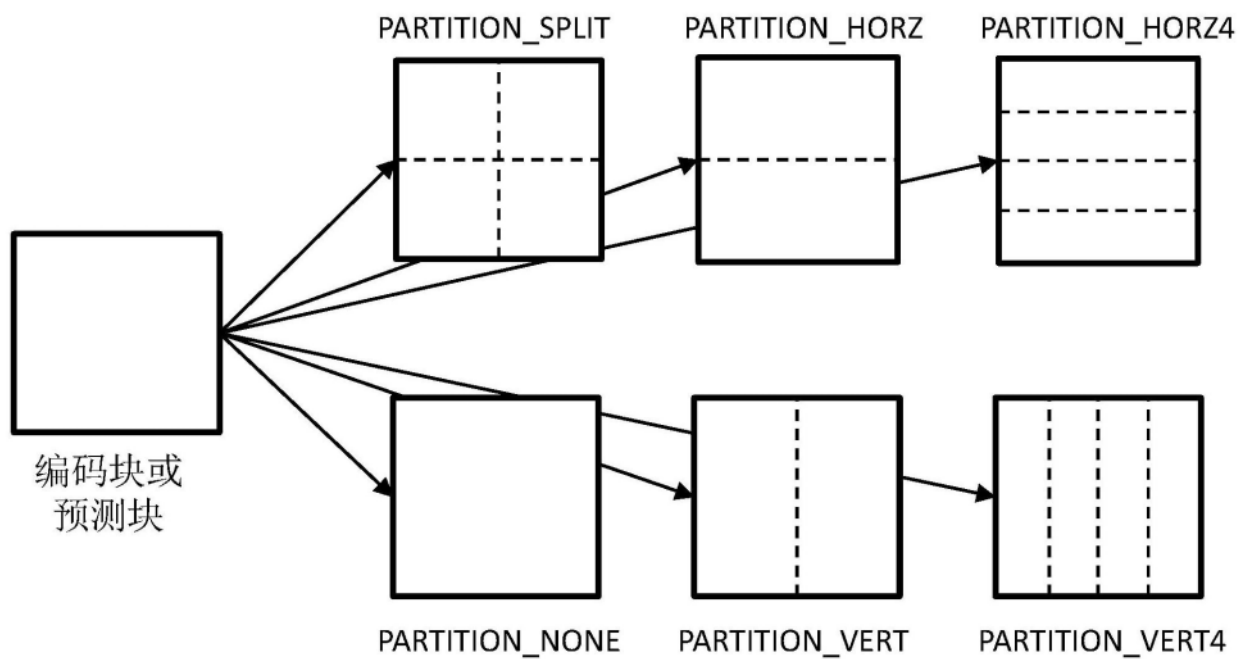


图17

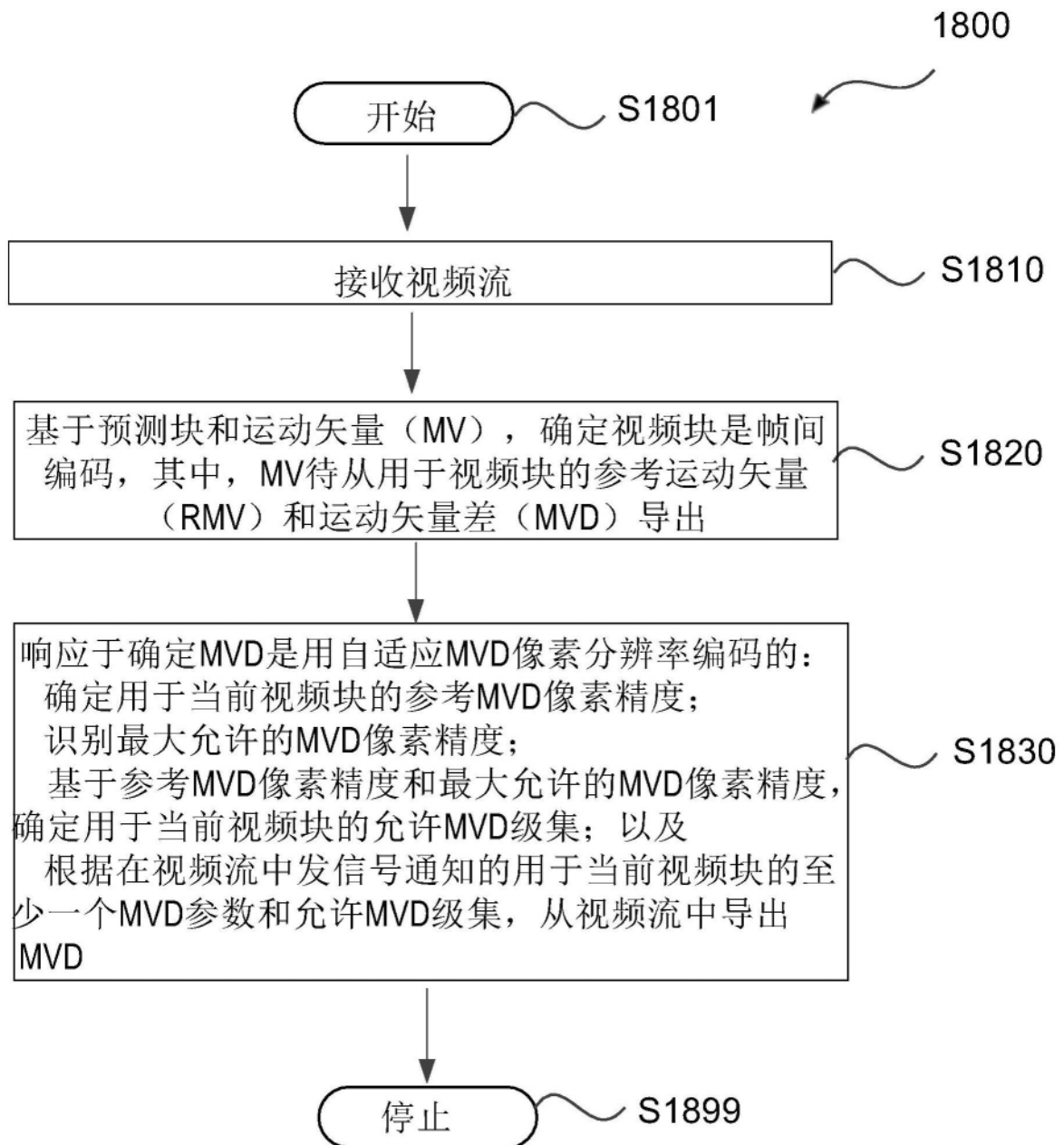


图18

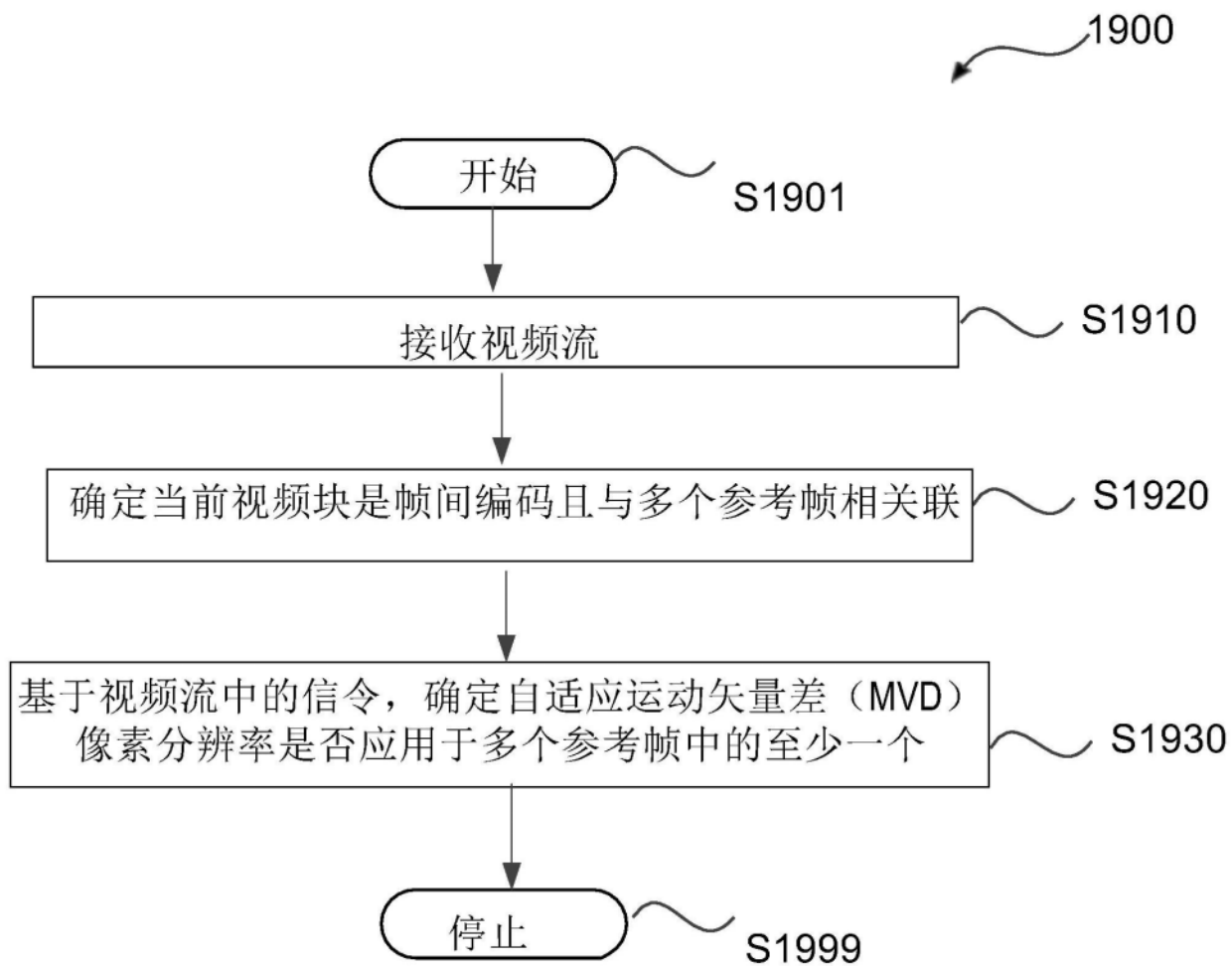


图19

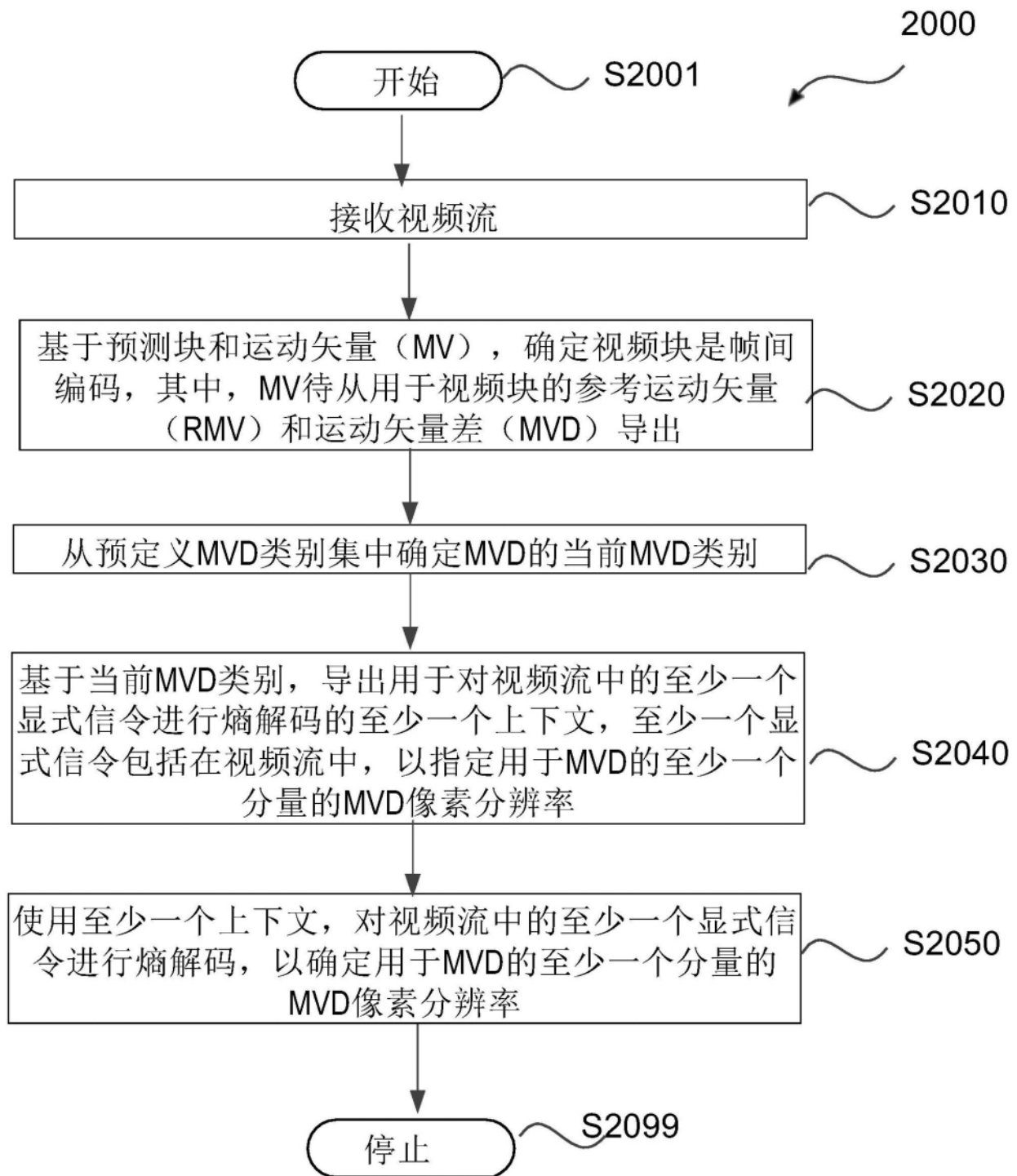


图20

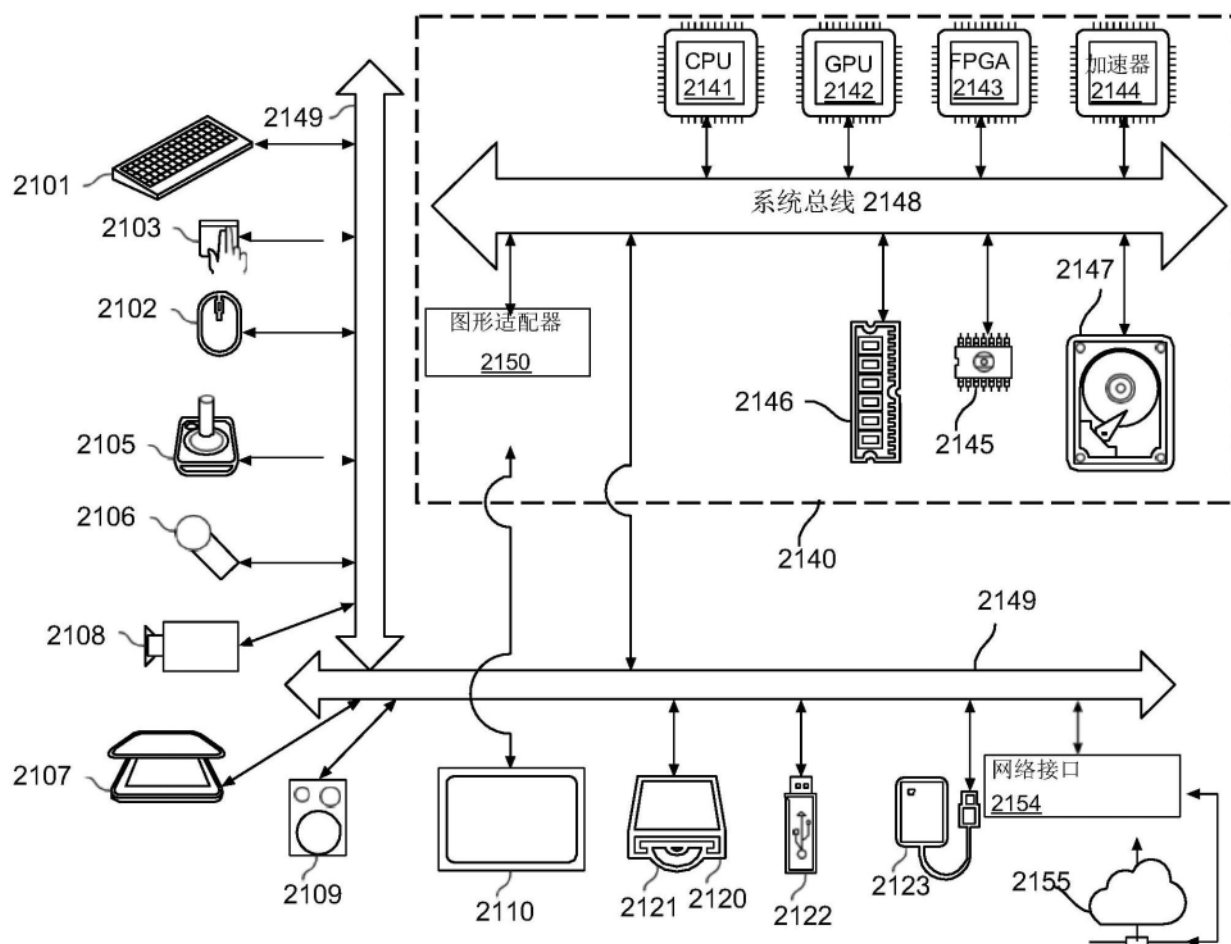


图21