(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 118433426 A
(43)申请公布日 2024.08.02

(21)申请号 202410517369.3

(22)申请日 2020.03.11

(30)优先权数据
62/817,517 2019.03.12 US

(62)分案原申请数据
202080018657.4 2020.03.11

(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 赵欣 许晓中 李翔 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 陈美娥 王琦

(51)Int.Cl.
H04N 19/96(2014.01)
H04N 19/107(2014.01)
H04N 19/159(2014.01)
H04N 19/176(2014.01)
H04N 19/513(2014.01)

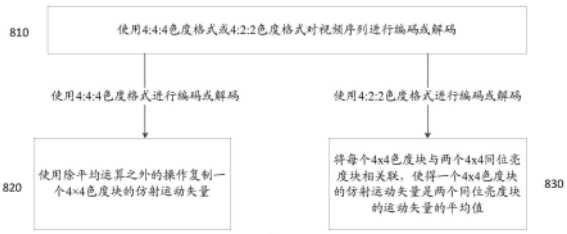
权利要求书2页 说明书17页 附图14页

(54)发明名称

对视频进行编码或解码的方法和装置

(57)摘要

一种对视频序列进行编码或解码的方法和装置包括:使用4:4:4色度格式对所述视频序列进行编码或解码,或者使用4:2:2色度格式对所述视频序列进行编码或解码。当使用所述4:4:4色度格式对所述视频序列进行编码或解码时,使用除平均运算之外的操作复制一个4×4亮度块的仿射运动矢量,并将所述仿射运动矢量与同位4×4色度块相关联。当使用所述4:2:2色度格式对所述视频序列进行编码或解码时,将每个4×4色度块与两个4×4同位亮度块相关联,使得一个4×4色度块的仿射运动矢量是两个同位亮度块的运动矢量的平均值。



CN 118433426 A

1.一种对已编码的视频序列进行解码的方法,其特征在于,所述方法包括:

使用4:4:4色度格式和4:2:2色度格式之一对所述视频序列进行解码;

当使用所述4:4:4色度格式对所述视频序列进行解码时,复制一个4×4亮度块的仿射运动矢量,而不对所述仿射运动矢量使用平均运算,并将所述仿射运动矢量与同位4×4色度块相关联;以及

当使用所述4:2:2色度格式对所述视频序列进行解码时,将每个4×4色度块与两个4×4同位亮度块相关联,使得一个4×4色度块的仿射运动矢量是两个同位亮度块的运动矢量的平均值;

对于不同的颜色分量,允许的最大变换大小是相同的;

当使用所述4:2:2色度格式对所述视频序列进行解码时,不同颜色分量中的最大垂直大小是相同的,并且色度分量的最大水平变换大小是亮度分量的最大水平变换大小的一半;

当使用所述4:2:2色度格式对所述视频序列进行解码时,对所述亮度分量应用8抽头插值滤波器,并对亮度分量和色度分量应用不可分离二次变换NSST。

2.根据权利要求1所述的方法,其特征在于,进一步包括:

将当前4×4色度块划分为四个2×2子块,

导出左上2×2色度子块的同位亮度块的第一仿射运动矢量,

导出右下2×2色度子块的同位亮度块的第二仿射运动矢量,以及

使用所述第一仿射运动矢量和所述第二仿射运动矢量的平均值导出所述当前4×4色度块的仿射运动矢量。

3.根据权利要求1所述的方法,其特征在于,进一步包括:

将分量Y、Cb和Cr编码为三个单独的树,

其中,所述三个单独的树中的每个树对所述分量Y、Cb和Cr中的一个分量进行编码。

4.根据权利要求3所述的方法,其特征在于,对I切片或I图块群组执行所述将分量Y、Cb和Cr编码为三个单独的树的操作。

5.根据权利要求1所述的方法,其特征在于,将位置相关预测器组合PDPC、多变换选择MTS、帧内子分区ISP和多参考行MRL帧内预测中的至少一个应用于亮度分量和色度分量。

6.根据权利要求5所述的方法,其特征在于,

当将所述多参考行MRL帧内预测应用于所述亮度分量和所述色度分量时,并且当使用所述4:4:4色度格式对所述视频序列进行解码时,所述方法进一步包括:选择用于帧内预测的第N个参考,并且使用相同的参考行而无需为色度分量显式发信号通知;

当将所述帧内子分区ISP应用于所述亮度分量和所述色度分量时,所述方法进一步包括:对于分量Y、Cb和Cr,在当前块的块级别应用所述帧内子分区ISP;以及

当不同的树用于不同的颜色分量时,所述方法进一步包括:从同位Y分量隐式地导出用于U分量和V分量的编解码参数而无需发信号通知。

7.一种对视频进行编码的方法,用于生成已编码的视频序列,其特征在于,所述方法包括:

确定使用4:4:4色度格式和4:2:2色度格式之一对所述视频进行编码;

当使用所述4:4:4色度格式对所述视频进行编码时,复制一个4×4亮度块的仿射运动

矢量,而不对所述仿射运动矢量使用平均运算,并将所述仿射运动矢量与同位4×4色度块相关联;以及

当使用所述4:2:2色度格式对所述视频进行编码时,将每个4×4色度块与两个4×4同位亮度块相关联,使得一个4×4色度块的仿射运动矢量是两个同位亮度块的运动矢量的平均值;

对于不同的颜色分量,允许的最大变换大小是相同的;

当使用所述4:2:2色度格式对所述视频序列进行编码时,不同颜色分量中的最大垂直大小是相同的,并且色度分量的最大水平变换大小是亮度分量的最大水平变换大小的一半;

当使用所述4:2:2色度格式对所述视频序列进行编码时,对所述亮度分量应用8抽头插值滤波器,并对亮度分量和色度分量应用不可分离二次变换NSST。

8.一种对视频序列进行解码的设备,其特征在于,所述设备包括:

至少一个存储器,被配置为存储程序代码;

至少一个处理器,被配置为读取所述程序代码,并按照所述程序代码的指示进行操作,执行权利要求1-6任一项所述的方法。

9.一种对视频进行编码的设备,用于生成已编码的视频序列,其特征在于,所述设备包括:

至少一个存储器,被配置为存储程序代码;

至少一个处理器,被配置为读取所述程序代码,并按照所述程序代码的指示进行操作,执行权利要求7所述的方法。

10.一种非易失性计算机可读介质,其特征在于,用于存储程序代码,所述程序代码包括一个或多个指令,当所述指令由设备的一个或多个处理器执行时,使所述一个或多个处理器执行权利要求1-7任一项所述的方法。

11.一种非易失性计算机可读介质,其特征在于,用于存储程序代码,所述程序代码包括一个或多个指令,当所述指令由设备的一个或多个处理器执行时,使所述一个或多个处理器执行权利要求1-8任一项所述的方法。

12.一种存储已编码的视频序列的方法,其特征在于,在非易失性计算机可读存储介质上存储视频序列,所述视频序列根据权利要求7所述的编码方法产生,或者基于权利要求1-6任一项所述的解码方法进行解码。

对视频进行编码或解码的方法和装置

[0001] 相关申请的交叉引用

[0002] 本申请根据35U.S.C.§119要求于2019年3月12日在美国专利商标局提交的、申请号为62/817,517的美国临时申请的优先权,其全部内容通过引用并入本文。

技术领域

[0003] 本公开实施例涉及视频处理,更具体地,涉及对视频序列进行编码或解码的方法和装置、以及计算机可读介质。

背景技术

[0004] 最近,国际电信联盟(ITU)电信标准化部门(ITU-T)的视频编解码专家组(VCEG)(ITU的一个部门)和ISO/IEC MPEG(JTC 1/SC 29/WG 11)(国际标准化组织(ISO)和国际电工委员会(IEC)联合技术委员会ISO/IEC JTC 1的标准化小组委员会)于2013年发布了H.265/高效视频编解码(HEVC)标准(第1版)。该标准于2014年更新为第2版,于2015年更新为第3版,于2016年更新为第4版。

[0005] 从那时起,他们一直在研究对于压缩能力明显超过HEVC标准(包括其扩展)的未来视频编解码技术标准化的潜在需求。2017年10月,他们发布了针对能力超出HEVC(CfP)的视频压缩提案的联合征集。截至2018年2月15日,分别提交了针对标准动态范围(SDR)的共22个CfP回复、针对高动态范围(HDR)的12个CfP回复和针对360种视频类别的12个CfP回复。2018年4月,在122运动图像专家组(MPEG)/第十次联合视频开发组-联合视频专家组(JVET)会议上,对所有收到的CfP回复进行了评估。通过仔细的评估,JVET正式启动了超越HEVC的下一代视频编解码的标准化,即所谓的通用视频编解码(VVC)。

[0006] 现在将描述HEVC块分区结构。在HEVC中,可以使用表示为编码树的四叉树结构将编码树单元(CTU)分割成多个编码单元(CU),以适应各种局部特征。可以在CU级别决定是使用帧间图片(时间)预测还是使用帧内图片(空间)预测来编码图片区域。根据预测单元(PU)分割类型,每个CU可以进一步被分割成一个、两个或四个PU。在一个PU内,可以应用相同的预测过程,并且可以基于PU将相关信息发送到解码器。通过应用基于PU分割类型的预测过程而获得残差块之后,可以根据与CU的编码树类似的另一种四叉树结构,将CU划分为多个变换单元(TU)。HEVC结构的一个特征是它可以包含多个分区概念,包括CU、PU和TU。在HEVC中,对于帧间预测块,CU或TU通常只能是正方形,而PU可以是正方形或矩形。在HEVC中,可以将一个编码块进一步分割为四个正方形子块,并且可以对每个子块(即TU)执行变换。每个TU可以被进一步递归分割(使用四叉树分割)成较小的TU。这称为残差四叉树(RQT)。

[0007] 在图片边界处,HEVC采用隐式四叉树分割,使得块保持四叉树分割,直到大小符合图片边界。

[0008] 现在将描述使用四叉树(QT)加二叉树(BT)的块分区结构。在HEVC中,可以使用表示为编码树的四叉树结构将CTU分割成多个CU,以适应各种局部特征。可以在CU级别决定是使用帧间图片(时间)预测还是使用帧内图片(空间)预测来编码图片区域。根据PU分割类

型,每个CU可以被进一步分割成一个、两个或四个PU。在一个PU内,可以应用相同的预测过程,并且可以基于PU将相关信息发送到解码器。通过应用基于PU分割类型的预测过程获得残差块之后,可以根据与CU的编码树类似的另一种四叉树结构,将CU划分为多个变换单元(TU)。HEVC结构的一个特征是它包含多个分区概念,包括CU、PU和TU。

[0009] 四叉树加二叉树(QTBT)结构消除了多个分区类型的概念,即,它消除了CU、PU和TU概念的分离,并且支持更灵活的CU分区形状。在QTBT块结构中,CU可以具有正方形或矩形形状。如图1A所示,首先采用四叉树结构对编码树单元(CTU)进行分区。然后,可以进一步采用二叉树结构对四叉树叶节点进行分区。在二叉树分割中,可以有对称水平分割和对称垂直分割两种分割类型。二叉树叶节点被称为编码单元(CU),并且该分割可用于预测和变换处理而无需任何进一步的分区。这意味着,CU、PU和TU在QTBT编码块结构中具有相同的块大小。在JEM中,CU有时可以由不同颜色分量的编码块(CB)组成,例如,在4:2:0色度格式的P切片和B切片的情况下,一个CU可以包含一个亮度CB和两个色度CB,并且该CU可以包含单个分量的CB,例如,在I切片的情况下,一个CU可以仅包含一个亮度CB或仅包含两个色度CB。

[0010] 为QTBT分区方案定义了以下参数:

[0011] CTU大小:四叉树的根节点大小,与HEVC中的概念相同;

[0012] MinQTSize:允许的最小四叉树叶节点大小;

[0013] MaxBTSize:允许的最大二叉树根节点大小;

[0014] MaxBTDepth:允许的最大二叉树深度;

[0015] MinBTSize:允许的最小二叉树叶节点大小。

[0016] 在QTBT分区结构的一个示例中,可以将CTU大小设置为具有两个相应的64×64色度样本块的128×128亮度样本。可以将MinQTSize设置为16×16,将MaxBTSize设置为64×64,将MinBTSize(对于宽度和高度)设置为4×4,将MaxBTDepth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。四叉树叶节点可以具有从16×16(即MinQTSize)到128×128(即CTU大小)的大小。如果四叉树叶节点是128×128,则由于其大小超过MaxBTSize(即64×64),因此将不会采用二叉树对其进一步分割。否则,四叉树叶节点会采用二叉树进一步分割。因此,四叉树叶节点也可以是二叉树的根节点,并且其二叉树深度为0。当二叉树深度达到MaxBTDepth(即4)时,不考虑进一步的分割。当二叉树节点的宽度等于MinBTSize(即4)时,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。通过预测和变换处理可以进一步处理二叉树的叶节点,而无需任何进一步的分区。在JEM中,最大CTU大小可以是256×256亮度样本。

[0017] 图1A示出了利用QTBT进行块分区的示例,图1B示出了相应的树表示。实线表示四叉树分割,虚线表示二叉树分割。在二叉树的每个分割(即,非叶)节点中,可以发信号通知一个标志,以指示可以使用哪种分割类型(即,水平或垂直),其中0表示水平分割,1表示垂直分割。对于四叉树分割,不需要指示分割类型,因为四叉树分割水平和垂直地分割块,以产生具有相等大小的4个子块。

[0018] 此外,QTBT方案支持亮度和色度具有单独的QTBT结构的灵活性。当前,对于P切片和B切片,一个CTU中的亮度和色度CTB共用相同的QTBT结构。然而,对于I切片,亮度CTB可以通过QTBT结构划分为多个CU,色度CTB可以通过另一个QTBT结构(即双树(DT)结构)划分为多个色度CU。这意味着,I切片中的CU由亮度分量的编码块或两个色度分量的编码块组成,

而P切片或B切片中的CU由所有三个颜色分量的编码块组成。

[0019] 在HEVC中,可以对小块的帧间预测进行限制,以减少运动补偿的存储器访问,这使得4×8块和8×4块不支持双向预测,而4×4块不支持帧间预测。在JEM-7.0中实施的QTBT方案中,可以消除这些限制。

[0020] 现在将描述使用三叉树(TT)的块分区。已经提出了一种多类型树(MTT)结构。MTT是比QTBT更灵活的树结构。在MTT中,除了四叉树和二叉树之外,还引入了水平和垂直中心-侧(center-side)三叉树,如图2A和2B所示。

[0021] 三叉树分区的一些优点如下:

[0022] 提供对四叉树分区和二叉树分区的补充,三叉树分区能够捕获位于块中心的对象,而四叉树和二叉树可以沿着块中心进行分割。

[0023] 三叉树的分区的宽度和高度可以是2的幂,因此不需要额外的变换。

[0024] 两级树的设计主要是为了降低复杂度。理论上,遍历树的复杂度为T^D,其中,T表示分割类型的数量,D是树的深度。

[0025] 现在将描述YUV格式。图3中示出了不同的YUV格式,即色度格式。不同的色度格式限定了不同颜色分量的不同下采样网格。

[0026] 现在将描述跨分量线性建模(CCLM)。在VTM中,对于帧内PU的色度分量,编码器在8种模式(包括平面、DC、水平、垂直、来自亮度分量的帧内预测模式(DM)的直接副本、左和上跨分量线性模式(LT_CCLM)、左跨分量线性模式(L_CCLM)和上跨分量线性模式(T_CCLM))中选择最佳色度预测模式。在这些模式中,可以将LT_CCLM、L_CCLM和T_CCLM归类为跨分量线性模式(CCLM)。这3种模式之间的区别在于,相邻样本的不同区域可用于导出参数α和β。对于LT_CCLM,左相邻样本和上相邻样本均可用于导出参数α和β。对于L_CCLM,只有左相邻样本用于导出参数α和β。对于T_CCLM,只有上相邻样本用于导出参数α和β。

[0027] 跨分量线性模型(CCLM)预测模式可用于减少跨分量冗余,其中,可以通过使用如下线性模型基于同一CU的重建亮度样本来预测色度样本:

[0028] $pred_c(i,j)=\alpha \cdot rec_L'(i,j)+\beta$.(等式1)

[0029] 在次,$pred_c(i,j)$表示CU中的预测色度样本,$rec_L'(i,j)$表示同一CU的下采样重建亮度样本。参数α和β可以通过直线方程(例如max-min方法)导出。该计算过程可以作为解码过程的一部分执行,而不仅仅作为编码器搜索操作,因此不使用语法来传达α和β值。

[0030] 对于色度4:2:0格式,CCLM预测应用六抽头插值滤波器,以获得与色度样本对应的下采样亮度样本,如图3所示。这里,可以根据重建亮度样本计算下采样亮度样本Rec'L[x,y]。

[0031] 下采样亮度样本可用于找到最大和最小样本点。这两个点(亮度和色度的对)(A,B)可以是如图4所示的一组相邻亮度样本内的最小值和最大值。其中,线性模型参数α和β可以根据以下等式获得:

[0032] $$\alpha = \frac{y_B - y_A}{x_B - x_A} \qquad (等式2)$$

[0033] $\beta = y_A - \alpha x_A$(等式3)

[0034] 在此,可以避免除法并且可以用乘法和移位来代替除法。查找表(LUT)可用来存储预先计算的值,最大亮度样本和最小亮度样本之间的绝对差值可用于指定LUT的条目索引,

并且LUT的大小可以为512。

[0035] 在T_CCLM模式中,仅可以使用图5A和图5B中示出的相邻样本(包括2*W个样本)来计算线性模型系数。

[0036] 在L_CCLM模式中,仅可以使用左相邻样本(包括2*H个样本)来计算线性模型系数,如图6A和图6B所示。

[0037] CCLM预测模式还包括两个色度分量之间的预测,即,可以从Cb分量预测Cr分量。不使用重建样本信号,而是可以将CCLM的Cb到Cr预测应用到残差域中。这可以通过将加权重建的Cb残差添加到原始Cr帧内预测以形成最终Cr预测来实现:

[0038] $pred_{cr}*(i,j)=pred_{cr}(i,j)+\alpha\cdot resi_{cb}'(i,j)$.(等式4)

[0039] 可以将CCLM的亮度到色度预测模式作为一个附加的色度帧内预测模式进行添加。在编码器侧,可以添加针对色度分量的一个或多个RD成本检查,以选择色度帧内预测模式。当除CCLM的亮度到色度预测模式之外的Cb帧内预测模式可用于CU的色度分量时,CCLM的Cb到Cr预测可用于Cr分量预测。

[0040] 多模型CCLM(MMLM)是CCLM的另一个扩展。顾名思义,MMLM中可以有多于一个的模型,例如可以使用两个模型。在MMLM中,当前块的相邻亮度样本和相邻色度样本可以划分为两组,每一组都可以用作训练集,以导出线性模型(即,可以为特定组导出特定的α和β)。此外,还可以基于与相邻亮度样本相同的分类规则对当前亮度块的样本进行分类。

[0041] 图8示出了将相邻样本划分为两组的示例。阈值可以计算为相邻重建亮度样本的平均值。可以将Rec'L[x,y]<=阈值(Rec'L[x,y]<=Threshold)的相邻样本划分到组1;而将Rec'L[x,y]>阈值(Rec'L[x,y]>Threshold)的相邻样本划分到组2:

[0042] $$\begin{cases} Pred_C[x,y]=\alpha_1\times Rec'_L[x,y]+\beta_1 & if\ Rec'_L[x,y]\le Threshold \\ Pred_C[x,y]=\alpha_2\times Rec'_L[x,y]+\beta_2 & if\ Rec'_L[x,y]>Threshold \end{cases}$$ (等式5)

[0043] 现在将描述仿射运动补偿预测。在HEVC中,可以将平移运动模型应用于运动补偿预测(MCP)。然而,可能存在许多种运动,例如,放大/缩小、旋转、透视运动和其它不规则运动。在VTM4中,可以应用基于块的仿射变换运动补偿预测。如图9所示,块的仿射运动场可以通过两个控制点运动矢量(4参数)或三个控制点运动矢量(6参数)的运动信息来描述。

[0044] 对于4参数仿射运动模型,在块中的样本位置(x,y)处的运动矢量可以导出为:

[0045] $$\begin{cases} mv_x=\frac{mv_{1x}-mv_{0x}}{W}x+\frac{mv_{1y}-mv_{0y}}{W}y+mv_{0x} \\ mv_y=\frac{mv_{1y}-mv_{0y}}{W}x+\frac{mv_{1y}-mv_{0x}}{W}y+mv_{0y} \end{cases}$$ (等式6)。

[0046] 对于6参数仿射运动模型,在块中的样本位置(x,y)处的运动矢量可以导出为:

[0047] $$\begin{cases} mv_x=\frac{mv_{1x}-mv_{0x}}{W}x+\frac{mv_{2x}-mv_{0x}}{H}y+mv_{0x} \\ mv_y=\frac{mv_{1y}-mv_{0y}}{W}x+\frac{mv_{2y}-mv_{0y}}{H}y+mv_{0y} \end{cases}$$ (等式7)。

[0048] 这里,mv_{0x}、mv_{0y}是左上角控制点的运动矢量,mv_{1x}、mv_{1y}是右上角控制点的运动矢量,mv_{2x}、mv_{2y}是左下角控制点的运动矢量。

[0049] 为了简化运动补偿预测,可以应用基于块的仿射变换预测。为了导出每个4×4亮度子块的运动矢量,可以根据上述等式来计算每个子块的中心样本的运动矢量,如图10所示,并将其四舍五入到1/16分数精度。然后,可以应用运动补偿插值滤波器,以利用导出的

运动矢量生成每个子块的预测。色度分量的子块大小也可以设置为4×4。可以将4×4色度子块的MV计算为四个对应的4×4亮度子块的MV的平均值。与平移运动帧间预测一样,也可以有两种仿射运动帧间预测模式:仿射合并模式和仿射高级运动矢量预测(AMVP)模式。

[0050] 现在将描述仿射合并预测。仿射合并(AF_MERGE)模式可以应用于宽度和高度都大于或等于8的CU。在这种模式中,当前CU的控制点运动矢量(CPMV)可以基于空间相邻CU的运动信息生成。最多可以有五个控制点运动矢量预测(CPMVP)候选,并且可以发信号通知索引,以指示要用于当前CU的一个CPMVP候选。以下三种类型的CPMV候选可用于形成仿射合并候选列表:

[0051] (1)从相邻CU的CPMV推断出的继承的仿射合并候选;

[0052] (2)从使用相邻CU的平移MV导出的构造的仿射合并候选CPMVP;以及

[0053] (3)零MV

[0054] 在VTM4中,最多可以有两个继承的仿射候选,这两个继承的仿射候选可以从相邻块的仿射运动模型导出,其中一个从左相邻CU导出,另一个从上相邻CU导出。图11中示出了候选块。对于左预测器(predictor),扫描顺序可以是A0->A1,对于上预测器,扫描顺序可以是B0->B1->B2。可以从每一侧选择第一个继承的候选。不需要在两个继承的候选之间执行修剪检查(pruning check)。当识别出相邻仿射CU时,相邻仿射CU的控制点运动矢量可用于导出当前CU的仿射合并列表中的CPMVP候选。如图12所示,如果以仿射模式对相邻的左下块A进行编码,则可以获得包含块A的CU的左上角、右上角和左下角的运动矢量v_2、v_3和v_4。当使用4参数仿射模型对块A进行编码时,可以根据v_2和v_3来计算当前CU的两个CPMV。在可以使用6参数仿射模型对块A进行编码的情况下,可以根据v_2、v_3和v_4来计算当前CU的三个CPMV。

[0055] 构造的仿射候选是指可以通过组合每个控制点的相邻平移运动信息而构造的候选。控制点的运动信息可以从图13所示的指定的空间相邻块和时间相邻块导出。$CPMV_k$(k=1,2,3,4)表示第k个控制点。对于CPMV1,可以检查B2->B3->A2块,并且可以使用第一个可用块的MV。对于CPMV2,可以检查B1->B0块,对于CPMV3,可以检查A1->A0块。如果TMVP可用,则将其用作CPMV4。

[0056] 在获得四个控制点的MV之后,可以基于该运动信息构造仿射合并候选。可以使用控制点MV的以下组合顺序地构造:{$CPMV_1$,$CPMV_2$,$CPMV_3$}、{$CPMV_1$,$CPMV_2$,$CPMV_4$}、{$CPMV_1$,$CPMV_3$,$CPMV_4$}、{$CPMV_2$,$CPMV_3$,$CPMV_4$}、{$CPMV_1$,$CPMV_2$}、以及{$CPMV_1$,$CPMV_3$}。

[0057] 3个CPMV的组合构造6参数仿射合并候选,2个CPMV的组合构造4参数仿射合并候选。为了避免运动缩放过程,如果控制点的参考索引不同,可以丢弃控制点MV的相关组合。

[0058] 在检查继承的仿射合并候选和构造的仿射合并候选之后,如果列表仍然未满,则可以将零MV插入到列表的末尾。

[0059] 现在将描述仿射AMVP预测。仿射AMVP模式可以用于宽度和高度都大于或等于16的CU。可以在比特流中发信号通知CU级别的仿射标志,以指示是否可以使用仿射AMVP模式,并且可以发信号通知另一标志,以指示使用4参数仿射还是6参数仿射。在这种模式中,可以在比特流中发信号通知当前CU的CPMV与其预测器CPMVP的差。仿射AMVP候选列表的大小可以是2,并且可以通过依次使用以下四种类型的CPMV候选来生成:

[0060] (1)从相邻CU的CPMV推断出的继承的仿射AMVP候选;

[0061] (2)使用相邻CU的平移MV导出的构造的仿射AMVP候选CPMVP;

[0062] (3)来自相邻CU的平移MV;以及

[0063] (4)零MV。

[0064] 继承的仿射AMVP候选的检查顺序可以与继承的仿射合并候选的检查顺序相同。区别是,对于AMVP候选,可以考虑具有与当前块中的参考图片相同的参考图片的仿射CU。当将继承的仿射运动预测值插入到候选列表中时,不需要应用修剪过程。

[0065] 构造的AMVP候选可以从图13所示的指定的空间相邻块导出。可以使用与在仿射合并候选构造中执行的相同的检查顺序。此外,还可以检查相邻块的参考图片索引。可以使用检查顺序中的第一个块,该第一个块可以是帧间编码的,并且具有与当前CU中参考图片相同的参考图片。当使用4参数仿射模式对当前CU进行编码并且mv_0和mv_1均可用时,这两者可以作为一个候选添加到仿射AMVP候选列表中。当使用6参数仿射模式对当前CU进行编码并且所有三个CPMV均可用时,可以将这三个CPMV作为一个候选添加到仿射AMVP候选列表中。否则,将构造的AMVP候选设置为不可用。

[0066] 如果在检查继承的仿射AMVP候选和构造的仿射AMVP候选之后,仿射AMVP候选列表仍然小于2,则按顺序将mv_0、mv_1和mv_2(如果可用的话)添加为平移MV,以预测当前CU的所有控制点MV。最后,如果仿射AMVP候选列表仍然未满,则可以使用零MV来填充仿射AMVP候选列表。

[0067] 现在将描述仿射运动信息的存储。在VTM4中,仿射CU的CPMV可以存储在单独的缓冲器中。存储的CPMV可用于以仿射合并模式和仿射AMVP模式为最近编码的CU生成继承的CPMVP。从CPMV导出的子块MV可用于运动补偿、平移MV的合并/AMVP列表的MV推导、以及解块。

[0068] 为了避免用于附加CPMV的图片行缓冲器,来自上CTU的CU的仿射运动数据继承可以与来自正常相邻CU的继承区别对待。如果用于仿射运动数据继承的候选CU在上CTU行(line)中,则可以将行缓冲器中的左下子块和右下子块MV而不是CPMV用于仿射MVP推导。这样,CPMV可以存储在本地缓冲器中。如果候选CU是6参数仿射编码的,则仿射模型可以降级为4参数模型。如图14中所示,沿着上CTU边界,CU的左下子块运动矢量和右下子块运动矢量可以用于下CTU中的CU的仿射继承。

[0069] 尽管有上述进展,在VTM-4.0中,仿射编码的编码块中的4×4色度子块的MV被计算为四个对应的4×4亮度子块的MV的平均值。然而,对于色度4:4:4格式和4:2:2格式,其中每个4×4色度子块仅与一个或两个4×4亮度子块相关联,用于4×4色度分量的MV推导的当前方案留下了改进的空间,以适应色度4:4:4格式和4:2:2格式的情况。

发明内容

[0070] 根据本公开的方面,一种对视频序列进行编码或解码的方法可以包括:使用4:4:4色度格式对所述视频序列进行编码或解码,或者使用4:2:2色度格式对所述视频序列进行编码或解码;当使用所述4:4:4色度格式对所述视频序列进行编码或解码时,使用除平均运算之外的操作复制一个4×4亮度块的仿射运动矢量,并将所述仿射运动矢量与同位4×4色度块相关联;以及当使用所述4:2:2色度格式对所述视频序列进行编码或解码时,将每个4×4色度块与两个4×4同位亮度块相关联,使得一个4×4色度块的仿射运动矢量是两个同位亮度块的运动矢量的平均值。

[0071] 根据本公开的方面,所述方法可以进一步包括:将当前4×4色度块划分为四个2×2子块,导出左上2×2色度子块的同位亮度块的第一仿射运动矢量,导出右下2×2色度子块的同位亮度块的第二仿射运动矢量,以及使用所述第一仿射运动矢量和所述第二仿射运动矢量的平均值导出所述当前4×4色度块的仿射运动矢量。

[0072] 根据本公开的方面,所述方法可以进一步包括:对准用于亮度分量和色度分量之间的运动补偿的插值滤波器。

[0073] 根据本公开的方面,当使用4:2:0色度格式输入视频序列时,所述方法可以进一步包括:对亮度分量和色度分量应用8抽头插值滤波器。

[0074] 根据本公开的方面,所述方法可以进一步包括:将分量Y、Cb和Cr编码为三个单独的树,其中,所述三个单独的树中的每个树对所述分量Y、Cb和Cr中的一个分量进行编码。

[0075] 根据本公开的方面,可以对I切片或I图块群组执行所述编码为三个单独的树。

[0076] 根据本公开的方面,对于不同的颜色分量,允许的最大变换大小可以是相同的。

[0077] 根据本公开的方面,当使用所述4:2:2色度格式对所述视频序列进行编码或解码时,不同颜色可以分量中的最大垂直大小是相同的,并且色度分量的最大水平变换大小可以是亮度分量的最大水平变换大小的一半。

[0078] 根据本公开的方面,可以将位置相关预测器组合(PDPC)、多变换选择(MTS)、不可分离二次变换(NSST)、帧内子分区(ISP)和多参考行(MRL)帧内预测中的至少一个应用于亮度分量和色度分量。

[0079] 根据本公开的方面,当将所述多参考行(MRL)帧内预测应用于所述亮度分量和所述色度分量时,并且当使用所述4:4:4色度格式对所述视频序列进行编码或解码时,所述方法可以进一步包括:选择用于帧内预测的第N个参考,并且使用相同的参考行而无需为色度分量显式发信号通知;当将所述帧内子分区(ISP)应用于所述亮度分量和所述色度分量时,所述方法可以进一步包括:对于分量Y、Cb和Cr,在当前块的块级别应用所述帧内子分区ISP;以及当不同的树用于不同的颜色分量时,所述方法可以进一步包括:从同位Y分量隐式地导出用于U分量和V分量的编解码参数而无需发信号通知。

[0080] 根据本公开的方面,一种对视频序列进行编码或解码的设备可以包括:至少一个存储器,被配置为存储程序代码;至少一个处理器,被配置为读取所述程序代码,并按照所述程序代码的指示进行操作,所述程序代码包括:第一编码或解码代码,被配置为使所述至少一个处理器使用4:4:4色度格式对所述视频序列进行编码或解码,或者使用4:2:2色度格式对所述视频序列进行编码或解码;当所述第一编码或解码代码被配置为使所述至少一个处理器使用所述4:4:4色度格式对所述视频序列进行编码或解码时,所述第一编码或解码代码可以进一步包括代码,被配置为使所述至少一个处理器使用除平均运算之外的操作复制一个4×4亮度块的仿射运动矢量,并将所述仿射运动矢量与同位4×4色度块相关联;以及当所述第一编码或解码代码被配置为使所述至少一个处理器使用所述4:2:2色度格式对所述视频序列进行编码或解码时,所述第一编码或解码代码可以进一步包括代码,被配置为使所述至少一个处理器将每个4×4色度块与两个4×4同位亮度块相关联,使得一个4×4色度块的仿射运动矢量是两个同位亮度块的运动矢量的平均值。

[0081] 根据本公开的方面,所述第一编码或解码代码可以进一步包括代码,被配置为使所述至少一个处理器将当前4×4色度块划分为四个2×2子块,导出左上2×2色度子块的同

位亮度块的第一仿射运动矢量,导出右下2×2色度子块的同位亮度块的第二仿射运动矢量,以及使用所述第一仿射运动矢量和所述第二仿射运动矢量的平均值导出所述当前4×4色度块的仿射运动矢量。

[0082] 根据本公开的方面,所述第一编码或解码代码可以进一步包括代码,被配置为使所述至少一个处理器对准用于亮度分量和色度分量之间的运动补偿的插值滤波器。

[0083] 根据本公开的方面,当所述第一编码或解码代码被配置为使所述至少一个处理器使用4:2:2色度格式输入视频序列时,所述第一编码或解码代码可以进一步包括代码,被配置为使所述至少一个处理器对亮度分量和色度分量应用8抽头插值滤波器。

[0084] 根据本公开的方面,所述第一编码或解码代码可以进一步包括代码,被配置为使所述至少一个处理器将分量Y、Cb和Cr编码为三个单独的树,其中,所述三个单独的树中的每个树对所述分量Y、Cb和Cr中的一个分量进行编码。

[0085] 根据本公开的方面,可以对I切片或I图块群组执行所述编码为三个单独的树。

[0086] 根据本公开的方面,所述第一编码或解码代码可以进一步包括代码,被配置为使所述至少一个处理器对于不同的颜色分量,允许最大变换大小是相同的。

[0087] 根据本公开的方面,当所述第一编码或解码代码被配置为使所述至少一个处理器使用所述4:2:2色度格式对所述视频序列进行编码或解码时,所述第一编码或解码代码可以进一步包括代码,被配置为使所述至少一个处理器将不同颜色分量中的最大垂直大小设置为相同,并且将色度分量的最大水平变换大小设置为亮度分量的最大水平变换大小的一半。

[0088] 根据本公开的方面,所述第一编码或解码代码可以进一步包括代码,被配置为使所述至少一个处理器将位置相关预测器组合(PDPC)、多变换选择(MTS)、不可分离二次变换(NSST)、帧内子分区(ISP)和多参考行(MRL)帧内预测中的至少一个应用于亮度分量和色度分量。

[0089] 根据本公开的方面,可以提供一种非易失性计算机可读介质,用于存储程序代码,所述程序代码包括一个或多个指令,当所述指令由设备的一个或多个处理器执行时,使所述一个或多个处理器执行本公开实施例所述的方法。

[0090] 尽管已经分别描述了上述方法,设备和非易失性计算机可读介质,但是这些描述并不旨在对其使用范围或功能提出任何限制。实际上,可以这些方法、设备和非易失性计算机可读介质可以在本公开的其他方面中进行组合。

附图说明

[0091] 根据以下详细描述和附图,所公开的主题的其他特征、性质和各种优点将进一步明确,其中:

[0092] 图1A是根据实施例的分区的编码树单元的示意图;

[0093] 图1B是根据实施例的编码树单元的示意图;

[0094] 图2A是根据实施例的编码树单元的示意图;

[0095] 图2B是根据实施例的编码树单元的示意图;

[0096] 图3是根据实施例的不同YUV格式的示意图;

[0097] 图4是根据实施例的不同亮度值的示意图;

[0098] 图5A是根据实施例的在跨分量线性建模(cross-component linear modeling,CCLM)中使用的样本的示意图;

[0099] 图5B是根据实施例的在跨分量线性建模中使用的样本的示意图;

[0100] 图6A是根据实施例的在跨分量线性建模中使用的样本的示意图;

[0101] 图6B是根据实施例的在跨分量线性建模中使用的样本的示意图;

[0102] 图7A是根据实施例的在跨分量线性建模中使用的样本的示意图;

[0103] 图7B是根据实施例的在跨分量线性建模中使用的样本的示意图;

[0104] 图8是根据实施例的使用多模型CCLM的分类的示例;

[0105] 图9A是根据实施例的块的仿射运动场的示例;

[0106] 图9B是根据实施例的块的仿射运动场的示例;

[0107] 图10是根据实施例的仿射运动矢量场的示例;

[0108] 图11是根据实施例的用于预测的候选块的示例;

[0109] 图12是根据实施例的用于预测的候选块的示例;

[0110] 图13是根据实施例的用于预测的候选块的示例;

[0111] 图14是根据实施例的运动矢量使用的示例;

[0112] 图15是根据实施例的通信系统的简化框图;

[0113] 图16是根据实施例的流式传输环境的示意图;

[0114] 图17是根据实施例的视频解码器的框图;

[0115] 图18是根据实施例的视频编码器的框图;

[0116] 图19是根据实施例的对视频序列进行编码或解码的示例过程的流程图;

[0117] 图20是根据实施例的计算机系统的示意图。

具体实施方式

[0118] 图15示出了根据本公开实施例的通信系统(400)的简化框图。通信系统(400)可以包括通过网络(450)互联的至少两个终端装置(410,420)。对于单向数据传输,第一终端装置(410)可在本地位置对视频数据进行编码,以通过网络(450)传输到另一终端装置(420)。第二终端装置(420)可从网络(450)接收另一终端装置的已编码视频数据,对已编码视频数据进行解码并显示恢复的视频数据。单向数据传输在媒体服务等应用中是较常见的。

[0119] 图15示出了支持已编码视频的双向传输的第二对终端装置(430,440),所述双向传输可例如在视频会议期间发生。对于双向数据传输,每个终端装置(430,440)可对在本地位置采集的视频数据进行编码,以通过网络(450)传输到另一终端装置。每个终端装置(430,440)还可接收由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并在本地显示设备上显示恢复的视频数据。

[0120] 在图15中,终端装置(410-440)可为服务器、个人计算机和智能手机,但本公开的原理不限于此。本公开的实施例适用于膝上型计算机、平板电脑、媒体播放器和/或专用视频会议设备。网络(450)表示在终端装置(410-440)之间传送已编码视频数据的任何数目的网络,包括例如有线和/或无线通信网络。通信网络(450)可在电路交换和/或分组交换信道中交换数据。代表性网络包括电信网络、局域网、广域网和/或互联网。出于本讨论的目的,除非在下文中有所解释,否则网络(450)的架构和拓扑对于本公开的操作来说可能是无关

紧要的。

[0121] 作为所公开主题的应用实施例,图16示出了视频解码器和编码器在流式传输环境中的放置方式。所公开的主题可同等地适用于其它支持视频的应用,包括例如视频会议、数字TV、在包括CD、DVD、存储棒等的数字介质上存储压缩视频等等。

[0122] 流式传输系统可包括采集子系统(513),所述采集子系统可包括诸如数码相机等视频源(501),所述视频源创建例如未压缩的视频样本流(502)。较于已编码的视频比特流,样本流(502)被描绘为粗线,以强调其为高数据量的视频样本流,样本流(502)可由耦接到相机(501)的编码器(503)处理。编码器(503)可包括硬件、软件或软硬件组合以实现或实施如下文更详细地描述的所公开主题的各方面。相较于样本流,已编码的视频比特流(504)被描绘为细线,以强调较低数据量的已编码的视频比特流,其可存储在流式传输服务器(505)上以供将来使用。一个或多个流式传输客户端(506,508)可访问流式传输服务器(505)以检索已编码的视频比特流(504)的副本(507,509)。客户端(506)可包括视频解码器(510)。视频解码器(510)对已编码的视频比特流的传入副本(507)进行解码,且产生可在显示器(512)或另一呈现装置(未示出)上呈现的输出视频样本流(511)。在一些流式传输系统中,可以根据某些视频编解码/压缩标准对视频比特流(504,507,509)进行编码。这些标准的示例包括H.265HEVC。正在开发的视频编解码标准被非正式地称为多功能视频编解码(VVC)。所公开的主题可以在VVC的上下文中使用。

[0123] 图17是根据本发明实施例的视频解码器(510)的功能框图。

[0124] 接收器(610)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,一次接收一个已编码视频序列,其中每个已编码视频序列的解码独立于其它已编码视频序列。可从信道(612)接收已编码视频序列,所述信道可以是通向存储已编码的视频数据的存储装置的硬件/软件链路。接收器(610)可接收已编码的视频数据以及其它数据,例如,可转发到它们各自的使用实体(未标示)的已编码音频数据和/或辅助数据流。接收器(610)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(615)可耦接在接收器(610)与熵解码器/解析器(620)(此后称为“解析器”)之间。当接收器(610)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(615),或可以将所述缓冲存储器做得较小。当然,为了在互联网等分组网络上使用,也可能需要缓冲存储器(615),所述缓冲存储器可相对较大且可有利地具有自适应性大小。

[0125] 视频解码器(510)可包括解析器(620)以根据熵编码视频序列重建符号(621)。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,以及用以控制诸如显示器(512)的显示装置的潜在信息,所述显示装置不是解码器的组成部分,但可耦接到解码器,如图17所示。用于显示装置的控制信息可以是辅助增强信息(Supplementary Enhancement Information,SEI消息)或视频可用性信息(Video Usability Information,VUI)的参数集片段(未标示)的形式。解析器(620)可对接收到的已编码视频序列进行解析/熵解码。已编码视频序列的编解码可根据视频编解码技术或标准进行,且可遵循本领域技术人员公知的原理,包括可变长度编解码、霍夫曼编解码(Huffman coding)、具有或不具有上下文灵敏度的算术编解码等等。解析器(620)可基于对应于群组的至少一个参数,从已编码视频序列提取用于视频解码器中的像素的子群中的至少一个子群的子群参数集。子群可包括图片群组

(Group of Pictures,GOP)、图片、图块(tile)、切片(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。熵解码器/解析器还可从已编码视频序列提取信息,例如变换系数、量化器参数(QP)值、运动矢量等等。

[0126] 解析器(620)可对从缓冲存储器(615)接收的视频序列执行熵解码/解析操作,从而创建符号(621)。解析器(620)可以接收已编码数据,并且选择性地解码特定符号(621)。此外,解析器(620)可以确定是否将特定符号(621)提供给运动补偿预测单元(653)、缩放器/逆变换单元(651)、帧内预测单元(652)或环路滤波器单元(656)。

[0127] 取决于已编码视频图片或一部分已编码视频图片(例如:帧间图片和帧内图片、帧间块和帧内块)的类型以及其它因素,符号(621)的重建可涉及多个不同单元。涉及哪些单元以及涉及方式可由解析器(620)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(620)与下文的多个单元之间的此类子群控制信息流。

[0128] 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中,这些单元中的许多单元彼此紧密交互并且可以至少部分地彼此集成。然而,出于描述所公开的主题的目的,概念上细分成下文的功能单元是适当的。

[0129] 第一单元是缩放器/逆变换单元(651)。缩放器/逆变换单元(651)从解析器(620)接收作为符号(621)的量化变换系数以及控制信息,包括使用哪种变换方式、块大小、量化因子、量化缩放矩阵等。缩放器/逆变换单元(651)可输出包括样本值的块,所述样本值可输入到聚合器(655)中。

[0130] 在一些情况下,缩放器/逆变换单元(651)的输出样本可属于帧内编码块;即:不使用来自先前重建的图片的预测性信息,但可使用来自当前图片的先前重建部分的预测性信息的块。此类预测性信息可由帧内图片预测单元(652)提供。在一些情况下,帧内图片预测单元(652)采用从(部分重建的)当前图片(658)提取的周围已重建的信息生成大小和形状与正在重建的块相同的块。在一些情况下,聚合器(655)基于每个样本,将帧内预测单元(652)生成的预测信息添加到由缩放器/逆变换单元(651)提供的输出样本信息中。

[0131] 在其它情况下,缩放器/逆变换单元(651)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(653)可访问参考图片存储器(657)以提取用于预测的样本。在根据符号(621)对提取的样本进行运动补偿之后,这些样本可由聚合器(655)添加到缩放器/逆变换单元的输出(在这种情况下被称作残差样本或残差信号),从而生成输出样本信息。运动补偿预测单元从参考图片存储器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(621)的形式而供运动补偿预测单元(653)使用,所述符号(621)例如是包括X、Y和参考图片分量。运动补偿还可包括在使用子样本精确运动矢量时,从参考图片存储器(657)提取的样本值的内插、运动矢量预测机制等等。

[0132] 聚合器(655)的输出样本可在环路滤波器单元(656)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比特流中的参数,且所述参数作为来自解析器(620)的符号(621)可用于环路滤波器单元(656)。然而,视频压缩技术还可响应于在解码已编码图片或已编码视频序列的先前(按解码次序)部分期间获得的元信息,以及响应于先前重建且经过环路滤波的样本值。

[0133] 环路滤波器单元(656)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(657),以用于后续的帧间图片预测。

[0134] 一旦完全重建,某些已编码图片就可用作参考图片以用于将来预测。一旦已编码图片被完全重建,且已编码图片(通过例如解析器(620))被识别为参考图片,则当前图片(658)可变为参考图片存储器(657)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存储器。

[0135] 视频解码器(510)可根据例如H.265HEVC标准中记录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在视频压缩技术文献或标准、特别是配置文件中所规定的视频压缩技术或标准的语法的意义上,已编码视频序列可符合所使用的视频压缩技术或标准指定的语法。对于合规性,还要求已编码视频序列的复杂度处于视频压缩技术或标准的层级所限定的范围内。在一些情况下,层级限制最大图片大小、最大帧率、最大重建取样率(以例如每秒兆(mega)个样本为单位进行测量)、最大参考图片大小等等。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。

[0136] 在实施例中,接收器(610)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、冗余切片、冗余图片、前向纠错码等形式。

[0137] 图18是根据本公开实施例的视频编码器(503)的功能框图。

[0138] 视频编码器(503)可从视频源(501)(并非该编码器的一部分)接收视频样本,所述视频源可采集将由视频编码器(503)编码的视频图像。

[0139] 视频源(501)可提供将由视频编码器(503)编码的呈数字视频样本流形式的源视频序列,所述数字视频样本流可具有任何合适位深度(例如:8位、10位、12位……)、任何色彩空间(例如BT.601Y CrCB、RGB……)和任何合适取样结构(例如Y CrCb 4:2:0、Y CrCb4:4:4)。在媒体服务系统中,视频源(501)可以是存储先前已准备的视频的存储装置。在视频会议系统中,视频源(501)可以为采集本地图像信息作为视频序列的相机。可将视频数据提供为多个单独的图片,当按顺序观看时,这些图片被赋予运动。图片自身可构建为空间像素阵列,其中取决于所用的取样结构、色彩空间等,每个像素可包括一个或多个样本。本领域技术人员可以容易地理解像素和样本之间的关系。下文侧重于描述样本。

[0140] 根据实施例,视频编码器(503)可实时或在由应用所要求的任何其它时间约束下,将源视频序列的图片编码且压缩成已编码视频序列(743)。施行适当的编码速度是控制器(750)的一个功能。控制器(750)控制如下文所描述的其它功能单元且在功能上耦接到这些单元。为了简洁起见,图中未标示耦接。由控制器设置的参数可包括速率控制相关参数(例如,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。本领域技术人员可以容易地识别控制器(750)的其他功能,因为它们可能属于针对特定系统设计而优化的视频编码器(503)。

[0141] 一些视频编码器以本领域技术人员容易识别为“编码环路”方式进行操作。作为简单的描述,编码环路可包括编码器(730)(下文称为“源编码器(source coder)”,其负责基于待编码的输入图片和参考图片创建符号)的编码部分、和嵌入于视频编码器(503)中的

(本地)解码器(733)。“本地”解码器(733)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(因为在所公开的主题所考虑的视频压缩技术中,符号与已编码视频码流之间的任何压缩是无损的)。将重建的样本流输入到参考图片存储器(734)。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中的内容在本地编码器与远程编码器之间也是按比特位精确对应的。换句话说,编码器的预测部分“看到”的参考图片样本与解码器将在解码期间使用预测时所“看到”的样本值完全相同。这种参考图片同步性基本原理(以及在例如因信道误差而无法维持同步性的情况下产生的漂移)是本领域技术人员公知的。

[0142] “本地”解码器(733)的操作可与已在上文结合图17详细描述的“远程”解码器(510)相同。然而,另外简要参考图17,当符号可用且熵编码器(745)和解析器(620)能够无损地将符号编码/解码为已编码视频序列时,视频解码器(510)的熵解码部分(包括信道(612)、接收器(610)、缓冲存储器(615)和解析器(620))可能无法完全在本地解码器(733)中实施。

[0143] 此时可以观察到,除存在于解码器中的解析/熵解码之外的任何解码器技术,也必定以基本上相同的功能形式存在于对应的编码器中。对编码器技术的描述可以简写,因为它们与全面描述的解码器技术互逆。仅在某些区域中需要更详细的描述,并且在下文提供。

[0144] 作为操作的一部分,源编码器(730)可执行运动补偿预测编码。参考来自视频序列中被指定为“参考帧”的一个或多个先前已编码帧,所述运动补偿预测编码对输入帧进行预测性编码。以此方式,编码引擎(732)对输入帧的像素块与参考帧的像素块之间的差异进行编码,所述参考帧可被选作所述输入帧的预测参考。

[0145] 本地视频解码器(733)可基于源编码器(730)创建的符号,对可指定为参考帧的帧的已编码视频数据进行解码。编码引擎(732)的操作可以有利地为有损过程。当已编码视频数据可在视频解码器(图17中未示)处被解码时,重建的视频序列通常可以是带有一些误差的源视频序列的副本。本地视频解码器(733)复制解码过程,所述解码过程可由视频解码器对参考帧执行,且可使重建的参考帧存储在参考图片存储器(734)中。以此方式,编码器(503)可在本地存储重建的参考帧的副本,所述副本与将由远端视频解码器获得的重建参考帧具有共同内容(不存在传输误差)。

[0146] 预测器(735)可针对编码引擎(732)执行预测搜索。即,对于将要编码的新帧,预测器(735)可在参考图片存储器(734)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(735)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(735)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(734)中存储的多个参考图片取得的预测参考。

[0147] 控制器(750)可管理源编码器(730)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。

[0148] 可在熵编码器(745)中对所有上述功能单元的输出进行熵编码。熵编码器可根据例如霍夫曼编码、可变长度编码、算术编码等本领域技术人员公知的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。

[0149] 传输器(740)可缓冲由熵编码器(745)创建的已编码视频序列,从而为通过通信信

道(760)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(740)可将来自源编码器(730)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)。

[0150] 控制器(750)可管理视频编码器(503)的操作。在编码期间,控制器(750)可以为每个已编码图片分配某一已编码图片类型,但这可能影响可应用于相应的图片的编码技术。例如,通常可将图片分配为以下任一种图片类型。

[0151] 帧内图片(I图片),其可以是不将序列中的任何其它帧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。本领域技术人员知道I图片的那些变体以及它们各自的应用和特征。

[0152] 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。

[0153] 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。

[0154] 源图片通常可在空间上细分成多个样本块(例如,4×4、8×8、4×8或16×16个样本的块),且逐块进行编码。这些块可参考其它(已编码)块进行预测编码,根据应用于块的相应图片的编码分配来确定所述其它块。举例来说,I图片的块可进行非预测编码,或所述块可参考同一图片的已经编码的块来进行预测编码(空间预测或帧内预测)。P图片的像素块可参考一个先前编码的参考图片通过空间预测或通过时域预测进行预测编码。B图片的块可参考一个或两个先前编码的参考图片通过空间预测或通过时域预测进行预测编码。

[0155] 视频编码器(503)可根据例如H.265HEVC的预定视频编码技术或标准执行编码操作。在操作中,视频编码器(503)可执行各种压缩操作,包括利用输入视频序列中的时间和空间冗余的预测编码操作。因此,已编码视频数据可符合所用视频编码技术或标准指定的语法。

[0156] 在实施例中,传输器(740)可传输附加数据和已编码的视频。源编码器(730)可以包括诸如已编码视频序列的一部分的此类数据。附加数据可包括时间/空间/SNR增强层、冗余图片和切片等其它形式的冗余数据、SEI消息、VUI参数集片段等。

[0157] 本公开针对几种块分区方法,其中在用于视频编码的树分割期间考虑运动信息。更具体地,本公开中的技术涉及用于基于运动场信息的灵活树结构的树分割方法。本公开中所提出的技术可应用于同质(homogenous)导出的运动场及异质(heterogeneous)导出的运动场。

[0158] 如果块的导出的运动场可用于该块中的所有子块,并且导出的运动场中的所有运动矢量相似(例如,运动矢量共用同一参考帧,并且运动矢量之间的绝对差均低于某个阈值),则将该块的导出的运动场定义为同质的。该阈值可以在比特流中发信号通知或预定义。

[0159] 如果块的导出的运动场可用于该块中的所有子块,并且导出的运动场中的运动矢量不相似(例如,至少一个运动矢量参考其它运动矢量未参考的参考帧,或者场中的两个运动矢量之间的至少一个绝对差大于发信号通知的或预定义的阈值),则将该块的导出的运

动场定义为异质的。

[0160] 图19是对视频序列进行编码或解码的示例过程(800)的流程图。在一些实施方式中,图19的一个或多个过程框可以由解码器(510)执行。在一些实施方式中,图19的一个或多个过程框可以由与解码器(510)分离或包括解码器(510)的另一设备或一组设备执行。

[0161] 如图19所示,过程(800)可以包括:使用4:4:4色度格式或4:2:2色度格式对视频序列进行编码或解码(810)。

[0162] 当过程(800)包括使用4:4:4色度格式对视频序列进行编码或解码时,如图19进一步所示,过程(800)可以进一步包括:使用除平均运算之外的操作复制一个4×4色度块的仿射运动矢量(820)。

[0163] 当过程(800)包括使用4:2:2色度格式对视频序列进行编码或解码时,如图19进一步所示,过程(800)可以进一步包括:将每个4x4色度块与两个4x4同位亮度块相关联,使得一个4x4色度块的仿射运动矢量是两个同位亮度块的运动矢量的平均值(830)。

[0164] 尽管图19示出了过程(800)的示例框,在一些实施方式中,过程(800)可以包括比图19中所描绘的更多的框、更少的框、不同的框或不同排列的框。另外或可替代地,过程(800)的两个或更多个框可以并行执行。

[0165] 此外,所提出的方法可以由处理电路(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易失性计算机可读介质中的程序,以执行所提出的方法中的一个或多个。

[0166] 上述技术可以通过计算机可读指令实现为计算机软件,并且物理地存储在一个或多个计算机可读介质中。例如,图20示出了计算机系统(900),其适于实现所公开主题的某些实施例。

[0167] 所述计算机软件可通过任何合适的机器代码或计算机语言进行编码,通过汇编、编译、链接等机制创建包括指令的代码,所述指令可由计算机中央处理单元(CPU),图形处理单元(GPU)等直接执行或通过译码、微代码等方式执行。

[0168] 所述指令可以在各种类型的计算机或其组件上执行,包括例如个人计算机、平板电脑、服务器、智能手机、游戏设备、物联网设备等。

[0169] 图20所示的用于计算机系统(900)的组件本质上是示例性的,并不用于对实现本公开实施例的计算机软件的使用范围或功能进行任何限制。也不应将组件的配置解释为与计算机系统(900)的示例性实施例中所示的任一组件或其组合具有任何依赖性或要求。

[0170] 计算机系统(900)可以包括某些人机界面输入设备。这种人机界面输入设备可以通过触觉输入(如:键盘输入、滑动、数据手套移动)、音频输入(如:声音、掌声)、视觉输入(如:手势)、嗅觉输入(未示出),对一个或多个人类用户的输入做出响应。所述人机界面设备还可用于捕获某些媒体,气与人类有意识的输入不必直接相关,如音频(例如:语音、音乐、环境声音)、图像(例如:扫描图像、从静止影像相机获得的摄影图像)、视频(例如二维视频、包括立体视频的三维视频)。

[0171] 人机界面输入设备可包括以下中的一个或多个(仅绘出其中一个):键盘(901)、鼠标(902)、触控板(903)、触摸屏(910)、数据手套、操纵杆(905)、麦克风(906)、扫描仪(907)、照相机(908)。

[0172] 计算机系统(900)还可以包括某些人机界面输出设备。这种人机界面输出设备可

以通过例如触觉输出、声音、光和嗅觉/味觉来刺激一个或多个人类用户的感觉。这样的人机界面输出设备可包括触觉输出设备(例如通过触摸屏(910)、数据手套或操纵杆(905)的触觉反馈,但也可以有不用作输入设备的触觉反馈设备)、音频输出设备(例如,扬声器(909)、耳机(未示出))、视觉输出设备(例如,包括阴极射线管(CRT)屏幕、液晶显示(LCD)屏幕、等离子屏幕、有机发光二极管(OLED)屏幕的屏幕(910),其中每一个都具有或没有触摸屏输入功能、每一个都具有或没有触觉反馈功能——其中一些可通过诸如立体画面输出的手段输出二维视觉输出或三维以上的输出;虚拟现实眼镜(未示出)、全息显示器和放烟箱(未示出))以及打印机(未示出)。

[0173] 计算机系统(900)还可以包括人可访问的存储设备及其相关介质,如包括具有CD/DVD的高密度只读/可重写式光盘(CD/DVD ROM/RW)(920)或类似介质(921)的光学介质、拇指驱动器(922)、可移动硬盘驱动器或固体状态驱动器(923),诸如磁带和软盘(未示出)的传统磁介质,诸如安全软件保护器(未示出)等的基于ROM/ASIC/PLD的专用设备,等等。

[0174] 本领域技术人员还应当理解,结合所公开的主题使用的术语“计算机可读介质”不包括传输介质、载波或其它瞬时信号。

[0175] 计算机系统(900)还可以包括通往一个或多个通信网络的一个或多个接口。例如,网络可以是无线的、有线的、光学的。网络还可为局域网、广域网、城域网、车载网络和工业网络、实时网络、延迟容忍网络等等。网络还包括以太网、无线局域网、蜂窝网络(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)等局域网、电视有线或无线广域数字网络(包括有线电视、卫星电视、和地面广播电视)、车载和工业网络(包括CANBus)等等。某些网络通常需要外部网络接口适配器,用于连接到某些通用数据端口或外围总线(949)(例如,计算机系统(900)的通用串行总线(USB)端口);其它系统通常通过连接到如下所述的系统总线集成到计算机系统(900)的核心(例如,以太网接口集成到PC计算机系统或蜂窝网络接口集成到智能电话计算机系统)。通过使用这些网络中的任何一个,计算机系统(900)可以与其它实体进行通信。所述通信可以是单向的,仅用于接收(例如,无线电视),单向的仅用于发送(例如CAN总线到某些CAN总线设备),或双向的,例如通过局域或广域数字网络到其它计算机系统。上述的每个网络和网络接口可使用某些协议和协议栈。

[0176] 上述的人机界面设备、人可访问的存储设备以及网络接口可以连接到计算机系统(900)的核心(940)。

[0177] 核心(940)可包括一个或多个中央处理单元(CPU)(941)、图形处理单元(GPU)(942)、以现场可编程门阵列(FPGA)(943)形式的专用可编程处理单元、用于特定任务的硬件加速器(944)等。这些设备以及只读存储器(ROM)(945)、随机存取存储器(RAM)(946)、内部大容量存储器(例如内部非用户可存取硬盘驱动器、固态驱动器(SSD)等)(947)等可通过系统总线(948)进行连接。在某些计算机系统中,可以以一个或多个物理插头的形式访问系统总线(948),以便可通过额外的中央处理单元、图形处理单元等进行扩展。外围装置可直接附接到核心的系统总线(948),或通过外围总线(949)进行连接。外围总线的体系结构包括外围组件互联(PCI)、USB等。

[0178] CPU(941)、GPU(942)、FPGA(943)和加速器(944)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(945)或RAM(946)中。过渡数

据也可以存储在RAM(946)中,而永久数据可以存储在例如内部大容量存储器(947)中。通过使用高速缓冲存储器可实现对任何存储器设备的快速存储和检索,高速缓冲存储器可与一个或多个CPU(941)、GPU(942)、大容量存储器(947)、ROM(945)、RAM(946)等紧密关联。

[0179] 所述计算机可读介质上可具有计算机代码,用于执行各种计算机实现的操作。介质和计算机代码可以是为本公开的目的而特别设计和构造的,也可以是计算机软件领域的技术人员所熟知和可用的介质和代码。

[0180] 作为实施例而非限制,具有体系结构(900)的计算机系统,特别是核心(940),可以作为处理器(包括CPU、GPU、FPGA、加速器等)提供执行包含在一个或多个有形的计算机可读介质中的软件的功能。这种计算机可读介质可以是与上述的用户可访问的大容量存储器相关联的介质,以及具有非易失性的核心(940)的特定存储器,例如核心内部大容量存储器(947)或ROM(945)。实现本公开的各种实施例的软件可以存储在这种设备中并且由核心(940)执行。根据特定需要,计算机可读介质可包括一个或一个以上存储设备或芯片。该软件可以使得核心(940)特别是其中的处理器(包括CPU、GPU、FPGA等)执行本文所述的特定过程或特定过程的特定部分,包括定义存储在RAM(946)中的数据结构以及根据软件定义的过程来修改这种数据结构。另外或作为替代,计算机系统可以提供逻辑硬连线或以其它方式包含在电路(例如,加速器(944))中的功能,该电路可以代替软件或与软件一起运行以执行本文所述的特定过程或特定过程的特定部分。在适当的情况下,对软件的引用可以包括逻辑,反之亦然。在适当的情况下,对计算机可读介质的引用可包括存储执行软件的电路(如集成电路(IC)),包含执行逻辑的电路,或两者兼备。本公开包括任何合适的硬件和软件组合。

[0181] 虽然本公开已对多个示例性实施例进行了描述,但实施例的各种变更、排列和各种等同替换均属于本公开的范围内。因此应理解,本领域技术人员能够设计多种系统和方法,所述系统和方法虽然未在本文中明确示出或描述,但其体现了本公开的原则,因此属于本公开的精神和范围之内。

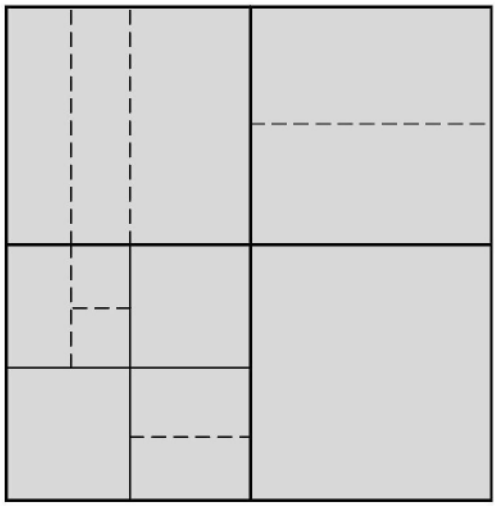


图1A

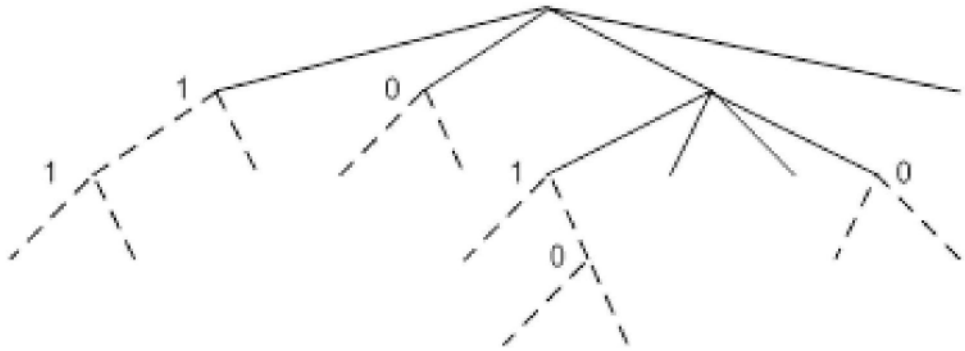


图1B

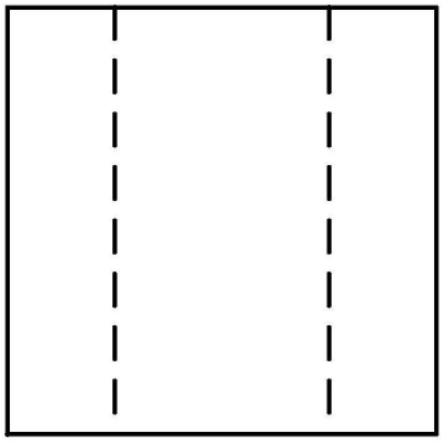


图2A

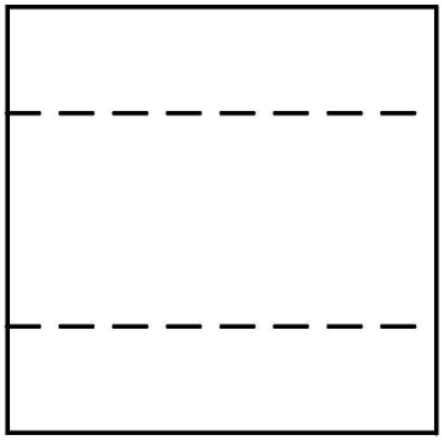


图2B

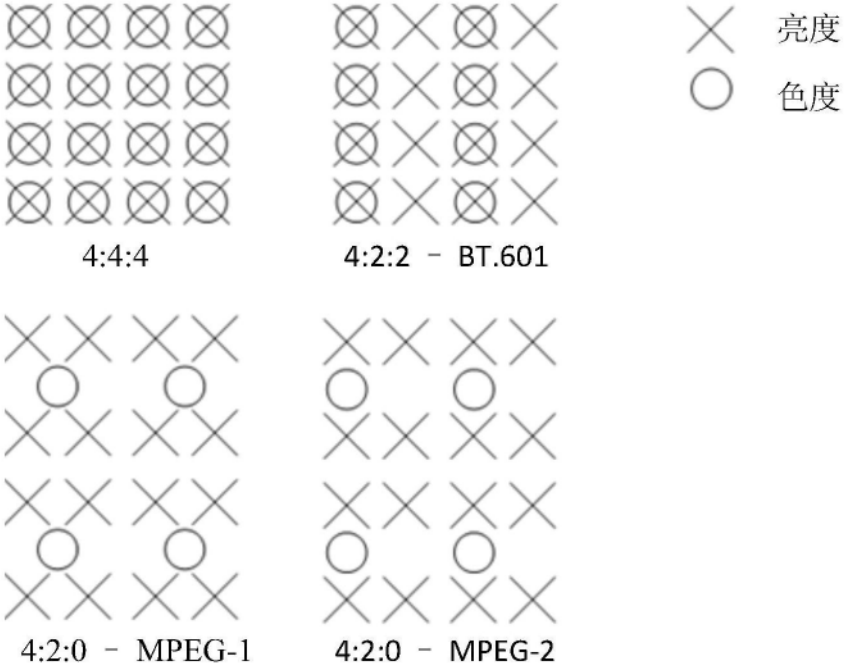


图3

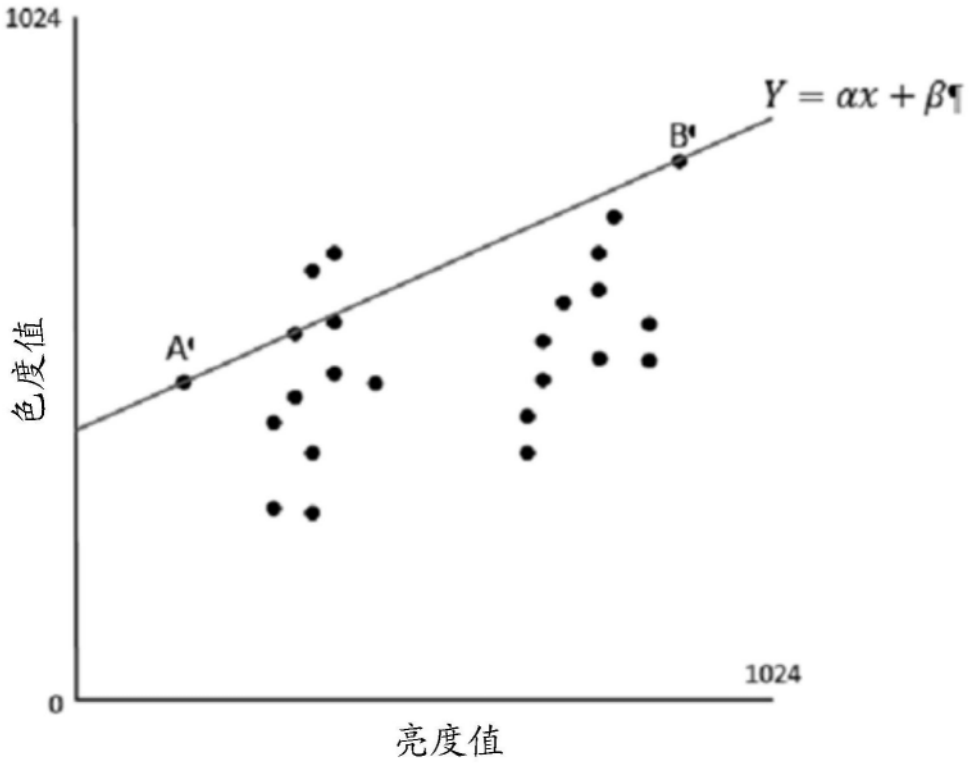


图4

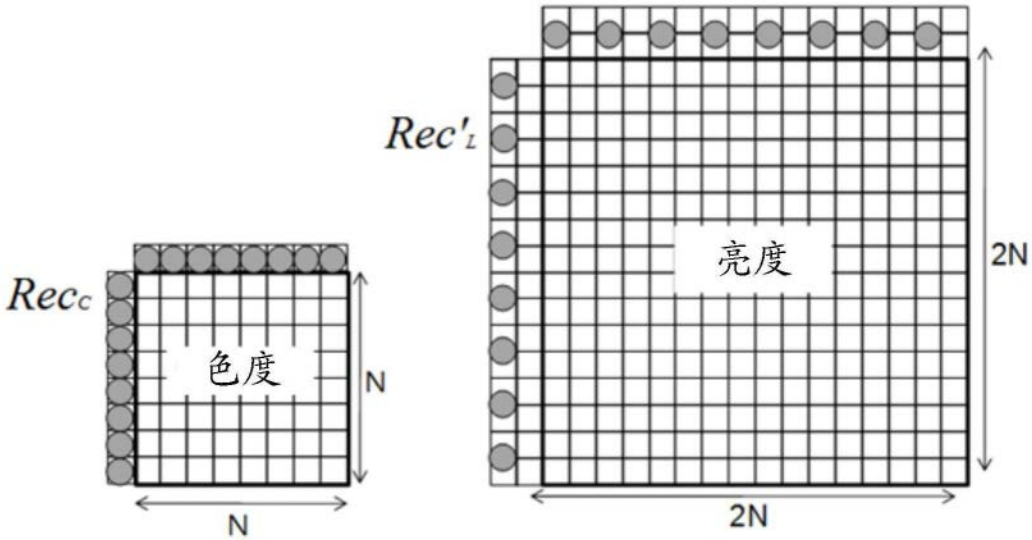


图 5A 图 5B

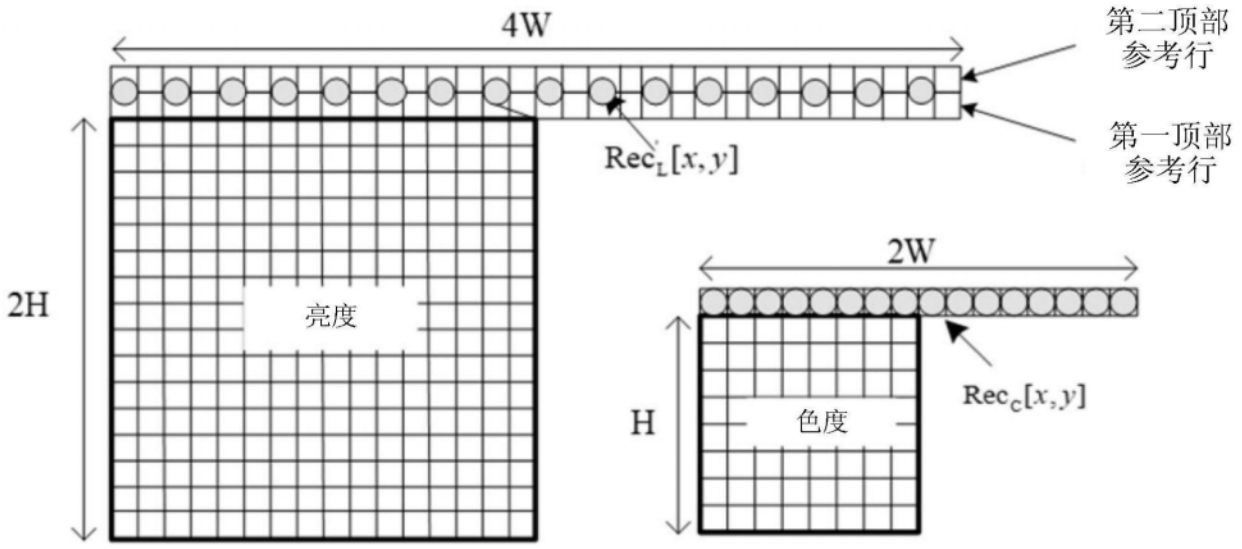

4W
第二顶部参考行
第一顶部参考行
$Rec_L'[x,y]$
2H
亮度
2W
H
色度
$Rec_C[x,y]$

图 6A 图 6B

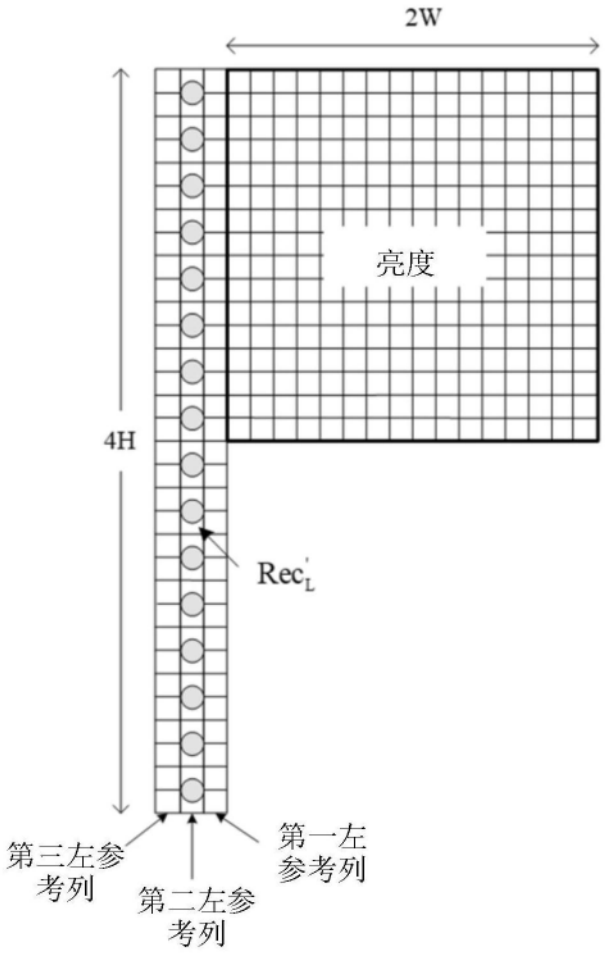

2W
亮度
4H
Rec_L'
第三左参考列
第二左参考列
第一左参考列

图7A

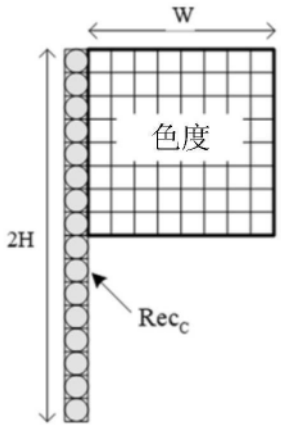


图7B

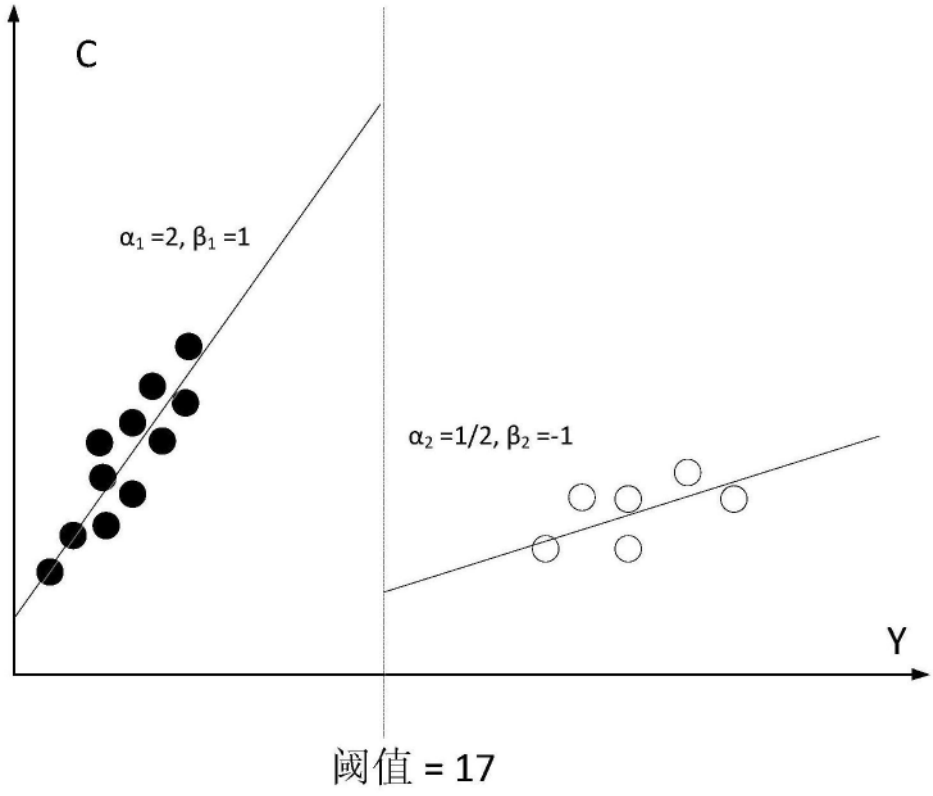


图8

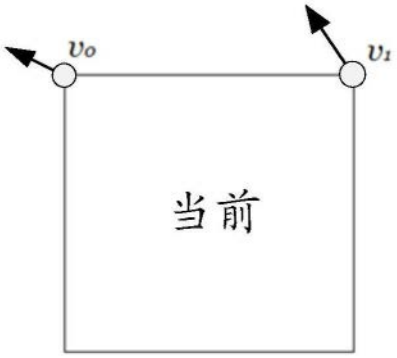


图9A

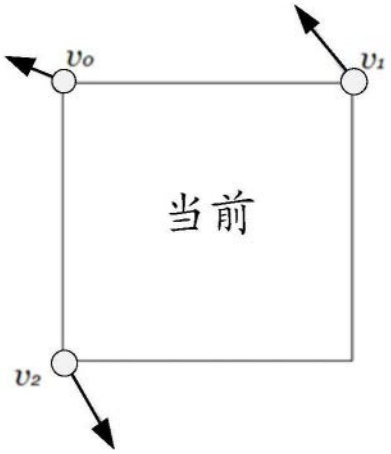


图9B

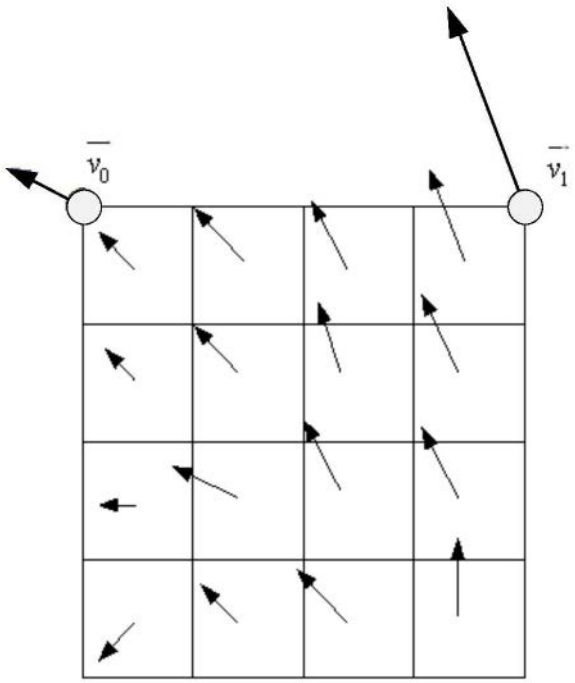


图10

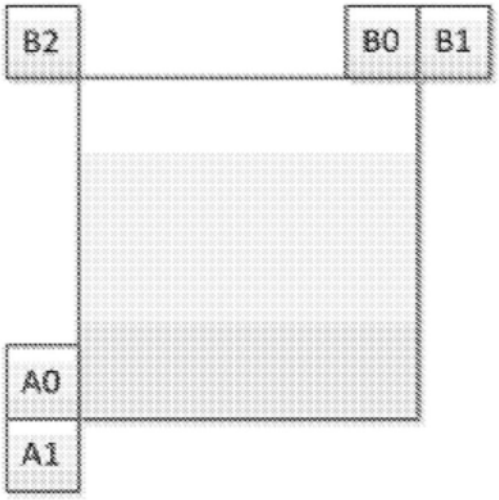


图11

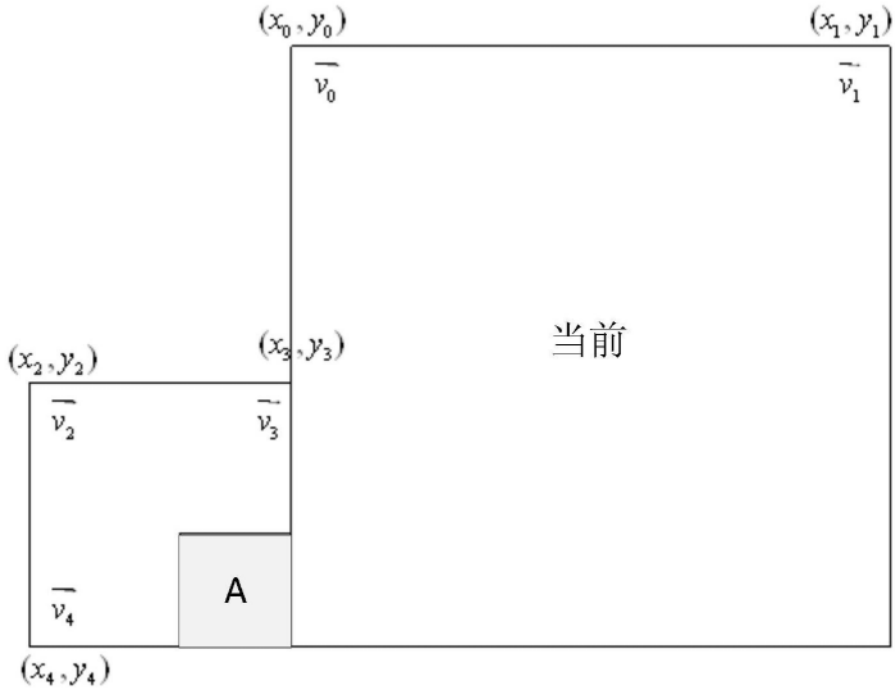

(x_0, y_0)
(x_1, y_1)
$\overrightarrow{v_0}$
$\overrightarrow{v_1}$
当前
(x_2, y_2)
(x_3, y_3)
$\overrightarrow{v_2}$
$\overrightarrow{v_3}$
A
$\overrightarrow{v_4}$
(x_4, y_4)

图12

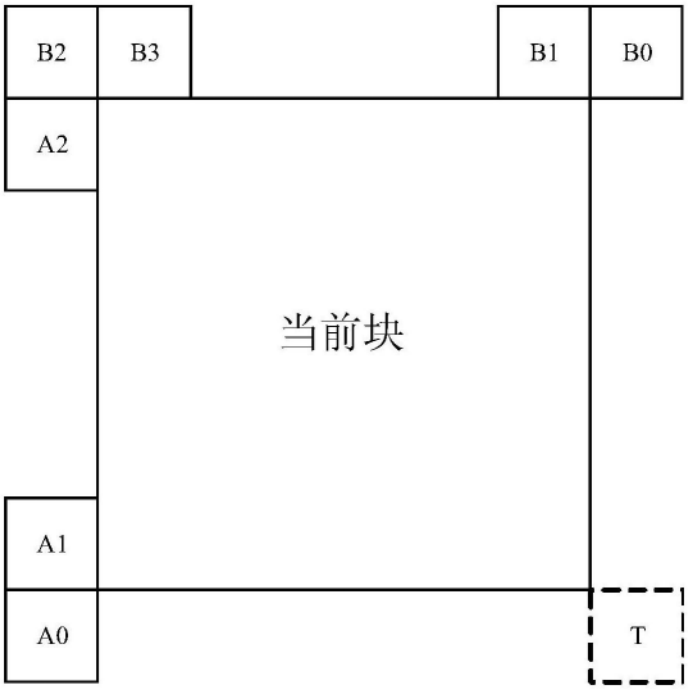

B2
B3
B1
B0
A2
当前块
A1
A0
T

图13

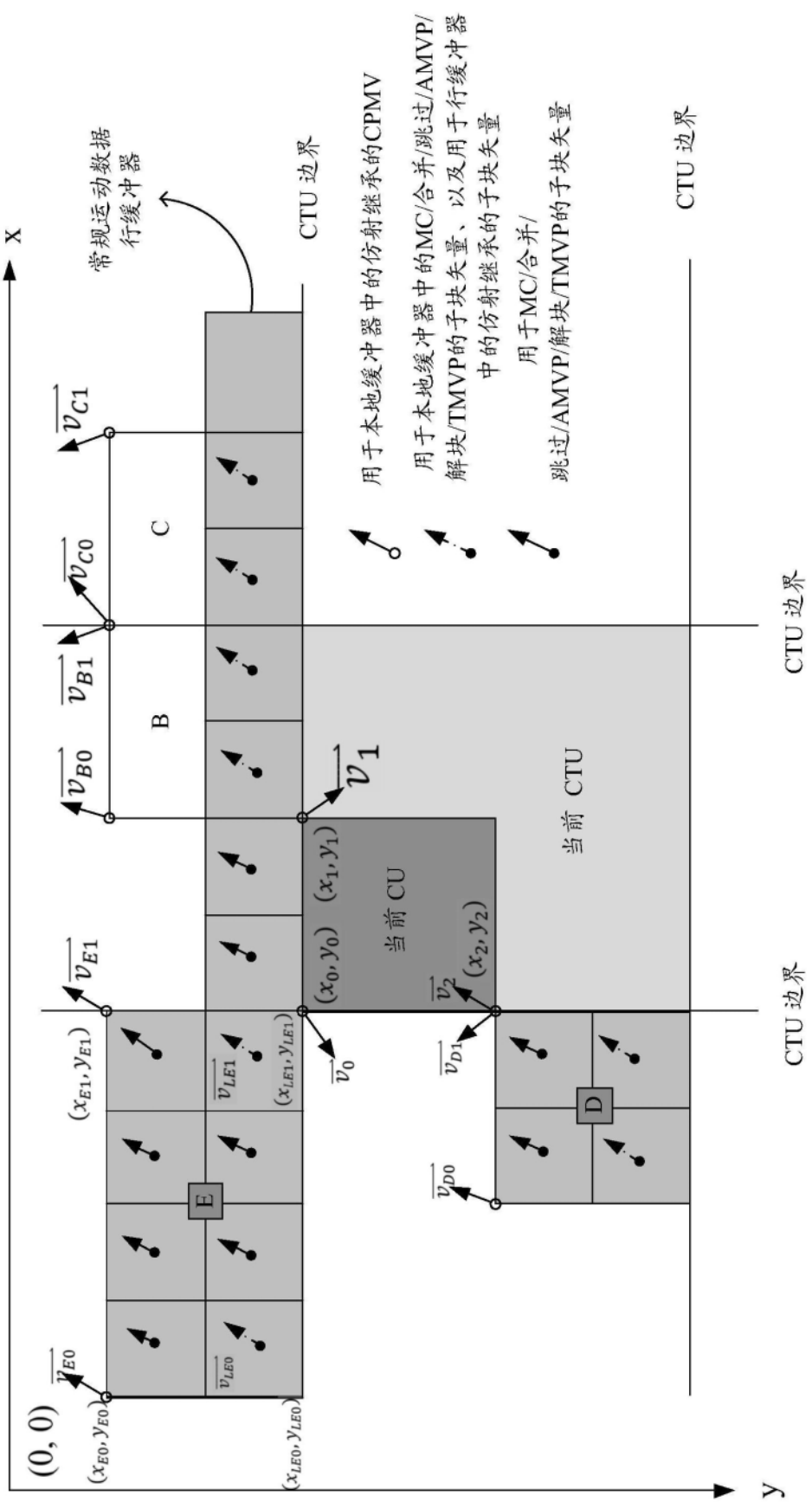

x
y
(0, 0)
常规运动数据
行缓冲器
CTU 边界
用于本地缓冲器中的仿射继承的CPMV
用于本地缓冲器中的MC/合并/跳过/AMVP/
解块/TMVP的子块矢量、以及用于行缓冲器
中的仿射继承的子块矢量
用于MC/合并/
跳过/AMVP/解块/TMVP的子块矢量
当前 CTU
当前 CU
B
C
D
E

图14

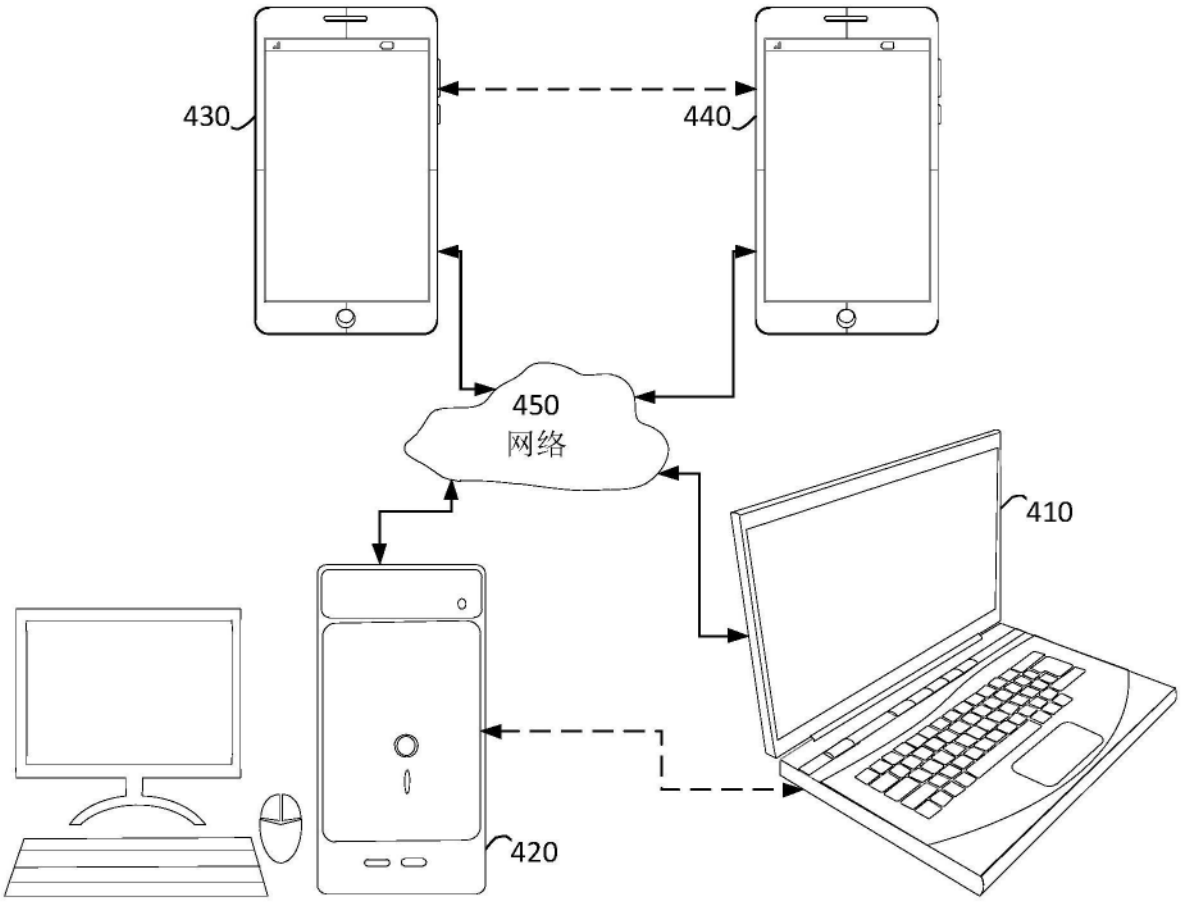


图15

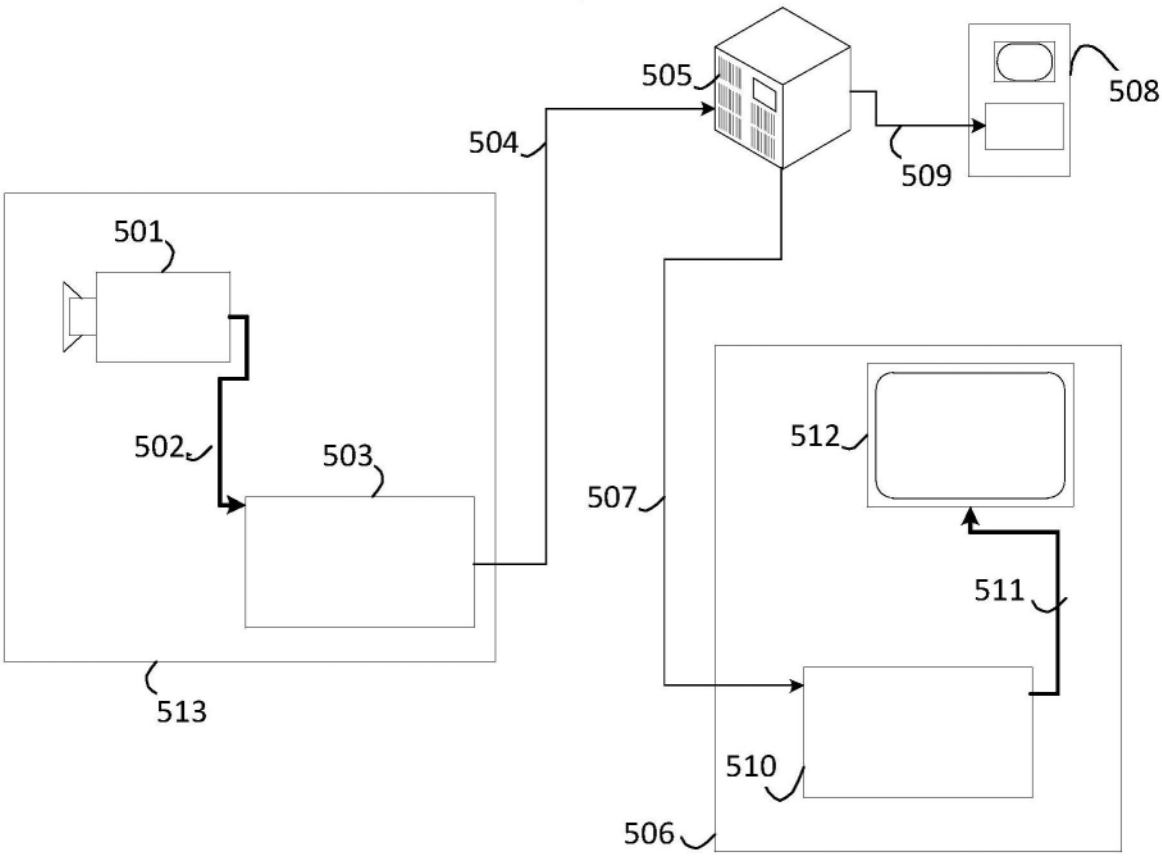


图16

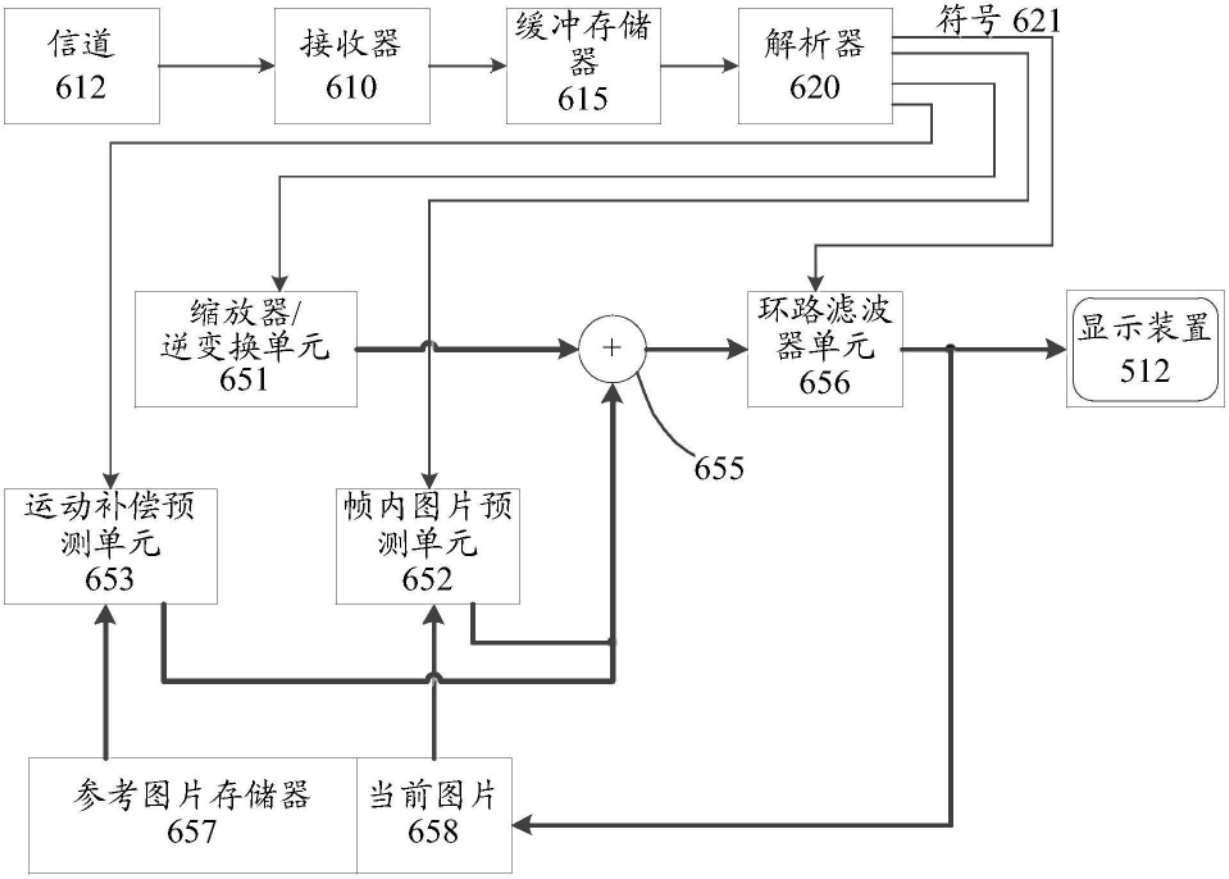

信道
612
接收器
610
缓冲存储器
615
解析器
620
符号 621
缩放器/
逆变换单元
651
+
环路滤波器单元
656
显示装置
512
655
运动补偿预测单元
653
帧内图片预测单元
652
参考图片存储器
657
当前图片
658

图17

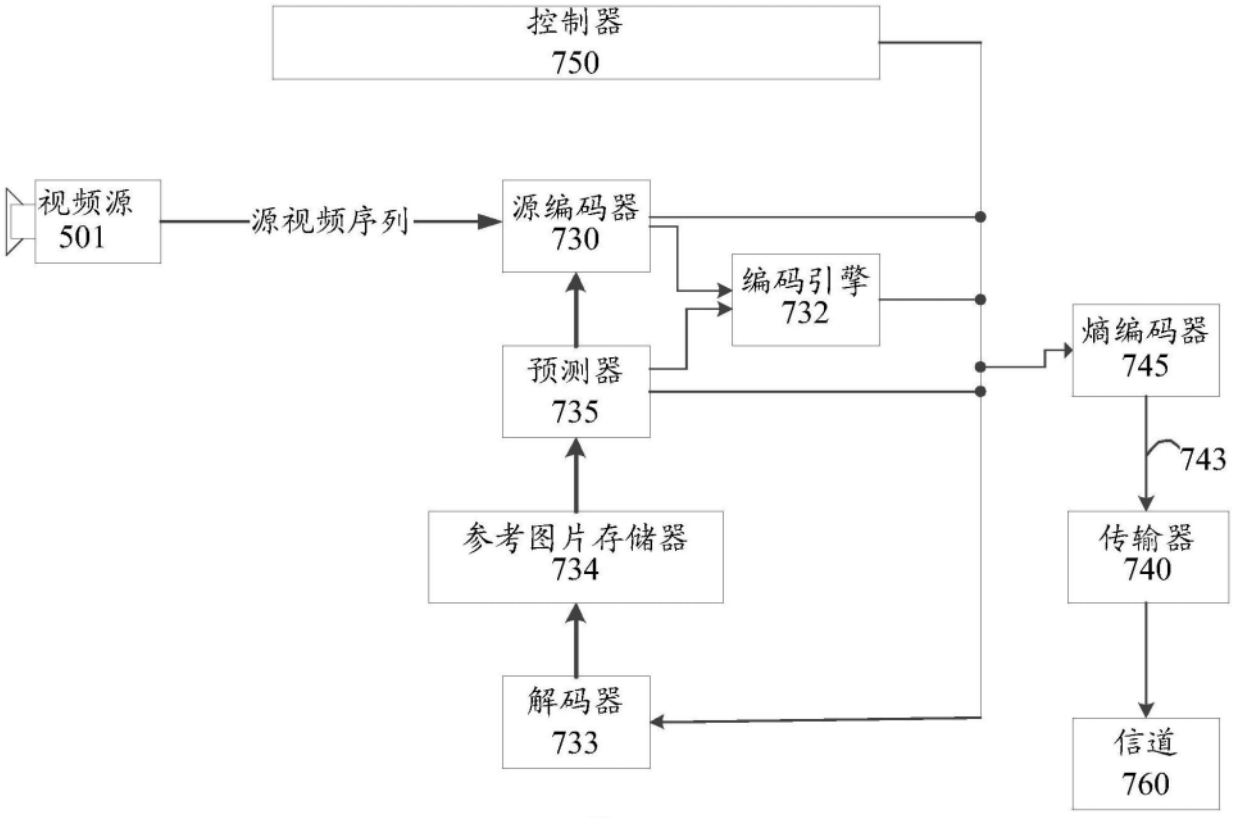

控制器
750
视频源
501
源视频序列
源编码器
730
编码引擎
732
熵编码器
745
预测器
735
743
参考图片存储器
734
传输器
740
解码器
733
信道
760

图18

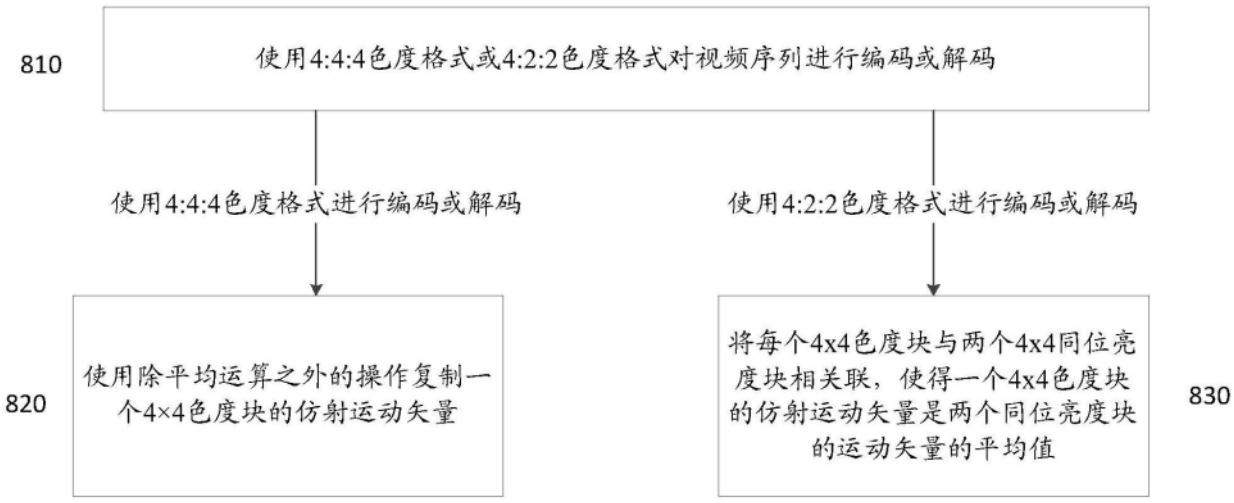

810
使用4:4:4色度格式或4:2:2色度格式对视频序列进行编码或解码
使用4:4:4色度格式进行编码或解码
使用4:2:2色度格式进行编码或解码
820
使用除平均运算之外的操作复制一个4×4色度块的仿射运动矢量
将每个4x4色度块与两个4x4同位亮度块相关联，使得一个4x4色度块的仿射运动矢量是两个同位亮度块的运动矢量的平均值
830

图19

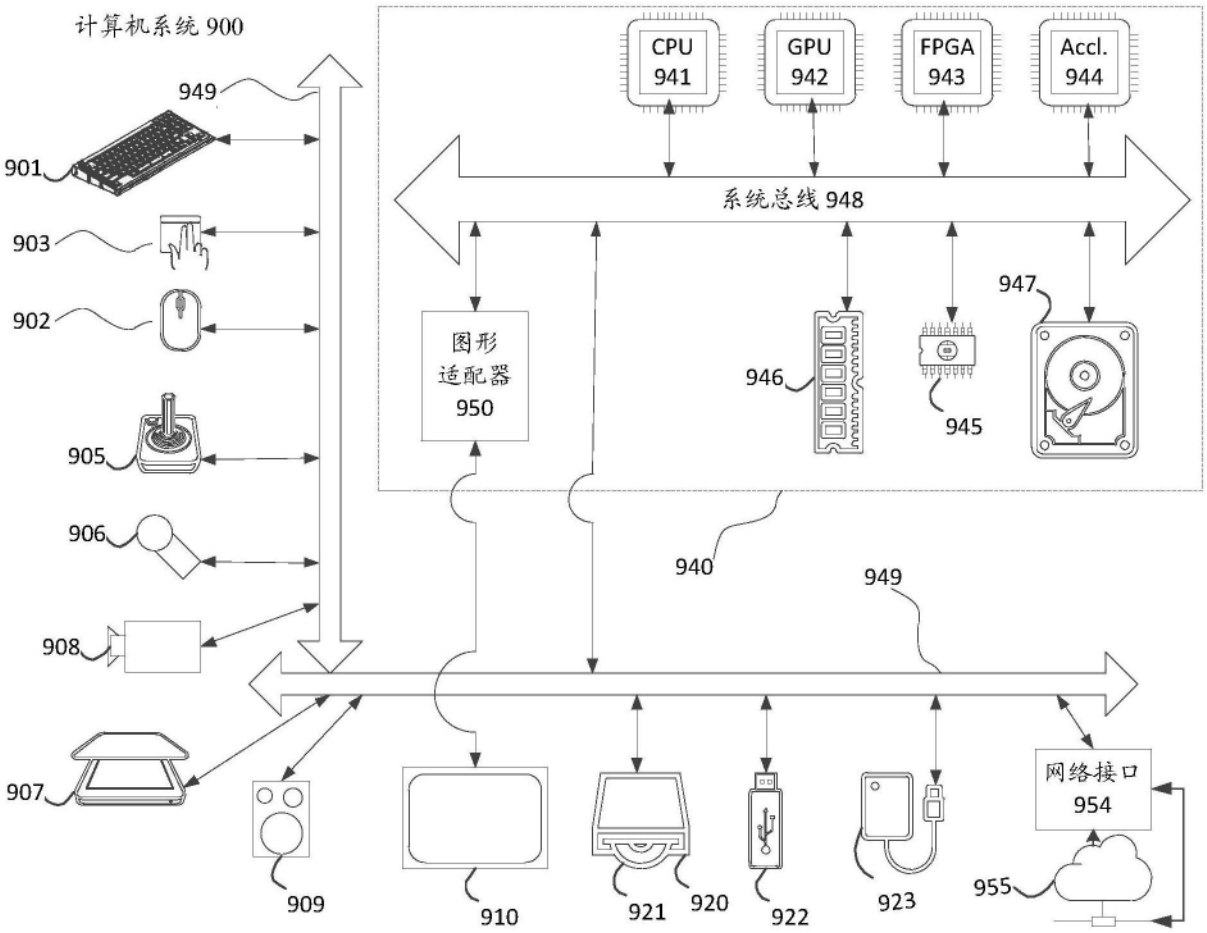

计算机系统 900
949
901
903
902
905
906
908
907
909
910
921
920
922
923
955
CPU 941
GPU 942
FPGA 943
Accl. 944
系统总线 948
图形适配器 950
946
945
947
940
949
网络接口 954

图20