



(12)发明专利

(10)授权公告号 CN 105917648 B

(45)授权公告日 2019.07.12

(21)申请号 201580004941.5

(22)申请日 2015.01.12

(65)同一申请的已公布的文献号  
申请公布号 CN 105917648 A

(43)申请公布日 2016.08.31

(30)优先权数据  
61/928,970 2014.01.17 US  
61/954,572 2014.03.17 US  
14/455,856 2014.08.08 US

(85)PCT国际申请进入国家阶段日  
2016.07.18

(86)PCT国际申请的申请数据  
PCT/US2015/010944 2015.01.12

(87)PCT国际申请的公布数据  
W02015/108793 EN 2015.07.23

(73)专利权人 微软技术许可有限责任公司  
地址 美国华盛顿州

(72)发明人 L·朱 Y·吴 G·J·沙利文

(74)专利代理机构 上海专利商标事务所有限公司 31100

代理人 段登新

(51)Int.Cl.  
H04N 19/105(2006.01)  
H04N 19/176(2006.01)  
H04N 19/70(2006.01)  
H04N 19/119(2006.01)  
H04N 19/61(2006.01)  
H04N 19/593(2006.01)  
H04N 19/11(2006.01)

(56)对比文件  
US 2011/0194613 A1,2011.08.11,说明书第【0052】-【0212】段;附图11,12.  
CN 102077594 A,2011.05.25,全文.  
C-C CHEN ET AL.AHG8: Line-based Intra Block Copy.《Joint CollaborativeTeam on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 15th Meeting: Geneva, CH, 23 Oct. - 1 Nov. 2013》.2013,第1-4页.

审查员 朱一雷

权利要求书2页 说明书32页 附图23页

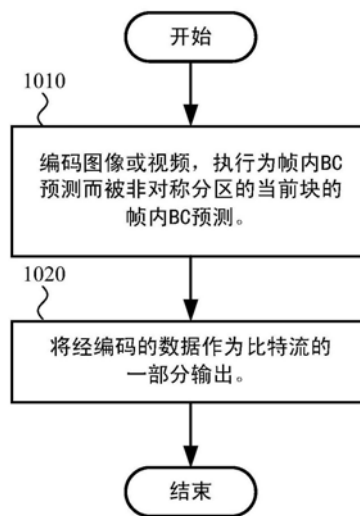
(54)发明名称

具有非对称分区的帧内块复制预测以及编码器侧搜索图案、搜索范围和用于分区的方法

(57)摘要

在帧内块复制(“BC”)预测中的创新以及在编码器侧搜索图案、和用于分区的方法中的创新。例如,一些创新涉及针对帧内BC预测使用非对称分区。其它创新涉及在块向量估计(用于帧内BC预测)或运动估计期间编码器使用的搜索图案或方法。又一些其它创新涉及使用在BV估计期间使用具有水平或垂直偏置的BV搜索范围。

1000



CN 105917648 B

1. 一种在实现图像或视频编码器的计算设备中的方法,该方法包括:

编码图像或视频以产生经编码的数据,包括执行当前块的帧内块复制(“BC”)预测,所述当前块为所述帧内BC预测被非对称分区,其中所述非对称分区包括将所述当前块划分为具有不同尺寸的两个分区;以及

将所述经编码的数据作为比特流的一部分来输出。

2. 如权利要求1所述的方法,其特征在于,所述编码进一步包括针对另一块执行帧内BC预测,所述另一块为所述帧内BC预测被对称分区,其中所述另一块是 $2N \times 2N$ 块,且其中所述另一块被分区为(1)两个 $2N \times N$ 块,(2)两个 $N \times 2N$ 块,或(3)四个 $N \times N$ 块,每个 $N \times N$ 块可被进一步分区为两个 $N \times N/2$ 块、两个 $N/2 \times N$ 块、或四个 $N/2 \times N/2$ 块。

3. 一种在实现图像或视频解码器的计算设备中的方法,该方法包括:

将经编码的数据作为比特流的一部分来接收;以及

解码所述经编码的数据以重构图像或视频,包括执行当前块的帧内块复制(“BC”)预测,所述当前块为所述帧内BC预测被非对称分区,其中所述非对称分区包括将所述当前块划分为具有不同尺寸的两个分区。

4. 如权利要求3所述的方法,其特征在于,所述解码进一步包括针对另一块执行帧内BC预测,所述另一块为所述帧内BC预测被对称分区,其中所述另一块是 $2N \times 2N$ 块,且其中所述另一块被分区为(1)两个 $2N \times N$ 块,(2)两个 $N \times 2N$ 块,或(3)四个 $N \times N$ 块,每个 $N \times N$ 块可被进一步分区为两个 $N \times N/2$ 块、两个 $N/2 \times N$ 块、或四个 $N/2 \times N/2$ 块。

5. 如权利要求1或3所述的方法,其特征在于,所述当前块是 $2N \times 2N$ 块,且其中所述当前块被分区为(1)一 $2N \times N/2$ 块和一 $N \times 3N/2$ 块或(2)一 $2N \times 3N/2$ 块和一 $2N \times N/2$ 块。

6. 如权利要求1或3所述的方法,其特征在于,所述当前块是 $2N \times 2N$ 块,且其中所述当前块被分区为(1)一 $N/2 \times 2N$ 块和一 $3N/2 \times 2N$ 块或(2)一 $3N/2 \times 2N$ 块和一 $N/2 \times 2N$ 块。

7. 如权利要求1或3所述的方法,其特征在于,所述当前块是 $64 \times 64$ 块、 $32 \times 32$ 块、 $16 \times 16$ 块或 $8 \times 8$ 块。

8. 如权利要求1或3所述方法,其特征在于,所述视频是人造视频。

9. 一个或多个存储计算机可执行指令的计算机可读介质,所述计算机可执行指令使得计算系统由此执行一种方法,所述方法包括:

编码图像或视频以产生经编码的数据,包括计算当前图片的当前块的预测,其中计算所述预测使用从下到上方法来标识所述当前块的分区,其中所述当前块被针对帧内BC预测而被非对称分区且其中所述非对称分区包括将所述当前块划分为具有不同尺寸的两个分区;以及

将所述经编码的数据作为比特流的一部分来输出。

10. 如权利要求9所述的一个或多个计算机可读介质,其特征在于,所述预测是运动估计的一部分或是用于帧内块复制预测的块向量估计的一部分。

11. 如权利要求9或10所述的一个或多个计算机可读介质,其特征在于,所述当前块是 $2N \times 2N$ 块,且其中所述从下到上方法包括:

检查所述 $2N \times 2N$ 块的每 $N \times N$ 块的模式;

选择所述 $2N \times 2N$ 块的各个 $N \times N$ 块的最佳模式;

高速缓存所述 $2N \times 2N$ 块的各个 $N \times N$ 块的向量值;

检查所述 $2N \times 2N$ 块的具有 $2N$ 尺寸的模式,包括使用被高速缓存的向量值;  
选择所述 $2N \times 2N$ 块的具有 $2N$ 尺寸的最佳模式;以及  
在所述 $2N \times 2N$ 块的具有 $2N$ 尺寸的所述最佳模式和所选择的所述 $2N \times 2N$ 块的各个 $N \times N$ 块的最佳模式之间做出选择。

12. 如权利要求9或10所述的一个或多个计算机可读介质,其特征在于,所述当前块是 $2N \times 2N$ 块,且其中所述从下到上方法包括:

检查所述 $2N \times 2N$ 块的每 $N \times N$ 块的模式子集;  
高速缓存所述 $2N \times 2N$ 块的各个 $N \times N$ 块的向量值;  
检查所述 $2N \times 2N$ 块的具有 $2N$ 尺寸的模式子集,包括使用被高速缓存的向量值;  
选择所述 $2N \times 2N$ 块的具有 $2N$ 尺寸的最佳模式;以及  
在所述 $2N \times 2N$ 块的具有 $2N$ 尺寸的所述最佳模式和所述 $2N \times 2N$ 块的各个 $N \times N$ 块的最佳模式之间做出选择。

13. 如权利要求12所述的一个或多个计算机可读介质,其特征在于,每 $N \times N$ 块的模式子集是具有两个 $N \times N/2$ 块的模式,并且其中具有 $2N$ 尺寸的模式子集是具有 $2N \times 2N$ 块的模式、具有两个 $2N \times N$ 块的模式、具有一 $2N \times N/2$ 块和一 $2N \times 3N/2$ 块的模式、以及具有一 $2N \times 3N/2$ 块和一 $2N \times N/2$ 块的模式。

14. 如权利要求12所述的一个或多个计算机可读介质,其特征在于,每 $N \times N$ 块的模式子集是具有两个 $N/2 \times N$ 块的模式,并且其中具有 $2N$ 尺寸的模式子集是具有 $2N \times 2N$ 块的模式、具有两个 $N \times 2N$ 块的模式、具有一 $N/2 \times 2N$ 块和一 $3N/2 \times 2N$ 块的模式、以及具有一 $3N/2 \times 2N$ 块和一 $N/2 \times 2N$ 块的模式。

## 具有非对称分区的帧内块复制预测以及编码器侧搜索图案、 搜索范围和用于分区的方法

[0001] 背景

[0002] 工程师使用压缩(也叫做源编码(source coding或source encoding))来降低数字视频的比特率。压缩通过将视频信息转换成较低比特率的形式来降低存储和传送该信息的成本。解压(也被称为解码)从压缩的形式中重构一种版本的原始信息。“编解码器”是编码器/解码器系统。

[0003] 在过去的二十年中,已采用了各种视频编解码器标准,包括ITU-T H.261、H.262(MPEG-2或ISO/IEC 13818-2)、H.263和H.264(MPEG-4AVC或ISO/IEC14496-10)标准、MPEG-1(ISO/IEC 11172 1 172-2)和MPEG-4视觉(ISO/IEC14496-2)标准以及SMPTE 421M(VC-1)标准。最近,H.265/HEVC标准(ITU-T H.265或ISO/IEC 23008-2)已被批准。目前,(例如,用于可缩放视频编码/解码、用于在样本比特深度或色度采样率方面具有较高保真度的视频的编码/解码、用于屏幕捕捉内容、或用于多视图编码/解码的)H.265/HEVC标准的扩展处于开发中。视频编解码器标准通常定义针对经编码的视频比特流的句法的选项,从而详述当在编码和解码时使用特定特征时该比特流中的参数。在许多情况下,视频编解码器标准还提供关于解码器应当执行以在解码时取得一致的结果的解码操作的细节。除了编解码器标准外,各种专用编解码器格式定义针对经编码的视频比特流的句法的其他选项以及相应的解码操作。

[0004] 帧内块复制(“BC”)是H.265/HEVC扩展的处于开发中的预测模式。对于帧内BC预测模式,图片的当前块的样本值是使用同一图片中的先前重构的样本值来预测的。块向量(“BV”)指示从当前块到图片的包括用于预测的先前重构的样本值的区域的位移。BV被信号化在比特流中。帧内BC预测是一种形式的图片内预测——针对图片的块的帧内BC预测不使用除了同一图片中的样本值以外的任何样本值。

[0005] 如当前在HEVC标准中规定并在针对HEVC标准的某一参考软件中实现的,帧内BC预测模式具有若干问题。特别是,用于帧内BC预测的块大小的选项在许多场景下过于受限,并且在许多场景下不能使关于块大小以及如何使用帧内BC预测的编码器侧决策高效。

[0006] 概述

[0007] 概言之,本详细描述呈现了在帧内块复制(“BC”)预测中的创新以及在编码器侧搜索图案、搜索范围和用于分区的方法中的创新。”例如,一些创新涉及针对帧内BC预测使用非对称分区(有时被称为“AMP)。”其它创新涉及在块向量(“BV”)估计(用于帧内BC预测)或运动估计期间编码器使用的搜索图案或方法。又一些其它创新涉及使用在BV估计期间使用具有水平或垂直偏置的BV搜索范围。

[0008] 根据本文中描述的创新中的第一方面,图像编码器或视频编码器对图像或视频进行编码以产生经编码的数据并将经编码数据作为比特率的一部分输出。作为编码的一部分,编码器执行为了帧内BC预测而被非对称分区的当前块的帧内BC预测。例如,所述当前块是 $2N \times 2N$ 块,且所述当前块被分区为(1)  $-2N \times N/2$ 块和 $-2N \times 3N/2$ 块或(2)  $-2N \times 3N/2$ 块和 $-2N \times N/2$ 块。或者,作为另一示例,所述当前块是 $2N \times 2N$ 块,且其中所述当前块被

分区为(1)  $-N/2 \times 2N$ 块和 $-3N/2 \times 2N$ 块或(2)  $-3N/2 \times 2N$ 块和 $-N/2 \times 2N$ 块。更一般而言,对于非对称分区,当前块可被划分为具有不同尺寸的两个分区。作为编码的一部分,编码器还可执行为了帧内BC预测而被对称分区的另一块的帧内BC预测。例如,另一块是被分区为(1)两个 $2N \times N$ 块,(2)两个 $N \times 2N$ 块,或(3)四个 $N \times N$ 块的 $2N \times 2N$ 块,每个块可被进一步分区为两个 $N \times N/2$ 块、两个 $N/2 \times N$ 块、或四个 $N/2 \times N/2$ 块。更一般而言,对于对称分区,该另一块可被划分为具有等同尺寸的分区。

[0009] 根据本文中描述的创新中的第二方面,图像解码器或视频解码器接收作为比特率的一部分的经编码数据并将经编码数据进行解码以重构图像或视频。作为解码的一部分,解码器执行为了帧内BC预测而被非对称分区的当前块的帧内BC预测。例如,所述当前块是 $2N \times 2N$ 块,且所述当前块被分区为(1)  $-2N \times N/2$ 块和 $-2N \times 3N/2$ 块或(2)  $-2N \times 3N/2$ 块和 $-2N \times N/2$ 块。或者,作为另一示例,所述当前块是 $2N \times 2N$ 块,且其中所述当前块被分区为(1)  $-N/2 \times 2N$ 块和 $-3N/2 \times 2N$ 块或(2)  $-3N/2 \times 2N$ 块和 $-N/2 \times 2N$ 块。更一般而言,对于非对称分区,当前块可被划分为具有不同尺寸的两个分区。作为解码的一部分,解码器还可执行为了帧内BC预测而被对称分区的另一块的帧内BC预测。例如,另一块是被分区为(1)两个 $2N \times N$ 块,(2)两个 $N \times 2N$ 块,或(3)四个 $N \times N$ 块的 $2N \times 2N$ 块,每个块可被进一步分区为两个 $N \times N/2$ 块、两个 $N/2 \times N$ 块、或四个 $N/2 \times N/2$ 块。更一般而言,对于对称分区,该另一块可被划分为具有等同尺寸的分区。

[0010] 根据本文中描述的创新中的第三方面,图像编码器或视频编码器对图像或视频进行编码以产生经编码的数据并将经编码数据作为比特率的一部分输出。作为编码的一部分,编码器计算当前图片的当前块的预测(例如,预测单元的预测块)。预测可针对用于帧内BC预测的运动估计或BV估计。在任何情况下,计算预测使用从下到上的方法来标识当前块的分区。一般而言,当前块的分区包括具有不同尺寸的两个或更多个分区。例如,当前块是 $2N \times 2N$ 块,且从下到上方法包括:(a)检查该 $2N \times 2N$ 块的每 $N \times N$ 块的模式;(b)选择各个 $N \times N$ 块的最佳模式;(c)高速缓存各个 $N \times N$ 块的向量值;(d)检查该 $2N \times 2N$ 块的具有 $2N$ 尺寸的模式,包括使用被高速缓存的向量值;(e)选择该 $2N \times 2N$ 块的具有 $2N$ 尺寸的最佳模式;以及(f)在该 $2N \times 2N$ 块的具有 $2N$ 尺寸的最佳模式和所选择的该 $2N \times 2N$ 块的各个 $N \times N$ 块的最佳模式之间做出选择。或者,作为另一示例,当前块是 $2N \times 2N$ 块,且从下到上方法包括:(a)检查该 $2N \times 2N$ 块的每 $N \times N$ 块的模式子集;(b)高速缓存各个 $N \times N$ 块的向量值;(c)检查该 $2N \times 2N$ 块的具有 $2N$ 尺寸的模式子集,包括使用被高速缓存的向量值;(d)选择该 $2N \times 2N$ 块的具有 $2N$ 尺寸的最佳模式;以及(e)在该 $2N \times 2N$ 块的具有 $2N$ 尺寸的最佳模式和各个 $N \times N$ 块的最佳模式之间做出选择。

[0011] 根据本文中描述的创新中的第四方面,图像编码器或视频编码器对图像或视频进行编码以产生经编码的数据并将经编码数据作为比特率的一部分输出。作为编码的一部分,编码器计算当前图片的当前块的预测。预测可针对用于帧内BC预测的运动估计或BV估计。在任何情况下,计算预测包括(a)通过当前最佳位置周围的小邻域(例如,水平或垂直紧邻当前最佳位置的位置)中的迭代评估来标识该预测的当前最佳位置,以及(b)通过当前最佳位置周围的后继更大大邻域(例如,小邻域外的环中的位置)中的迭代评估来确认该当前最佳位置。例如,如果当前最佳位置比更大邻域之一中的位置差,则编码器取代当前最佳位置并重复所述标识和所述确认。如果达到后继更大大邻域中的评估的阈值迭代次数,所述确

认阶段可停止。

[0012] 根据本文描述的创新第五方面,图像编码器或视频编码器确定图片的当前块的BV、使用该BV执行当前块的帧内BC预测,并编码该BV。该BV指示到图片内的一区域的位移。在确定该BV时,编码器检查该区域位于具有水平偏置或垂直偏置的BV搜索范围内的约束。编码器可从多个可用BV搜索范围中选择BV搜索范围,例如至少部分取决于一个或多个先前块的BV值,其可在直方图数据结构中被跟踪。

[0013] 根据本文中描述的创新中的第六方面,图像编码器或视频编码器使用帧内BC预测来编码图片的数据,并将经编码数据作为比特率的一部分输出。作为编码的一部分,编码器使用具有水平或垂直偏置的BV搜索范围来执行BV估计操作。编码器可从多个可用BV搜索范围中选择BV搜索范围,例如至少部分取决于一个或多个先前块的BV值,其可在直方图数据结构中被跟踪。

[0014] 这些创新可以被实现为方法的一部分、被适配成执行该方法的计算系统的一部分或存储用于使得计算系统执行该方法的计算机可执行指令的有形计算机可读介质的一部分。各创新可以结合地或分开地使用。

[0015] 参考附图阅读以下详细描述,将更清楚本发明的前述和其他目标、特征和优点。

[0016] 附图简述

[0017] 图1是其中可实现所描述的一些实施例的示例计算系统的示意图。

[0018] 图2a和2b是其中可实现所描述的一些实施例的示例网络环境的示意图。

[0019] 图3是结合其可实现所描述的一些实施例的示例编码器系统的示意图。

[0020] 图4是结合其可实现所描述的一些实施例的示例解码器系统的示意图。

[0021] 图5a和5b是示出结合其可实现所描述的一些实施例的示例视频编码器的示意图。

[0022] 图6是示出结合其可实现所描述的一些实施例的示例视频解码器的示意图。

[0023] 图7和8是示出对图片的块和该块在块匹配中的候选块的帧内BC预测的示意图。

[0024] 图9是示出用于帧内BC预测的示例块大小的图示,其包括一些非对称分区和一些对称分区。

[0025] 图10和11分别是包括具有非对称分区的帧内BC预测的用于编码和解码的一般化技术。

[0026] 图12是示出针对图片的块的运动估计的示意图。

[0027] 图13是示出用于对经帧内BC预测的块进行分区的从上到下的方法的流程图和附图。

[0028] 图14是示出用于分区的从下到上的方法的一般化技术的流程图。

[0029] 图15是示出用于对经帧内BC预测的块进行分区的从下到上的方法的流程图和附图。

[0030] 图16是示出用于对经帧内BC预测的块进行分区的甚至更快的从下到上的方法的流程图和附图。

[0031] 图17是示出用于使用小邻域中的位置的迭代估计和更大邻域中的该位置的迭代确认来搜索块的BV值或MV值的一般化技术的流程图。

[0032] 图18a和18b是示出在搜索块的BV值或MV值时的小邻域中的位置的迭代估计的图示。

[0033] 图19是示出在搜索块的BV值或MV值时在一个或多个更大邻域中的位置的迭代确认的图示。

[0034] 图20是示出用于使用小邻域中的位置的迭代估计和更大邻域中的该位置的迭代确认来搜索块的BV值或MV值的示例技术的流程图。

[0035] 图21a-21e是示出针对BV值的搜索范围的示例约束的示图。

[0036] 图22是示出用于采用帧内BC预测模式来进行编码的通用技术的流程图,其中该帧内BC预测模式受制于针对BV值的选择的一个或多个约束。

[0037] 图23是示出对图片的块的示例z-扫描次序的示图。

[0038] 详细描述

[0039] 本详细描述呈现了在帧内块复制(“BC”)预测中的创新以及在编码器侧搜索图案、搜索范围和用于分区的方法中的创新。例如,一些创新涉及在编码和/或解码期间针对帧内BC预测使用非对称分区(有时被称为“AMP”)。其它创新涉及在块向量(“BV”)估计(用于帧内BC预测)或运动估计期间编码器使用的搜索图案或方法。又一些其它创新涉及使用在BV估计期间使用具有水平或垂直偏置的BV搜索范围。

[0040] 虽然在此所述的操作是被适当描述为由视频编码器或视频解码器来执行,但在许多情况中,这些操作可以由另一类型的媒体处理工具(例如图像编码器或图像解码器)来执行。

[0041] 本文中描述的一些创新是参考专用于H.265/HEVC标准的句法元素和操作来示出的。例如,对H.265/HEVC标准的草稿版本JCTVC-P1005—即2014年2月的“高效视频编码(HEVC)范围扩展文本规范:第6稿”JCTVC-P1005\_v1作出参考。本文中描述的各创新还可以被实现为用于其它标准或格式。

[0042] 本文中描述的许多创新可以在当编码诸如屏幕捕捉内容之类的某种“人工创建的”视频内容时改善速率-失真性能。一般来说,与仅相机捕捉的视频内容相反,屏幕捕捉视频(也称为屏幕内容视频)是包含在向计算机显示器渲染时捕捉的经渲染的文本、计算机图形、动画生成的内容或其它类似类型的内容的视频。屏幕捕捉内容通常包括重复的结构(例如图形、文本字符)。屏幕捕捉内容一般以具有高色度采样分辨率的格式(例如,YUV 4:4:4或RGB 4:4:4)来编码,但是也可用具有较低色度采样分辨率的格式(例如,YUV 4:2:0)来编码。用于编码/解码屏幕捕捉内容的常见场景包括远程桌面会议和对自然视频或其它“混合内容”视频上的图形覆盖的编码/解码。在此所述的几种发明适用于编码屏幕内容视频或其它人工创建的视频。这些创新也可以被用于自然视频,但可能不像那样有效。本文描述的其它创新在自然视频或人类创建的视频的编码中是有效的。

[0043] 更一般地,本文中描述的各示例的各种替代是可能的。例如,本文中描述的一些方法可以通过改变描述的方法动作的顺序、通过拆分、重复或忽略某些方法动作等来更改。所公开的技术的各方面能够被组合地或分开地使用。不同的实施例使用所描述的创新中的一个或多个。本文中描述的一些创新解决了背景中指出的一个或多个问题。通常,所给出的技术/工具并不解决所有这些问题。

[0044] I. 示例计算系统

[0045] 图1示出了其中可实现所描述的若干创新的合适计算系统(100)的概括示例。计算系统(100)并不旨对使用范围或功能提出任何限制,因为这些创新可以在不同的通用或专

用计算系统中实现。

[0046] 参考图1, 计算系统(100)包括一个或多个处理单元(110、115)和存储器(120、125)。处理单元(110、115)执行计算机可执行指令。处理单元可以是通用中央处理单元(“CPU”)、专用集成电路(“ASIC”)中的处理器或任何其它类型的处理器。在多处理系统中, 多个处理单元执行计算机可执行指令以提高处理能力。例如, 图1示出中央处理单元(110)以及图形处理单元或协处理单元(115)。有形存储器(120、125)可以是易失性存储器(例如, 寄存器、高速缓存、RAM)、非易失性存储器(例如, ROM、EEPROM、闪存等)或两者的某种组合, 可通过(一个或多个)处理单元存取。存储器(120、125)存储实现用于具有非对称分区的帧内预测的一个或多个创新和/或用于编码器侧搜索图案、具有水平或竖直偏置的搜索范围和/或用于分区的方法的一个或多个创新的软件(180), 其形式为适于由(一个或多个)处理单元执行的计算机可执行指令。

[0047] 计算系统可具有附加的特征。例如, 计算系统(100)包括存储(140)、一个或多个输入设备(150)、一个或多个输出设备(160)以及一个或多个通信连接(170)。诸如总线、控制器或网络之类的互连机制(未示出)将计算系统(100)的各组件互连。通常, 操作系统软件(未示出)为在计算系统(100)中执行的其它软件提供操作环境, 并协调计算系统(100)的各组件的活动。

[0048] 有形存储(140)可以是可移动或不可移动的, 并包括磁盘、磁带或磁带盒、CD-ROM、DVD或可用于储存信息并可在计算系统(100)内访问的任何其他介质。存储(140)存储用于实现具有非对称分区的帧内BC预测的一个或多个创新和/或用于编码器侧搜索图案、搜索范围和/或用于分区的方法的一个或多个创新的软件(180)的指令。

[0049] (诸)输入设备(150)可以是触摸输入设备(诸如键盘、鼠标、笔或跟踪球)、语音输入设备、扫描设备或向计算系统(100)提供输入的另一设备。对于视频, (一个或多个)输入设备(150)可以是相机、视频卡、TV调谐卡、屏幕捕捉模块或接受模拟或数字形式的视频输入的类似设备、或将视频输入读到计算系统(100)中的CD-ROM或CD-RW。(一个或多个)输出设备(160)可以是显示器、打印机、扬声器、CD刻录机或提供来自计算系统(100)的输出的另一设备。

[0050] (一个或多个)通信连接(170)允许通过通信介质与另一计算实体通信。通信介质传达诸如计算机可执行指令、音频或视频输入或输出、或已调制数据信号中的其他数据之类的信息。已调制数据信号是使其一个或多个特征以在信号中编码信息的方式设置或改变的信号。作为示例而非限制, 通信介质可以使用电的、光学的、RF或其它载体。

[0051] 各创新可以在计算机可读介质的一般上下文中描述。计算机可读介质是可在计算环境内访问的任何可用有形介质。作为示例而非局限, 对于计算系统(100), 计算机可读介质包括存储器(120、125)、存储(140)、和以上任意的组合。

[0052] 各创新可在计算机可执行指令(诸如包括在程序模块中的在目标现实或虚拟处理器上在计算系统中执行的那些计算机可执行指令)的一般上下文中描述。一般而言, 程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、库、对象、类、组件、数据结构等。如各实施例中所描述的, 这些程序模块的功能可以被组合, 或者在这些程序模块之间拆分。针对各程序模块的计算机可执行指令可以在本地或分布式计算系统中执行。

[0053] 术语“系统”和“设备”在此被互换地使用。除非上下文明确指示, 否则, 术语并不暗



示对计算系统或计算设备的类型的任何限制。一般说来,计算系统或计算设备可以是本地的或分布式的,并且可以包括具有实现本文中描述的功能的软件的专用硬件和/或通用硬件的任意组合。

[0054] 所公开的方法还可使用被配置成执行所公开的方法中的任一者的专用计算硬件来实现。例如,所公开的方法可以由被专门设计或配置成实现所公开的方法中的任一者的集成电路(例如,诸如ASIC数字信号处理器(“DSP”)之类的ASIC、图像处理单元(“GPU”)、或诸如场可编程门阵列(“FPGA”)之类的可编程逻辑器件(“PLD”))来实现。

[0055] 为了呈现起见,本详细描述使用了如“确定”和“使用”等术语来描述计算系统中的计算机操作。这些术语是对由计算机执行的操作的高级抽象,且不应与人类所执行的动作混淆。对应于这些术语的实际的计算机操作取决于实现而不同。如本文用于描述编码选项时所使用的,术语“最佳”(如在“最佳位置”、用于分区的“最佳模式”或“最佳组合”)指示:相较于其它编码选项,关于估计的编码效率或实际编码效率,在失真成本、比特率成本或失真成本和比特率成本的某种组合方面优选的编码选项。任何可用失真度量可被用于失真成本。任何可用比特率度量可被用于比特率成本。其它因素(诸如算法编码复杂度、算法解码复杂度、资源使用和/或延迟)也可影响关于哪些编码选项为“最佳”的决策。

[0056] II. 示例网络环境

[0057] 图2a和2b示出了包括视频编码器(220)和视频解码器(270)的示例网络环境(201, 202)。编码器(220)和解码器(270)使用合适的通信协议通过网络(250)连接。网络(250)可包括因特网或另一计算机网络。

[0058] 在图2a所示的网络环境(201)中,每个实时通信(“RTC”)工具(210)都包括用于双向通信的编码器(220)和解码器(270)两者。给定的编码器(220)可以产生输出,该输出遵循H.265/HEVC标准、SMPTE 421M标准、ISO/IEC14496-10标准(也称为H.264或AVC)、另一标准的变体或扩展,或遵循针对接受来自编码器(220)的经编码的数据的对应的解码器(270)的专用格式。双向通信可以是视频会议、视频电话呼叫或其它双方或多方通信场景的部分。虽然,图2a中的网络环境(201)包括两个实时通信工具(210),但网络环境(201)可改为包括参与多方通信的三个或更多个实时通信工具(210)。

[0059] 实时通信工具(210)管理编码器(220)做出的编码。图3示出可以被包括在实时通信工具(210)中的示例编码器系统(300)。替换地,实时通信工具(210)使用另一编码器系统。实时通信工具(210)还管理解码器(270)做出的解码。图4示出可以被包括在实时通信工具(210)中的示例解码器系统(400)。替换地,实时通信工具(210)使用另一解码器系统。

[0060] 在图2b中示出的网络环境(202)中,编码工具(212)包括编码供递送给多个回放工具(214)的视频的编码器(220),此多个回放工具(214)包括解码器(270)。单向通信可被提供用于视频监视系统、web相机监视系统、屏幕捕捉模块、远程桌面会议演示或在其中编码视频并将其从一个位置发送到一个或多个其它位置的其它场景。虽然在图2b中的网络环境(202)包括两个回放工具(214),但该网络环境(202)可以包括更多或更少的回放工具(214)。一般来说,回放工具(214)与编码工具(212)通信以确定回放工具(214)要接收的视频流。回放工具(214)接收该流、缓冲所接收的经编码数据达合适的时间段并开始解码和回放。

[0061] 图3示出可以被包括在编码工具(212)中的示例编码器系统(300)。替换地,编码工

具(212)使用另一编码器系统。编码工具(212)还可以包括用于管理与一个或多个回放工具(214)的连接的服务侧控制器逻辑。图4示出可以被包括在回放工具(214)中的示例解码器系统(400)。替换地,回放工具(214)使用另一解码器系统。回放工具(214)还可以包括用于管理与编码工具(212)的连接的客户侧控制器逻辑。

[0062] III. 示例编码器系统。

[0063] 图3是结合其可实现所描述的一些实施例的示例编码器系统(300)的框图。编码器系统(300)可以是能够用多种编码模式中的任一者(诸如用于实时通信的低等待时间编码模式、转码模式和用于从文件或流中产生供回放的媒体的较高等待时间编码模式)操作的通用编码工具,或它可以是适用于一种这样的编码模式的专用编码工具。编码器系统(300)可以适用于编码特定类型的内容(例如屏幕捕捉内容)。编码器系统(300)可以被实现为操作系统模块、应用库的部分、独立的应用。总体上,编码器系统(300)从视频源(310)接收源视频帧序列(311)并产生经编码数据作为到信道(390)的输出。输出到信道的经编码的数据可包括使用帧内BC预测模式来编码的内容。

[0064] 视频源(310)可以是相机、调谐卡、存储介质、屏幕捕捉模块或其它数字视频源。视频源(310)以例如每秒30帧的帧速率产生视频帧序列。如本文中所使用的,术语“帧”一般指代源、经编码的或经重构的图像数据。对于逐行扫描视频,帧是逐行扫描视频帧。对于隔行视频,在各示例实施例中,隔行视频帧可以在编码之前被去隔行。替换地,两个互补的隔行视频场可以被编码在一起作为单个视频帧或者被编码成两个经分开编码的场。除了指示逐行扫描视频帧或隔行扫描视频帧之外,术语“帧”或“图片”可以指示单个非成对的视频场、互补的成对视频场、表示在给定时间的视频对象的视频对象平面、或较大图像中的感兴趣区域。视频对象平面或区域可以是包括场景的多个对象或区域的较大图像的一部分。

[0065] 抵达的源帧(311)被存储在包括多个帧缓冲器存储区域(321、322、...、32n)的源帧临时存储器存储区域(320)中。帧缓冲区(321、322等)在源帧存储区域(320)中保持一个源帧。在一个或多个源帧(311)已被存储在帧缓冲区(321、322等)中后,帧选择器(330)从源帧存储区域(320)中选择一个体源帧。帧选择器(330)选择帧以供输入到编码器(340)的次序可不同于视频源(310)产生这些帧的次序,例如某些帧的编码可以被按序延迟,从而允许一些更后面的帧被先编码并由此促成时间上向后的预测。在编码器(340)之前,编码器系统(300)可包括预处理器(未示出),该预处理器在编码之前执行对选中的帧(331)的预处理(例如滤波)。预处理可包括将色彩空间转换成主要(例如亮度)和次要(例如偏向红色和偏向蓝色的色度差)分量以及对编码的重采样处理(例如以减少色度分量的空间分辨率)。通常,在编码之前,视频已被转换成诸如YUV的色彩空间,其中亮度(Y)分量的样本值表示明亮度或强度值,而色度(U,V)分量的样本值表示色差值。色差值(以及从YUV色彩空间到诸如RGB的另一色彩空间和/或从另一色彩空间到YUV色彩空间的转换操作)的精确定义取决于实现。通常,如在此所用,术语YUV指示具有亮度(或照度)分量和一个或多个色度(或色差)分量的任意色彩空间,包括Y'UV、YIQ、Y'IQ和YDbDr以及诸如YCbCr和YCoCg之类的变体。色度采样值可以被子采样到较低的色度采样率(例如对于YUV 4:2:0格式),或者色度采样值可以具有与亮度采样值相同的分辨率(例如对于YUV4:4:4格式)。或者,视频可用另一格式(例如RGB 4:4:4格式、GBR 4:4:4格式或BGR 4:4:4格式)来编码。

[0066] 编码器(340)编码选中的帧(331)以产生经编码的帧(341)并且还产生存储器管理

控制操作 (“MMCO”) 信号 (342) 或参考图片集 (“RPS”) 信息。RPS是在针对当前帧或任意后续帧的运动补偿中可用于供参考的帧集。如果当前帧不是已被编码的第一帧,则在执行其编码处理时,编码器 (340) 可以使用已经被存储在经解码帧的临时存储器存储区域 (360) 中的一个或多个先前被编码/解码的帧 (369)。这样的存储的经解码帧 (369) 被作用于当前源帧 (331) 的内容的帧间预测的参考帧。MMCO/RPS信息 (342) 向解码器指示哪些经重构的帧可被用作参考帧,并且因此应该被存储在帧存储区域中。

[0067] 一般来说,编码器 (340) 包括执行编码任务的多个编码模块,编码任务为诸如分割成小块、帧内预测估计和预测、运动估计和补偿、频率变换、量化和熵编码。由编码器 (340) 执行的确切操作可以取决于压缩格式而变化。输出的经编码数据的格式可以是H. 265/HEVC格式、Windows媒体视频格式、VC-1格式、MPEG-x格式 (例如,MPEG-1、MPEG-2或MPEG-4)、H. 26x格式 (例如,H. 261、H. 262、H. 263或H. 264) 或其它格式的变型或扩展。

[0068] 编码器 (340) 可以将帧分割成相同尺寸或不同尺寸的多个小块。例如,编码器 (340) 沿小块行和小块列来拆分帧,这些小块行和小块列利用帧边界定义在该帧内的小块的水平和垂直边界,其中每个小块是矩形区域。小块通常被用于提供并行处理的选项。帧还可以被组织成一个或多个片,其中一个片可以是整个帧或该帧的区域。片可以独立于帧中的其它片编码,这改善了错误复原性。出于编码和解码的目的,片或小块的内容被进一步分隔成块或其它采样值集。

[0069] 对于根据H. 265/HEVC标准的句法来说,编码器将帧 (或片或小块) 的内容拆分成编码树单元。编码树单元 (“CTU”) 包括被组织为亮度编码树块 (“CTB”) 的亮度样本值,并且对应的被组织为两个色度CTB的色度样本值。CTU (和其CTB) 的尺寸由编码器选择。亮度CTB可包含例如64x64、32x32或16x16亮度样本值。CTU包括一个或多个编码单元。编码单元 (“CU”) 具有亮度编码块 (“CB”) 和两个相应的色度CB。例如,具有64x64亮度CTB和两个64x64色度CTB (YUV 4:4:4格式) 的CTU可以被拆分成4个CU,其中每个CU包括一32x32亮度CB和两个32x32色度CB,并且每个CU可能被进一步拆分成更小的CU。或者,作为另一示例,具有64x64亮度CTB和两个32x32色度CTB (YUV 4:2:0格式) 的CTU可以被拆分成4个CU,其中每个CU包括一32x32亮度CB和两个16x16色度CB,并且每个CU可能被进一步拆分成更小的CU。CU的最小允许尺寸 (例如8×8、16×16) 可被信号化在比特流中。

[0070] 通常,CU具有诸如帧间或帧内之类的预测模式。出于信号化预测信息 (例如预测模式细节、位移值等) 和/或预测处理的目的,CU包括一个或多个预测单元。预测单元 (“PU”) 具有亮度预测块 (“PB”) 和两个色度PB。根据H. 265/HEVC标准,对于帧内预测CU,PU与CU具有相同的尺寸,除非CU具有最小尺寸 (例如8x8)。在那个情况下,CU可被拆分成四个较小的PU (例如,对于图片内预测,如果最小CU尺寸为8x8,则每个PU为4x4),或者PU可具有最小CU尺寸,如CU的句法元素所指示的。然而,对于在帧内BC预测中使用的非对称分区,CU可被划分为多个PU,如在图9中所示。在此情况下,较大的CU (例如,64x64、32x32或16x16) 或最小大小的CU可被划分为多个PU。

[0071] 出于残差编码和解码的目的,CU还具有一个或多个变换单元,其中变换单元 (“TU”) 具有亮度变换块 (“TB”) 和两个色度TB。帧内预测的CU中的PU可包含单个TU (尺寸上等于PU) 或多个TU。编码器决定如何将视频分割成CTU、CU、PU、TU等。

[0072] 在H. 265/HEVC实现中,片可以包括单个片段 (独立的片段) 或被划分成多个片段

(独立的片段和一个或多个从属的片段)。片段是包含在单个网络抽象层(“NAL”)单元中的、在小块扫描中连续排序的整数数目的CTU。对于独立的片段,片段头部包括适用于该独立的片段的句法元素的值。对于从属片段,截断的片段头部包括适用于该从属片段的句法元素的几个值,并且从属片段的其它句法元素的值是以解码顺序从在前的独立的片段的值中推导出来的。

[0073] 如本文中所使用的,术语“块”可以指示宏块、预测单元、残留数据单元、或者CB、PB或TB,或者某些其它样本值集,这取决于上下文。

[0074] 回到图3,编码器根据来自源帧(331)中的其它、先前重构的样本值的预测来表示该帧(331)内编码的块。对于帧内BC预测,图片内估计器估计块相对于其它、先前重构的采样值的块的位移。帧内预测参考区域是帧中被用于生成块的BC预测值的样本值区域。帧内预测区域可以用(在BV估计中确定的)块向量(“BV”)值来指示。以下描述了在图片内编码期间作出判定的示例方法。取决于实现,编码器可以使用输入采样值或重构采样值(在同一图片中先前经编码的采样值)对块执行BV估计。更多细节参看第V节对BV估计的描述。

[0075] 对于块的空间内预测,图片内估计器估计相邻的经重构的样本值到该块的外插。图片内估计器可以输出经熵编码的预测信息(例如帧内BC预测的BV值或帧内空间预测的预测模式(方向))。帧内预测预测器应用预测信息来确定帧内预测值。

[0076] 编码器(340)依据来自参考帧的预测来表示源帧(331)的帧间编码的预测的块。运动估计器估计块相对于一个或多个参考帧(369)的运动。当使用多个参考帧时,这多个参考帧可来自不同的时间方向或相同的时间方向。经运动补偿的预测参考区域是在参考帧中被用于生成当前帧的采样值的块的经运动补偿的预测值的采样值的区域。运动估计器输出诸如运动向量(“MV”)信息之类的运动信息,该运动信息被熵编码。运动补偿器将MV应用于参考帧(369)以确定用于帧间预测的经运动补偿的预测值。以下描述了在图片内编码期间作出判定的示例方法。

[0077] 编码器可以确定在块预测值(帧内或帧间)和对应的原始值之间的差值(如果有的话)。这些预测残留值将使用频率变换(如果该频率变换不被跳过)、量化和熵编码来被进一步编码。例如,编码器(340)为图片、小块、片和/或视频中的其它部分设置量化参数(“QP”)的值,并相应地量化变换系数。编码器(340)的熵编码器压缩经量化的变换系数值以及某些辅助信息(例如MV信息、BV预测值的索引值、BV差、QP值、模式判定、参数选择)。典型的熵编码技术包括指数-Golomb编码、Golomb-Rice编码、算术编码、差分编码、Huffman编码、行程长度编码、可变长度到可变长度(“V2V”)编码、可变长度到固定长度(“V2F”)编码、Lempel-Ziv(“LZ”)编码、字典编码、概率区间划分熵编码(“PIPE”)和上述编码的组合。熵编码器可针对不同种类的信息使用不同的编码技术,并可组合地应用多个技术(例如,通过应用Golomb-Rice编码,随后应用算术编码),并可从特定编码技术内的多个码表中进行选择。在一些实现中,可以跳过频率变换。在这种情况下,可以对预测残留值进行量化和熵编码。

[0078] 自适应去块滤波器可以被包括在编码器(340)中的运动补偿环路(也即“内环路”滤波)中以平滑经解码的帧中的块边界行和/或列上的间断。可替换地或另外地应用其它滤波(诸如去振铃滤波、自适应环路滤波(“ALF”)或样本自适应偏移(“SAO”)滤波;未示出)作为内环路滤波操作。

[0079] 由编码器(340)产生的经编码的数据包括针对各层比特流句法的句法元素。对于

根据H.265/HEVC的标准的句法,例如,图片参数集(“PPS”)是包含了与图片相关联的句法元素的句法结构。PPS可被用于单个图片,或者PPS可被重用于序列中的多个图片。PPS典型地与图片的经编码的数据分开地发送信号(例如PPS的一个NAL单元和用于图片的经编码的数据的一个或多个其它NAL单元)。在图片的经编码数据内,句法元素指示要为该图片使用哪个PPS。类似地,对于根据H.265/HEVC的标准的句法,序列参数集(“SPS”)是包含了与图片的序列相关联的句法元素的句法结构。比特流可包括单个SPS或多个SPS。SPS通常被信号化成与序列的其它数据分开,并且其它数据中的句法元素指示要使用哪个SPS。

[0080] 经编码的帧(341)和MMCO/RPS信息(342)(或与MMCO/RPS信息(342)等价的信息,因为在编码器(340)处已经知道各帧的依赖关系和排序结构)由解码处理仿真器(350)处理。解码处理仿真器(350)实现了解码器的一些功能,例如对任务进行解码以重构参考帧。以与MMCO/RPS信息(342)相一致的方式,解码处理仿真器(350)确定给定的经编码帧(341)是否需要被重构并被存储以供在对要编码的后续帧的帧间预测中用作参考帧。如果经编码的帧(341)需要被存储,则解码处理仿真器(350)对将由解码器进行的解码处理建模,该解码器接收经编码的帧(341)并产生相应的经解码的帧(351)。通过这么做,当编码器(340)已经使用已被存储在经解码帧存储区域(360)中的经解码的帧(369)时,解码处理仿真器(350)还使用来自存储区域(360)的经解码的帧(369)作为解码处理的一部分。

[0081] 经解码帧临时存储器存储区域(360)包括多个帧缓冲存储区域(361,362,⋯,36n)。以与MMCO/RPS信息(342)相一致的方式,解码处理仿真器(350)管理存储区域(360)中的内容,以便标识出具有编码器(340)不再需要将其用作参考帧的帧的任何帧缓冲区(361、362等)。在对解码处理进行建模之后,解码处理仿真器(350)在帧缓冲区(361、362等)中存储已经以此方式标识出的新解码的帧(351)。

[0082] 经编码的帧(341)和MMCO/RPS信息(342)被缓冲在临时的经编码数据区域(370)中。被聚集在经编码数据区域(370)中的经编码数据包含一个或多个图片的经编码数据作为基本经编码视频位流的句法的一部分。在经编码数据区域(370)中被聚集的经编码数据还可包括与经编码的视频数据相关的媒体元数据(例如作为一个或多个补充增强信息(“SEI”)消息或视频可用性信息(“VUI”)消息中的一个或多个参数)。

[0083] 来自临时的经编码数据区域(370)的经聚集的数据(371)由信道编码器(380)处理。信道编码器(380)可以分组化和/或复用经聚集的数据以供作为媒体流传输或存储(例如根据媒体程序流或传输流格式,例如ITU-T H.222.0|ISO/IEC 13818-1或因特网实时传输协议格式(例如IETF RFC 3550)),在这种情况下,信道编码器(380)可以添加句法元素作为媒体传输流的句法的一部分。或者,信道编码器(380)可以组织经聚集的数据以供存储成文件(例如根据媒体容器格式,诸如ISO/IEC 14496-12),在这种情况下信道编码器(380)可添加句法元素作为媒体存储文件的句法的一部分。或者,更一般地,信道编码器(380)可以实现一个或多个媒体系统复用协议或传输协议,在这种情况下,信道编码器(380)可以添加句法元素作为一个或多个协议的句法的一部分。信道编码器(380)将输出提供给信道(390),该信道(390)表示存储、通信连接或该输出的另一信道。信道编码器(380)或信道(390)还可以包括例如用于前向纠错(“FEC”)编码和模拟信号调制的其它元素(未示出)。

[0084] IV. 示例解码器系统。

[0085] 图4是结合其可实现所描述的一些实施例的示例解码器系统(400)的框图。解码器

系统(400)可以是能够在多种解码模式(例如针对实时通信的低等待时间解码模式以及针对来自文件或流中的媒体回放的较高等待时间解码模式)中的任一者操作的通用解码工具,或它可以是适用于一种这样的解码模式的专用解码工具。解码器系统(400)可以被实现为操作系统模块、应用库的一部分或独立的应用。总体上,解码器系统(400)从信道(410)接收经编码数据并产生经重构的帧作为针对输出目的地(490)的输出。

[0086] 解码器系统(400)包括信道(410),该信道(410)可以表示存储、通信连接或针对作为输入的经编码数据的另一信道。信道(410)产生已经被信道编码的经编码数据。信道解码器(420)可以处理经编码的数据。例如,信道解码器(420)可以(例如根据诸如ITU-T H.222.0|ISO/IEC 13818-1的媒体程序流或传输流格式,或诸如IETF RFC 3550的因特网实时传输协议格式)去分组化和/或去复用已被聚集以供作为媒体流来传输或存储的数据,在这种情况下,信道解码器(420)可以解析所添加的作为媒体传输流的句法的一部分的句法元素。或者,信道解码器(420)可以(例如根据诸如ISO/IEC 14496-12的媒体容器格式)来将已被聚集以供作为文件来存储的经编码的视频数据分开,在这种情况下,信道解码器(420)可以解析所添加的作为媒体存储文件的句法的一部分的句法元素。或者,更一般地,信道解码器(420)可以实现一个或多个媒体系统去复用协议或传输协议,在这种情况下,信道解码器(420)可以解析所添加的作为(一个或多个)协议的句法的一部分的句法元素。信道(410)或信道解码器(420)还可以包括例如用于FEC解码和模拟信号解调的其它元素(未示出)。

[0087] 从信道解码器(420)输出的经编码的数据(421)被存储在临时的经编码数据区域(430)中,直到已经接收到足够数量的这样的数据。经编码数据(421)包括经编码的帧(431)和MMCO/RPS信息(432)。在经编码数据区域(430)中的经编码数据(421)包含一个或多个图片的经编码数据作为基本编码的视频比特流的句法的一部分。在经编码数据区域(430)中的经编码数据(421)还可包括与经编码的视频数据相关的媒体元数据(例如作为一个或多个SEI消息或VUI消息中的一个或多个参数)。

[0088] 一般来说,经编码数据区域(430)临时存储经编码数据(421),直到这样的经编码数据(421)被解码器(450)使用。此时,经编码帧(431)和MMCO/RPS信息(432)的经编码数据被从经编码数据区域(430)传输到解码器(450)。当解码继续时,新的经编码数据被添加到经编码数据区域(430)并且保留在经编码数据区域(430)中的最旧的经编码数据被传输到解码器(450)。

[0089] 解码器(450)解码经编码帧(431)以产生相应的经解码帧(451)。在适当时,当执行其解码处理时,解码器(450)可以将一个或多个先前解码的帧(469)用作帧间预测的参考帧。解码器(450)从经解码帧临时存储器存储区域(460)中读取这样的先前解码的帧(469)。通常,解码器(450)包括执行诸如熵解码、帧内预测、运动补偿帧间预测、逆量化、逆频率变换(如果未被跳过)和合并小块之类的解码任务的多个解码模块。由解码器(450)执行的确切操作可以取决于压缩格式而变化。

[0090] 例如,解码器(450)接收经压缩的帧或帧序列的经编码数据,并产生包括经解码的帧(451)的输出。在解码器(450)中,缓冲器接收针对经压缩的帧的经编码数据,并在适当时,使得接收到的经编码数据对熵解码器可用。熵解码器对经熵编码的量化数据以及经熵编码的辅助信息进行熵解码,通常应用编码器中执行的熵编码的逆来进行熵解码。运动补偿器将运动信息应用于一个或多个参考帧以形成针对正被重构的帧的任何帧间编码块的

经运动补偿的预测值。帧内预测模块可从相邻的、先前重构的样本值中空间地预测当前块的样本值,或者对于帧内BC预测,使用帧中的帧内预测区域的先前重构的样本值来预测当前块的样本值。帧内预测区域可用BV值来指示。解码器(450)还重构预测残留值。逆量化器对经熵解码的数据进行逆量化。例如,解码器(450)基于比特流中的句法元素来设置图片、小块、片和/或视频中的其他部分的QP,并相应地对变换系数进行逆量化。逆频率变换器将量化的频域数据转换成空间域数据。在一些实现中,可以跳过频率变换,在这样的情况中逆频率变换也被跳过。如果这样,预测残留值可以被熵解码并逆量化。对于帧间预测的块,解码器(450)将重构的预测残留值与经运动补偿的预测值组合。解码器(450)可以类似地将预测残留值与来自帧内预测的预测值组合。自适应去块滤波器被包括在视频解码器(450)中的运动补偿环路内以平滑经解码的帧(451)中的块边界行和/或列上的间断。可替换地或另外地应用其它滤波(例如去振铃滤波、ALF或SAO滤波;未示出)作为内环路滤波操作。

[0091] 经解码帧临时存储器存储区域(460)包括多个帧缓冲器存储区域(461、462、...、46n)。经解码帧存储区域(460)是经解码图片缓冲区的一个示例。解码器(450)使用MMC0/PS信息(432)来标识该解码器可将经解码帧(451)存储在其中的帧缓冲器(461、462等)。解码器(450)将经解码的帧(451)存储在那个帧缓冲器中。

[0092] 输出序列发生器(480)标识按输出次序将产生的下一帧何时可在经编码帧存储区域(460)中获得。当按输出顺序将产生的下一帧(481)可在经解码帧存储区域(460)中获得时,输出序列发生器(480)读取该下一帧并将其输出到输出目的地(490)(例如显示器)。一般来说,输出序列发生器(480)将帧从经解码帧存储区域(460)中输出的次序可以与解码器(450)解码这些帧的次序不同。

[0093] V. 示例视频编码器。

[0094] 图5a和5b是可结合其实现所描述的一些实施例的通用视频编码器(500)的框图。编码器(500)接收包括当前图片的视频图片序列作为输入视频信号(505)并在经编码视频比特流(595)中产生经编码数据作为输出。

[0095] 编码器(500)是基于块的并使用取决于实现的块格式。块还可在不同的阶段上被进一步细分,例如在预测、频率变换和/或熵编码阶段。例如,图片可以被划分成64x64块、32x32块或16x16块,这些块进而可以被划分成更小的样本值块以用于编码和解码。在针对H.265/HEVC标准的编码的实现中,编码器将图片分割成CTU(CTB)、CU(CB)、PU(PB)和TU(TB)。块(例如CU)可被非对称地分区为更小的块(例如,PU)以用于帧内BC预测的目的,如图9中所示。

[0096] 编码器(500)使用图片内编码和/或图片间编码来压缩图片。编码器(500)的许多组件被用于图片内编码和图片间编码两者。由这些组件执行的确切操作可取决于所压缩的信息的类型而变化。

[0097] 小块化模块(510)可选地将图片分割成相同尺寸或不同尺寸的多个小块。例如,小块化模块(510)沿小块行和小块列来拆分图片,所述小块行和小块列利用图片边界定义在图片内的小块的水平和垂直边界,其中每个小块是矩形区域。在H.265/HEVC实现中,编码器(500)将图片分割成一个或多个片,其中每个片包括一个或多个片段。

[0098] 通用编码控件(520)接收输入视频信号(505)的图片以及来自编码器(500)的各个模块的反馈(未示出)。整体上,通用编码控件(520)将控制信号(未示出)提供给其它模块



(例如小块化模块(510)、变换器/缩放器/量化器(530)、缩放器/逆变换器(535)、图片内估计器(540)、运动估计器(550)以及帧内/帧间切换)以设置和改变编码期间的编码参数。特别是,通用编码控制(520)可在编码期间管理关于分区的决策。更一般而言,通用编码控件(520)可以决定在编码期间是否和如何使用帧内BC预测。通用编码控件(520)还可以估计编码期间的中间结果,例如执行速率-失真分析来估计。通用编码控件(520)产生指示在编码期间作出的判定的通用控制数据(522),使得对应的解码器可以作出一致的判定。通用控制数据(522)被提供给头部格式化器/熵编码器(590)。

[0099] 如果使用图片间预测来预测当前的图片,则运动估计器(550)估计输入视频信号(505)的当前图片的采样值块相对于一个或多个参考图片的运动。经解码的图片缓冲器(570)缓冲一个或多个经重构的先前编码的图片以供用作参考图片。在确定如何针对运动估计来划分块时,运动估计器(550)可应用从上到下方法或从下到上方法,如下所述。运动估计器(550)可使用如下所述的搜索图案或其它搜索图案。当使用多个参考图片时,这多个参考图片可以来自不同的时间方向或相同的时间方向。运动估计器(550)产生诸如MV数据、合并模式索引值和参考图片选择数据之类的辅助信息运动数据(552)。运动数据(552)被提供给头部格式化器/熵编码器(590)以及运动补偿器(555)。

[0100] 运动补偿器(555)将MV应用于来自经解码图片缓冲器(570)的(一个或多个)经重构的参考图片。运动补偿器(555)产生针对当前图片的经运动补偿的预测。

[0101] 在编码器(500)内的分开的路径中,图片内估计器(540)确定如何执行对输入视频信号(505)的当前图片的样本值块的图片内预测。当前图片可全部或部分使用图片内编码来编码。对于空间内预测,使用当前图片的重构(538)的值,图片内估计器(540)确定如何从当前图片的邻近的、先前重构的样本值中空间地预测当前图片的当前块的样本值。

[0102] 或者,对于使用BV值的帧内BC预测,图片内估计器(540)估计当前块的样本值到当前图片内的不同候选参考区域的位移。在确定如何针对BV估计(以及帧内BC预测)来划分块时,图片内估计器(540)可应用从上到下方法或从下到上方法,如下所述。图片内估计器(540)可使用如下所述的搜索图案或其它搜索图案。对于帧内BC预测,帧内预测估计器(540)可使用以下所描述的一个或多个约束来约束BV选择过程。

[0103] 取决于实现,编码器可以使用在循环内滤波之前的输入采样值、经重构的采样值,或在循环内滤波之后的经重构的采样值,来执行当前块的BV估计。通常,通过使用输入采样值或未经滤波的经重构的采样值用于BV估计,编码器可以避免序列化处理的瓶颈(该瓶颈可能由在BV估计/帧内BC预测之前滤波参考区域的经重构的采样值而导致)。另一方面,存储所述未经滤波的、经重构的样本值使用了额外的存储器。而且,如果在BV估计之前应用了循环内滤波,可能存在一个在解码当前区域之后将应用的滤波处理和正被用于BV估计/帧内BC预测的区域之间重叠的影响区域。在这样的情况中,BV估计/帧内BC预测将在滤波操作的该方面之前被应用。在一些实现中,编码器可以在BV估计/帧内BC预测之前应用某些循环内滤波操作,并且在随后的处理阶段中执行附加的或替换的滤波。

[0104] 或者,对于图片内字典编码模式,块中的像素是使用存储在字典或其他位置的先前的样本值来编码的,其中一像素是各共置样本值的集合(例如,RGB三元组或YUV三元组)。例如,编码器(500)可计算先前重构的样本值(例如,1个像素、2个像素、4个像素、8个像素并以此类推的群组)的散列值并将那些散列值与正被编码的当前像素集合的散列值进行比



较。长度为一或更多的匹配可在先前重构的样本值中基于散列比较来标识出。(一个或多个)当前像素(或样本值)可使用标识在先前像素内(例如,字典中)的位置的偏移以及指示从该偏移预测出的像素数的长度以各种1-D和伪2-D字典模式来编码。通常,没有计算出针对用图片内字典编码模式编码的块的残留值。

[0105] 图片内估计器(540)产生帧内预测数据(542)作为副信息,帧内预测数据(442)为诸如指示帧内预测是否使用空间预测、帧内BC预测或字典模式、预测模式方向(用于空间内预测)、BV值(用于帧内BC预测)以及偏移和长度(用于字典模式)的信息。帧内预测数据(542)被提供给头部格式化器/熵编码器(590)以及图片内预测器(545)。

[0106] 根据帧内预测数据(542),图片内预测器(545)在空间上从当前图片的邻近的先前重构的采样值中预测当前图片的当前块的采样值。或者,对于帧内BC预测,图片内预测器(545)使用图片内预测参考区域的先前重构的样本值来预测当前块的样本值,图片内预测参考区域由针对当前块的BV值指示。在一些情况中,BV值可以是BV预测值(预测的BV值)。在其它情况中,BV值可以不同于其被预测的BV值,在该情况中BV差异指示了在被预测的值BV和BV值之间的差值。或者,对于图片内字典模式,图片内预测器(545)使用偏移和长度重构各像素。

[0107] 帧内/帧间切换选择针对给定块的预测(558)将是经运动补偿的预测还是图片内预测。

[0108] 对于非字典模式,当残留值编码没有被跳过时,预测(558)的块和输入视频信号(505)的原始当前图片的对应部分之间的差(如果有的话)提供各残留(518)的值。在当前图片的重构期间,当残留值已经被编码/信号化,经重构的残留值与预测(558)组合以从视频信号(505)中生成对原始内容的近似重构或准确重构(538)。(在有损压缩中,一些信息从视频信号(505)中丢失。)

[0109] 在变换器/缩放器/量化器(530)中,对于非字典模式,当不跳过频率变换时,频率变换器将空间域视频信息转换为频域(即频谱、变换)数据。对于基于块的视频编码,频率变换器将离散余弦变换(“DCT”)、其整数近似、或其它类型的前向块变换(例如离散正弦变换或其整数近似)应用于预测残留数据的块(如果预测(558)为空则应用于采样值数据),产生频率变换系数的块。变换器/缩放器/量化器(530)可以应用具有可变块尺寸的变换。在这种情况下,变换器/缩放器/量化器(530)可确定要对当前块的残留值使用哪些块尺寸的变换。缩放器/量化器对变换系数进行缩放和量化。例如,量化器将以逐图片基础、逐小块基础、逐片基础、逐块基础、频率专用基础或其它基础来变化的量化步进尺寸将死区标量化应用于频域数据。经量化的变换系数数据(532)被提供给头部格式化器/熵编码器(590)。如果频率变换被跳过,则缩放器/量化器可对预测残留数据块(或如果预测(558)为空则对样本值)进行缩放和量化,从而生成要提供给头部格式化器/熵编码器(590)的经量化的值。

[0110] 在缩放器/逆变换器(535)中,对于非字典模式,缩放器/逆量化器对经量化的变换系数执行逆缩放和逆量化。当变换阶段没有被跳过时,逆频率变换器执行逆频率变换,生成经重构的预测残留值或采样值的块。如果变换阶段已经被跳过,则也跳过逆频率变换。在这种情况下,缩放器/逆量化器可对预测残留值(或样本值数据)的块执行逆缩放和逆量化,从而产生经重构的值。当残留值已经被编码/信号化时,编码器(500)将经重构的残留值与预测(558)的值(例如,经运动补偿的预测器、图片内预测器)组合以形成重构(538)。当残留值

尚未被编码/信号化时,编码器(500)将预测(558)的值用作重构(538)。

[0111] 对于图片内预测,重构(538)的值可以被馈送回图片内估计器(540)和图片内预测器(545)。重构(538)的值可以被用于对后续图片的经运动补偿的预测。重构(538)的值可被进一步滤波。对于视频信号(505)的给定图片,滤波控件(560)确定如何对重构(538)的值执行去块滤波和SAO滤波。滤波控件(560)生成滤波控制数据(562),它被提供给头部格式化器/熵编码器(590)和合并器/(一个或多个)滤波器(565)。

[0112] 在合并器/(一个或多个)滤波器(565)中,编码器(500)将来自不同小块的内容合并到图片的经重构版本中。编码器(500)根据滤波器控制数据(562)选择性地执行去块滤波和SAO滤波,以便自适应地平滑跨各图片中的各边界上的间断。可替换地或另外地应用其它滤波(例如去振铃滤波或ALF;未示出)。取决于编码器(500)的设置,小块边界可以被选择性地滤波或根本不被滤波,并且编码器(500)可以在经编码的位流内提供指示是否应用这样的滤波的句法。经解码的图片缓冲器(570)缓冲经重构的当前图片以供在后续的经运动补偿的预测中使用。

[0113] 头部格式化器/熵编码器(590)对通用控制数据(522)、经量化的变换系数数据(532)、图片内预测数据(542)、运动数据(552)以及滤波器控制数据(562)进行格式化和/或熵编码。对于运动数据(552),头部格式化器/熵编码器(590)可以选择并熵编码合并模式索引值,或可使用默认的MV预测器。在一些情况中,头部格式化器/熵编码器(590)还确定MV值(相对于MV值的MV预测器)的MV差异,随后例如使用上下文自适应二进制算术编码对该MV差异进行熵编码。对于帧内预测数据(542),可使用BV预测来编码BV值。BV预测可使用默认BV预测器(例如来自一个或多个邻近块)。当多个BV预测器是可能的时,BV预测器索引可以指示多个BV预测器中的哪个用于BV预测。头部格式化器/熵编码器(590)可以选择并熵编码BV预测器索引值(用于帧内BC预测),或可以使用默认BV预测器。在一些情况中,头部格式化器/熵编码器(590)还(相对于BV值的BV预测值)确定BV值的BV差,随后对BV差进行熵编码,例如使用上下文自适应二进制算术编码来对BV差进行熵编码。

[0114] 头部格式化器/熵编码器(590)在经编码的视频位流(595)中提供经编码数据。经编码的视频比特流的格式(595)可以是H.265/HEVC格式、Windows媒体视频格式、VC-1格式、MPEG-x格式(例如,MPEG-1、MPEG-2或MPEG-4)、H.26x格式(例如,H.261、H.262、H.263或H.264)或其它格式的变型或扩展。

[0115] 取决于实现和期望的压缩类型,编码器(500)的模块可被添加、省略、拆分成多个模块、与其它模块组合、和/或用相似的模块来替代。在替换实施例中,具有不同模块和/或其它配置的模块的编码器执行一个或多个所描述的技术。编码器的具体实施例通常使用编码器(500)的变型或补充版本。所示的编码器(500)内的各模块之间的关系指示了信息在编码器中的一般流动;为简明起见,未示出其它关系。

[0116] VI. 示例视频解码器。

[0117] 图6是结合其可实现所描述的一些实施例的通用解码器(600)的框图。解码器(600)接收经编码的视频比特流(605)中的经编码的数据并产生包括经重构视频(695)的图片的输出。经编码的视频比特流(605)的格式可以是HEVC格式、Windows媒体视频格式、VC-1格式、MPEG-x格式(例如,MPEG-1、MPEG-2或MPEG-4)、H.26x格式(例如,H.261、H.262、H.263或H.264)或其它格式的变体或扩展。

[0118] 解码器(600)是基于块的并使用取决于实现的块格式。块可以在不同的阶段被进一步细分。例如,图片可以被划分成64x64块、32x32块或16x16块,这些块随后可以被划分成更小的样本值块。在针对HEVC标准的解码实现中,将图片分割成CTU(CTB)、CU(CB)、PU(PB)和TU(TB)。块(例如CUI)可被非对称地分区为更小的块(例如,PU)以用于帧内BC预测的目的,如图9中所示。

[0119] 解码器(600)使用图片内解码和/或图片间解码来对图片进行解压缩。解码器(600)的许多组件被用于图片内解码和图片间解码这两者。由这些组件执行的确切操作可以取决于正被解压缩的信息的类型而变化。

[0120] 缓冲器接收经编码的视频比特流(605)中的经编码数据,并使得接收到的经编码数据可用于解析器/熵解码器(610)。解析器/熵解码器(610)通常应用在编码器(500)中执行的熵编码的逆(例如上下文自适应二进制算术解码)对经熵编码的数据进行熵解码。作为解析和熵解码的结果,解析器/熵解码器(610)产生通用控制数据(622)、经量化的变换系数数据(632)、图片内预测数据(642)、运动数据(652)以及滤波器控制数据(662)。对于帧内预测数据(642),如果BV预测器索引值被信号化,解析器/熵解码器(610)可以例如使用上下文自适应二进制算术解码来熵解码DV预测器索引值。在一些情况中,解析器/熵解码器(610)还可以(例如使用上下文自适应二进制算术解码)对BV值的BV差进行熵解码,随后将该BV差与相应的BV预测值组合以重构BV值。在其它情况中,该BV差异可以自比特流中忽略,并且BV值仅仅是(例如用BV预测器索引值所指示的)BV预测器。

[0121] 通用解码控件(620)接收通用控制数据(622)并将控制信号(未示出)提供给其它模块(例如缩放器/逆变换器(635)、图片内预测器(645)、运动补偿器(655)以及帧内/帧间切换)以设置和改变解码期间的解码参数。

[0122] 如果当前图片是使用图片间预测来预测的,则运动补偿器(655)接收运动数据(652),例如MV数据、参考图片选择数据、以及合并模式索引值。运动补偿器(655)将MV应用于来自经解码图片缓冲器(670)的一个或多个经重构的参考图片。运动补偿器(655)产生针对当前图片中的经帧间编码的块的运动补偿的预测。经解码图片缓冲器(670)存储一个或多个先前重构的图片以供用作参考图片。

[0123] 在解码器(600)内的分开路径中,帧内预测预测器(645)接收帧内预测数据(642),诸如指示帧内预测是否使用空间预测或帧内BC预测或字典模式、以及预测模式方向(对于帧内空间预测)、BV值(对于帧内BC预测)或偏移和长度(对于字典模式)的信息。对于帧内空间内预测,根据预测模式数据使用当前图片的重构(638)的值,图片内预测器(645)从当前图片的相邻的、先前重构的样本值中空间地预测当前图片的当前块的样本值。或者,对于使用BV值的帧内BC预测,图片内预测器(645)使用帧内预测区域的先前重构的样本值来预测当前块的样本值,图片内预测区域的先前重构的样本值由针对当前块的BV值来指示。或者,对于图片内字典模式,图片内预测器(645)使用偏移和长度重构各像素。

[0124] 帧内/帧间切换选择运动补偿的预测或图片内预测的值用作给定块的预测(658)。例如,当遵循HEVC句法时,可以基于为图片中的CU编码的句法元素来控制图片内/图片间切换,该CU可以包括图片内预测的CU和图片间预测的CU。当残留值已经被编码/信号化时,解码器(600)将预测(658)与重构的残留值组合以产生来自视频信号的内容的重构(638)。当残留值尚未被编码/信号化时,解码器(600)将预测(658)的值用作重构(638)。

[0125] 为了在残留值已被编码/信号化时重构残留值,缩放器/逆变换器(635)接收并处理经量化的变换系数数据(632)。在缩放器/逆变换器(635)中,缩放器/逆量化器对经量化的变换系数执行逆缩放和逆量化。逆频率变换器执行逆频率变换,从而产生经重构的预测残留值或样本值的块。例如,逆频率变换器将逆块变换应用到频率变换系数,从而产生样本值数据或预测残留数据。逆频率变换可以是逆DCT、其整数近似、或另一种类型的逆频率变换(例如逆离散正弦变换或其整数近似)。如果频率变换在编码期间被跳过,则逆频率变换也被跳过。在这种情况下,缩放器/逆量化器可对预测残留值(或样本值数据)的块执行逆缩放和逆量化,从而产生经重构的值。

[0126] 对于图片内预测,重构(638)的值可以被馈送回图片内预测器(645)。对于图片间预测,重构(638)的值可以被进一步滤波。在合并器/(一个或多个)滤波器(665)中,解码器(600)将来自不同小块的内容合并到图片的经重构版本中。解码器(600)根据滤波器控制数据(662)和滤波自适应规则来执行去块滤波和SAO滤波,以便自适应地平滑各帧中的各边界上的间断。可替换地或另外地应用其它滤波(例如去振铃滤波或ALF;未示出)。小块边界可以被选择性地滤波或根本不滤波,这取决于解码器(600)的设置或经编码的比特流数据内的句法指示。经解码的图片缓冲器(670)缓冲经重构的当前图片以供在后续的运动补偿的预测中使用。

[0127] 解码器(600)还可包括后处理滤波器。后处理滤波器(608)可以包括去块滤波、去振铃滤波、自适应维纳滤波、膜颗粒再现滤波、SAO滤波或另一种类的滤波。尽管内环路滤波是在运动补偿环路中对帧的经重构的样本值执行的,并因此影响参考帧的样本值,但后处理滤波器(608)是在经重构的样本值被输出以供显示之前在运动补偿环路外部被应用于经重构的样本值的。

[0128] 取决于实现和期望的解压缩类型,解码器(600)的模块可被添加、被省略、被拆分成多个模块、被与其他模块组合和/或被替换为类似的模块。在替换实施例中,具有不同模块和/或其它配置的模块的解码器执行一个或多个所描述的技术。解码器的具体实施例通常使用解码器(600)的变型或补充版本。所示的解码器(600)内的模块之间的关系指示信息在解码器中的一般流动;为简明起见,未示出其它关系。

[0129] VII. 带有非对称分区的块内复制预测

[0130] 本节呈现了经帧内BC预测的块的非对称分区的示例。使用非对称分区可以允许编码器适配分区到块内的强定义但是不规则的图像图案,所述图案在文本、网页、以及屏幕捕捉视频和其它人造视频的其它部分中很常见。

[0131] A. 帧内BC预测模式和BV值——介绍。

[0132] 对于帧内BC预测,图片的当前块的样本值是使用同一图片中的样本值来预测的。BV值指示从当前块到图片的包括用于预测的样本值的区域(“参考区域”)的位移。参考区域提供了当前块的预测值。用于预测的样本值是先前重构的样本值,由此这些样本值可在编码期间在编码器处获得并在解码期间在解码器处获得。BV值在比特流中被信号化,并且解码器可以使用BV值来确定图片的参考区域以用于预测,其也在解码器处被重构。帧内BC预测是一种形式的图片内预测——针对图片的块的帧内BC预测不使用除了同一图片中的样本值以外的任何样本值。

[0133] 图7示出针对当前图片(710)的当前块(730)的帧内BC预测。当前块可以是编码单

元(“CU”)的编码块(“CB”)、预测单元(“PU”)的预测块(“PB”)、变换单元(“TU”)的变换块(“TB”)或其它块。当前块的尺寸可以是 $64 \times 64$ 、 $32 \times 32$ 、 $16 \times 16$ 、 $8 \times 8$ 或某个其它尺寸。更一般地,当前块的尺寸为 $m \times n$ ,其中 $m$ 和 $n$ 中的每一者都是整数,并且其中 $m$ 和 $n$ 可彼此相等或具有不同的值。这样,当前块可以是正方形或矩形。替换地,当前块可具有某种其它形状。

[0134] BV(740)指示当前块(730)到图片的包括用于预测的样本值的参考区域(750)的位移(或偏移)。由BV(740)指示的参考区域(750)有时被称为当前块(730)的“匹配块”。匹配块可与当前块(730)相同,或它可以是当前块(730)的近似。假设当前块的左上位置是在当前图片中的位置 $(x_0, y_0)$ 处,并假设参考区域的左上位置在当前图片中的位置 $(x_1, y_1)$ 处。BV指示位移 $(x_1 - x_0, y_1 - y_0)$ 。例如,如果当前块的左上位置是在位置(256, 128)处,并且参考区域的左上位置在位置(126, 104)处,BV值是(-130, -24)。在本示例中,负水平位移指示在当前块的左边的位置,而负垂直位移指示在当前块之上的位置。

[0135] 帧内BC预测可使用BC操作通过利用冗余性(例如在图片内部的重复的图案)来改善编码效率。然而,考虑到编码器可能估计的候选块的数量,为当前块查找匹配块可以在计算上是复杂的并且耗时的。图8示出在块匹配操作中用于当前图片(810)的当前块(830)的某些候选块。四个BV(841, 842, 843, 844)指示了四个候选块的位移。候选块可在当前图片(810)的经重构内容内的任何地方。(通常从左向右随后从上到下地编码块。)候选块可以与其它候选块重叠,如由BV(843, 844)指示的候选块所示。

[0136] 在一些示例实现中,帧内预测区域(850)被约束到在与当前块(830)相同的片或小块内。这样的帧内BC预测不使用在其它片或小块中的采样值。帧内预测区域(850)的位置可受制于一个或多个其它约束(例如,对于搜索范围的约束,与对帧间编码的块的重构的采样值的使用有关的约束)。替换地,帧内预测区域(850)的位置不被约束在当前图片(810)的重构内容内(即,全部搜索范围)。

[0137] 具有帧内BC预测的预测模式的块可以是CB、PB或其它块。当块是CB时,可以在CU级上用信号通知块的BV值(而CU中的其它CB使用相同的BV或其经缩放的版本)。或者,当块是PB时,可以在PU级上用信号通知块的BV值(而PU中的其它PB使用相同的BV或其经缩放的版本)。更一般地,在块的适当的句法级上用信道通知帧内-BC预测块的BV。

[0138] 根据帧内BC预测模式,预测的块复制操作可以在CB级(当根据CB信号化BV时)或PB(当根据PB信号化BV时)被执行。例如,假设 $16 \times 16$ CB具有单个 $16 \times 16$ PB。(PB的)BV被应用于块复制 $16 \times 16$ 区域。当帧内预测区域被约束为不与正被预测的 $16 \times 16$ 块重叠时,BV在水平或垂直上具有至少16的幅值(绝对值)。

[0139] 替换地,即使当为PB或CB信号化BV时,可以在PB或CB内的TB级执行块复制操作。在这种方式中,BV,就像应用于TB一样,可参考相同PB或CB中的其它TB的位置。例如,假设 $16 \times 16$ CB具有单个 $16 \times 16$ PB,但是出于残留编码/解码的目的被拆分成十六个 $4 \times 4$ TB。(PB的)BV被应用以按照光栅扫描次序对第一TB的 $4 \times 4$ 区域进行块复制,随后同一BV被应用以按照光栅扫描次序对第二TB的 $4 \times 4$ 区域进行块复制,依此类推。在将残留值与对那些之前重构的TB的预测值组合之后,用于TB的BC操作中的 $4 \times 4$ 区域可包括在同一CB中之前被重构的TB中的位置。(BV仍然不参考正被预测的同一TB中的位置。)在TB级应用BC操作促进了对于具有相对小的幅值的BV的使用。

[0140] TB级重叠隐式地创造了在重构PU内的TU时的串行依赖。这减少了并行计算的机会

并潜在地减少了硬件级的吞吐量。PU级重叠可消除TB中的这种串行依赖。在此情况下,块复制操作在可重叠的PB级执行。以此方式,BV,就像应用于PB一样,可参考相同CB中的其它PB的位置。假定出于帧内预测之目的,一 $16 \times 16$ CB被化分为两个PB(例如,两个 $16 \times 8$ PB,或两个 $8 \times 16$ PB,或一 $4 \times 16$ PB和一 $12 \times 16$ PB等)。一个PB的BV被应用以块复制该PB的区域,随后另一PB的BV被应用以块复制该另一PB的区域。在将残留值与第一PB的预测值组合之后,在第二PB的BC操作中使用的区域可包括在同一CB中之前被重构的第一PB中的位置。(BV仍然不参考正被预测的同一PB中的位置。)在PB级应用BC操作促进了对具有相对小量的BV的使用(相较于在CB级应用BC操作)。而且,当在PB级应用BC操作时,对于PB内的TB,TU级并行处理仍旧被允许。

[0141] 一CU的色度块的帧内BC预测操作大体对应于该CU的亮度块的帧内BC预测操作。正常情况下,色度PB和色度TB的分段直接对应于该CU中的亮度PB和亮度TB的分段。当视频的格式为YUV 4:4:4时,色度PB和TB的大小与相应PB和TB的大小匹配。当视频的格式是YUV 4:2:0时,色度PB和TB的宽度和高度是对应的亮度PB和TB的一半。然而,如果亮度TB具有最小变换尺寸,则使用具有该最小变换尺寸的单个色度TB。当视频的格式是YUV4:2:2时,色度PB和TB的宽度是对应的亮度PB和TB的一半。

[0142] 在一些实现中,对于帧内BC预测的CU,对PU中的色度块的帧内BC预测使用与对于PU中的亮度块的帧内BC预测相同的BV值,这通常是当色度数据已减少了相对于亮度数据的分辨率时(例如,当格式是YUV 4:2:0格式时,对于水平和垂直分量BV值被除以2;或者,当格式是YUV 4:2:2格式时,对于水平分量BV值被除以2)在缩放和舍入之后进行的。替代地,可针对PU的亮度块和色度块用信号指示不同的BV值。

[0143] 在一些实现中,编码器在BV估计或MV估计期间在标识BV值或MV值时考虑亮度样本值。例如,编码器尝试将当前块(例如一PU的PB)的亮度样本值匹配到经重构的亮度样本值。然而,所得到的BV值或MV值也被应用到相应色度块的色度样本值。替代地,编码器在BV估计或MV估计期间在标识BV值或MV值时考虑亮度样本值和相应的色度样本值。

[0144] 在一些实现中,如果PU的亮度块的预测模式是帧内BC预测,则该PU的色度块的预测模式也是帧内BC预测。例如,用信号指示该PU的预测模式。替代地,PU的亮度块或色度块的预测模式可以是帧内BC预测,但不都是帧内BC预测。

[0145] B. 非对称分区

[0146] 图9示出了在一些示例实现中用于帧内BC预测的块的分区的示例(900)。使用帧内BC预测来编码一 $2N \times 2N$ 块。例如,该 $2N \times 2N$ 块为 $64 \times 64$ 块、 $32 \times 32$ 块、 $16 \times 16$ 块或 $8 \times 8$ 块。该 $2N \times 2N$ 块可以是在无分区情况下经帧内BC预测的。或者,该 $2N \times 2N$ 块可以按各种方式分区,如在图9中所示。

[0147] 该 $2N \times 2N$ 块可以被水平分区为两个分区。这两个分区可具有相同的尺寸——两个 $2N \times N$ 块——以用于对称分区。或者,这两个分区可以不对称。例如,上分区为 $2N \times N/2$ 块,而下分区为 $2N \times 3N/2$ 块。或者,作为另一示例,上分区为 $2N \times 3N/2$ 块,而下分区为 $2N \times N/2$ 块。从而,一 $64 \times 64$ 块可被分区为两个 $64 \times 32$ 块、一 $64 \times 16$ 块和一 $64 \times 48$ 块、或一 $64 \times 48$ 块和一 $64 \times 16$ 块。一 $32 \times 32$ 、 $16 \times 16$ 或 $8 \times 8$ 块可被类似地水平分区。

[0148] 该 $2N \times 2N$ 块可以相反被垂直分区为两个分区。这两个分区可具有相同的尺寸——两个 $N \times 2N$ 块——以用于对称分区。或者,这两个分区可以不对称。例如,左分区为

$N/2 \times 2N$ 块,而右分区为 $3N/2 \times 2N$ 块。或者,作为另一示例,左分区为 $3N/2 \times 2N$ 块,而右分区为 $N/2 \times 2N$ 块。从而,一 $64 \times 64$ 块可被分区为两个 $32 \times 64$ 块、一 $16 \times 64$ 块和一 $48 \times 64$ 块、或一 $48 \times 64$ 块和一 $16 \times 64$ 块。一 $32 \times 32$ 、 $16 \times 16$ 或 $8 \times 8$ 块可被类似地垂直分区。

[0149] 或者,该 $2N \times 2N$ 块可被分区为四个 $N \times N$ 分区,所述分区可以被进一步细分。例如,如图9中所示,给定 $N \times N$ 分区可被进一步分区为两个 $N \times N/2$ 块、两个 $N/2 \times N$ 块、或四个 $N/2 \times N/2$ 块。从而, $64 \times 64$ 块可被分区为四个 $32 \times 32$ 块,每个 $32 \times 32$ 块可被进一步分区为两个 $32 \times 16$ 块、两个 $16 \times 32$ 块、或四个 $16 \times 16$ 块。 $32 \times 32$ 、 $16 \times 16$ 或 $8 \times 8$ 块可类似地通过四叉树划分来分区为四个分区。

[0150] 在大多数情况下,在本文中使用的术语“ $N \times N$ ”分区也可被认为是 $2N \times 2N$ 分区。在H.265/HEVC实现中,术语 $N \times N$ 通常被用来描述PU或PB,而不是CU或CB。如本文所使用的,术语“ $N \times N$ 分区”或“ $N \times N$ 块”指示一 $2N \times 2N$ 当前块的分区(例如,作为分区模式的从上到下评估或分区模式的从下到上评估的一部分)。在此上下文中, $N \times N$ 分区或 $N \times N$ 块本身可被认为是 $2N \times 2N$ 块,并且在进一步分区中可被这样对待,除非该 $N \times N$ 分区或 $N \times N$ 块具有最小大小且不被进一步分区。 $N \times N$ 分区或 $N \times N$ 块内的分区的注释可被类似地调整。

[0151] 将非对称分区大小限制为 $N/2$ 的倍数可减少评估在编码期间使用哪些分区模式时的复杂度。替换地,编码器可考虑其它分区大小(例如, $N/4$ 的倍数或更一般地分区大小 $m \times n$ )。考虑其它分区大小可略微改善编码增益,但是也增加了编码期间搜索过程的复杂度,并且可能增加信令开销。

[0152] C. 包括帧内BC的编码或解码的示例技术具有非对称分区的预测

[0153] 图10示出包括针对具有非对称分区的帧内BC预测的编码的一般化技术(1000)。图像编码器或视频编码器(诸如参考图3或图5a-5b描述的编码器)可执行该技术(1000)。图11示出包括针对具有非对称分区的帧内BC预测的解码的一般化技术(1100)。图像解码器或视频解码器(诸如参考图4或图6描述的解码器)可执行该技术(1100)。

[0154] 参考图10,编码器对图像或视频进行编码(1010)以产生经编码的数据。作为编码的一部分,编码器执行为了帧内BC预测而被非对称分区的当前块的帧内BC预测。特别是,当前块可被划分为具有不同尺寸的两个分区。编码器将经编码的数据作为比特流的一部分来输出(1020)。

[0155] 参考图11,解码器将经编码的数据作为比特流的一部分来接收(1110)。解码器对经编码的数据进行解码(1120)以重构图像或视频。作为解码的一部分,解码器执行为了帧内BC预测而被非对称分区的当前块的帧内BC预测。特别是,当前块可被划分为具有不同尺寸的两个分区。

[0156] 例如,当前 $2N \times 2N$ 块被水平分区为(1)  $2N \times N/2$ 块和 $2N \times 3N/2$ 块,或(2)  $2N \times 3N/2$ 块和 $2N \times N/2$ 块。或者,当前 $2N \times 2N$ 块被垂直分区为(1)  $N/2 \times 2N$ 块和 $3N/2 \times 2N$ 块,或(2)  $3N/2 \times 2N$ 块和 $N/2 \times 2N$ 块。替换地,以某一其它方式来非对称地分区当前块。

[0157] 另一块可被对称地分区以用于帧内BC预测。例如,另一块是被分区为(1) 两个 $2N \times N$ 块,(2) 两个 $N \times 2N$ 块,或(3) 四个 $N \times N$ 块的 $2N \times 2N$ 块,每个块可被进一步分区为两个 $N \times N/2$ 块、两个 $N/2 \times N$ 块、或四个 $N/2 \times N/2$ 块。更一般而言,对于对称分区,该另一块可被划分为具有等同尺寸的分区。

[0158] VIII. 用于分区的编码器侧方法

[0159] 本节呈现了用于在运动估计或块向量估计(对于帧内BC预测)期间标识分区的各种编码器侧方法。

[0160] 例如,在章节V、VII.A、IX和X中呈现了帧内BC预测和BV估计的示例。考虑到当前块的大量可能候选块,BV估计可以是计算上复杂的。当BV值的搜索范围包括图片的所有先前重构的区域时,由于当前块可与其进行比较的候选块的较大数目,BV估计的计算复杂性特别成问题。对于整个帧的各候选块,操作的数量甚至更高。当编码器对非对称分区求值时,BV估计的复杂度被进一步增加。

[0161] 运动估计的计算复杂度也可能很高,如在下章节解释的。

[0162] A. 运动估计和MV值——介绍

[0163] 对于运动估计,当前图片中的当前块的样本值是使用另一图片(其被称为参考图片)中的样本值来预测的。运动矢量(“MV”)值指示从当前块在参考图片中的位置到参考图片的包括用于预测的样本值的区域(“参考区域”)的位移。参考区域提供了当前块的预测值。用于预测的样本值是先前重构的样本值,由此这些样本值可在编码期间在编码器处获得并在解码期间在解码器处获得。MV值在比特流中被信号化,并且解码器可以使用该MV值来确定参考图片的参考区域以用于预测,其也在解码器处被重构。当多个参考图片可用时,比特流还可包括关于要使用哪个参考图片来找到参考区域的指示。

[0164] 图12示出针对当前图片(1210)的当前块(1230)的运动预测。当前块可以是编码单元(“CU”)的编码块(“CB”)、预测单元(“PU”)的预测块(“PB”)、变换单元(“TU”)的变换块(“TB”)或其它块。当前块的尺寸可以是 $64 \times 64$ 、 $32 \times 32$ 、 $16 \times 16$ 、 $8 \times 8$ 或某个其它尺寸。更一般地,当前块的尺寸为 $m \times n$ ,其中 $m$ 和 $n$ 中的每一者都是整数,并且其中 $m$ 和 $n$ 可彼此相等或可具有不同的值。这样,当前块可以是正方形或矩形。替换地,当前块可具有某种其它形状。

[0165] MV值(1240)指示从当前块(1230)的位置到参考图片(1250)中包括用于预测的样本值的参考区域(有时被称为参考块)的位移(或偏移)。该MV值(1240)所指示的参考区域有时被称为当前块(1230)的“匹配块”。匹配块可与当前块(1230)相同,或它可以是当前块(1230)的近似。假设当前(1230)块的左上位置在当前图片(1210)中的位置 $(x_0, y_0)$ 处,并假设参考区域的左上位置在参考图片(1250)中的位置 $(x_1, y_1)$ 处。MV值(1240)指示位移 $(x_1 - x_0, y_1 - y_0)$ 。例如,如果当前块的左上位置在位置(256, 128)处,并且参考区域的左上位置在位置(126, 104)处,该MV值为(-130, -24)。在本示例中,负水平位移指示在当前块的左边的位置,而负垂直位移指示在当前块之上的位置。

[0166] 当前块(1230)的参考区域是在运动估计期间从多个候选块之中选出的。图12还示出运动估计中针对当前图片(1210)的当前块(1230)的某些候选块。四个MV值(1241, 1242, 1243, 1244)指示四个候选参考块的位移。一般来说,这些候选块可在参考图片(1250)内的任何地方。候选块可与其它候选块重叠,如针对MV值(1243, 1244)所指示的候选块示出的。运动估计的计算复杂度在MV值的搜索范围包含参考图片的全部区域时因当前块与大量候选块比较,而特别成问题。当运动估计针对多个参考图片执行时,该计算复杂性更加复杂。

[0167] B. BV值和MV值的精度

[0168] 在本文描述的许多示例中,BV值具有整数样本精度(integer-sample precision)。这样的BV值指示整数样本偏移。帧内BC预测常在编码人造视频(诸如屏幕内容视频)时使用,对于人造视频而言,分数样本位移很少见,所以整数样本精度对于BV值已足



够。替换地, BV值可指示分数样本偏移。例如, 具有1/2样本精度(1/2-sample precision)的BV值可指示具有1/2样本偏移(诸如1.5个样本、2.5个样本等等)的水平和/或垂直位移。或者, 具有1/4样本精度的BV值可指示具有1/4、1/2或3/4样本偏移的水平和/或垂直位移。或者, 具有1/8样本精度的BV值可指示具有1/8、1/4、3/8、1/2、5/8、3/4或7/8样本偏移的水平或垂直位移。或者, BV值可具有某种其它精度。

[0169] 另一方面, MV值通常具有分数样本精度, 因为分数样本精度位移在编码自然视频时是常见的。例如, MV值可具有1/2样本精度、1/4样本精度、1/8样本精度或某种其它精度。替换地, MV值具有整数样本精度。

[0170] C. 用于标识分区的方法

[0171] 一般而言, 编码器在运动估计或BV估计期间(用于帧内BC预测)在标识分区时可使用从上到下方法或从下到上方法。从下到上方法初始地评估最小大小分区的选项, 随后在评估后续更大大小分区的选项时使用来自该初始评估的结果。从下到上方法可能在计算上是昂贵的。另一方面, 从上到下方法初始地评估较大大小分区的选项, 随后在评估后续更小小分区的选项是使用来自该初始评估的结果。从上到下方法较不可能标识分区块的最优方式, 但是通常比从下到上方法在计算上更简单。

[0172] 对于关于分区方法的更多细节, 参见例如: (1) Sullivan等人在加拿大多伦多的国际声学、语言和信号处理会议(ICASSP) IEEE汇刊1991年5月第4卷第2661-2664页中的“Efficient Quadtree Coding of Images and Video(图像和视频的高效二叉树编码)”, 其描述了如何针对树结构编码(至少在某些场景下)在速率失真的意义上标识最佳树, (2) Sullivan等人在亚利桑那凤凰城的国际电信会议(GLOBECOM) IEEE汇刊1991年12月第85-90页中的“Rate-Distortion Optimized Motion Compensation for Video Compression using Fixed or Variable Size Blocks(针对使用固定或可变大块的视频压缩的速率失真优化的运动补偿)”, 其描述了如何将速率失真优化应用到运动补偿(带或不带树); 以及(3) Sullivan等人在IEEE图像处理汇刊1994年5月第IP-3卷第3号第327-331页中的“Efficient Quadtree Coding of Images and Video(图像和视频的高效二叉树编码)”, 其进一步描述了如何对于树结构编码在速率失真意义上标识最优树。

[0173] 在一些实现中, 应用以下描述的从上到下方法或从下到上方法之一, 编码器标识PU的分区。编码器可基于对PU的亮度PB的分析来设置PU的分区。或者, 编码器可基于对PU的亮度PB和色度PB两者的分析来设置PU的分区。在任一情况下, 以PU为基础设置的分区随后被应用到该PU的PB。替换地, 编码器标识另一类型的单元或块的分区。

[0174] 1. 用于标识分区的从上到下方法

[0175] 编码器在BV估计(对于帧内BC预测)或运动估计期间在标识当前块的分区时可使用从上到下方法。当前块可以是PU的PB, 或其它类型的块。图13是示出用于对经帧内BC预测的块进行分区的从上到下的方法的流程图和附图。图像编码器或视频编码器(诸如参考图3或图5a-5b描述的编码器)可使用该方法(1300)。在图13中所示的示例中, 编码器标识具有非对称分区的经帧内BC预测的 $2N \times 2N$ 块的至少一些分区。

[0176] 编码器检查(1310)具有 $2N$ 尺寸的模式。例如, 编码器检查单一 $2N \times 2N$ 块的模式、两个 $2N \times N$ 块的模式、两个 $N \times 2N$ 块的模式、一个 $N/2 \times 2N$ 块和一个 $3N/2 \times 2N$ 块的模式(图

13中示出的两个选项,在左侧或右侧具有窄块)、以及一个 $2N \times N/2$ 块和一个 $2N \times 3N/2$ 块的模式(在图13中示出的两个选项,在顶部或底部具有较短的块)。例如,对于 $16 \times 16$ CU,编码器检查具有大小 $16 \times P$ 的所有PU和具有大小 $P \times 16$ 的所有PU,其中在所允许的组合中,P可以是4、8、12和16。

[0177] 对于当前 $2N \times 2N$ 块的给定模式,编码器根据该模式确定当前 $2N \times 2N$ 块的(一个或多个)块的(一个或多个)BV值。对于当前 $2N \times 2N$ 块的一块,编码器可例如基于(1)当前图片中的(一个或多个)相邻块所使用的(一个或多个)BV值,(2)前一图片的并列块所使用的BV值,或(3)较早运动估计中针对当前 $2N \times 2N$ 块的该块所标识的MV值来选择起始BV值编码器随后针对该当前 $2N \times 2N$ 块的该块寻找适当的BV值。

[0178] 编码器选择(1320)具有 $2N$ 尺寸的最佳模式。选择准则可以是失真成本、比特率成本或失真成本和比特率成本的某种组合,或选择准则可使用某种其它度量(例如,使用方差阈值或边缘检测器)。例如,在图13中,编码器选择具有左 $N/2 \times 2N$ 块和右 $3N/2 \times 2N$ 块的模式。

[0179] 在将 $2N \times 2N$ 块作为二叉树划分之后,编码器还检查(1330)该 $2N \times 2N$ 块的每 $N \times N$ 块的模式。例如,对于给定 $N \times N$ 块,编码器检查单一 $N \times N$ 块的模式、两个 $N \times N/2$ 块的模式、两个 $N/2 \times N$ 块的模式、以及四个 $N/2 \times N/2$ 块的模式。编码器可独立地检查每个 $N \times N$ 块。对于例如 $8 \times 8$ U,编码器检查 $8 \times 8$ PU、两个 $8 \times 4$ PU、两个 $4 \times 8$ PU、以及四个 $4 \times 4$ PU。

[0180] 对于给定 $N \times N$ 块的给定模式,编码器根据该模式确定该给定 $N \times N$ 块的(一个或多个)块的(一个或多个)BV值。编码器可例如基于针对该当前 $2N \times 2N$ 块标识的BV值来选择起始BV值。编码器随后根据该模式针对该给定 $N \times N$ 块的(一个或多个)块寻找适当的(一个或多个)BV值。

[0181] 编码器选择(1340)各个 $N \times N$ 块的最佳模式组合。选择准则可以是失真成本、比特率成本或失真成本和比特率成本的某种组合,或选择准则可使用某种其它度量(例如,使用方差阈值或边缘检测器)。如图13中所示,不同 $N \times N$ 块可具有相同模式或不同模式。

[0182] 对于 $2N \times 2N$ 块,编码器随后在具有 $2N$ 尺寸的最佳模式和各个 $N \times N$ 块的最佳模式组合之间进行选择(1350)。选择准则可以是失真成本、比特率成本或失真成本和比特率成本的某种组合,或选择准则可使用某种其它度量(例如,使用方差阈值或边缘检测器)。

[0183] 2. 用于标识分区的从下到上的方法

[0184] 取代使用从上到下方法,编码器在BV估计(对于帧内BC预测)或运动估计期间在标识当前块的分区时可使用从下到上方法。图14是示出用于分区的从下到上的方法的一般化技术的流程图。图像编码器或视频编码器(诸如参考图3或图5a-5b描述的编码器)可执行技术(1400)。

[0185] 编码器对图像或视频进行编码(1410)以产生经编码数据。作为编码的一部分(例如,作为运动估计的一部分或者作为用于帧内BC预测的块向量估计的一部分),编码器计算当前图片的当前块的预测。当前块可以是PU的PB,或其它类型的块。在计算预测时,编码器使用从下到上的方法来标识当前块的分区。编码器可不对称地分区当前块,以使得当前块的分区是具有不同尺寸的两个分区,或对称地分区当前块,以使得当前块的分区具有等同的尺寸。

[0186] 一般而言,对于从下到上方法,对于当前块,编码器检查当前块内的更小块至少

一些分区模式。编码器高速缓存更小块的结果,诸如BV值或MV值。编码器随后使用(来自更小块的)被高速缓存的结果来检查当前块的至少一些分区模式以减少检查当前块的分区模式的计算复杂度。例如,编码器使用来自更小块的被高速缓存的结果来标识当前块的分区模式的起始BV值(在BV估计期间)或MV值(在运动估计期间)。在许多情况下,起始BV值(或MV值)将被用于当前块,这显著缩短了搜索过程。图15和16示出了用于标识当前块的分区的示例从下到上方法。替换地,编码器使用另一从下到上方法。

[0187] 编码器随后将经编码的数据作为比特流的一部分来输出(1420)。

[0188] 图15是示出用于对帧内BC预测的块进行分区的从下到上方法(1500)的流程图和附图。图像编码器或视频编码器(诸如参考图3或图5a-5b描述的编码器)可使用该方法(1500)。

[0189] 对于被划分为四叉树的 $2N \times 2N$ 块,编码器检查(1510)该 $2N \times 2N$ 块的每 $N \times N$ 块的模式。例如,对于给定 $N \times N$ 块,编码器检查单一 $N \times N$ 块的模式、两个 $N \times N/2$ 块的模式、两个 $N/2 \times N$ 块的模式、以及四个 $N/2 \times N/2$ 块的模式。编码器可独立地检查每个 $N \times N$ 块。对于例如 $8 \times 8$ U,编码器检查 $8 \times 8$ PU、两个 $8 \times 4$ PU、两个 $4 \times 8$ PU、以及四个 $4 \times 4$ PU的模式。

[0190] 对于给定 $N \times N$ 块的给定模式,编码器根据该模式确定该给定 $N \times N$ 块的(一个或多个)块的(一个或多个)BV值。对于给定 $N \times N$ 块的一块,编码器可例如基于(1)当前图片中的(一个或多个)相邻块所使用的(一个或多个)BV值,(2)前一图片的并列块所使用的BV值,或(3)较早运动估计中针对给定 $N \times N$ 块的该块所标识的MV值来选择起始BV值编码器随后针对该给定 $N \times N$ 块的该块寻找适当的BV值。

[0191] 编码器选择(1520)各个 $N \times N$ 块的最佳模式组合。选择准则可以是失真成本、比特率成本或失真成本和比特率成本的某种组合,或选择准则可使用某种其它度量(例如,使用方差阈值或边缘检测器)。如图15中所示,不同 $N \times N$ 块可具有相同模式或不同模式。

[0192] 编码器高速缓存(1530)向量值、分区模式信息、和/或检查(1510) $2N \times 2N$ 块的每 $N \times N$ 块的模式的结果。例如,编码器高速缓存BV估计期间的BV值,以及针对 $2N \times 2N$ 块的 $N \times N$ 块的分区模式选择。

[0193] 编码器使用被高速缓存的信息来检查(1540)具有 $2N$ 尺寸的模式,以通过在许多情况下缩短搜索过程来减少计算复杂度。编码器可使用被高速缓存的结果来标识 $2N \times 2N$ 块的起始BV值(在BV估计期间)。在许多情况下,起始BV值将被用于 $2N \times 2N$ 块的 $2N$ 尺寸分区模式,其显著缩短搜索过程。

[0194] 例如,编码器检查单一 $2N \times 2N$ 块的模式、两个 $2N \times N$ 块的模式、两个 $N \times 2N$ 块的模式、一个 $N/2 \times 2N$ 块和一个 $3N/2 \times 2N$ 块的模式(图15中示出的两个选项,在左侧或右侧具有窄块)、以及一个 $2N \times N/2$ 块和一个 $2N \times 3N/2$ 块的模式(在图15中示出的两个选项,在顶部或底部具有较短的块)。例如,对于 $16 \times 16$ CU,编码器重用在检查该 $16 \times 16$ CU的四个 $8 \times 8$ CU的分区模式之后高速缓存的信息。在许多情况下, $16 \times 16$ CU的分区模式终止于使用来自 $8 \times 8$ CU的被高速缓存的BV值,这显著缩短了搜索过程。

[0195] 编码器选择(1550)具有 $2N$ 尺寸的最佳模式。选择准则可以是失真成本、比特率成本或失真成本和比特率成本的某种组合,或选择准则可使用某种其它度量(例如,使用方差阈值或边缘检测器)。例如,在图15中,编码器选择具有左 $N/2 \times 2N$ 块和右 $3N/2 \times 2N$ 块的模式。

[0196] 对于 $2N \times 2N$ 块,编码器随后在具有 $2N$ 尺寸的最佳模式和各个 $N \times N$ 块的模式组合之间进行选择(1560)。选择准则可以是失真成本、比特率成本或失真成本和比特率成本的某种组合,或选择准则可使用某种其它度量(例如,使用方差阈值或边缘检测器)。

[0197] 参考图15描述的技术(1500)也可被用于经帧间编码的块的从下到上分区。在此情况下,来自 $N \times N$ 块的MV值和其它结果被高速缓存。编码器可使用被高速缓存的结果来标识 $2N \times 2N$ 块的起始MV值(在运动估计期间)。在许多情况下,起始MV值将被用于 $2N \times 2N$ 块的 $2N$ 尺寸分区模式,其显著缩短搜索过程。

[0198] 图16是示出用于对经帧内BC预测的块进行分区的甚至更快的从下到上的方法(1600)的流程图和附图。图像编码器或视频编码器(诸如参考图3或图5a-5b描述的编码器)可使用这些方法(1600)之一。图16中的方法(1600)类似于图15中的方法(1500),但是已在若干地方被修改以进一步缩短搜索过程。

[0199] 对于被划分为二叉树的 $2N \times 2N$ 块,编码器检查(1610)该 $2N \times 2N$ 块的每 $N \times N$ 块的模式子集。即,编码器检查每 $N \times N$ 块的部分而非全部模式。例如,对于给定 $N \times N$ 块,编码器仅检查具有两个 $N \times N/2$ 块的模式。或者,作为另一示例,编码器仅检查具有两个 $N/2 \times N$ 块的模式。编码器可独立地检查每个 $N \times N$ 块。然而,通过检查更少的模式,搜索过程被缩短。替换地,编码器检查每 $N \times N$ 块的其它和/或附加模式。

[0200] 如果多个模式被检查(1610),则编码器选择(1620)各个 $N \times N$ 块的最佳模式组合。选择准则可以是失真成本、比特率成本或失真成本和比特率成本的某种组合,或选择准则可使用某种其它度量(例如,使用方差阈值或边缘检测器)。如果仅单一模式被检查(1610),则编码器简单地每 $N \times N$ 块使用该模式。

[0201] 编码器高速缓存(1630)向量值、分区模式信息、和/或检查(1610) $2N \times 2N$ 块的每 $N \times N$ 块的模式其它结果。例如,编码器高速缓存BV估计期间的BV值,以及针对 $2N \times 2N$ 块的 $N \times N$ 块的分区模式选择。

[0202] 编码器使用被高速缓存的信息来检查(1640)具有 $2N$ 尺寸的模式子集以减少计算复杂度。例如,如果编码器仅检查(1610) $N \times N$ 块的 $N \times N/2$ 块,则编码器检查单一 $2N \times 2N$ 块的模式、两个 $2N \times N$ 块的模式、以及一个 $2N \times N/2$ 块和一个 $2N \times 3N/2$ 块的模式(在图16中示出的两个选项,其中在顶部或底部具有较短的块)。或者,如果编码器仅检查(1610) $N \times N$ 块的 $N/2 \times N$ 块,则编码器检查单一 $2N \times 2N$ 块的模式、两个 $N \times 2N$ 块的模式、以及一个 $N/2 \times 2N$ 块和一个 $3N/2 \times 2N$ 块的模式(在图16中示出的两个选项,其中在左侧或右侧具有窄块)。

[0203] 或者,作为另一示例(在图16中未示出),编码器仅检查(1610) $2N \times 2N$ 块的每 $N \times N$ 块的 $N \times N$ 分区。如果 $2N \times 2N$ 块是经帧内编码的块,则编码器检查(1640)单一 $2N \times 2N$ 块的模式、两个 $N \times 2N$ 块的模式、以及一个 $N/2 \times 2N$ 块和一个 $3N/2 \times 2N$ 块的模式(在左侧或右侧具有窄块)。如果 $2N \times 2N$ 块是经帧间编码的块,则编码器检查(1640)单一 $2N \times 2N$ 块的模式、两个 $2N \times N$ 块的模式、以及一个 $2N \times N/2$ 块和一个 $2N \times 3N/2$ 块的模式(在顶部或底部具有更短的块)。

[0204] 编码器选择(1650)具有 $2N$ 尺寸的最佳模式。选择准则可以是失真成本、比特率成本或失真成本和比特率成本的某种组合,或选择准则可使用某种其它度量(例如,使用方差阈值或边缘检测器)。对于 $2N \times 2N$ 块,编码器随后在具有 $2N$ 尺寸的最佳模式和各个 $N \times N$ 块

的模式组合之间进行选择(1660)。选择准则可以是失真成本、比特率成本或失真成本和比特率成本的某种组合,或选择准则可使用某种其它度量(例如,使用方差阈值或边缘检测器)。

[0205] 参考图16描述的技术(1600)也可被用于经帧间编码的块的从下到上分区。在此情况下,来自 $N \times N$ 块的MV值和其它结果被高速缓存。编码器可使用被高速缓存的结果来标识 $2N \times 2N$ 块的起始MV值(在运动估计期间)。在许多情况下,起始MV值将被用于 $2N \times 2N$ 块的 $2N$ 维分区模式,其显著缩短搜索过程。

[0206] 在图15和16中示出的方法(1500、1600)中,编码器可限制被高速缓存的BV值或MV值的数量。例如,编码器仅存储 $2N \times 2N$ 块的每 $N \times N$ 块的最佳分区模式的(一个或多个)BV值或MV值。替换地,编码器还存储其它BV值或MV值。编码器约束可控制编码器存储多少BV值或MV值。

[0207] IX. 搜索图案

[0208] 本章节呈现可在运动估计或帧内BC预测中使用的各种搜索图案。特别是,搜索图案利用屏幕捕捉视频或其它人造视频中的共同的运动类型。在这种视频中,块的运动通常是纯水平运动或纯垂直运动(例如,来自计算机桌面环境中的网页内容或应用内容的滚动,或来自混合内容视频中的跑马灯图形(ticker graphic)的滚动)。在此意义上,搜索图案被适配于屏幕捕捉视频或其它人造视频,但是它们也在编码自然视频时被使用。

[0209] 图17是示出用于使用小邻域中的位置的迭代估计和更大邻域中的该位置的迭代确认来搜索块的BV值或MV值的一般化技术(1700)的流程图。图像编码器或视频编码器(诸如参考图3或图5a-5b描述的编码器)可执行该技术(1700)。

[0210] 编码器对图像或视频进行编码以产生经编码的数据,并将该经编码的数据作为比特流的一部分来输出。在编码期间,编码器计算当前图片的当前块的预测(例如,使用BV估计或运动估计)。在计算预测时,编码器标识(1710)通过该预测的当前最佳位置周围的小邻域中的迭代估计来预测的当前最佳位置。例如,小邻域包括水平或垂直地紧邻当前最佳位置的位置。如果小邻域中的位置之一提供比当前最佳位置更好的结果,则用该小邻域中的该更好位置替换该当前最佳位置,且在该新位置周围的小邻域中检查该新位置。以此方式,使用小邻域图案的搜索过程可迭代地重复,直到当前最佳位置是其周围的小邻域中的最佳位置为止。这可在第一迭代中或在多次迭代之后发生。

[0211] 图18a和18b示出在搜索块的BV值或MV值时的小邻域中的位置的迭代估计。在图18a中,当前最佳位置被示出为黑色圆。菱形图案中的邻近位置(被示出为灰色圆)被评估。四个邻近位置的结果分别被与当前最佳位置的结果进行比较(在失真成本、比特率成本或失真成本和比特率成本的某种组合方面)。在图18a和18b的示例中,图18a中的当前最佳位置下方的位置提供了更好的结果并变成当前最佳位置。在下一次迭代中,如图18b中所示,新当前最佳位置周围的菱形图案中的邻近位置被评估。替换地,小邻域可具有另一种形状。

[0212] 编码器可使用阈值来限制标识(1710)阶段中的迭代次数。阈值取决于实现,且例如为4次迭代。如果达到阈值迭代次数,则编码器可执行另一搜索过程(例如,完整搜索过程或散列化过程)来确定该预测的最佳位置。

[0213] 返回图17,在标识小邻域中的当前最佳位置后(在阈值迭代次数内),编码器通过该预测的当前最佳位置周围的相继更大邻域中的迭代评估来确认(1720)该预测的当前最

佳位置。例如,每个更大邻域包括小邻域外的环中的位置。相继更大邻域可在每一侧按照一个样本、两个样本等递增增长。或者,相继更大邻域可按照某种因子来放大。例如,更大邻域的半径在第一次之后的每次迭代中按因子2放大。

[0214] 如果更大邻域中的位置之一提供比当前最佳位置更好的结果,则用该更大邻域中的该更好位置替换该当前最佳位置,且编码器在新当前最佳位置处重新开始该过程。否则(更大邻域中的位置均不提供比当前最佳位置更好的结果),编码器用下一更大邻域重复该确认过程,直到达到确认(1720)过程的阈值迭代次数。阈值取决于实现,且例如为对于相继更大邻域而言为4到7个阶段的检查中的任一者。

[0215] 在最大邻域被成功检查之后,编码器终止搜索过程。否则(如决策1730中所示),编码器在新当前最佳位置处重新开始该过程。

[0216] 图19示出在搜索块的BV值或MV值时在一个或多个更大邻域中的位置的迭代确认。在图19中,当前最佳位置被示出为黑色圆。编码器检查当前最佳位置周围的方形(内部的方形)的角和中点处的八个位置。这八个位置在图19中被示出为灰色圆。这八个位置的结果分别被与当前最佳位置的结果进行比较(在失真成本、比特率成本或失真成本和比特率成本的某种组合方面)。如果内部的方形中的这八个位置均不比当前最佳位置更好,则编码器检查更大的方形(图19中的外部的方形)的八个位置(角和中点)。替换地,更大邻域可具有另一形状(例如,位置的圆圈),包括更少被评估的位置(例如,仅角)或包括更多被评估的位置(例如,在形状外围处的位置)。

[0217] 第一阈值的计数器(小阈值迭代)可在每次标识(1710)阶段被开始或重新开始时被重置。或者,编码器可仅在该过程(1700)开始时将第一阈值的计数器重置一次。类似地,第二阈值的计数器(较大阈值迭代)可在每次确认(1720)阶段被开始或重新开始时被重置。或者,编码器可仅在该过程(1700)开始时将第二阈值的计数器重置一次。如果在过程(1700)内重置计数器,则编码器可使用另一约束来限制搜索过程持续多久,从而确保其在合理时间量内终止。

[0218] 图20是示出用于使用小邻域中的位置的迭代估计和更大邻域中的该位置的迭代确认来搜索块的BV值或MV值的示例技术的流程图。

[0219] 编码器重置(2010)第一和第二计数器。编码器设置(2020)当前最佳位置并评估该当前最佳位置(如果该当前最佳位置的结果尚未从先前评估可用)。编码器随后评估(2030)该当前最佳位置周围的小邻域中的邻近位置(如果邻近位置的结果尚未从先前评估可用)。编码器检查(2040)新最佳位置是否被发现(通过比较结果)。如果是,则编码器使用第一计数器检查(2042)是否达到第一阈值。如果否,则编码器递增第一计数器(未示出),将当前最佳位置设置(2010)为新最佳位置(来自邻近位置中)并从此处继续。以此方式,编码器可迭代地检查当前最佳位置周围的邻近位置的小邻域。

[0220] 如果达到第一阈值,则编码器执行(2080)完整搜索。替换地,编码器使用另一搜索过程,诸如散列化过程。

[0221] 如果没有发现新最佳位置(在决策(2040)处),则编码器评估(2050)当前最佳位置周围的位置的环。编码器检查(2060)新最佳位置是否被发现(通过比较结果)。如果是,则编码器重置(2090)第一和第二计数器,将当前最佳位置设置(2010)为新最佳位置(来自邻近位置中)并从此处继续。以此方式,编码器重新开始该过程(2000)。

[0222] 否则(在决策(2060)处没有发现新最佳位置),编码器使用第二计数器检查(2062)是否达到第二阈值。如果是,则编码器(成功地)终止搜索过程。如果不是,则编码器递增第二计数器(未示出),扩展位置的环(2070)并评估(2050)当前最佳位置周围的位置的(经扩展)环。

[0223] 在本章节中描述的示例中,BV值和MV值指示整数样本偏移。替换地,BV值和/或MV值可指示分数样本偏移。当分数样本偏移被允许时,编码器可标识具有整数样本偏移的BV值或MV值,如参考图17-20描述的。随后,编码器可标识整数样本BV或MV值周围(例如,在距整数样本BV或MV值单一样本偏移内)的邻域中的BV值或MV值。

[0224] 替换地,当分数样本偏移被允许时,编码器可在参考图17-20描述的每个阶段处标识具有带有所允许的精度的分数样本偏移的BV值或MV值(即,标识小邻域中具有分数样本偏移的BV值或MV值、标识更大邻域中具有分数样本偏移的BV值或MV值等等)。

[0225] X. 约束针对帧内BC预测的BV搜索范围

[0226] 在一些示例实现中,编码器使用完整搜索范围以用于BV估计。经重构的样本值的整个区域被搜索以标识当前块的BV值。尽管使用完整搜索范围可帮助标识要在帧内BC预测中使用的最佳BV值,然而这也可能增加BV估计的复杂度。

[0227] 在其它示例实现中,编码器根据一个或多个约束来限制BV搜索范围。通过限制BV搜索范围,编码和解码期间为了帧内BC预测而进行的快速存储器存取所参考的经重构的样本值的区域可被缩小,这有助于降低实现成本。

[0228] 在本章节的示例中,编码器在BV估计期间标识BV值时考虑当前亮度块的亮度样本值。编码器尝试将当前块的亮度样本值匹配到先前亮度块的经重构的亮度样本值。然而,所得到的BV值也被应用到相应色度块的色度样本值。

[0229] 图21a示出了针对BV值的搜索范围的示例约束。除了当前帧(2110)的当前块(2130),图21a还示出了通过两个CTB(2120、2122)来限定的搜索范围。当前CTB(2120)是当前CTU的一部分并且包括当前块(2130)。在CTB(2122)在其左侧的情况下,当前CTB(2120)限定了能够在其中找到当前块(2130)的可允许BV的搜索范围。BV(2142、2144)参考区域位于该搜索范围之外,因此那些BV值(2142、2144)不被允许。

[0230] 在一些示例实现中,当前块的BV值的搜索范围是当前CTB和在其左侧的CTB。例如,CTB可具有64x64、32x32或16x16个样本值的大小,这产生128x64、64x32或32x16个样本值的搜索范围。只有在当前CTB和左侧CTB中的样本值被用于当前块的帧内BC预测。这通过对搜索过程进行约束而简化了编码器实现。还通过限制解码器在快速存储器中缓冲的用于帧内预测的样本值的数量而简化了解码器实现。(即使选择更小的CTU/CTB大小,解码器也具有足够的缓冲能力来存储最大可能尺寸的两个CTB的样本值)。另一约束是帧内预测不能参考来自另一片或小块的样本值。对于具有位于 $(x_0, y_0)$ 处的左上位置的当前的块 $m \times n$ 以及各自具有尺寸 $CTB_{sizeY} \times CTB_{sizeY}$ 的CTB,编码器可如下地为具有水平分量BV[0]和垂直分量BV[1]的二维BV检查这些约束。

[0231]  $BV[0] \geq -((x_0 \% CTB_{sizeY}) + CTB_{sizeY})$

[0232]  $BV[1] \geq -(y_0 \% CTB_{sizeY})$

[0233] 位于位置 $(x_0, y_0)$ 、 $(x_0 + BV[0], y_0 + BV[1])$ 和 $(x_0 + BV[0] + m - 1, y_0 + BV[1] + n - 1)$ 处的样本值应当在同一片中。

[0234] 位于位置  $(x_0, y_0)$ 、 $(x_0 + BV[0], y_0 + BV[1])$  和  $(x_0 + BV[0] + m - 1, y_0 + BV[1] + n - 1)$  处的样本值应当在同一小块中。

[0235] 在实践中,如果最佳BV值趋向于水平朝向或垂直朝向,则评估大的方形搜索范围(诸如 $S \times S$ 搜索范围,其中 $S$ 为 $CTB_{sizeY}$ )或大的矩形搜索范围(诸如 $2S \times S$ 搜索范围,其中 $S$ 为 $CTB_{sizeY}$ )中的候选BV值没有意义。相反,编码器可使用仍旧包括可能的候选BV值的较小搜索范围,其中较小BV搜索范围是水平定向(例如,具有尺寸 $2S \times 1/4S$ 或 $2S \times 3/4S$ )或垂直定向(例如,具有尺寸 $1/4S \times 2S$ 或 $3/4S \times 2S$ )的。在大多数场景中,编码器在BV估计期间检查更少的BV值,但仍旧能找到最适当的BV值。

[0236] 具有水平偏置(或水平定向)的BV搜索范围包括具有比垂直BV分量值的范围更大的水平BV分量的候选BV值。相反,具有垂直偏置(或垂直定向)的BV搜索范围包括具有比水平BV分量值的范围更大的垂直BV分量的候选BV值。

[0237] BV搜索范围也可经受其它约束。例如,BV搜索范围可经受以下约束:当前块的任何BV值参考位于当前CTB和/或当前CTB左侧的CTB内的区域。即,BV搜索范围适合放入当前CTB或其左侧的CTB内。或者,作为另一示例,BV搜索范围可经受以下约束:当前块的任何BV值参考位于当前CTB和/或当前CTB上方的CTB内的区域内。即,BV搜索范围适合放入当前CTB或其上方的CTB内。BV搜索范围也可被约束为适合放入当前图片内。在一些实现中,BV搜索范围被进一步约束为适合放入当前切片(slice)和/或当前小块(tile)内。

[0238] 除了在BV搜索范围的远边界(远离当前块)处的约束外,BV搜索范围还可在其近边界(靠近当前块)处被约束。例如,在一些实现中,对于具有大小 $2N \times 2N$ 的CU,PU分区可具有大小 $N \times N$ 和 $2N \times N$ 或 $N \times 2N$ ,其中每个PU具有其自己的BV,或某种其它分区大小。然而,PU的BV不被允许参考同一CU内的其它PU区域。对BV搜索范围的这一约束相对于允许PU级重叠在某种程度上降低了性能,但是允许每个PU具有其自己的BV并允许每个CU内的PU被并行地重构,这可促进高效解码。

[0239] 替换地,如果在CB和相应帧内预测区域之间的重叠被允许,则编码器可执行重叠处理。在此情况下,具有水平偏置或垂直偏置的BV搜索范围可延伸到当前CU中。

[0240] 图21b示出具有水平偏置(水平定向)的第一替换搜索范围(2181)。搜索范围(2181)具有最多 $2S \times 1/4S$ 的尺寸。搜索范围(2181)可被截断以适合放入当前CTB及其左侧的CTB内,如在搜索范围(2181)的交叉阴影和阴影部分中所示。或者,搜索范围(2181)可被进一步约束为不包括当前CTB内的任何部分(在图21b被示出为搜索范围(2181)的阴影部分)。

[0241] 图21c示出具有水平偏置的第二替换搜索范围(2182)。搜索范围(2182)具有最多 $2S \times 3/4S$ 的尺寸。搜索范围(2182)可被截断以适合放入当前CTB及其左侧的CTB内,如在搜索范围(2182)的交叉阴影和阴影部分中所示。或者,搜索范围(2182)可被进一步约束为不包括当前CTB内的任何部分(在图21c被示出为搜索范围(2182)的阴影部分)。

[0242] 图21b示出具有垂直偏置(垂直定向)的第三替换搜索范围(2183)。搜索范围(2183)具有最多 $1/4S \times 2S$ 的尺寸。搜索范围(2183)可被截断以适合放入当前CTB及其上方的CTB内,如在搜索范围(2183)的交叉阴影和阴影部分中所示。或者,搜索范围(2183)可被进一步约束为不包括当前CTB内的任何部分(在图21d中被示出为搜索范围(2183)的阴影部分)。



[0243] 图21e示出具有垂直偏置的第四替换搜索范围(2184)。搜索范围(2184)具有最多 $3/4S \times 2S$ 的尺寸。搜索范围(2184)可被截断以适合放入当前CTB及其上方的CTB内,如在搜索范围(2184)的交叉阴影和阴影部分中所示。或者,搜索范围(2184)可被进一步约束为不包括当前CTB内的任何部分(在图21d中被示出为搜索范围(2184)的阴影部分)。

[0244] 在BV估计期间,视频编码器或图像编码器可执行包括如下的BV估计的编码。

[0245] 编码器确定图片的当前块的BV。该BV指示到图片内的一区域的位移。当前块是具有大小S的当前CTB。作为确定BV的一部分,编码器检查该区域在具有水平偏置或垂直偏置的BV搜索范围内的约束。编码器使用BV来执行对于当前块的帧内BC预测。编码器还对BV进行编码。例如,编码器执行BC预测并编码BV,如在本申请中别处所述。

[0246] 更一般地,编码器使用帧内BC预测来编码图片的数据。编码包括使用具有水平或垂直偏置的BV搜索范围来执行BV估计操作。编码器输出该图片的经编码的数据。

[0247] BV搜索范围可具有水平偏置,其具有尺寸 $2S \times 1/4S$ 或 $2S \times 3/4S$ 。或者,更一般地的,水平偏置的BV搜索范围具有在S和2S之间(包括S和2S)的宽度,并具有在 $1/4S$ 和 $3/4S$ 之间的高度(包括 $1/4S$ 和 $3/4S$ )。或者,BV搜索范围可具有垂直偏置,其具有尺寸 $1/4S \times 2S$ 或 $3/4S \times 2S$ 。或者,更一般地的,垂直偏置的BV搜索范围具有在S和2S之间(包括S和2S)的高度,并具有在 $1/4S$ 和 $3/4S$ 之间的宽度(包括 $1/4S$ 和 $3/4S$ )。

[0248] 编码器可从多个可用BV搜索范围中选择BV搜索范围。例如,编码器从具有水平偏置的多个搜索范围(诸如 $2S \times 1/4S$ 和 $2S \times 3/4S$ 搜索范围)中选择。或者,编码器从具有垂直偏置的多个搜索范围(诸如 $1/4S \times 2S$ 和 $3/4S \times 2S$ 搜索范围)中选择。或者,编码器从各自具有水平偏置或垂直偏置的多个搜索范围中选择。

[0249] 编码器可至少部分基于一个或多个先前块的BV值来选择BV搜索范围。例如,(一个或多个)先前块在当前图片中。或者,(一个或多个)先前块在一个或多个先前图片中。或者,先前块在当前图片中并且在一个或多个先前图片中。通过考虑(一个或多个)先前块的(一个或多个)BV值,编码器可标识BV值中的趋势(例如,大多数BV值具有强水平BV分量但是具有小垂直BV分量或没有垂直BV分量)并选择适当的BV搜索范围。BV搜索范围的选择也可取决于其它因素(例如,用户设置)。

[0250] (一个或多个)先前块的(一个或多个)BV值可被跟踪,例如使用将(一个或多个)BV值组织为直方图的数据结构,其中不同类别(或“槽”)对应于不同BV值范围并且每个类别/槽存储一计数器。从而,直方图数据结构可提供关于不同BV值的使用频率的统计数据。或者,可按某一其他方式来跟踪(一个或多个)BV值。例如,编码器跟踪当前图片的块的BV值,随后评估当前块周围的邻域中的先前块的BV值来确定要使用哪个BV搜索范围。

[0251] 使用具有水平或垂直偏置的较小BV搜索范围与使用较大的 $S \times S$ 或 $2S \times S$ 搜索范围相比可能略微低效(在速率失真性能方面)。对于许多编码场景,BV估计的计算复杂度的减少使这种罚分有意义。

[0252] 图22示出用于采用帧内BC预测模式的编码的技术(2200),其中该帧内BC预测模式受制于针对BV值的选择的一个或多个约束。编码器(诸如参考图3或图5a-5b描述的编码器)可执行该技术(2200)。

[0253] 作为开始,编码器确定(2210)图片的当前块的BV。当前块可以是CB、PB或其他块。该BV指示到图片内的一区域的位移。在确定BV时,编码器检查一个或多个约束。

[0254] 根据一个可能的约束,编码器检查用于帧内BC预测的样本值的范围。编码器可检查候选帧内预测区域是否位于当前CTB和一个或多个其它CTB(例如,当前CTB左侧的CTB)所限定的范围内。例如,当BV具有第一分量 $BV[0]$ 和第二分量 $BV[1]$ ,当前块具有位置 $(x_0, y_0)$ 处的左上位置,并且每一个CTB具有宽度 $CTB_{width}$ 和高度 $CTB_{height}$ 时,如果 $BV[0] \geq -((x_0 \% CTB_{width}) + CTB_{width})$ 且 $BV[1] \geq -(y_0 \% CTB_{height})$ ,则满足约束。编码器可类似地检查在以下搜索范围内的 $BV[0]$ 和 $BV[1]$ 的值的上限: $BV[0] < (CTB_{width} - m - (x_0 \% CTB_{width}))$ 且 $BV[1] < (CTB_{height} - n - (y_0 \% CTB_{height}))$ 。替换地,该搜索范围包括更多或更少的CTB,或者该搜索范围以某种其它方式来定义。

[0255] 根据另一可能约束,编码器将搜索限制于当前片和小块(即,当前块和区域是图片的不大于1个片和图片的不大于1个小块的一部分)。编码器可检查当前块的左上位置、候选帧内预测区域的左上位置以及候选帧内预测区域的右下位置是否是单个片和单个小块的一部分。例如,如果 $(x_0, y_0)$ 、 $(x_0 + BV[0], y_0 + BV[1])$ 和 $(x_0 + BV[0] + m - 1, y_0 + BV[1] + n - 1)$ 是单个片和单个小块的一部分,则满足约束。

[0256] 替代地,编码器检查其它和/或附加约束。

[0257] 编码器使用BV来执行(2220)对于当前块的帧内BC预测。例如,编码器对整个当前块执行帧内BC预测。或者,编码器对与当前块相关联的多个块(例如,对于逐TB为基础的多个TB,其中TB与具有该BV的当前PB相关联)执行帧内BC预测。

[0258] 编码器对BV进行编码(2230)。编码器可为另一帧内BC预测模式块重复该技术(2200)。

[0259] 对于帧内BC预测,编码器和解码器使用重构的样本值。未重构的样本值可被呈现作为尚未被编码和重构的图片的部分。为了避免将未重构的样本值用于帧内BC预测,编码器可设置针对可允许的BV值的约束,使得只有真实的、之前被重构的样本值被用于根据BV进行的帧内BC预测。

[0260] 在一些示例实现中,编码器通过考虑当前块和包含候选帧内预测区域的右下位置的块的z-扫描次序来检查BV值。更具体地,编码器检查包含位置 $(x_0 + BV[0] + m - 1, y_0 + BV[1] + n - 1)$ 的块的z-扫描次序是否小于包含 $(x_0, y_0)$ 的块的z-扫描次序。如果是,则包含帧内预测区域的右下位置的块已经在先前被重构(并且因此具有该帧内预测区域的剩余部分)。BV还满足条件 $V[0] + m \leq 0$ 和 $BV[1] + n \leq 0$ 中的至少一个,确保帧内预测区域不覆盖当前块。

[0261] z-扫描次序遵循分隔图片的顺序指定的块排序。图23示出用于当前块(2330)和可能包括候选BV的帧内预测区域的右下位置的块的示例z-扫描次序(2300)。当前块(2330)可以是CB、PB或其他块。z-扫描次序通常被以在行中从左到右顺序分派给块,在从上到下的连续行中重复。当块被拆分时,可以在被拆分的块内递归地分派z-扫描次序。对于用于HEVC标准的编码/解码的实现,z-扫描次序通过CTB光栅扫描模式(在CTB行中从左到右,从上到下在连续CTB行中重复)从CTB到CTB开始行进。如果CTB被拆分,z-扫描次序遵循在经拆分的CTB内的四叉树的CB的光栅扫描模式。并且,如果CB被拆分(例如,拆成多个CB或拆成多个PB),z-扫描次序遵循在经拆分的CB内的块的光栅扫描模式。

[0262] 替代地,当帧内BC预测可在逐TB的基础上执行时,编码器和解码器可检查帧内预测区域和当前块(TB)之间的可能重叠,随后使用检查的结果来决定当前TB是否应当被拆分成更小的TB来应用帧内BC预测操作。假设当前TB具有 $m \times n$ 的大小,其中 $m$ 和 $n$ 可以彼此相等

或可具有不同值。如果 $BV[0] > -m$ 并且 $BV[1] > -n$ ,则帧内预测区域与当前的 $m \times n$  TB重叠,这会产生问题,除非该当前的 $m \times n$  TB被拆分成更小的TB来应用帧内BC预测操作。因此,如果 $BV[0] > -m$ 并且 $BV[1] > -n$ ,则编码器和解码器将当前TB拆分成更小的TB。对于更小的TB检查同一条件(例如,递归地检查),如果对于拆分后的更小的 $m$ 和 $n$ 值,还是 $BV[0] > -m$ 且 $BV[1] > -n$ ,则这些更小的TB可被进一步拆分。

[0263] 例如,假设PB的BV是 $(-9, -5)$ ,并且当前TB是 $32 \times 32$ 的块。编码器和解码器确定 $-9 > -32$ 且 $-5 > -32$ ,这表示帧内预测区域(其左上角位移 $-9, -5$ )将与当前的 $32 \times 32$  TB重叠。编码器和解码器将该 $32 \times 32$  TB拆分成四个 $16 \times 16$  TB。对于每一个 $16 \times 16$  TB,编码器和解码器确定 $-9 > -16$ 且 $-5 > -16$ ,这表示帧内预测区域(其左上角位移 $-9, -5$ )将与当前的 $16 \times 16$  TB重叠。编码器和解码器继续将每个 $16 \times 16$  TB拆分成四个 $8 \times 8$  TB。对于 $8 \times 8$  TB, $(-9, -5)$ 的BV不是问题,因此 $8 \times 8$  TB不被进一步执行拆分。

[0264] 在这一场景中,当TB由于BV值和TB的大小的缘故而被拆分时,编码器可省略对于标志值的信令,该标志值原本会指示当前TB是否要拆分成更小的TB。编码数据的比特流缺少指导解码器将当前TB拆分成更小TB的标志值。替代地,解码器可根据BV值和TB的大小来推导TB应当被拆分。这可节省原先被花费在指示关于拆分TB的信息的比特。

[0265] 作为BV估计的一部分,编码器可使用若干方法中的任一方法。编码器可使用全面搜索,从而评估搜索范围中允许的每一候选BV值。或者,编码器可使用部分搜索,从而评估搜索范围中允许的仅一些候选BV值。例如,编码器可在当前块的预测的BV值(例如,基于一个或多个相邻块的BV值来预测的)处开始部分搜索。在为部分搜索评估了在开始位置处的候选BV值后,编码器可(例如,根据螺旋搜索模式或某一其他模式)评估在距开始位置递增的距离处的一个或多个其他候选BV值。或者,编码器可使用如在前节中描述的搜索图案。在评估给定候选BV值时,编码器可比较帧内预测区域中的所有样本值和当前块。或者,编码器可评估样本值的子集(即,值被评估的子样本)。在比较帧内预测区域和当前块之间的样本值以确定失真成本时,编码器可计算均方差、差方和(“SSD”)、绝对差之和(“SAD”)或某一其他失真测量。编码器还可确定与对候选BV值的编码相关联的比特率成本。

[0266] XI. 替换和变型

[0267] 在在此所述的许多示例中,在分开的组件或过程中实现帧内BC预测和运动补偿,并且在分开的组件或过程中实现BV估计和运动估计。或者,帧内BC预测可以被实现为运动补偿的特定情况,而BV估计可以被实现为运动估计的特定情况,所述情况中当前图片被用作参考图片。在这样的实现中,BV值可以被信号通知为MV值,但用于(在当前图片内)的帧内BC预测,而不是图片间预测。作为术语被用在此处时,“帧内BC预测”指示在当前图片内的预测,无论该预测是使用图片内预测模块、运动补偿模块、或某种其它模块提供的。类似地,可以使用MV值或使用不同类型的参数或语法元素来表示BV值,并且可以使用图片内预测模块、运动补偿模块、或某种其它模块提供BV估计。

[0268] 鉴于可应用所公开的本发明的原理的许多可能的实施例,应当认识到,所示实施例仅是本发明的优选示例,并且不应认为是限制本发明的范围。相反,本发明的范围由后续的权利要求来界定。

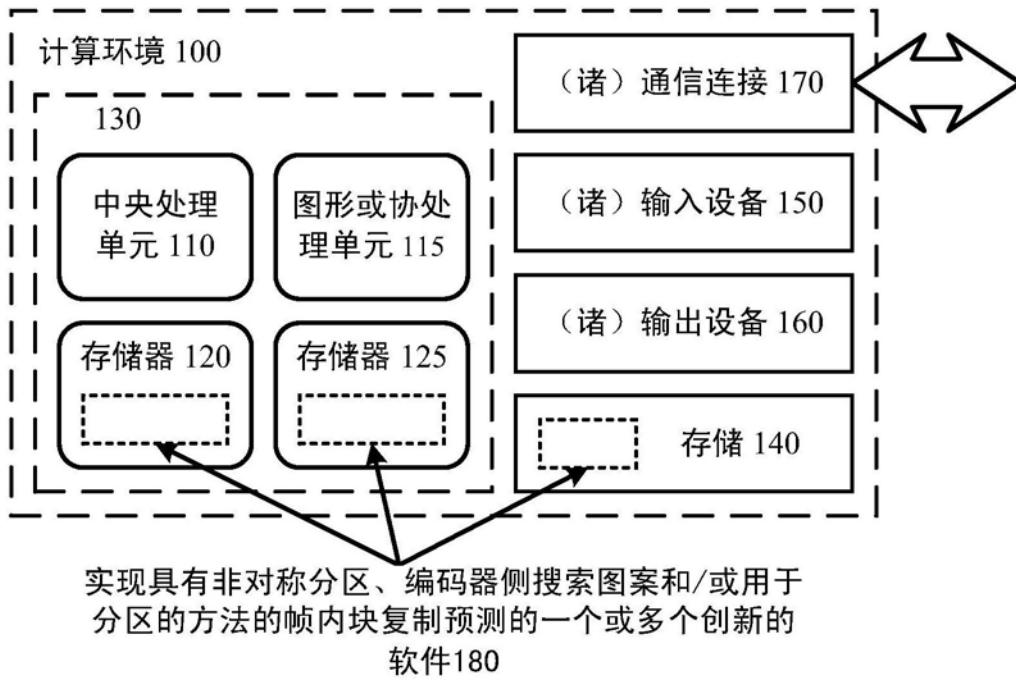


图1

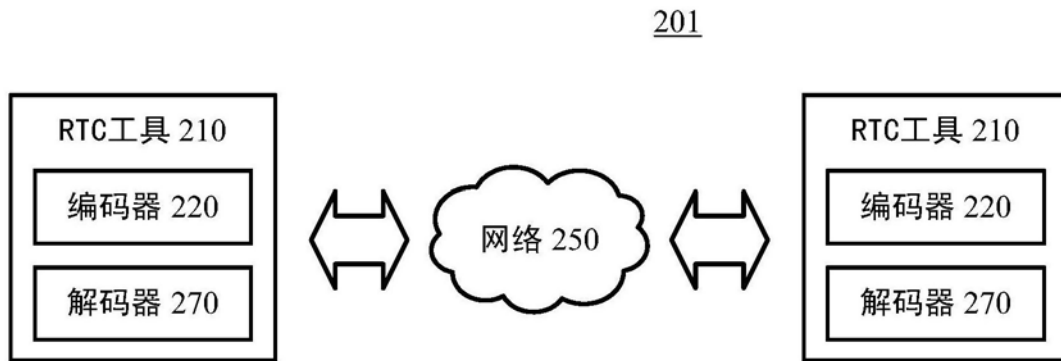


图2a

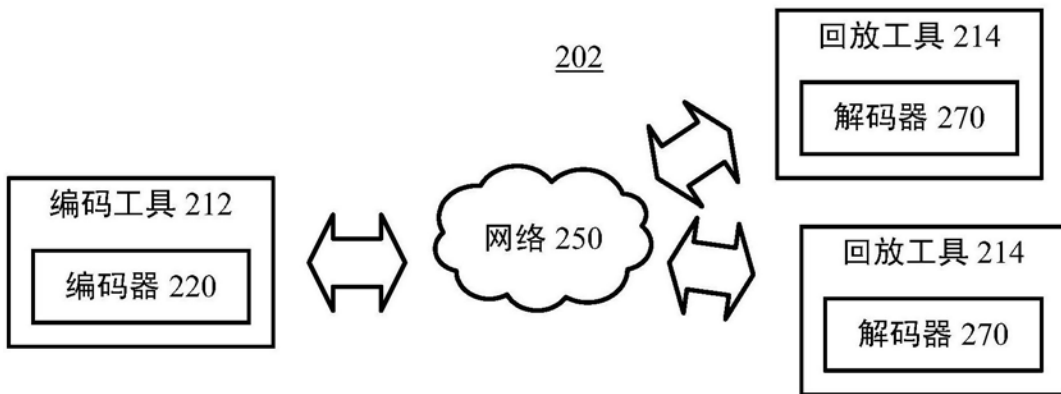


图2b

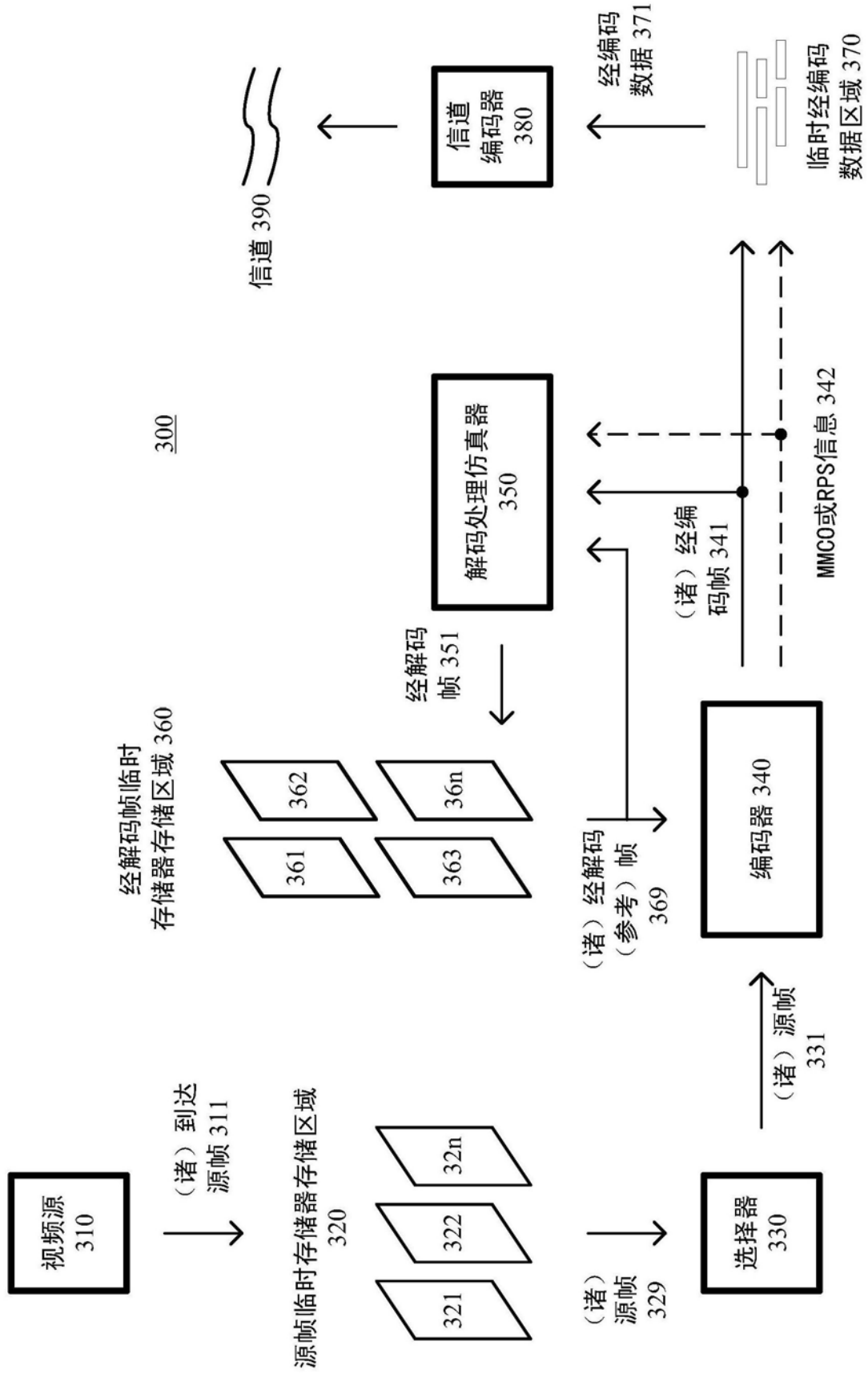


图3

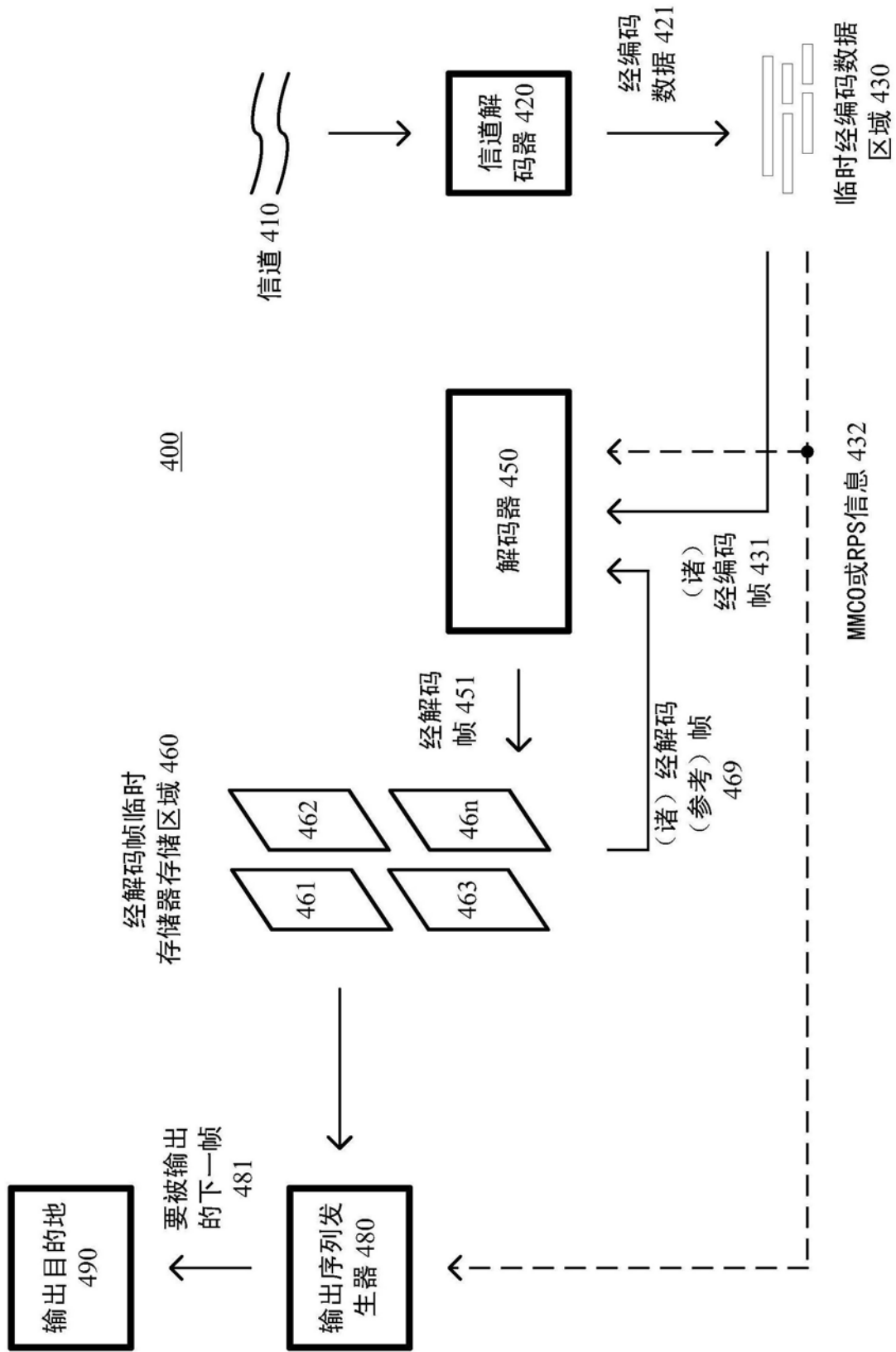


图4

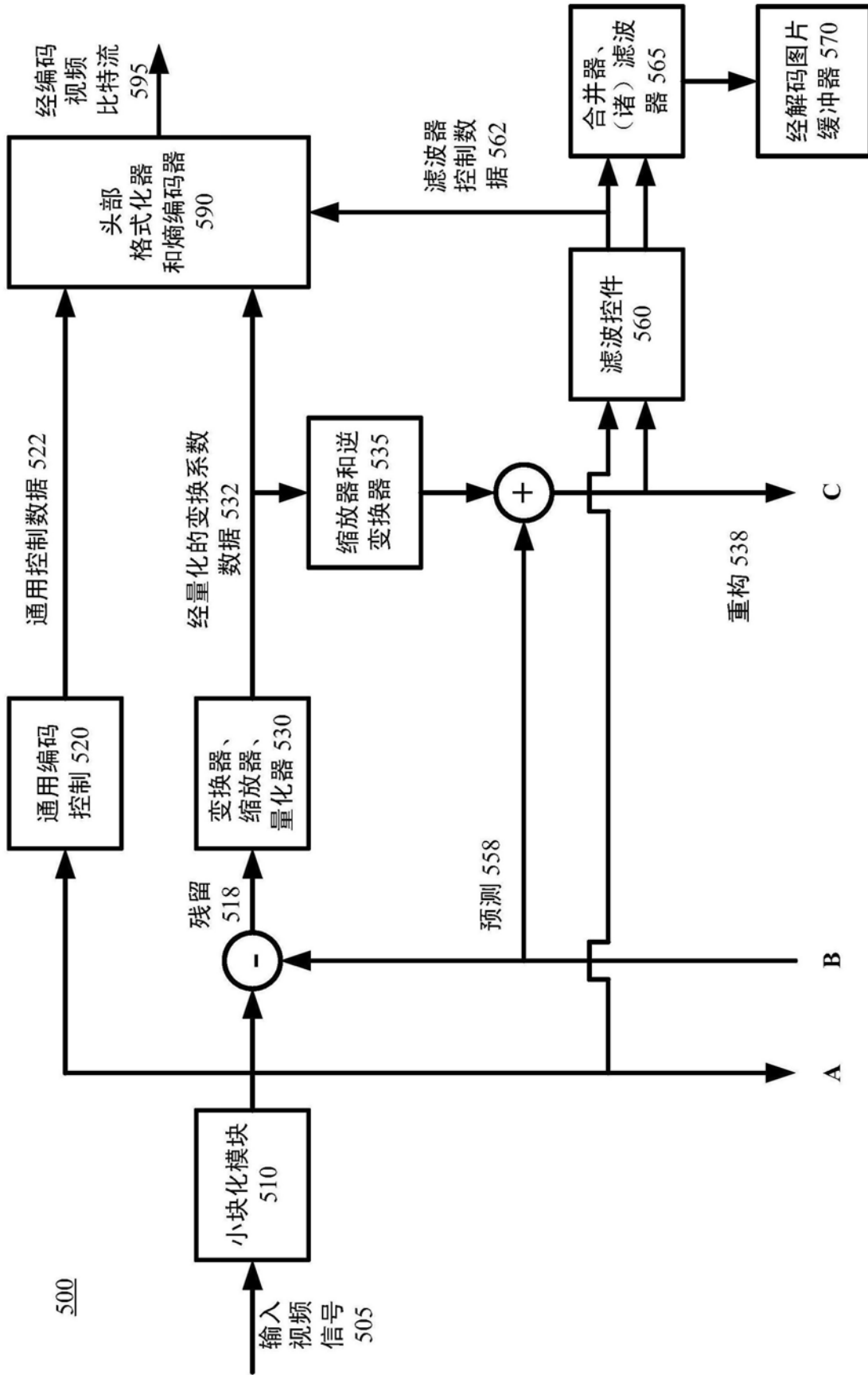
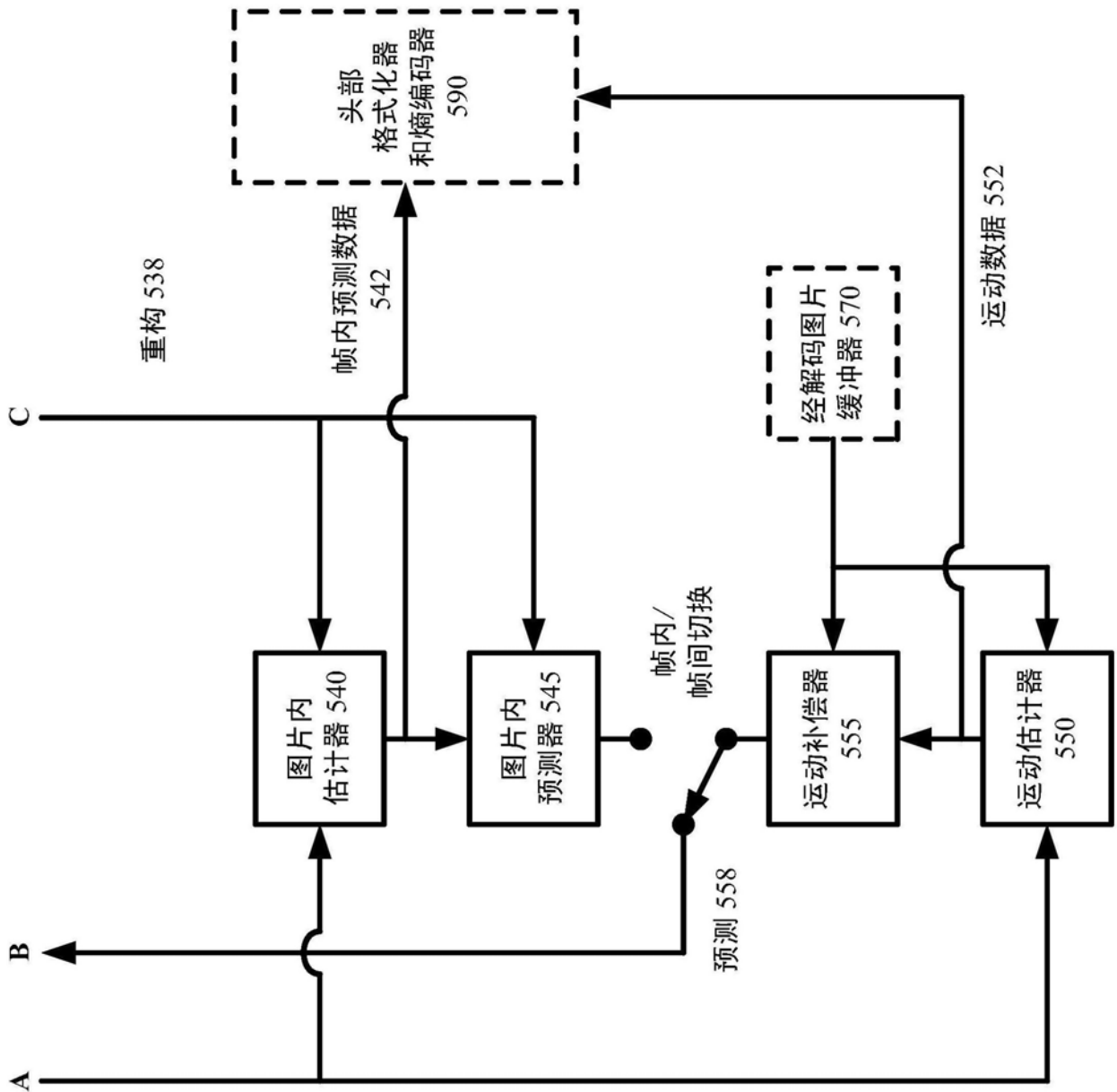


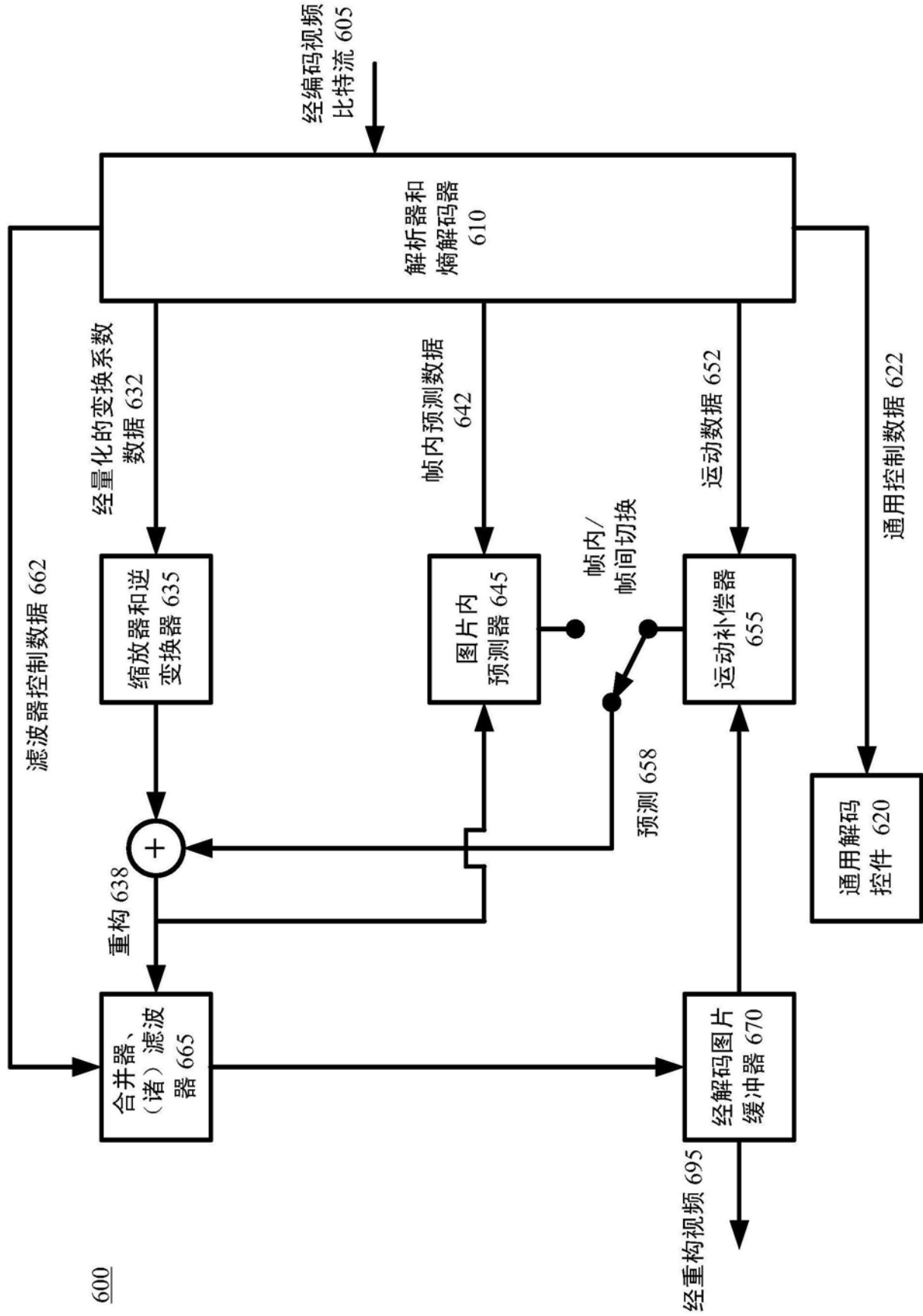
图5a



500

图5b

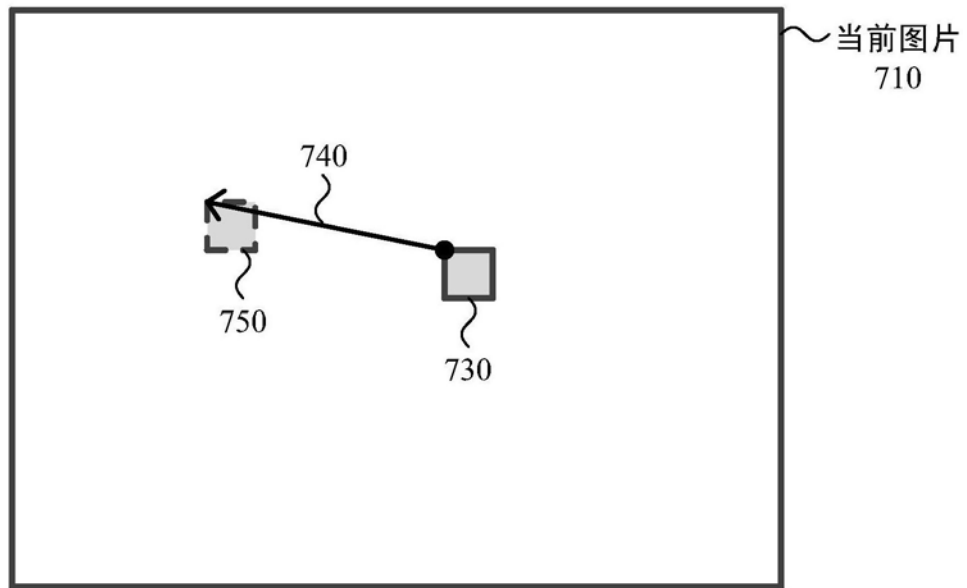




600

图6

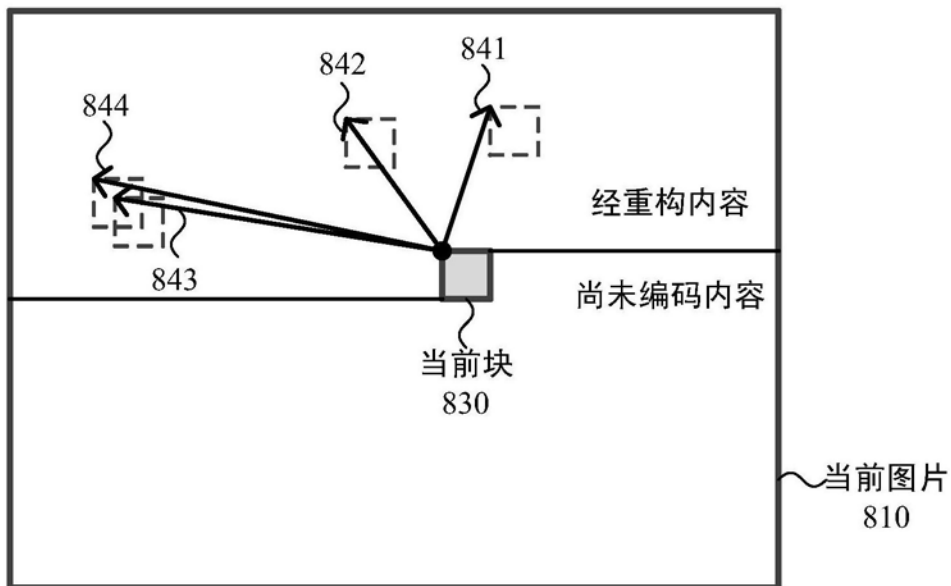
700



当前图片 (710) 的当前块 (730) 的块向量 (740)，指示当前图片 (710) 中的参考区域 (750) 的位移

图7

800



指示当前图片 (810) 的当前块 (830) 的候选块的位移的块向量 (841、842、843、844)

图8

900

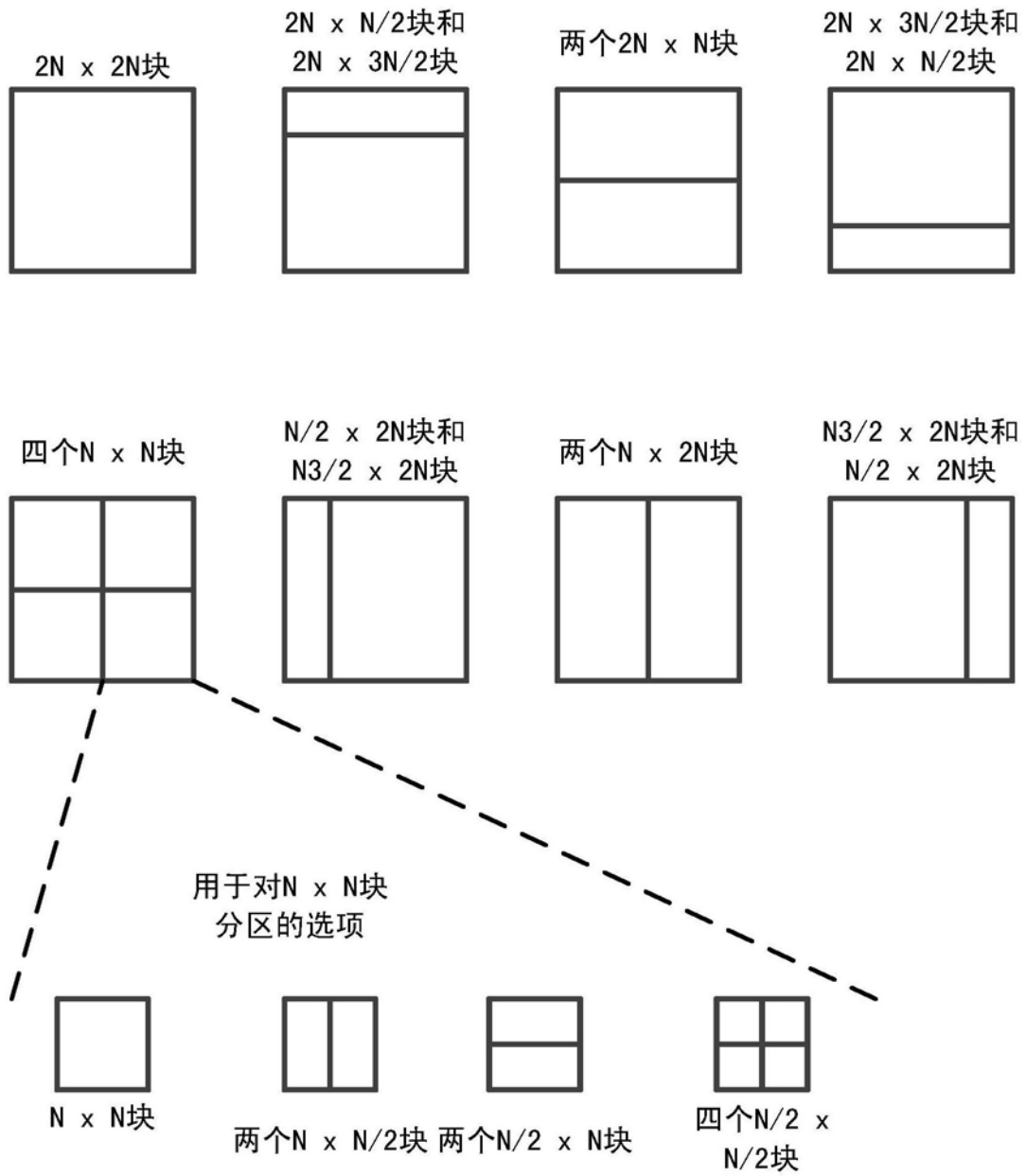


图9

1000

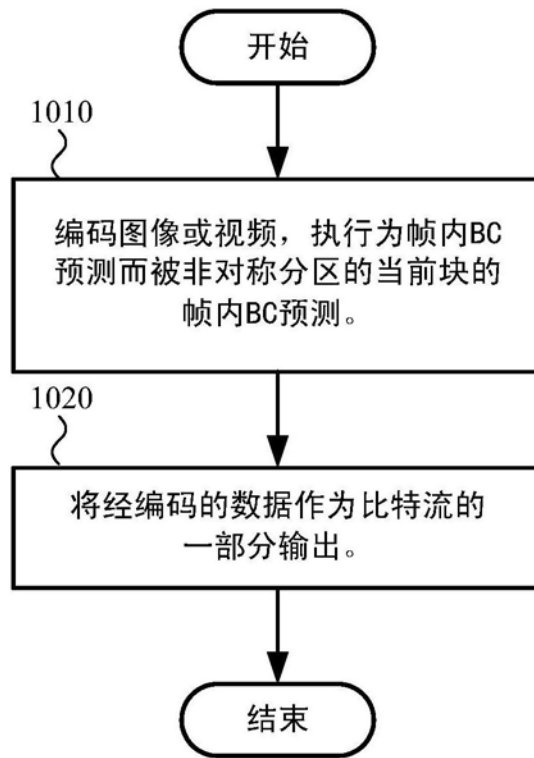


图10

1100

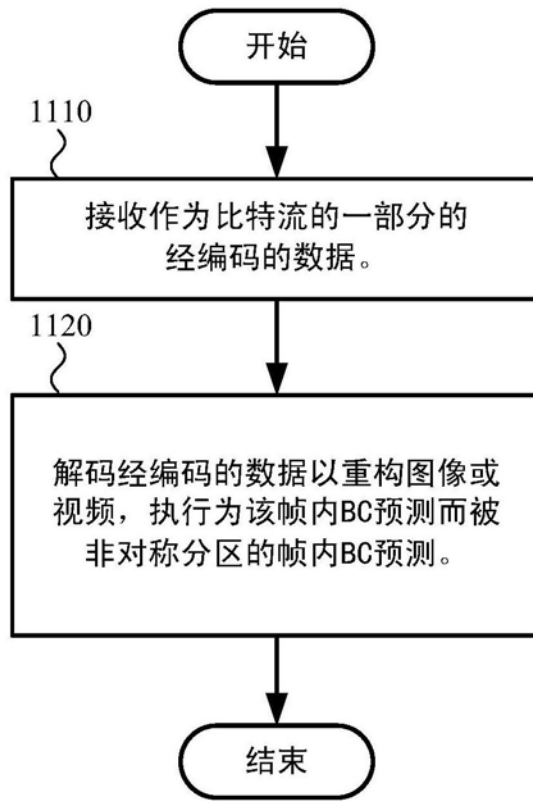
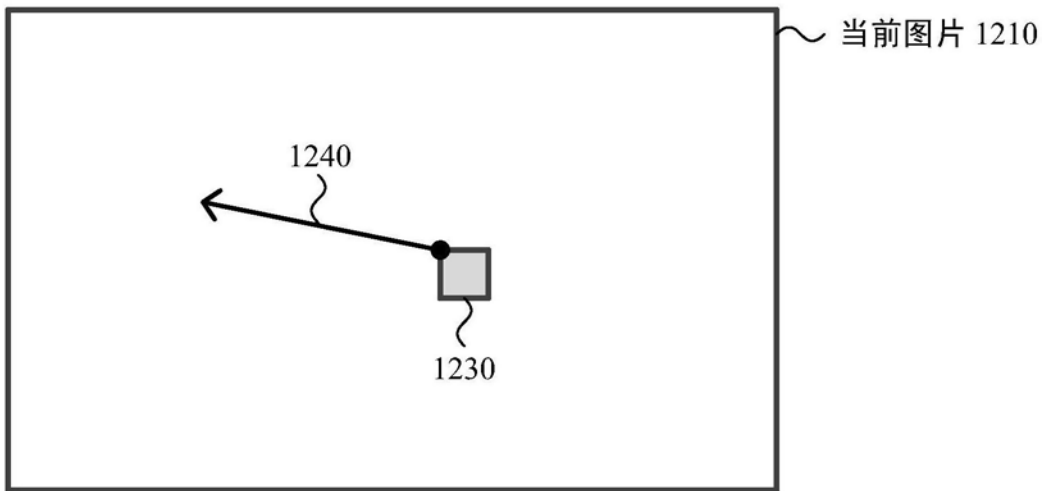
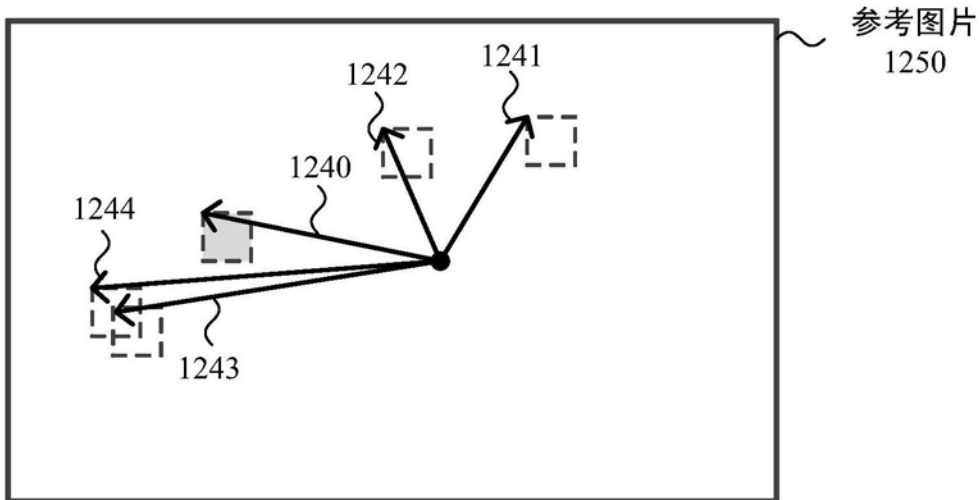


图11

1200



当前图片 (1210) 的当前块 (1230) 的MV值 (1240), 其指示参考图片 (1250) 的参考块的位移



指示参考图片 (1250) 中的候选块相对于当前块 (1230) 的左上角的坐标的位移的MV值 (1241、1242、1243、1244)

图12

用于分区具有分对称分区的2Nx2N帧内BC预测块的  
从上到下方法

1300

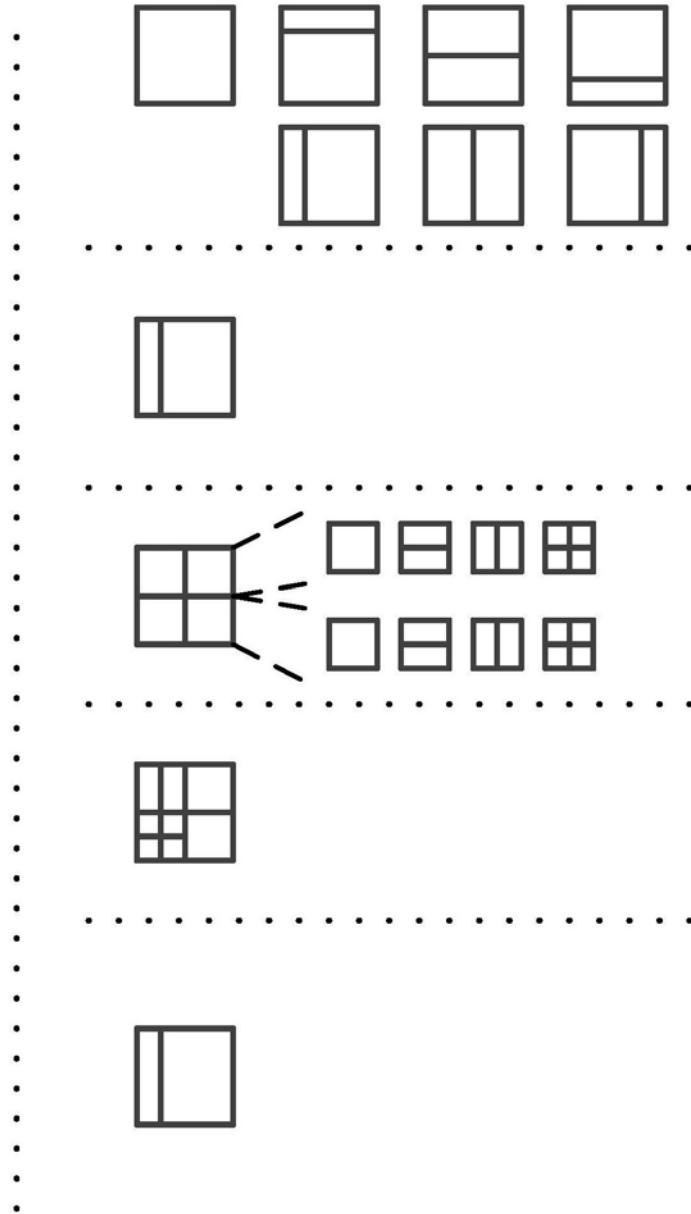
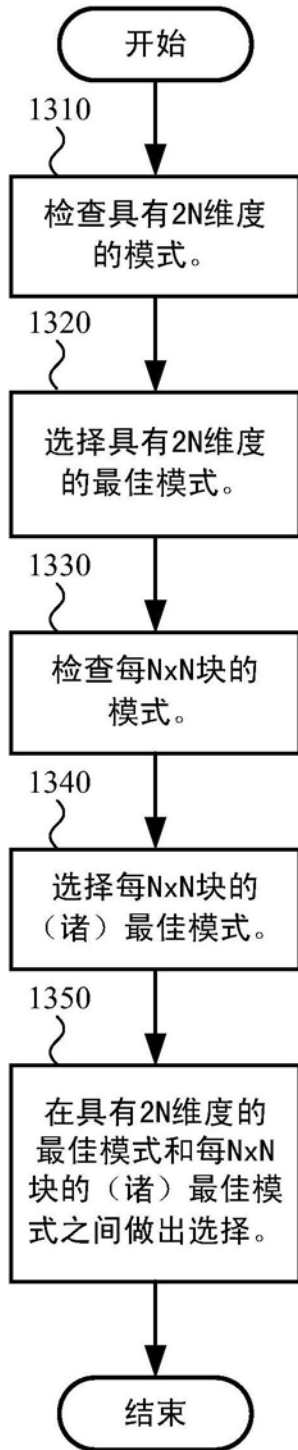


图13

1400

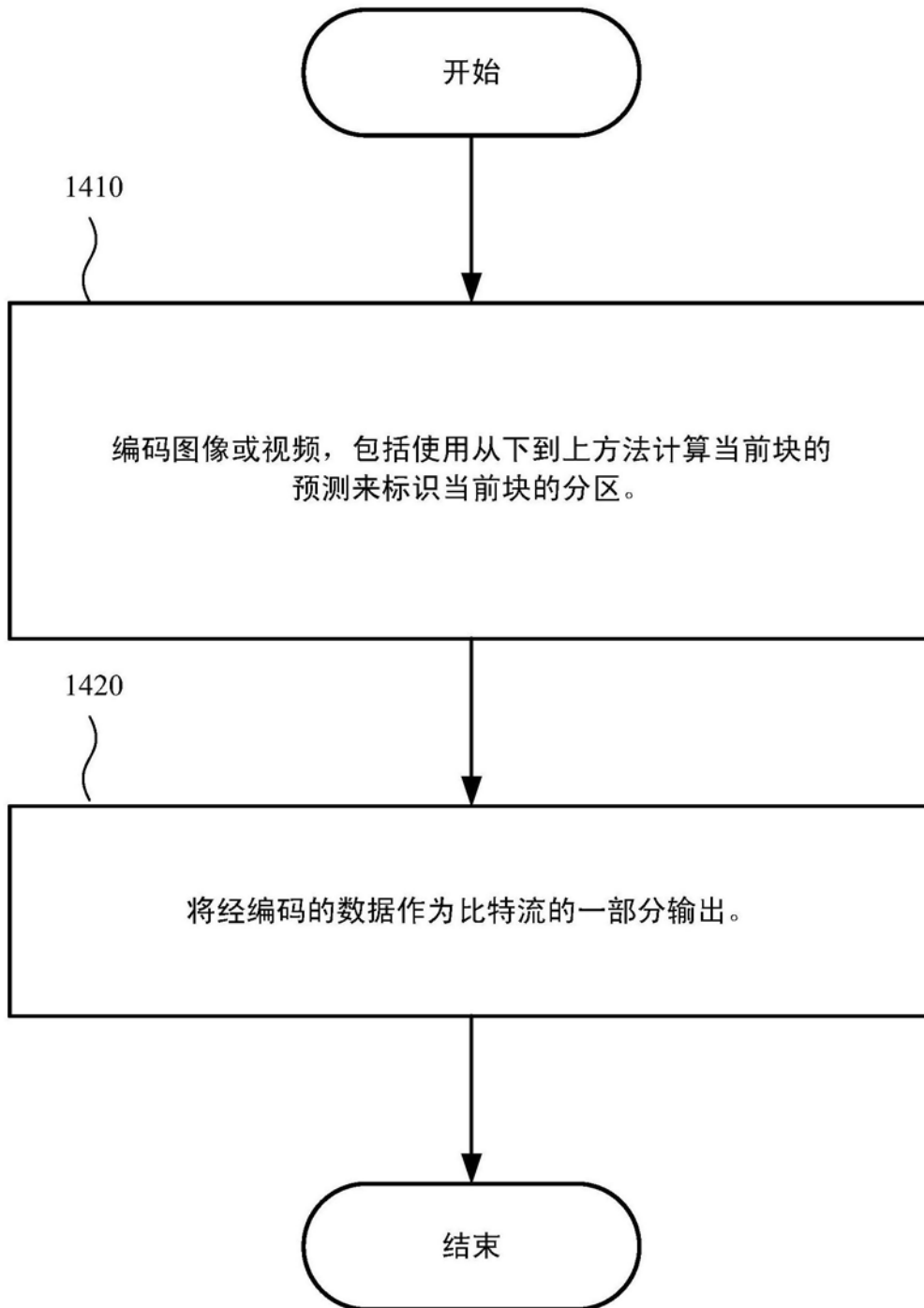


图14



1500 用于分区具有非对称分区的2N×2N块的从下到上方法

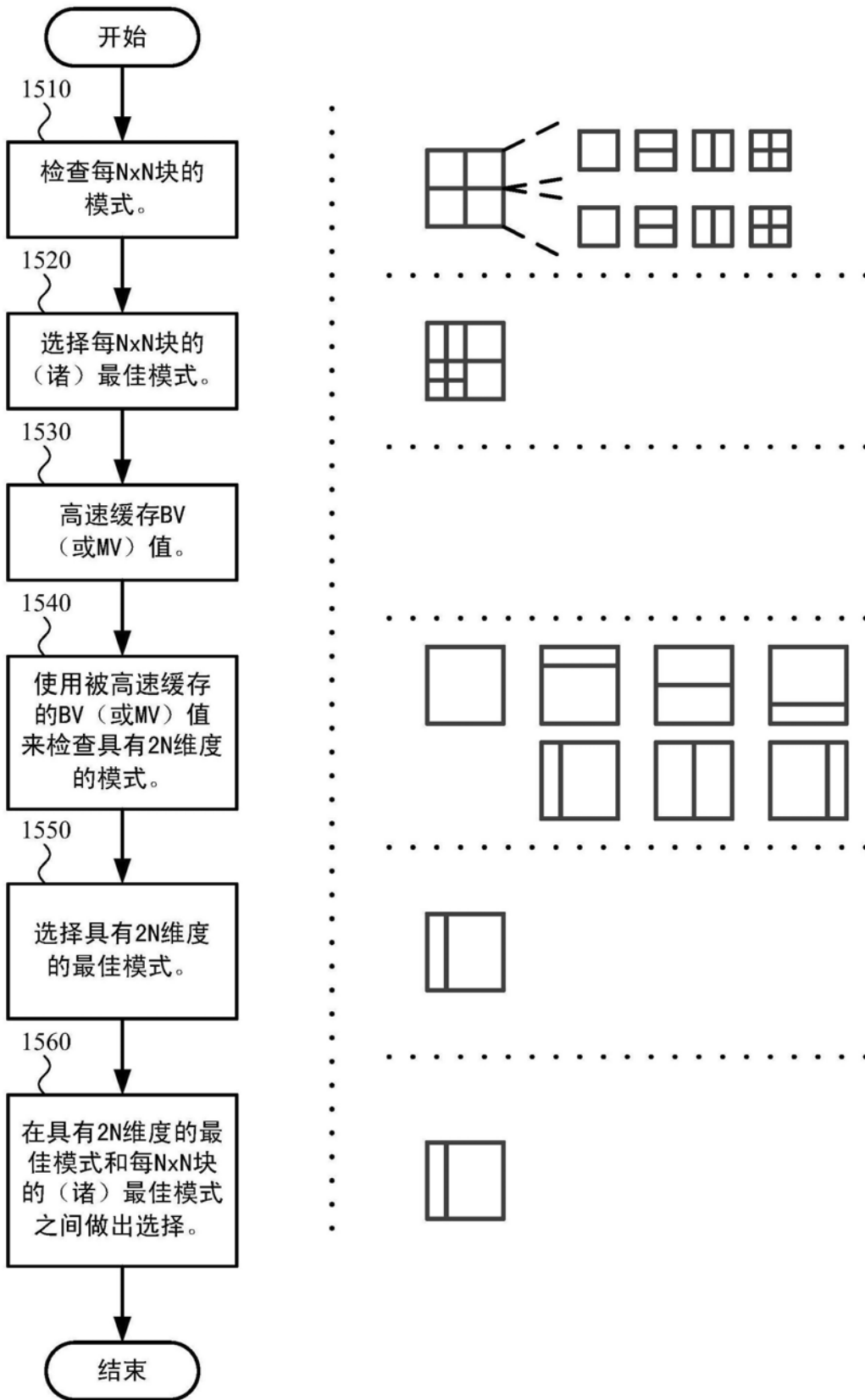


图15

用于分区具有非对称分区的2N×2N块的更快的  
从下到上方法

1600

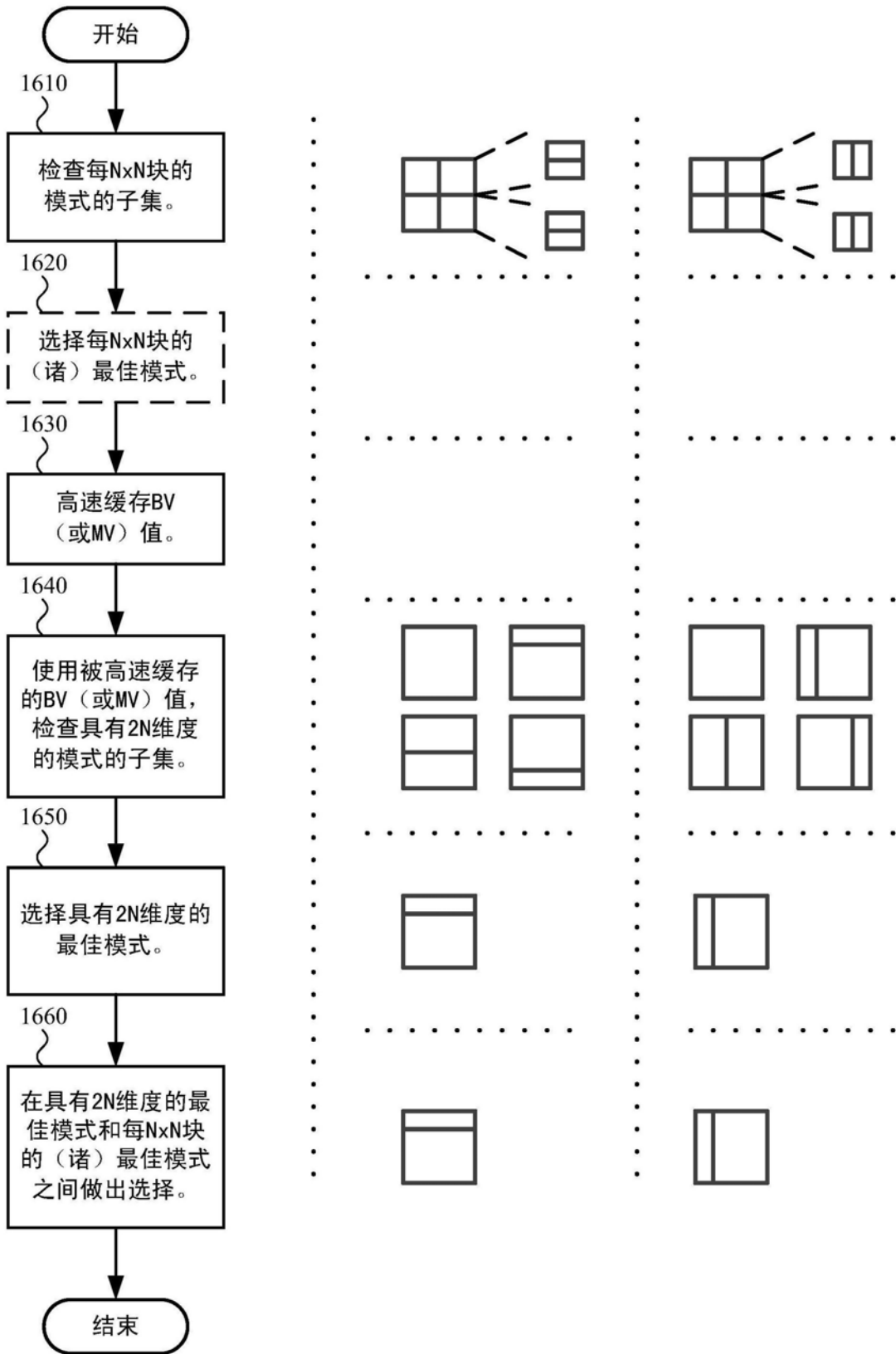


图16

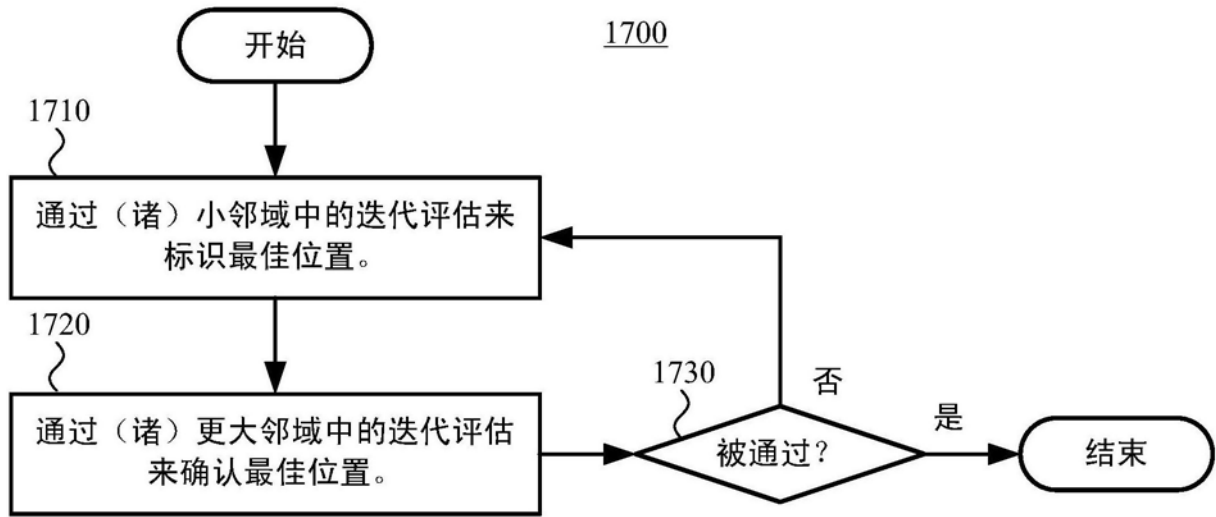


图17

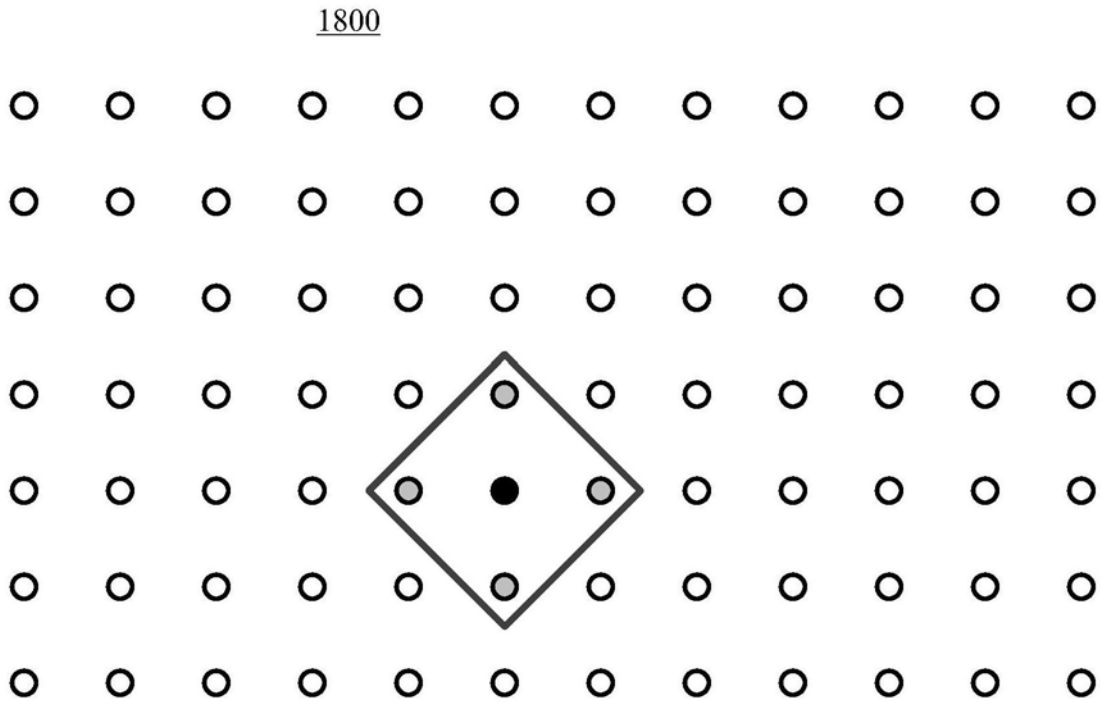


图18a

1801

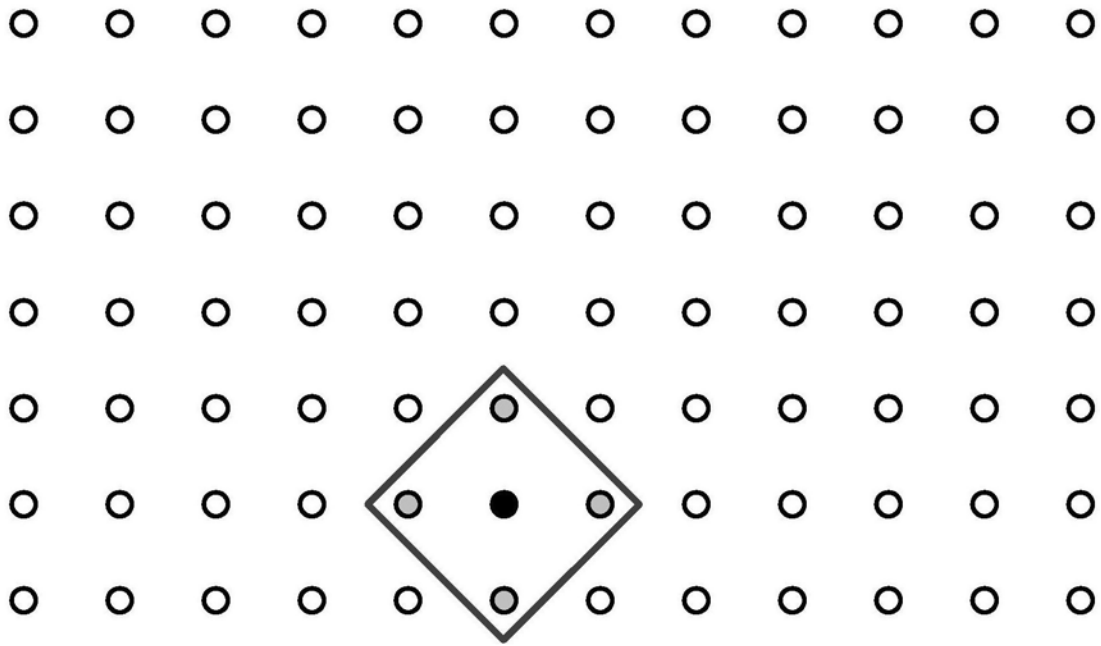


图18b

1900

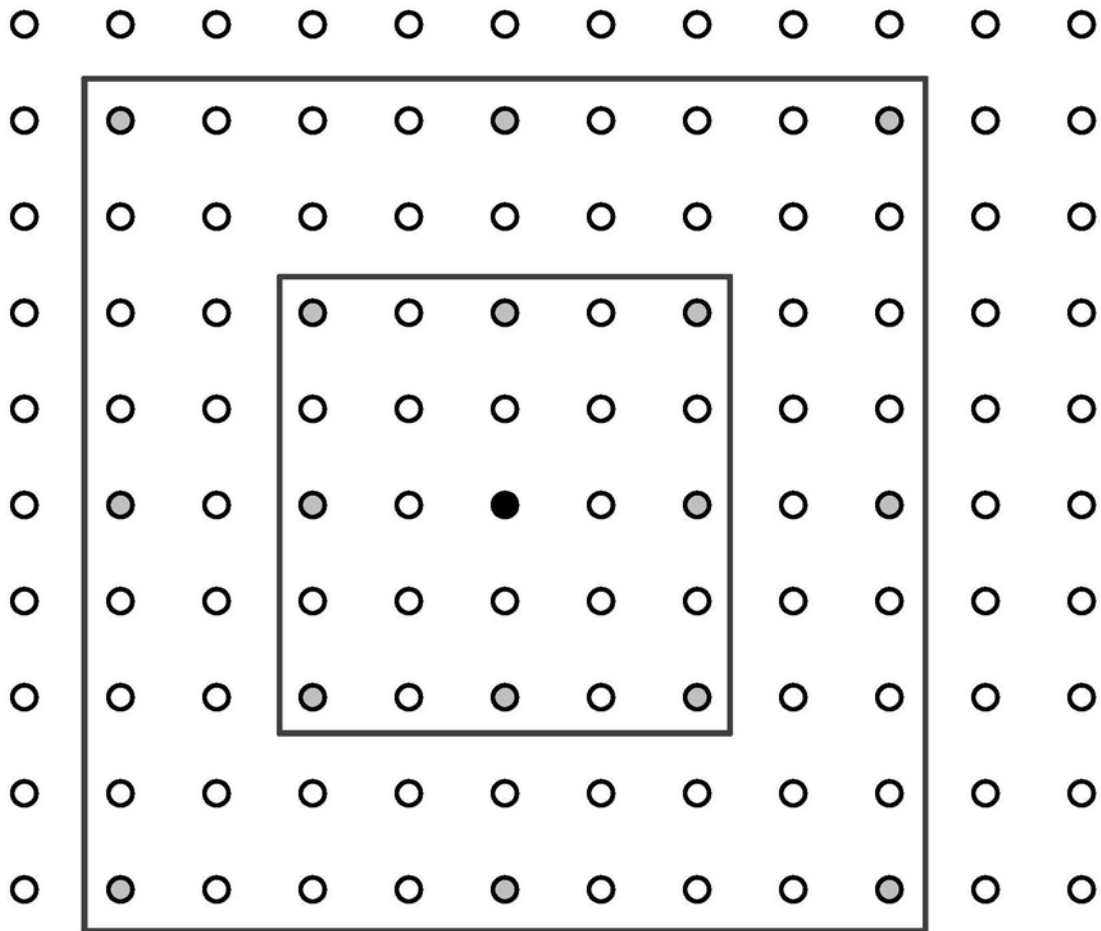


图19

2000

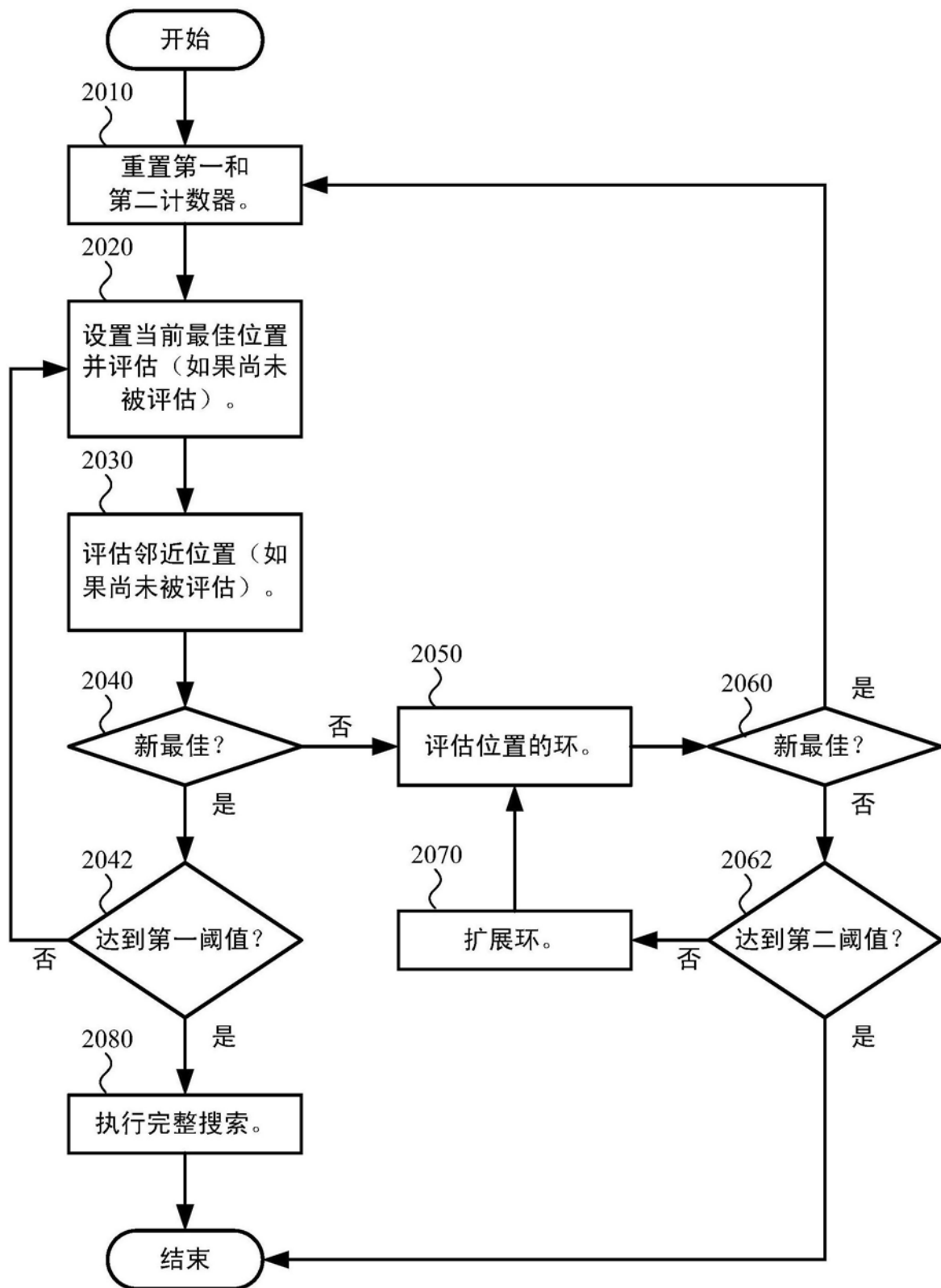
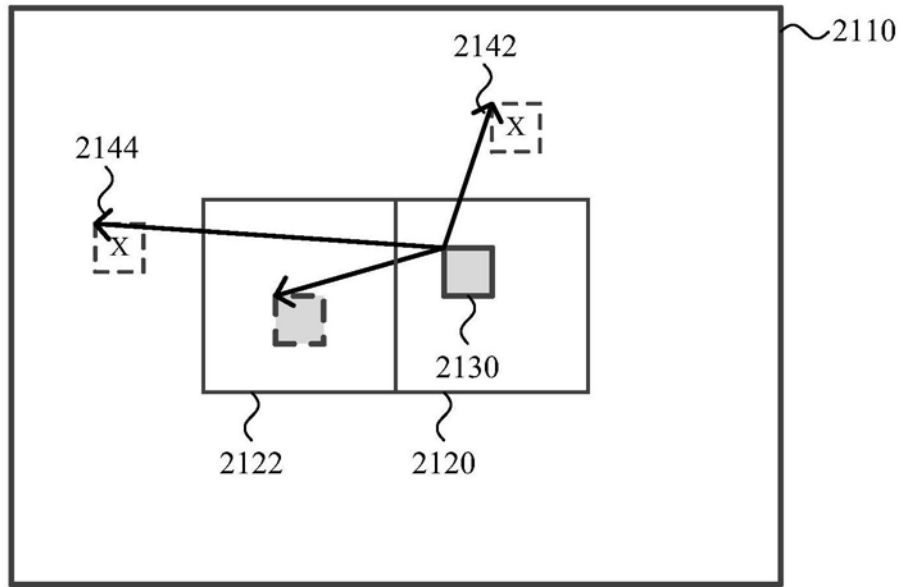
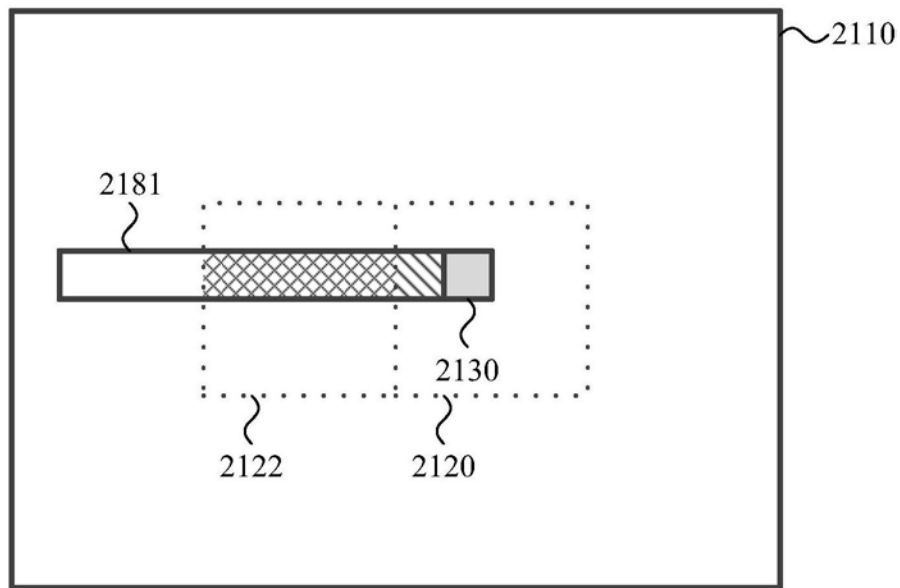


图20



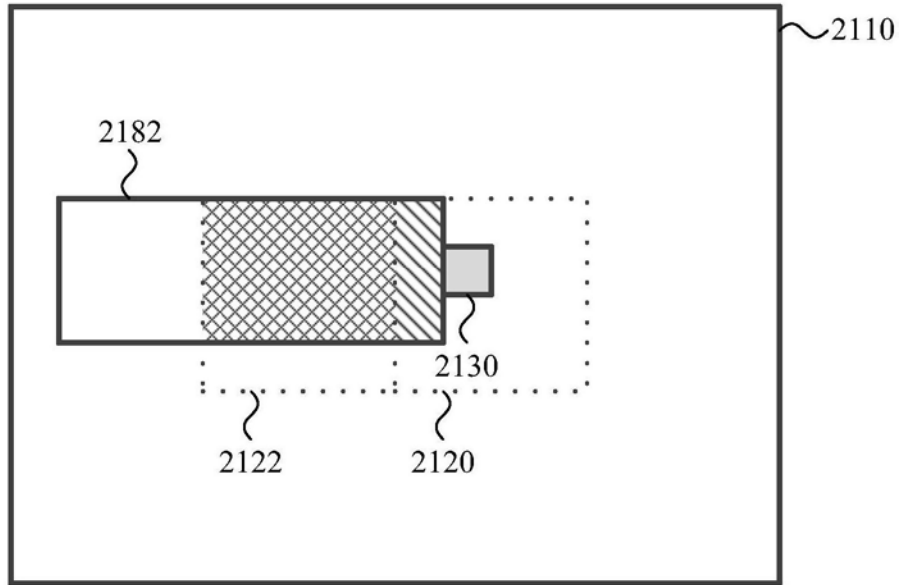
指示在当前帧（2110）的当前块（2130）的搜索范围外的区域的位移的候选块向量（2142、2144）

图21a



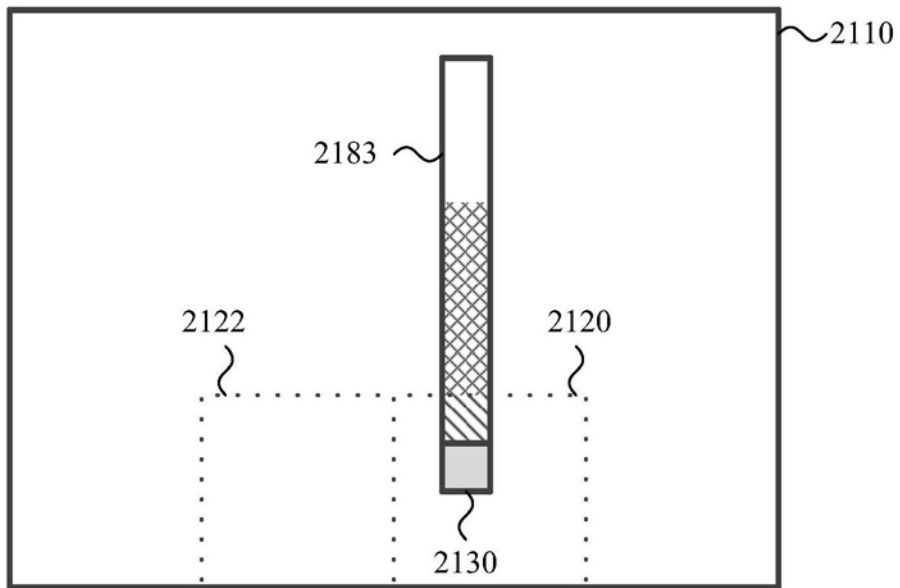
当前帧（2110）的当前块（2130）的具有维度 $2N \times \frac{1}{4}N$ 的水平偏置的搜索范围（2181）

图21b



当前帧 (2110) 的当前块 (2130) 的具有维度  $2N \times \frac{3}{4}N$  的水平偏置的搜索范围 (2182)

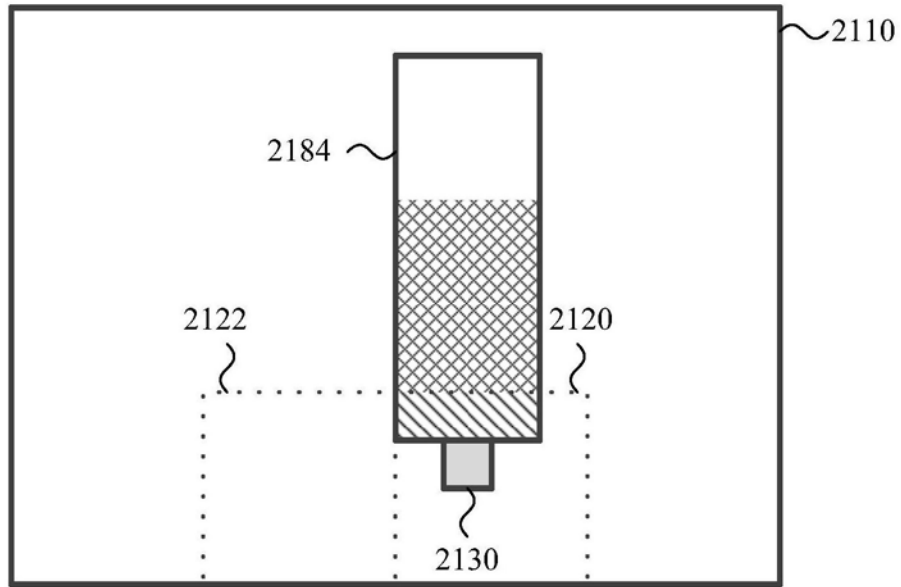
图21c



当前帧 (2110) 的当前块 (2130) 的具有维度  $\frac{1}{4}N \times 2N$  的垂直偏置的搜索范围 (2183)

图21d





当前帧（2110）的当前块（2130）的具有维度 $\frac{3}{4}N \times 2N$ 的垂直偏置的搜索范围（2184）

图21e

2200

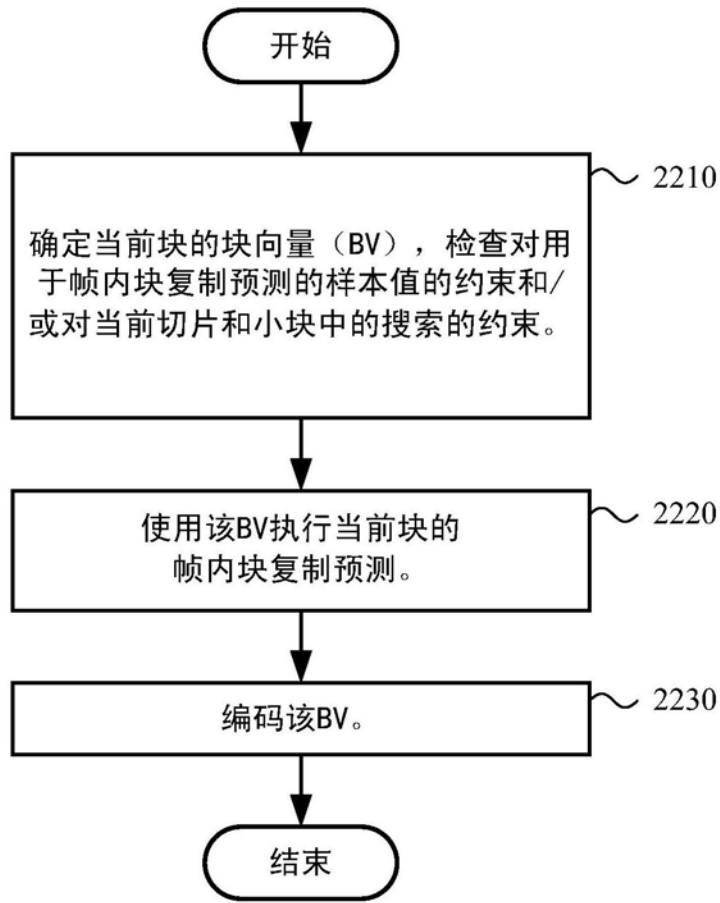


图22

2300

当前块和可包括BV的区域的右下位置的块的z-扫描次序

0	1	2	8	9	12
	3	4   5 6   7	10	11	
13	14	15	18	? 当前块 2330	
	16	17			

图23