



(10) 授权公告号 CN 113678452 B

(45) 授权公告日 2024. 09. 13

(21) 申请号 202080016719.8

(22) 申请日 2020.02.28

(65) 同一申请的已公布的文献号
申请公布号 CN 113678452 A

(43) 申请公布日 2021.11.19

(30) 优先权数据

62/812,828 2019.03.01 US

16/803,520 2020.02.27 US

(85) PCT国际申请进入国家阶段日
2021.08.25(86) PCT国际申请的申请数据
PCT/US2020/020392 2020.02.28(87) PCT国际申请的公布数据
W02020/180685 EN 2020.09.10(73) 专利权人 高通股份有限公司
地址 美国加利福尼亚州(72) 发明人 H.王 W-J.钱 M.卡尔切维茨
H.黄(74) 专利代理机构 北京市柳沈律师事务所
11105
专利代理师 安之斐

(51) Int.Cl.

H04N 19/44 (2006.01)

H04N 19/513 (2006.01)

H04N 19/577 (2006.01)

H04N 19/523 (2006.01)

(56) 对比文件

CN 110169073 A, 2019.08.23

US 2018184117 A1, 2018.06.28

US 2018192071 A1, 2018.07.05

US 2018278951 A1, 2018.09.27

审查员 于雷

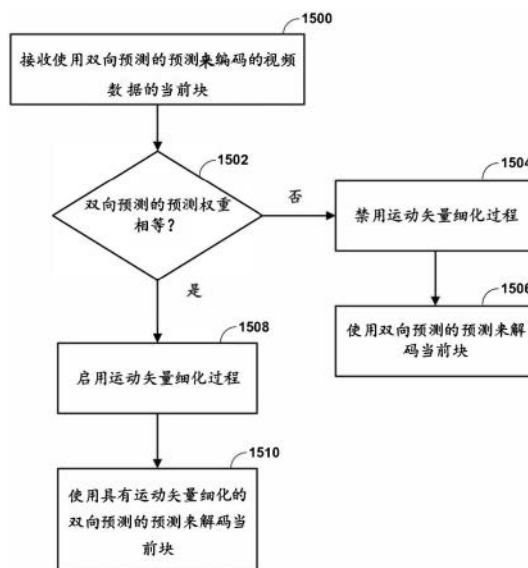
权利要求书3页 说明书29页 附图16页

(54) 发明名称

对解码器侧运动矢量细化的约束

(57) 摘要

一种视频编码器, 可经配置以基于用于双向预测的预测的权重是否相等来确定使用解码器侧运动矢量细化过程, 包括双边模板匹配。在一个示例中, 当用于双向预测的预测的权重不相等时, 可以禁用解码器侧运动矢量细化。



1. 一种用于对视频数据进行解码的方法,所述方法包括:
确定使用双向预测的预测来编码的视频数据的当前块的运动矢量;
确定用于所述当前块的所述双向预测的预测的来自两个参考图片列表的预测器的权重不相等;
基于确定用于所述当前块的所述双向预测的预测的来自所述两个参考图片列表的所述预测器的所述权重不相等,确定对所述运动矢量禁用运动矢量细化过程;以及
在没有所述运动矢量细化过程的情况下使用所述双向预测的预测来对视频数据的所述当前块进行解码。
2. 如权利要求1所述的方法,其中,所述双向预测的预测是加权的双向预测。
3. 如权利要求1所述的方法,其中,所述双向预测的预测是广义双向预测。
4. 如权利要求1所述的方法,还包括:
确定用于在图片级别对所述当前块进行所述双向预测的预测的来自两个参考图片列表的预测器的权重。
5. 如权利要求1所述的方法,还包括:
确定用于在块级别对所述当前块进行所述双向预测的预测的来自两个参考图片列表的预测器的权重。
6. 如权利要求1所述的方法,其中,所述运动矢量细化过程是双边模板匹配。
7. 如权利要求1所述的方法,还包括:
确定使用双向预测的预测进行编码的视频数据的第二块的运动矢量;
确定用于所述第二块的所述双向预测的预测的来自两个参考图片列表的预测器的权重相等;
确定由所述第二块的所述运动矢量识别的预测块之间的差;
基于用于所述第二块的所述双向预测的预测的来自所述两个参考图片列表的所述预测器的所述权重和所确定的由所述第二块的所述运动矢量识别的所述预测块之间的差来确定对所述第二块的所述运动矢量应用所述运动矢量细化过程。
8. 如权利要求1所述的方法,确定由所述运动矢量识别的预测块之间的差,其中,基于用于所述当前块的所述双向预测的预测的来自所述两个参考图片列表的所述预测器的所述权重来确定对所述运动矢量禁用所述运动矢量细化过程还基于所确定的由所述运动矢量识别的所述预测块之间的差,并且还包括:
在由所述运动矢量识别的所述预测块之间的差小于阈值的情况下,确定禁用所述运动矢量细化过程。
9. 如权利要求7所述的方法,其中,确定由所述第二块的所述运动矢量识别的预测块之间的差包括:
使用所述预测块之间的绝对差的和或者所述预测块之间的平方误差的和来确定由所述第二块的所述运动矢量识别的所述预测块之间的差。
10. 如权利要求8所述的方法,还包括:
基于所述预测块中样本的数量和所述预测块中样本的位深度来确定所述阈值。
11. 如权利要求7所述的方法,其中,所述运动矢量细化过程是双边模板匹配。
12. 如权利要求11所述的方法,其中,对所述第二块的所述运动矢量应用所述双边模板

匹配包括：

使用所述双边模板匹配为所述第二块确定两个预测块，其中，所述两个预测块中的第一预测块是基于所述第二块的所述运动矢量被确定的，其中，所述第二块的所述运动矢量包括非整数运动矢量，并且其中，确定所述第一预测块包括：

基于非整数运动矢量来执行水平插值；

以比内部深度更高的位深度存储水平插值的输出；以及

使用存储在较高位深度的水平插值的输出来执行垂直插值。

13. 一种被配置为对视频数据进行解码的装置，所述装置包括：

存储器，被配置为存储视频数据的当前块；以及

一个或多个处理器，在电路中实现并与所述存储器通信，其中，所述一个或多个处理器被配置为：

确定使用双向预测的预测来编码的视频数据的当前块的运动矢量；

确定用于所述当前块的所述双向预测的预测的来自两个参考图片列表的预测器的权重不相等；

基于确定用于所述当前块的所述双向预测的预测的来自所述两个参考图片列表的所述预测器的所述权重不相等，确定对所述运动矢量禁用运动矢量细化过程；以及

在没有所述运动矢量细化过程的情况下使用所述双向预测的预测来对视频数据的所述当前块进行解码。

14. 如权利要求13所述的装置，其中，所述双向预测的预测是加权的双向预测。

15. 如权利要求13所述的装置，其中，所述双向预测的预测是广义双向预测。

16. 如权利要求13所述的装置，其中，所述一个或多个处理器还被配置为：

确定用于在图片级别对所述当前块进行所述双向预测的预测的来自两个参考图片列表的预测器的权重。

17. 如权利要求13所述的装置，其中，所述一个或多个处理器还被配置为：

确定用于在块级别对所述当前块进行所述双向预测的预测的来自两个参考图片列表的预测器的权重。

18. 如权利要求13所述的装置，其中，所述运动矢量细化过程是双边模板匹配。

19. 如权利要求13所述的装置，其中，所述一个或多个处理器还被配置为：

确定使用双向预测的预测进行编码的视频数据的第二块的运动矢量；

确定用于所述第二块的所述双向预测的预测的来自两个参考图片列表的预测器的权重相等；

确定由所述第二块的所述运动矢量识别的预测块之间的差；

基于用于所述第二块的所述双向预测的预测的来自所述两个参考图片列表的所述预测器的所述权重和所确定的由所述运动矢量识别的所述预测块之间的差来确定对所述第二块的所述运动矢量应用所述运动矢量细化过程。

20. 如权利要求13所述的装置，其中，所述一个或多个处理器还被配置为确定由所述运动矢量识别的预测块之间的差，其中，基于用于所述当前块的所述双向预测的预测的来自所述两个参考图片列表的所述预测器的所述权重来确定对所述运动矢量禁用所述运动矢量细化过程还基于所确定的由所述运动矢量识别的所述预测块之间的差，所述一个或多个

处理器还被配置为：

在由所述运动矢量识别的所述预测块之间的差小于阈值的情况下，确定禁用所述运动矢量细化过程。

21. 根据权利要求19所述的装置，其中为了确定由所述第二块的所述运动矢量识别的预测块之间的差，所述一个或多个处理器还被配置为：

使用所述预测块之间的绝对差的和或者所述预测块之间的平方误差的和来确定由所述第二块的所述运动矢量识别的所述预测块之间的差。

22. 如权利要求20所述的装置，其中，所述一个或多个处理器还被配置为：

基于所述预测块中样本的数量和所述预测块中样本的位深度来确定所述阈值。

23. 如权利要求19所述的装置，其中，所述运动矢量细化过程是双边模板匹配。

24. 如权利要求23所述的装置，其中，为了对所述第二块的所述运动矢量应用所述双边模板匹配，所述一个或多个处理器还被配置为：

使用所述双边模板匹配为所述第二块确定两个预测块，其中，所述两个预测块中的第一预测块是基于所述第二块的所述运动矢量被确定的，其中，所述第二块的所述运动矢量包括非整数运动矢量，并且其中，为了确定所述第一预测块，所述一个或多个处理器还被配置为：

基于非整数运动矢量来执行水平插值；

以比内部位深度更高的位深度存储水平插值的输出；以及

使用存储在较高位深度的水平插值的输出来执行垂直插值。

25. 如权利要求13所述的装置，其中，所述装置是无线通信设备。

26. 一种存储指令的非暂时性计算机可读存储介质，所述指令在执行时使一个或多个处理器被配置为对视频数据进行解码以：

确定使用双向预测的预测来编码的视频数据的当前块的运动矢量；

确定用于所述当前块的所述双向预测的预测的来自两个参考图片列表的预测器的权重不相等；

基于确定用于所述当前块的所述双向预测的预测的来自所述两个参考图片列表的所述预测器的所述权重不相等，确定对所述运动矢量禁用运动矢量细化过程；以及

在没有所述运动矢量细化过程的情况下使用所述双向预测的预测来对视频数据的所述当前块进行解码。

27. 一种被配置为对视频数据进行解码的装置，所述装置包括：

用于确定使用双向预测的预测来编码的视频数据的当前块的运动矢量的部件；

用于确定用于所述当前块的所述双向预测的预测的来自两个参考图片列表的预测器的权重不相等的部件；

用于基于确定用于所述当前块的所述双向预测的预测的来自所述两个参考图片列表的所述预测器的所述权重不相等，确定对所述运动矢量禁用运动矢量细化过程的部件；以及

用于在没有所述运动矢量细化过程的情况下使用所述双向预测的预测来对视频数据的所述当前块进行解码的部件。

对解码器侧运动矢量细化的约束

[0001] 本申请要求于2020年2月27日提交的美国专利申请第16/803,520号的优先权,后者要求于2019年3月1日提交的美国临时申请第62/812,828号的优先权,上述每个申请的全部内容通过引用并入本文。

技术领域

[0002] 本公开涉及视频编码和视频解码。

背景技术

[0003] 数字视频能力可以被整合到广泛的设备中,包括数字电视、数字直接广播系统、无线广播系统、个人数字助理(PDA)、笔记本电脑或台式电脑、平板电脑、电子书阅读器、数码相机、数字记录设备、数字媒体播放器、视频游戏设备、视频游戏机、蜂窝或卫星无线电话、所谓的“智能电话”、视频电话会议设备、视频流传送设备等。数字视频设备实施视频编码技术,诸如MPEG-2、MPEG-4、ITU-T H.263、ITU-T H.264/MPEG-4、第10部分、高级视频编码(AVC)、ITU-T H.265/高效视频编码(HEVC)所定义的标准或者此类标准的扩展中所描述的那些视频编码技术。通过实施这样的视频编码技术,视频设备可以更有效地发送、接收、编码、解码和/或存储数字视频信息。

[0004] 视频编码技术包括空间(图片内)预测和/或时间(图片间)预测,以减少或去除视频序列中固有的冗余。对于基于块的视频编码,可以将视频切片(例如,视频图片或视频图片的一部分)划分为视频块,这些视频块也可以称为编码树单元(CTU)、编码单元(CU)和/或编码节点。图片的帧内编码(I)切片中的视频块使用相对于同一图片中邻近块中的参考样本的空间预测进行编码。图片的帧间编码(P或B)切片中的视频块可以使用相对于同一图片中邻近块中的参考样本的空间预测或者相对于其他参考图片中的参考样本的时间预测。图片可以称为帧,并且参考图片可以称为参考帧。

发明内容

[0005] 总体上,本公开描述了用于解码器侧运动矢量细化(DMVR)的技术。特别是,本公开描述了用于确定何时将运动矢量细化过程(诸如双边模板匹配)应用于使用双向预测的预测来编码的视频数据块的技术。在一个示例中,当用于双向预测的预测的权重不相等时,视频编码器可禁用双向预测块的运动矢量细化过程。在本公开的另一示例中,当由双向预测的运动矢量识别的预测块之间的差小于阈值时,视频编码器还可以禁用运动矢量细化过程。本公开的技术可以应用于任何现有视频编解码器的扩展,与开发中的VVC(通用视频编码)标准或者是任何未来视频编码标准中的有效编码工具一起使用,或者成为任何未来视频编码标准中的有效编码工具。

[0006] 在一个示例中,一种方法包括确定使用双向预测的预测来编码的视频数据的当前块的运动矢量,基于用于当前块的双向预测的预测的权重来确定对运动矢量使用运动矢量细化过程,以及基于使用运动矢量细化过程的确定来对运动矢量应用运动矢量细化过程。

[0007] 在另一示例中,一种被配置为对视频数据进行解码的装置,该装置包括被配置为存储视频数据的当前块的存储器,以及在电路中实现并与存储器通信的一个或多个处理器,一个或多个处理器被配置为确定使用双向预测的预测来编码的视频数据的当前块的运动矢量,基于用于当前块的双向预测的预测的权重来确定对运动矢量使用运动矢量细化过程,以及基于使用运动矢量细化过程的确定来对运动矢量应用运动矢量细化过程。

[0008] 在另一示例中,一种被配置为对视频数据进行解码的装置包括用于确定使用双向预测的预测来编码的当前视频数据块的运动矢量的部件、用于基于用于当前块的双向预测的预测的权重来确定对运动矢量使用运动矢量细化过程的部件,以及用于基于使用运动矢量细化过程的确定来对运动矢量应用运动矢量细化过程的部件。

[0009] 在另一示例中,本公开描述了一种存储指令的非暂时性计算机可读存储介质,指令在执行时使一个或多个处理器被配置为对视频数据进行解码以确定使用双向预测的预测来编码的视频数据的当前块的运动矢量,基于用于当前块的双向预测的预测的权重来确定对运动矢量使用运动矢量细化过程,以及基于使用运动矢量细化过程的确定来对运动矢量应用运动矢量细化过程。

[0010] 在附图和以下描述中阐述了一个或多个示例的细节。从描述、附图和权利要求中,其他特征、目的和优点将是显而易见的。

附图说明

[0011] 图1是说明可以执行本公开的技术的示例视频编码和解码系统的框图。

[0012] 图2是说明可以执行本公开的技术的示例视频编码器的框图。

[0013] 图3是说明可以执行本公开的技术的示例视频解码器的框图。

[0014] 图4A是说明用于合并模式的空间邻近运动矢量候选的概念图。

[0015] 图4B为说明用于高级运动矢量预测 (AMVP) 模式的空间邻近运动矢量候选的概念图。

[0016] 图5A是说明时间运动矢量候选的概念图。

[0017] 图5B是说明运动矢量缩放的概念图。

[0018] 图6A是说明合并运动矢量细化 (MMVR) 过程的概念图。

[0019] 图6B是说明MMVR搜索点的概念图。

[0020] 图7是说明偏移合并候选的概念图。

[0021] 图8是说明双边模板匹配的概念图。

[0022] 图9是说明解码器侧运动矢量推导 (DMVD) 技术的示例流水线状态的概念图。

[0023] 图10是说明示例光流轨迹的概念图。

[0024] 图11是说明8x4块的示例双向光流 (BIO) 的概念图。

[0025] 图12是说明示例编码方法的流程图。

[0026] 图13是说明示例解码方法的流程图。

[0027] 图14是说明用于应用解码器侧运动矢量细化的示例解码方法的流程图。

[0028] 图15是说明用于应用解码器侧运动矢量细化的另一示例解码方法的流程图。

具体实施方式

[0029] 总体上,本公开描述了用于解码器侧运动矢量细化(DMVR)的技术。特别是,本公开描述了用于确定何时将运动矢量细化过程(诸如双边模板匹配)应用于使用双向预测的预测来编码的视频数据块的技术。根据本公开的技术,视频编码器可被配置为确定何时应用运动矢量细化过程而无需对指示是否将应用运动矢量细化的语法元素进行显式编码(例如,编码或解码)。以此方式,视频编码器可以选择性地应用运动矢量细化而不增加经编码视频位流中的开销信令。在一个示例中,当用于双向预测的权重不相等时,视频编码器(例如,视频编码器和/或视频解码器)可以禁用双向预测块的运动矢量细化过程。在本公开的另一示例中,当由双向预测的运动矢量识别的预测块之间的差小于阈值时,视频编码器还可以禁用运动矢量细化过程。

[0030] 图1是说明可为DMVR执行本公开的技术的示例视频编码及解码系统100的框图。本公开的技术总体上针对对视频数据进行编码(编码和/或解码)。通常,视频数据包括用于处理视频的任何数据。因此,视频数据可以包括原始的、未编码的视频、经编码视频、经解码(例如,重建的)视频和视频元数据(诸如信令数据)。如下文将更详细解释的,本公开描述了用于确定何时将用于DMVR的运动矢量细化过程(诸如双边模板匹配)应用于使用双向预测的预测来编码的视频数据块的技术。在一个示例中,当用于双向预测的预测的权重不相等时,视频编码器可以禁用用于双向预测块的运动矢量细化过程。在本公开的另一示例中,当由双向预测的运动矢量识别的预测块之间的差小于阈值时,视频编码器还可以禁用运动矢量细化过程。本公开的技术可提高编码效率和/或简化DMVR技术的实现。

[0031] 如图1所示,在此示例中,系统100包括提供将由目的地设备116进行解码和显示的经编码视频数据的源设备102。特别地,源设备102经由计算机可读介质110将视频数据提供给目的地设备116。源设备102和目的地设备116可以包括多种设备中的任何一种,包括台式计算机、笔记本(即膝上型计算机)、平板电脑、机顶盒、手持电话(如智能手机)、电视、照相机、显示设备、数字媒体播放器、视频游戏机、视频流设备等。在一些情况下,源设备102和目的地设备116可以被配备用于无线通信,并且因此可以称为无线通信设备。

[0032] 在图1的示例中,源设备102包括视频源104、存储器106、视频编码器200和输出接口108。目的地设备116包括输入接口122、视频解码器300、存储器120和显示设备118。根据本公开,源设备102的视频编码器200和目的地设备116的视频解码器300可以配置为应用用于解码器侧运动矢量细化的技术。因而,源设备102表示视频编码设备的示例,而目的地设备116表示视频解码设备的示例。在其他示例中,源设备和目的地设备可以包括其他组件或布置。例如,源设备102可以从诸如外部摄像机的外部视频源接收视频数据。同样,目的地设备116可以与外部显示设备通过接口相连接,而不包括集成的显示设备。

[0033] 如图1所示的系统100仅是一个示例。总的来说,任何数字视频编码和/或解码设备都可以执行用于解码器侧运动矢量细化的技术。源设备102和目的地设备116仅仅是此类编码设备的示例,其中,源设备102生成用于传输到目的地设备116的编码视频数据。本公开将“编码”设备称为执行数据编码(编码和/或解码)的设备。从而,视频编码器200和视频解码器300表示编码设备的示例,特别地,分别是视频编码器和视频解码器。在一些示例中,设备102和设备116可以用基本上对称的方式操作,使得设备102和设备116中的每一个均包括视频编码和解码组件。于是,系统100可以支持视频设备102和设备116之间的单向或双向视频

传输,例如用于视频流、视频回放、视频广播或者视频电话。

[0034] 总的来说,视频源104表示视频数据(即原始的未编码视频数据)源,并将视频数据的顺序系列的图片(也称为“帧”)提供给视频编码器200,其对图片的数据进行编码。源设备102的视频源104可以包括视频捕获设备,诸如摄像机,包括先前捕获的原始视频的视频档案和/或从视频内容提供商接收视频的视频馈送接口。作为进一步的替代方案,视频源104可以生成基于计算机图形的数据作为源视频,或者实况视频、存档视频和计算机生成视频的组合。在每种情况下,视频编码器200对捕获的、预捕获的或计算机生成的视频数据进行编码。视频编码器200可以将图片从接收顺序(有时称为“显示顺序”)重新排列为用于编码的编码顺序。视频编码器200可以生成包括经编码视频数据的比特流。然后,源设备102可以经由输出接口108将经编码视频数据输出到计算机可读介质110上,通过例如目的地设备116的输入接口122进行接收和/或检索。

[0035] 源设备102的存储器106和目的地设备116的存储器120表示通用存储器。在一些示例中,存储器106、120可以存储原始视频数据,例如来自视频源104的原始视频和来自视频解码器300的原始的经解码视频数据。附加地或可替代地,存储器106、120可以分别存储可由例如视频编码器200和视频解码器300执行的软件指令。尽管在此示例中视频编码器200和视频解码器300分开示出,但是应当理解的是,视频编码器200和视频解码器300还可以包括实现功能上相似或等效目的的内部存储器。进一步来说,存储器106、120可以存储例如从视频编码器200输出并输入到视频解码器300的编码视频数据。在一些示例中,存储器106、120的部分可以分配为一个或多个视频缓冲器,例如用来存储原始的经解码和/或经编码视频数据。

[0036] 计算机可读介质110可以表示能够将经编码视频数据从源设备102传输到目的地设备116的任何类型的介质或设备。在一个示例中,计算机可读介质110表示通信介质以使源设备102能够例如经由射频网络或基于计算机的网络将经编码视频数据实时地直接发送到目的地设备116。根据诸如无线通信协议的通信标准,输出接口108可以对包括经编码视频数据的传输信号进行调制,并且输入接口122可以对接收到的传输信号进行调制。通信介质可以包括任何无线或有线通信介质,诸如射频(RF)频谱或一条或多条物理传输线。通信介质可以形成诸如局域网、广域网或诸如因特网的全球网络的基于分组的网络的一部分。通信介质可以包括路由器、交换机、基站或可能有助于从源设备102到目的地设备116的通信的任何其他装备。

[0037] 在一些示例中,源设备102可以将经编码数据从输出接口108输出到存储设备116。类似地,目的地设备116可以经由输入接口122访问来自存储设备116的编码数据。存储设备116可以包括各种分布式或本地访问的数据存储介质中的任何一种,诸如硬盘、蓝光光盘、DVD、CD-ROM、闪存、易失性或非易失性存储器,或者用于存储编码视频数据的任何其他合适的数字存储介质。

[0038] 在一些示例中,源设备102可以将经编码视频数据输出到文件服务器114或可以存储由源设备102生成的经编码视频数据的另一中间存储设备。目的地设备116可以经由流或下载来访问来自文件服务器114的存储的视频数据。文件服务器114可以是能够存储经编码视频数据并将经编码视频数据发送到目的地设备116的任何类型的服务器设备。文件服务器114可以表示网络服务器(例如,用于网站)、文件传输协议(FTP)服务器、内容交付网络设

备或网络附加存储 (NAS) 设备。目的地设备116可以通过包括因特网连接的任何标准数据连接来访问来自文件服务器114的经编码视频数据。这可以包括无线信道 (例如Wi-Fi连接)、有线连接 (例如DSL、电缆调制解调器等) 或者适合访问存储在文件服务器114上的编码视频数据的二者的组合。文件服务器114和输入接口122可以配置为根据流传输协议、下载传输协议或其组合来操作。

[0039] 输出接口108和输入接口122可以表示无线发射机/接收机、调制解调器、有线网络组件 (例如, 以太网卡)、根据各种IEEE 802.11标准中的任何一种进行操作的无线通信组件, 或者其他物理组件。在输出接口108和输入接口122包括无线组件的示例中, 输出接口108和输入接口122可以配置为根据诸如4G、4G-LTE (长期演进)、LTE高级、5G或类似标准的蜂窝通信标准来传输诸如经编码视频数据的数据。在输出接口108包括无线发射机的某些示例中, 输出接口108和输入接口122可以配置为根据其他无线标准, 诸如IEEE 802.11规范、IEEE 802.15规范 (例如ZigBee™)、Bluetooth™标准等来传输诸如经编码视频数据的数据。在一些示例中, 源设备102和/或目的地设备116可以包括各自的片上系统 (SoC) 设备。例如, 源设备102可以包括SoC设备来执行归属于视频编码器200和/或输出接口108的功能, 并且目的地设备116可以包括SoC设备来执行归属于视频解码器300和/或输入接口122的功能。

[0040] 本公开的技术可以应用于支持各种多媒体应用中的任何一种的视频编码, 诸如空中电视广播、有线电视传输、卫星电视传输、诸如基于HTTP的动态自适应流 (DASH) 的互联网流视频传输、被编码到数据存储介质上的数字视频、对存储在数据存储介质上的数字视频进行解码或者其他应用。

[0041] 目的地设备116的输入接口122从计算机可读介质110 (例如, 存储设备112、文件服务器114等) 接收经编码视频比特流。经编码视频比特流计算机可读存储介质110可以包括由视频编码器200定义的、也由视频解码器300使用的信令信息, 诸如语法元素, 该语法元素具有描述视频块或其他经编码单元 (例如, 切片、图片、图片组、序列等) 的特点和/或处理的值。显示设备118向用户显示经解码视频数据的解码图片。显示设备118可以表示诸如阴极射线管 (CRT)、液晶显示器 (LCD)、等离子显示器、有机发光二极管 (OLED) 显示器或另一类显示设备的各种显示设备中的任何一种。

[0042] 尽管未在图1中示出, 但在一些示例中, 视频编码器200和视频解码器300可以各自与音频编码器和/或音频解码器集成在一起, 并且可以包括适当的MUX-DEMUX单元或其他硬件和/或软件, 来处理公共数据流中包括音频和视频的多路复用流。

[0043] 视频编码器200和视频解码器300可以各自实现为各种合适的编码器和/或解码器电路中的任何一种, 诸如一个或多个微处理器、数字信号处理器 (DSP)、专用集成电路 (ASIC)、现场可编程门阵列 (FPGA)、离散逻辑、软件、硬件、固件或其任何组合。当这些技术部分地以软件实现时, 设备可以将用于软件的指令存储在合适的非暂时计算机可读介质中, 并使用一个或多个处理器在硬件中执行该指令来执行本公开的技术。视频编码器200和视频解码器300中的每一个可以被包括在一个或多个编码器或解码器中, 这两者中的任一者都可以集成为各自设备中组合编码器/解码器 (CODEC) 的一部分。包括视频编码器200和/或视频解码器300的设备可以包括集成电路、微处理器和/或诸如蜂窝电话的无线通信设备。

[0044] 视频编码器200和视频解码器300可以根据视频编码标准进行操作,诸如ITU-T H.265,也被称为高效视频编码(HEVC)或其扩展,诸如多视图和/或可缩放的视频编码扩展。替代地,视频编码器200和视频解码器300可以根据其他专有或工业标准(诸如联合探索测试模型(JEM)或ITU-T H.266(也称为通用视频编码(VVC)))进行操作。Bross等人描述了VVC标准的草案(“Versatile Video Coding (Draft 4),” Joint Video Experts Team (JVET) of ITU-T SG 16WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 13th Meeting: Marrakech, MA, 9-18 January 2019, JVET-M1001-v5), (以下简称“VVC草案4”)。然而,本公开的技术不限于任何特定的编码标准。

[0045] 总的来说,视频编码器200和视频解码器300可以执行图片的基于块的编码。术语“块”通常上是指包括待处理(例如,编码、解码或以其他方式在编码和/或解码过程中使用)的数据的结构。例如,块可以包括亮度和/或色度数据的样本的二维矩阵。一般来说,视频编码器200和视频解码器300可以对以YUV(例如Y、Cb、Cr)格式表示的视频数据进行编码。也就是说,视频编码器200和视频解码器300可以对亮度和色度分量进行编码,其中色度分量可以包括红色和蓝色色度分量,而不是对图片的样本的红色、绿色和蓝色(RGB)数据进行编码。在一些示例中,视频编码器200在编码之前将接收到的RGB格式的数据转换成YUV表示,并且视频解码器300将YUV表示转换成RGB格式。可替代地,预处理和后处理单元(未示出)可以执行这些转换。

[0046] 本公开一般可以指图片的编码(例如,编码和解码)以包括对图片数据进行编码或解码的过程。类似地,本公开可以涉及对图片的块进行编码以包括对块的数据进行编码或解码的过程,例如预测和/或残差编码。经编码视频比特流一般包括用于代表编码决策(例如,编码模式)的语法元素和将图片分割为块的一系列值。从而,对图片或块进行编码的引用一般应理解为对形成图片或块的语法元素的值进行编码。

[0047] HEVC定义各种的块,包括编码单元(CU)、预测单元(PU)和变换单元(TU)。根据HEVC,视频编码器(诸如视频编码器200)根据四叉树结构将编码树单元(CTU)分割为CU。也就是说,视频编码器将CTU和CU分割为四个相等的非重叠正方形,并且四叉树的每一个节点具有零个或四个子代节点。没有子节点的节点可以称为“叶节点”,并且这种叶节点的CU可以包括一个或多个PU和/或一个或多个TU。视频编码器可以进一步分割PU和TU。例如,在HEVC中,残差四叉树(RQT)表示TU的分割。在HEVC中,PU表示帧间预测数据,而TU表示残差数据。帧内预测的CU包括帧内预测信息,诸如帧内模式指示。

[0048] 作为另一示例,视频编码器200和视频解码器300可以配置为根据JEM或VVC进行操作。根据JEM或VVC,视频编码器(诸如视频编码器200)将图片分割为多个编码树单元(CTU)。视频编码器200可以根据诸如四叉树-二叉树(QTBT)结构或多类型树(MTT)结构的树结构来对CTU进行分割。QTBT结构消除了多个分割类型的概念,诸如HEVC的CU、PU和TU之间的分离。QTBT结构包括两个级别:根据四叉树分割而分割出来的第一级,以及根据二叉树分割而分割出来的第二级。QTBT结构的根节点对应于CTU。二叉树的叶节点对应于编码单元(CU)。

[0049] 在MTT分割结构中,可以使用四叉树(QT)分割、二叉树(BT)分割和一种或多种类型的三叉树(triple tree)(TT)(也称为三元树(ternary)(TT))分割对块进行分割。三叉树或三元树分割是一个块被分成三个子块的分割。在一些示例中,三叉树或三元树分割将块划分为三个子块而不通过中心划分原始块。MTT中的分割类型(例如QT、BT和TT)可以是对称的

或非对称的。

[0050] 在一些示例中,视频编码器200和视频解码器300可以使用单个QTBT结构来表示亮度分量和色度分量中的每一个,而在其他示例中,视频编码器200和视频解码器300可以使用两个或更多个QTBT或MTT结构,诸如用于亮度分量的一个QTBT/MTT结构和用于两个色度分量的另一QTBT/MTT结构(或用于各自色度分量的两个QTBT/MTT结构)。

[0051] 视频编码器200和视频解码器300可以配置为使用根据HEVC的四叉树分割、QTBT分割或其他分割结构。出于说明的目的,针对QTBT分割给出了本公开的技术的描述。然而,应该理解的是,本公开的技术还可以应用于配置使用四叉树分割或其它类型的分割的视频编码器。

[0052] 块(例如,CTU或CU)可以以各种方式在图片中分组。作为一个示例,区块可以指图片中特定图块内的CTU行的矩形区域。图块可以是图片中特定图块列和特定图块行内的CTU的矩形区域。图块列指的是高度等于图片高度且宽度由(例如,诸如在图片参数集中的)语法元素指定的CTU的矩形区域。图块行指的是具有由语法元素(例如,诸如在图片参数集中)指定的高度和等于图片宽度的宽度的CTU的矩形区域。

[0053] 本公开可以互换地使用“ $N \times N$ ”和“ N 乘 N ”来指块(诸如CU或其他视频块)在垂直和水平维度方面的样本维度,例如 16×16 样本或16乘16样本。一般来说, 16×16 的CU在垂直方向上将有16个样本($y=16$),在水平方向上将有16个样本($x=16$)。同样地, $N \times N$ 的CU一般地在垂直方向上具有 N 个样本,在水平方向上具有 N 个样本,其中 N 表示非负整数值。CU中的样本可以按行和列来排列。此外,CU在水平方向上不必具有与垂直方向上相同数量的样本。例如,CU可以包含 $N \times M$ 个样本,其中 M 不一定等于 N 。

[0054] 视频编码器200对表示预测和/或残差信息以及其他信息的CU的视频数据进行编码。预测信息指示将如何预测CU以便形成用于CU的预测块。残差信息总体上表示编码前CU的样本与预测块之间的逐样本差。

[0055] 为了预测CU,视频编码器200一般可以通过帧间预测或帧内预测来形成用于CU的预测块。帧间预测一般是指从先前编码的图片的数据中预测CU,而帧内预测总体上是指出从同一图片的先前编码数据中预测CU。为了执行帧间预测,视频编码器200可以使用一个或多个运动矢量来生成预测块。视频编码器200一般可以执行运动搜索以识别例如在CU和参考块之间的差的方面与CU紧密地匹配的参考块。视频编码器200可以使用绝对差之和(SAD)、平方差之和(SSD)、平均绝对差(MAD)、均方差(MSD)或其他此类差值计算来计算差值度量以确定参考块是否与当前CU紧密地匹配。在一些示例中,视频编码器200可以使用单向预测或双向预测来预测当前CU。

[0056] JEM和VVC的一些示例还提供仿射运动补偿模式,其可以被认为是帧间预测模式。在仿射运动补偿模式中,视频编码器200可以确定表示非平移运动的两个或更多个运动矢量,例如放大或缩小、旋转、透视运动或其他不规则运动类型。

[0057] 为了执行帧内预测,视频编码器200可以选择帧内预测模式来生成预测块。JEM和VVC的一些示例提供了六十七种帧内预测模式,包括各种方向的模式以及平面模式和DC模式。一般来说,视频编码器200选择帧内预测模式,其描述当前块(例如,CU的块)的邻近样本,以从其预测当前块的样本。假设视频编码器200以光栅扫描顺序(从左到右、从上到下)对CTU和CU进行编码,则此类样本一般可以在与当前块相同的图片中在当前块的上方、上左

侧或左侧。

[0058] 视频编码器200对表示当前块的预测模式的数据进行编码。举例来说,对于帧间预测模式,视频编码器200可以对数据进行编码,该数据表示使用了各种可用帧间预测模式中的哪一种以及对应模式的运动信息。对于单向或双向帧间预测,例如视频编码器200可以使用高级运动矢量预测 (AMVP) 或合并模式来对运动矢量进行编码。视频编码器200可以使用类似模式来对仿射运动补偿模式的运动矢量进行编码。

[0059] 在预测 (诸如块的帧内预测或帧间预测) 之后,视频编码器200可以计算块的残差数据。残差数据 (诸如残差块) 表示块与该块使用对应预测模式形成的预测块之间的逐样本差。视频编码器200可以将一个或多个变换应用于残差块以在变换域而非样本域中产生变换数据。举例来说,视频编码器200可以将离散余弦变换 (DCT)、整数变换、小波变换或概念上类似的变换应用于残差视频数据。另外,视频编码器200可以在一次变换之后应用二次变换,诸如取决于模式的不可分二次变换 (MDNSST)、取决于信号的变换、Karhunen-Loeve变换 (KLT) 等。视频编码器200在应用一个或多个变换之后产生变换系数。

[0060] 如上所述,在进行任何变换以产生变换系数之后,视频编码器200可以对变换系数执行量化。量化一般是指对变换系数进行量化来可能地减少用于表示变换系数的数据量,从而提供进一步的压缩的过程。通过执行量化过程,视频编码器200可以减小与系数中的一些或全部相关联的位深度。例如,视频编码器200可以在量化期间将n比特值四舍五入为m比特值,其中n大于m。在一些示例中,为了执行量化,视频编码器200可以对待量化的值执行按位右移。

[0061] 量化之后,视频编码器200可以扫描变换系数,从包括量化变换系数的二维矩阵中产生一维矢量。可以将扫描设计为将较高能量 (因此频率较低) 的系数放在矢量的前面,并将较低能量 (因此频率较高) 的变换系数放在矢量的后面。在一些示例中,视频编码器200可以利用预定义的扫描顺序来对量化变换系数进行扫描以产生序列化的矢量,然后对矢量的量化变换系数进行熵编码。在其他示例中,视频编码器200可以执行自适应扫描。在对量化变换系数进行扫描以形成一维矢量之后,视频编码器200可以例如根据上下文自适应二进制算术编码 (CABAC) 来对一维矢量进行熵编码。视频编码器200还可以熵编码用于语法元素的值,语法元素描述与视频解码器300在解码视频数据中使用的编码视频数据相关联的元数据。

[0062] 为了执行CABAC,视频编码器200可以将上下文模型内的上下文分配给待发送的符号。例如,上下文可以涉及符号的邻近值是否是零值。概率确定可以是基于分配给符号的上下文。

[0063] 视频编码器200可以进一步地生成例如在图片标头、块标头、切片标头中到视频解码器300的语法数据,诸如基于块的语法数据、基于图片的语法数据以及基于序列的语法数据,或者生成其他语法数据,诸如序列参数集 (SPS)、图片参数集 (PPS) 或视频参数集 (VPS)。视频解码器300可以类似地对此类语法数据进行解码以确定如何解码对应视频数据。

[0064] 以此方式,视频编码器200可以生成包括经编码视频数据的比特流,例如,描述将图片分割成块 (例如CU) 的语法元素以及块的预测和/或残差信息。最终,视频解码器300可以接收比特流并且对编码视频数据进行解码。

[0065] 一般来说,视频解码器300执行由视频编码器200执行的相反过程以对比特流的经

编码视频数据进行解码。例如,视频解码器300可以使用CABAC,以与视频编码器200的CABAC编码过程基本相似(尽管与之相反)的方式,对比特流的语法元素的值进行解码。语法元素可以定义图片的分割信息为CTU,并根据诸如QTBT结构的对应分割结构对每一个CTU进行分割以定义CTU的CU。语法元素可以进一步定义视频数据的块(例如,CU)的预测和残差信息。

[0066] 例如,可以由量化变换系数表示残差信息。视频解码器300可以逆量化和逆变换块的量化变换系数以再现该块的残差块。视频解码器300使用信号通知的预测模式(帧内或帧间预测)和相关的预测信息(例如,用于帧间预测的运动信息)来形成块的预测块。然后,视频解码器300可以组合预测块和残差块(在逐样本的基础上)以再现原始块。视频解码器300可以执行附加过程(诸如执行解块过程)来减少沿块边界的视觉伪影。

[0067] 根据本公开的技术,视频编码器200和视频解码器300可以被配置为确定使用双向预测编码的当前视频数据块的运动矢量,基于用于当前块的双向预测的预测的权重来确定对运动矢量使用运动矢量细化过程,并且基于使用运动矢量细化过程的确定对运动矢量应用运动矢量细化过程。

[0068] 本公开一般可以指“用信令通知”诸如语法元素的某些信息。术语“用信令通知”一般可以指语法元素的值和/或用于对编码的视频数据进行解码的其他数据的通信。也就是说,视频编码器200可以在比特流中发信号通知语法元素的值。通常,发信号通知是指在比特流中生成值。如上所述,源设备102可以基本上实时地或不实时地将比特流传输到目的地设备116,例如在将语法元素存储到存储设备112以供稍后由目的地设备116检索时可能发生。

[0069] 图2是说明可执行本公开的技术的示例视频编码器200的框图。提供图2是为了解释的目的,而不应被认为是对本公开中广泛例示和描述的技术的限制。为了解释的目的,本公开在诸如HEVC视频编码标准和开发中的H.266视频编码标准的视频编码标准的上下文中描述了视频编码器200。然而,本公开的技术不限于这些视频编码标准,并且通常可应用于视频编码和解码。

[0070] 在图2的示例中,视频编码器200包括视频数据存储器230、模式选择单元202、残差生成单元204、变换处理单元206、量化单元208、逆量化单元210、逆变换处理单元212、重建单元214、滤波器单元216、解码图片缓冲器(DPB)218和熵编码单元220。视频数据存储器230、模式选择单元202、残差生成单元204、变换处理单元206、量化单元208、逆量化单元210、逆变换处理单元212、重建单元214、滤波器单元216、DPB218和熵编码单元220中的任一个或全部可以在一个或多个处理器或处理电路中实现。此外,视频编码器200可以包括附加的或替代的处理器或处理电路以执行这些和其他功能。

[0071] 视频数据存储器230可以存储将由视频编码器200的组件编码的视频数据。视频编码器200可以从例如视频源104(图1)接收存储在视频数据存储器230中的视频数据。DPB 218可以充当参考图片存储器,该参考图片存储器存储参考视频数据,以供视频编码器200预测后续视频数据时使用。视频数据存储器230和DPB 218可以由诸如动态随机存取存储器(DRAM)(包括同步DRAM(SDRAM)、磁阻式RAM(MRAM)、电阻式RAM(RRAM)或其他类型的存储设备)的多种存储设备中的任何一个形成。视频数据存储器230和DPB 218可以由相同的存储设备或分离的存储设备提供。在各种示例中,视频数据存储器230可以与视频编码器200的其他组件一起在芯片上,如图所示,或者相对于那些组件在芯片外。

[0072] 在本公开中,对视频数据存储器230的引用不应被解释为限于视频编码器200内部的存储器,除非明确地这样描述,或者不应解释为限于视频数据存储器200外部的存储器,除非明确地这样描述。相反,对视频数据存储器230的引用应该被理解为指存储视频编码器200所接收的用于编码的视频数据(例如,要编码的当前块的视频数据)的存储器。图1的存储器106也提供对于来自视频编码器200的各种单元的输出的临时存储。

[0073] 图2的各种单元被说明以帮助理解由视频编码器200执行的操作。这些单元可以被实现为固定功能电路、可编程电路或其组合。固定功能电路是指提供特定功能并被预设在其可以执行的操作上的电路。可编程电路是指能够被编程以执行各种任务并在可以执行的操作中提供灵活功能性的电路。例如,可编程电路可以执行软件或固件,使可编程电路以软件或固件的指令所定义的方式操作。固定功能电路可以执行软件指令(例如,以接收参数或输出参数),但是固定功能电路执行的操作类型通常是不可变的。在一些示例中,一个或多个单元可以是不同的电路块(固定功能或可编程),并且在一些示例中,一个或多个单元可以是集成电路。

[0074] 视频编码器200可包括由可编程电路形成的算术逻辑单元(ALU)、基本功能单元(EFU)、数字电路、模拟电路和/或可编程核。在使用由可编程电路执行的软件来执行视频编码器200的操作的示例中,存储器106(图1)可以存储视频编码器200接收并执行的软件的目标代码,或者在视频编码器200内的另一个存储器(未示出)可以存储此类指令。

[0075] 视频数据存储器230被配置为存储接收到的视频数据。视频编码器200可以从视频数据存储器230检索视频数据的图片,并将视频数据提供给残差生成单元204和模式选择单元202。视频数据存储器230中的视频数据可以是将被编码的原始视频数据。

[0076] 模式选择单元202包括运动估计单元222、运动补偿单元224和帧内预测单元226。模式选择单元202可以包括附加功能单元,以根据其他预测模式来执行视频预测。作为示例,模式选择单元202可以包括调色板单元、块内复制单元(其可以是运动估计单元222和/或运动补偿单元224的一部分)、仿射单元、线性模型(LM)单元等。

[0077] 模式选择单元202通常协调多个编码通道(pass),以测试编码参数的组合以及针对这种组合的结果速率失真值。编码参数可以包括将CTU划分为CU、CU的预测模式、CU的残差数据的变换类型、CU的残差数据的量化参数等。模式选择单元202可以最终选择具有比其他测试的组合更好的速率失真值的编码参数的组合。

[0078] 视频编码器200可以将视频数据存储器230检索的图片划分为一系列CTU,并将一个或多个CTU封装在切片中。模式选择单元202可以根据诸如上述的QTBT结构或HEVC的四叉树结构的树结构来划分图片的CTU。如上所述,视频编码器200可以根据树结构通过划分CTU来形成一个或多个CU。这样的CU也通常可以被称为“视频块”或“块”。

[0079] 通常,模式选择单元202还控制其组件(例如,运动估计单元222、运动补偿单元224和帧内预测单元226)以生成针对当前块(例如,当前CU,或HEVC中,PU和TU的重叠部分)的预测块。对于当前块的帧间预测,运动估计单元222可执行运动搜索以识别一个或多个参考图片(例如,存储在DPB 218中的一个或多个先前编码的图片)中的一个或多个紧密匹配的参考块。特别地,运动估计单元222可以例如根据绝对差之和(SAD)、平方差之和(SSD)、平均绝对差(MAD)、均方差(MSD)等来计算表示潜在参考块与当前块有多相似的值。运动估计单元222通常可以使用当前块与所考虑的参考块之间的逐样本差来执行这些计算。运动估计单

元222可识别具有由这些计算产生的最低值的参考块,其指示与当前块最紧密匹配的参考块。

[0080] 运动估计单元222可以形成一个或多个运动矢量(MV),其定义参考图片中的参考块相对于当前图片中的当前块的位置。然后,运动估计单元222可以将运动矢量提供给运动补偿单元224。例如,对于单向帧间预测,运动估计单元222可提供单个运动矢量,而对于双向帧间预测(双向预测的预测),运动估计单元222可以提供两个运动矢量。运动补偿单元224然后可以使用运动矢量来产生预测块。例如,运动补偿单元224可以使用运动矢量来检索参考块的数据。作为另一示例,如果运动矢量具有分数采样精度,则运动补偿单元224可以根据一个或多个插值滤波器为预测块插值。此外,对于双向帧间预测,运动补偿单元224可以例如通过逐样本平均或加权平均来检索由相应运动矢量识别的两个参考块的数据,并组合所检索的数据。

[0081] 根据将在下文更详细描述的本公开的技术,运动补偿单元224可以经配置以使用双向预测的预测来编码视频数据块。如上所述,双向预测的预测可以使用来自两个参考图片列表的两个运动矢量。在一些示例中,运动补偿单元224还可将运动矢量细化过程(例如,DMVR)应用于双向预测的预测中使用的运动矢量中的一个或多个。示例DMVR技术在下面被讨论,并且可以包括合并运动矢量细化、合并偏移扩展、双边模板匹配、参数子像素运动矢量推导和双向光流中的一个或多个。本公开描述用于在使用双向预测的预测对视频数据块进行编码时确定何时使用DMVR技术(例如,双边模板匹配)的技术。

[0082] 在本公开的一个示例中,运动补偿单元224可以被配置为确定使用双向预测的预测来编码的当前视频数据块的运动矢量,基于用于当前块的双向预测的预测的权重来确定对运动矢量使用运动矢量细化过程,并且基于使用运动矢量细化过程的确定对运动矢量应用运动矢量细化过程。举例来说,当用于当前块的双向预测的预测的权重不相等时,运动补偿单元224可以确定禁用运动矢量细化过程。在其他示例中,运动补偿单元224可以被配置为确定由运动矢量识别的预测块之间的差,并且基于用于当前块的双向预测的预测的权重和由运动矢量识别的预测块之间的确定的差来确定对运动矢量使用运动矢量细化过程。举例来说,运动补偿单元224可经配置以在由运动矢量识别的预测块之间的差小于阈值的情况下确定禁用运动矢量细化过程。

[0083] 作为另一示例,对于帧内预测或帧内预测编码,帧内预测单元226可以从邻近当前块的样本来生成预测块。例如,对于定向模式,帧内预测单元226通常可以在数学上组合邻近样本的值,并跨当前块上沿定义的方向填充这些计算出的值以产生预测块。作为另一示例,对于DC模式,帧内预测单元226可以计算到当前块的邻近样本的平均值,并且生成预测块以包括针对预测块的每个样本的该所得平均值。

[0084] 模式选择单元202将预测块提供给残差生成单元204。残差生成单元204从视频数据存储单元230接收当前块的原始的未编码版本,并从模式选择单元202接收预测块。残差生成单元204计算当前块与预测块之间的逐样本差。所得的逐样本差定义了当前块的残差块。在一些示例中,残差生成单元204还可以确定残差块中的样本值之间的差以使用残差差分脉冲编码调制(RDPCM)来生成残差块。在一些示例中,可以使用执行二进制减法的一个或多个减法器电路来形成残差生成单元204。

[0085] 在模式选择单元202将CU分割为PU的示例中,每个PU可以与亮度预测单元和对应

的色度预测单元相关联。视频编码器200和视频解码器300可以支持具有各种尺寸的PU。如上所述, CU的尺寸可以指CU的亮度编码块的尺寸, 而PU的尺寸可以指PU的亮度预测单元的尺寸。假设特定CU的尺寸为 $2N \times 2N$, 则视频编码器200可支持用于帧内预测的 $2N \times 2N$ 或 $N \times N$ 的PU尺寸, 以及用于帧间预测的 $2N \times 2N$ 、 $2N \times N$ 、 $N \times 2N$ 、 $N \times N$ 或类似的对称PU尺寸。视频编码器200和视频解码器300还可针对用于帧间预测的 $2N \times nU$ 、 $2N \times nD$ 、 $nL \times 2N$ 和 $nR \times 2N$ 的PU尺寸支持非对称分割。

[0086] 在模式选择单元不将CU进一步分割成PU的示例中, 每个CU可以与亮度编码块和对应的色度编码块相关联。如上所述, CU的尺寸可以指CU的亮度编码块的尺寸。视频编码器200和视频解码器300可以支持 $2N \times 2N$ 、 $2N \times N$ 或 $N \times 2N$ 的CU尺寸。

[0087] 对于其他视频编码技术, 例如作为几个例子的块内复制模式编码、仿射模式编码和线性模型(LM)模式编码, 模式选择单元202经由与编码技术相关联的各个单元生成用于正被编码的当前块的预测块。在一些示例中, 例如调色板模式编码, 模式选择单元202可以不生成预测块, 而是生成语法元素, 该语法元素指示基于所选调色板重建块的方式。在这样的模式中, 模式选择单元202可以将这些语法元素提供给熵编码单元220以进行编码。

[0088] 如上所述, 残差生成单元204接收当前块和对应的预测块的视频数据。残差生成单元204然后为当前块生成残差块。为了生成残差块, 残差生成单元204计算预测块与当前块之间的逐样本差。

[0089] 变换处理单元206将一个或多个变换应用于残差块以生成变换系数的块(这里称为“变换系数块”)。变换处理单元206可以将各种变换应用于残差块以形成变换系数块。例如, 变换处理单元206可以将离散余弦变换(DCT)、方向变换、Karhunen-Loeve变换(KLT)或概念上类似的变换应用于残差块。在一些示例中, 变换处理单元206可以对残差块执行多个变换, 例如主变换和诸如旋转变换的副变换。在一些示例中, 变换处理单元206不将变换应用于残差块。

[0090] 量化单元208可以量化变换系数块中的变换系数, 以产生量化的变换系数块。量化单元208可以根据与当前块相关联的量化参数(QP)值来量化变换系数块的变换系数。视频编码器200(例如, 经由模式选择单元202)可以通过调整与CU相关联的QP值来调整应用于与当前块相关联的系数块的量化程度。量化可能会引入信息丢失, 因此, 量化的变换系数的精度可能会比变换处理单元206产生的原始变换系数的精度低。

[0091] 逆量化单元210和逆变换处理单元212可以将逆量化和逆变换分别应用于量化的变换系数块, 以从变换系数块重建残差块。重建单元214可以基于重建的残差块和由模式选择单元202生成的预测块来产生与当前块相对应的重建块(尽管可能具有一定程度的失真)。例如, 重建单元214可将重建的残差块的样本添加到来自由模式选择单元202所产生的预测块的对应样本中, 以生成重建块。

[0092] 滤波器单元216可以对重建块执行一个或多个滤波器操作。例如, 滤波器单元216可以执行解块操作以减少沿着CU的边缘的块效应伪影(blockiness artifact)。在一些示例中, 可以跳过滤波器单元216的操作。

[0093] 视频编码器200将重建块存储在DPB 218中。例如, 在不需要滤波器单元216的操作的示例中, 重建单元214可以将重建块存储到DPB 218中。在需要滤波器单元216的操作的示例中, 滤波器单元216可以将滤波后的重建块存储到DPB 218中。运动估计单元222和运动补

偿单元224可以从DPB 218检索参考图片,该参考图片由重建(并可能被滤波)的块形成,以对随后编码的块进行帧间预测。另外,帧内预测单元226可以使用当前图片的DPB 218中的重建块来对当前图片中的其他块进行帧内预测。

[0094] 通常,熵编码单元220可以对从视频编码器200的其他功能组件接收的语法元素进行熵编码。例如,熵编码单元220可以对来自量化单元208的量化的变换系数块进行熵编码。作为另一示例,熵编码单元220可以对来自模式选择单元202的预测语法元素(例如,用于帧间预测的运动信息或用于帧内预测的模式内信息)进行熵编码。熵编码单元220可以对作为视频数据的另一个示例的语法元素执行一个或多个熵编码操作,以产生熵编码的数据。例如,熵编码单元220可以对数据执行上下文自适应可变长度编码(CAVLC)操作、CABAC操作、可变至可变(V2V)长度编码操作、基于语法的上下文自适应二进制算术编码(SBAC)操作、概率区间分割熵(PIPE)编码操作、指数哥伦布编码操作或另一类型的熵编码操作。在一些示例中,熵编码单元220可以在语法元素未被熵编码的绕过模式中操作。

[0095] 视频编码器200可以输出包括重建切片或图片的块所需的熵编码语法元素的比特流。特别地,熵编码单元220可以输出比特流。

[0096] 以上描述的操作是针对块描述的。这样的描述应该被理解为用于亮度编码块和/或色度编码块的操作。如上所述,在一些示例中,亮度编码块和色度编码块是CU的亮度和色度分量。在一些示例中,亮度编码块和色度编码块是PU的亮度和色度分量。

[0097] 在一些示例中,不需要针对色度编码块重复针对亮度编码块执行的操作。作为一个示例,不需要重复识别用于亮度编码块的运动矢量(MV)和参考图片的操作来识别用于色度编码块的MV和参考图片。而是可以缩放用于亮度编码块的MV以确定用于色度块的MV,并且参考图片可以是相同的。作为另一示例,对于亮度编码块和色度编码块,帧内预测处理可以是相同的。

[0098] 视频编码器200表示被配置为对视频数据进行编码的设备的示例,该设备包括被配置为存储视频数据的存储器,以及在电路中实现的一个或多个处理单元,该一个或多个处理单元被配置为确定视频的当前块的运动矢量,基于用于当前块的双向预测的预测的权重来确定对运动矢量使用运动矢量细化过程,并基于使用运动矢量细化过程的确定来细化运动矢量。

[0099] 图3为说明可执行本公开的技术的示例视频解码器300的框图。提供图3是为了解释的目的,并且不限制本公开中广泛示例和描述的技术。为了解释的目的,本公开描述了根据JEM和HEVC的技术的视频解码器300。然而,本公开的技术可以由被配置为其他视频编码标准的视频编码设备来执行。

[0100] 在图3的示例中,视频解码器300包括编码图片缓冲器(CPB)存储器320、熵解码单元302、预测处理单元304、逆量化单元306、逆变换处理单元308、重建单元310、滤波器单元312和解码图片缓冲器(DPB)314。CPB存储器320、熵解码单元302、预测处理单元304、逆量化单元306、逆变换处理单元308、重建单元310、滤波器单元312和DPB 314中的任何一个或全部可以在一个或多个处理器或处理电路中实现。此外,视频解码器300可以包括附加的或替代的处理器或处理电路以执行这些和其他功能。

[0101] 预测处理单元304包括运动补偿单元316和帧内预测单元318。预测处理单元304可包括附加单元以根据其他预测模式执行预测。作为示例,预测处理单元304可以包括调色板

单元、帧内块复制单元(其可以形成运动补偿单元316的一部分)、仿射单元、线性模型(LM)单元等。在其他示例中,视频解码器300可以包括更多、更少或不同的功能组件。

[0102] CPB存储器320可以存储将由视频解码器300的组件解码的视频数据,例如经编码的视频比特流。例如,可以从计算机可读介质110(图1)获得存储在CPB存储器320中的视频数据。CPB存储器320可以包括存储来自编码的视频比特流的编码的视频数据(例如,语法元素)的CPB。而且,CPB存储器320可以存储除了编码的图像的语法元素之外的视频数据,例如表示来自视频解码器300的各单元的输出的临时数据。DPB 314通常存储解码图片,该解码图片可以被视频解码器300输出和/或在对编码的视频比特流的后续数据或图片进行解码时,作为参考视频数据。CPB存储器320和DPB 314可以由诸如动态随机存取存储器(DRAM)(包括同步DRAM(SDRAM)、磁阻式RAM(MRAM)、电阻式RAM(RRAM)或其他类型的存储设备)的各种存储设备中的任何一种形成。CPB存储器320和DPB 314可以由相同的存储设备或分离的存储设备提供。在各种示例中,CPB存储器320可以与视频解码器300的其他组件在芯片上,或者相对于那些组件在芯片外。

[0103] 附加地或可替代地,在一些示例中,视频解码器300可以从存储器120(图1)检索编码的视频数据。也就是说,存储器120可像上文关于存储器320讨论的那样存储数据。同样地,当视频解码器300的一些或全部功能在软件中实现以通过视频解码器300的处理电路来执行时,存储器120可存储将由视频解码器300执行的指令。

[0104] 图3中所示的各种单元被说明以帮助理解由视频解码器300执行的操作。这些单元可以被实现为固定功能电路、可编程电路或其组合。与图2类似,固定功能电路是指提供特定功能性并且被预设在其可以执行的操作上的电路。可编程电路是指能够编程以执行各种任务并在可以执行的操作中提供灵活功能性的电路。例如,可编程电路可以执行软件或固件,使可编程电路以软件或固件的指令所定义的方式操作。固定功能电路可以执行软件指令(例如,以接收参数或输出参数),但是固定功能电路执行的操作类型通常是不可变的。在一些示例中,一个或多个单元可以是不同的电路块(固定功能或可编程),并且在一些示例中,一个或多个单元可以是集成电路。

[0105] 视频解码器300可以包括ALU、EFU、数字电路、模拟电路和/或由可编程电路形成的可编程核。在视频解码器300的操作由在可编程电路上执行的软件执行的示例中,片上或片外存储器可以存储视频解码器300接收并执行的软件的指令(例如,目标代码)。

[0106] 熵解码单元302可从CPB接收编码的视频数据,并对视频数据进行熵解码以再现语法元素。预测处理单元304、逆量化单元306、逆变换处理单元308、重建单元310和滤波器单元312可以基于从比特流中提取的语法元素来生成解码的视频数据。

[0107] 通常,视频解码器300在逐块的基础上重建图片。视频解码器300可单独地对每个块执行重建操作(其中当前正在被重建(即,被解码)的块可被称为“当前块”)。

[0108] 熵解码单元302可对定义量化的变换系数块的量化的变换系数的语法元素以及诸如量化参数(QP)和/或(一个或多个)变换模式指示的变换信息进行熵解码。逆量化单元306可以使用与量化的变换系数块相关联的QP来确定量化度,并且同样地,确定逆量化度以供逆量化单元306应用。逆量化单元306可以例如执行按位左移运算以对量化的变换系数进行逆量化。逆量化单元306从而可以形成包括变换系数的变换系数块。

[0109] 在逆量化单元306形成变换系数块之后,逆变换处理单元308可以将一个或多个逆

变换应用于变换系数块以生成与当前块相关联的残差块。例如,逆变换处理单元308可以将逆DCT、逆整数变换、逆Karhunen-Loeve变换(KLT)、逆旋转变换、逆方向变换或另一逆变换应用于系数块。

[0110] 此外,预测处理单元304根据由熵解码单元302熵解码的预测信息语法元素来生成预测块。例如,如果预测信息语法元素指示当前块是帧间预测的,则运动补偿单元316可以生成预测块。在这种情况下,预测信息语法元素可以指示DPB 314中的从其检索参考块的参考图片,以及识别参考图片中参考块相对于当前图片中的当前块的位置的运动矢量。运动补偿单元316通常可以以与关于运动补偿单元224(图2)所描述的方式基本相似的方式来执行帧间预测处理。

[0111] 根据将在下文更详细描述的本公开的技术,运动补偿单元316可经配置以使用双向预测的预测来解码视频数据块。如上所述,双向预测的预测可以使用来自两个参考图片列表的两个运动矢量。在一些示例中,运动补偿单元316还可将运动矢量细化过程(例如,DMVR)应用于双向预测的预测中使用的运动矢量中的一个或多个。示例DMVR技术在下面被讨论,并且可以包括合并运动矢量细化、合并偏移扩展、双边模板匹配、参数子像素运动矢量推导和双向光流中的一个或多个。本公开描述用于在使用双向预测的预测对视频数据块进行解码时确定何时使用DMVR技术(例如,双边模板匹配)的技术。

[0112] 在本公开的一个示例中,运动补偿单元316可以被配置为确定使用双向预测编码的视频数据的当前块的运动矢量,基于用于当前块的双向预测的预测的权重来确定对运动矢量使用运动矢量细化过程,并且基于使用运动矢量细化过程的确定对运动矢量应用运动矢量细化过程。举例来说,当用于当前块的双向预测的权重不相等时,运动补偿单元316可以确定禁用运动矢量细化过程。在其他示例中,运动补偿单元316可以被配置为确定由运动矢量识别的预测块之间的差,并且基于用于当前块的双向预测的预测的权重和所确定的由运动矢量识别的预测块之间的差来确定对运动矢量使用运动矢量细化过程。举例来说,运动补偿单元316可经配置以在由运动矢量识别的预测块之间的差小于阈值的情况下确定禁用运动矢量细化过程。

[0113] 作为另一示例,如果预测信息语法元素指示当前块是帧内预测的,则帧内预测单元318可以根据由预测信息语法元素指示的帧内预测模式来生成预测块。再次,帧内预测单元318通常可以以与关于帧内预测单元226(图2)所描述的方式基本上相似的方式来执行帧内预测处理。帧内预测单元318可以从DPB 314检索当前块的邻近样本的数据。

[0114] 重建单元310可以使用预测块和残差块来重建当前块。例如,重建单元310可以将残差块的样本添加到预测块的对应样本以重建当前块。

[0115] 滤波器单元312可以对重建块执行一个或多个滤波器操作。例如,滤波器单元312可执行解块操作以减少沿着重建块的边缘的块效应伪影。不一定在所有示例中都执行滤波器单元312的操作。

[0116] 视频解码器300可将重建的块存储在DPB 314中。如上所述,DPB 314可将参考信息(例如用于帧内预测的当前图片的样本以及用于后续运动补偿的先前解码图片)提供给预测处理单元304。此外,视频解码器300可从DPB输出解码图片,以用于随后在显示设备(诸如图1的显示设备118)上呈现。

[0117] 以这种方式,视频解码器300表示视频解码设备的示例,该视频解码设备包括被配

置为存储视频数据的存储器,以及在电路中实现的一个或多个处理单元,该一个或多个处理单元被配置为确定用于使用双向预测的预测来编码的视频数据的当前块的运动矢量,基于用于当前块的双向预测的预测的权重确定对运动矢量使用运动矢量细化过程,并基于确定使用运动矢量细化过程来对运动矢量应用运动矢量细化过程。

[0118] 在以下部分中,描述与本公开相关的HEVC和JEM技术以及多功能视频编码(VVC)中的正在进行的工作。

[0119] HEVC中的CU结构和运动矢量预测

[0120] 在HEVC中,切片中最大的编码单元称为编码树块(CTB)或编码树单元(CTU)。CTB包含四叉树,其节点是编码单元。

[0121] 在HEVC主配置文件中,CTB的尺寸可以从 16×16 到 64×64 (尽管从技术上讲,可以支持 8×8 CTB尺寸)。编码单元(CU)的尺寸可以与CTB的尺寸相同,也可以小至 8×8 。每个编码单元可以以一种模式编码。当CU被帧间编码时,CU可被进一步分割为2或4个预测单元(PU),或当不应用进一步的分割时,CU仅变为一个PU。当一个CU中存在两个PU时,两个PU能是一半尺寸的矩形,也可以是分别为CU的尺寸的 $1/4$ 或 $3/4$ 的两个矩形。

[0122] 当CU被帧间编码时,针对每个PU存在一组运动信息。另外,利用独特的帧间预测模式对每个PU进行编码以导出运动信息集合。

[0123] 运动矢量预测

[0124] 在HEVC标准中,对于预测单元(PU),存在两种帧间预测模式,分别称为合并(跳过被认为是无残差合并的特殊情况)和高级运动矢量预测(AMVP)模式。在AMVP或合并模式中,视频编码器200和视频解码器300可以构建和维护多个运动矢量预测器的运动矢量(MV)候选列表。视频编码器200和视频解码器300可以通过从MV候选列表中获取一个候选的运动信息来产生当前PU的(多个)运动矢量以及合并模式下的参考索引。

[0125] 在HEVC中,MV候选列表包括用于合并模式的多达5个候选和用于AMVP模式的仅两个候选。合并候选可以包含一组运动信息,例如,与参考图片列表(列表0和列表1)和参考索引两者相对应的运动矢量。如果通过合并索引识别了合并候选,则视频编码器200和视频解码器300被配置成使用参考图片用于当前块的预测。此外,视频编码器200和视频解码器300被配置来确定与合并索引相关联的运动矢量。但是,在AMVP模式下,对于来自列表0或列表1的每个潜在预测方向,由于AMVP候选仅包含运动矢量,因此可以将参考索引与到MV候选列表的MV预测器(MVP)索引一起明确地用信号通知。在AMVP模式中,可以进一步细化预测的运动矢量。

[0126] 如上文所见,合并候选对应于运动信息的完整集合,而AMVP候选仅包含用于特定预测方向和参考索引的一个运动矢量。两种模式的候选都类似地从相同的空间和时间邻近块中得出。

[0127] 空间邻近候选

[0128] 在HEVC中,对于特定的PU(PU_0),视频编码器200和视频解码器300可从邻近块导出空间MV候选,如图4A和图4B所示,尽管用于从块生成候选的技术对于合并和AMVP模式不同。

[0129] 图4A是示出用于合并模式的块500的空间邻近候选的概念图。图4B是示出用于AMVP模式的块502的空间邻近候选的概念图。在合并模式中,视频编码器200和视频解码器300可以按照图4A所示的顺序导出多达四个空间MV候选。顺序如下:左块(0,A1)、上方块(1,

B1)、右上方块(2,B0)、左下方块(3,A0)和右上方(4,B2)块。

[0130] 在AMVP模式下,视频编码器200和视频解码器300可以将邻近块分成两组:左侧组包括块0和1,上方组包括块2、3和4,如图4B所示。对于每个组,参考与用信号通知的参考索引指示的参考图片相同的参考图片的邻近块中的潜在候选具有最高优先级,以被选择来形成该组的最终候选。可能所有邻近块都不包含指向同一参考图片的运动矢量。因此,如果找不到这样的候选,则视频编码器200和视频解码器300可以缩放第一可用候选以形成最终候选。因此,能够补偿时间距离差。

[0131] HEVC中的时间运动矢量预测

[0132] 现在将讨论HEVC中的时间运动矢量预测。视频编码器200和视频解码器300可经配置以在空间运动矢量候选之后将时间运动矢量预测器(TMVP)候选(如果启用且可用)添加到MV候选列表中。TMVP候选的运动矢量推导过程对于合并和AMVP模式是相同的。但是,在HEVC中,合并模式中TMVP候选的目标参考索引设置为0。

[0133] 图5A示出了块604(PU0)的示例TMVP候选,并且图5B示出运动矢量缩放过程606。用于TMVP候选推导的主要块位置是并置PU之外的右下块。该候选示在图5A中示为块“T”。块T的位置用于补偿对用于生成空间邻近候选的上块和左块的偏差。然而,如果该块位于当前CTB行之外或运动信息不可用,则该块将替换为PU的中心块。

[0134] TMVP候选的运动矢量是从在切片级别中指示的并置图片的并置PU导出的。位于并置PU的运动矢量称为并列MV。类似于AVC中的时间直接模式,为了导出TMVP候选运动矢量,可以缩放并置MV以补偿时间距离差,如图5B所示。

[0135] 合并运动矢量细化

[0136] S.Jeong等人的“CE4 Ultimate Motion Vector Expression(Test 4.5.4),Joint Video Experts Team(JVET)of ITU-T SG 16WP 3and ISO/IEC JTC 1/SC 29/WG 11,12th Meeting,Macao,CN,3-12,October 2018,JVET-L0054,Merge Motion Vector Refinement(MMVR,a.k.a.Ultimate Motion Vector Expression,UMVE)”提出了基于发信号通知的运动矢量差来细化合并候选的运动矢量的技术。UMVE提供了具有简化信令的替代运动矢量表达,包括起点、运动幅度和运动方向。合并运动可以在以未细化合并运动矢量指向的位置为中心的十字形图案上使用图6A和图6B中所描绘的偏移之一来被细化。此外,指向列表L0中参考图片的MV偏移(例如,细化后的MV减去原始MV)被缩放到列表L1中的参考图片。

[0137] 图6A是说明示例UMVE搜索过程的概念图。如图6A所示,当前帧702中的当前块706的运动矢量708可以在运动幅度和运动方向两者方面被修改。运动幅度和运动方向的不同变化可以产生不同的运动矢量(例如,运动矢量710或运动矢量712)。运动矢量710和712可以被称为具有UMVE(或简称为UMVE候选)的合并候选(在这种情况下为运动矢量708)。即,运动矢量710和712是通过用UMVE改变运动矢量708来创建的。

[0138] 运动矢量708还可具有不同的预测方向。预测方向信息表示L0(帧700)、L1(帧704)、L0和L1预测中的预测方向。对于B切片,使用所提出的UMVE技术,视频编码器200和视频解码器300可以通过使用镜像技术从具有单预测的合并候选生成双预测候选,如图6A和图6B所示。图6B是说明示例UMVE搜索点的概念图。例如,如果合并候选是具有参考列表L1的单预测候选,则视频编码器200和视频解码器300可以通过搜索列表0中的参考图片来确定参考列表L0的参考索引,该参考图片与列表1的参考图片镜像。如果没有对应的图片,则视

频编码器200和视频解码器300使用与当前图片最近的参考图片。视频编码器200和视频解码器300可通过缩放L1 (帧704) 的运动矢量来导出L0 (帧700) 的运动矢量。视频编码器200和视频解码器300可使用图片顺序计数 (POC) 距离来计算缩放因子。

[0139] 合并偏移扩展

[0140] X.Chen等人的“CE 4:Merge Offset Extension (Test 4.4.8), Joint Video Experts Team (JVET) of ITU-T SG 16WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 12th Meeting, Macao, CN, 3-12, October 2018, JVET-L0176”描述了仅基于合并候选列表的第一个候选构建新扩展MV偏移候选的技术。新扩展MV偏移候选仅包括对第一候选的当前MV的MV偏移。其他预测信息与第一候选相同。新添加的候选在时间候选之后被推入合并候选列表中。支持的运动矢量偏移量在图7中被描述, 包括相对于当前样本800具有偏移 (0或 ± 1 , 0或 ± 1) 的点802和相对于当前样本800具有偏移 (0或 ± 2 , 0或 ± 2) 的点804。

[0141] 对于来自过去 (例如, 列表0) 的参考图片的一个MV (例如, MV0) 和来自未来 (例如, 列表1) 的参考图片的另一个 (例如, MV1) 的双向预测, 视频编码器200和视频解码器300可以将选择的MV偏移添加到第一候选MV0, 并且将逆MV偏移添加到第一候选MV1。其他双向预测情况包括将相同的MV偏移量分别添加到第一候选MV0和MV1。

[0142] 基于历史的运动矢量预测

[0143] 基于历史的运动矢量预测 (HMVP) (例如, 如L.Zhang等人的“CE4-related: History-based Motion Vector Prediction,” Joint Video Experts Team (JVET) of ITU-T SG 16WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 11th Meeting, Ljubljana, SI, 10-18, July 2018 [JVET-K0104]) 是基于历史的技术, 它允许每个块找到它的来自除了直接邻近的因果邻近运动场中的那些MV以外还有过去解码的MV的列表的MV预测器。在编码/解码过程中维护一个具有多个HMVP候选的表。当遇到新切片时, 该表被清空。每当有帧间编码块时, 相关的运动信息就会以先进先出 (FIFO) 的方式作为新的HMVP候选插入到表中。然后, 可以应用约束FIFO规则。在向表中插入HMVP时, 首先应用冗余检查以查找表中是否存在相同的HMVP。如果找到相同的HMVP, 则从表中移除该特定HMVP, 然后移除所有HMVP候选。

[0144] HMVP候选也可用于合并候选列表构建过程中。表中从最后一个条目到第一个条目的所有HMVP候选都插入在TMVP候选之后。修剪适用于HMVP候选。一旦可用合并候选的总数达到用信号通知的最大允许合并候选数, 合并候选列表构建过程就终止。

[0145] 类似地, HMVP候选也可用于AMVP候选列表构建过程中。表中最后K个HMVP候选的运动矢量插入在TMVP候选之后。仅使用与AMVP目标参考图片具有相同参考图片的HMVP候选来构建AMVP候选列表。修剪适用于HMVP候选。

[0146] HEVC中运动预测的其他方面

[0147] 合并和AMVP模式的几个方面如下值得一提。

[0148] 运动矢量缩放: 假设运动矢量的值与呈现时间中图片的距离成比例。运动矢量将两个图片、参考图片和包含运动矢量的图片 (即包含图片) 相关联。当运动矢量被用于预测另一运动矢量时, 可以基于的图片顺序计数 (POC) 值来计算包含图片与参考图片之间的距离。

[0149] 对于要被预测的运动矢量, 其相关联的包含图片和参考图片两者可以不同。因此, 可以计算 (基于POC的) 新的距离。并且可以基于这两个POC距离来缩放运动矢量。对于空间

邻近候选,用于这两个运动矢量的包含图片是相同的,而参考图片可能不同。在HEVC中,运动矢量缩放适用于TMVP和AMVP的空间和时间邻近候选。

[0150] 人工运动矢量候选生成:如果运动矢量候选列表不完整(例如,具有比预定最大值少的候选),则生成人工运动矢量候选并将其插入到列表的末尾,直到列表满为止。

[0151] 在合并模式中,存在两种类型的人工MV候选:仅针对B切片导出的组合候选以及如果第一类型不提供足够的人工候选情况下的仅用于AMVP的零候选。

[0152] 对于已经在候选列表中并且具有必要运动信息的每对候选,通过参考列表0中的图片的第一候选的运动矢量与参考列表1中的图片的第二候选的运动矢量的组合来导出双向组合运动矢量候选。

[0153] 候选插入的修剪过程:来自不同块的候选可能恰好相同,这降低了合并/AMVP候选列表的效率。可以应用修剪过程来解决这个问题。修剪过程将当前候选列表中的一个候选与其他候选进行比较,以在一定程度上避免插入相同的候选。为了降低复杂性,只应用有限数量的修剪过程,而不是将每个潜在候选与所有其他现有候选进行比较。

[0154] 双边模板匹配

[0155] 图8是说明双边模板匹配的概念图。双边匹配是解码器侧运动矢量细化(DMVR)技术的一种变体,用于避免基于模板的细化过程。当前块900的双边模板匹配包括直接计算由初始双向预测MV(例如,图8中的 v_0 和 v_1)指向的单向预测参考块902与904之间的双边匹配成本(表示为 $I_0(x+v_0)$ 和 $I_1(x+v_1)$, x 为当前块900内的像素)。

[0156] 视频编码器200和视频解码器300可以基于在初始双向预测MV周围的预定义搜索范围内的双边匹配来执行局部搜索。具体来说,假设初始MV是 $v_0^{(0)}$ 和 $v_1^{(0)}$,在第一次搜索迭代时,若干MV对(例如 $v_0^{(0)} + \Delta$ 和 $v_1^{(0)} - \Delta$,其中 $\Delta \in \{(0,0), (-1,1), (0,1), (1,1), (1,0), (1,-1), (0,-1), (-1,-1), (-1,0)$ 等})被测试以确定 Δ^* 的最佳值,这导致最低的双边匹配成本。在本公开中,成本函数被定义为 $I_0(x+v_0^{(0)} + \Delta)$ 与 $I_1(x+v_1^{(0)} - \Delta)$ 之间的失真加上运动成本。失真函数可以是平均去除绝对差和(MRSAD)或绝对差和(SAD)。图8示出了在第一次搜索迭代中由测试的MV对之一指示的示例搜索块906和908。

[0157] 在找到最佳 Δ^* 之后,由视频编码器200和视频解码器300执行的迭代过程通过使用 Δ^* 来更新初始MV($v_0^{(0)}$ 和 $v_1^{(0)}$)的值。具体来说, $v_0^{(1)} = v_0^{(0)} + \Delta^*$ 和 $v_1^{(1)} = v_1^{(0)} - \Delta^*$ 。然后,在将上述描述中的所有上标推进1后,重复相同的迭代过程,直到达到 Δ^* 等于(0,0)。输出MV对(表示为 $v_0^{(n)}$ 和 $v_1^{(n)}$, $n \geq 1$)然后可以以子像素精度再次细化。然后将产生的MV对替换合并块的原始MV($v_0^{(0)}$ 和 $v_1^{(0)}$)。然后基于细化的MV执行运动补偿。

[0158] 参数子像素运动矢量推导

[0159] 在S.Sethurman等人的“Decoder Side MV Refinement/Derivation with CTB-level concurrency and other normative complexity reduction techniques,Joint Video Experts Team(JVET)of ITU-T SG 16WP 3and ISO/IEC JTC 1/SC 29/WG 11,11th Meeting,Ljubljana,SI,10-18,July 2018,JVET-K0041”中,二次参数函数用于为每个可能的分数像素MV形成预测误差曲面。基本上,二次参数函数是将预测误差的值作为估计量进行插值的插值函数。基于整数搜索的准确预测误差值,推导出二次参数函数的参数,从而可以找到该误差搜索的最佳运动采样位置。然后,原始MV被调整到这个精确的运动采样位置,而不是实际执行子像素运动矢量估计。该参数函数以五个点的成本值为参考,形成一个误

差曲面,并在该曲面上找到成本值最低的最佳位置。这五个点形成一个十字形,每两个邻近点之间的间隙为2个像素,其中中心/左/右/上/下点坐标为(0,0)/(-1,0)/(1,0)/(0,-1)/(0,1)。该参数误差曲面函数是一个二维抛物线误差曲面方程:

$$[0160] \quad E_{x,y} = A(x - \Delta x)^2 + B(y - \Delta y)^2 + C,$$

[0161] 其中, $(\Delta x, \Delta y)$ 对应成本最低的位置, C 对应成本值最小的位置。通过求解5个未知数 $(\Delta x, \Delta y)$ 中的5个方程,计算如下:

$$[0162] \quad \Delta x = \text{floor}((\alpha(E_{-1,0} - E_{1,0})) / (2(E_{-1,0} + E_{1,0} - 2E_{0,0}))),$$

$$[0163] \quad \Delta y = \text{floor}((\alpha(E_{0,-1} - E_{0,1})) / (2(E_{0,-1} + E_{0,1} - 2E_{0,0}))),$$

[0164] 其中 α 是一个整数比例因子,用于表示某个子像素精度中的 $(\Delta x, \Delta y)$,例如,16表示1/16精度,4表示1/4精度。

[0165] 使用解码器侧导出的运动矢量的运动矢量预测

[0166] 尽管解码器侧运动矢量推导(DMVD)在减少运动开销方面是有效的,但是由于空间邻近CU的编码之间的相互依赖性,现有DMVD设计(诸如DMVR)可能遇到解码延迟问题。如果从使用DMVR编码的其空间邻居预测CU的MV,则解码过程必须等待,直到邻近CU的细化MV变得可用。在多功能视频编码标准的开发过程中,研究了若干技术,这些技术可以为某些解码器侧MV推导(DMVD)方法实现低延迟设计。下面详细介绍其中一些技术。

[0167] 没有来自紧接在前的DMVD块的空间MV预测

[0168] 对于通过使用DMVD技术编码的块,解码过程可以解释为三个步骤:

[0169] 1. 初始运动场的重建和参考块的预取;

[0170] 2. 块运动的细化过程以获得最终的MV;和

[0171] 3. 利用最终MV进行运动补偿。

[0172] 在步骤2中的细化过程之后,最终MV被写回到图片运动场,因此细化后的MV可以用于空间MV预测、时间MV预测和边界强度计算。图9示出了DMVD技术(诸如DMVR)的流水线阶段的某个实现。在图9中,三个主要模块重建和预取1000、MV细化1002和运动补偿1004表示DMVD技术的三个解码步骤。模块1000、1002和1004可以是视频解码器300的运动补偿单元316的一部分(见图3)。如图所9示,在时间 T_0 ,对具有原始运动矢量 V_0 的 CU_0 进行重建和预取。在时间 T_1 ,原始运动矢量 V_0 被细化以创建细化运动矢量 V_0 。在时间 T_3 ,细化运动矢量 V_0 用于 CU_1 的重建和预取。同样在时间 T_3 ,使用细化运动矢量 V_0 对 CU_0 执行运动补偿。对于未来的CU,该级联过程将继续进行。

[0173] 一般而言, CU_{prev} 是在当前CU(CU_{cur})之前的先前编码的CU。在重建 CU_{cur} 的原始MV时,如果MV预测器来自 CU_{prev} ,而 CU_{prev} 恰好是DMVD编码块,则该预测器被标记为对 CU_{cur} 不可用。因此, CU_{cur} 的初始MV的重建不再依赖于 CU_{prev} 的细化MV,并且在某种程度上消除了MV细化与MV预测之间的相互依赖性。

[0174] 没有来自当前帧中的DMVD块的空间MV预测

[0175] 该技术不使用细化的MV,而是使用每个DMVR CU的原始MV来导出空间MV预测器。对于时间MV预测,可以使用细化的MV,而不会有解码延迟问题,因为并置图片已被完全重建。相应地,由于空间邻近CU之间的编码依赖性不再存在,因此,DMVR的解码延迟问题可以完全被解决。然而,可以预期编码性能的下降。

[0176] 没有来自当前CTU行中的DMVD块的空间MV预测

[0177] 如果这些邻近块与当前块一起落入同一CTU行,则该技术将未细化的MV作为来自直接邻近块的参考以执行空间MV预测。(注意,一些其他技术可能会将来自此类邻近块的MV预测器标记为不可用。)相反,仅当它们的关联块落在位于紧接上方CTU和左上方CTU的邻近因果CTU内时,细化MV才可用于空间MV预测。因此,该技术打破了CTU行内MV细化和空间MV预测之间的相互依赖性。

[0178] JEM中的双向光流

[0179] 双向光流(BIO)是逐像素运动细化,在双向预测的情况下,它在逐块运动补偿之上执行。由于BIO补偿块内的精细运动,因此启用BIO会导致运动补偿的块尺寸增大。样本级运动细化不需要详尽的搜索或信令,因为有一个明确的方程可以给出每个样本的精细运动矢量。图10是说明光流轨迹1100的概念图。

[0180] $I^{(k)}$ 是补偿块运动后参考k ($k=0,1$) 的亮度值,并且 $\partial I^{(k)}/\partial x, \partial I^{(k)}/\partial y$ 分别是 $I^{(k)}$ 梯度的水平和垂直分量。假设光流有效,则运动矢量场 (v_x, v_y) 由以下方程给出:

$$[0181] \quad \partial I^{(k)}/\partial t + v_x \partial I^{(k)}/\partial x + v_y \partial I^{(k)}/\partial y = 0. \quad (1)$$

[0182] 将光流方程与每个样本的运动轨迹的Hermite插值相结合,最后获得一个唯一的三阶多项式,该多项式与函数值 $I^{(k)}$ 和导数 $\partial I^{(k)}/\partial x, \partial I^{(k)}/\partial y$ 相匹配。该多项式在 $t=0$ 时的值是BIO预测:

$$[0183] \quad pred_{BIO} = 1/2 \cdot (I^{(0)} + I^{(1)} + v_x/2 \cdot (\tau_1 \partial I^{(1)}/\partial x - \tau_0 \partial I^{(0)}/\partial x) + v_y/2 \cdot (\tau_1 \partial I^{(1)}/\partial y - \tau_0 \partial I^{(0)}/\partial y)) \quad (2)$$

[0184] 这里 τ_0 和 τ_1 表示到参考帧的距离,如图10所示。基于Ref0和Ref1的POC计算距离 τ_0 和 τ_1 :

[0185] $\tau_0 = \text{POC}(\text{current}) - \text{POC}(\text{Ref0})$, $\tau_1 = \text{POC}(\text{Ref1}) - \text{POC}(\text{current})$ 。如果两个预测都来自相同的时间方向(都来自过去或都来自未来),则符号不同 $\tau_0 \cdot \tau_1 < 0$ 。在这种情况下,仅当预测不是来自同一时刻($\tau_0 \neq \tau_1$)时,才应用BIO,两个参考区域都具有非零运动:

[0186] $(MV_{x_0}, MV_{y_0}, MV_{x_1}, MV_{y_1} \neq 0)$,

[0187] 并且块运动矢量与时间距离成正比:

[0188] $(MV_{x_0}/MV_{x_1} = MV_{y_0}/MV_{y_1} = -\tau_0/\tau_1)$ 。

[0189] 运动矢量场 (v_x, v_y) 是通过最小化点A和B(图10上的运动轨迹和参考帧平面的交叉点)中的值之间的差 Δ 来确定的。该模型仅使用 Δ 的局部泰勒展开的第一个线性项:

$$[0190] \quad \Delta = (I^{(0)} - I^{(1)}_0 + v_x (\tau_1 \partial I^{(1)}/\partial x + \tau_0 \partial I^{(0)}/\partial x) + v_y (\tau_1 \partial I^{(1)}/\partial y + \tau_0 \partial I^{(0)}/\partial y)) \quad (3)$$

[0191] 等式(1)中的所有值都取决于样本位置 (i', j') , 到目前为止已被省略。假设运动在局部周围样本中是一致的,以当前预测点 (i, j) 为中心的 $(2M+1) \times (2M+1)$ 方窗 Ω 内的 Δ 可以最小化如下:

$$[0192] \quad (v_x, v_y) = \underset{v_x, v_y}{\operatorname{argmin}} \sum_{[i', j'] \in \Omega} \Delta^2[i', j'] \quad (4)$$

[0193] 对于这个优化问题,可以使用首先在垂直方向上进行最小化然后在水平方向上进行最小化的简化解方案,这导致:

$$[0194] \quad v_x = (s_1 + r) > m? \text{clip3}\left(-thBIO, thBIO, -\frac{s_3}{(s_1+r)}\right): 0 \quad (5)$$

$$[0195] \quad v_y = (s_5 + r) > m? \text{clip3}\left(-thBIO, thBIO, -\frac{s_6-v_x s_2/2}{(s_5+r)}\right): 0 \quad (6)$$

[0196] 其中,

$$[0197] \quad s_1 = \sum_{[i',j] \in \Omega} (\tau_1 \partial I^{(1)} / \partial x + \tau_0 \partial I^{(0)} / \partial x)^2; s_3 = \sum_{[i',j] \in \Omega} (I^{(1)} - I^{(0)}) (\tau_1 \partial I^{(1)} / \partial x + \tau_0 \partial I^{(0)} / \partial x);$$

$$[0198] \quad s_2 = \sum_{[i',j] \in \Omega} (\tau_1 \partial I^{(1)} / \partial x + \tau_0 \partial I^{(0)} / \partial x) (\tau_1 \partial I^{(1)} / \partial y + \tau_0 \partial I^{(0)} / \partial y);$$

$$[0199] \quad s_5 = \sum_{[i',j] \in \Omega} (\tau_1 \partial I^{(1)} / \partial y + \tau_0 \partial I^{(0)} / \partial y)^2; s_6 = \sum_{[i',j] \in \Omega} (I^{(1)} - I^{(0)}) (\tau_1 \partial I^{(1)} / \partial y + \tau_0 \partial I^{(0)} / \partial y) \quad (7)$$

[0200] 为了避免除以零或除以很小的值,可以在等式(2)、(3)中引入正则化参数r和m,例如如下:

$$[0201] \quad r = 500 \cdot 4^{d-8} \quad (8)$$

$$[0202] \quad m = 700 \cdot 4^{d-8} \quad (9)$$

[0203] 这里d是输入视频的内部位深度。

[0204] 在某些情况下,由于噪音或不规则运动,BIO的MV团可能不可靠。因此,在BIO中,MV团的大小被剪裁到某个阈值thBIO。根据当前图片的所有参考图片是否都来自一个方向来确定阈值。如果当前图片的所有参考图片都来自一个方向,则阈值设置为 $12 \times 2^{14-d}$;否则,阈值设置为 $12 \times 2^{13-d}$ 。

[0205] 使用与HEVC运动补偿过程(2D可分离FIR)一致的操作,与运动补偿插值同时计算BIO的梯度。根据块运动矢量的小数部分,此2D可分离FIR的输入是与运动补偿过程和小数位置(fracX, fracY)相同的参考帧样本。在水平梯度的情况下,首先使用BIOfilterS垂直插值 $\partial I / \partial x$ 信号,该BIOfilterS对应于具有去缩放位移d-8的分数位置fracY;然后梯度过滤器BIOfilterG在水平方向上应用,对应于分数位置fracX与18-d的去缩放位移。在垂直梯度 $\partial I / \partial y$ 的情况下,使用BIOfilterG垂直应用第一梯度滤波器,该滤波器对应于具有去缩放位移d-8的分数位置fracY;然后在水平方向上使用BIOfilterS执行信号位移,该位移对应于具有18-d的去缩放位移的分数位置fracX。用于梯度计算BIOfilterG和信号位移BIOfilterF的插值滤波器的长度较短(6抽头),以保持较低的实现复杂度,同时实现理想的比特率和/或失真。

[0206] 图11是说明8x4电流块1200的双向光流(BIO)的概念图。图11特别示出了当前块1200的梯度计算的示例。对于8x4当前块1200,视频编码器获取运动补偿预测器并计算当前块内的所有像素的HOR/VER梯度以及外部的两行像素。这是因为求解每个像素的vx和vy需要HOR/VER梯度值和以每个像素为中心的窗口 Ω 内像素的运动补偿预测值,如等式(4)所示。在JEM中,此窗口的大小设置为5x5。因此,视频编码器获取运动补偿预测器并计算外部两行像素的梯度。

[0207] 在JEM中,当两个预测来自不同的参考图片时,BIO被应用于所有双向预测块。当为CU启用LIC时,禁用BIO。

[0208] 广义双向预测(GBi)

[0209] GBi由C.Chen提出(“Generalized bi-prediction for inter coding,”Joint Video Experts Team(JVET)of ITU-T SG 16WP 3and ISO/IEC JTC 1/SG 29/WG 11,3rd Meeting:Geneva,CH,26May-1June 2016,JVET-C0047.In Y.Su,et.al.,“CE4.4.1: Generalized bi-prediction for inter coding,”Joint Video Experts Team(JVET)of ITU-T SG 16WP 3and ISO/IEC JTC 1/SC 29/WG 11,11th Meeting,Ljubljana,SI,10-18, July 2018,JVET-K0248),该提案改进了GBi的增益复杂度权衡,并被VVC的Benchmark Set 2.1(BMS2.1)采用。BMS2.1 GBi在双向预测模式下对来自L0和L1的预测器应用不相等的权重。在帧间预测模式下,基于速率失真优化(RDO)评估多个权重对,包括相等权重对(1/2,1/2),并将所选权重对的GBi索引信号发送给视频解码器300。在合并模式下,GBi索引继承自邻近CU。在BMS2.1 GBi中,双向预测模式下的预测器生成如下所示。

$$P_{GBi} = (w_0 * P_{L0} + w_1 * P_{L1} + \text{RoundingOffset}_{GBi}) \gg \text{shiftNum}_{GBi}, \quad (10)$$

[0211] 其中 P_{GBi} 是GBi的最终预测器。权重 w_0 和 w_1 是选定的GBi权重对,分别应用于列表0(L0)和列表1(L1)的预测器。 $\text{RoundingOffset}_{GBi}$ 和 shiftNum_{GBi} 用于标准化GBi中的最终预测器。支持的 w_1 权重集为 $\{-1/4, 3/8, 1/2, 5/8, 5/4\}$,其中五个权重对应一个相等权重对和四个不相等权重对。混合增益,即 w_1 和 w_0 的总和,固定为1.0。因此,对应的 w_0 权重集为 $\{5/4, 5/8, 1/2, 3/8, -1/4\}$ 。权重对选择在CU级别。

[0212] 对于非低延迟图片,权重集大小从5个减少到3个,其中 w_1 权重集为 $\{3/8, 1/2, 5/8\}$, w_0 权重集为 $\{5/8, 1/2, 3/8\}$ 。非低延迟图片的权重集大小缩减被应用于BMS2.1 GBi和JVET-K0248中的所有GBi测试。

[0213] 本公开描述以下用于解码器侧运动矢量细化的技术。

[0214] 1.1受限解码器侧运动矢量细化

[0215] 本公开提出如果用于双向预测的两个权重不相等,则禁用解码器侧运动细化。即,对于使用两个运动矢量编码的视频数据的双向预测块,视频编码器200和视频解码器300可以被配置为确定用于双向预测的权重是否相等。视频编码器200和视频解码器300可以确定要用于双向预测的权重。如果权重不相等,则视频编码器200和视频解码器300可能不对用于双向预测的预测的运动矢量应用(例如,禁用)运动矢量细化过程(例如,双边模板匹配)。即,视频编码器200和视频解码器300可在没有运动矢量细化的情况下使用双向预测的预测来编码视频数据块。如果权重相等,则视频编码器200和视频解码器300可以对用于双向预测的运动矢量应用(例如,启用)运动矢量细化过程(例如,双边模板匹配)。即,视频编码器200和视频解码器300可使用具有运动矢量细化的双向预测来编码视频数据块。

[0216] 例如,如果GBi被应用作为双向预测的预测并且权重 w_0 和 w_1 不相等,则视频编码器200和视频解码器300可以禁用解码器侧运动矢量细化过程。在另一示例中,如果加权双向预测的预测用于编码当前块并且权重 w_0 和 w_1 不相等,则视频编码器200和视频解码器300可禁用解码器侧运动矢量细化过程。在以上示例中,双边模板匹配可以用作运动矢量细化(例如,DMVR)过程。然而,本公开的技术可与任何运动矢量细化过程一起使用。

[0217] 因此,在本公开的一个示例中,视频编码器200和视频解码器300可以被配置为确定使用双向预测的预测来编码的视频数据的当前块的运动矢量,基于用于当前块的双向预测预测的权重来确定对运动矢量使用运动矢量细化过程,并且基于使用运动矢量细化过程的确定来对运动矢量应用运动矢量细化过程。在一个示例中,用于当前块的双向预测的权

重不相等。在该示例中,视频编码器200和视频解码器300可经配置以确定禁用运动矢量细化过程。

[0218] 1.2解码器侧运动矢量细化的自适应开/关

[0219] 第1.2节的技术可以与第1.1节的技术一起应用或可以独立应用。用于当前块的双向预测预测的(例如,来自列表0和列表1的)两个运动矢量指向一个或多个参考图片中的两个不同块。在本公开的另一示例中,如果使用两个原始运动矢量(即,在任何运动矢量细化之前)的两个预测块之间的差小于阈值,则视频编码器200和视频解码器300可不应用(例如,禁用)当前编码块的运动矢量细化过程(例如,DMVR)。

[0220] 视频编码器200和视频解码器300可经配置以使用两个预测块的绝对差之和(SAD)或平方误差之和(SSE)来测量预测块之间的“差”。当然,可以使用其他技术。在一些示例中,视频编码器200和视频解码器300可使用一些简化技术来跳过预测块中的一些样本以计算差。例如,视频编码器200和视频解码器300可以使用预测块的每隔一行的样本来计算预测块的SAD/SSE。

[0221] 视频编码器200和视频解码器300可以根据预测块中的样本数量和样本的位深度来确定用于确定是否应用运动矢量细化的“阈值”。例如,视频编码器200和视频解码器300可以将阈值计算为: $(N \ll (BD-8)) * W * H$ 。W和H分别是块的宽度和高度。BD是样本的位深度。N是预定义的值,例如1、2、4、8等。N也可以被用高级句法用信号通知。运算符 \ll 是按位右移。

[0222] 在另一种技术中,视频编码器200和视频解码器300可以根据用于计算SAD/SSE的预测块中的样本数量和样本的位深度来确定阈值。例如,如果SAD/SSE被每隔一行计算一次,则阈值是: $(N \ll (BD-8)) * W * H / 2$ 。

[0223] 因此,在本公开的另一示例中,视频编码器200和视频解码器300可经配置以确定由运动矢量识别的预测块之间的差。为了基于用于当前块的双向预测预测的权重确定对运动矢量使用运动矢量细化过程,视频编码器200和视频解码器300可以被配置为基于用于当前块的双向预测的权重的矢量以及由运动矢量识别的预测块之间的确定差来确定对运动矢量使用运动矢量细化过程。

[0224] 在另一示例中,为了基于用于当前块的双向预测预测的权重和由运动矢量识别的预测块之间的确定差来确定对运动矢量使用运动矢量细化过程,视频编码器200和视频解码器300可以被配置为在由运动矢量识别的预测块之间的差小于阈值的情况下确定禁用运动矢量细化过程。

[0225] 在另一示例中,为了确定由运动矢量识别的预测块之间的差,视频编码器200和视频解码器300可以被配置为使用预测块之间的绝对差的总和或预测块之间的平方差总和来确定由运动矢量识别的预测块之间的差。

[0226] 在本公开的另一示例中,视频编码器200和视频解码器300可以被配置为基于预测块中的样本数量和预测块中的样本的位深度来确定阈值。

[0227] 1.3双边模板匹配的运动补偿

[0228] 在双边模板匹配中,如上所述,视频编码器200和视频解码器300获得两个预测块。在非整数运动矢量的情况下,视频编码器200和视频解码器300通过插值获得预测块。在需要插值的一些示例中,诸如在HEVC中,视频编码器200和视频解码器300通过首先执行水平插值然后随后进行垂直插值来获得预测块。水平插值的输出是垂直插值的输入。视频编码

器200和视频解码器300可以通过对某个数量的样本进行加权平均来执行插值,即,将插值系数与对应的样本相乘,然后通过除以权重之和来归一化。

[0229] 在水平插值之后获得的中间数据可以存储在更高的位深度中以减少舍入误差。用于编码的内部位深度为 bitDepthI ,插值时(例如水平插值后)获得的中间数据的位深度为 bitDepthIF 。在一些示例中,视频编码器200和视频解码器300移位水平插值的结果(样本的加权和)以实现内部插值位深度 bitDepthIF 。在一些示例中,视频编码器200和视频解码器300可在移位之前应用一些偏移。这个移位数值可以称为 shift1 。垂直插值的结果也可以被移位以实现内部 bitDepth bitDepthI 。这个移位数值可以称为 shift2 。选择变量 precIF 的数量,使得权重之和小于或等于2的 precIF 次方。

[0230] 技术1

[0231] 在本公开的第一种技术中,双边模板匹配中预测块的位深度是固定的,不一定与视频编码的内部位深度相同。然后,视频编码器200和视频解码器300可以通过预测块的固定位深度来确定第1.2节中描述的“阈值”。例如,位深度为10,并且阈值为 $(1 \ll 2) * W * H / 2$,即 $4 * W * H / 2$ 。因此,阈值与用于编码视频的内部位深度无关。视频编码器200和视频解码器300可移位垂直插值的结果以实现内部插值位深度 bitDepthIF 而不是内部位深度 bitDepthI 。

[0232] 在一个示例中,数值 shift1 被设置为 $\text{precIF} - \min(\text{precIF}, \text{bitDepthIF} - \text{bitDepthI})$,并且数值 shift2 被设置为 precIF 。如果移位数大于0,则视频编码器200和视频解码器300可在右移位之前应用偏移。对于 shift1 ,如果 shift1 大于0,则偏移为 $1 \ll (\text{shift1} - 1)$ 。对于 shift2 ,如果 shift2 大于0,则偏移量 $1 \ll (\text{shift2} - 1)$ 。如果不应用插值,即运动矢量不是分数,则视频编码器200和视频解码器300可应用移位以实现相同的中间位深度和输出位深度。在一个例子中,当使用双线性插值时, precIF 为4, bitDepthIF 为10。数 shift1 为 $4 - \min(4, 10 - \text{bitDepthI})$,并且 shift2 为4。如果运动矢量的x分量为零,即,不执行水平插值,则将 shift2 重置为 $\text{precIF} - \min(\text{precIF}, \text{bitDepthIF} - \text{bitDepthI})$ 。

[0233] 在另一个例子中,将数值 shift1 设置为 $\text{precIF} - \min(\text{precIF} - 1, \text{bitDepthIF} - \text{bitDepthI})$,并且将数值 shift2 设置为 precIF 。如果运动矢量的x分量为零,即不执行水平插值,则 shift2 被重置为 $\text{precIF} - \min(\text{precIF}, \text{bitDepthIF} - \text{bitDepthI})$ 。

[0234] 在又一示例中,数值 shift1 被设置为 $\text{precIF} - (\text{bitDepthIF} - \text{bitDepthI})$,并且数值 shift2 被设置为 precIF 。这是因为 $\text{bitDepthIF} - \text{bitDepthI}$ 通常小于 precIF 。

[0235] 双边模板匹配中预测块的插值滤波器可以与正常运动补偿的插值滤波器不同。因此,双边模板匹配的 precIF 和 bitDepthIF 可能与正常运动补偿的不同。

[0236] 鉴于前述内容,为了对运动矢量应用双边模板匹配,视频编码器200和视频解码器300可以被配置为使用双边模板匹配确定当前块的两个预测块,其中基于当前块的运动矢量确定两个预测块的第一预测块,并且其中运动矢量包括非整数运动矢量。为了确定第一预测块,视频编码器200和视频解码器300可以被配置为基于非整数运动矢量执行水平插值,以比内部位深度更高的位深度存储水平插值的输出,并且使用存储在较高位深度的水平插值的输出以执行垂直插值。本公开的技术可提高编码效率和/或简化DMVR技术的实现。

[0237] 技术2

[0238] 在第二种技术中,预测块的位深度被设置为与用于编码视频的内部位深度相同。因此,预测块的位深度取决于用于编码视频的内部位深度。第1.2节中描述的“阈值”由用于

编码视频的内部位深度确定。在一个示例中,双边模板匹配的precIF和bitDepthIF可能与正常运动补偿的不同。

[0239] 在另一个示例中,双边模板匹配和正常运动补偿的插值过程是统一的。在该示例中,双边模板匹配的precIF和bitDepthIF与正常运动补偿中的相同。

[0240] 图12是说明用于编码当前块的示例方法的流程图。当前块可以包括当前CU。尽管关于视频编码器200(图1和图2)进行了描述,但是应当理解,其他设备也可以被配置为执行类似于图12的方法的方法。

[0241] 在该示例中,视频编码器200最初预测当前块(350)。例如,视频编码器200可以形成当前块的预测块。视频编码器200然后可以计算当前块的残差块(352)。为了计算残差块,视频编码器200可计算原始未编码块与当前块的预测块之间的差。视频编码器200然后可以变换和量化残差块的系数(354)。接下来,视频编码器200可扫描残差块的量化变换系数(356)。在扫描期间或在扫描之后,视频编码器200可对系数进行熵编码(358)。例如,视频编码器200可使用CAVLC或CABAC对系数进行编码。视频编码器200然后可以输出块的熵编码数据(360)。

[0242] 图13是说明用于解码视频数据的当前块的示例方法的流程图。当前块可以包括当前CU。尽管关于视频解码器300(图1和图3)进行了描述,但是应当理解,其他设备页可以被配置为执行类似于图13的方法的方法。

[0243] 视频解码器300可以接收当前块的熵编码数据,诸如与当前块相对应的残差块的系数的熵编码预测信息和熵编码数据(370)。视频解码器300可对熵编码数据进行熵解码以确定当前块的预测信息并再现残差块的系数(372)。视频解码器300可预测当前块(374),例如,使用由当前块的预测信息指示的帧内或帧间预测模式来计算当前块的预测块。本公开的用于确定是否执行例如如下图14和图15中所描述的运动矢量细化的技术可以在图13的过程374之前和/或期间由视频解码器300执行。视频解码器300然后可以逆扫描再现的系数(376),以创建量化变换系数块。视频解码器300然后可以对系数进行逆量化和逆变换以产生残差块(378)。视频解码器300可以通过组合预测块和残差块来最终解码当前块(380)。

[0244] 图14是说明用于应用解码器侧运动矢量细化的示例解码方法的示例的流程图。将参考包括运动补偿单元316在内的视频解码器300来描述图14的技术。然而,应当理解,也可以由包括运动补偿单元224的视频编码器200以相反的方式执行图14的技术。

[0245] 在本公开的一个示例中,视频解码器300可以被配置为接收使用双向预测的预测来编码的视频数据的当前块(1500)。视频解码器300然后可以确定双向预测的预测权重是否相等(1502)。如果双向预测的预测权重不相等,则视频解码器300可以禁用运动矢量细化过程(1504),并且使用没有运动矢量细化的双向预测的预测来解码当前块(1506)。如果视频解码器300确定双向预测的预测权重相等,则视频解码器300可以启用运动矢量细化过程(1508),并且使用具有运动矢量细化的双向预测的预测来解码当前块(1510)。

[0246] 因此,鉴于以上示例,视频解码器300可以被配置为确定使用双向预测的预测来编码的视频数据的当前块的运动矢量,基于用于当前块的双向预测的预测的权重来确定对运动矢量使用运动矢量细化过程,并且基于使用运动矢量细化过程的确定来对运动矢量应用运动矢量细化过程。在一个示例中,当用于当前块的双向预测的预测的权重不相等时,视频解码器300可经配置以确定禁用运动矢量细化过程。在一个示例中,运动矢量细化过程是双

边模板匹配。

[0247] 在一个示例中,双向预测是加权双向预测。在另一个示例中,双向预测是广义双向预测。

[0248] 在另一个示例中,视频解码器300可以被配置为确定用于在图片级别对当前块进行双向预测的预测的权重。在另一示例中,视频解码器300可经配置以在块级别确定用于当前块的双向预测的预测的权重。

[0249] 图15是说明用于应用解码器侧运动矢量细化的另一示例解码方法的流程图。将参考包括运动补偿单元316在内的视频解码器300来描述图15的技术。然而,应当理解,也可以由包括运动补偿单元224的视频编码器200以相反的方式执行图15的过程。

[0250] 在本公开的一个示例中,视频解码器300可经配置以接收使用双向预测的预测来编码的视频数据的当前块(1600)。视频解码器300然后可以确定双向预测的预测权重是否相等(1602)。如果双向预测的预测权重不相等,则视频解码器300可以禁用运动矢量细化过程(1604),并且使用没有运动矢量细化的双向预测的预测来解码当前块(1606)。

[0251] 如果视频解码器300确定双向预测的预测权重相等,则视频解码器300可以进一步确定由双向预测块的运动矢量识别的预测块之间的差是否小于阈值(1608)。如果是,则视频解码器300可以禁用运动矢量细化过程(1604),并且使用没有运动矢量细化的双向预测来解码当前块(1606)。如果否,则视频解码器300可以启用运动矢量细化过程(1610),并且使用具有运动矢量细化的双向预测的预测来解码当前块(1612)。

[0252] 因此,鉴于以上示例,视频解码器300可以被配置为确定使用双向预测编码的视频数据的当前块的运动矢量,确定由运动矢量识别的预测块之间的差,基于用于当前块的双向预测的预测的权重和由运动矢量识别的预测块之间的确定差来确定对运动矢量使用运动矢量细化过程,并且基于确定使用运动矢量细化过程来在运动矢量上应用运动矢量细化过程。

[0253] 本公开的说明性示例包括以下内容:

[0254] 示例1:一种用于对视频数据进行编码的方法,该方法包括确定视频数据的当前块的运动矢量,基于用于当前块的双向预测的预测的权重来确定对运动矢量使用运动矢量细化过程,以及基于该确定来细化运动矢量。

[0255] 示例2:示例1的方法,其中双向预测的预测是加权双向预测。

[0256] 示例3:示例1的方法,其中双向预测的预测是广义双向预测。

[0257] 示例4:示例1的方法,其中确定视频数据的当前块的运动矢量包括使用双边模板匹配确定运动矢量。

[0258] 示例5:一种用于对视频数据进行编码的方法,该方法包括确定视频数据的当前块的运动矢量,基于两个预测块之间的差确定对运动矢量使用运动矢量细化过程,以及基于该确定来细化运动矢量。

[0259] 示例6:一种对视频数据进行编码的方法,该方法包括接收要使用双边模板匹配编码的视频数据的当前块,使用双边模板匹配确定当前块的两个预测块,其中两个预测块的第一块基于非整数运动矢量被确定,其中确定第一块包括基于非整数运动矢量执行水平插值,以比内部位深度更高的位深度存储水平插值的输出,并使用存储在较高位深度的水平插值的输出来执行垂直插值。

- [0260] 示例7:示例1-6中任一项的方法,其中编码包括解码。
- [0261] 示例8:示例1-6中任一项的方法,其中编码包括编码。
- [0262] 示例9:一种用于对视频数据进行编码的设备,该设备包括一个或多个用于执行示例1-8中任一项的方法的部件。
- [0263] 示例10:示例9的设备,其中一个或多个部件包括在电路中实现的一个或多个处理器。
- [0264] 示例11:示例9和10中任一个的设备,还包括用于存储视频数据的存储器。
- [0265] 示例12:示例9-11中任一个的设备,还包括被配置为显示解码的视频数据的显示器。
- [0266] 示例13:示例9-12中任一项的设备,其中所述设备包括照相机、计算机、移动设备、广播接收器设备或机顶盒中的一个或多个。
- [0267] 示例14:示例9-13中任一项的设备,其中所述设备包括视频解码器。
- [0268] 示例15:示例9-14中任一项的设备,其中所述设备包括视频编码器。
- [0269] 示例16:一种其上存储有指令的计算机可读存储介质,该指令在被执行时使一个或多个处理器执行示例1-8中任一个的方法。
- [0270] 应当认识到,根据示例,本文描述的任何技术的某些动作或事件可以以不同的顺序执行,可以被添加、合并或完全省略(例如,并非所有描述的动作或事件是实施该技术所必需的)。此外,在某些示例中,动作或事件可以例如通过多线程处理、中断处理或多个处理器并发地而不是顺序地执行。
- [0271] 在一个或多个示例中,可以以硬件、软件、固件或其任何组合来实现所描述的功能。如果以软件实现,则功能可以作为一个或多个指令或代码存储在计算机可读介质上或通过计算机可读介质发送,并由基于硬件的处理单元执行。计算机可读介质可以包括计算机可读存储介质,其对应于诸如数据存储介质的有形介质,或者通信介质,包括例如根据通信协议来促进将计算机程序从一个地方转移到另一个地方的任何介质。以这种方式,计算机可读介质通常可以对应于(1)非暂时性的有形计算机可读存储介质,或者(2)诸如信号或载波的通信介质。数据存储介质可以由一台或多台计算机或一个或多个处理器访问以检索指令、代码和/或数据结构以实现本公开中描述的技术的任何可用介质。计算机程序产品可以包括计算机可读介质。
- [0272] 作为示例而非限制,这种计算机可读存储介质可包括RAM、ROM、EEPROM、EEPROM、CD-ROM或其他光盘存储器、磁盘存储器或其他磁性存储设备、闪存存储器或任何其他可用于以指令或数据结构形式存储所需程序代码且可由计算机访问的介质中的一个或多个。而且,任何连接都适当地称为计算机可读介质。例如,如果使用同轴电缆、光纤电缆、双绞线、数字用户线(DSL)或诸如红外、无线电和微波的无线技术从网站、服务器或其他远程源发送指令,则介质的定义包括同轴电缆、光纤电缆、双绞线、DSL或诸如红外、无线电和微波的无线技术。然而,应当理解,计算机可读存储介质和数据存储介质不包括连接、载波、信号或其他瞬时介质,而是针对非瞬时的有形存储介质。本文使用的磁盘和光盘包括压缩光盘(CD)、激光光盘、光盘、数字多功能光盘(DVD)、软盘和蓝光光盘,其中磁盘通常以磁性方式复制数据,而光盘则通过激光光学方式复制数据。上述的组合也应包括在计算机可读介质的范围内。

[0273] 指令可以由一个或多个处理器执行,例如一个或多个DSP、通用微处理器、ASIC、FPGA或其他等效的集成或离散逻辑电路。因此,如本文所使用的,术语“处理器”可以指任何前述结构或适合于实现本文描述的技术的任何其他结构。另外,在一些方面中,本文描述的功能性可以在被配置用于编码和解码的专用硬件和/或软件模块内提供,或结合在组合编解码器中。而且,该技术可以在一个或多个电路或逻辑元件中完全实现。

[0274] 本公开的技术可以在包括无线手持机、集成电路(IC)或一组IC(例如,芯片组)的多种设备或装置中实现。在本公开中描述了各种组件、模块或单元以强调经配置以执行所揭示技术的设备的功能方面,但不一定需要由不同硬件单元来实现。而是,如上所述,各种单元可以组合在编解码器硬件单元中,或者由互操作硬件单元的集合来提供,包括与合适的软件和/或固件结合的如上所述的一个或多个处理器。

[0275] 已经描述了各种示例。这些和其他示例在所附权利要求的范围内。

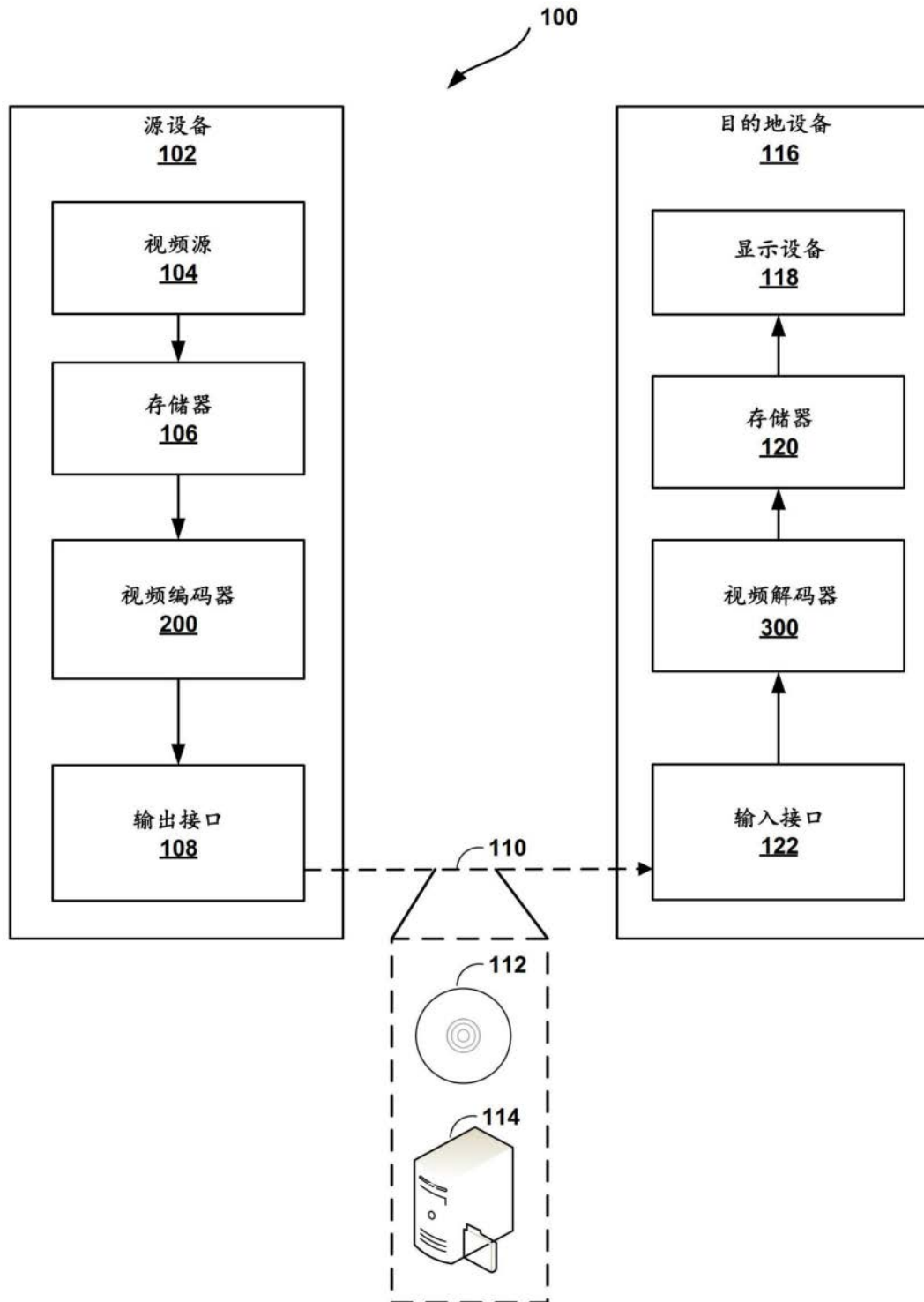


图1

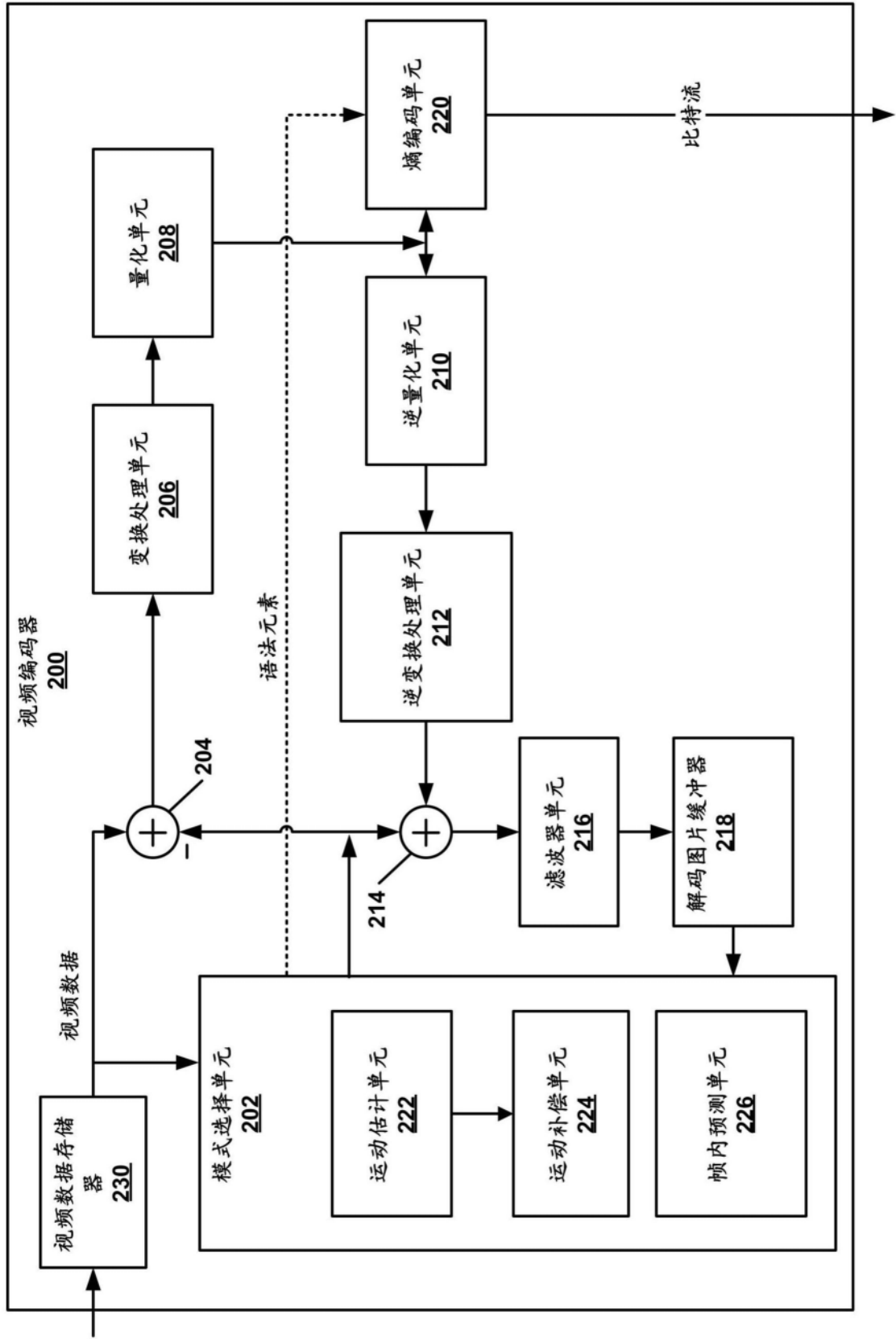


图2

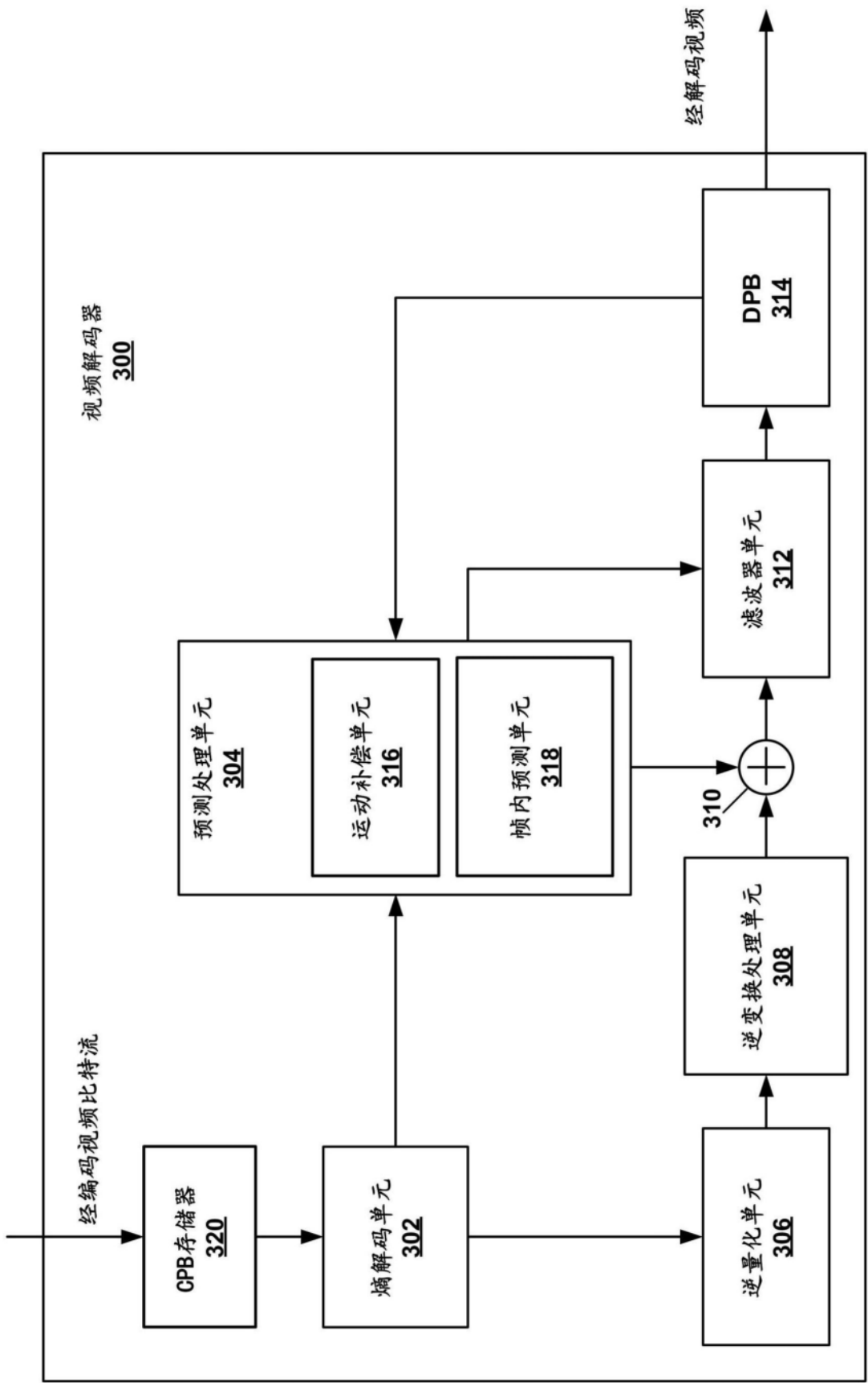


图3

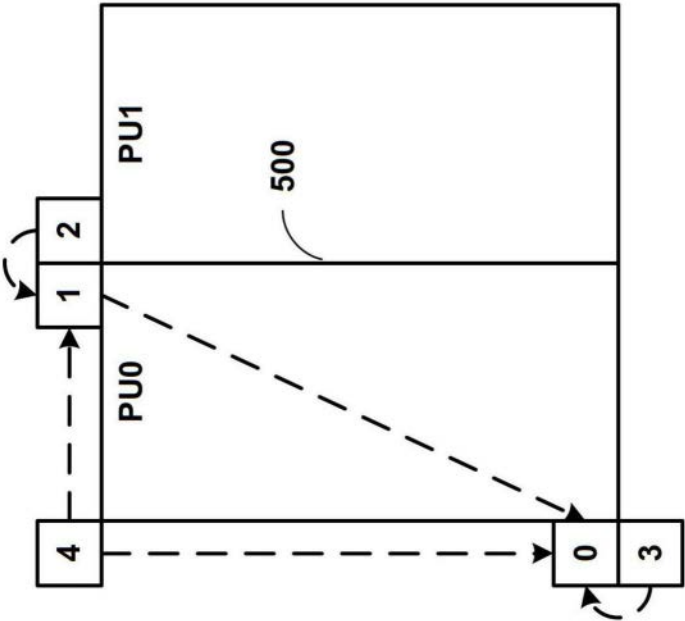


图4A

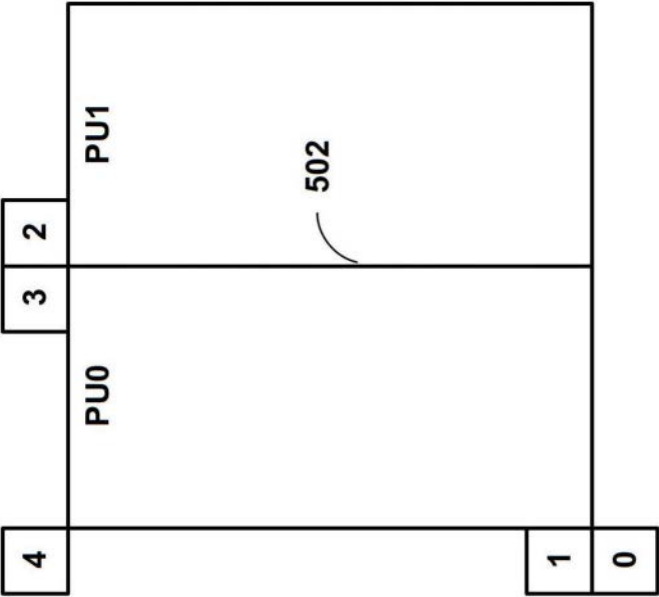


图4B

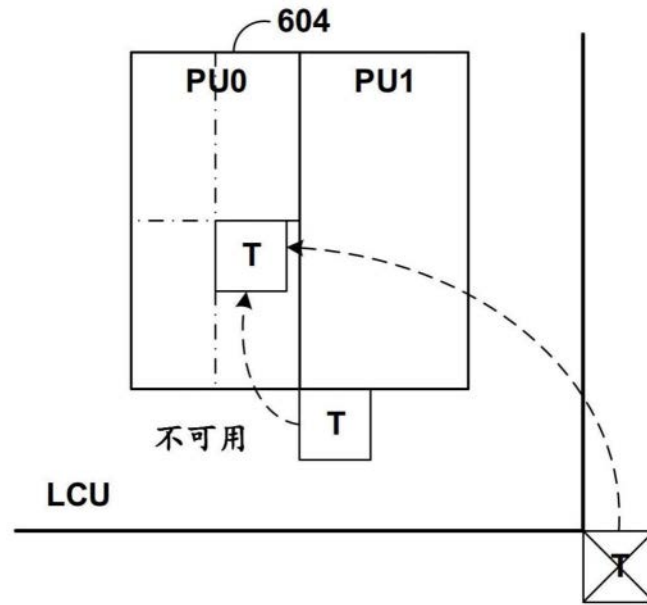


图5A

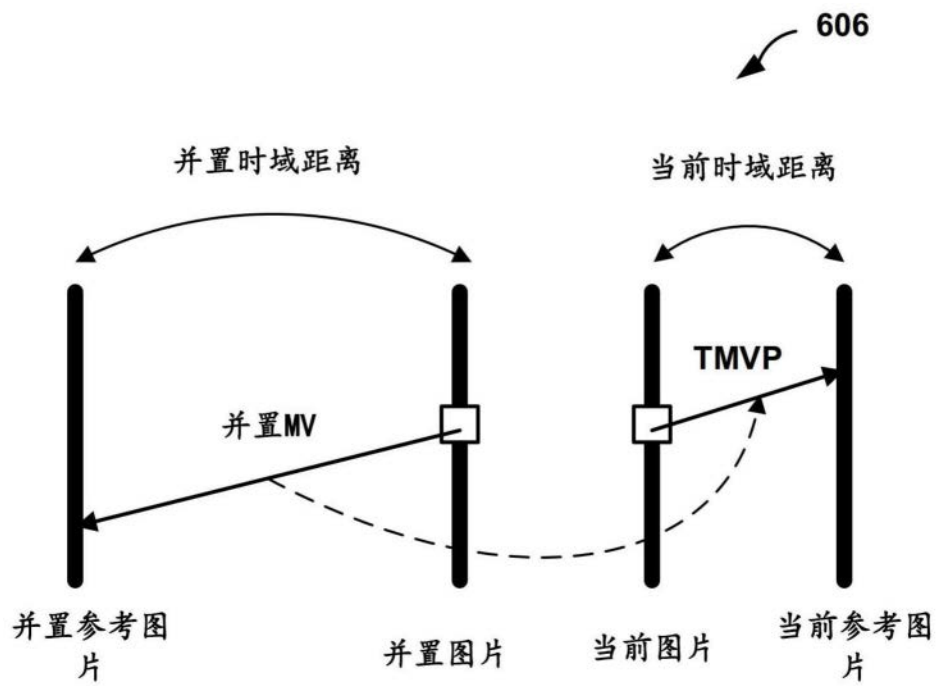


图5B

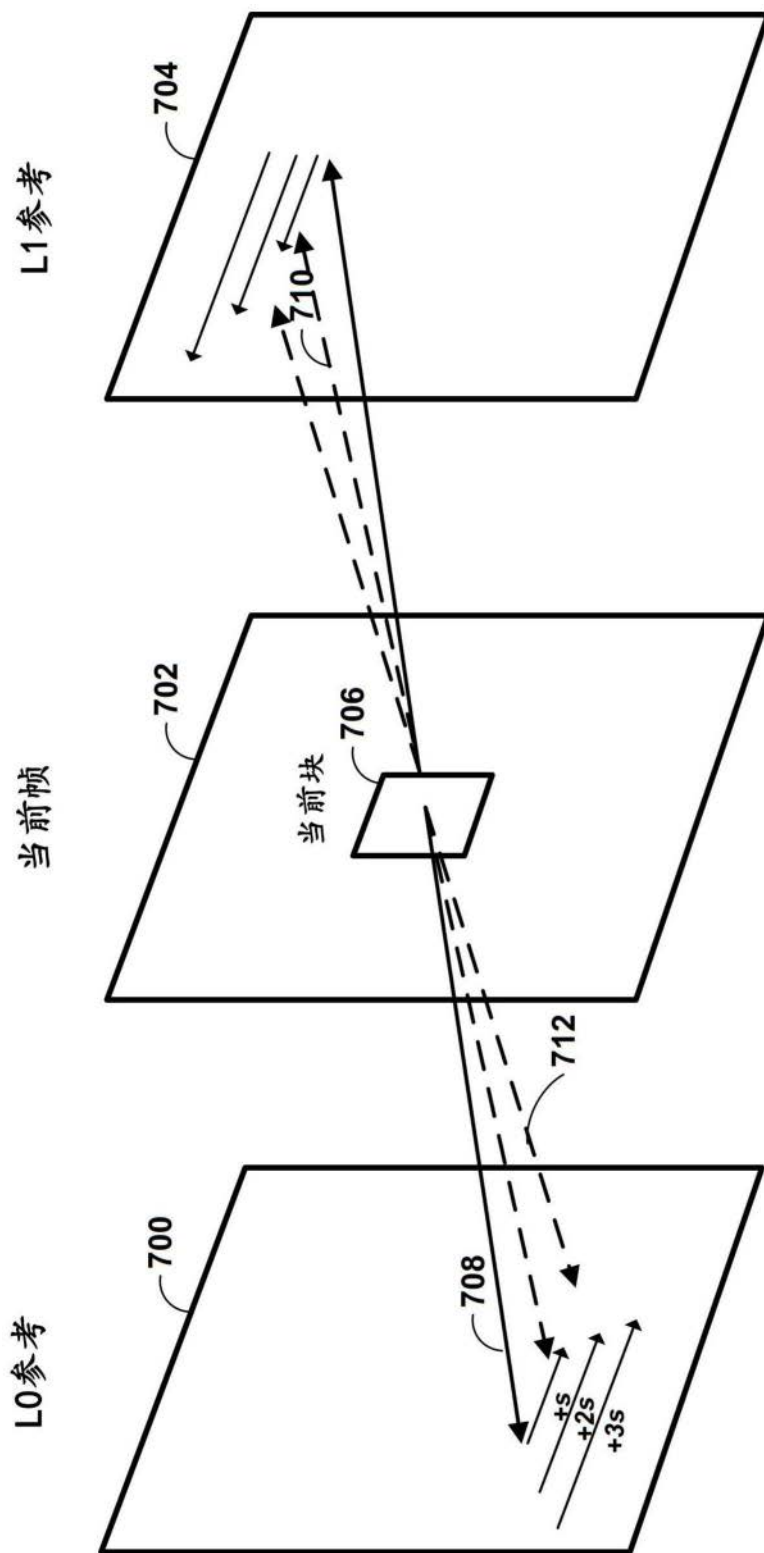


图6A

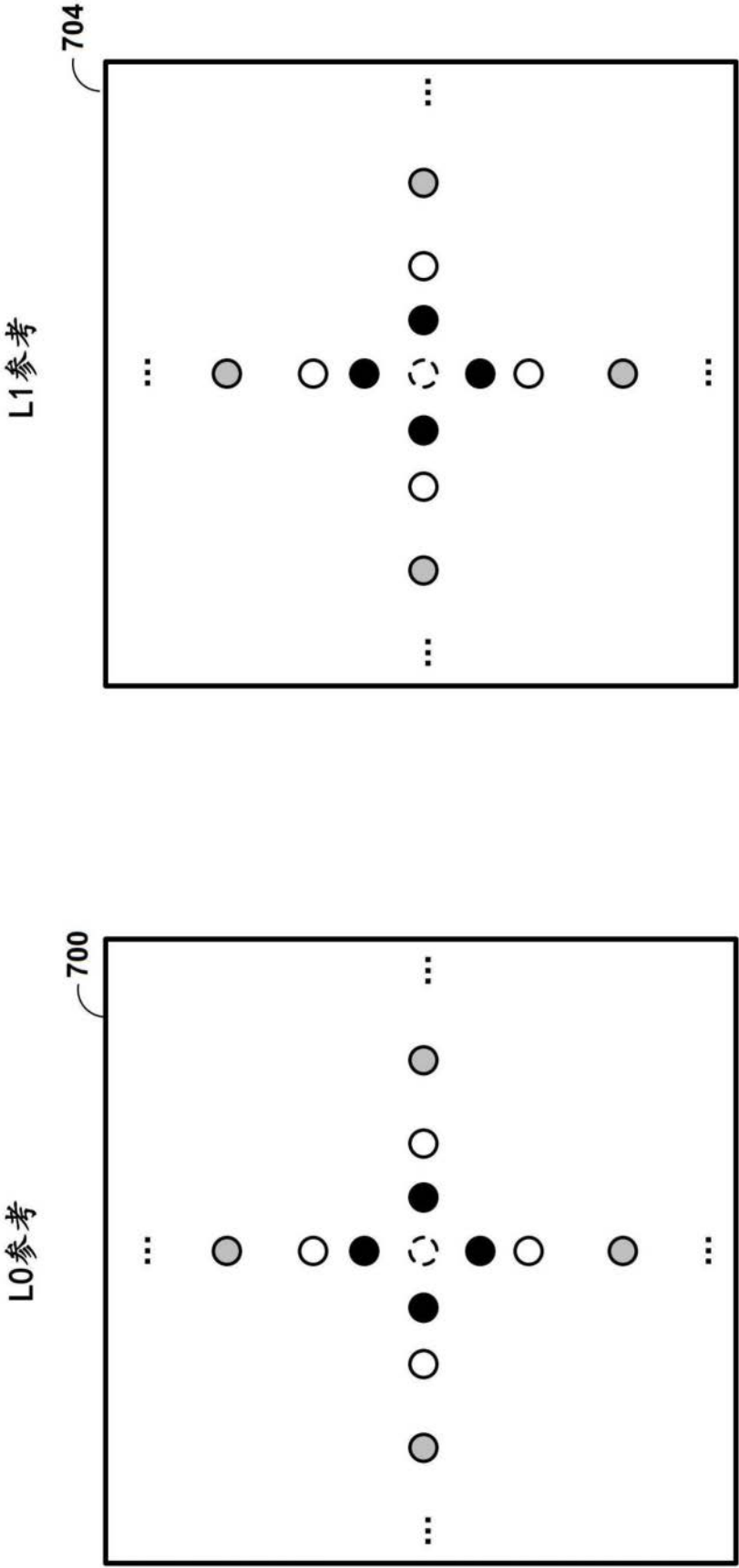


图6B

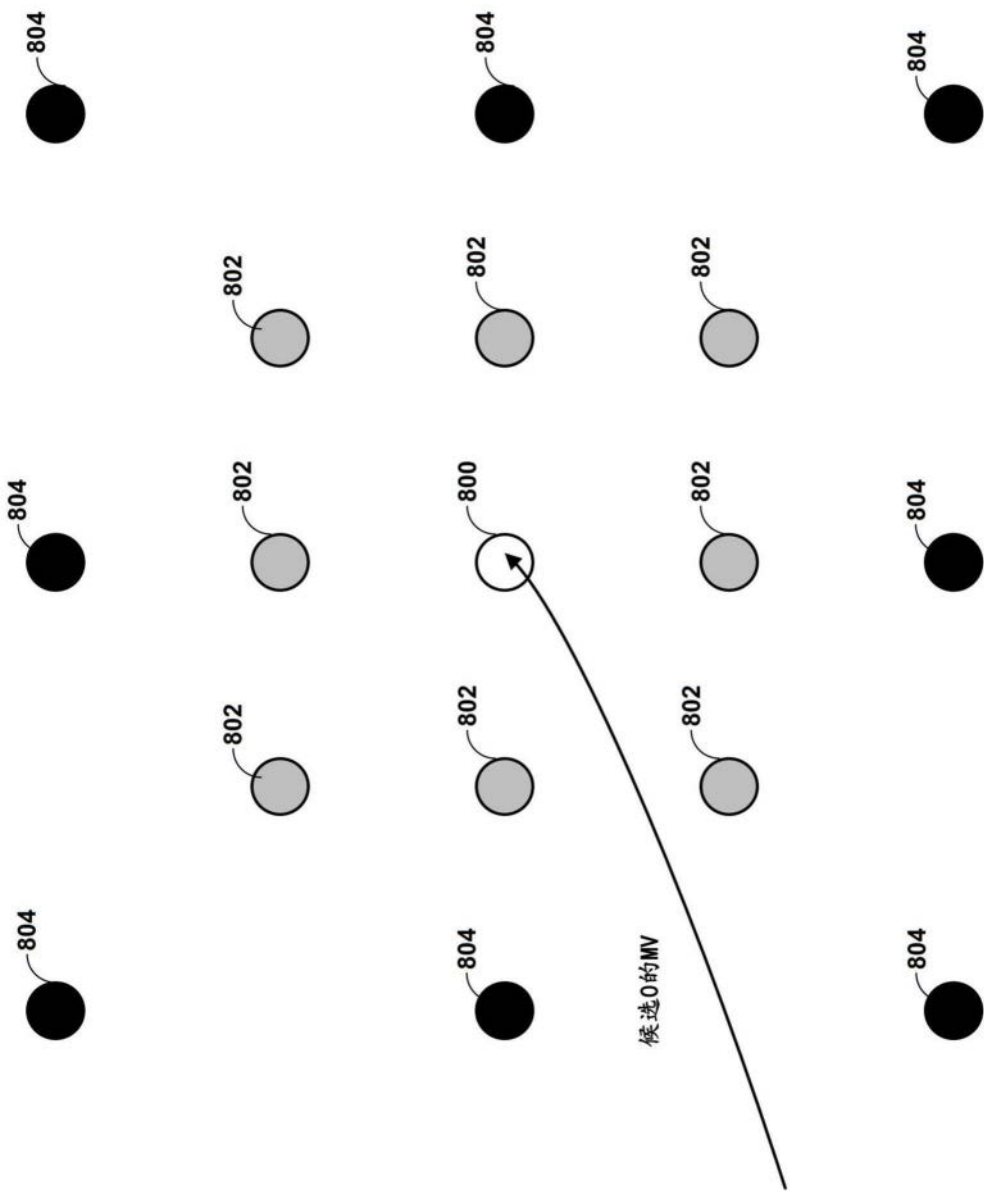


图7

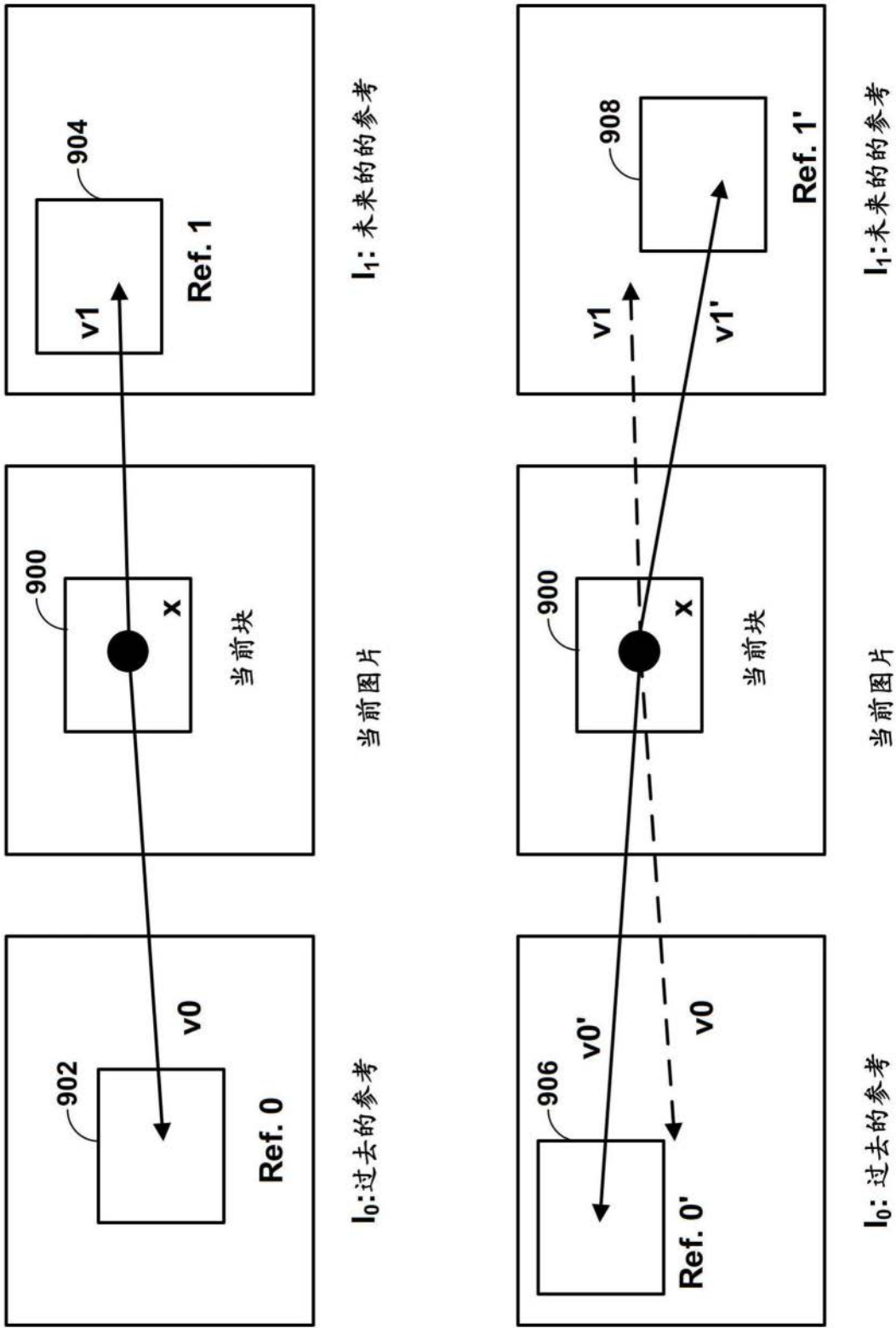


图8

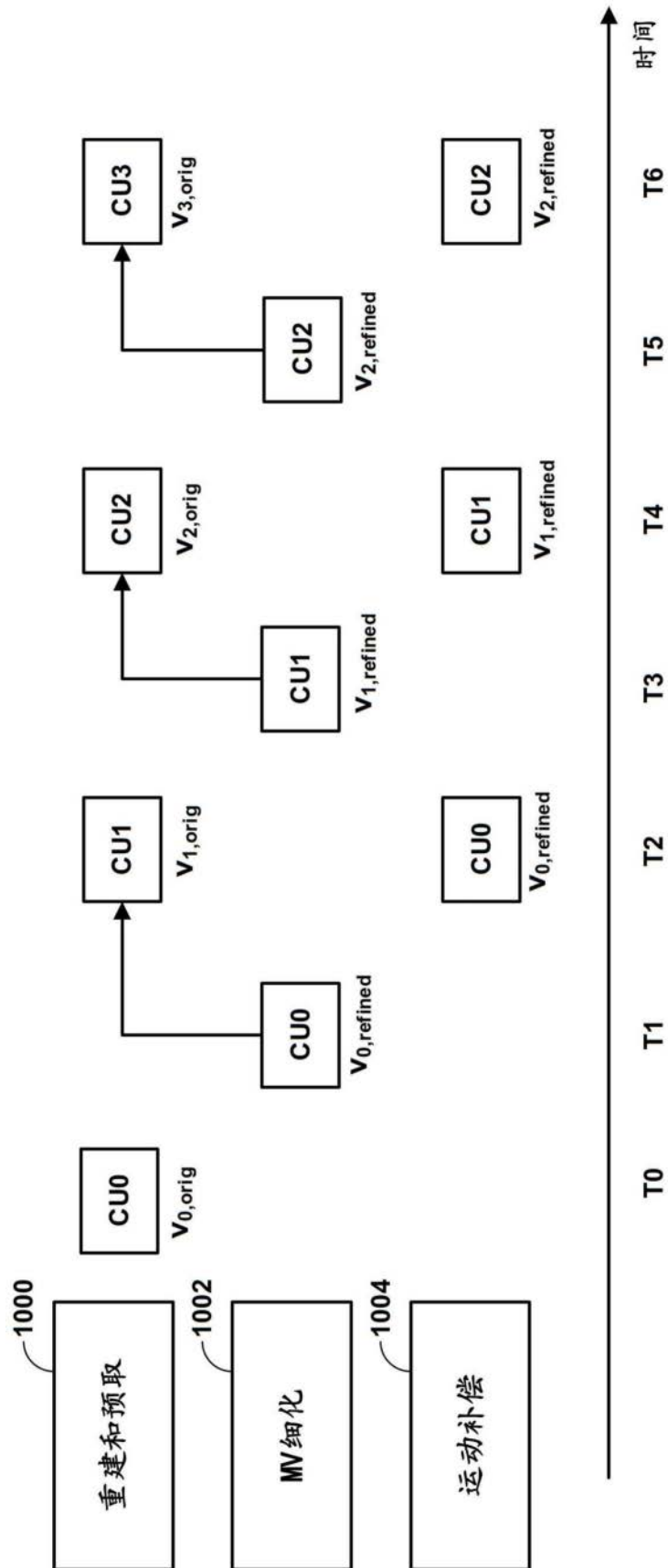


图9

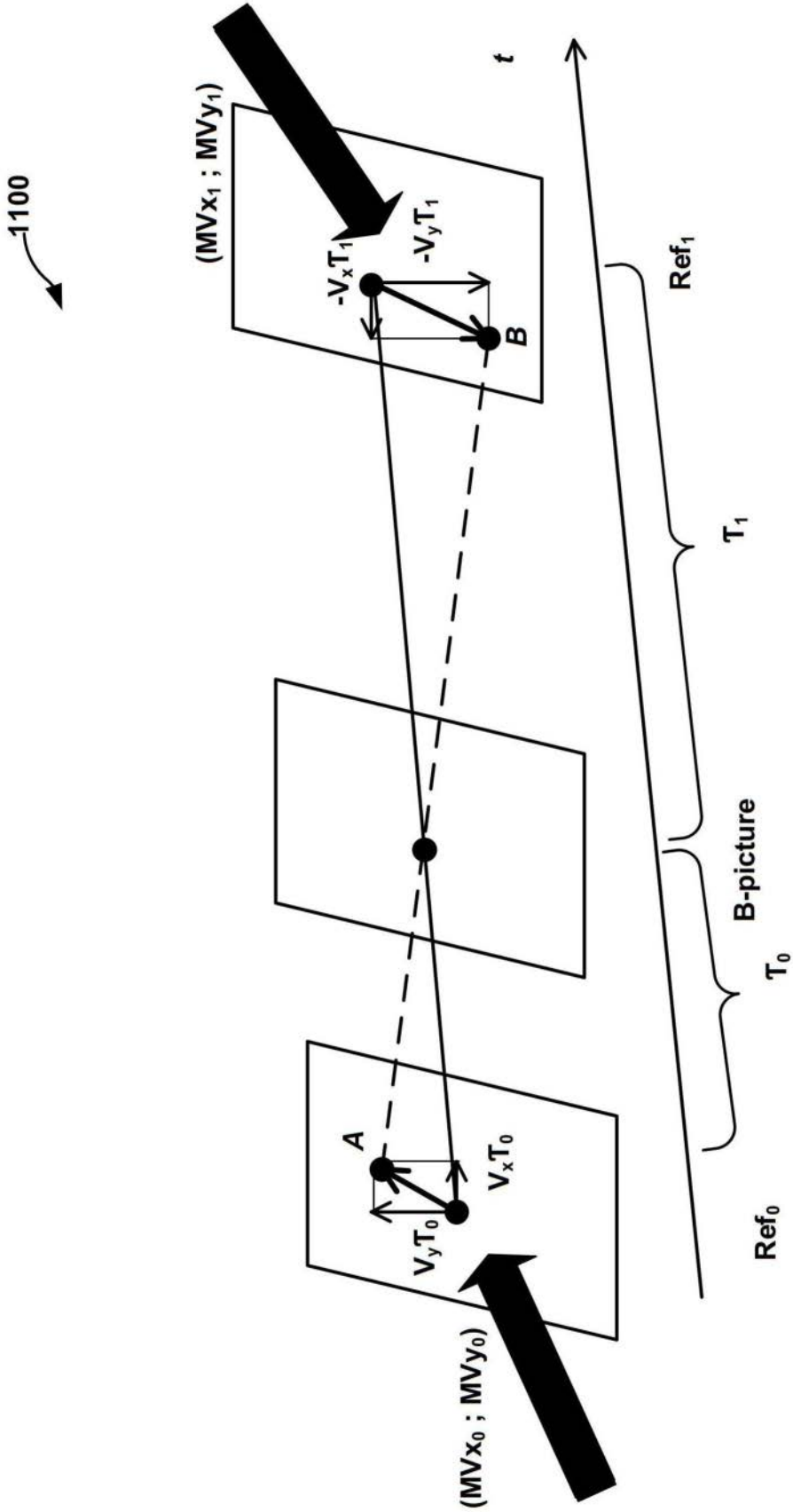


图10

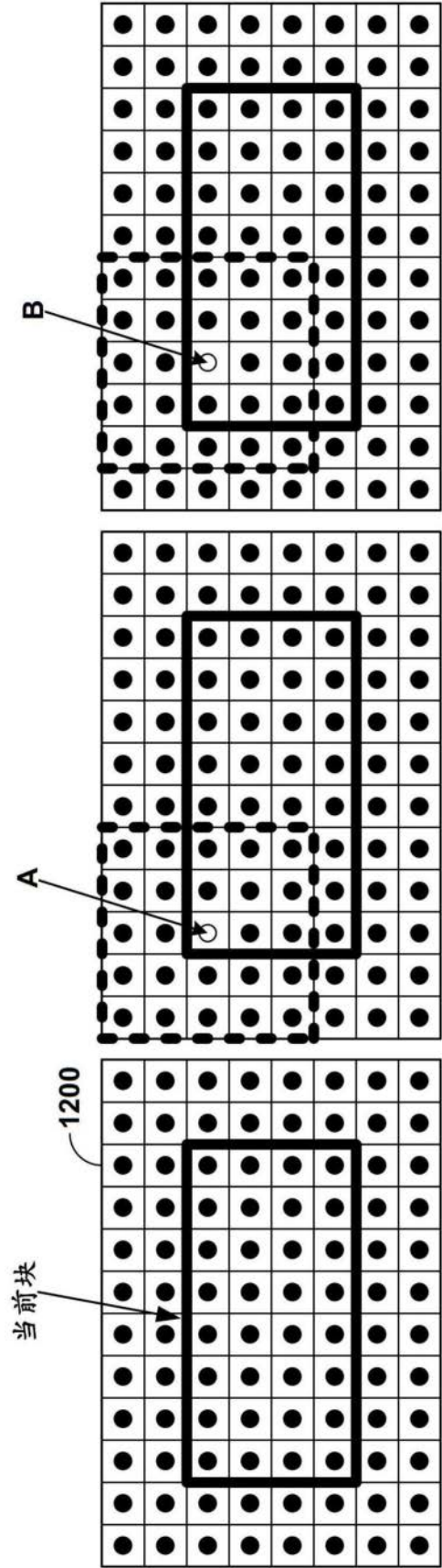


图11

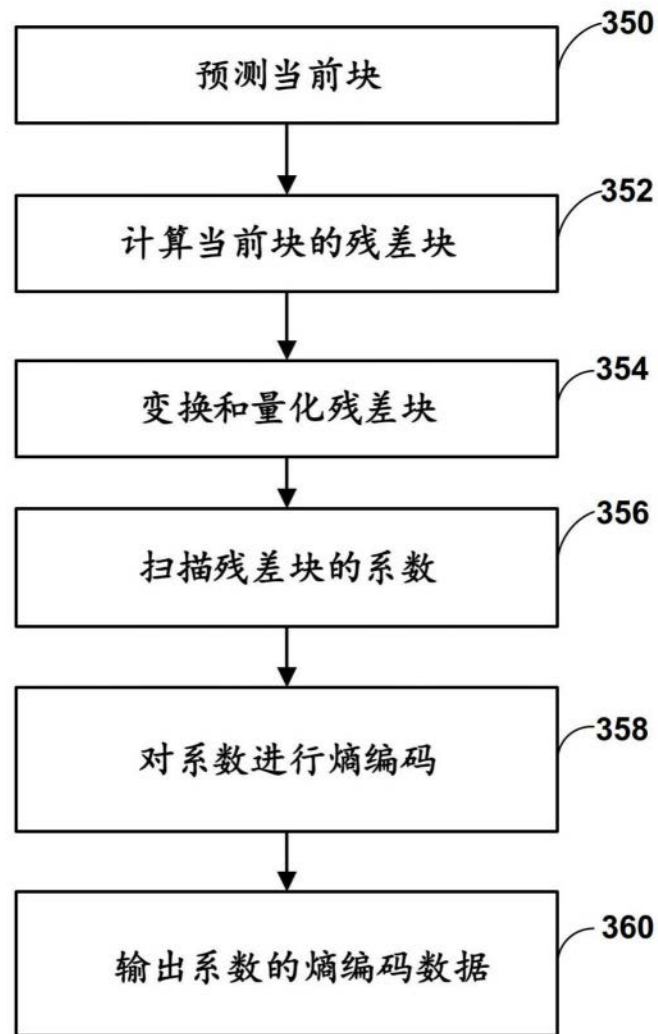


图12

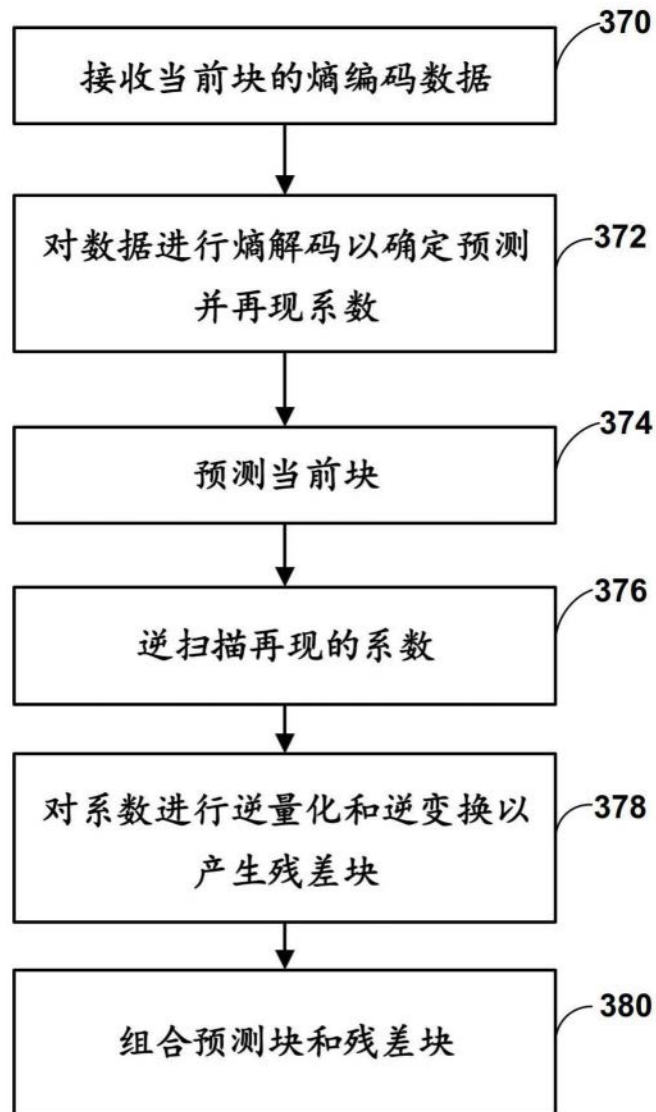


图13

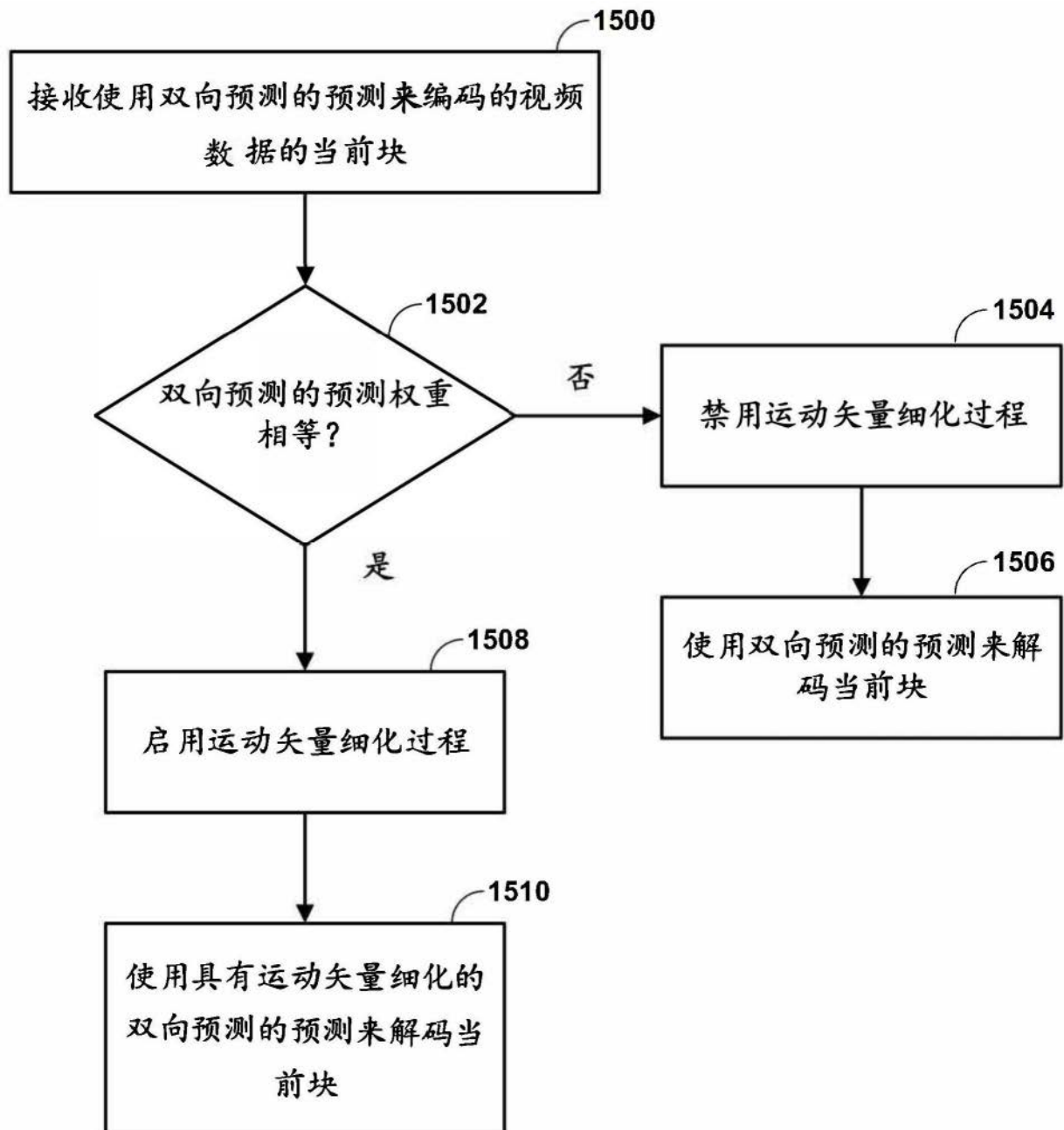


图14

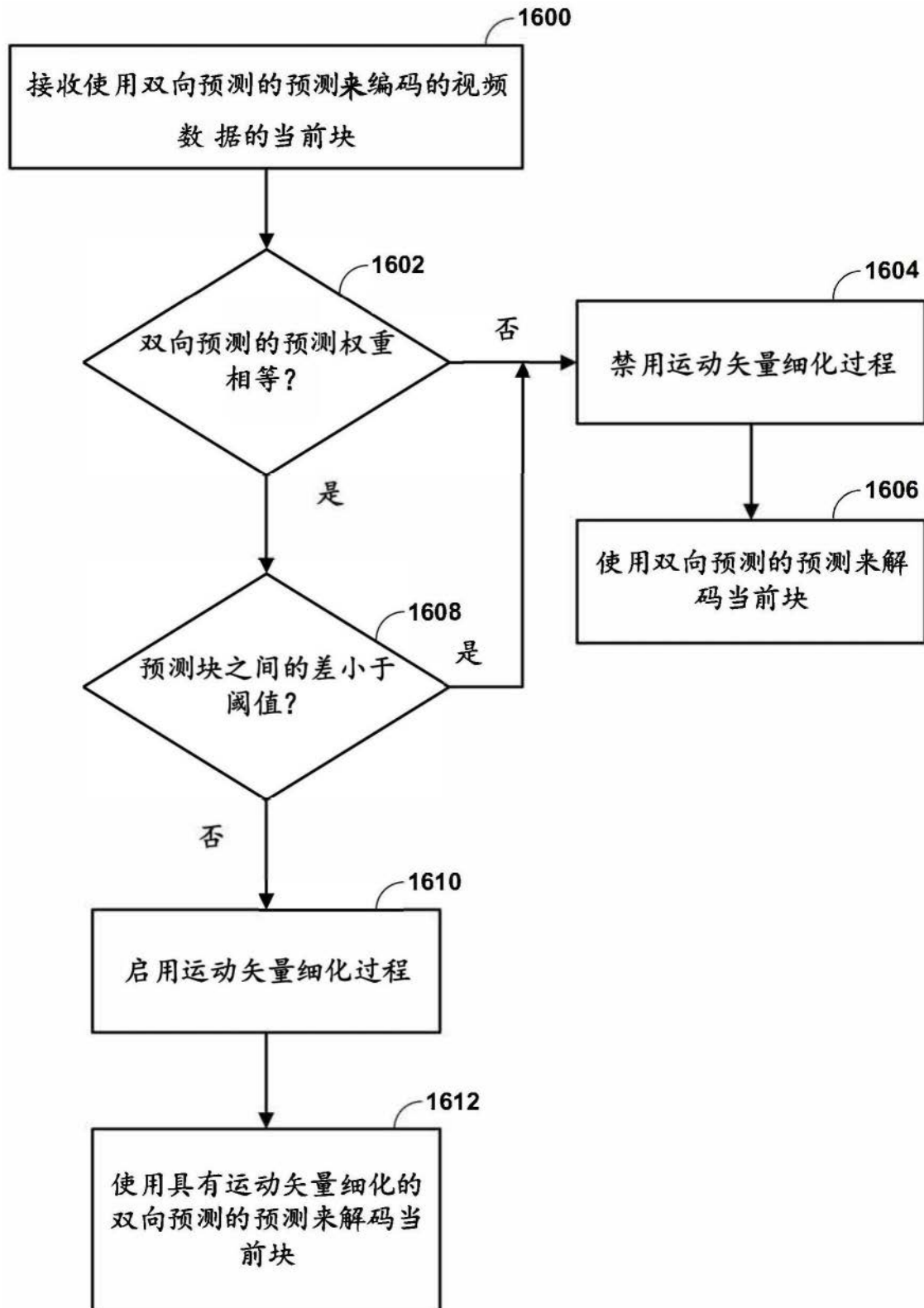


图15