



## (12)发明专利

(10)授权公告号 CN 107690809 B

(45)授权公告日 2020.05.01

(21)申请号 201680033592.4

(22)申请日 2016.06.09

(65)同一申请的已公布的文献号  
申请公布号 CN 107690809 A

(43)申请公布日 2018.02.13

(30)优先权数据

62/174,393 2015.06.11 US

62/295,329 2016.02.15 US

15/176,790 2016.06.08 US

(85)PCT国际申请进入国家阶段日  
2017.12.08

(86)PCT国际申请的申请数据  
PCT/US2016/036682 2016.06.09

(87)PCT国际申请的公布数据  
WO2016/201094 EN 2016.12.15

(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 钱威俊 王翔林 张莉 刘鸿彬  
陈建乐 马尔塔·卡切维奇

(74)专利代理机构 北京律盟知识产权代理有限公司  
11287

代理人 杨林勳

(51)Int.Cl.

H04N 19/51(2014.01)

H04N 19/52(2014.01)

H04N 19/70(2014.01)

H04N 19/61(2014.01)

H04N 19/172(2014.01)

H04N 19/176(2014.01)

审查员 朱一雷

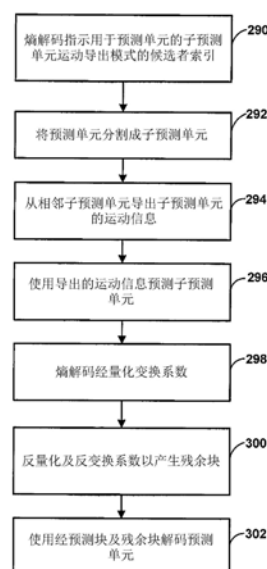
权利要求书3页 说明书30页 附图13页

(54)发明名称

使用空间及/或时间运动信息的子预测单元  
运动向量预测

(57)摘要

在一个实例中,一种用于解码视频数据的装置包含经配置以存储视频数据的存储器,及经配置以执行以下操作的视频解码器:确定当前视频数据块的运动预测候选者指示将针对所述当前块的子块导出运动信息;响应于所述确定:将所述当前块分割成所述子块,针对所述子块中的每一者,使用至少两个相邻块的运动信息来导出运动信息,并使用所述相应导出的运动信息解码所述子块。



1. 一种解码视频数据的方法,所述方法包括:

确定识别当前视频数据块的运动预测候选者列表中的运动预测候选者的候选者索引的值指示将针对所述当前块的子块导出运动信息;

响应于所述确定:

基于所述候选者索引确定所述子块而无需解码除所述候选者索引外指示所述当前块分裂成所述子块的语法元素;

针对所述子块中的每一者,使用上方相邻块、左相邻块和时间上相邻块的运动信息的平均值导出运动信息;及

使用所述相应导出的运动信息解码所述子块。

2. 根据权利要求1所述的方法,其中所述上方相邻块包括在所述当前块内的上方相邻子块。

3. 根据权利要求1所述的方法,其中所述左相邻块包括在所述当前块内的左相邻子块。

4. 根据权利要求1所述的方法,其中所述上方相邻块包括在所述当前块外部的上方相邻子块。

5. 根据权利要求1所述的方法,其中所述左相邻块包括在所述当前块外部的左相邻子块。

6. 根据权利要求1所述的方法,其中所述时间上相邻的块包括先前经解码图片中的相邻于与所述先前经解码图片中的所述当前块共置的块的块。

7. 根据权利要求1所述的方法,其中所述上方相邻、左相邻和时间上相邻的块具有等于或小于所述子块的大小的数据。

8. 根据权利要求1所述的方法,其进一步包括解码表示所述子块的大小的数据。

9. 根据权利要求8所述的方法,其中解码表示所述子块的所述大小的所述数据包括解码在切片标头、序列参数集SPS或图片参数集PPS中的至少一者中的所述数据。

10. 根据权利要求1所述的方法,其中导出所述运动信息包括按比例缩放对于共同参考图片的所述上方相邻、左相邻和时间上相邻的块的运动信息。

11. 根据权利要求1所述的方法,其中导出所述运动信息包括根据以下公式导出所述子块中的每一者的运动向量MV,其中x和y指示所述运动向量MV的水平分量及垂直分量,且i在{0,1,2}中:

$$(MV_x, MV_y) = ((\sum_{i=0}^2 MV_{xi} + \text{sign}(\sum_{i=0}^2 MV_{xi}) * 1) * 43/128, (\sum_{i=0}^2 MV_{yi} + \text{sign}(-\sum_{i=0}^2 MV_{yi}) * 1) * 43/128))。$$

12. 根据权利要求1所述的方法,其进一步包括将所述运动预测候选者添加到所述运动预测候选者列表。

13. 根据权利要求1所述的方法,其进一步包括在解码所述子块之前编码所述子块,所述方法进一步包含编码识别所述当前块的所述运动预测候选者的数据。

14. 根据权利要求1所述的方法,所述方法可在无线通信装置上执行,其中所述装置包括:

存储器,其经配置以存储所述视频数据;

处理器,其经配置以执行指令以处理存储于所述存储器中的所述视频数据;及  
接收器,其经配置以接收所述视频数据的经编码版本。

15. 根据权利要求14所述的方法,其中所述无线通信装置是蜂窝式电话,且所述经编码视频数据是通过所述接收器接收且根据蜂窝式通信标准调制。

16. 一种用于解码视频数据的装置,所述装置包括:

存储器,其经配置以存储视频数据;及

视频解码器,其经配置以:

确定识别当前视频数据块的运动预测候选者列表中的运动预测候选者的候选者索引的值指示将针对所述当前块的子块导出运动信息;

响应于所述确定:

基于所述候选者索引确定所述子块而无需解码除所述候选者索引外指示所述当前块分裂成所述子块的语法元素;

对于所述子块中的每一者,使用上方相邻块、左相邻块和时间上相邻的块的运动信息的平均值导出运动信息;及

使用所述相应导出的运动信息解码所述子块。

17. 根据权利要求16所述的装置,

其中所述上方相邻子块包括在所述当前块内或在所述当前块外部的上方相邻子块,且其中所述左相邻子块包括在所述当前块内或在所述当前块外部的左相邻子块。

18. 根据权利要求16所述的装置,其中所述视频解码器进一步经配置以从切片标头、序列参数集SPS或图片参数集PPS中的至少一者解码表示所述子块的大小的数据。

19. 根据权利要求16所述的装置,其中所述视频解码器进一步经配置以按比例缩放对于共同参考图片的所述上方相邻、左相邻和时间上相邻的块的运动信息。

20. 根据权利要求16所述的装置,其进一步包括经配置以在所述视频解码器解码所述子块之前编码所述子块的视频编码器。

21. 根据权利要求16所述的装置,其中所述装置是无线通信装置,其进一步包括:

接收器,其经配置以接收所述视频数据的经编码版本。

22. 根据权利要求21所述的装置,其中所述无线通信装置是蜂窝式电话,且所述经编码视频数据是通过所述接收器接收且根据蜂窝式通信标准调制。

23. 一种用于解码视频数据的装置,所述装置包括:

用于确定识别当前视频数据块的运动预测候选者列表中的运动预测候选者的候选者索引的值指示将针对所述当前块的子块导出运动信息的装置;

用于响应于所述确定基于所述候选者索引确定所述子块而无需解码除所述候选者索引外指示所述当前块分裂成所述子块的语法元素的装置;

用于响应于所述确定针对所述子块中的每一者使用上方相邻块、左相邻块和时间上相邻的块的运动信息的平均值导出运动信息的装置;及

用于响应于所述确定使用所述相应导出的运动信息解码所述子块的装置。

24. 根据权利要求23所述的装置,

其中所述上方相邻子块包括在所述当前块内或在所述当前块外部的上方相邻子块,且其中所述左相邻子块包括在所述当前块内或在所述当前块外部的左相邻子块。

25. 根据权利要求23所述的装置,其进一步包括用于在用于解码的所述装置解码所述子块之前编码所述子块的装置。

26. 一种上面存储有指令的计算机可读存储媒体,所述指令在执行时使得处理器进行以下操作:

确定识别当前视频数据块的运动预测候选者列表中的运动预测候选者的候选者索引的值指示将针对所述当前块的子块导出运动信息;

响应于所述确定:

基于所述候选者索引确定所述子块而无需解码除所述候选者索引外指示所述当前块分裂成所述子块的语法元素;

针对所述子块中的每一者,使用上方相邻块、左相邻块和时间上相邻的块的运动信息的平均值导出运动信息;及

使用所述相应导出的运动信息解码所述子块。

## 使用空间及/或时间运动信息的子预测单元运动向量预测

[0001] 本申请案主张2015年6月11日申请的美国临时申请案第62/174,393号及2016年2月15日申请的美国临时申请案第62/295,329号的权益,所述申请案中的每一者的全部内容特此以引用的方式并入。

### 技术领域

[0002] 本发明涉及视频译码。

### 背景技术

[0003] 数字视频能力可并入到广泛范围的装置中,包含数字电视、数字直播系统、无线广播系统、个人数字助理(PDA)、膝上型或台式计算机、平板计算机、电子书阅读器、数码相机、数字记录装置、数字媒体播放器、视频游戏装置、视频游戏控制台、蜂窝式或卫星无线电话、所谓的“智能型电话”、视频电话会议装置、视频流式传输装置及其类似者。数字视频装置实施视频译码技术,例如在由MPEG-2、MPEG-4、ITU-T H.263、ITU-T H.264/MPEG-4第10部分高级视频译码(AVC)所定义的标准、高效视频译码(HEVC)标准(还称作ITU-T H.265)及这些标准的扩展中所描述的那些视频译码技术。视频装置可通过实施此类视频译码技术来更有效地发射、接收、编码、解码及/或存储数字视频信息。

[0004] 视频译码技术包含空间(图片内)预测及/或时间(图片间)预测以减少或去除视频序列中固有的冗余。对于基于块的视频译码,可将视频切片(例如,视频帧或视频帧的一部分)分割成视频块,对于一些技术来说,视频块还可被称作树块、译码单元(CU)及/或译码节点。使用相对于同一图片中的相邻块中的参考样本的空间预测来编码图片的帧内译码(I)的切片中的视频块。图片的帧间译码(P或B)切片中的视频块可使用关于同一图片中的相邻块中的参考样本的空间预测或关于其它参考图片中的参考样本的时间预测。图片可被称作帧,且参考图片可被称作参考帧。

[0005] 空间或时间预测产生待译码的块的预测性块。残余数据表示待译码的原始块与预测性块之间的像素差。根据指向形成预测性块的参考样本的块的运动向量及指示经译码块与预测性块之间的差异的残余数据来编码帧间译码块。帧内译码块根据帧内译码模式及残余数据来编码。为了进行进一步压缩,可将残余数据从像素域变换到变换域,从而产生残余变换系数,可接着量化所述残余变换系数。可扫描最初布置成二维阵列的经量化变换系数以便产生变换系数的一维向量,且可应用熵译码以实现甚至更多的压缩。

### 发明内容

[0006] 一般来说,本发明的技术涉及视频数据块的子块的运动信息(例如,运动向量)的导出。举例来说,所述技术可用于导出预测单元(PU)或PU的子预测单元(子PU)的运动信息。一般来说,这些技术包含从相邻子块的运动信息导出所述子块中的每一者的运动信息。相邻子块可包含在空间上及/或在时间上相邻的子块。举例来说,对于给定子块,视频译码器(例如视频编码器或视频解码器)可通过组合(例如,平均)左相邻子块、上方相邻子块及/或

时间上相邻的子块(例如,右下时间上相邻的子块)的运动信息而导出运动信息。另外,子块的此运动信息的导出可使用用于运动信息预测的候选者列表的特定候选者来用信号表示。

[0007] 在一个实例中,一种解码视频数据的方法包含确定当前视频数据块的运动预测候选者指示将针对当前块的子块导出运动信息,及响应于所述确定:将当前块分割成子块,针对所述子块中的每一者,使用至少两个相邻块的运动信息来导出运动信息,及使用相应导出的运动信息解码所述子块。

[0008] 在另一实例中,一种用于解码视频数据的装置包含经配置以存储视频数据的存储器及经配置以执行以下操作的视频解码器:确定当前视频数据块的运动预测候选者指示将针对当前块的子块导出运动信息;及响应于所述确定:将当前块分割成子块,针对所述子块中的每一者,使用至少两个相邻块的运动信息导出运动信息,并使用相应导出的运动信息解码所述子块。

[0009] 在另一实例中,一种用于解码视频数据的装置包含用于确定当前视频数据块的运动预测候选者指示将针对当前块的子块导出运动信息的装置,用于响应于所述确定将当前块分割成子块的装置,用于响应于所述确定对于所述子块中的每一者使用至少两个相邻块的运动信息导出运动信息的装置,及用于响应于所述确定使用相应导出的运动信息解码所述子块的装置。

[0010] 在另一实例中,一种计算机可读存储媒体上面存储有指令,所述指令在执行时使得处理器执行以下操作:确定当前视频数据块的运动预测候选者指示将针对当前块的子块导出运动信息;及响应于所述确定:将当前块分割成子块,针对所述子块中的每一者,使用至少两个相邻块的运动信息导出运动信息,及使用相应导出的运动信息解码所述子块。

[0011] 在以下随附图式及描述中阐述一或多个实例的细节。其它特征、目标及优点将从所述描述及图式以及从权利要求书而显而易见。

## 附图说明

[0012] 图1是说明可利用用于实施高级时间运动向量预测(ATMVP)的技术的实例视频编码及解码系统的框图。

[0013] 图2是说明可实施用于高级时间运动向量预测(ATMVP)的技术的视频编码器的实例的框图。

[0014] 图3是说明可实施用于高级时间运动向量预测(ATMVP)的技术的视频解码器的实例的框图。

[0015] 图4是说明高效率视频译码(HEVC)中的空间相邻候选者的概念图。

[0016] 图5是说明HEVC中的时间运动向量预测(TMVP)的概念图。

[0017] 图6说明3D-HEVC的实例预测结构。

[0018] 图7是说明3D-HEVC中的基于子PU的视图间运动预测的概念图。

[0019] 图8是说明从参考图片的子PU运动预测的概念图。

[0020] 图9是说明ATMVP(类似于TMVP)中的相关图片的概念图。

[0021] 图10是说明实例空间时间运动向量预测符(STMVP)导出程序的流程图。

[0022] 图11A及11B是说明PU的子PU以及PU的相邻子PU的实例的概念图。

[0023] 图12是说明根据本发明的技术的编码视频数据的实例方法的流程图。

[0024] 图13是根据本发明的技术的解码视频数据的方法的实例。

### 具体实施方式

[0025] 一般来说,本发明涉及视频编解码器中的运动向量预测。更特定地说,高级运动向量预测可通过从空间及时间相邻块导出给定块(例如,预测单元(PU))的子块(例如,子预测单元(PU))的运动向量而实现。在一个实例中,视频译码器(例如视频编码器或视频解码器)可将当前块(例如,当前PU)分割成子块(例如,子PU),并针对每一子PU从相邻块导出子PU中的每一者的运动信息(包含运动向量),所述相邻块可包含空间及/或时间上相邻的块。举例来说,针对子块中的每一者,视频译码器可从左相邻空间块、上方相邻空间块及/或右下相邻时间块导出运动信息。空间上相邻的块可以是直接邻近于子块或在包含子块的当前块外部的子块。使用在当前块外部的子块可允许子块的运动信息被并行导出。

[0026] 视频译码标准包含ITU-T H.261、ISO/IEC MPEG-1 Visual、ITU-T H.262或ISO/IEC MPEG-2 Visual、ITU-T H.263、ISO/IEC MPEG-4 Visual及ITU-T H.264(还称为ISO/IEC MPEG-4AVC),包含其可伸缩视频译码(SVC)及多视图视频译码(MVC)扩展。MVC的最新联合草案描述于2010年3月的“用于通用视听服务的高级视频译码(Advanced video coding for generic audiovisual services)”(ITU-T标准H.264)中。

[0027] 另外,存在新开发的视频译码标准,即ITU-T视频译码专家群组(VCEG)及ISO/IEC运动图片专家群组(MPEG)的视频译码联合合作小组(JCT-VC)已开发的高效率视频译码(HEVC)。HEVC的最新草案可从[phenix.int-evry.fr/jct/doc\\_end\\_user/documents/12\\_Geneva/wg11/JCTVC-L1003-v34.zip](http://phenix.int-evry.fr/jct/doc_end_user/documents/12_Geneva/wg11/JCTVC-L1003-v34.zip)获得。HEVC标准还在推荐标准ITU-T H.265及国际标准ISO/IEC 23008-2中联合呈现,两者都题为“高效率视频译码”且两者都在2014年10月公开。

[0028] 运动信息:对于每一块,运动信息的集合可以是可用的。运动信息的集合含有用于前向及后向预测方向的运动信息。此处,前向及后向预测方向是对应于当前图片或切片的参考图片列表0(RefPicList0)及参考图片列表1(RefPicList1)的两个预测方向。术语“前向”及“后向”不必具有几何含义。替代地,其用以区分运动向量是基于哪一参考图片列表。前向预测意味着基于参考列表0形成的预测,而后向预测意味着基于参考列表1形成的预测。在参考列表0及参考列表1两者都用以形成给定块的预测的情况下,其被称为双向预测。

[0029] 对于给定图片或切片,如果使用仅一个参考图片列表,那么在图片或切片内部的每一块经前向预测。如果两个参考图片列表都用于给定图片或切片,那么在图片或切片内部的块可经前向预测,或后向预测,或双向预测。

[0030] 对于每一预测方向,运动信息含有参考索引及运动向量。参考索引用以识别对应参考图片列表中(例如,RefPicList0或RefPicList1)的参考图片。运动向量具有水平分量及垂直分量两者,其中每一分量分别指示沿水平方向及垂直方向的偏移值。在一些描述中,为简单起见,术语“运动向量”可与运动信息互换地使用以指示运动向量及其相关联参考索引两者。

[0031] 视频译码标准中广泛地使用图片次序计数(POC)以识别图片的显示次序。尽管存在一个经译码视频序列内的两个图片可具有相同POC值的情况,但经译码视频序列内通常不发生此类情况。当位流中存在多个经译码视频序列时,具有同一POC值的图片就解码次序来说可更接近于彼此。图片的POC值通常用于参考图片列表构建、如HEVC中的参考图片集的

导出及运动向量按比例缩放。

[0032] 高级视频译码 (AVC) (H.264) 中的宏块 (MB) 结构: 在 H.264/AVC 中, 每一帧间宏块 (MB) 可被分割成四个不同方式:

[0033] • 一个  $16 \times 16$  MB 分区

[0034] • 两个  $16 \times 8$  MB 分区

[0035] • 两个  $8 \times 16$  MB 分区

[0036] • 四个  $8 \times 8$  MB 分区

[0037] 一个 MB 中的不同 MB 分区针对每一方向可具有不同参考索引值 (RefPicList0 或 RefPicList1)。

[0038] 当 MB 不被分割成四个  $8 \times 8$  MB 分区时, 对于每一方向上的每一 MB 分区, MB 仅具有一个运动向量。

[0039] 当 MB 被分割成四个  $8 \times 8$  MB 分区时, 每一  $8 \times 8$  MB 分区可进一步被分割成子块, 所述子块中的每一者在每一方向上可具有不同运动向量。存在从  $8 \times 8$  MB 分区得到子块的四种不同方式:

[0040] • 一个  $8 \times 8$  子块

[0041] • 两个  $8 \times 4$  子块

[0042] • 两个  $4 \times 8$  子块

[0043] • 四个  $4 \times 4$  子块

[0044] 每一子块在每一方向上可具有不同运动向量。因此, 运动向量以等于或高于子块的电平呈现。

[0045] AVC 中的时间直接模式: 在 AVC 中, 可以针对 B 切片中的跳跃或直接模式的 MB 或 MB 分区电平启用时间直接模式。对于每一 MB 分区, 与当前块的 RefPicList1[0] 中的当前 MB 分区共置的块的运动向量用以导出运动向量。共置块中的每一运动向量基于 POC 距离而按比例缩放。

[0046] AVC 中的空间直接模式: 在 AVC 中, 直接模式还可从空间相邻者预测运动信息。

[0047] 高效率视频译码 (HEVC) 中的译码单元 (CU) 结构: 在 HEVC 中, 切片中的最大译码单元被称作译码树块 (CTB) 或译码树型单元 (CTU)。CTB 含有四分树, 所述四分树的节点是译码单元。

[0048] CTB 的大小范围在 HEVC 主规范中可介于  $16 \times 16$  到  $64 \times 64$  之间 (尽管技术上可支持  $8 \times 8$  CTB 大小)。译码单元 (CU) 可与 CTB 大小相同, 且小达  $8 \times 8$ 。每一译码单元是用一个模式译码。当 CU 经帧间译码时, 其可进一步分割成 2 个或 4 个预测单元 (PU), 或当不应用另一分区时变为仅一个 PU。当两个 PU 存在于一个 CU 中时, 所述两个 PU 可以是一半大小的矩形或大小是 CU 的  $1/4$  或  $3/4$  大小的两个矩形。

[0049] 当 CU 经帧间译码时, 针对每一 PU 呈现运动信息的一个集合。另外, 每一 PU 通过独特帧间预测模式译码以导出所述运动信息集合。

[0050] HEVC 中的运动预测: 在 HEVC 标准中, 存在针对预测单元 (PU) 的两个运动向量预测模式, 其分别称为合并 (跳跃被视为合并的特殊情况) 模式及高级运动向量预测 (AMVP) 模式。

[0051] 在 AMVP 或合并模式中, 针对多个运动向量预测符维持运动向量 (MV) 候选者列表。



当前PU的运动向量以及合并模式中的参考索引是通过从MV候选者列表获取一个候选者而产生。

[0052] MV候选者列表含有用于合并模式的达5个候选者且仅两个候选者用于AMVP模式。合并候选者可含有运动信息的集合,例如对应于两个参考图片列表(列表0及列表1)的运动向量及参考索引。如果由合并索引来识别合并候选者,那么参考图片用于当前块的预测,以及确定相关联的运动向量。然而,在AMVP模式下,对于从列表0或列表1的每一潜在预测方向,需要将参考索引连同对MV候选者列表的MVP索引一起明确地用信号表示,这是因为AMVP候选者含有仅一运动向量。在AMVP模式中,可进一步改进经预测运动向量。

[0053] 如可从上文看出,合并候选者对应于运动信息的整个集合,而AMVP候选者含有仅用于特定预测方向的一个运动向量及参考索引。

[0054] 以类似方式从相同空间及时间相邻块导出两个模式的候选者。

[0055] 图1是说明可利用用于实施高级时间运动向量预测(ATMVP)的技术的实例视频编码及解码系统10的框图。如图1中所示,系统10包含源装置12,其提供待在稍后时间由目的地装置14解码的经编码视频数据。具体地说,源装置12经由计算机可读媒体16将视频数据提供到目的地装置14。源装置12及目的地装置14可包括广泛范围的装置中的任一者,包含台式计算机、笔记型(即,膝上型)计算机、平板计算机、机顶盒、例如所谓的“智能”电话的电话手机、所谓的“智能”平板计算机、电视、相机、显示装置、数字媒体播放器、视频游戏控制台、视频流式传输装置或类似者。在一些情况下,源装置12及目的地装置14可经装备以用于无线通信。

[0056] 目的地装置14可经由计算机可读媒体16接收待解码的经编码视频数据。计算机可读媒体16可包括能够将经编码视频数据从源装置12移动到目的地装置14的任一类型的媒体或装置。在一个实例中,计算机可读媒体16可包括通信媒体以使源装置12能够实时地将经编码视频数据直接传输到目的地装置14。可根据例如无线通信协议的通信标准调制经编码视频数据,且将其传输到目的地装置14。通信媒体可包括任何无线或有线通信媒体,例如,射频(RF)频谱或一或多个物理发射线。通信媒体可形成基于包的网路(例如局域网、广域网或例如因特网的全球网路)的一部分。通信媒体可包含路由器、交换器、基站或可用于促进从源装置12到目的地装置14的通信的任何其它设备。

[0057] 在一些实例中,经编码数据可从输出接口22输出到存储装置。类似地,可由输入接口从存储装置存取经编码数据。存储装置可包含多种分散式或本地存取的数据存储媒体中的任一者,例如,硬盘驱动器、蓝光碟片、DVD、CD-ROM、快闪存储器、易失性或非易失性存储器或用于存储经编码视频数据的任何其它合适数字存储媒体。在再一实例中,存储装置可对应于文件服务器或可存储由源装置12产生的经编码视频的另一中间存储装置。目的地装置14可经由流式传输或下载从存储装置存取所存储的视频数据。文件服务器可以是能够存储经编码视频数据且将所述经编码视频数据传输到目的地装置14的任何类型的服务器。实例文件服务器包含网页服务器(例如,用于网站)、FTP服务器、网络附接存储(NAS)装置或本地磁盘驱动器。目的地装置14可通过任何标准数据连接(包含因特网连接)而存取经编码视频数据。此数据连接可包含无线信道(例如,Wi-Fi连接)、有线连接(例如,DSL、电缆调制解调器,等等),或两者的适合于存取存储于文件服务器上的经编码视频数据的组合。从存储装置的经编码视频数据的传输可为流式传输发射、下载发射或其组合。

[0058] 本发明的技术不必限于无线应用或设定。所述技术可应用于支持多种多媒体应用中的任一者的视频译码,例如,空中电视广播、有线电视发射、卫星电视发射、因特网流式视频发射(例如,经由HTTP的动态自适应流式传输(DASH))、经编码到数据存储媒体上的数字视频、存储在数据存储媒体上的数字视频的解码或其它应用。在一些实例中,系统10可经配置以支持单向或双向视频发射以支持例如视频流式传输、视频重放、视频广播及/或视频电话的应用。

[0059] 在图1的实例中,源装置12包含视频源18、视频编码器20及输出接口22。目的地装置14包含输入接口28、视频解码器30及显示装置32。根据本发明,源装置12的视频编码器20可经配置以应用用于高级时间运动向量预测(ATMVP)的技术。在其它实例中,源装置及目的地装置可包含其它组件或布置。举例来说,源装置12可从外部视频源18(例如外部相机)接收视频数据。同样地,目的地装置14可与外部显示装置间接,而非包含集成式显示装置。

[0060] 图1的所说明系统10仅为一个实例。用于高级时间运动向量预测(ATMVP)的技术可通过任何数字视频编码及/或解码装置来执行。尽管本发明的技术一般由视频编码装置执行,但所述技术还可由视频编码器/解码器(通常被称作“编解码器”)执行。此外,本发明的技术还可由视频预处理器执行。源装置12及目的地装置14仅为源装置12产生经译码视频数据用于发射到目的地装置14的这些译码装置的实例。在一些实例中,装置12、14可以大体上对称的方式操作,使得装置12、14中的每一者包含视频编码及解码组件。因此,系统10可支持视频装置12、14之间的单向或双向视频发射以用于(例如)视频流式传输、视频播放、视频广播或视频电话。

[0061] 源装置12的视频源18可包含例如相机的视频捕获装置、含有先前捕获的视频的视频存档及/或用以从视频内容提供者接收视频的视频馈入接口。作为另一替代方案,视频源18可产生基于计算机图形的数据作为源视频,或实况视频、存档视频及计算机产生的视频的组合。在一些情况下,如果视频源18是相机,那么源装置12及目的地装置14可形成所谓的相机电话或视频电话。然而,如上文所提及,本发明中所描述的技术一般可适用于视频译码,且可适用于无线及/或有线应用。在每一情况下,经捕获、预先捕获或计算机产生的视频可由视频编码器20编码。经编码视频信息可接着由输出接口22输出到计算机可读媒体16上。

[0062] 计算机可读媒体16可包含例如无线广播或有线网络发射的瞬时媒体,或例如硬盘、闪存驱动器、紧密光盘、数字影音光盘、蓝光光盘的存储媒体(即,非暂时性存储媒体),或其它计算机可读媒体。在一些实例中,网络服务器(未图示)可从源装置12接收经编码视频数据,且(例如)经由网络发射将经编码视频数据提供到目的地装置14。类似地,媒体产生设施(例如光盘冲压设施)的计算装置可从源装置12接收经编码视频数据且生产含有经编码视频数据的光盘。因此,在各种实例中,可理解计算机可读媒体16包含各种形式的一或多个计算机可读媒体。

[0063] 目的地装置14的输入接口28从计算机可读媒体16接收信息。计算机可读媒体16的信息可包含由视频编码器20定义的语法信息,所述语法信息还由视频解码器30使用,包含描述块及其它经译码单元(例如,GOP)的特性及/或处理的语法元素。显示装置32将经解码视频数据显示给用户,且可包括多种显示装置中的任一者,例如,阴极射线管(CRT)、液晶显示器(LCD)、等离子体显示器、有机发光二极管(OLED)显示器或另一类型的显示装置。

[0064] 视频编码器20及视频解码器30可根据例如高效率视频译码 (HEVC) 标准、HEVC标准的扩展或后续标准 (例如, ITU-T H.266) 的视频译码标准操作。替代地, 视频编码器20及视频解码器30可根据例如替代地被称作MPEG-4第10部分高级视频译码 (AVC) 的ITU-T H.264标准的其它专有或行业标准或这些标准的扩展而操作。然而, 本发明的技术不限于任何特定译码标准。视频译码标准的其它实例包含MPEG-2及ITU-T H.263。尽管图1中未展示, 但在一些方面中, 视频编码器20及视频解码器30可各自与音频编码器及解码器集成, 且可包含适当多路复用器-多路分用器单元或其它硬件及软件以处置共同数据流或单独数据流中的音频及视频两者的编码。如果适用, 那么多路复用器-多路分用器单元可遵照ITU H.223多路复用器协议或例如用户数据报协议 (UDP) 的其它协议。

[0065] 视频编码器20及视频解码器30各自可实施为多种合适的编码器电路中的任一者, 例如, 一或多个微处理器、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。当所述技术以软件部分实施时, 装置可将针对软件的指令存储于合适的非暂时性计算机可读媒体中, 且在硬件中使用一或多个处理器执行指令以执行本发明的技术。视频编码器20及视频解码器30中的每一者可包含于一或多个编码器或解码器中, 编码器或解码器中的任一者可集成为相应装置中的组合式编码器/解码器 (编解码器) 的部分。

[0066] 一般来说, HEVC标准描述视频帧或图片可被划分成包含明度样本及色度样本两者的一连串树块或最大译码单元 (LCU)。位流内的语法数据可定义LCU的大小, LCU就像像素的数目来说是最大译码单元。切片包含按译码次序的许多连续树块。视频帧或图片可分割成一或多个切片。每一树块可根据四分树而分裂成若干译码单元 (CU)。一般来说, 四分树数据结构每CU包含一个节点, 其中根节点对应于树块。如果CU分裂成四个子CU, 那么对应于所述CU的节点包含四个叶节点, 所述四个叶节点中的每一者对应于所述子CU中的一者。

[0067] 所述四分树数据结构中的每一节点可提供对应CU的语法数据。举例来说, 所述四分树中的节点可包含分裂旗标, 从而指示对应于所述节点的CU是否分裂成子CU。针对CU的语法元素可经递回地定义, 且可取决于所述CU是否分裂成子CU。如果CU未经进一步分裂, 那么所述CU被称作叶CU。在本发明中, 即使不存在原始叶CU的明显分裂, 叶CU的四个子CU也将被称作叶CU。举例来说, 如果 $16 \times 16$ 大小的CU未进一步分裂, 那么四个 $8 \times 8$ 子CU也将被称作叶CU, 尽管所述 $16 \times 16$ CU从未分裂。

[0068] 除CU不具有大小区别外, H.265中的CU具有与H.264标准的宏块类似的用途。举例来说, 可将树块分裂成四个子节点 (还称作子CU), 且每一子节点又可为父代节点且可被分裂成另外四个子节点。被称作四分树的叶节点的最终的未分裂子节点包括译码节点, 所述译码节点也被称作叶CU。与经译码位流相关联的语法数据可定义被称作最大CU深度的可分裂一树块的最大次数, 且还可定义所述译码节点的最小大小。因此, 位流亦可定义最小译码单元 (SCU)。本发明使用术语“块”来指HEVC的上下文中的CU、PU或TU中的任一者, 或其它标准 (例如, H.264/AVC中的宏块及其子块) 的上下文中的类似数据结构。

[0069] CU包含译码节点以及与所述译码节点相关联的预测单元 (PU) 及变换单元 (TU)。CU的大小对应于译码节点的大小, 且形状必须为正方形。CU的大小可在从 $8 \times 8$ 像素达到具有最大 $64 \times 64$ 像素或大于 $64 \times 64$ 像素的树块的大小的范围内。每一CU可含有一或多个PU及一或多个TU。与CU相关联的语法数据可描述 (例如) CU到一或多个PU的分割。分割模式可在CU

经跳过或直接模式编码、帧内预测模式编码抑或帧间预测模式编码之间不同。PU可分割成非正方形形状。与CU相关联的语法数据还可描述(例如)CU根据四分树到一或多个TU的分割。TU的形状可为正方形或非正方形(例如,矩形)。

[0070] HEVC标准允许根据TU进行变换,所述变换对于不同CU可不同。通常基于针对经分割LCU所定义的给定CU内的PU的大小来对TU设定大小,尽管可能并非总是此状况。TU的大小通常与PU相同或比PU小。在一些实例中,可使用被称为“残余四分树”(RQT)的四分树结构而将对应于CU的残余样本再分为更小单元。可将RQT的叶节点称作变换单元(TU)。与TU相关联的像素差值可经变换以产生可加以量化的变换系数。

[0071] 叶CU可包含一或多个预测单元(PU)。一般来说,PU表示对应于对应CU的全部或一部分的空间区域,且可包含用于检索PU的参考样本的数据。此外,PU包含与预测有关的数据。举例来说,当PU经帧内模式编码时,PU的数据可包含于残余四分树(RQT)中,所述RQT可包含描述用于对应于PU的TU的帧内预测模式的数据。作为另一实例,当PU经帧间模式编码时,PU可包含定义PU的一或多个运动向量的数据。定义PU的运动向量的数据可描述(例如)运动向量的水平分量、运动向量的垂直分量、运动向量的分辨率(例如,四分之一像素精度或八分之一像素精度)、运动向量指向的参考图片,及/或运动向量的参考图片列表(例如,列表0、列表1或列表C)。

[0072] 具有一或多个PU的叶CU还可包含一或多个变换单元(TU)。如上文所论述,可使用RQT(还称作TU四分树结构)来指定所述变换单元。举例来说,分裂旗标可指示叶CU是否分裂成四个变换单元。接着,可将每一变换单元进一步分裂为其它子TU。当TU未进一步分裂时,可将其称作叶TU。一般来说,对于帧内译码来说,属于叶CU的所有叶TU共享同一帧内预测模式。也就是说,通常应用同一帧内预测模式来计算叶CU的所有TU的预测值。对于帧内译码,视频编码器可使用帧内预测模式将每一叶TU的残余值计算为CU对应于TU的部分与原始块之间的差。TU不必限于PU的大小。因此,TU可大于或小于PU。对于帧内译码,PU可与用于同一CU的对应叶TU共置。在一些实例中,叶TU的最大大小可对应于对应叶CU的大小。

[0073] 此外,叶CU的TU还可与称作残余四分树(RQT)的相应四分树数据结构相关联。也就是说,叶CU可包含指示叶CU如何被分割成TU的四分树。TU四分树的根节点通常对应于叶CU,而CU四分树的根节点通常对应于树块(或LCU)。将RQT的未被分裂的TU称作叶TU。一般来说,除非另有注释,否则本发明分别使用术语CU及TU来指叶CU及叶TU。

[0074] 视频序列通常包含一系列视频帧或图片。图片群组(GOP)通常包括一系列视频图片中的一或多者。GOP可包含语法数据于GOP的标头、图片中的一或多者的标头中或别处,所述语法数据描述包含于GOP中的图片的数目。图片的每一切片可包含描述所述相应切片的编码模式的切片语法数据。视频编码器20通常对个别视频切片内的视频块进行操作,以便编码视频数据。视频块可对应于CU内的译码节点。视频块可具有固定或变化的大小,且可根据指定译码标准而大小不同。

[0075] 作为一实例,HM支持以各种PU大小的预测。假定特定CU的大小为 $2N \times 2N$ ,那么HM支持以 $2N \times 2N$ 或 $N \times N$ 的PU大小的帧内预测,及以 $2N \times 2N$ 、 $2N \times N$ 、 $N \times 2N$ 或 $N \times N$ 的对称PU大小的帧间预测。HM还支持以 $2N \times nU$ 、 $2N \times nD$ 、 $nL \times 2N$ 及 $nR \times 2N$ 的PU大小的帧间预测的不对称分割。在不对称分割中,CU的一个方向未分割,而另一方向分割成25%及75%。CU的对应于25%分区的部分由“n”随后“上(Up)”、“下(Down)”、“左(Left)”或“右(Right)”的指示来指

示。因此,例如,“ $2N \times nU$ ”指水平方向上以顶部 $2N \times 0.5N$  PU及底部 $2N \times 1.5N$  PU分割的 $2N \times 2N$  CU。

[0076] 在本发明中,“ $N \times N$ ”与“ $N$ 乘 $N$ ”可互换地使用以指视频块在垂直尺寸与水平尺寸方面的像素尺寸,例如, $16 \times 16$ 像素或16乘16像素。一般来说, $16 \times 16$ 块在垂直方向上将具有16个像素( $y=16$ )且在水平方向上将具有16个像素( $x=16$ )。同样地, $N \times N$ 块通常在垂直方向上将具有 $N$ 个像素且在水平方向上将具有 $N$ 个像素,其中 $N$ 表示非负整数值。可按行及列来布置块中的像素。此外,块未必需要在水平方向上与垂直方向上具有相同数目个像素。举例来说,块可包括 $N \times M$ 像素,其中 $M$ 未必等于 $N$ 。

[0077] 在使用CU的PU的帧内预测性或帧间预测性译码之后,视频编码器20可计算CU的TU的残余数据。PU可包括描述在空间域(还被称作像素域)中产生预测性像素数据的方法或模式的语法数据,且TU可包括在对残余视频数据应用变换(例如离散余弦变换(DCT)、整数变换、小波变换或概念上类似的变换)之后变换域中的系数。残余数据可对应于未经编码图片的像素与对应于PU的预测值之间的像素差。视频编码器20可形成包含CU的残余数据的TU,且接着变换所述TU以产生CU的变换系数。

[0078] 在进行用以产生变换系数的任何变换之后,视频编码器20可执行对变换系数的量化。量化通常指变换系数经量化以可能减少用以表示变换系数的数据的量,从而提供进一步压缩的过程。量化过程可减小与一些或所有系数相关联的位深度。举例来说,可在量化期间将 $n$ 位值下舍入到 $m$ 位值,其中 $n$ 大于 $m$ 。

[0079] 在量化之后,视频编码器可扫描变换系数,从而从包含经量化变换系数的二维矩阵产生一维向量。所述扫描可经设计以将较高能量(且因此较低频率)系数置于阵列前部,且将较低能量(且因此较高频率)系数置于阵列后部。在一些实例中,视频编码器20可利用预定义扫描次序来扫描经量化的变换系数以产生可经熵编码的串行化向量。在其它实例中,视频编码器20可执行自适应扫描。在扫描经量化变换系数以形成一维向量之后,视频编码器20可(例如)根据上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码方法来对一维向量进行熵编码。视频编码器20还可熵编码与经编码视频数据相关联的语法元素以供视频解码器30用于解码视频数据中。

[0080] 为了执行CABAC,视频编码器20可将上下文模型内的上下文分配给待发射的符号。上下文可涉及(例如)符号的邻近值是否为非零。为了执行CAVLC,视频编码器20可选择用于待发射的符号的可变长度码。可将VLC中的码字构建成使得相对较短的码对应于更有可能的符号,而较长码对应于较不可能的符号。以此方式,相对于(例如)针对待发射的每一符号使用相等长度码字,使用VLC可实现位节省。概率确定可以是基于分配给符号的上下文。

[0081] 2016年1月25日申请的美国申请案第15/005,564号(下文中,“’564申请案”)描述除本发明的技术以外的以下技术,所述技术可通过视频编码器20及/或视频解码器30单独或以任何组合方式执行。具体地说,’564申请案描述与ATMVP候选者在经插入(例如)作为合并候选者列表的情况下的位置相关的技术。假定空间候选者及TMVP候选者以某一次序插入到合并候选者列表中。ATMVP候选者可被插入到那些候选者的任何相对固定位置中。在一个替代方案中,例如ATMVP候选者可在前两个空间候选者(例如,A1及B1)之后被插入到合并候选者列表中。在一个替代方案中,例如ATMVP候选者可被插入到前三个空间候选者(例如,A1

及B1及B0)之后。在一个替代方案中,例如ATMVP候选者可被插入到前四个空间候选者(例如,A1、B1、B0及A0)之后。在一个替代方案中,例如ATMVP候选者可紧先于TMVP候选者而被插入。在一个替代方案中,例如ATMVP候选者可紧后于TMVP候选者被插到。替代地,候选者列表中的ATMVP候选者的位置可在位流中用信号表示。其它候选者(包含TMVP候选者)的位置可另外用信号表示。

[0082] '564申请案还描述与ATMVP候选者的可用性检查可通过存取仅运动信息的一个集合而应用有关的技术,视频编码器20及/或视频解码器30可经配置以执行所述技术。当此信息集合不可用(例如,一个块经帧内译码)时,整个ATMVP候选者被视为不可用。在所述情况下,ATMVP将不被插入到合并列表中。中心位置或中心子PU仅仅用于检查ATMVP候选者的可用性。当使用中心子PU时,选择中心子PU为涵盖中心位置(例如,中心3位置,具有对PU的左上样本的相对坐标( $W/2, H/2$ ),其中 $W \times H$ 为PU的大小)的一个位置。此位置或中心子PU可连同时间向量一起使用以识别运动源图片中的对应块。识别来自涵盖对应块的中心位置的块的运动信息集合。

[0083] '564申请案还描述来自子PU的经ATMVP译码的PU的运动信息的代表集合的技术,视频编码器20及/或视频解码器30可经配置以执行所述技术。为了形成ATMVP候选者,首先形成运动信息的代表集合。可从固定位置或固定子PU导出运动信息的此代表集合。如上文所描述,可以与用于确定ATMVP候选者的可用性的运动信息的集合的所述方式相同的方式来选择所述固定位置或固定子PU。当子PU已识别其自身运动信息的集合且其为不可用时,其经设定成等于运动信息的代表集合。如果运动信息的代表集合经设定成子PU的所述代表运动信息集合,那么在最差情境下,在解码器侧并不需要额外运动存储以用于当前CTU或切片。当解码过程需要整个PU由一个运动信息集合表示(包含修剪)时,此代表运动信息集合用于所有情境中,以使得过程用于产生组合式双预测性合并候选者。

[0084] '564申请案还描述与可如何利用TMVP候选者修剪ATMVP候选者以及可如何考虑TMVP与ATMVP之间的交互有关的技术,可通过视频编码器20及/或视频解码器30执行所述技术。基于子PU的候选者(例如,具有普通候选者的ATMVP候选者)的修剪可通过使用此基于子PU的候选者的代表运动信息集合(如项目符号#3中)来进行。如果此运动信息集合与普通合并候选者相同,那么两个候选者被视为相同的。替代地或另外,执行检查以确定ATMVP是否含有多个子PU的多个不同运动信息集合;如果识别出至少两个不同集合,那么基于子PU的候选者并不用于修剪(即,被视为与任何其它候选者不同);否则,其可用于修剪(例如,可在修剪过程期间被修剪)。替代地或另外,可用空间候选者(例如,仅左及上候选者)修剪ATMVP候选者,其中位置指明为A1及B1。替代地,仅一个候选者由时间参考形成,从而为ATMVP候选者或TMVP候选者。当ATMVP可用时,候选者为ATMVP;否则,候选者为TMVP。此候选者在类似于TMVP的位置的位置中经插入到合并候选者列表中。在此情况下,候选者的最大数目可保持不变。替代地,即使当ATMVP不可用时,TMVP仍始终被禁用。替代地,仅当ATMVP不可用时使用TMVP。替代地,当ATMVP可用且TMVP不可用时,一个子PU的一个运动信息集合被用作TMVP候选者。此外,在此情况下,不应用ATMVP与TMVP之间的修剪过程。替代地或另外,用于ATMVP的时间向量还可用于TMVP,以使得并不需要使用如用于HEVC中的当前TMVP的右下位置或中心3位置。替代地,由时间向量识别的位置及右下及中心3位置联合地被考虑以提供可用的TMVP候选者。

[0085] '564申请案还描述可如何支持ATMVP的多个可用性检查以给予较高机会得到更准确且有效的ATMVP候选者,前述操作可通过视频编码器20及/或视频解码器30执行。当来自如由第一时间向量(例如,如图9中所展示)识别的运动源图片的当前ATMVP候选者不可用时,其它图片可被视为运动源图片。当考虑另一图片时,其可与不同第二时间向量相关联,或可仅与第二时间向量相关联,所述第二时间向量从指向不可用ATMVP候选者的第一时间向量按比例缩放。第二时间向量可识别第二运动源图片中的ATMVP候选者,且可应用相同可用性检查。如果如从第二运动源图片导出的ATMVP候选者为可用的,那么ATMVP候选者被导出且不需要检查其它图片;否则,需要检查作为运动源图片的其它图片。待检查的图片可为当前图片的参考图片列表中的具有给定次序的那些图片。对于每一列表,以参考索引的升序来检查图片。首先检查列表X且接着检查列表Y(为1-X)中的图片。选择列表X以使得列表X为含有用于TMVP的共置图片的列表。替代地,X仅设定成1或0。待检查的图片可以给定次序包含通过空间相邻者的运动向量所识别的那些图片。当前ATMVP适用于的PU的分区可为 $2N \times 2N$ 、 $N \times N$ 、 $2N \times N$ 、 $N \times 2N$ 或其它AMP分区(例如 $2N \times N/2$ )。替代地或另外,如果可允许其它分区大小,那么还可支持ATMVP且可包含此大小(例如,  $64 \times 8$ )。替代地,所述模式可仅应用于某些分区,例如 $2N \times 2N$ 。

[0086] '564申请案还描述可如何使用不同类型合并模式标记ATMVP候选者,视频编码器20及/或视频解码器30可经配置以执行以上操作。

[0087] 当从相邻者识别向量(如在第一阶段中的时间向量)时,可按次序检查多个相邻位置,例如用于合并候选者列表构建中的那些位置。对于相邻者中的每一者,可按次序检查对应于参考图片列表0(列表0)或参考图片列表1(列表1)的运动向量。当两个运动向量可用时,可首先检查列表X中的运动向量,且接着检查列表Y(其中Y等于1-X)中的运动向量,以使得列表X为含有用于TMVP的共置图片的列表。在ATMVP中,时间向量经使用、作为子PU的任何中心位置的移位而添加,其中时间向量的分量可需要移位到整数。此经移位中心位置用于识别运动向量可经分配给的最小单元,例如,具有涵盖当前中心位置的 $4 \times 4$ 大小的最小单元。替代地,可在对应于列表1的那些运动向量之前检查对应于列表0的运动向量。替代地,可在对应于列表0的那些运动向量之前检查对应于列表1的运动向量。替代地,按次序检查所有空间相邻者中的对应于列表X的所有运动向量,接着检查对应于列表Y(其中Y等于1-X)的运动向量。此处,X可为指示共置图片属于的列表,或仅仅被设定成0或1。空间相邻者的次序可与用于HEVC合并模式中的所述次序相同。

[0088] '564申请案还描述可通过视频编码器20及/或视频解码器30执行的与以下操作相关的技术:当识别时间向量的第一阶段中不包含识别参考图片时,如图9中所示的运动源图片可仅设定成固定图片,例如,用于TMVP的共置图片。在此情况下,可仅从指向此固定图片的运动向量来识别向量。在此情况下,向量可仅从指向任一图片的运动向量来识别,但进一步朝向固定图片按比例缩放。当在识别向量的第一阶段中由识别参考图片(如图9中所示的运动源图片)组成时,以下额外检查中的一或多者可适用于候选者运动向量。如果运动向量与经帧内译码的图片或切片相关联,那么此运动向量被视为不可用且不可用于转换为向量。如果运动向量识别相关联图片中的帧内块(通过例如将当前中心坐标与运动向量相加),那么此运动向量被视为不可用且不可用于转换为向量。

[0089] '564申请案还描述可通过视频编码器20及/或视频解码器30执行的与以下操作相



关的技术:当在识别向量的第一阶段中时,向量的分量可设定成(当前PU的半宽度,当前PU的半高度),使得其识别运动源图片中的右下像素位置。此处,(x,y)指示一个运动向量的水平分量及垂直分量。替代地,向量的分量可设定成(总和(当前PU的半宽度M)、总和(当前PU的半高度N)),其中函数总和(a,b)传回a及b的总和。在一个实例中,当运动信息存储于 $4 \times 4$ 单元中时,M及N两者都被设定成等于2。在另一实例中,当运动信息存储于 $8 \times 8$ 单元中时,M及N两者被设定成等于4。

[0090] '564申请案还描述与当ATMVP应用时的子块/子PU大小在参数集(例如,图片参数集的序列参数集)中用信号表示相关的技术,可通过视频编码器20及/或视频解码器30执行所述技术。所述大小范围介于最小PU大小到CTU大小。还可预定义或用信号表示所述大小。所述大小可(例如)小达 $4 \times 4$ 。替代地,可基于PU或CU的大小导出子块/子PU大小。举例来说,子块/子PU可被设定成等于最大( $4 \times 4$ 、(CU的宽度) $\gg M$ )。可在位流中预定义或用信号表示M的值。

[0091] '564申请案还描述与合并候选者的最大数目归因于ATMVP可被视为新的合并候选者而增加1有关的技术,所述技术可通过视频编码器20及/或视频解码器30执行。举例来说,相较于在修剪之后占据合并候选者列表中的5个候选者的HEVC,合并候选者的最大数目可增加到6。替代地,可针对ATMVP执行用常规TMVP候选者进行修剪或用常规TMVP候选者进行统一,以使得合并候选者的最大数目可保持不变。替代地,当ATMVP被识别为可用时,从合并候选者列表排除空间相邻候选者,例如,排除按提取次序的最后空间相邻候选者。

[0092] '564申请案还描述可通过视频编码器20及/或视频解码器30执行的与以下操作相关的技术:当多个空间相邻运动向量经考虑以导出时间向量时,可基于当前PU的相邻运动向量以及通过被设定为等于运动向量的特定时间向量识别的相邻运动向量来计算运动向量类似性。可选择产生最高运动类似性的一个时间向量作为最终时间向量。在一个替代方案中,对于来自相邻位置N的每一运动向量,运动向量识别运动源图片中的块(与当前PU相同的大小),其中其相邻位置N含有运动信息的集合。此运动向量集合与如当前块的相邻位置N中的运动信息的集合进行比较。在另一替代方案中,对于来自相邻位置N的每一运动向量,所述运动向量识别运动源图片中的块,其中其相邻位置含有多个运动信息集合。这些多个运动向量集合与来自相同相对位置中的当前PU的相邻位置的多个运动信息集合进行比较。

[0093] 可根据以上技术计算运动信息类似性。举例来说,当前PU具有指明为MIA1、MIB1、MIA0及MIB0的来自A1、B1、A0及B0的以下运动信息集合。对于时间向量TV,其识别对应于运动源图片中的PU的块。此块具有来自相同相对A1、B1、A0及B0位置且指明为TMIA1、TMIB1、TMIA0及TMIB0的运动信息。如通过TV所确定的运动类似性被计算为 $MStv = \sum_{N \in \{A1, B1, A0, B0\}} MVS_{im}(MI_N, TMI_N)$ ,其中 $MVS_{im}()$ 定义运动信息的两个集合(MIN, TMIN)之间的类似性。在以上两个情况中,可使用运动类似性 $MVS_{im}$ ,其中两个输入参数为两个运动信息,每一运动信息有达两个运动向量及两个参考索引。由于列表X中的每一对运动向量实际上与不同图片的不同列表X中的参考图片、当前图片及运动源图片相关联。

[0094] 对于两个运动向量MVXN及TMVXN(其中X等于0或1)中的每一者,运动向量差MVDXN可根据上述技术计算为 $MVXN - TMVXN$ 。随后,差 $MVS_{imX}$ 计算为(例如) $abs(MVDX_N[0]) + abs(MVDX_N[1])$ ,或 $(MVDX_N[0] * MVDX_N[0] + MVDX_N[1] * MVDX_N[1])$ 。如果两个运动信息集合含有可用



运动向量,那么MVS<sub>Sim</sub>设定为等于MVS<sub>Sim0</sub>+MVS<sub>Sim1</sub>。为了具有运动差的统一计算,运动向量中的两者需要向相同固定图片按比例缩放,所述固定图片可例如为当前图片的列表X的第一参考图片RefPicListX[0]。如果来自第一集合的列表X中的运动向量的可用性及来自第二集合的列表X中的运动向量的可用性不同,即一个参考索引为-1而另一个并非-1,那么此两个运动信息集合被视为在方向X上不类似。

[0095] 如果两个集合在两者设定上不类似,那么最终MVS<sub>Sim</sub>函数可根据上述技术传回大值T,所述大值T可(例如)被视为无穷大。替代地,对于一对运动信息集合,如果从列表X(X等于0或1)但并非从列表Y(Y等于1-X)来预测一集合且另一集合具有相同状态,那么可使用1与2之间的加权(例如,MVS<sub>Sim</sub>等于MVS<sub>SimX</sub>\*1.5)。当一个集合仅从列表X预测且另一个仅从列表Y预测时,MVS<sub>Sim</sub>经设定成大值T。替代地,对于任何运动信息集合,只要一个运动向量为可用的,便将产生两个运动向量。在仅一个运动向量为可用的情况下(对应于列表X),所述运动向量按比例缩放以形成对应于另一列表Y的运动向量。替代地,可基于当前PU的相邻像素与由运动向量识别的块(与当前PU大小相同)的相邻像素之间的差来测量运动向量。可选择产生最小差的运动向量作为最终时间向量。

[0096] '564申请案还描述可通过视频编码器20及/或视频解码器30执行的与以下操作相关的技术:当导出当前块的时间向量时,来自用ATMVP译码的相邻块的运动向量及/或时间向量可具有比从其它相邻块的运动向量高的优先权。在一个实例中,首先仅检查相邻块的时间向量,且第一可用的时间向量可设定成当前块的时间向量。仅当不存在此类时间向量时,进一步检查普通运动向量。在此情况下,需要存储用于经ATMVP译码的块的时间向量。在另一实例中,首先仅检查来自经ATMVP译码的相邻块的运动向量,且第一可用的运动向量可被设定成当前块的时间向量。仅当不存在此类时间向量时,进一步检查普通运动向量。在另一实例中,首先仅检查来自经ATMVP译码的相邻块的运动向量,且第一可用运动向量可被设定成当前块的时间向量。如果此类运动向量不可用,那么时间向量的检查类似于上文所论述方式而继续。在另一实例中,首先检查来自相邻块的时间向量,第一可用的时间向量可被设定成当前块的时间向量。如果此类运动向量不可用,那么时间向量的检查类似于上文所论述方式而继续。在另一实例中,首先检查经ATMVP译码的相邻块的时间向量及运动向量,第一可用的时间向量及运动向量可被设定成当前块的时间向量。仅当不存在此类时间向量及运动向量时,进一步检查普通运动向量。

[0097] '564申请案还描述与如下操作相关的技术:当多个空间相邻运动向量经考虑以导出时间向量时,可选择一运动向量使得其最小化从像素域计算的失真,例如,模板匹配可用于导出时间向量以使得产生最小匹配成本的一者经选择为最终时间向量。这些技术也可通过视频编码器20及/或视频解码器30执行。

[0098] '564申请案还描述可通过视频编码器20及/或视频解码器30执行的针对以下操作的技术:从对应块(在运动源图片中)导出运动信息集合正以一方式执行,在所述方式中,当运动向量在任何列表X的对应块中可用(指明运动向量为MVX)时,对于ATMVP候选者的当前子PU,运动向量被视为可供用于列表X(通过按比例缩放MVX)。如果运动向量不可用于任何列表X的对应块中,那么运动向量被视为不可用于列表X。替代地,当对应块中的运动向量不可用于列表X但可用于列表1-X(将1-X指明为Y且指明运动向量为MVY)时,运动向量仍被视为可供用于列表X(通过向列表X中的目标参考图片按比例缩放MVY)。替代地或另外,当列表

X及列表Y(等于1-X)的对应块中的两个运动向量为可用时,来自列表X及列表Y的运动向量不必直接用于按比例缩放以通过按比例缩放产生当前子PU的两个运动向量。在一个实例中,当阐述ATMVP候选者时,TMVP中完成的低延迟检查应用于每一子PU。如果对于当前切片的每一参考图片列表中的每一图片(由refPic指明),refPic的图片次序计数(POC)值小于当前切片的POC,那么当前切片被考虑为具有低延迟模式。在此低延迟模式中,来自列表X及列表Y的运动向量按比例缩放以分别产生列表X及列表Y的当前子PU的运动向量。当不处于低延迟模式时,仅来自MVX或MVY的一个运动向量MVZ被选择且按比例缩放以产生用于当前子PU的两个运动向量。类似于TMVP,在此情况下,Z经设定为等于collocated\_from\_l0\_flag,从而意味着其取决于如TMVP中的共置图片是在当前图片的列表X中抑或在列表Y中。替代地,Z经设定如下:如果从列表X识别出运动源图片,那么Z设定成X。替代地,另外,当运动源图片属于两个参考图片列表,且RefPicList0[idx0]为首先存在于列表0中的运动源图片且RefPicList(1)[idx1]为首先存在于列表1中的运动源图片时,Z在idx0小于或等于idx1情况下设定成0,且否则被设定成1。

[0099] '564申请案还描述用于用信号表示运动源图片的技术,所述技术可通过视频编码器20及/或视频解码器30执行。具体地说,指示运动源图片是来自列表0抑或来自列表1的旗标经用信号表示用于B切片。替代地,另外,对当前图片的列表0或列表1的参考索引可用信号表示以识别运动源图片。

[0100] '564申请案还描述可通过视频编码器20及/或视频解码器30执行的与以下操作相关的技术:当识别时间向量时,如果向量指向相关联运动源图片中的经帧内译码块,那么向量被视为不可用(因此可考虑其它向量)。

[0101] 根据本发明的技术,视频编码器20及/或视频解码器30可经配置以从空间及时间相邻块导出用于块(例如,PU)的子块(例如,子PU)的运动向量。如下文所论述,视频译码器(例如视频编码器20或视频解码器30)可从三维域中的相邻块的信息导出用于PU的每一子PU的运动向量。这意味着相邻块可以是当前图片中的空间相邻者或先前经译码图片中的时间相邻者。下文更详细地论述的图10为说明实例空间时间运动向量预测符(STMVP)导出程序的流程图。另外,上文关于项目符号#1、#2、#3、#4、#6、#7、#12及#13所描述的方法可直接扩展到STMVP。

[0102] 在以下描述中,术语“块”用以指用于存储预测相关信息(例如帧间或帧内预测、帧内预测模式、运动信息等)的块单元。此预测信息经保存且可用于译码将来块,例如预测用于将来块的预测模式信息。在AVC及HEVC中,此块的大小为 $4 \times 4$ 。

[0103] 应注意,在以下描述中,“PU”指示经帧间译码块单元及用以指示从相邻块导出运动信息的单元的子PU。

[0104] 视频编码器20及/或视频解码器30可经配置以单独或以任何组合方式应用以下方法中的任一者。

[0105] 子PU及相邻块的大小:考虑具有多个子PU的PU,子PU的大小通常等于或大于所述相邻块大小。在一个实例中,如图11A中所示,加阴影正方形表示在当前PU外部的相邻块(使用小写字母a、b...i表示),且剩余正方形(使用大写字母A、B...P表示)表示当前PU中的子PU。子PU及其相邻块的大小是相同的。举例来说,所述大小等于 $4 \times 4$ 。图11B展示子PU大于相邻块的另一实例。以此方式,用于运动信息导出的相邻块的大小可等于或小于导出运动信息

所针对的子块的大小。替代地,子PU可采用非正方形形状,例如矩形或三角形形状。此外,子PU的大小可在切片标头中用信号表示。在一些实例中,上文关于(例如)在参数集中用信号表示子块或子PU大小所论述的过程可扩展到这些技术。举例来说,子PU大小可在参数集(例如序列参数集(SPS)或图片参数集(PPS))中用信号表示。

[0106] 关于图11A的实例,假定视频译码器将光栅扫描次序(A、B、C、D、E等)应用于子PU以导出子块的运动预测。然而,还可应用其它扫描次序,且应注意这些技术不受限于仅光栅扫描次序。

[0107] 相邻块可分类成两种不同类型:空间及时间。空间相邻块为已经译码的块或在当前图片或切片中并与当前子PU相邻的已经扫描的子PU。时间相邻块为在先前经译码图片中并与当前子PU的共置块相邻的块。在一个实例中,视频译码器使用与当前PU相关联的所有参考图片以获得时间相邻块。在另一实例中,视频译码器将参考图片的子集用于STMVP导出,例如,每一参考图片列表的仅第一条目用于STMVP导出。

[0108] 在这些定义之后,对于子PU(A),进一步参看图11A,所有白块(a、b...i)及其在先前经译码图片中的共置块为视作可用的空间及时间相邻块。根据光栅扫描次序,块B、C、D、E...P并非在空间上可用于子PU(A)。然而,所有子PU(从A到P)为子PU(A)的时间上可用的相邻块,这是因为其运动信息可在先前经译码图片中的其共置块中发现。采用子PU(G)作为另一实例:其可用的空间相邻块包含从a、b...到i以及从A到F的那些块。此外,在一些实例中,某一约束可应用于空间相邻块,例如,空间相邻块(即,从a、b...到i)可约束为在同一LCU/切片/图块中。

[0109] 根据本发明的技术,视频译码器(视频编码器20或视频解码器30)可选择所有可用相邻块的子集以导出用于每一子PU的运动信息或运动字段。用于导出每一PU的子集可经预定义;替代地,视频编码器20可用信号表示所述子集(且视频解码器30可接收指示所述子集的用信号表示的数据)为切片标头、PPS、SPS或其类似者中的高电平语法。为优化译码性能,子集对于每一子PU可不同。实际上,为简单起见,子集的位置的固定图案是优选的。举例来说,每一子PU可使用其直接上方空间相邻者,其直接左空间相邻者及其直接右下时间相邻者作为子集。关于图11A的实例,当考虑子PU(J)(水平地散列)时,上方块(F)及左块(I)(左下对角地散列)为在空间上可用的相邻块,且右下块(O)(在两个方向上对角地散列)为在时间上可用的相邻块。在此子集情况下,在当前PU中的子PU归因于处理相依性而将顺序地(以所定义次序,例如光栅扫描次序)处理。

[0110] 另外或替代地,当考虑子PU(J)时,视频编码器20及视频解码器30可将上方块(F)及左块(I)看作在空间上可用的相邻块,且将底部块(N)及右块(K)看作在时间上可用的相邻块。在此子集的情况下,视频编码器20及视频解码器30可归因于处理相依性而顺序地处理当前PU中的子PU。

[0111] 为允许并行处理当前PU中的每一子PU,视频编码器20及视频解码器30可将一些子PU的不同相邻块子集用于运动预测导出。在一个实例中,可定义仅含有不属于当前PU的空间相邻块(例如块a、b...i)的子集。在此情况下,并行处理将是可能的。

[0112] 在另一实例中,对于给定子PU,如果子PU空间相邻块在当前PU内,那么彼空间相邻块的共置块可置于子集中并用以导出当前子PU的运动信息。举例来说,当考虑子PU(J)时,将上方块(F)及左块(I)及右下块(O)的时间共置块选择为子集以导出子PU(J)的运动。在此

情况下,子PU(J)的子集含有三个时间相邻块。在另一实例中,可启用部分并行过程,其中一个PU分裂成若干区且每一区(涵盖若干子PU)可经独立地处理。

[0113] 有时相邻块经帧内译码,其中需要具有用以确定针对那些块的替代运动信息用于更好地运动预测及译码效率的规则。举例来说,考虑子PU(A),可能存在块b、c及/或f经帧内译码,且a、d、e、g、h及i经帧间译码的情况。对于空间相邻者,视频编码器20及视频解码器30可使用预定义次序来用首先发现的帧间译码块的所述运动信息填充经帧内译码块的运动信息。举例来说,上方相邻者的搜索次序可经设定为从直接上方相邻者开始向右直到最右相邻者为止,从而意味着b、c、d及e的次序。左相邻者的搜索次序可经设定为从直接左相邻者开始向下直到最下相邻者为止,从而意味着f、g、h及i的次序。如果通过搜索过程未发现经帧间译码块,那么上方或左空间相邻者被考虑为不可用。对于时间相邻者,可使用与TMVP导出中指定的规则相同的规则。然而,应注意还可使用其它规则,例如基于运动方向、时间距离(在不同参考图片中的搜索)及空间位置等的规则。

[0114] 视频编码器20及视频解码器30可使用根据本发明的技术的用于导出给定子PU的运动信息的以下方法。视频编码器20及视频解码器30可首先确定目标参考图片,并执行运动向量按比例缩放。对于每一相邻块,运动向量按比例缩放可基于每一参考图片列表应用于其运动向量以便将所有相邻块的运动向量映射到每一列表中的相同参考图片。存在两个步骤:首先,确定待用于按比例缩放的源运动向量。第二,确定源运动向量投影到的目标参考图片。对于第一步,可使用若干方法:

[0115] a) 对于每一参考列表,运动向量按比例缩放独立于另一参考列表中的运动向量;对于给定块的运动信息,如果在参考列表中不存在运动向量(例如,替代双向预测模式的单向预测模式),那么无运动向量按比例缩放经执行用于所述列表。

[0116] b) 运动向量按比例缩放并不独立于另一参考列表中的运动向量;对于给定块的运动信息,如果无运动向量在参考列表中不可用,那么其可从另一参考列表中的运动向量按比例缩放。

[0117] c) 两个运动向量从一个预定义参考列表(如在TMVP中)按比例缩放

[0118] 在一个实例中,根据本发明的技术,视频编码器20及视频解码器30将上述方法a)用于按比例缩放空间相邻块的运动向量,及上述方法c)用于按比例缩放时间相邻块的运动向量。然而,可在其它实例中使用其它组合。

[0119] 对于第二步骤,可基于可用的空间相邻块的运动信息(例如参考图片)根据某一规则来选择目标参考图片。此规则的一个实例为大部分规则,即,选择通过大部分块共享的参考图片。在此情况下,不存在目标参考图片从编码器到解码器的所需要的用信号表示,这是因为还可在解码器侧使用相同规则推断相同信息。替代地,此参考图片还可在切片标头中经明确地规定,或在一些其它方法中用信号表示到解码器。在一个实例中,目标参考图片经确定为每一参考列表的第一参考图片(refidx=0)。

[0120] 在确定目标参考图片及视需要按比例缩放运动向量之后,视频编码器20及视频解码器30导出给定子PU的运动信息。假定对于给定子PU存在具有运动信息的N个可用相邻块。首先,视频编码器20及视频解码器30确定预测方向(InterDir)。用于确定预测方向的一个简单方法如下:

[0121] a. InterDir经初始化为零,接着循环穿过N个可用相邻块的运动信息;

[0122] b. 如果列表0中存在至少一个运动向量,那么 $\text{InterDir} = (\text{InterDir} \text{ bitwiseOR} 1)$ ;

[0123] c. 如果列表1中存在至少一个运动向量,那么 $\text{InterDir} = (\text{InterDir} \text{ bitwiseOR} 2)$ 。

[0124] 此处“bitwiseOR”表示逐位或操作。在此实例中,InterDir的值经定义为:0(非帧间预测),1(基于列表0的帧间预测)、2(基于列表1的帧间预测),及3(基于列表0及列表1两者的帧间预测)。

[0125] 替代地,类似于关于上文所描述的运动向量按比例缩放的目标参考图片的确定,大部分规则可用于基于所有可用相邻块的运动信息确定给定子PU的InterDir的值。

[0126] 在确定InterDir之后,可导出运动向量。对于基于导出的InterDir的每一参考列表,可存在如上文所论述的通过对目标参考图片的运动向量按比例缩放可获得的M个运动向量( $M \leq N$ )。参考列表的运动向量可导出为:

$$[0127] \quad (MV_x, MV_y) = ((\sum_{i=0}^M w_i * MV_{xi} + O_i) / \sum_{i=0}^M w_i, (\sum_{j=0}^M w_j * MV_{yj} + O_j) / \sum_{j=0}^M w_j) \quad (1)$$

[0128] 其中x和y指示运动向量的水平分量及垂直分量, $w_i$ 及 $w_j$ 分别为水平及垂直运动分量的加权因子,且 $O_i$ 及 $O_j$ 为取决于加权因子的偏移值。

[0129] 可基于各种因子确定加权因子。在一个实例中,相同规则可应用于一个PU内的所有子PU。规则可定义如下:

[0130] • 举例来说,可基于当前子PU及对应相邻块的位置距离确定加权因子。

[0131] • 在另一实例中,还可在按比例缩放之前基于目标参考图片与关联于对应相邻块的运动向量的参考图片之间的POC距离来确定加权因子。

[0132] • 在又一实例中,可基于运动向量差或一致性而确定加权因子。

[0133] • 为简单起见,所有加权因子还可设定成1。

[0134] 替代地,不同规则可应用于一个PU内的若干子PU。举例来说,可应用上述规则,另外,对于位于第一行/第一列处的子PU,从时间相邻块导出的运动向量的加权因子设定成0,而对于剩余块,从空间相邻块导出的运动向量的加权因子设定成0。

[0135] 应注意,实际上,上述方程可按原样实施,或经简化而易于实施。举例来说,为避免除法或浮点运算,固定点运算可用于模拟上述方程。一个例子是,为了避免除以3,可替代地选择乘以43/128以用乘法及位移位替换除法运算。实施方案中的那些变化应考虑为涵盖于本发明的技术的相同精神下。

[0136] 另外或替代地,当过程调用两个运动向量时,方程(1)可被以下方程(2)取代:

$$[0137] \quad (MV_x, MV_y) = ((\sum_{i=0}^1 MV_{xi})/2, (\sum_{i=0}^1 MV_{yi})/2) \quad (2)$$

[0138] 另外或替代地,当过程调用三个运动向量时,方程(1)可被以下方程(3)取代:

$$[0139] \quad (MV_x, MV_y) = ((\sum_{i=0}^2 MV_{xi} + \text{sign}(\sum_{i=0}^2 MV_{xi}) * 1) * 43/128, (\sum_{i=0}^2 MV_{yi} + \text{sign}(-\sum_{i=0}^2 MV_{yi}) * 1) * 43/128)) \quad (3)$$

[0140] 另外或替代地,当过程调用四个运动向量时,方程(1)可被以下方程(4)取代:

$$[0141] \quad (MV_x, MV_y) = ((\sum_{i=0}^3 MV_{xi} + \text{sign}(\sum_{i=0}^3 MV_{xi}) * 2) / 4,$$

$$[0142] \quad (\sum_{i=0}^3 MV_{yi} + \text{sign}(\sum_{i=0}^3 MV_{yi}) * 2) / 4) \quad (4)$$

[0143] 其中如果t是正值,那么sign(t)是1;且如果t是负值,那么sign(t)是-1。

[0144] 另外或替代地,还可应用非线性操作以导出运动向量,例如中值滤波。

[0145] 视频编码器20及视频解码器30可进一步确定这些技术的运动向量可用性。即使当每一子PU的运动向量预测符可用时,STMVP模式可经复位为对于一个PU不可用。举例来说,一旦每一子PU的运动向量预测符经导出用于给定PU,一些可用性检查便经执行以确定是否应使STMVP模式可供用于给定PU。此操作用以消除STMVP模式经最终选择用于给定PU极其不大可能的情况。当STMVP模式不可用时,模式用信号表示不包含STMVP。在STMVP模式通过在合并列表中插入SMTVP而实施情况下,当STMVP模式经确定为不可用时,合并列表不包含此STMVP候选者。因此,可减少用信号表示额外负担。

[0146] 考虑一个PU分割成M个子PU。在一个实例中,如果M个子PU当中的N1 ( $N1 \leq M$ ) 个子PU具有相同运动向量预测符(即,相同运动向量及相同参考图片索引),那么仅在N1小于阈值或预测符不同于合并列表中的其它运动向量预测符(具有较小合并索引)时使STMVP可用。在另一实例中,如果在STMVP模式下的N2 ( $N2 \leq M$ ) 个子PU共享与在ATMVP情况下的对应子PU相同的运动向量预测符,那么仅在N2小于另一阈值时使STMVP可用。在一个实例中,用于N1及N2的两个阈值都设定为等于M。

[0147] 视频编码器20及视频解码器30接着可将导出的运动预测符插入到候选者列表(例如合并列表)中。如果STMVP候选者可用,那么视频编码器20及视频解码器30可将STMVP候选者插入到候选者列表(例如,合并列表)中。在上述项目符号#1中的过程可经扩展且STMVP候选者可在ATMVP候选者之前或之后被插入。在一个实例中,视频编码器20及视频解码器30就在合并列表中的ATMVP候选者之后插入STMVP。

[0148] 视频编码器20可(例如)在帧标头、块标头、切片标头或GOP标头中进一步将语法数据(例如,基于块的语法数据、基于帧的语法数据,及基于GOP的语法数据)发送到视频解码器30。GOP语法数据可描述相应GOP中的帧数目,且帧语法数据可指示用以编码对应帧的编码/预测模式。

[0149] 视频编码器20及视频解码器30各自可实施为可适用的多种合适编码器或解码器电路中的任一者,例如一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、离散逻辑电路、软件、硬件、固件或其任何组合。视频编码器20及视频解码器30中的每一者可包含在一或多个编码器或解码器中,编码器或解码器中的任一者可经集成为组合式编码器/解码器(编解码器)的部分。包含视频编码器20及/或视频解码器30的装置可包括集成电路、微处理器及/或无线通信装置(例如蜂窝式电话)。

[0150] 图2是说明可实施用于高级时间运动向量预测(ATMVP)的技术的视频编码器20的实例的框图。视频编码器20可执行视频切片内的视频块的帧内译码及帧间译码。帧内译码依赖于空间预测以减小或去除给定视频帧或图片内的视频的空间冗余。帧间译码依赖于时间预测以减少或去除视频序列的相邻帧或图片内的视频中的时间冗余。帧内模式(I模式)可指若干基于空间的译码模式中的任一者。例如单向预测(P模式)或双向预测(B模式)的帧

间模式可指若干基于时间的译码模式中的任一者。

[0151] 如图2中所示,视频编码器20接收待编码的视频帧内的当前视频块。在图2的实例中,视频编码器20包含模式选择单元40、参考图片存储器64、求和器50、变换处理单元52、量化单元54及熵编码单元56。模式选择单元40又包含运动补偿单元44、运动估计单元42、帧内预测单元46及分割单元48。对于视频块重建,视频编码器20还包含反量化单元58、反变换单元60及求和器62。还可包含解块滤波器(图2中未展示)以对块边界进行滤波以从重建的视频去除块效应伪影。如果需要,那么解块滤波器将通常对求和器62的输出滤波。除了解块滤波器外,还可使用额外滤波器(回路中或回路后)。为简洁起见未展示这些滤波器,但如果需要,这些滤波器可对求和器50的输出滤波(作为回路内滤波器)。

[0152] 在编码程序期间,视频编码器20接收待译码的视频帧或切片。可将所述帧或切片划分成多个视频块。运动估计单元42及运动补偿单元44执行接收的视频块相对于一或多个参考帧中的一或多个块的帧间预测性译码以提供时间预测。帧内预测单元46可替代地执行接收的视频块相对于与待译码块相同的帧或切片中的一或多个相邻块的帧内预测性译码以提供空间预测。视频编码器20可执行多个译码遍次,(例如)以选择用于每一视频数据块的适当译码模式。

[0153] 此外,分割单元48可基于对先前译码遍次中的先前分割方案的评估而将视频数据的块分割为子块。举例来说,分割单元48可最初将一帧或切片分割成多个LCU,且基于率失真分析(例如,率失真优化)来将所述LCU中的每一者分割成子CU。模式选择单元40可进一步产生指示将LCU分割为子CU的四分树数据结构。四分树的叶节点CU可包含一或多个PU及一或多个TU。

[0154] 模式选择单元40可(例如)基于误差结果选择译码模式(帧内或帧间)中的一者,且将所得经帧内或经帧间译码块提供到求和器50以产生残余块数据,及提供到求和器62以重建用作参考帧的经编码块。模式选择单元40还将例如运动向量、帧内模式指示符、分区信息及其它此类语法信息的语法元素提供到熵编码单元56。

[0155] 运动估计单元42及运动补偿单元44可高度集成,但为概念目的而分开来说明。由运动估计单元42执行的运动估计为产生运动向量的过程,所述运动向量估计视频块的运动。举例来说,运动向量可指示当前视频帧或图片内的视频块的PU相对于在参考帧(或其它经译码单元)内的预测性块的位移,所述预测性块是相对于所述当前帧(或其它经译码单元)内正经译码的当前块。预测性块是就像素差来说被发现紧密地匹配于待译码块的块,所述像素差可通过绝对差和(SAD)、平方差和(SSD)或其它差度量予以确定。在一些实例中,视频编码器20可计算存储于参考图片存储器64中的参考图片的子整数像素位置的值。举例来说,视频编码器20可内插参考图片的四分之一像素位置、八分之一像素位置或其它分率像素位置的值。因此,运动估计单元42可相对于全像素位置及分率像素位置执行运动搜索且输出具有分率像素精确度的运动向量。

[0156] 运动估计单元42通过比较PU的位置与参考图片的预测性块的位置来计算经帧间译码切片中的视频块的PU的运动向量。参考图片可选自第一参考图片列表(列表0)或第二参考图片列表(列表1),其中的每一者识别存储于参考图片存储器64中的一或多个参考图片。运动估计单元42将所计算运动向量发送到熵编码单元56及运动补偿单元44。

[0157] 由运动补偿单元44执行的运动补偿可涉及基于由运动估计单元42确定的运动向

量提取或产生预测性块。而且,在一些实例中,运动估计单元42与运动补偿单元44可在功能上集成。在接收到当前视频块的PU的运动向量之后,运动补偿单元44可在参考图片列表中的一者中定位运动向量指向的预测性块。求和器50通过从正经译码的当前视频块的像素值减去预测性块的像素值来形成残余视频块,从而形成像素差值,如下文所论述。一般来说,运动估计单元42相对于明度分量执行运动估计,且运动补偿单元44将基于明度分量计算的运动向量用于色度分量与明度分量两者。模式选择单元40还可产生与视频块及视频切片相关联的语法元素以供视频解码器30在解码视频切片的视频块中使用。

[0158] 模式选择单元40还可选择子块(例如,子PU)运动导出模式用于块(例如,PU)。也就是说,模式选择单元40可在一系列编码遍次之间比较多种编码因子(包含预测模式)以确定哪一编码遍次(且因此,哪一因子集合,包含哪一预测模式)产生所需要率失真优化(RDO)特性。当模式选择单元40选择子块运动信息导出模式用于视频数据块(例如,PU)时,运动补偿单元44可使用本发明的技术来预测块。

[0159] 具体地说,使用子块运动信息导出模式,运动补偿单元44可导出块的子块的运动信息。举例来说,运动补偿单元44可针对每一子块确定两个或更多个相邻子块的运动信息并从相邻子块的运动信息导出子块的运动信息。举例来说,相邻子块可包含空间及/或时间相邻子块。在一个实例中,运动补偿单元44通过平均左相邻空间子块、上方相邻空间子块及右下时间相邻子块的运动信息(例如,运动向量)而导出每一子块的运动信息,如下文关于图11A更详细地论述。在其它实例中,运动补偿单元44可使用(例如)式(1)到(4)中的一者导出每一子块的运动信息。运动补偿单元44可将导出的运动信息用于子块中的每一者以确定用于子块的预测数据。通过检索用于子块中的每一者的此预测数据,运动补偿单元44使用子块运动信息导出模式产生当前块的经预测块。

[0160] 如上文所描述,作为由运动估计单元42及运动补偿单元44执行的帧间预测的替代方案,帧内预测单元46可对当前块进行帧内预测。具体地说,帧内预测单元46可确定待用以编码当前块的帧内预测模式。在一些实例中,帧内预测单元46可(例如)在单独编码遍次期间使用各种帧内预测模式来编码当前块,且帧内预测单元46(或在一些实例中模式选择单元40)可从所测试模式中选择适当帧内预测模式来使用。

[0161] 举例来说,帧内预测单元46可使用对于各种所测试的帧内预测模式的率失真分析来计算率失真值,且在所测试的模式的中选择具有最优选率失真特性的帧内预测模式。率失真分析大体上确定经编码块与原始、未编码块(其经编码以产生经编码块)之间的失真(或误差)量,以及用以产生经编码块的位率(即,位的数目)。帧内预测单元46可根据各种经编码块的失真及速率来计算比率以确定哪一帧内预测模式展现所述块的最优选率失真值。

[0162] 在选择用于块的帧内预测模式后,帧内预测单元46可将指示块的所选帧内预测模式的信息提供到熵编码单元56。熵编码单元56可编码指示所选帧内预测模式的信息。视频编码器20可在所发射的位流中包含以下各者:配置数据,其可包含多个帧内预测模式索引表及多个经修改的帧内预测模式索引表(还称作码字映射表);各种块的编码上下文的定义;及待用于所述上下文中的每一者的最有可能的帧内预测模式、帧内预测模式索引表及经修改的帧内预测模式索引表的指示。

[0163] 视频编码器20通过从正被译码的原始视频块减去来自模式选择单元40的预测数据而形成残余视频块。求和器50表示执行此减法运算的一或多个组件。变换处理单元52将



变换(例如离散余弦变换(DCT)或概念上类似的变换)应用于残余块,从而产生包括残余变换系数值的视频块。变换处理单元52可执行概念上类似于DCT的其它变换。还可使用小波变换、整数变换、子频带变换或其它类型的变换。在任何情况下,变换处理单元52将变换应用于残余块,从而产生残余变换系数块。所述变换可将残余信息从像素值域转换到变换域,例如频域。变换处理单元52可将所得变换系数发送到量化单元54。量化单元54量化变换系数以进一步减小位率。量化过程可减小与系数中的一些或所有相关联的位深度。可通过调整量化参数来修改量化程度。在一些实例中,量化单元54可接着对包含经量化变换系数的矩阵执行扫描。替代地,熵编码单元56可执行所述扫描。

[0164] 在量化之后,熵编码单元56熵译码经量化的变换系数。举例来说,熵编码单元56可执行上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵译码技术。在基于上下文的熵译码的情况下,上下文可以是基于相邻块。在由熵编码单元56进行熵译码之后,可将经编码位流发射到另一装置(例如,视频解码器30)或存档以供稍后发射或检索。

[0165] 此外,熵编码单元56可编码各种视频数据块的各种其它语法元素。举例来说,熵编码单元56可编码表示视频数据的每一CU的每一PU的预测模式的语法元素。当帧间预测经指示用于PU时,熵编码单元56可编码运动信息,运动信息可包含是使用合并模式还是高级运动向量预测(AMVP)来编码运动向量。在任一情况下,视频编码器20形成包含可借以预测运动信息的候选者(对于PU的空间及/或时间相邻块)的候选者列表。根据本发明的技术,候选者列表可包含指示子块运动信息导出模式将用于PU的候选者。此外,熵编码单元56可将指示将使用哪一候选者的候选者索引编码到候选者列表中。因此,如果选定子块运动信息导出模式,那么熵编码单元56编码候选者索引,所述候选者索引指表示子块运动信息导出模式的候选者。

[0166] 反量化单元58及反变换单元60分别应用反量化及反变换以在像素域中重建残余块(例如)以供稍后用作参考块。运动补偿单元44可通过将残余块添加到参考图片存储器64的帧中的一者的预测性块来计算参考块。运动补偿单元44还可将一或多个内插滤波器应用到经重建的残余块以计算用于运动估计中的次整数像素值。求和器62将经重建的残余块添加到由运动补偿单元44产生的运动补偿预测块,以产生用于存储于参考图片存储器64中的经重建的视频块。经重建的视频块可由运动估计单元42及运动补偿单元44用作参考块,以对后续视频帧中的块进行帧间译码。

[0167] 以此方式,视频编码器20表示经配置以执行以下操作的视频编码器的实例:确定当前视频数据块的运动预测候选者指示运动信息将经导出用于当前块的子块,及响应于所述确定:将当前块分割成子块,针对所述子块中的每一者,使用至少两个相邻块的运动信息导出运动信息,并使用相应导出的运动信息解码所述子块。也就是说,视频编码器20使用本发明的技术既编码且又解码视频数据的块。

[0168] 图3是说明可实施用于高级时间运动向量预测(ATMVP)的技术的视频解码器30的实例的框图。在图3的实例中,视频解码器30包含熵解码单元70、运动补偿单元72、帧内预测单元74、反量化单元76、反变换单元78、参考图片存储器82和求和器80。在一些实例中,视频解码器30可执行与关于视频编码器20(图2)描述的编码遍次大体上互逆的解码遍次。运动

补偿单元72可基于从熵解码单元70接收的运动向量产生预测数据,而帧内预测单元74可基于从熵解码单元70接收的帧内预测模式指示符产生预测数据。

[0169] 在解码过程期间,视频解码器30从视频编码器20接收表示经编码视频切片的视频块及相关联的语法元素的经编码视频位流。视频解码器30的熵解码单元70熵解码位流以产生经量化系数、运动向量或帧内预测模式指示符及其它语法元素。熵解码单元70将运动向量及其它语法元素转递到运动补偿单元72。视频解码器30可在视频切片层级及/或视频块层级接收语法元素。

[0170] 当视频切片经译码为帧内译码(I)切片时,帧内预测单元74可基于用信号表示的帧内预测模式及来自当前帧或图片的先前经解码块的数据而产生当前视频切片的视频块的预测数据。当视频帧经译码为经帧间译码(即,B、P或GPB)切片时,运动补偿单元72基于运动向量及从熵解码单元70接收的其它语法元素产生用于当前视频切片的视频块的预测性块。预测性块可从参考图片列表中的一者内的参考图片中的一者产生。视频解码器30可基于存储于参考图片存储器82中的参考图片使用默认构建技术来构建参考帧列表,列表0及列表1。

[0171] 运动补偿单元72通过剖析运动向量及其它语法元素来确定用于当前视频切片的视频块的预测信息,并使用所述预测信息以产生经解码的当前视频块的预测性块。举例来说,运动补偿单元72使用所接收语法元素中的一些来确定用以译码视频切片的视频块的预测模式(例如,帧内或帧间预测)、帧间预测切片类型(例如,B切片、P切片或GPB切片)、用于所述切片的参考图片列表中的一或多者的构建信息、用于所述切片的每一经帧间编码视频块的运动向量、用于所述切片的每一经帧间译码视频块的帧间预测状态及用以解码当前视频切片中的视频块的其它信息。

[0172] 运动补偿单元72还可执行基于内插滤波器的内插。运动补偿单元72可使用如由视频编码器20在视频块的编码期间所使用的内插滤波器,以计算参考块的次整数像素的内插值。在此情况下,运动补偿单元72可根据接收的语法元素确定由视频编码器20使用的内插滤波器且使用所述内插滤波器来产生预测性块。

[0173] 根据本发明的技术,当使用帧间预测来预测例如PU的块时,熵解码单元70解码参考候选者列表的候选者索引的值,并将候选者索引的值传递到运动补偿单元72。候选者索引的值可参考候选者列表中的候选者,所述候选者表示使用子块运动信息导出模式预测块。如果候选者索引的值确实参考候选者列表中的表示使用子块运动信息导出模式预测块的候选者,那么运动补偿单元72可使用子块运动信息导出模式产生块的经预测块。

[0174] 更明确地说,使用子块运动信息导出模式,运动补偿单元72可导出块的子块的运动信息。举例来说,运动补偿单元72可针对每一子块确定两个或更多个相邻子块的运动信息并从相邻子块的运动信息导出子块的运动信息。举例来说,相邻子块可包含空间及/或时间相邻子块。在一个实例中,运动补偿单元72通过平均左相邻空间子块、上方相邻空间子块及右下时间相邻子块的运动信息(例如,运动向量)而导出每一子块的运动信息,如下文关于图11A更详细地论述。在其它实例中,运动补偿单元72可使用例如公式(1)到(4)中的一者导出每一子块的运动信息。运动补偿单元72可使用用于子块中的每一者的导出的运动信息以确定用于子块的预测数据。通过检索用于子块中的每一者的此预测数据,运动补偿单元72使用子块运动信息导出模式产生当前块的经预测的块。

[0175] 反量化单元76反量化(即,解量化)位流中所提供并由熵解码单元70解码的经量化变换系数。反量化过程可包含使用由视频解码器30针对视频切片中的每一视频块计算的量化参数QPY以确定应应用的量化程度及(同样地)反量化程度。

[0176] 反变换单元78将反变换(例如,反DCT、反整数变换或概念上类似的反变换程序)应用于变换系数,以便在像素域中产生残余块。

[0177] 在运动补偿单元72基于运动向量及其它语法元素产生当前视频块的预测性块后,视频解码器30通过对来自反变换单元78的残余块与由运动补偿单元72产生的对应预测性块求和而形成经解码的视频块。求和器80表示执行此求和运算的一或多个组件。如果需要,还可应用解块滤波器来对经解码块滤波以便去除块效应伪影。还可使用其它回路滤波器(在译码回路中抑或在译码回路之后)以使像素转变平滑,或以其它方式改进视频质量。接着将给定帧或图片中的经解码的视频块存储于参考图片存储器82中,所述参考图片存储器存储用于后续运动补偿的参考图片。参考图片存储器82还存储经解码视频以用于稍后在显示装置(例如,图1的显示装置32)上呈现。

[0178] 以此方式,视频解码器30表示经配置以执行以下操作的视频解码器的实例:确定当前视频数据块的运动预测候选者指示运动信息将经导出用于当前块的子块,及响应于所述确定:将当前块分割成子块,针对所述子块中的每一者,使用至少两个相邻块的运动信息导出运动信息,并使用相应导出的运动信息解码所述子块。

[0179] 图4是说明HEVC中的空间相邻候选者的概念图。尽管从块产生候选者的方法对于合并模式及AMVP模式来说不同,但对于特定PU(PU0),空间MV候选者是从图4上展示的相邻块导出。

[0180] 在合并模式中,可用图4(a)中展示的具有数字的次序导出多达四个空间MV候选者,且所述次序如下:左(0,A1)、上(1,B1)、右上(2,B0)、左下(3,A0)及左上(4,B2),如图4(a)中所示。也就是说,在图4(a)中,块100包含PU0 104A及PU1 104B。当视频译码器使用合并模式译码用于PU0 104A的运动信息时,视频译码器以所述次序将来自空间相邻块108A、108B、108C、108D及108E的运动信息添加到候选者列表。如同HEVC中一样,块108A、108B、108C、108D及108E还可分别被称作块A1、B1、B0、A0及B2。

[0181] 在AMVP模式中,如图4(b)上所展示,相邻块被分成两个群组:包含块0及1的左群组及包含块2、3及4的上群组。这些块分别被标记为图4(b)中的块110A、110B、110C、110D及110E。具体地说,在图4(b)中,块102包含PU0 106A及PU1 106B,且块110A、110B、110C、110D及110E表示对PU0 106A的空间相邻者。对于每一群组,参考与由用信号表示的参考索引指示的参考图片相同的参考图片的相邻块中的潜在候选者具有待选择的最高优先权以形成所述群组的最终候选者。有可能的是,所有相邻块均不含指向同一参考图片的运动向量。因此,如果不可发现此候选者,那么将按比例缩放第一可用候选者以形成最终候选者;因此,可补偿时间距离差。

[0182] 图5是说明HEVC中的时间运动向量预测的概念图。具体地说,图5(a)说明包含PU0 122A及PU1 122B的实例CU 120。PU0 122A包含PU 122A的中心块126及对于PU0 122A的右下块124。图5(a)还展示可从PU0 122A的运动信息预测运动信息所针对的外部块128,如下文所述。图5(b)说明包含将预测运动信息所针对的当前块138的当前图片130。具体地说,图5(b)说明对当前图片130的共置图片134(包含对当前块138的共置块140)、当前参考图片

132及共置参考图片136。使用运动向量144预测共置块140,所述运动向量144用作块138的运动信息的时间运动向量预测符(TMVP)142。

[0183] 如果TMVP经启用且TMVP候选者可用,那么视频译码器可在任何空间运动向量候选者之后添加TMVP候选者(例如, TMVP候选者142)到MV候选者列表中。对于合并模式及AMVP模式两者, TMVP候选者的运动向量导出的过程是相同的。然而,根据HEVC,合并模式中的TMVP候选者的目标参考索引设定成0。

[0184] TMVP候选者导出的初始块位置是共置PU外部的右下块,如图5(a)中展示为对PU0122A的块124,以补偿用于产生空间相邻候选者的上方块及左块的偏差。然而,如果块124位于当前CTB列的外部或运动信息并不可用于块124,那么块被如图5(a)中所展示的PU的中心块126取代。

[0185] 如切片层级信息所指示,从共置图片134的共置块140导出TMVP候选者142的运动向量。

[0186] 类似于AVC中的时间直接模式, TMVP候选者的运动向量可经受运动向量按比例缩放,所述按比例缩放经执行以补偿当前图片130与当前参考图片132及共置图片134与共置参考图片136之间的图片次序计数(POC)距离差。也就是说,可按比例缩放运动向量144以基于这些POC差产生TMVP候选者142。

[0187] 下文论述HEVC的合并及AMVP模式的若干方面。

[0188] 运动向量按比例缩放:假定运动向量的值在呈现时间上与图片之间的距离成比例。运动向量使两个图片相关联:参考图片及含有运动向量的图片(即含有图片)。当视频编码器20或视频解码器30使用运动向量预测另一运动向量时,基于图片次序计数(POC)值计算含有图片与参考图片之间的距离。

[0189] 对于待预测的运动向量,其相关联的含有图片及参考图片不同。也就是说,对于两个截然不同的运动向量存在两个POC差值:经预测的第一运动向量,及用以预测第一运动向量的第二运动向量。此外,第一POC差为第一运动向量的当前图片与参考图片之间的差,且第二POC差为含有第二运动向量的图片与第二运动向量参考的参考图片之间的差。可基于这些两个POC距离按比例缩放第二运动向量。对于空间相邻候选者,用于两个运动向量的含有图片相同,而参考图片不同。在HEVC中,运动向量按比例缩放适用于空间及时间相邻候选者的TMVP及AMVP两者。

[0190] 人造运动向量候选者产生:如果运动向量候选者列表不完整,那么可产生人造运动向量候选者并将其插入到所述列表的末端,直到所述列表包含预定数目个候选者。

[0191] 在合并模式中,存在两个类型的人造MV候选者:仅针对B切片导出的组合式候选者,及在第一类型并未提供足够人造候选者情况下仅针对AMVP使用的零候选者。

[0192] 对于已在候选者列表中且具有必要运动信息的每一对候选者,双向组合运动向量候选者通过参考列表0中的图片的第一候选者的运动向量与参考列表1中的图片的第二候选者的运动向量的组合导出。

[0193] 下文为用于候选者插入的实例修剪过程的描述。来自不同块的候选者可恰巧相同,此情形降低合并/AMVP候选者列表的效率。可应用修剪过程以解决此问题。根据修剪过程,在某一程度上,视频译码器将当前候选者列表中的一个候选者与其它候选者相比较以避免插入相同候选者。为减小复杂度,仅应用有限数目个修剪过程,而不是将每一潜在候选

者与列表中已经存在的所有其它现有候选者相比较。

[0194] 图6说明3D-HEVC的实例预测结构。3D-HEVC是通过JCT-3V开发的HEVC的3D视频扩展。在此子章节中描述与本发明的技术相关的关键技术。

[0195] 图6展示针对三视图情况的多视图预测结构。V3指明基础视图,且可从同一时间例项的附属(基础)视图中的图片预测非基础视图(V1或V5)中的图片。

[0196] 值得提及的是,在MV-HEVC中支持视图间样本预测(从重构建的样本),图8中展示MV-HEVC的典型预测结构。

[0197] MV-HEVC及3D-HEVC两者以基础(纹理)视图可由HEVC(版本1)解码器解码的方式与HEVC相容。MV-HEVC及3D-HEVC的测试模型描述于Zhang等人的截至2015年1月26日可从[mpeg.chiariglione.org/standards/mpeg-h/high-efficiency-video-coding/test-model-6-3d-hevc-and-mv-hevc](http://mpeg.chiariglione.org/standards/mpeg-h/high-efficiency-video-coding/test-model-6-3d-hevc-and-mv-hevc)获得的“3D-HEVC及MV-HEVC的测试模型6(Test Model 6 of 3D-HEVC and MV-HEVC)”(JCT-3V文献ISO/IEC JTC1/SC29/WG11N13940)中。

[0198] 在MV-HEVC中,可由具有相同时间例项的同一视图中的图片及参考视图中的图片两者通过将这些图片的全部放置于图片的参考图片列表中而预测非基础视图中的当前图片。因此,当前图片的参考图片列表含有时间参考图片及视图间参考图片两者。

[0199] 与对应于时间参考图片的参考索引相关联的运动向量指明为时间运动向量。

[0200] 与对应于视图间参考图片的参考索引相关联的运动向量指明为视差运动向量。

[0201] 3D-HEVC支持MV-HEVC中的所有特征;因此,实现如上文所提及的视图间样本预测。

[0202] 另外,支持更高级的仅纹理译码工具及深度相关/相依译码工具。

[0203] 仅纹理译码工具常常要求可属于同一目标的对应块(视图之间)的识别。因此,视差向量导出为3D-HEVC中的基本技术。

[0204] 图7是说明3D-HEVC中的基于子PU的视图间运动预测的概念图。图7展示当前视图(V1)的当前图片160及参考视图(V0)中的共置图片162。当前图片160包含当前PU 164,所述当前PU 164包含四个子PU 166A到166D(子PU 166)。相应视差向量174A到174D(视差向量174)识别与共置图片162中的子PU 166对应的子PU 168A到168D。3D-HEVC描述用于视图间合并候选者(即,从参考视图中的参考块导出的候选者)的子PU层级视图间运动预测方法。

[0205] 当启用此模式时,当前PU 164可对应于参考视图中的参考区域(具有与由视差向量识别的当前PU相同的大小),且所述参考区域可具有比产生通常用于PU的运动信息的一个集合所需的运动信息更丰富的运动信息。因此,可使用子PU层级视图间运动预测(SPIVMP)方法,如图7中所展示。

[0206] 还可作为特殊合并候选者用信号表示此模式。子PU中的每一者含有运动信息的全集。因此,PU可含有多个运动信息集合。

[0207] 3D-HEVC中的基于子PU的运动参数继承(MPI):类似地,在3D-HEVC中,还可以类似于子PU层级视图间运动预测的方式扩展MPI候选者。举例来说,如果当前深度PU具有含有多个PU的共置区,那么当前深度PU可被分成若干子PU,且每一PU可具有不同运动信息集合。此方法被称作子PU MPI。也就是说,对应子PU 168A到168D的运动向量172A到172D可由子PU 166A到166D继承(如运动向量170A到170D一般),如图7中所展示。

[0208] 用于2D视频译码的子PU相关信息:在2014年9月25日申请、2015年3月26日公开为美国公开案第2015/0086929号的美国申请案第14/497,128号中,描述基于子PU的高级TMVP

设计。在单层译码中,提议两阶段高级时间运动向量预测设计。

[0209] 第一阶段将导出识别参考图片中的当前预测单元(PU)的对应块的向量,且第二阶段将从对应块提取多个运动信息集合且将其分配给PU的子PU。PU的每一子PU因此经单独地运动补偿。ATMVP的概念概述如下:

[0210] 1. 第一阶段中的向量可从当前PU的空间及时间相邻块导出。

[0211] 2. 可随着启动所有其它合并候选者当中的合并候选者而实现此过程。

[0212] 适用于单层译码及子PU时间运动向量预测,PU或CU可具有待预测符顶部上传送的运动细化数据。

[0213] 14/497,128申请案的若干设计方面强调如下:

[0214] 1. 向量导出的第一阶段还可由仅零向量简化。

[0215] 2. 向量导出的第一阶段可包含联合识别运动向量及其相关联图片。已提议选择相关联图片及进一步决定运动向量为第一阶段向量的各种方式。

[0216] 3. 如果运动信息在以上程序期间不可用,那么“第一阶段向量”用于取代。

[0217] 4. 从时间相邻者识别的运动向量必须以类似于TMVP中的运动向量按比例缩放的方式按比例缩放以用于当前子PU。然而,可使用以下方式中的一者设计可按比例缩放此运动向量针对的参考图片:

[0218] a. 由当前图片的固定参考索引识别图片。

[0219] b. 如果还可用于当前图片的参考图片列表中,那么所述图片经识别为对应时间相邻者的参考图片。

[0220] c. 图片经设定成第一阶段中所识别及来自运动向量被检索之处的共置图片。

[0221] 图8是说明从参考图片的子PU运动预测的概念图。在此实例中,当前图片180包含当前PU 184(例如,PU)。在此实例中,运动向量192识别参考图片182的相对于PU 184的PU 186。PU 186被分割成子PU 188A到188D,每一子PU具有相应运动向量190A到190D。因此,虽然当前PU 184实际上未分割成独立子PU,但在此实例中,可使用来自子PU 188A到188D的运动信息预测当前PU 184。具体地说,视频译码器可使用相应运动向量190A到190D译码当前PU 184的子PU。然而,视频译码器无需译码指示当前PU 184分裂成子PU的语法元素。以此方式,可使用从相应子PU 188A到188D继承的多个运动向量190A到190D有效地预测当前PU 184,而不用信号表示用于将当前PU 184分裂成多个子PU的语法元素的额外负担。

[0222] 图9是说明ATMVP(类似于TMVP)中的相关图片的概念图。具体地说,图9说明当前图片204、运动源图片206及参考图片200、202。更特定地说,当前图片204包含当前块208。时间运动向量212识别相对于当前块208的运动源图片206的对应块210。对应块210又包含运动向量214,所述运动向量214对参考图片202进行参考且充当当前块208的至少一部分(例如,当前块208的子PU)的高级时间运动向量预测符。也就是说,可添加运动向量214作为当前块208的候选者运动向量预测符。如果被选定,那么可使用对应运动向量(即,参考参考图片200的运动向量216)预测当前块208的至少一部分。

[0223] 图10是说明根据本发明的技术的实例方法的流程图。图10的方法可通过视频编码器20及/或视频解码器30执行。出于通用性,图10的方法经解释为通过“视频译码器”执行,所述视频译码器可再次对应于视频编码器20或视频解码器30中的任一者。

[0224] 最初,视频译码器从PU的当前子PU的空间或时间相邻块获得可用运动字段(230)。

视频译码器接着从所获得相邻运动字段导出运动信息 (232)。视频译码器接着确定是否已导出了PU的所有子PU的运动信息 (234)。如果并未导出 (234的“否”分支), 那么视频译码器导出剩余子PU的运动信息 (230)。另一方面, 如果已导出了所有子PU的运动信息 (234的“是”分支), 那么视频译码器例如如上文所解释确定空间时间子PU运动预测符的可用性 (236)。如果空间时间子PU运动预测符可用, 那么视频译码器将空间时间子PU运动预测符插入到合并列表中 (238)。

[0225] 尽管图10的方法中未展示, 但视频译码器接着可使用合并候选者列表译码PU (例如, PU的子PU中的每一者)。举例来说, 当通过视频编码器20执行时, 视频编码器20可使用子PU作为预测符计算PU (例如, 每一子PU) 的残余块, 变换及量化所述 (等) 残余块, 并熵编码所得经量化变换系数。类似地, 视频解码器30可熵解码所接收的数据以再生经量化变换系数, 反量化及反变换这些系数以再生所述残余块, 且接着组合所述残余块与对应子PU以解码对应于PU的块。

[0226] 图11A及11B是说明包含使用导出的运动信息预测的子块的块的实例的概念图。具体地说, 图11A说明包含子块254A到254P (子块254) 的块250 (例如, PU), 子块254A到254P在块250为PU时可表示子PU。对于块250的相邻子块256A到256I (相邻子块256) 在图11A中同样予以展示并加浅灰色阴影。

[0227] 一般来说, 视频译码器 (例如视频编码器20或视频解码器30) 可使用来自两个或更多个相邻块的运动信息导出用于块250的子块254的运动信息。相邻块可包含在空间上相邻及/或在时间上相邻的块。举例来说, 视频译码器可从在空间上相邻的子块254F及254I和从对应于子块2540的位置的在时间上相邻的块导出用于子块254J的运动信息。在时间上相邻的块可能来自与子块2540共置的先前经译码图片。为了导出用于子块254J的运动信息的运动向量, 视频译码器可平均用于子块254F、子块254I及与子块2540共置的在时间上相邻的块的运动向量。替代地, 视频译码器可使用如上文所论述的公式 (1) 到 (4) 中的一者导出运动向量。

[0228] 在一些实例中, 视频译码器可经配置以始终从在块250外部的子块 (例如, 相邻子块256及/或在时间上相邻的子块) 导出运动信息。此配置可允许子块254将被并行译码。举例来说, 视频译码器可从子块256B及256F以及与子块254F共置的时间上相邻子块的运动信息导出子块254A的运动信息。视频译码器还可使用子块256C、256B、256F及与子块254F及254G共置的时间上相邻的子块的运动信息与子块254A并行地导出子块254B的运动信息。

[0229] 图11B说明包含子块264A到264D (子块264) 的块260 (例如, PU), 子块264A到264D再次可表示子PU。图11B还说明相邻子块266A到266I (相邻子块266)。一般来说, 图11B的实例指示块 (例如块260) 的子块可具有多种大小, 且可大于用以导出运动信息的相邻块。在此实例中, 子块264大于相邻子块266。然而, 视频译码器 (例如视频编码器20及视频解码器30) 可经配置以将类似于以上关于子块254所论述的那些技术的技应用于子块264。

[0230] 图12是说明根据本发明的技术的编码视频数据的实例方法的流程图。出于解释及实例的目的, 关于视频编码器20 (图1及2) 及其组件描述图12的方法。然而, 应理解, 其它视频编码装置可经配置以执行这些或类似技术。此外, 某些步骤可被省去, 以不同次序执行, 及/或并行执行。

[0231] 最初, 视频编码器20将译码单元 (CU) 分割成一或多个预测单元 (PU) (270)。视频编

码器20接着可针对PU中的每一者测试多种预测模式(例如,空间或帧内预测、时间或帧间预测及子块运动导出预测)(272)。具体地说,模式选择单元40可测试多种预测模式,并选择用于PU的模式中的一者,所述所选择模式产生用于PU的最优选率失真特性。出于实例的目的假定,视频编码器20选择用于CU的PU的子PU运动导出模式(274)。

[0232] 根据子PU运动导出模式,视频编码器20将PU分割成子PU(276)。一般来说,子PU与PU的可区分之处在于,例如运动信息的独立信息未经译码用于子PU。替代地,根据本发明的技术,视频编码器20从相邻子PU导出用于子PU的运动信息(278)。相邻子PU可包含空间及/或时间相邻子PU。举例来说,可如关于图11A所论述而选择相邻子PU。也就是说,在此实例中,对于每一子PU,视频编码器20从包含上方相邻空间子PU、左相邻空间子PU及右下时间相邻子PU的相邻子PU导出运动信息。视频编码器20可将导出的运动信息计算为相邻子PU的运动信息的平均值,或根据上文所论述的公式(1)到(4)计算导出的运动信息。

[0233] 视频编码器20接着可使用导出的运动信息预测子PU(280)。也就是说,视频编码器20的运动补偿单元44可使用相应子PU的导出的运动信息检索用于PU的子PU中的每一者的预测的信息。视频编码器20可将PU的经预测块形成为经预测子PU中的每一者在PU的其相应位置中的组合件。

[0234] 视频编码器20接着可计算PU的残余块(282)。举例来说,求和器50可逐像素计算PU的原始版本与经预测块之间的差,从而形成残余块。接着,视频编码器20的变换处理单元52及量化单元54可分别变换及量化残余块以产生经量化变换系数(284)。熵编码单元56接着可熵编码经量化变换系数(286)。

[0235] 此外,熵编码单元56可熵编码PU的指示使用子PU运动导出模式预测PU的候选者索引(286)。具体地说,熵编码单元56可构建包含多个运动预测候选者以及表示子PU运动导出模式的候选者的候选者列表。因此,当视频编码器20选择子PU运动信息导出模式时,熵编码单元56熵编码表示索引的值,所述索引识别在PU的候选者列表中的表示子PU运动导出模式的候选者的位置。

[0236] 在以上文所描述的方式编码PU之后,视频编码器20还以大体上类似(尽管互逆)方式解码PU。尽管图12中未图示,但视频编码器20还反变换及反量化经量化变换系数以再生残余块,并组合残余块与预测块以解码PU,以在后续预测(例如,帧内及/或帧间预测)期间用作参考块。

[0237] 以此方式,图12的方法表示方法的实例,所述方法包含确定当前视频数据块的运动预测候选者指示运动信息将经导出用于当前块的子块,及响应于所述确定:将当前块分割成子块,针对所述子块中的每一者,使用至少两个相邻块的运动信息导出运动信息,并使用相应导出的运动信息编码(及解码)所述子块。

[0238] 图13是根据本发明的技术的解码视频数据的方法的实例。出于解释及实例的目的,关于视频解码器30(图1及3)及其组件描述图13的方法。然而,应理解,其它视频解码装置可经配置以执行这些或类似技术。此外,某些步骤可被省去,以不同次序执行,及/或并行执行。

[0239] 最初,视频解码器30的熵解码单元70熵解码候选者列表中的候选者的指示子PU运动导出模式用于预测单元的候选者索引(290)。尽管未图示,但应理解,最初视频解码器30构建候选者列表并添加候选者到候选者列表。在此实例中,出于解释的目的,候选者索引参



考表示子PU运动导出模式的候选者。然而,一般来说,应理解,候选者索引可参考用于PU的候选者列表中的候选者中的任一者。

[0240] 在此实例中,因为候选者索引是指表示子PU运动导出模式将用于PU的候选者,所以视频解码器30将PU分割成若干子PU (292)。视频解码器30的运动补偿单元72接着从相邻子PU导出所述子PU中的每一者的运动信息 (294)。相邻子PU可包含空间及/或时间相邻子PU。举例来说,可如关于图11A所论述而选择相邻子PU。也就是说,在此实例中,对于每一子PU,视频解码器30从包含上方相邻空间子PU、左相邻空间子PU及右下时间相邻子PU的相邻子PU导出运动信息。视频解码器30可将导出的运动信息计算为相邻子PU的运动信息的平均值,或根据上文所论述的公式 (1) 到 (4) 计算导出的运动信息。

[0241] 视频解码器30接着可使用导出的运动信息预测子PU (296)。也就是说,视频解码器30的运动补偿单元72可使用相应子PU的导出的运动信息检索用于PU的子PU中的每一者的所预测的信息。视频解码器30可将PU的经预测的块形成为经预测子PU中的每一者在PU的其相应位置中的组合件。

[0242] 视频解码器30的熵解码单元70可进一步熵解码PU的经量化变换系数 (298)。反量化单元76及反变换单元78可分别反量化及反变换经量化变换系数以产生PU的残余块 (300)。视频解码器30接着可使用经预测块及残余块解码预测单元 (302)。具体地说,求和器80可在逐像素基础上组合经预测的块与残余块以解码预测单元。

[0243] 以此方式,图13的方法表示解码视频数据的方法的实例,所述方法包含确定当前视频数据块的运动预测候选者指示运动信息将经导出用于当前块的子块,及响应于所述确定:将当前块分割成子块,针对所述子块中的每一者,使用至少两个相邻块的运动信息导出运动信息,并使用相应导出的运动信息解码所述子块。

[0244] 应认识到,取决于实例,本文中所描述的技术中的任一者的某些动作或事件可以不同序列执行、可经添加、合并或完全省略 (例如,对于实践所述技术来说并非所有所描述的动作或事件都是必要的)。此外,在某些实例中,可例如经由多线程处理、中断处理或多个处理器同时而非顺序执行动作或事件。

[0245] 在一或多个实例中,所描述的功能可以硬件、软件、固件或其任何组合实施。如果以软件实施,那么所述功能可作为一或多个指令或代码而存储于计算机可读媒体上或经由计算机可读媒体进行发射,且通过基于硬件的处理单元执行。计算机可读媒体可包含计算机可读存储媒体 (其对应于例如数据存储媒体的有形媒体) 或通信媒体,通信媒体包含 (例如) 根据通信协议促进计算机程序从一处传送到另一处的任何媒体。以此方式,计算机可读媒体大体可对应于 (1) 为非暂时性的有形计算机可读存储媒体,或 (2) 例如信号或载波的通信媒体。数据存储媒体可为可由一或多个计算机或一或多个处理器存取以检索用于实施本发明中所描述的技术的指令、代码及/或数据结构的任何可用媒体。计算机程序产品可包含计算机可读媒体。

[0246] 通过实例且非限制,这些计算机可读存储媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储器、磁盘存储器或其它磁性存储装置、快闪存储器或可用于存储呈指令或数据结构形式的所要代码且可由计算机存取的任何其它媒体。另外,任何连接被恰当地称为计算机可读媒体。举例来说,如果使用同轴电缆、光缆、双绞线、数字用户线 (DSL) 或无线技术 (例如红外线、无线电及微波) 从网站、服务器或其它远程源发射指令,那么同轴电缆、光

缆、双绞线、DSL或无线技术(例如红外线、无线电及微波)包含于媒体的定义中。然而,应理解,计算机可读存储媒体及数据存储媒体不包含连接、载波、信号或其它暂时性媒体,而是针对非暂时性有形存储媒体。如本文中所使用,磁盘及光盘包含紧密光盘(CD)、激光光盘、光学光盘、数字影音光盘(DVD)、软盘及蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘使用激光以光学方式再现数据。以上的组合也应包含于计算机可读媒体的范围内。

[0247] 指令可由一或多个处理器执行,所述一或多个处理器是例如一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其它等效的集成或离散逻辑电路。因此,如本文中所使用的术语“处理器”可指上述结构或适合于实施本文中所描述的技术的任何其它结构中的任一者。另外,在一些方面中,本文中所描述的功能性可提供于经配置用于编码及解码的专用硬件及/或软件模块内,或并入组合式编解码器中。此外,所述技术可完全实施于一或多个电路或逻辑元件中。

[0248] 本发明的技术可在包含无线手机、集成电路(IC)或一组IC(例如,芯片组)的广泛多种装置或设备中实施。本发明中描述各种组件、模块或单元以强调经配置以执行所公开技术的装置的功能方面,但未必要求由不同硬件单元来实现。确切地说,如上文所描述,可将各种单元组合于编解码器硬件单元中,或通过互操作性硬件单元(包含如上文所描述的一或多个处理器)的集合结合合适的软件及/或固件来提供所述单元。

[0249] 已描述了各种实例。这些及其它实例在以下权利要求书的范围内。

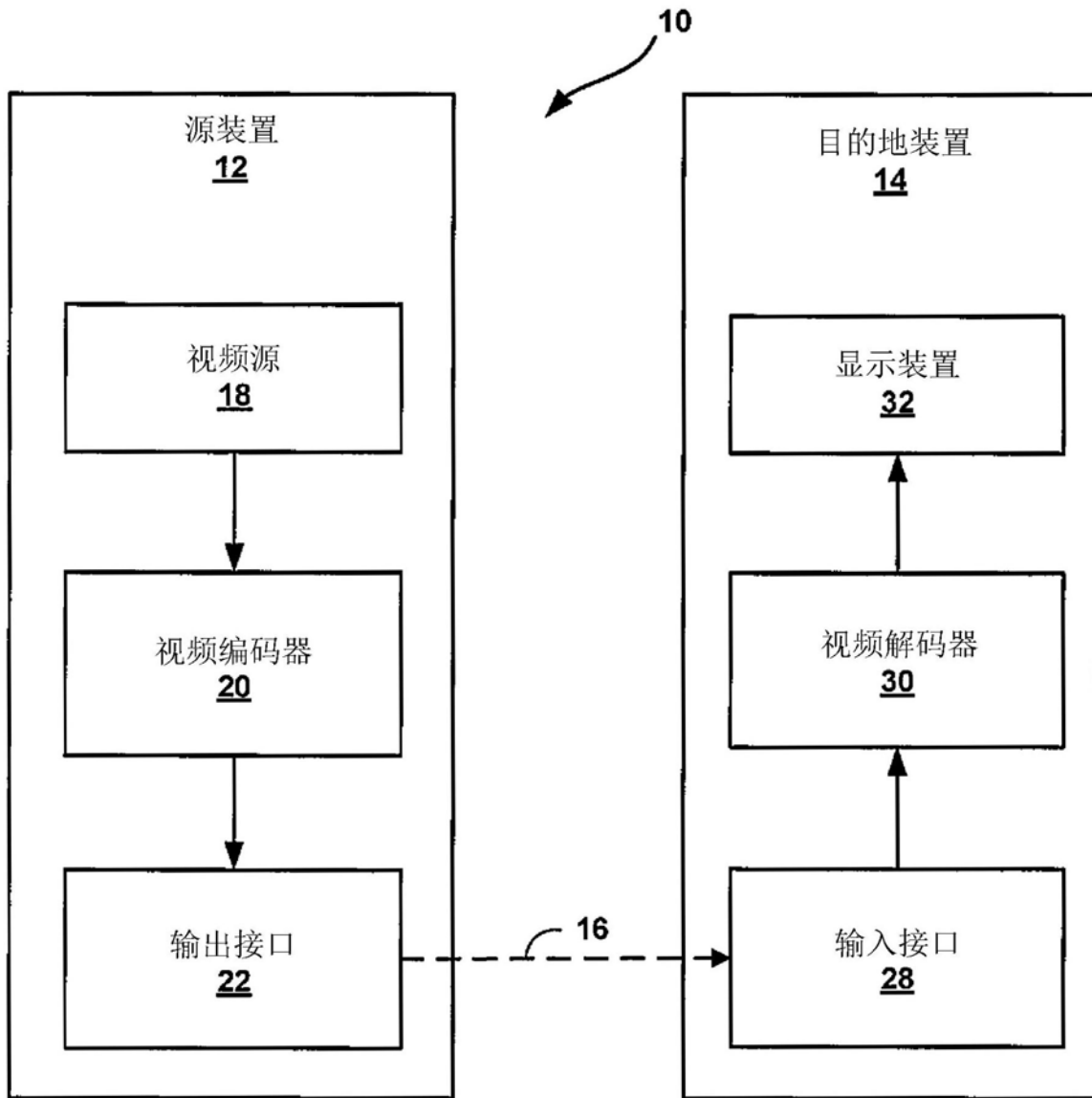


图1

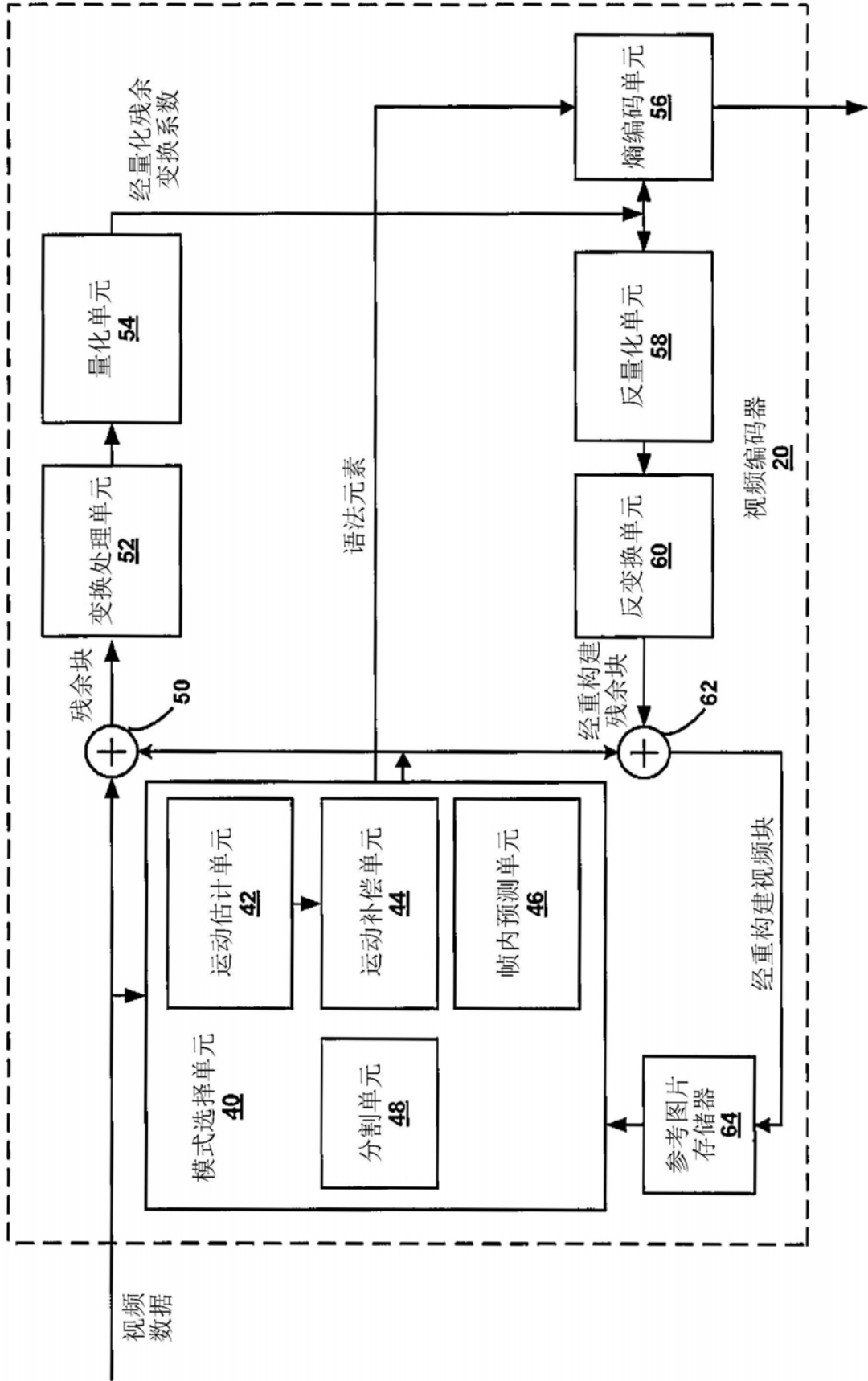


图2

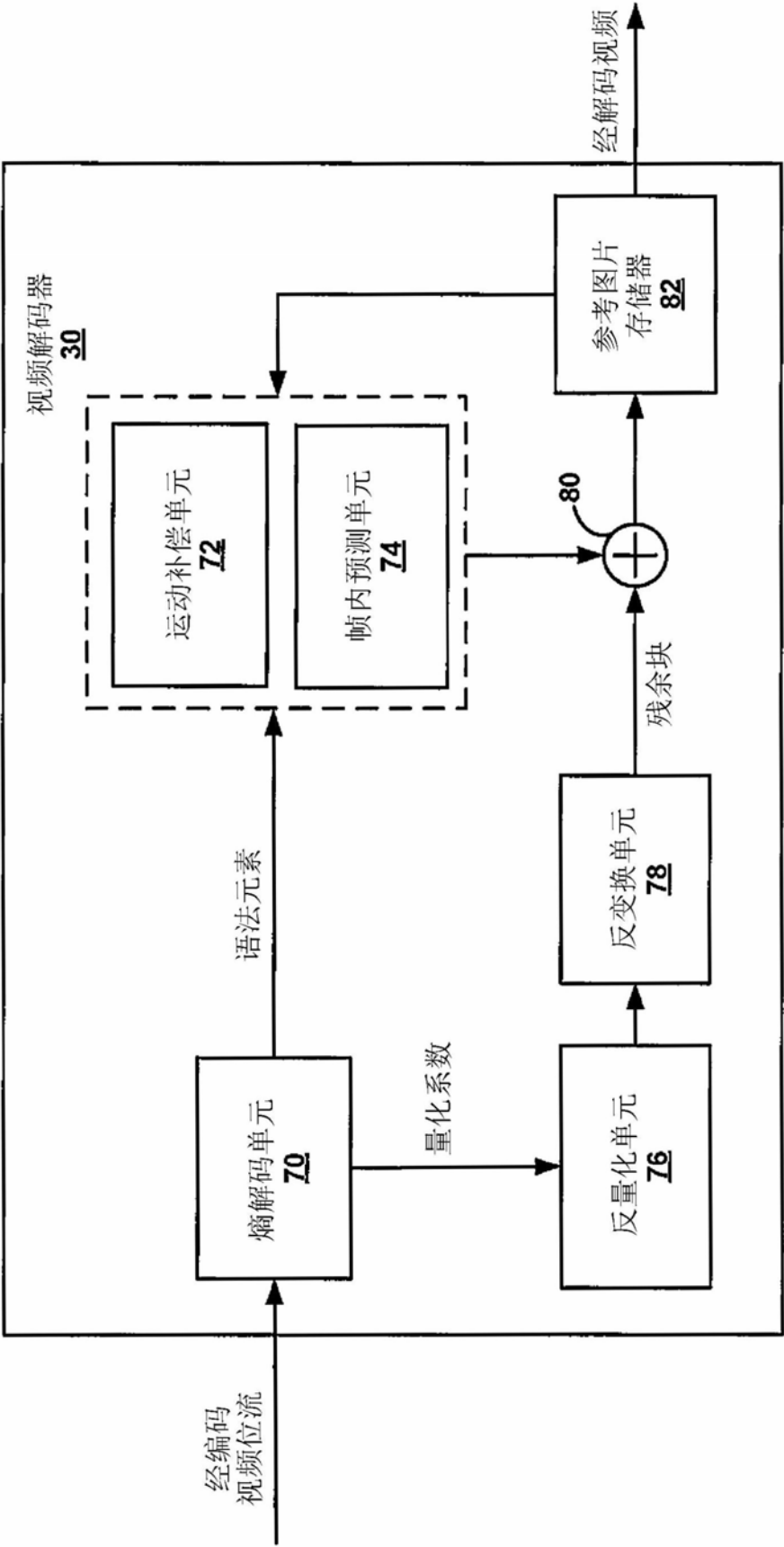


图3

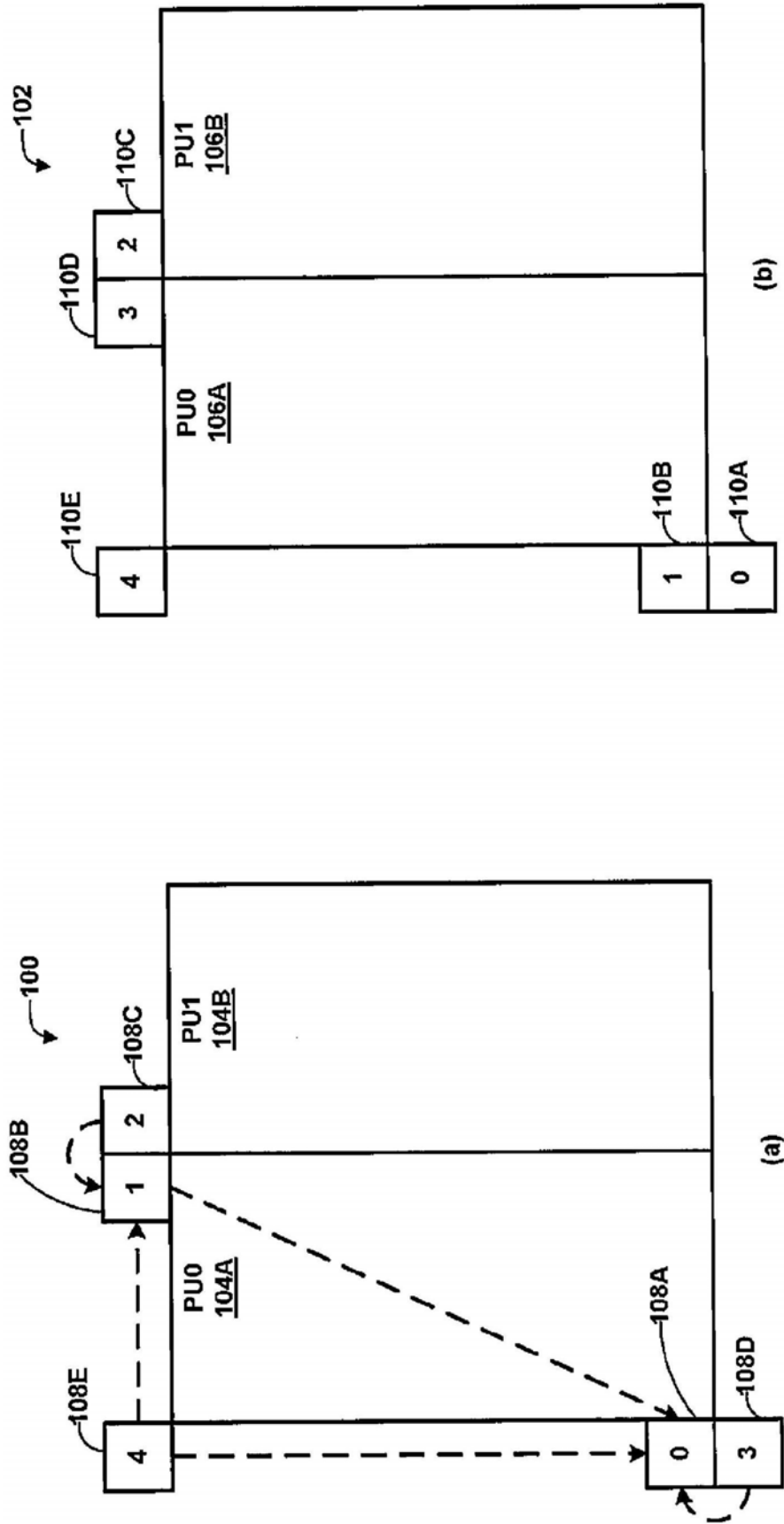


图4

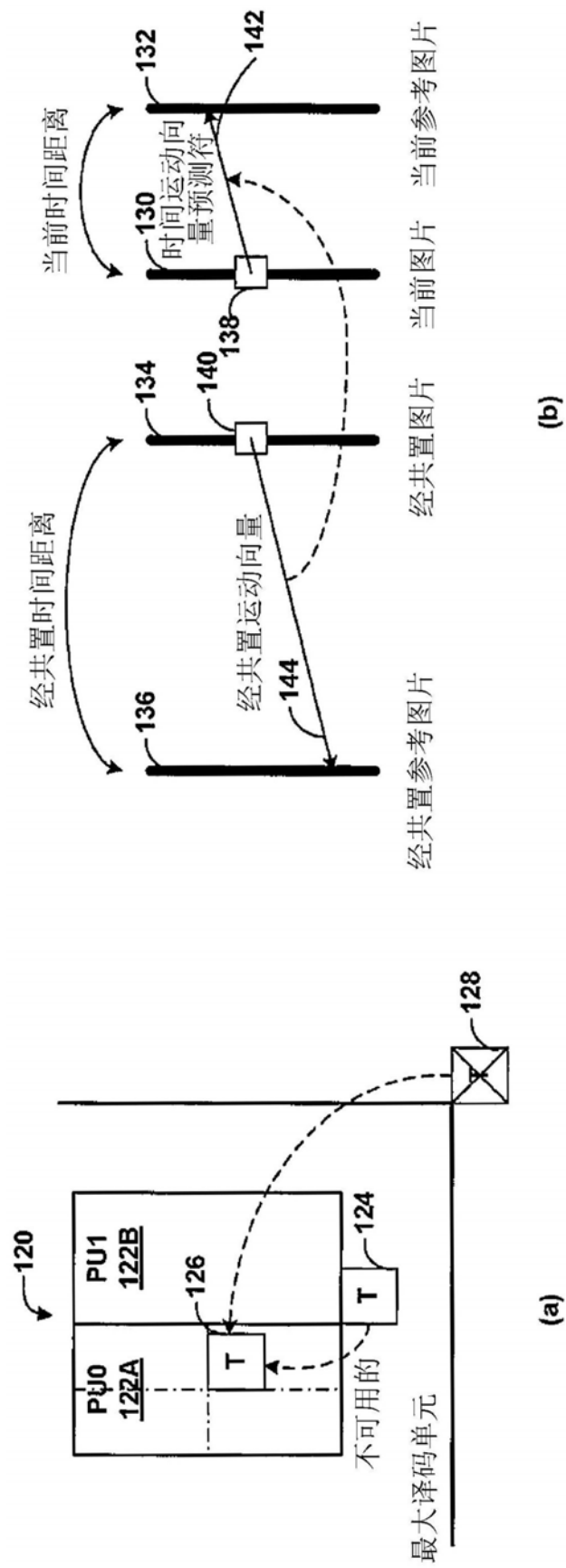


图5

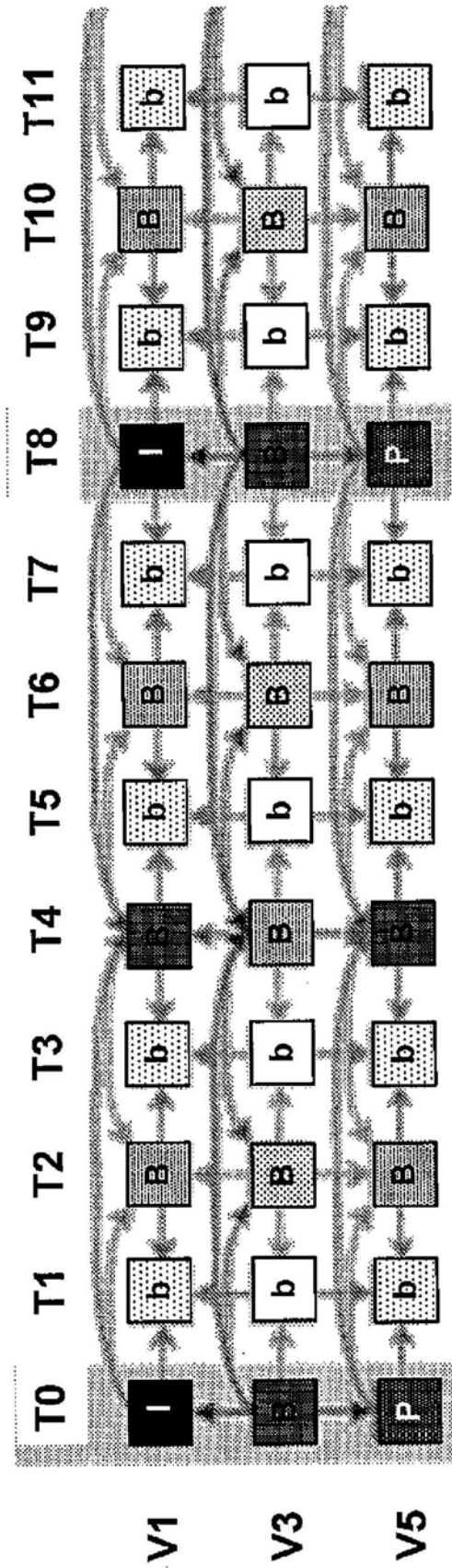


图6



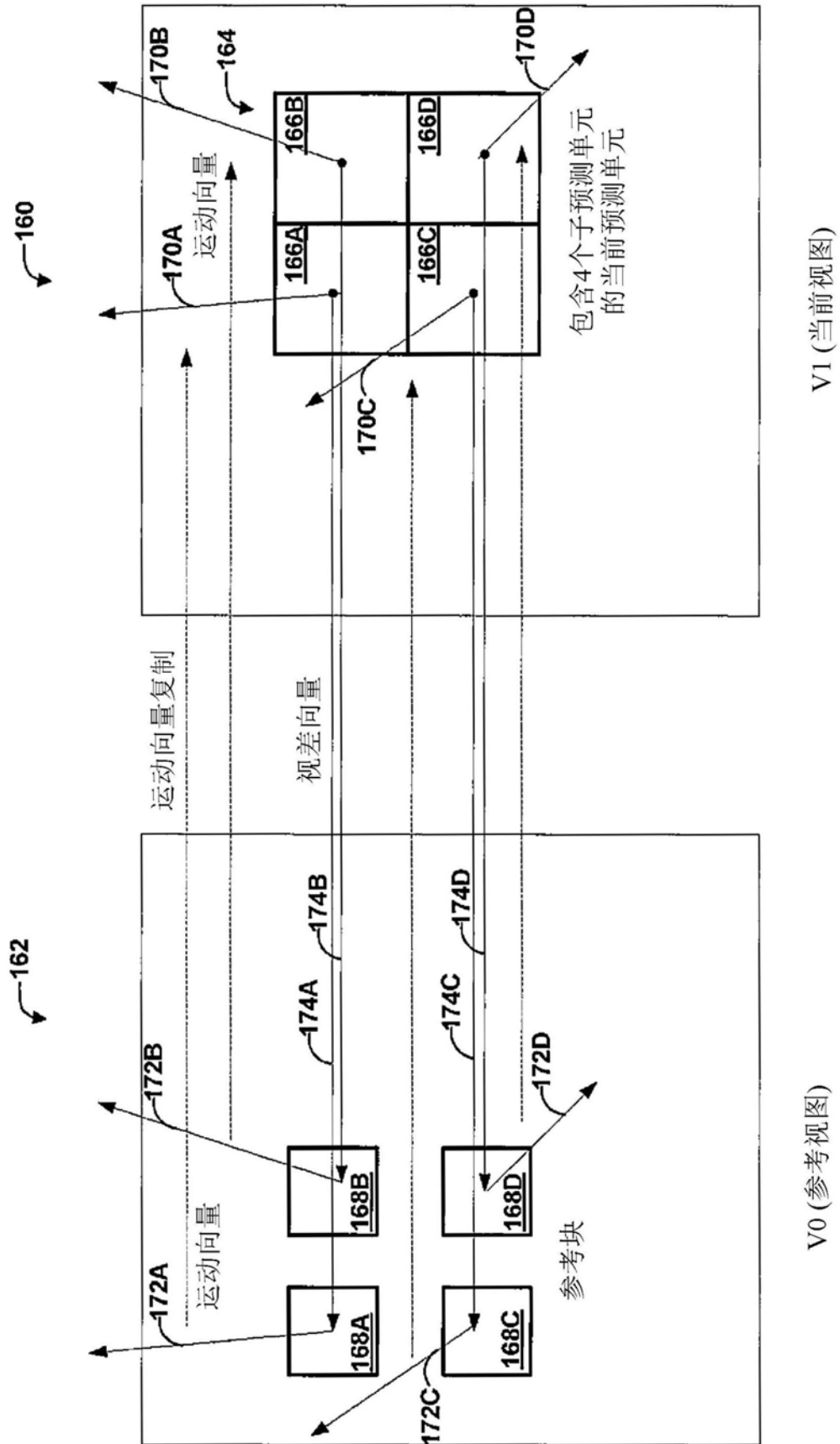


图7

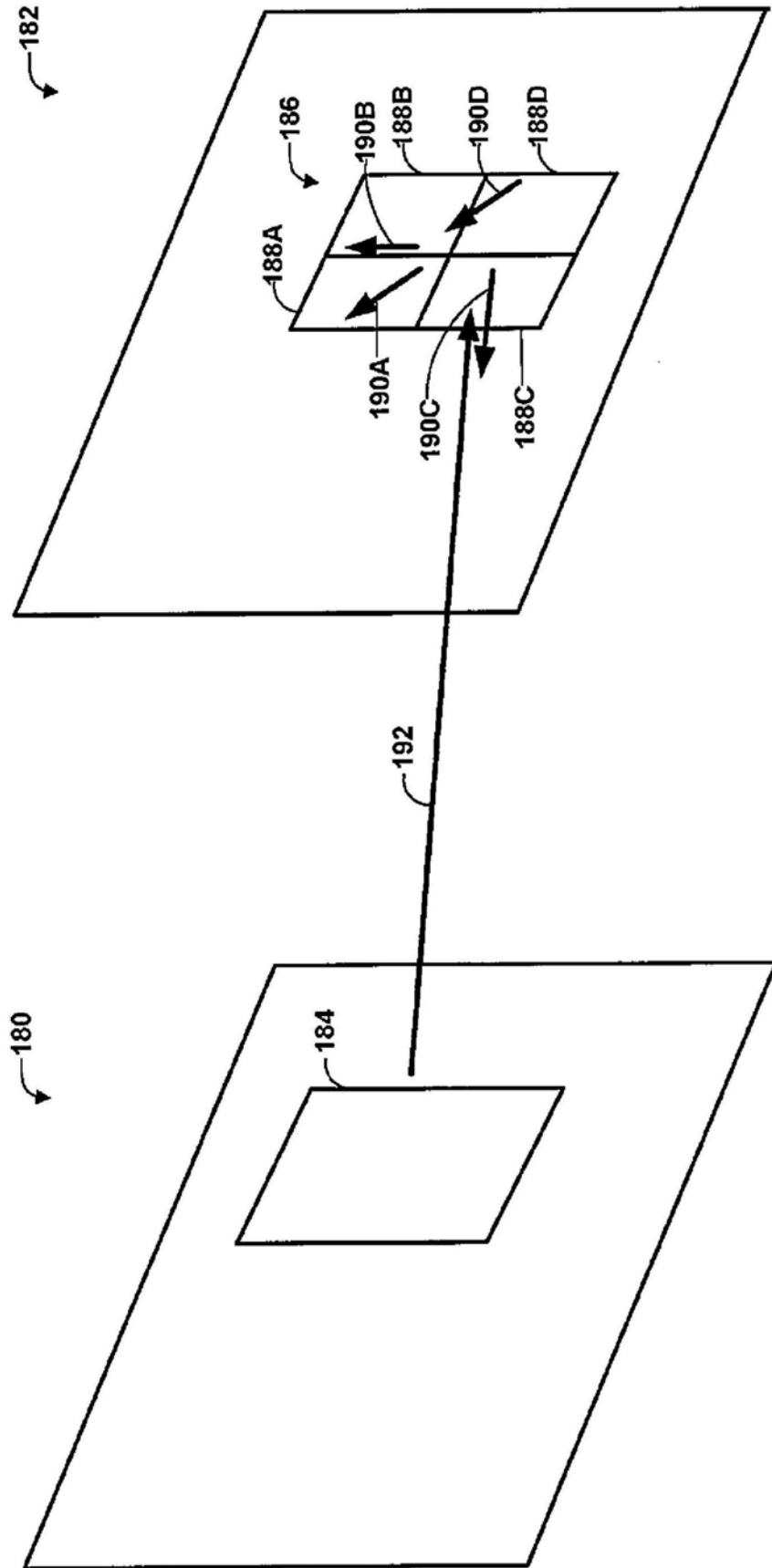


图8

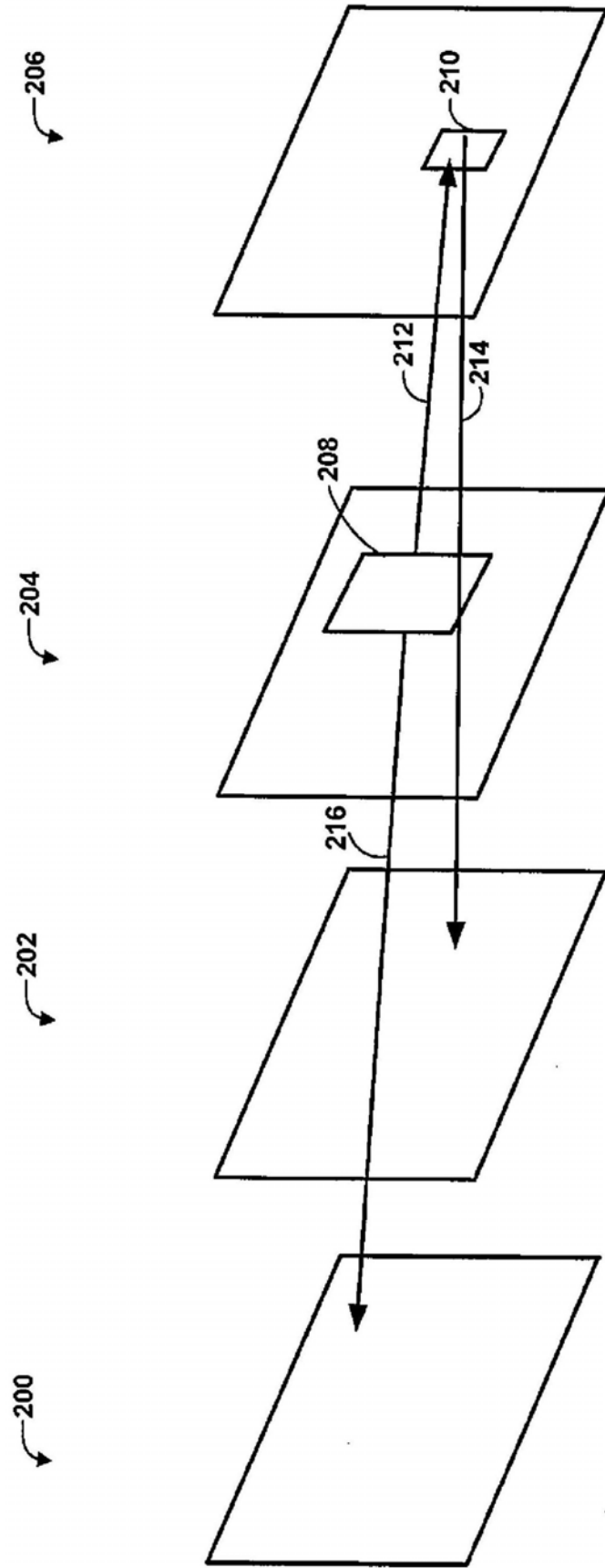


图9

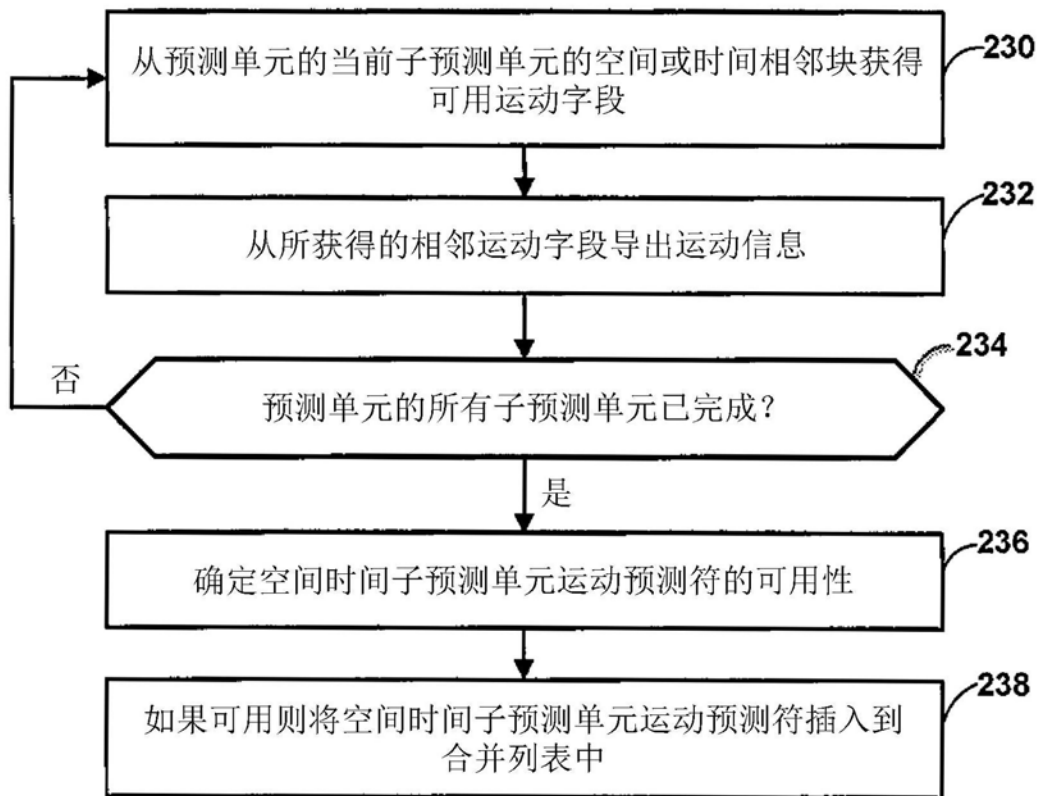


图10

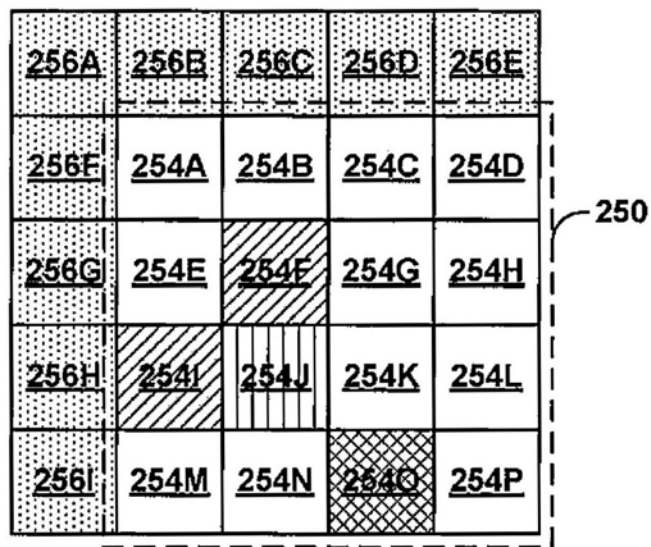


图11A

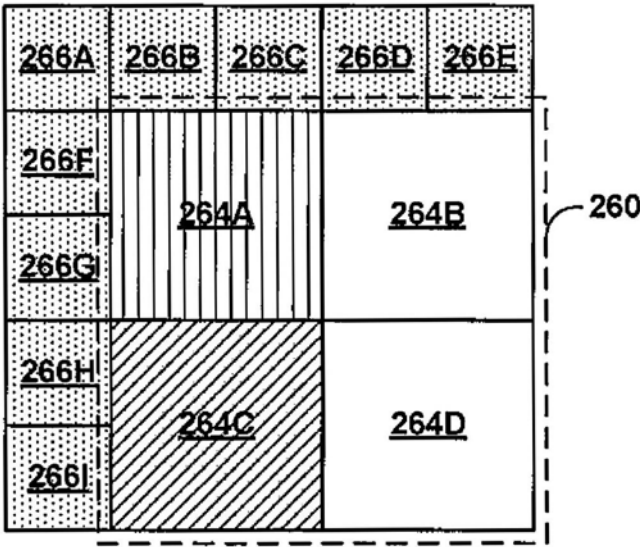


图11B

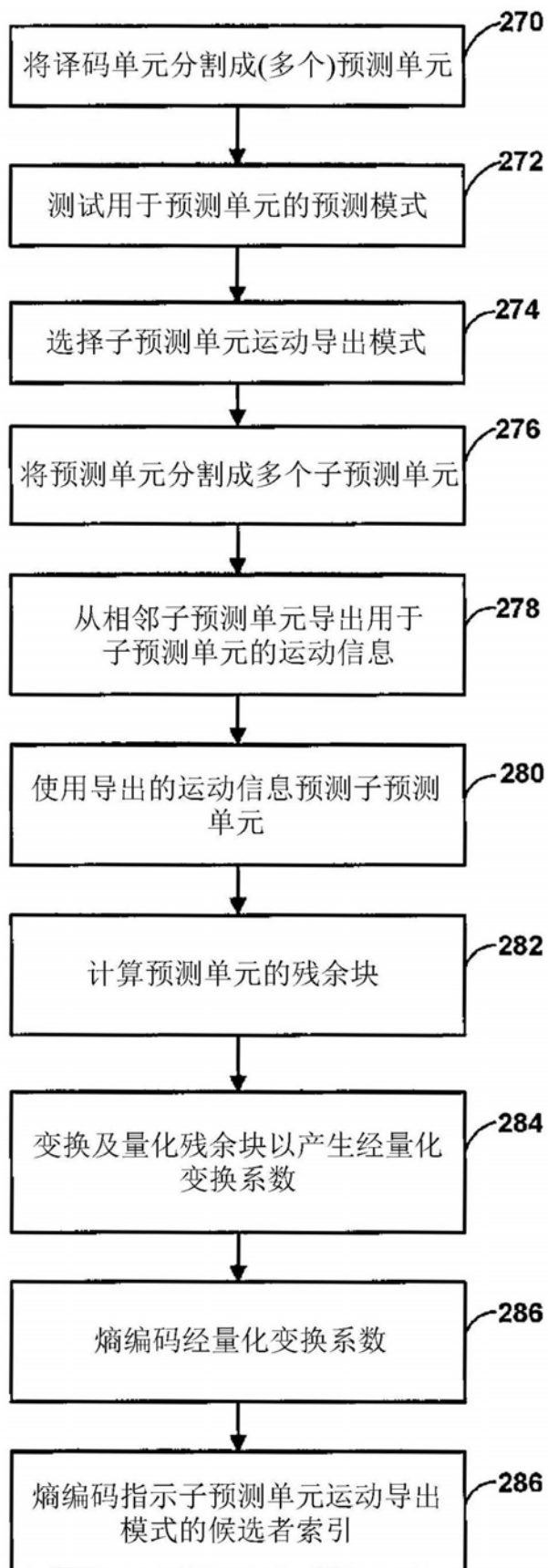


图12

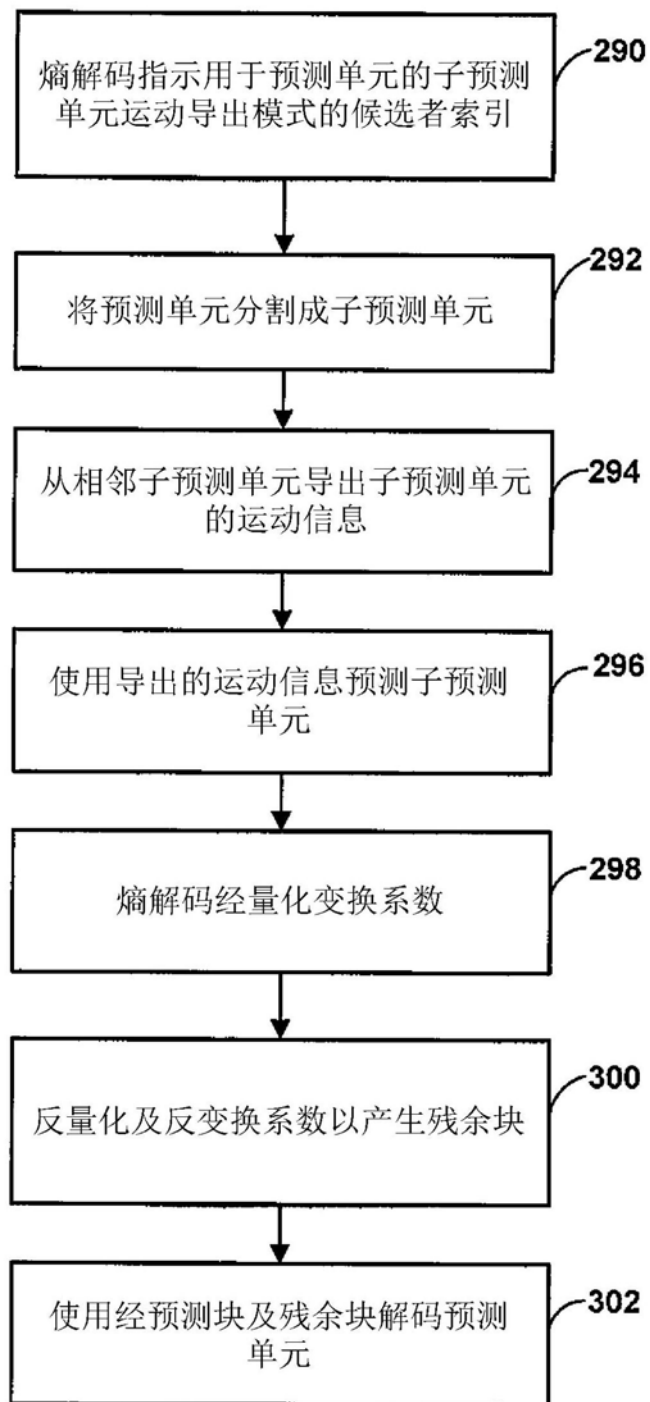


图13