



(12) 发明专利申请

(10) 申请公布号 CN 119422373 A

(43) 申请公布日 2025. 02. 11

(21) 申请号 202380048700.5

(22) 申请日 2023.06.13

(30) 优先权数据

10-2022-0075990 2022.06.22 KR

10-2023-0074954 2023.06.12 KR

(85) PCT国际申请进入国家阶段日

2024.12.20

(86) PCT国际申请的申请数据

PCT/KR2023/008133 2023.06.13

(87) PCT国际申请的公布数据

W02023/249306 KO 2023.12.28

(71) 申请人 现代自动车株式会社

地址 韩国首尔

申请人 起亚株式会社

梨花女子大学产学研协力团

(72) 发明人 姜制远 许镇 朴胜煜

(74) 专利代理机构 北京戈程知识产权代理有限公司 11314

专利代理师 程伟 甄雁翔

(51) Int.Cl.

H04N 19/137 (2014.01)

H04N 19/11 (2014.01)

H04N 19/593 (2014.01)

H04N 19/70 (2014.01)

H04N 19/176 (2014.01)

H04N 19/119 (2014.01)

H04N 19/96 (2014.01)

H04N 19/124 (2014.01)

G06T 9/00 (2006.01)

G06N 3/08 (2023.01)

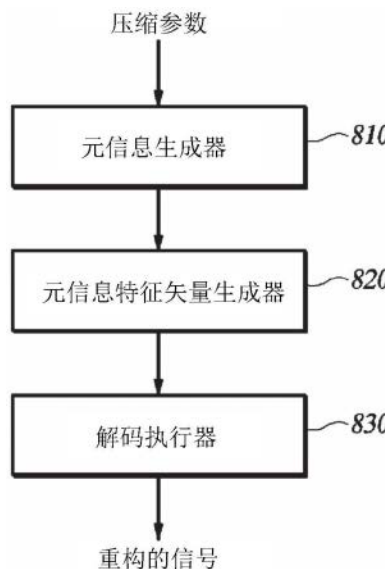
权利要求书2页 说明书21页 附图11页

(54) 发明名称

利用元信息的视频编解码的方法和装置

(57) 摘要

本实施方案公开了利用元信息的视频编解码的方法和装置。在本实施方案中,图像解码装置解码用于当前块的压缩参数,并且然后通过组合压缩参数的一个或多个来生成元信息。这里,元信息包括原来的每个压缩参数,或者包括从每个压缩参数生成的元信息矢量。图像解码装置将元信息输入到基于深度学习的提取模型中以生成元信息特征矢量,并且然后利用元信息特征矢量来对当前块执行解码处理。



1. 一种由视频解码装置执行的用于重构当前块的方法,所述方法包括:
从比特流解码用于当前块的压缩参数;
通过将压缩参数的一个或更多个进行组合来生成元信息,其中,所述元信息包括未经修改的每个压缩参数,或者包括从每个压缩参数生成的元信息矢量;
通过将元信息输入到基于深度学习的提取模型中来生成元信息的特征矢量;以及
通过利用特征矢量对当前块执行解码处理。
2. 根据权利要求1所述的方法,其中,所述压缩参数包括在当前块的更高级别处解码的参数以及当前块的级别处解码的信息,其中,所述解码的信息包括与当前块的块结构、当前块的大小、分区信息、帧内预测、帧间预测和运动矢量相关的全部或部分信息。
3. 根据权利要求1所述的方法,其中,生成元信息包括:
当压缩参数是当前块的大小时,基于当前块的宽度和高度或与宽度和高度相关的条件来生成元信息矢量。
4. 根据权利要求1所述的方法,其中,生成元信息包括:
当压缩参数是量化参数并且量化参数是属于0到 $Q-1$ 的范围的单个数字时,生成大小为 $\text{floor}(\log_2 Q)+1$ 的矢量,并且将矢量的一个映射到量化参数的元信息矢量,其中, Q 是自然数。
5. 根据权利要求1所述的方法,其中,生成元信息包括:
当压缩参数是量化参数时,生成与通过对量化参数的范围进行分区而获得的区段相对应的矢量,并且将矢量的一个映射到量化参数的元信息矢量。
6. 根据权利要求1所述的方法,其中,生成元信息包括:
当压缩参数是帧内预测模式的方向,并且方向是属于0到 $P-1$ 的范围的单个数字时,生成大小为 $\text{floor}(\log_2 P)+1$ 的矢量,并且将矢量的一个映射到方向的元信息矢量,其中, P 是自然数。
7. 根据权利要求1所述的方法,其中,生成元信息包括:
当压缩参数是帧内预测模式的方向时,生成与通过对帧内预测模式的范围进行分区而获得的区段相对应的矢量,并且将矢量的一个映射到方向的元信息矢量。
8. 根据权利要求1所述的方法,其中,生成元信息包括:
当压缩参数是运动矢量时,由元信息矢量代表运动矢量,
其中,元信息矢量的构成元素包括运动矢量的方向、运动矢量的分辨率、运动矢量的参考图片索引、运动矢量的参考图片列表、运动矢量的符号和运动矢量的大小。
9. 根据权利要求1所述的方法,其中,生成元信息包括:
在不进行修改的情况下,利用帧内预测模式、应用帧内预测还是帧间预测、帧内子分区(ISP)模式的信息、量化参数、多重变换选择(MTS)索引、用于MTS的标志或用于环内滤波器的参数集的索引作为元信息。
10. 根据权利要求1所述的方法,其中,生成元信息的特征矢量包括:
通过将元信息矢量的一个或更多个进行组合并且将输入矢量输入到提取模型以生成特征矢量来生成输入矢量。
11. 根据权利要求10所述的方法,其中,生成元信息的特征矢量包括:
除了输入矢量之外,通过将当前块的相邻参考像素输入到提取模型来生成特征矢量。

12. 根据权利要求1所述的方法, 其中, 当基于矩阵加权的帧内预测进行解码处理时, 所述元信息包括当前块的块结构、当前块的大小、当前块的参考样本、量化参数以及与帧内预测相关的参数; 并且所述特征矢量是指示集的一个的索引, 其中, 所述集包括用于应用矩阵加权的帧内预测的矩阵和矢量。

13. 根据权利要求1所述的方法, 其中, 当基于矩阵加权的帧内预测进行解码处理时, 所述元信息包括当前块的参考样本, 并且所述特征矢量是指示帧内预测模式的一个的替换索引。

14. 根据权利要求13所述的方法, 其中, 对当前块执行解码处理包括:

当基于当前块的帧内预测进行解码处理, 并且根据矩阵加权的帧内预测来预测位于当前块的左侧或上侧的亮度块时, 应用相应亮度块的替换索引以构建当前块的最可能模式(MPM) 列表。

15. 一种由视频编码装置执行的用于编码当前块的方法, 所述方法包括:

获得用于当前块的压缩参数;

通过将压缩参数的一个或更多个进行组合来生成元信息, 其中, 所述元信息包括未经修改的每个压缩参数, 或者包括从每个压缩参数生成的元信息矢量;

通过将元信息输入到基于深度学习的提取模型中来生成元信息的特征矢量; 以及

通过利用特征矢量对当前块执行编码处理。

16. 根据权利要求15所述的方法, 其中, 所述压缩参数包括在当前块的更高级别处配置的参数以及在当前块的级别处确定的信息, 其中, 所述确定的信息包括与当前块的块结构、当前块的大小、分区信息、帧内预测、帧间预测和运动矢量相关的全部或部分信息。

17. 一种计算机可读记录介质, 其存储由视频编码方法生成的比特流, 所述视频编码方法包括:

获得用于当前块的压缩参数;

通过将压缩参数的一个或更多个进行组合来生成元信息, 其中, 所述元信息包括未经修改的每个压缩参数, 或者包括从每个压缩参数生成的元信息矢量;

通过将元信息输入到基于深度学习的提取模型中来生成元信息的特征矢量; 以及

通过利用特征矢量对当前块执行编码处理。

利用元信息的视频编解码的方法和装置

技术领域

[0001] 本发明涉及利用元信息的视频编解码方法和装置。

背景技术

[0002] 本部分中的陈述仅提供与本发明有关的背景技术信息,并不一定构成现有技术。

[0003] 由于视频数据与音频数据或静止影像数据相比具有大量的数据,所以视频数据需要大量的硬件资源(其包括存储器)来存储或发送未经压缩处理的视频数据。

[0004] 相应地,编码器通常用于压缩并存储或发送视频数据。解码器接收压缩的视频数据,解压缩接收到的压缩的视频数据,并且播放解压缩的视频数据。视频压缩技术包括 H.264/高级视频编解码(AVC)、高效率视频编解码(High Efficiency Video Coding, HEVC)和多功能视频编解码(VVC),所述多功能视频编解码(VVC)比HEVC的编解码效率提高了大约30%或更多。

[0005] 然而,由于影像大小、分辨率和帧速率逐渐增加,要编码的数据量也会增加。相应地,需要提供比现有压缩技术更高的编解码效率和改善的图像增强效果的新的压缩技术。

[0006] 元信息被定义为在视频编码和解码过程中利用的一个或更多个参数的组合,例如块结构、帧内预测、帧间预测、量化参数和运动矢量。从元信息推导的特征矢量然后可以用于视频编码/解码。因此,为了提高视频编码效率和视频质量,需要考虑从元信息有效地推导特征矢量的方法以及有效地利用特征矢量的方法。

发明内容

[0007] 技术问题

[0008] 本发明寻求提供通过利用基于深度学习的提取模型从元信息提取特征矢量的视频编解码方法和装置。视频编解码方法和装置利用提取的特征矢量用于视频编码/解码。

[0009] 技术方案

[0010] 本发明的至少一个方面提供了由视频解码装置执行的用于重构当前块的方法。方法包括从比特流解码用于当前块的压缩参数。该方法还包括通过将压缩参数的一个或更多个进行组合来生成元信息。这里,元信息包括未经修改的每个压缩参数,或者包括从每个压缩参数生成的元信息矢量。该方法还包括通过将元信息输入到基于深度学习的提取模型中来生成元信息的特征矢量。该方法还包括通过利用特征矢量对当前块执行解码处理。

[0011] 本发明的另一个方面提供了由视频编码装置执行的用于编码当前块的方法。该方法包括:获得用于当前块的压缩参数。该方法还包括通过将压缩参数的一个或更多个进行组合来生成元信息。这里,元信息包括未经修改的每个压缩参数,或者包括从每个压缩参数生成的元信息矢量。该方法还包括通过将元信息输入到基于深度学习的提取模型中来生成元信息的特征矢量。该方法还包括通过利用特征矢量对当前块执行编码处理。

[0012] 本发明的又一方面提供了计算机可读记录介质,其存储由视频编码方法生成的比特流。视频编码方法包括获得用于当前块的压缩参数。视频编码方法还包括通过将压缩参

数的一个或更多个进行组合来生成元信息。这里,元信息包括未经修改的每个压缩参数,或者包括从每个压缩参数生成的元信息矢量。视频编码方法还包括通过将元信息输入到基于深度学习的提取模型中来生成元信息的特征矢量。视频编码方法还包括通过利用特征矢量对当前块执行编码处理。

[0013] 有益效果

[0014] 如上所述,本发明提供了这样一种视频编解码方法和装置,其通过利用基于深度学习的提取模型从元信息提取特征矢量,并且将提取的特征矢量用于视频编码/解码。因此,视频编解码方法和装置提高了视频编解码效率并且增强了视频质量。

附图说明

[0015] 图1是可以实现本发明技术的视频编码装置的框图。

[0016] 图2示出利用四叉树加二叉树三叉树(QTBT TT)结构来对块进行分区的方法。

[0017] 图3a和图3b示出包括宽角度帧内预测模式的多个帧内预测模式。

[0018] 图4示出当前块的相邻块。

[0019] 图5是可以实现本发明技术的视频解码装置的框图。

[0020] 图6示出矩阵加权的帧内预测。

[0021] 图7示出作为基于深度学习的神经网络的前馈网络。

[0022] 图8示出根据本发明的一个实施方案的利用元信息的视频解码装置。

[0023] 图9示出根据本发明的一个实施方案的元信息的利用。

[0024] 图10示出根据本发明的另一个实施方案的元信息的利用。

[0025] 图11示出根据本发明的一个实施方案的由视频编码装置执行的编码当前块的方法。

[0026] 图12示出根据本发明的一个实施方案的由视频解码装置执行的解码当前块的方法。

具体实施方式

[0027] 在下文中,将参考所附说明性附图对本发明的一些实施方案进行详细描述。在下面的描述中,相同的附图标记表示相同的元件,尽管元件在不同的附图中示出。此外,在一些实施方案的以下描述中,当认为相关的已知组件和功能的详细描述模糊了本发明的主题时,为了清楚和简洁起见,可以省略对相关的已知组件和功能的详细描述。

[0028] 图1是可以实现本发明技术的视频编码装置的框图。在下文中,参考图1的图示,对视频编码装置以及该装置的组件进行描述。

[0029] 编码装置可以包括:图像分割器110、预测器120、减法器130、变换器140、量化器145、重排单元150、熵编码器155、逆量化器160、逆变换器165、加法器170、环路滤波单元180和存储器190。

[0030] 编码装置的每个组件可以实现为硬件或软件,或者实现为硬件和软件的组合。另外,每个组件的功能可以实现为软件,并且微处理器也可以实现为执行对应于每个组件的软件的功能。

[0031] 一个视频由包括多个图像的一个或更多个序列组成。每个图像分割为多个区域,

并且对每个区域执行编码。例如,一个图像分割为一个或多个瓦片(tile)或/和切片(slice)。这里,一个或多个瓦片可以定义为瓦片组。每个瓦片或/和切片分割为一个或多个编码树单元(coding tree unit,CTU)。另外,每个CTU通过树结构分割为一个或多个编码单元(coding unit,CU)。应用于每个编码单元(CU)的信息被编码为CU的语法,并且共同应用于包括在一个CTU中的CU的信息被编码为CTU的语法。另外,共同应用于一个切片中的所有块的信息被编码为切片头的语法,而应用于构成一个或多个图像的所有块的信息被编码为图像参数集(Picture Parameter Set,PPS)或图像头。此外,多个图像共同参考的信息被编码为序列参数集(Sequence Parameter Set,SPS)。另外,一个或多个SPS共同参考的信息被编码为视频参数集(Video Parameter Set,VPS)。此外,共同应用于一个瓦片或瓦片组的信息也可以被编码为瓦片或瓦片组头的语法。包括在SPS、PPS、切片头、瓦片或瓦片组头中的语法可以称为高级语法。

[0032] 图像分割器110确定编码树单元(CTU)的大小。关于CTU的大小(CTU尺寸)的信息被编码为SPS或PPS的语法,并且被传送至视频解码装置。

[0033] 图像分割器110将构成视频的每个图像分割为具有预定大小的多个编码树单元(CTU),然后通过利用树结构递归地分割CTU。树结构中的叶节点成为编码单元(CU),所述CU是编码的基本单元。

[0034] 树结构可以是四叉树(quadtree,QT),其中较高节点(或父节点)被分割为具有相同大小的四个较低节点(或子节点)。树结构也可以是二叉树(binarytree,BT),其中较高的节点分割为两个较低的节点。树结构还可以是三叉树(ternarytree,TT),其中较高的节点以1:2:1的比率分割为三个较低的节点。树结构还可以是QT结构、BT结构和TT结构中的两个或更多个结构混合的结构。例如,可以利用四叉树加二叉树(quadtree plus binarytree,QTBT)结构,或者可以利用四叉树加二叉树三叉树(quadtree plus binarytree ternarytree,QTBTBT)结构。这里,二叉树三叉树(BTBT)被添加到树结构以称为多类型树(multiple-type tree,MTT)。

[0035] 图2是用于描述通过利用QTBTBT结构来对块进行分割的方法的示意图。

[0036] 如图2所示,CTU可以首先分割为QT结构。四叉树分割可以是递归的,直到分割块的大小达到QT中允许的叶节点的最小块大小(MinQTSize)。由熵编码器155对指示QT结构的每个节点是否被分割为下层的四个节点的第一标志(QT_split_flag)进行编码,并将其用信号通知视频解码装置。当QT的叶节点不大于BT中允许的根节点的最大块大小(MaxBTSize)时,可以将叶节点进一步分割为BT结构或TT结构的至少一个。在BT结构和/或TT结构中可能存在多个分割方向。例如,可以存在两个方向,即,水平地分割相应节点的块的方向以及竖直地分割相应节点的块的方向。如图2所示,当MTT分割开始时,由熵编码器155对指示节点是否被分割的第二标志(mtt_split_flag)、以及在节点被分割的情况下额外指示分割方向(竖直或水平)的标志和/或指示分割类型(二叉或三叉)的标志进行编码,并将其用信号通知视频解码装置。

[0037] 替代地,在对指示每个节点是否被分割为下层的四个节点的第一标志(QT_split_flag)进行编码之前,还可以对指示节点是否被分割的CU分割标志(split_cu_flag)进行编码。当CU分割标志(split_cu_flag)的值指示出每个节点没有被分割时,相应节点的块成为分割树结构中的叶节点并成为CU,CU是编码的基本单元。当CU分割标志(split_cu_flag)的

值指示出每个节点被分割时,视频编码装置以上述方案对第一标志首先开始编码。

[0038] 当QTBT用作树结构的另一个示例时,可以存在两种类型,即,将相应节点的块水平地分割为具有相同大小的两个块的类型(即,对称水平分割)以及将相应节点的块竖直地分割为具有相同大小的两个块的类型(即,对称竖直分割)。由熵编码器155对指示BT结构的每个节点是否被分割为下层的块的分割标志(split_flag)和指示分割类型的分割类型信息进行编码,并将其传送至视频解码装置。另一方面,可以另外存在相应节点的块被分割为彼此非对称的两个块的类型。非对称形式可以包括相应节点的块被分割为大小比率为1:3的两个矩形块的形式,或者还可以包括相应节点的块在对角线方向上被分割的形式。

[0039] CU可以根据从CTU分割的QTBT或QTBT TT而具有各种大小。在下文中,与要编码或解码的CU(即,QTBT TT的叶节点)相对应的块称为“当前块”。当采用QTBT TT分割时,除了正方形形状之外,当前块的形状也可以是矩形形状。

[0040] 预测器120对当前块进行预测以生成预测块。预测器120包括帧内预测器122和帧间预测器124。

[0041] 通常,可以对图像中的当前块的每个进行预测地编码。通常,可以通过利用帧内预测技术(其利用来自包括当前块的图像的数据)或帧间预测技术(其利用来自在包括当前块的图像之前编码的图像的数据)执行当前块的预测。帧间预测包括单向预测和双向预测两者。

[0042] 帧内预测器122通过利用在包括当前块的当前图像中位于当前块邻近的像素(参考像素)来预测当前块中的像素。根据预测方向,存在多个帧内预测模式。例如,如图3a所示,多个帧内预测模式可以包括包含平面(Planar)模式和DC模式的两种非方向模式,并且可以包括65种方向模式。根据每种预测模式不同地定义要使用的相邻像素和算法等式。

[0043] 为了对具有矩形形状的当前块进行高效的方向预测,可以额外地利用图3b中虚线箭头所示的方向模式(#67至#80,帧内预测模式#-1至#-14)。方向模式可以称为“宽角度帧内预测模式(wide angle intra-prediction modes)”。在图3b中,箭头指示用于预测的相应参考样本,而非代表预测方向。预测方向与箭头指示的方向相反。在当前块具有矩形形状时,宽角度帧内预测模式是在没有额外比特传输的情况下在与特定方向模式相反的方向上执行预测的模式。在这种情况下,在宽角度帧内预测模式中,可以通过具有矩形形状的当前块的宽度与高度的比率来确定可用于当前块的一些宽角度帧内预测模式。例如,在当前块具有高度小于宽度的矩形形状时,具有小于45度的角度的宽角度帧内预测模式(帧内预测模式#67至#80)是可用的。在当前块具有宽度大于高度的矩形形状时,具有大于-135度的角度的宽角度帧内预测模式是可用的。

[0044] 帧内预测器122可以确定要用于对当前块进行编码的帧内预测。在一些示例中,帧内预测器122可以通过利用多个帧内预测模式来对当前块进行编码,并且还可以从测试模式中选择要利用的适当的帧内预测模式。例如,帧内预测器122可以通过利用对多个测试的帧内预测模式的率失真(rate-distortion)分析来计算率失真值,并且还可以在测试模式中选择具有最佳率失真特性的帧内预测模式。

[0045] 帧内预测器122在多个帧内预测模式中选择一个帧内预测模式,并且通过利用根据选择的帧内预测模式确定的相邻像素(参考像素)和算法等式来预测当前块。由熵编码器155对关于选择的帧内预测模式的信息进行编码并将其传送至视频解码装置。

[0046] 帧间预测器124通过利用运动补偿处理来生成当前块的预测块。帧间预测器124在比当前图像更早已被编码和解码的参考图像中搜索与当前块最相似的块,并且通过利用搜索到的块来生成当前块的预测块。另外,生成运动矢量(motion vector,MV),所述运动矢量与当前图像中的当前块和参考图像中的预测块之间的位移(displacement)相对应。通常,对亮度(luma)分量执行运动估计,并且基于亮度分量计算的运动矢量用于亮度分量和色度分量两者。由熵编码器155对包括参考图像的信息和关于用于预测当前块的运动矢量的信息的运动信息进行编码,并将其传送至视频解码装置。

[0047] 帧间预测器124还可以执行参考图像或参考块的插值,以增加预测的准确性。换句话说,通过向包括两个整数样本的多个连续整数样本应用滤波器系数来在两个连续整数样本之间插值子样本。在对插值的参考图像执行搜索与当前块最相似的块的处理时,对于运动矢量可以表示小数单位精度而不是整数样本单位精度。对于要编码的每个目标区域,例如,诸如切片、瓦片、CTU、CU等的单元,可以不同地设置运动矢量的精度或分辨率。当应用这种自适应运动矢量分辨率(adaptive motion vector resolution,AMVR)时,应该针对每个目标区域用信号通知关于要应用于每个目标区域的运动矢量分辨率的信息。例如,当目标区域是CU时,用信号通知关于应用于每个CU的运动矢量分辨率的信息。关于运动矢量分辨率的信息可以是代表要在以下描述的运动矢量差的精度的信息。

[0048] 另一方面,帧间预测器124可以通过利用双向预测来执行帧间预测。在双向预测的情况下,利用两个参考图像和代表与每个参考图像中的当前块最相似的块位置的两个运动矢量。帧间预测器124分别从参考图像列表0(RefPicList0)和参考图像列表1(RefPicList1)选择第一参考图像和第二参考图像。帧间预测器124还搜索与相应参考图像中的当前块最相似的块,以生成第一参考块和第二参考块。此外,通过对第一参考块和第二参考块进行平均或加权平均来生成当前块的预测块。此外,包括关于用于预测当前块的两个参考图像的信息并且包括关于两个运动矢量的信息的运动信息被传送至熵编码器155。这里,参考图像列表0可以由预重构图像中按显示顺序在当前图像之前的图像构成,并且参考图像列表1可以由预重构图像中按显示顺序在当前图像之后的图像构成。然而,尽管不特别限于此,但可以将显示顺序中在当前图像之后的预重构图像额外地包括在参考图像列表0中。相反,在当前图像之前的预重构图像也可以额外地包括在参考图像列表1中。

[0049] 为了最小化用于编码运动信息所消耗的比特量,可以利用各种方法。

[0050] 例如,在当前块的参考图像和运动矢量与相邻块的参考图像和运动矢量相同时,能够识别相邻块的信息被编码以将当前块的运动信息传送至视频解码装置。这种方法称为合并模式(merge mode)。

[0051] 在合并模式中,帧间预测器124从当前块的相邻块选择预定数量的合并候选块(在下文中,称为“合并候选”)。

[0052] 作为用于推导合并候选的相邻块,可以利用与当前图像中的当前块邻近的左侧块A0、左下侧块A1、上侧块B0、右上侧块B1和左上侧块B2的全部或一些,如图4所示。此外,除了当前块所在的当前图像之外,位于参考图像内的块(其可以与用于预测当前块的参考图像相同或不同)也可以用作合并候选。例如,参考图像内的当前块的同位块(co-located block)或与该同位块邻近的块可以额外地用作合并候选。如果通过上述方法选择的合并候选的数量小于预设数量,则将零矢量添加到合并候选。

[0053] 帧间预测器124通过利用相邻块来配置包括预定数量的合并候选的合并列表。从合并列表中包括的合并候选之中选择要用作当前块的运动信息的合并候选,并且生成用于识别选择的候选的合并索引信息。由熵编码器155对生成的合并索引信息进行编码并将其传送至视频解码装置。

[0054] 合并跳过模式是合并模式的特例。在量化之后,在用于熵编码的所有变换系数都接近零时,仅传输相邻块选择信息而不传输残差信号。通过利用合并跳过模式,对于具有轻微运动的影像、静止影像、屏幕内容影像等可以实现相对较高的编码效率。

[0055] 此后,合并模式和合并跳过模式统称为合并/跳过模式。

[0056] 用于编码运动信息的另一种方法是高级运动矢量预测(advanced motion vector prediction, AMVP)模式。

[0057] 在AMVP模式中,帧间预测器124通过利用当前块的相邻块来推导用于当前块的运动矢量的运动矢量预测候选。作为用于推导运动矢量预测候选的相邻块,可以利用与图4所示的当前图像中的当前块邻近的左侧块A0、左下侧块A1、上侧块B0、右上侧块B1和左上侧块B2的全部或一些。此外,除了当前块所在的当前图像之外,位于参考图像(其可以与用于预测当前块的参考图像相同或不同)内的块也可以用作用于推导运动矢量预测候选的相邻块。例如,可以利用参考图像内的当前块的同位块或与该同位块邻近的块。如果通过上述方法选择的运动矢量候选的数量小于预设数量,则将零矢量添加到运动矢量候选。

[0058] 帧间预测器124通过利用相邻块的运动矢量来推导运动矢量预测候选,并且通过利用运动矢量预测候选来确定当前块的运动矢量的运动矢量预测。另外,通过将当前块的运动矢量减去运动矢量预测来计算运动矢量差。

[0059] 可以通过将预定义函数(例如,中值和平均值计算等)应用于运动矢量预测候选来获取运动矢量预测。在这种情况下,视频解码装置还知道预定义功能。此外,由于用于推导运动矢量预测候选的相邻块是已经完成编码和解码的块,所以视频解码装置也可能已经知道相邻块的运动矢量。因此,视频编码装置不需要对用于识别运动矢量预测候选的信息进行编码。相应地,在这种情况下,对于关于运动矢量差的信息和关于用于预测当前块的参考图像的信息进行编码。

[0060] 另一方面,还可以通过选择运动矢量预测候选的任何一个的方案来确定运动矢量预测。在这种情况下,用于识别选择的运动矢量预测候选的信息与用于预测当前块的关于运动矢量差的信息以及关于参考图像的信息共同进行额外的编码。

[0061] 减法器130通过将当前块减去由帧内预测器122或帧间预测器124生成的预测块来生成残差块。

[0062] 变换器140将具有空域的像素值的残差块中的残差信号变换为频域的变换系数。变换器140可以通过利用残差块的整个大小作为变换单元来变换残差块中的残差信号,或者也可以将残差块分割为多个子块,并且可以通过利用子块作为变换单元来执行变换。替代地,将残差块划分成两个子块,即变换区域和非变换区域,以通过仅利用变换区域子块作为变换单元来变换残差信号。这里,变换区域子块可以是基于水平轴线(或竖直轴线)的具有大小比率为1:1的两个矩形块的一个。在这种情况下,由熵编码器155对指示出仅变换子块的标志(cu_sbt_flag)、以及方向(竖直/水平)信息(cu_sbt_horizontal_flag)和/或位置信息(cu_sbt_pos_flag)进行编码,并将其用信号通知视频解码装置。另外,变换区域子

块的大小基于水平轴线(或竖直轴线)可以具有1:3的大小比率。在这种情况下,由熵编码器155对划分相应分割的标志(cu_sbt_quad_flag)进行额外地编码,并将其用信号通知视频解码装置。

[0063] 另一方面,变换器140可以在水平方向和竖直方向上单独地执行残差块的变换。对于该变换,可以利用各种类型的变换函数或变换矩阵。例如,用于水平变换和竖直变换的成对变换函数可以定义为多变换集(multiple transform set,MTS)。变换器140可以选择MTS中具有最高变换效率的一个变换函数对,并且可以在水平方向和竖直方向的每个上变换残差块。由熵编码器155对关于MTS中的变换函数对的信息(mts_idx)进行编码,并将其用信号通知视频解码装置。

[0064] 量化器145利用量化参数对从变换器140输出的变换系数进行量化,并且将量化的变换系数输出到熵编码器155。量化器145还可以对相关的残差块立即进行量化,而无需对任何块或帧进行变换。量化器145还可以根据变换块中的变换系数的位置来应用不同的量化系数(缩放值)。可以对应用于布置为二维的量化的变换系数的量化矩阵进行编码,并将其用信号通知视频解码装置。

[0065] 重排单元150可以对量化的残差值执行系数值的重新排列。

[0066] 重排单元150可以通过利用系数扫描将2D系数阵列改变为1D系数序列。例如,重排单元150可以利用锯齿形扫描(zig-zag scan)或对角线扫描(diagonal scan)来扫描DC系数到高频区域的系数以输出1D系数序列。根据变换单元的大小和帧内预测模式,也可以利用在列方向上扫描2D系数阵列的竖直扫描和在行方向上扫描2D块类型系数的水平扫描来代替锯齿形扫描。换句话说,根据变换单元的大小和帧内预测模式,可以在锯齿形扫描、对角线扫描、竖直扫描和水平扫描中确定要利用的扫描方法。

[0067] 熵编码器155通过利用包括基于上下文的自适应二进制算术编码(Context-based Adaptive Binary Arithmetic Code,CABAC)、指数哥伦布(Exponential Golomb)等的各种编码方案来对从重排单元150输出的1D量化的变换系数的序列进行编码,以生成比特流。

[0068] 此外,熵编码器155对与块分割相关的信息(例如,CTU大小、CTU分割标志、QT分割标志、MTT分割类型和MTT分割方向等)进行编码,以使得视频解码装置能够等同于视频编码装置来分割块。此外,熵编码器155对关于指示当前块是由帧内预测编码还是由帧间预测编码的预测类型的信息进行编码。熵编码器155根据预测类型对帧内预测信息(即,关于帧内预测模式的信息)或帧间预测信息(在合并模式的情况下的合并索引,以及在AMVP模式的情况下的关于参考图像索引和运动矢量差的信息)进行编码。此外,熵编码器155对与量化相关的信息(即,关于量化参数的信息和关于量化矩阵的信息)进行编码。

[0069] 逆量化器160对从量化器145输出的量化的变换系数进行逆量化,以生成变换系数。逆变换器165将从逆量化器160输出的变换系数从频域变换到空域,以重构残差块。

[0070] 加法器170将重构的残差块和由预测器120生成的预测块相加,以重构当前块。在对下一个块进行帧内预测时,重构的当前块中的像素用作参考像素。

[0071] 环路滤波单元180对重构的像素执行滤波,以减少由于基于块的预测和变换/量化而发生的块伪影(blocking artifacts)、振铃伪影(ringing artifacts)、模糊伪影(blurring artifacts)等。作为环内滤波器的环路滤波单元180可以包括去块滤波器182、样本自适应偏移(sample adaptive offset,SAO)滤波器184和自适应环路滤波器

(adaptive loop filter, ALF) 186的全部或一些。

[0072] 去块滤波器182对重构的块之间的边界进行滤波,以去除由于块单元编码/解码而发生的块伪影(blocking artifacts),并且SA0滤波器184和ALF 186对去块滤波的视频进行额外的滤波。SA0滤波器184和ALF 186是用于补偿由于有损编码(lossy coding)而发生重构的像素与原始像素之间的差异的滤波器。SA0滤波器184以CTU为单位应用偏移,以增强主观影像质量和编码效率。另一方面,ALF 186执行块单元滤波,并且通过划分相应块的边界和变化量的程度来应用不同的滤波器以补偿失真。可以对关于要用于ALF的滤波器系数的信息进行编码,并将其用信号通知视频解码装置。

[0073] 通过去块滤波器182、SA0滤波器184和ALF 186滤波的重构的块存储在存储器190中。当一个图像中的所有块被重构时,重构的图像可以用作用于对随后要被编码的图像内的块进行帧间预测的参考图像。

[0074] 视频编码装置可以将编码的视频数据的比特流存储在非易失性存储介质中或者通过通信网络将比特流发送到视频解码装置。

[0075] 图5是可以实现本发明技术的视频解码装置的功能框图。在下文中,参考图5,描述了视频解码装置和该装置的组件。

[0076] 视频解码装置可以包括熵解码器510、重排单元515、逆量化器520、逆变换器530、预测器540、加法器550、环路滤波单元560和存储器570。

[0077] 类似于图1的视频编码装置,视频解码装置的每个组件可以实现为硬件或软件,或者实现为硬件和软件的组合。另外,每个组件的功能可以实现为软件,并且微处理器也可以实现为执行对应于每个组件的软件的功能。

[0078] 熵解码器510通过解码由视频编码装置生成的比特流来提取与块分割相关的信息,以确定要解码的当前块,并且提取重构当前块所需的预测信息和关于残差信号的信息。

[0079] 熵解码器510通过从序列参数集(SPS)或图像参数集(PPS)提取关于CTU大小的信息来确定CTU的大小,并且将图像分割为具有确定大小的CTU。此外,CTU被确定为树结构的最高层(即,根节点),并且可以提取CTU的分割信息以通过利用树结构来分割CTU。

[0080] 例如,当通过利用QTBT结构来分割CTU时,首先提取与QT的分割相关的第一标志(QT_split_flag),以将每个节点分割为下层的四个节点。另外,相对于与QT的叶节点相对应的节点来提取与MTT的分割相关的第二标志(mtt_split_flag)、分割方向(竖直/水平)和/或分割类型(二叉/三叉),以将相应的叶节点分割为MTT结构。结果,QT的叶节点下方的每个节点被递归地分割为BT或TT结构。

[0081] 作为另一个示例,当通过利用QTBT结构来分割CTU时,提取指示是否分割CU的CU分割标志(split_cu_flag)。当分割相应块时,还可以提取第一标志(QT_split_flag)。在分割过程期间,对于每个节点,在递归QT分割0次或更多次之后可能发生0次或更多次的递归MTT分割。例如,对于CTU,MTT分割可以立即发生,或者相反,也可以仅发生多次QT分割。

[0082] 作为另一个示例,当通过利用QTBT结构来分割CTU时,提取与QT的分割相关的第一标志(QT_split_flag),以将每个节点分割为下层的四个节点。另外,提取指示将与QT的叶节点相对应的节点是否进一步分割为BT的分割标志(split_flag)以及分割方向信息。

[0083] 另一方面,当熵解码器510通过利用树结构的分割来确定要解码的当前块时,熵解码器510提取关于指示当前块是被帧内预测还是被帧间预测的预测类型的信息。当预测类

型信息指示帧内预测时,熵解码器510提取用于当前块的帧内预测信息(帧内预测模式)的语法元素。当预测类型信息指示帧间预测时,熵解码器510提取代表帧间预测信息的语法元素的信息,即,运动矢量和运动矢量参考的参考图像。

[0084] 此外,熵解码器510提取量化相关的信息并且提取关于当前块的量化的变换系数的信息作为关于残差信号的信息。

[0085] 重排单元515可以以由视频编码装置执行的系数扫描次序的相反顺序,将由熵解码器510熵解码的1D量化的变换系数的序列再次改变为2D系数阵列(即,块)。

[0086] 逆量化器520对量化的变换系数进行逆量化,并且通过利用量化参数对量化的变换系数进行逆量化。逆量化器520还可以向布置为2D的量化的变换系数应用不同的量化系数(缩放值)。逆量化器520可以通过将来自视频编码装置的量化系数(缩放值)的矩阵应用于量化的变换系数的2D阵列来执行逆量化。

[0087] 逆变换器530通过将逆量化的变换系数从频域逆变换到空域来重构残差信号,以生成当前块的残差块。

[0088] 此外,当逆变换器530对变换块的部分区域(子块)进行逆变换时,逆变换器530提取仅对变换块的子块进行变换的标志(cu_sbt_flag)、子块的方向(竖直/水平)信息(cu_sbt_horizontal_flag)和/或子块的位置信息(cu_sbt_pos_flag)。逆变换器530还将相应子块的变换系数从频域逆变换到空域以重构残差信号,并且用值“0”填充未被逆变换的区域作为残差信号,以生成当前块的最终残差块。

[0089] 此外,当应用MTS时,逆变换器530通过利用从视频编码装置用信号通知的MTS信息(mts_idx)来确定在水平方向和竖直方向的每个上要应用的变换索引或变换矩阵。逆变换器530还通过利用确定的变换函数在水平方向和竖直方向上对变换块中的变换系数执行逆变换。

[0090] 预测器540可以包括帧内预测器542和帧间预测器544。在当前块的预测类型是帧内预测时,激活帧内预测器542,而在当前块的预测类型是帧间预测时,激活帧间预测器544。

[0091] 帧内预测器542根据从熵解码器510提取的帧内预测模式的语法元素来确定多个帧内预测模式中当前块的帧内预测模式。帧内预测器542还根据帧内预测模式通过利用当前块的相邻参考像素来预测当前块。

[0092] 帧间预测器544通过利用从熵解码器510提取的帧间预测模式的语法元素来确定当前块的运动矢量和运动矢量参考的参考图像。

[0093] 加法器550通过将从逆变换器530输出的残差块与从帧间预测器544或帧内预测器542输出的预测块相加来重构当前块。在对后续要被解码的块进行帧内预测时,重构的当前块内的像素用作参考像素。

[0094] 作为环内滤波器的环路滤波单元560可以包括去块滤波器562、SAO滤波器564和ALF 566。去块滤波器562对重构的块之间的边界执行去块滤波,以便去除由于块单元解码而发生的块伪影。SAO滤波器564和ALF 566在去块滤波之后对重构的块执行额外的滤波,以便补偿由于有损编码而发生的重构的像素与原始像素之间的差异。通过利用关于从比特流解码的滤波器系数的信息来确定ALF的滤波器系数。

[0095] 通过去块滤波器562、SAO滤波器564和ALF 566滤波的重构的块存储在存储器570

中。当一个图像中的所有块被重构时,重构的图像可以用作用于对随后要被编码的图像内的块进行帧间预测的参考图像。

[0096] 在一些实施方案中本发明涉及如上所述的编码和解码视频影像。更具体地,本发明提供了这样一种视频编解码方法和装置,其利用基于深度学习的提取模型从元信息提取特征矢量,并且将提取的特征矢量用于视频编码/解码。

[0097] 以下实施方案可以由视频编码装置中的各种构成元件来执行。以下实施方案也可以由视频解码装置中的各种构成元件来执行。

[0098] 视频编码装置在编码当前块时可以从优化率失真的角度生成与本实施方案相关联的信令信息。视频编码装置可以利用熵编码器155对信令信息进行编码,并且将编码的信令信息发送到视频解码装置。视频解码装置可以利用熵解码器510从比特流解码与当前块的解码相关联的信令信息。

[0099] 在下面的描述中,术语“目标块”可以与当前块或编码单元(CU)互换利用,或者可以指代编码单元的一些区域。

[0100] 此外,一个标志的值为真(true)指示标志被设置为1的情况。另外,一个标志的值为假(false)指示标志被设置为0的情况。

[0101] I. 矩阵加权的帧内预测(Matrix weighted Intra Prediction, MIP)

[0102] 图6示出矩阵加权的帧内预测。

[0103] 当对具有宽度为W并且高度为H的当前块执行帧内预测时,可以利用当前块的相邻像素和当前块的编码信息,基于预定义的矩阵运算来生成预测因子,如图6所示。这种基于规则的预测方法被称为矩阵加权的帧内预测(Matrix Weighted Intra Prediction, MIP)。

[0104] MIP利用预定义的矩阵运算来生成全部或部分帧内预测因子。如果生成部分预测因子,则MIP可以额外地执行利用部分预测因子进行上采样或放大的插值,从而生成具有与当前块相同大小的最终帧内预测样本。

[0105] 另一方面,MIP可以选择性地选择与当前块在空间上邻近的部分像素,并且可以将部分像素用作当前块的相邻像素。在另一个实施方案中,MIP可以利用从诸如子采样和缩小的运算中推导的值用于矩阵运算。

[0106] 图6示出利用从运算推导的值和具有小于当前块的大小的矩阵来生成当前块的部分预测因子的示例。在下文中,将参考图6的示例描述MIP的实施方案。

[0107] 在第一步骤,利用平均运算从当前块的边界样本生成特定数量的样本。例如,根据块大小利用预定义的规则从上边界像素 $\text{bdry}_{\text{red}}^{\text{top}}$ 和左边界像素 $\text{bdry}_{\text{red}}^{\text{left}}$ 生成减小的边界像素 $\text{bdry}_{\text{red}}^{\text{top}}$ 和 $\text{bdry}_{\text{red}}^{\text{left}}$ 。这里,预定义的规则可以是平均和下采样的组合。此外,通过组合根据预定义的规则减小的 $\text{bdry}_{\text{red}}^{\text{top}}$ 和 $\text{bdry}_{\text{red}}^{\text{left}}$ 来生成减小的边界矢量 bdry_{red} 。根据当前块的大小和模式,可以如等式1所示生成减小的边界矢量 bdry_{red} 。

[0108] [等式1]

$$[0109] \quad bdry_{red} = \begin{cases} [bdry_{red}^{top}, bdry_{red}^{left}] & \text{对于 } W = H = 4 \text{ 并且模式} < 18 \\ [bdry_{red}^{left}, bdry_{red}^{top}] & \text{对于 } W = H = 4 \text{ 并且模式} \geq 18 \\ [bdry_{red}^{top}, bdry_{red}^{left}] & \text{对于 } \max(W, H) = 8 \text{ 并且模式} < 10 \\ [bdry_{red}^{left}, bdry_{red}^{top}] & \text{对于 } \max(W, H) = 8 \text{ 并且模式} \geq 10 \\ [bdry_{red}^{top}, bdry_{red}^{left}] & \text{对于 } \max(W, H) > 8 \text{ 并且模式} < 6 \\ [bdry_{red}^{left}, bdry_{red}^{top}] & \text{对于 } \max(W, H) > 8 \text{ 并且模式} \geq 6 \end{cases}$$

[0110] 这里,模式代表后面要描述的MIP模式。根据等式1,当 $W=H=4$ 并且模式小于18时,是通过连接 $bdry_{red}^{top}$ 和 $bdry_{red}^{left}$ 来生成 $bdry_{red}$ 。另一方面,当模式为18或更高时,通过按顺序地连接 $bdry_{red}^{left}$ 和 $bdry_{red}^{top}$ 的次序来生成 $bdry_{red}$ 。另一方面,当 $W=H=4$ 时, $bdry_{red}$ 的大小是4;在所有其他情况下, $bdry_{red}$ 的大小是8。

[0111] 在第二步骤,如图6所示,对减小的边界矢量 $bdry_{red}$ 应用预定义的矩阵运算,以生成当前块的特定部分的减小的预测因子 $pred_{red}$ 。这里, $pred_{red}$ 是具有宽度 W_{red} 和高度 H_{red} 的块,其代表当前块的下采样的大小。可以根据当前块的大小来确定宽度 W_{red} 和高度 H_{red} ,如等式2所示。

[0112] [等式2]

$$[0113] \quad \begin{aligned} W_{red} &= \begin{cases} 4 & \text{对于 } \max(W, H) \leq 8 \\ \min(W, 8) & \text{对于 } \max(W, H) > 8 \end{cases} \\ H_{red} &= \begin{cases} 4 & \text{对于 } \max(W, H) \leq 8 \\ \min(H, 8) & \text{对于 } \max(W, H) > 8 \end{cases} \end{aligned}$$

[0114] 另一方面,可以利用矩阵和偏移矢量根据等式3来计算减小的预测因子 $pred_{red}$ 。

[0115] [等式3]

$$[0116] \quad pred_{red} = A_k \cdot bdry_{red} + b_k$$

[0117] 这里, A_k 是预定义的矩阵,它具有 $W_{red} \cdot H_{red}$ 的行数和等于维数 $bdry_{red}$ 的列数。因此,当 $W=H=4$ 时, A_k 具有4列;否则,它具有8列。另一方面,偏移矢量 b_k 是预定义的矢量,其维数为 $W_{red} \cdot H_{red}$ 。 A_k 和 b_k 的下标 k 是指示预定义的矩阵和矢量的一个的索引。

[0118] A_k 和 b_k 可以选自集 S_0 、 S_1 和 S_2 的一个。此时,如等式4所示,基于块大小提取指示特定集的索引 $idx(W, H)$ 。

[0119] [等式4]

$$[0120] \quad idx(W, H) = \begin{cases} 0 & \text{对于 } W = H = 4 \\ 1 & \text{对于 } \max(W, H) = 8 \\ 2 & \text{对于 } \max(W, H) > 8 \end{cases}$$

[0121] 集 S_0 包括18个矩阵 $A_0^i, i \in \{0, \dots, 17\}$,每个矩阵具有16行和4列以及18个偏移矢量 $b_0^i, i \in \{0, \dots, 17\}$,每个矩阵的大小为16,并且用于 4×4 个块。集 S_1 包括10个矩阵 $A_1^i, i \in \{0, \dots, 9\}$,每个矩阵具有16行和8列以及10个偏移矢量 $b_1^i, i \in \{0, \dots, 9\}$,每个矩阵的大小为16,并且用于 4×4 、 8×4 和 8×8 个块。此外,集 S_2 包括6个矩阵 $A_2^i, i \in \{0, \dots, 5\}$,每个

矩阵具有64行和8列以及6个偏移矢量 $b_2^i, i \in \{0, \dots, 5\}$,每个矩阵的大小为64,并且用于各种大小的块。

[0122] 对于基于块大小的每个索引 $\text{idx}(W, H)$,MIP模式的数量分别为35、19和11。另一方面,为了减少存储器消耗,一对MIP模式可以利用相同的矩阵和偏移矢量。因此,在等式3的 A_k 和 b_k 中, k 可以由等式5表达。

[0123] [等式5]

$$[0124] \quad k = \begin{cases} \text{模式} & \text{对于 } W = H = 4 \text{ 并且模式} < 18 \\ \text{模式} - 17 & \text{对于 } W = H = 4 \text{ 并且模式} \geq 18 \\ \text{模式} & \text{对于 } \max(W, H) = 8 \text{ 并且模式} < 10 \\ \text{模式} - 9 & \text{对于 } \max(W, H) = 8 \text{ 并且模式} \geq 10 \\ \text{模式} & \text{对于 } \max(W, H) > 8 \text{ 并且模式} < 6 \\ \text{模式} - 5 & \text{对于 } \max(W, H) > 8 \text{ 并且模式} \geq 6 \end{cases}$$

[0125] 在第三步骤,将线性插值应用于减小的预测因子 pred_{red} ,以生成用于当前块的剩余位置的预测样本。首先在水平方向上应用线性插值,然后在竖直方向上进行,而不管块的大小和形状如何。

[0126] 例如,为了指示MIP是否被激活,视频编码装置可以编码基于矩阵的预测标志,并且然后可以将编码的标志发送到视频解码装置。当应用MIP模式时,用信号发出最可能模式(most probable mode,MPM)标志,以指示预测模式是否是MPM模式的一个。在MIP中,应用3-MPM,使得MPM模式用截断二进制码进行编码,并且非MPM模式用固定长度码(fixed length code,FLC)进行编码。此时,如果在应用MIP模式的块周围存在未应用MIP的块(被称为“常规块”),则可以利用映射表来执行应用MIP模式的块的MPM推导。换句话说,映射表用于从应用常规模式的相邻块推导具有相似特性的MIP模式。以这种方式推导的MIP模式用于推导应用MIP模式的块的MPM。

[0127] 此外,如果在常规块周围存在应用MIP的相邻块,则可以利用映射表来执行用于常规块的MPM推导。换句话说,映射表用于从应用MIP的相邻块的MIP模式推导具有相似特性的常规模式。以这种方式推导的常规模式用于推导常规块的MPM。类似地,即使在色度DM的推导中利用的相同位置的亮度块应用MIP,也可以利用映射表推导常规模式,并且推导的常规模式可以用于色度DM的推导。

[0128] 利用映射表的常规模式与MIP模式之间的映射可以如等式6中所示来表达。

[0129] [等式6]

$$[0130] \quad \text{predmode}_{\text{ALWIP}} = \text{map_angular_to_alwip}_{\text{idx}}[\text{predmode}_{\text{Angular}}]$$

$$[0131] \quad \text{predmode}_{\text{Angular}} = \text{map_alwip_to_angular}_{\text{idx(PU)}}[\text{predmode}_{\text{ALWIP}}]$$

[0132] 这里,仿射线性加权的帧内预测(affine linear weighted intra prediction,ALWIP)是MIP的另一个代表。在等式6中,前面的等式将常规模式映射到MIP模式,而后面的等式将MIP模式映射到常规模式。

[0133] 为每个索引 $\text{idx}(W, H)$ 将常规模式映射到MIP模式的映射表的一部分基于块大小可以如表1中所示来表达。

[0134] [表1]

[0135]

IntraPredModeY[xNbX][yNbX]	MipSizeId		
	0	1	2
0	17	0	5
1	17	0	1
2, 3	17	10	3
4, 5	9	10	3
6, 7	9	10	3
8, 9	9	10	3
10, 11	9	10	0
12, 13	17	4	0
...	...	...	...
56, 57	34	13	9
58, 59	26	1	8
60, 61	26	1	8
62, 63	26	1	8
64, 65	26	1	8
66	26	1	8

[0136] 在表1中,MipSizeID是与索引idx(W,H)相同的值。[xNbX][yNbX]代表相邻块的位置,并且IntraPreModeY[xNbX][yNbX]代表相应位置处的常规模式。根据表1,常规帧内预测模式(图3a中的0到66)可以被映射到相应的MIP模式。

[0137] 为每个索引idx(W,H)将MIP模式映射到常规模式的映射表的一部分基于块大小可以如表2中所示来表达。

[0138] [表2]

[0139]

IntraPredModeY[xNbX][yNbX]	MipSizeId		
	0	1	2
0	0	0	1
1	18	1	1
2	18	0	1
3	0	1	1
...	...	...	...
9	2	0	50
10	18	1	0
11	12	0	
12	18	1	
13	18	0	
14	1	44	
...	...	...	
17	0	1	
18	0	0	
19	50		
20	0		
21	50		
22	0		
...	...		
33	50		
34	50		

[0140] 在表2中,MipSizeID是与索引idx(W,H)相同的值。[xNbX][yNbX]代表相邻块的位置,并且IntraPreModeY[xNbX][yNbX]代表相应位置处的常规模式。根据表2,对应于各个索引idx(W,H)的MIP模式可以映射到常规帧内预测模式(图3a中的0到66)。

[0141] 在另一个示例中,为了指示MIP是否被激活,视频编码装置可以编码基于矩阵的预测标志,并且然后将编码的标志发送到视频解码装置。此时,可以与利用MPM的常规帧内预测模式分开执行MIP。另外,视频编码装置可以编码MIP模式以指示预定义的矩阵的一个以及预定义的矢量的一个,并且然后将编码的MIP模式发送到视频解码装置。另一方面,如果在常规块周围存在应用MIP的块,则假设应用MIP的相邻块处于平面模式,并且可以生成常规块的MPM候选列表。

[0142] II. 多重变换选择(Multiple transform selection, MTS)

[0143] 当在HEVC中应用变换时,DCT-II被用作变换核(在下文中,与变换类型可互换使用)以变换变换块(transform block, TB)的残差信号。然而,为了应用取决于残差信号特性的多样性的更合适的变换技术,可以利用多重变换选择(MTS)。MTS确定多个变换类型中的一个或两个最佳类型,并且然后根据确定的变换类型变换块。例如,在VVC中,如表3所示,除

了DCT-II之外,还添加了另外两种变换类型DCT-VIII和DST-VII,以便可以以各种方式变换残差信号。

[0144] [表3]

变换类型	基函数 $T_i(j)$, $i, j = 0, 1, \dots, N-1$
DCT-II	$T_i(j) = \omega_0 \cdot \sqrt{\frac{2}{N}} \cdot \cos\left(\frac{\pi \cdot i \cdot (2j+1)}{2N}\right)$ 其中, $\omega_0 = \begin{cases} \sqrt{\frac{2}{N}} & i = 0 \\ 1 & i \neq 0 \end{cases}$
DCT-VIII	$T_i(j) = \sqrt{\frac{4}{2N+1}} \cdot \cos\left(\frac{\pi \cdot (2i+1) \cdot (2j+1)}{4N+2}\right)$
DST-VII	$T_i(j) = \sqrt{\frac{4}{2N+1}} \cdot \sin\left(\frac{\pi \cdot (2i+1) \cdot (j+1)}{2N+1}\right)$

[0146] 这里,基函数构成定义每个变换类型的变换矩阵。在下文中,DCT-II、DCT-VIII和DST-VII将分别与DCT2、DCT8和DST7互换使用。

[0147] 可以在CU级别确定用于MTS的标志,并且可以根据预测模式是否是帧内/帧间预测以及预测模式是否在水平/竖直方向上,以各种方式用信号发出变换类型,如表4中所示。

[0148] [表4]

MTS_CU_flag	MTS_Hor_flag	MTS_Ver_flag	帧内/帧间	
			水平	竖直
0			DCT2	DCT2
1	0	0	DST7	DST7
	0	1	DCT8	DST7
	1	0	DST7	DCT8
	1	1	DCT8	DCT8

[0150] 首先,如果MTS_CU_flag为0,则在水平和竖直方向上应用DCT2变换类型。另一方面,如果MTS_CU_flag不为0,则可以根据MTS_Hor_flag和MTS_Ver_flag的组合在水平和竖直方向上应用不同的变换类型。

[0151] 另一方面,可以根据表4推导MTS索引(mts_idx)。例如,如果MTS_CU_flag为0,则mts_idx可以被设置为0。此外,如果MTS_CU_flag为1,则可以根据MTS_Hor_flag和MTS_Ver_flag的组合以表4的次序设置MTS索引。例如,如果MTS_Hor_flag=0并且MTS_Ver_flag=0,则mts_idx可以被设置为1;如果MTS_Hor_flag=1并且MTS_Ver_flag=1,则mts_idx可以被设置为4。

[0152] III. 前馈网络

[0153] 基于深度学习的神经网络由多个神经元和连接神经元的边缘组成,如图7的示例

中所示。神经网络的输入层和输出层由单层组成,并且隐藏层可以包括一层或更多层。一个隐藏层包括一个或更多个隐藏单元(或隐藏节点)。此外,每个边缘具有不同的权重。此外,激活函数可以用于将输出值从一层传播到下一层的过程中。典型的激活函数包括S型(sigmoid)函数、正切双曲函数和修正线性单元(Rectified linear unit, Relu)函数。

[0154] 如图7的示例中所示,具有基本前向网络结构的两层神经网络模型通过将输入矢量 x 传递通过中间隐藏单元来提供输出矢量 y 。在前馈网络结构中,在位于输出层中的任意节点处产生的值可以如等式7中所示表示。

[0155] [等式7]

$$[0156] \quad y_k(x, W) = \sigma \left(\sum_{j=0}^M w_{kj}^{(2)} \cdot h \left(\sum_{i=0}^D w_{ji}^{(1)} \cdot x_i \right) \right)$$

[0157] 此时, w^1 和 w^2 是权重矩阵,其分别对应于连接输入矢量和隐藏单元的边缘以及连接隐藏单元和输出矢量的边缘。 $w_{ji}^{(1)}$ 和 $w_{kj}^{(2)}$ 分别代表 w^1 和 w^2 的元素。 y_k 代表输出矢量 y 的元素。 x_0 和 z_0 分别代表输入层和隐藏层的偏置单元。因此, D 代表输入矢量 x 的维数, M 代表隐藏单元的数量,并且 K 代表输出矢量 y 的维数。此外, h 和 σ 是分别应用于隐藏层和输出层的激活函数。另一方面,激活函数不一定应用于隐藏层和输出层的输出。如果不利用激活函数,则等式7包括与 w^1 和 w^2 之间的权重矩阵乘法相对应的运算。

[0158] 基于训练数据和标签计算边缘的权重的过程被称为训练。通常可以利用基于反向传播算法的随机梯度下降(stochastic gradient descent, SGD)算法来执行训练。利用权重(其是根据训练计算的参数)计算前馈方向上的输出的过程被称为推理或测试。

[0159] 以下实施方案是参考视频解码装置描述的,但是它们也可以如上所述由视频编码装置执行。

[0160] IV. 根据本发明的实施方案

[0161] 在下文中,将在当前块的压缩处理中利用的一个或更多个参数的组合定义为元信息,并且将对将元信息用于视频编码/解码的方法进行描述。本实施方案包括生成元信息的方法、利用元信息生成元信息特征矢量的方法,以及利用元信息特征矢量编码/解码当前块的方法。

[0162] 图8示出根据本发明的一个实施方案的利用元信息的视频解码装置。

[0163] 为了利用元信息解码当前块,视频解码装置可以包括元信息生成器810、元信息特征矢量生成器820和解码执行器830。

[0164] 如上所述,元信息生成器810生成在当前块的压缩处理中利用的一个或更多个参数的组合作为元信息。压缩参数由语法元素及其衍生物组成,包括块大小、块结构、参考样本、帧内预测模式、帧间预测模式、运动矢量和量化参数。另一方面,压缩参数可以用于当前块的编码和解码两者。

[0165] 元信息特征矢量生成器820将元信息输入到基于深度学习的提取模型中,以生成大小为 $K \times 1$ 的元信息特征矢量。这里,作为神经网络的提取模型可以采用如上所述的前馈网络。此时,可以在排除激活功能的情况下利用前馈网络。元信息特征矢量可以由现有VVC中利用的映射表、根据块编码条件的模式等的结果组成。

[0166] 解码执行器830利用元信息特征矢量来解码当前块。例如,解码执行器830可以执

行预测器540的操作。或者,解码执行器830可以执行逆变换器530的操作。换句话说,解码执行器830可以是视频解码装置内的任何组件,其可以利用元信息特征矢量来解码当前块。

[0167] 此外,也可以通过例如在视频编码装置内的解码路径上的构成元件来利用图8所示的元信息的生成和利用。例如,解码执行器830可以执行视频编码装置内的预测器120的操作。替选地,解码执行器830可以执行视频编码装置内的逆变换器165的操作。

[0168] 根据本实施方案,与基于在现有VVC中利用的映射表来确定编码模式和索引的方法相比,可以替换现有过程以允许利用更多样的压缩参数而不涉及映射表。

[0169] 例如,如图9的顶部示例所示,现有VVC技术中的MIP预测推导如等式4所示的索引 idx ,并且然后选择集 S_{idx} ,所述集 S_{idx} 包括矩阵和根据推导的索引的偏移矢量。此时,如等式4所示,可以根据块大小推导 idx 值。然而,如果基于块结构、参考样本、量化参数以及可以在帧内预测过程中利用的其他参数来设置不同的 idx ,则情况的数量可能会显著增加。此时,如图9的底部示例所示,本实施方案能够高效地处理上述各种情况。

[0170] 此外,本实施方案可以替换在MIP中利用的映射表。例如,如表1和表2所示,本实施方案可以基于应用MIP的块的大小来替换MIP模式与帧内预测模式之间的映射关系。换句话说,尽管在表1和表2中仅利用应用MIP模式的块的大小,本实施方案可以通过利用更多样的压缩参数来实现映射关系。

[0171] 此外,本实施方案可以生成替换MIP模式的帧内预测模式。本实施方案可以利用前馈网络将元信息映射到替换的索引 $repIdx$ 。这里,元信息是当前块的参考样本,并且 $repIdx$ 可以指示根据图3a的示例的67个帧内预测模式的一个。稍后,当对块执行帧内预测时,映射到替换的索引的帧内预测模式可以用于推导MPM。例如,如果利用MIP预测位于当前块的左侧或上侧的亮度块,则视频解码装置可以将相应亮度块的替换的索引应用于构建MPM列表的过程。

[0172] 在另一个示例中,本实施方案可以替换VVC的MTS选择过程,如图10的顶部示例所示。例如,表4示出与VVC中的变换核选择相关的映射关系。换句话说,检查是否选择帧间预测和帧内预测模式,并且取决于是否在水平方向和竖直方向上应用变换,可以选择DCT2、DST7和DCT8的一个作为变换核。另一方面,如图10的底部示例所示,通过应用本实施方案,可以通过利用更多样的压缩参数来建立压缩参数与变换核之间的映射关系。

[0173] 本实施方案对MIP和MTS的应用仅仅是示例;本实施方案可以广泛地应用于现有VVC中利用的映射表、根据块编码条件的模式设置等。

[0174] 另一方面,可以通过利用训练元信息和相应的标签来预先训练元信息特征矢量生成器820内的提取模型。标签可以是例如如上所述的MIP中的 idx 或MTS中的变换核。在训练过程中,可以利用基于由提取模型生成的输出与标签之间的差异的损失函数来更新提取模型的权重。由于视频编码装置将权重传递到视频解码装置,提取模型可以在视频编码装置与视频解码装置之间共享。

[0175] 在下文中,将对用于生成元信息的方法进行描述。元信息将视频解码过程中利用的视频压缩参数表示为矢量。此时,视频压缩参数包括在较高级别处解码的参数,诸如从比特流解码的VPS、SPS、自适应参数集(adaptive parameter set, APS)、视频图片级别和视频切片级别。此外,视频压缩参数可以包括与在当前块级别处解码的视频块的大小、分区信息、帧内预测、帧间预测、运动矢量等相关的全部或部分信息。

[0176] 例如,将对将压缩参数表示为矢量的示例进行描述。换句话说,视频解码装置可以通过将压缩参数表示为矢量来生成元信息。在下文中,矢量代表一个或多个比特的组合。

[0177] 首先,视频解码过程中利用的所有块结构表示为矢量。例如,当利用 $2N \times 2N$ 、 $2N \times N$ 、 $N \times 2N$ 和 $N \times N$ 的块结构时,每个块结构表示为大小为2的矢量,诸如00、01、10和11。在VVC中,基于四叉树(QT)、二叉树(BT)和三叉树(TT)结构来对块进行分区。元信息将视频编码装置和视频解码装置可用的所有块结构表示为矢量。

[0178] 此外,在视频解码过程中确定的块的大小可以表示为矢量。可以利用块的宽度和高度,也可以利用关于宽度/高度的条件。例如,块的宽度和高度的一个大于或等于16的情况可以表示为0,而所有其他情况可以由1表示。替选地,块的宽度和高度相同的情况可以由0表示,而宽度和高度不同的情况可以由1表示。

[0179] 此外,在视频解码过程中利用的量化参数(quantization parameter, QP)表示为矢量。例如,当量化参数是属于0到 $Q-1$ (其中, Q 是自然数)的范围的单个数字时,可以生成大小为 $\text{floor}(\log_2 Q) + 1$ 的矢量,并且可以将量化参数映射到矢量的一个。这里,向下取整(floor)函数将实数四舍五入为最小的自然数。在另一个示例中,代替将量化参数表示为大小为 $\text{floor}(\log_2 Q) + 1$ 的矢量, Q 的范围可以分区为相等的区段,并且分区的区段可以表示为矢量。然后,量化参数可以被映射到矢量的一个。例如,当将范围分区为8个相等的部分时, Q 的范围分区为 $[0, Q/8-1]$ 、 $[Q/8, 2Q/8-1]$ 、 $\dots$ 、 $[7Q/8, Q-1]$,并且每个区段由大小为3的矢量表示。例如,如果 Q 的范围是64,则 $QP=27$ 由矢量011表示。

[0180] 此外,在视频解码过程中利用的帧内预测模式的方向表示为矢量。例如,当方向是属于0到 $P-1$ (其中, P 是自然数)的范围的单个数字时,可以生成大小为 $\text{floor}(\log_2 P) + 1$ 的矢量,并且可以将帧内预测模式映射到矢量的一个。在另一个示例中,代替将帧内预测模式的方向表示为大小为 $\text{floor}(\log_2 P) + 1$ 的矢量, P 的范围可以被分区为相等的区段,并且分区的区段可以表示为矢量。然后,帧内预测模式的方向可以被映射到矢量的一个。例如,当将范围分区为8个相等的部分时, P 的范围分区为 $[0, P/8-1]$ 、 $[P/8, 2P/8-1]$ 、 $\dots$ 、 $[7P/8, P-1]$,并且每个区段由大小为3的矢量表示。例如,如果 P 的范围是64,则帧内预测模式12由矢量001表示。

[0181] 此外,在视频解码过程中利用的运动矢量表示为矢量。例如,当运动矢量由具有6个元素的矢量表示时, $MV = (m_0, m_1, m_2, m_3, m_4, m_5, m_6)$ 。这里, m_0 指示运动矢量是在 x 方向上还是在 y 方向上。例如,当 m_0 为0时,其指示出运动矢量在 x 方向上(即,更靠近 x 轴的运动矢量,其中 x 的绝对值大于或等于 y 的绝对值);当 m_0 为1时,其指示出运动矢量在 y 方向上(即,更靠近 y 轴的运动矢量,其中 y 的绝对值大于 x 的绝对值)。 m_1 用于确定运动矢量的分辨率。例如,当 m_1 为0时,运动矢量具有整数精度的分辨率;当 m_1 为1时,运动矢量具有分数精度的分辨率。

[0182] 此外, m_2 指示运动矢量的参考图片索引。 m_3 指示运动矢量的参考图片列表。例如,当 m_3 为0时,其指示出参考图片列表为 L_0 列表;当 m_3 为1时,其指示出参考图片列表为 L_1 列表。 m_4 指示运动矢量的符号。换句话说,可以基于运动矢量的符号将 m_4 设置为0或1。 m_5 指示运动矢量的大小。

[0183] 在用于运动矢量的元信息的矢量元素中, m_0 、 m_1 、 m_3 和 m_4 可以由1比特表示,并且 m_2 和 m_5 可以由多个比特表示。

[0184] 在另一个示例中,在视频解码过程期间解析的压缩参数可以在不进行修改的情况下用作元信息。换句话说,视频解码装置可以在不进行修改的情况下,利用在解码过程中利用的编码模式、标志、索引信息和推导的参数作为元信息。

[0185] 在帧内预测的情况下,将解码的方向预测模式 p 用作元信息。

[0186] 对当前块应用帧内预测还是帧间预测由0或1指示。

[0187] 当应用帧内预测并且利用帧内子分区(Intra sub-partition,ISP)模式时,将ISP模式信息用作元信息。

[0188] 例如,如果量化参数是24,则将24用作元信息。

[0189] 如果在环内滤波器的滤波器参数集 A 内的特定索引被用于解码,则将相应的索引用作元信息。

[0190] 如果MTS模式用于变换,则将MTS索引或MTS的标志用作元信息。

[0191] 在下文中,将对从元信息生成元信息特征矢量的示例进行描述。

[0192] 例如,视频解码装置通过级联一条或更多条元信息来生成元信息输入矢量 X 。例如,在上述示例中,如果当前块大小为 32×32 ,则将结构设置为 $2N \times 2N$,并且选择矢量00。此外,由于块的宽度和高度的一个大于或等于16,第三文件被选择为0,并且因此,最终输入元信息矢量可以被表示为000。例如,在图7的示例中,可以将0、0、0输入到 x_1 、 x_2 和 x_3 。

[0193] 视频解码装置可以将元信息矢量输入到神经网络中以生成元信息特征矢量 y 。 y 可以用于在编码和解码过程期间选择模式,如在图9和图10示例中,或者表达映射表内的查询。例如,在图9的示例中,如果 y 为0,则从MIP的集 S_0 、 S_1 和 S_2 选择 S_0 。此外,在图10的示例中,如果 y 为00,则可以选择水平和垂直方向上的核作为DCT2。

[0194] 在另一个示例中,如上所述,视频解码装置可以在不进行修改的情况下,利用解析的压缩参数作为元信息。例如,如果量化参数是27,则当前块的长度是16,并且帧内预测模式是12、27、16,并且可以将12输入到在图7的示例中的 x_1 、 x_2 和 x_3 。

[0195] 在另一个示例中,除了输入元信息之外,视频解码装置可以将当前块的相邻参考像素输入到神经网络中以生成元信息特征矢量。例如,除了将000输入到在图7的示例中的 x_1 、 x_2 和 x_3 ,可以将位于当前块上侧的 W 参考像素输入到剩余输入节点中的 W 输入节点。此外,可以将位于当前块的左侧的 H 参考像素输入到 H 输入节点。

[0196] 在下文中,将利用图11和图12的图示对基于元信息的当前块的编码/解码的方法进行描述。

[0197] 图11示出根据本发明的一个实施方案的由视频编码装置执行的编码当前块的方法。

[0198] 视频编码装置获得用于当前块的压缩参数 $S1100$ 。

[0199] 这里,压缩参数包括考虑到当前块的较高级别处的率失真优化而配置的那些参数。此外,压缩参数可以包括考虑到当前块的级别处的率失真优化而确定的信息。该信息可以包括与当前块的块结构、当前块的大小、分区信息、帧内预测、帧间预测和运动矢量相关的全部或部分信息。

[0200] 视频编码装置通过将压缩参数的一个或更多个进行组合来生成元信息 $S1102$ 。这里,元信息可以包括未经修改的每个压缩参数,或者可以包括从每个压缩参数生成的矢量。

[0201] 视频编码装置将元信息输入到基于深度学习的提取模型中以生成元信息特征矢

量S1104。

[0202] 视频编码装置可以通过级联每个压缩参数的一个或更多个矢量来生成输入矢量,并且可以将生成的输入矢量输入到提取模型中。这里,提取模型可以是如上所述的前馈网络。元信息特征矢量可以由现有VVC中利用的映射表、根据块编码条件的模式等的结果组成。

[0203] 另一方面,可以通过利用训练元信息和相应的标签来预先训练提取模型。标签可以是例如MIP中的idx。由于视频编码装置将权重传递到视频解码装置,提取模型可以在视频编码装置与视频解码装置之间共享。

[0204] 视频编码装置通过利用元信息特征矢量对当前块执行编码处理S1106。

[0205] 例如,当基于MIP进行编码处理时,元信息包括当前块的块结构、当前块的大小、当前块的参考样本、量化参数以及与帧内预测相关的参数;并且元信息特征矢量可以是指示集的一个的idx。此时,集包括用于MIP的应用的矩阵和矢量。

[0206] 此外,可以生成替换MIP模式的帧内预测模式。这里,元信息可以是当前块的参考样本,并且元信息特征矢量可以由指示帧内预测模式的替换索引repIdx组成。稍后,当对块执行帧内预测时,映射到替换的索引的帧内预测模式可以用于推导MPM。例如,如果利用MIP预测位于当前块的左侧或上侧的亮度块,则视频编码装置可以将相应亮度块的替换的索引应用于构建MPM列表的过程。

[0207] 图12示出根据本发明的一个实施方案的由视频解码装置执行的解码当前块的方法。

[0208] 视频解码装置从比特流解码用于当前块的压缩参数S1200。

[0209] 这里,压缩参数包括在当前块的较高级别处解码的那些参数。此外,压缩参数可以包括在当前块的级别处解码的信息。此时,解码的信息可以包括与当前块的块结构、当前块的大小、当前块的参考样本、分区信息、帧内预测、帧间预测和运动矢量相关的全部或部分信息。

[0210] 视频解码装置通过将压缩参数的一个或更多个进行组合来生成元信息S1202。这里,元信息可以包括未经修改的每个压缩参数,或者可以包括从每个压缩参数生成的矢量。

[0211] 视频解码装置将元信息输入到基于深度学习的提取模型中以生成元信息特征矢量S1204。

[0212] 视频解码装置可以通过级联每个压缩参数的一个或更多个矢量来生成输入矢量,并且可以将生成的输入矢量输入到提取模型中。这里,提取模型可以是如上所述的前馈网络。元信息特征矢量可以由现有VVC中利用的映射表、根据块编码条件的模式等的结果组成。

[0213] 视频解码装置通过利用元信息特征矢量对当前块执行解码处理S1206。

[0214] 例如,当基于MIP进行解码处理时,元信息包括当前块的块结构、当前块的大小、当前块的参考样本、量化参数以及与帧内预测相关的参数;并且元信息特征矢量可以是指示集的一个的idx。此时,集包括用于MIP的应用的矩阵和矢量。

[0215] 此外,可以生成替换MIP模式的帧内预测模式。这里,元信息可以是当前块的参考样本,并且元信息特征矢量可以由指示帧内预测模式的替换索引repIdx组成。稍后,当对块执行帧内预测时,映射到替换的索引的帧内预测模式可以用于推导MPM。例如,如果利用MIP

预测位于当前块的左侧或上侧的亮度块,则视频解码装置可以将相应亮度块的替换的索引应用于构建MPM列表的过程。

[0216] 尽管描述了顺序执行的各个流程图中的步骤,但这些步骤仅仅例示了本发明的一些实施方案的技术思想。因此,本发明所属领域的普通技术人员可以通过改变在各个附图中描述的顺序或通过并行地执行两个或更多个步骤来执行步骤。因此,各个流程图中的步骤不限于所示的按发生时间排列的顺序。

[0217] 应当理解,上述描述呈现了可以以各种其他方式实现的说明性实施方案。在一些实施方案中描述的功能可以通过硬件、软件、固件和/或它们的组合来实现。还应当理解,本发明中描述的功能组件标记为“……单元”,以突出强调它们独立实现的可能性。

[0218] 另一方面,在一些实施方案中描述的各种方法或功能可以实现为存储在非易失性记录介质中的指令,所述指令可以由一个或更多个处理器读取和执行。非易失性记录介质可以包括例如以计算机系统可读取的形式存储数据的各种类型的记录装置。例如,非易失性记录介质可以包括存储介质,例如可擦除可编程只读存储器 (EPROM)、闪存驱动器、光盘驱动器、磁性硬盘驱动器和固态驱动器 (SSD) 等等。

[0219] 尽管出于说明的目的描述了本发明的示例性实施方案,但是本发明所属领域的普通技术人员应当理解,在不脱离本发明的思想和范围的情况下,可以进行各种修改、添加和替换。因此,出于简洁和清楚起见描述了本发明的实施方案。本发明的实施方案的技术思想的范围不受例示的限制。相应地,本发明所属领域的普通技术人员应当理解,本发明的范围不应受上述明确描述的实施方案的限制,而是受权利要求及其等同形式的限制。

[0220] 附图标记

[0221] 530:逆变换器

[0222] 540:预测器

[0223] 810:元信息生成器

[0224] 820:元信息特征矢量生成器

[0225] 830:解码执行器。

[0226] 相关申请的交叉引用

[0227] 本申请要求2022年6月22日提交的韩国专利申请No.10-2022-0075990和2023年6月12日提交的韩国专利申请No.10-2023-0074954的优先权和权益,其全部内容通过引用合并于本文中。

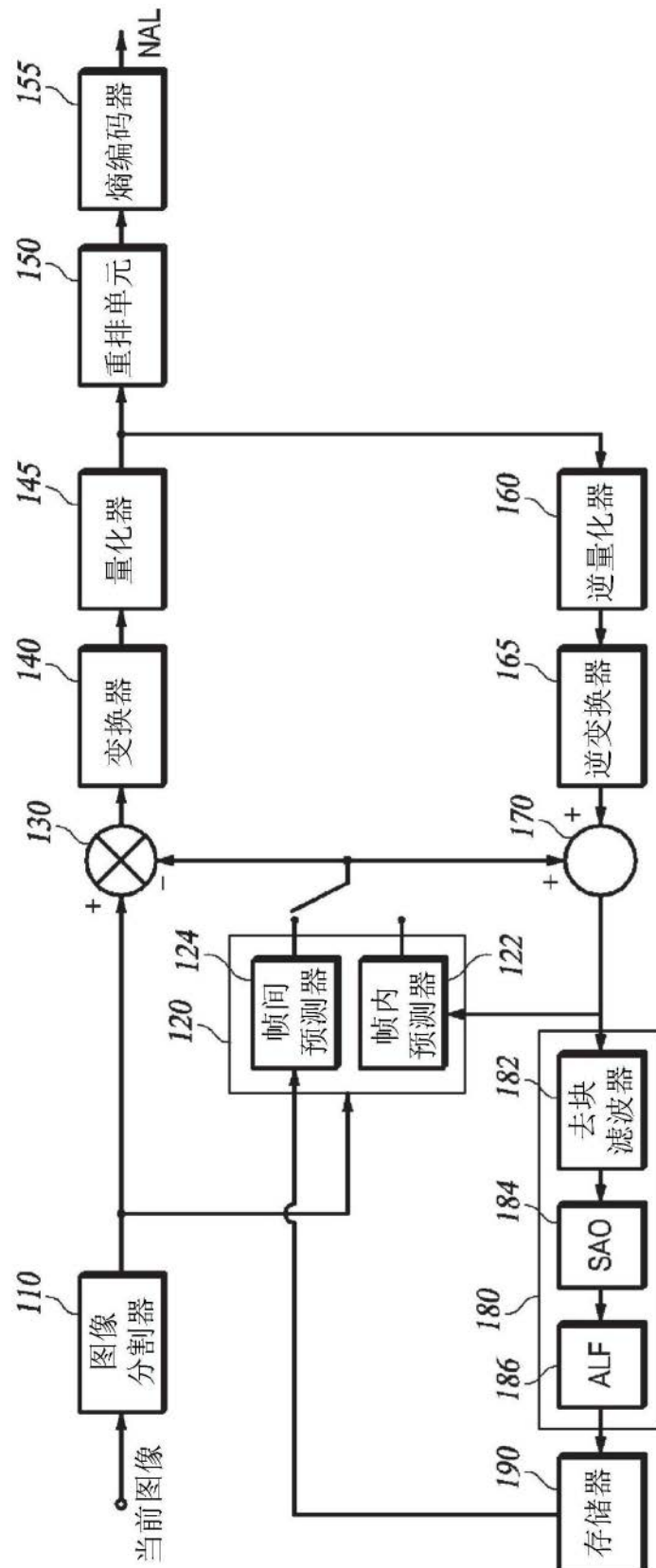


图1

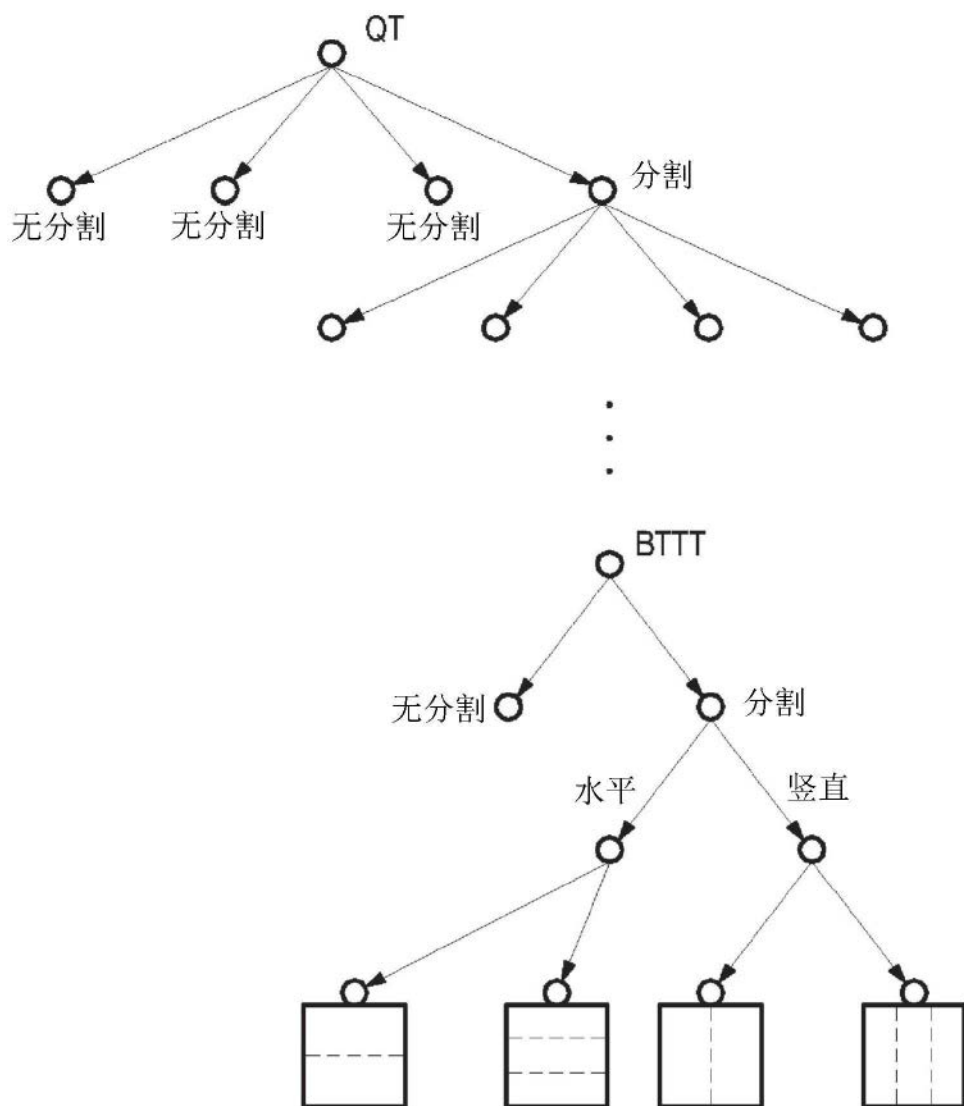


图2

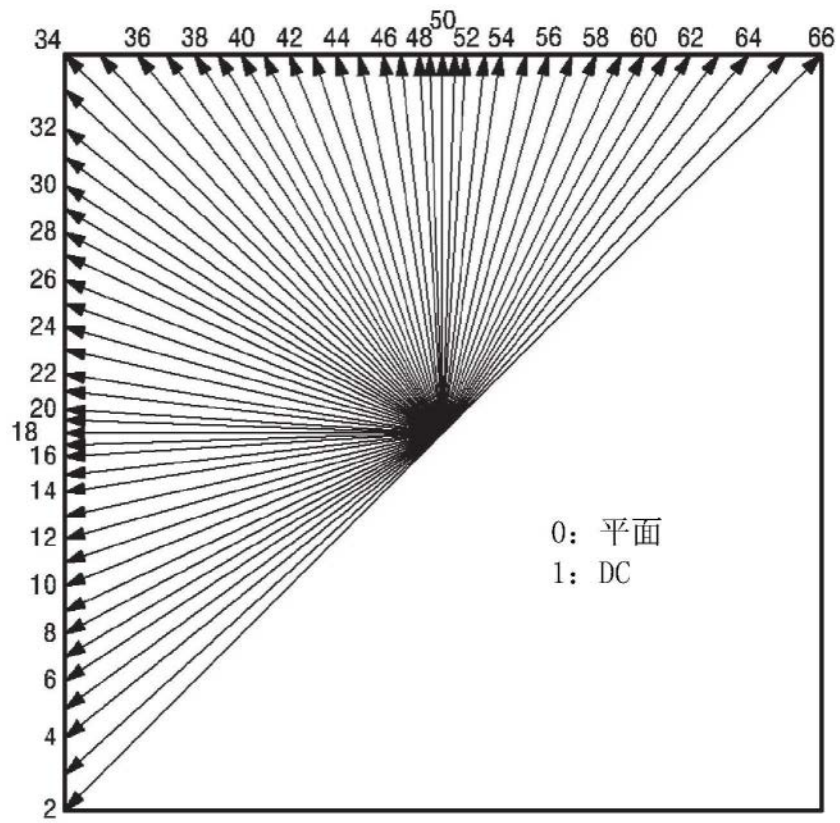


图3a

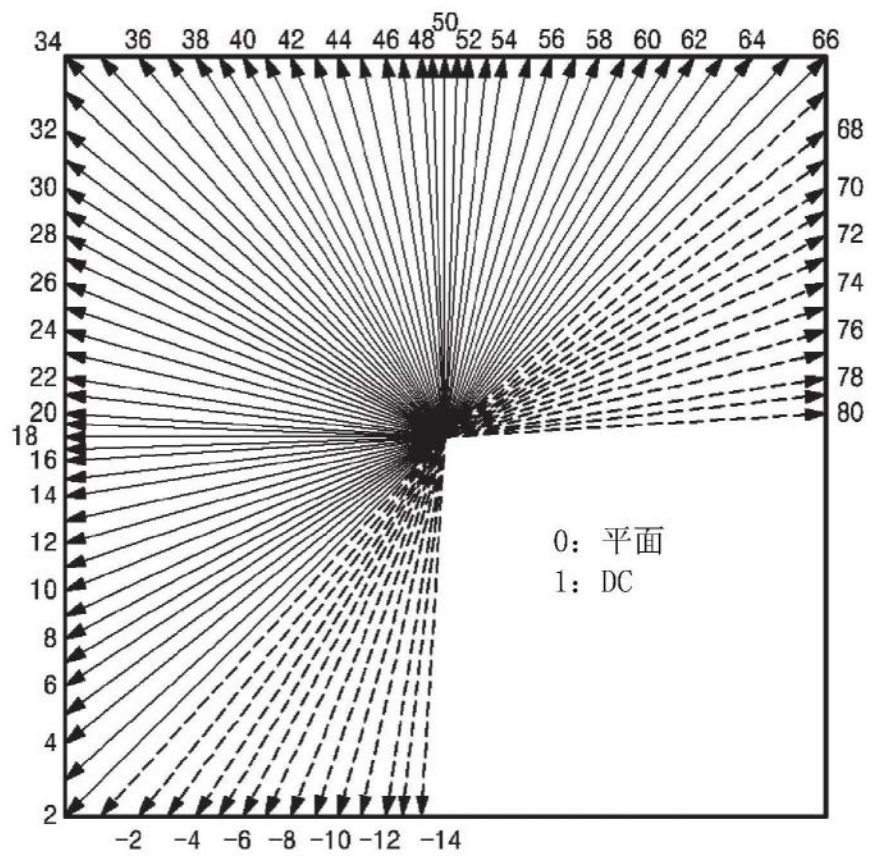


图3b

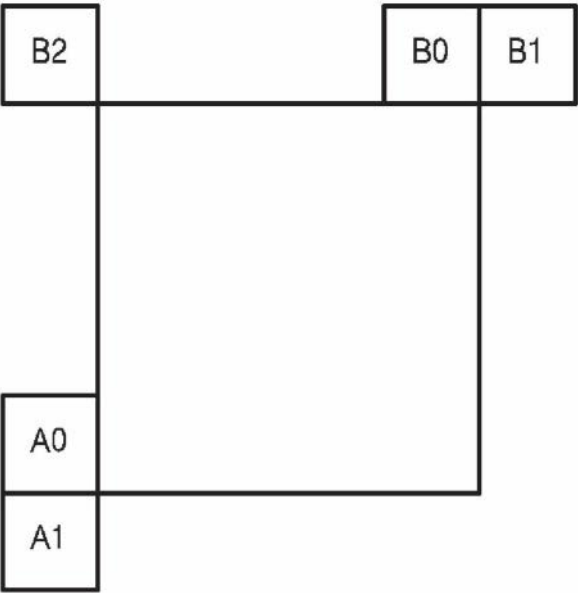


图4

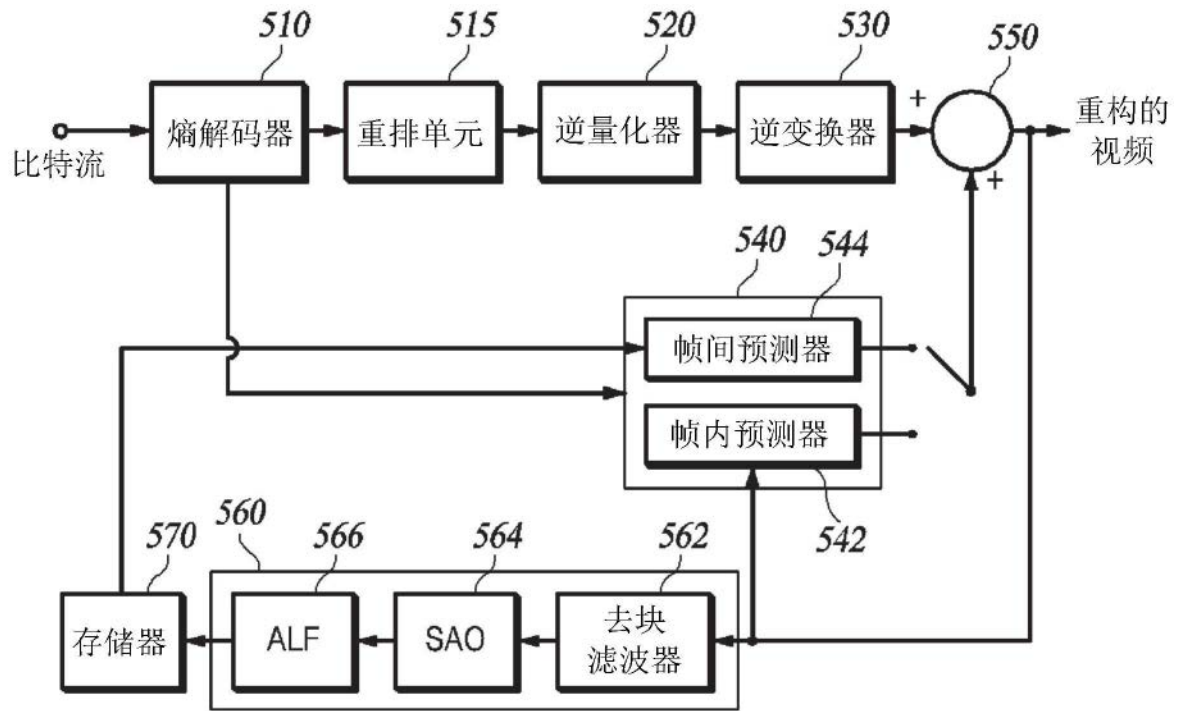


图5

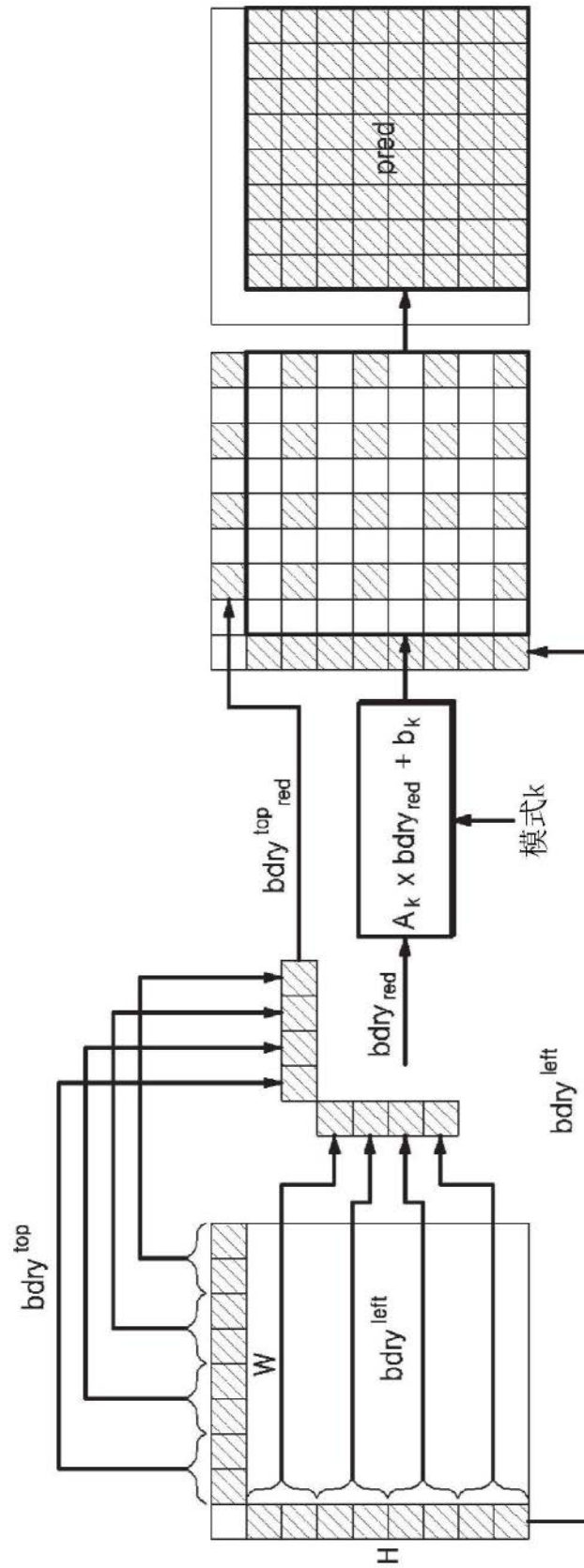


图6

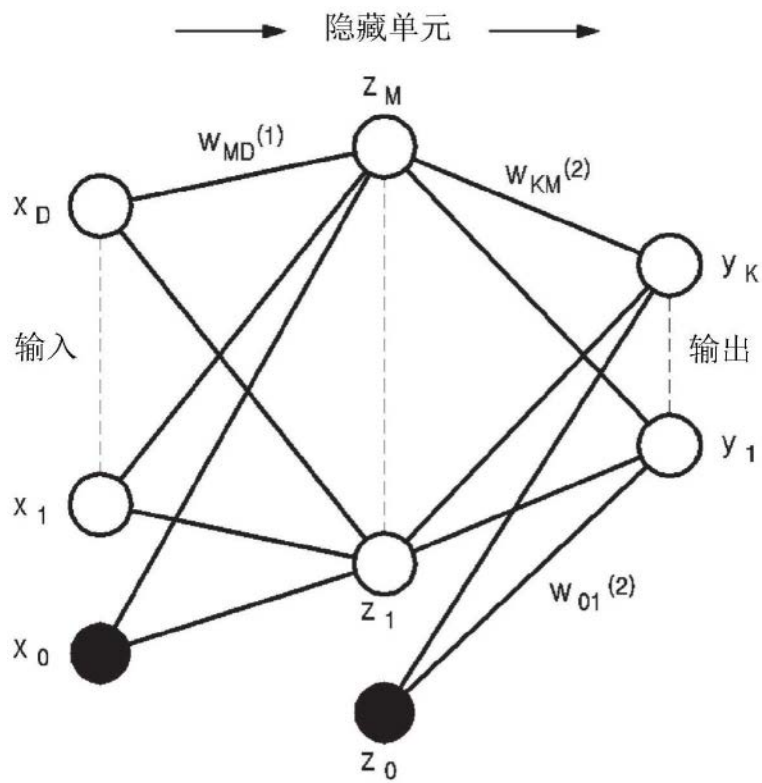


图7

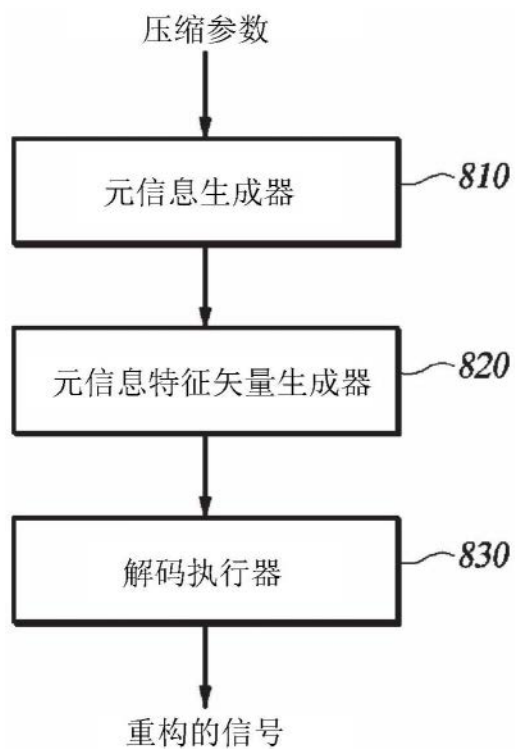


图8

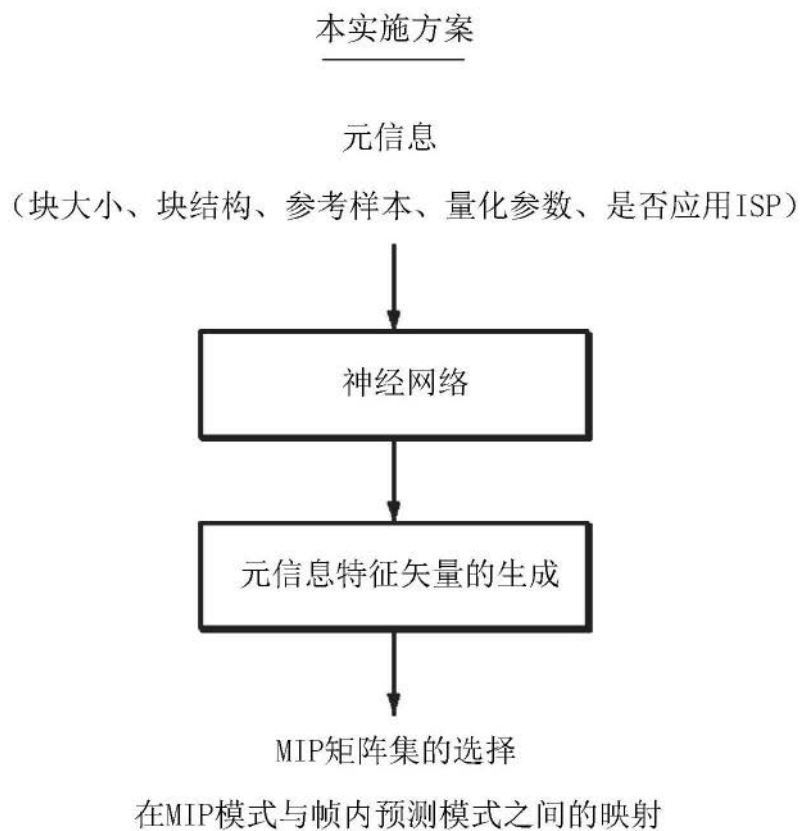
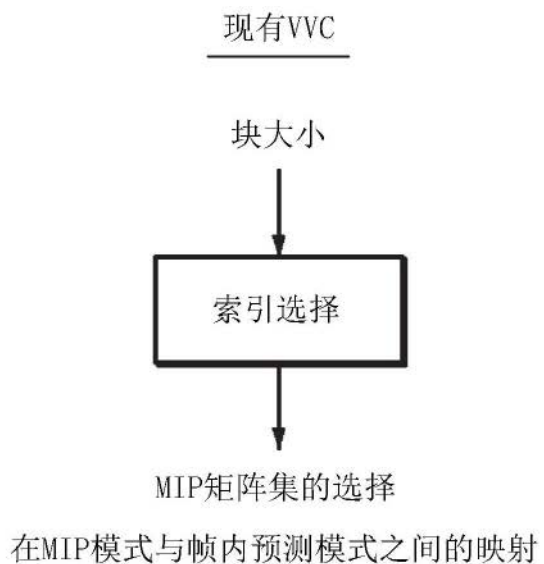


图9

现有VVC

是否应用帧内/帧间预测



在竖直/水平
方向上的变换核



变换核的选择

(DST7、DCT8、DCT2等)

本实施方案

元信息

(块大小、块结构、参考样本、量化参数、是否应用帧内/帧间预测)



神经网络



元信息特征矢量的生成



变换核的选择

(DST7、DCT8、DCT2等)

图10

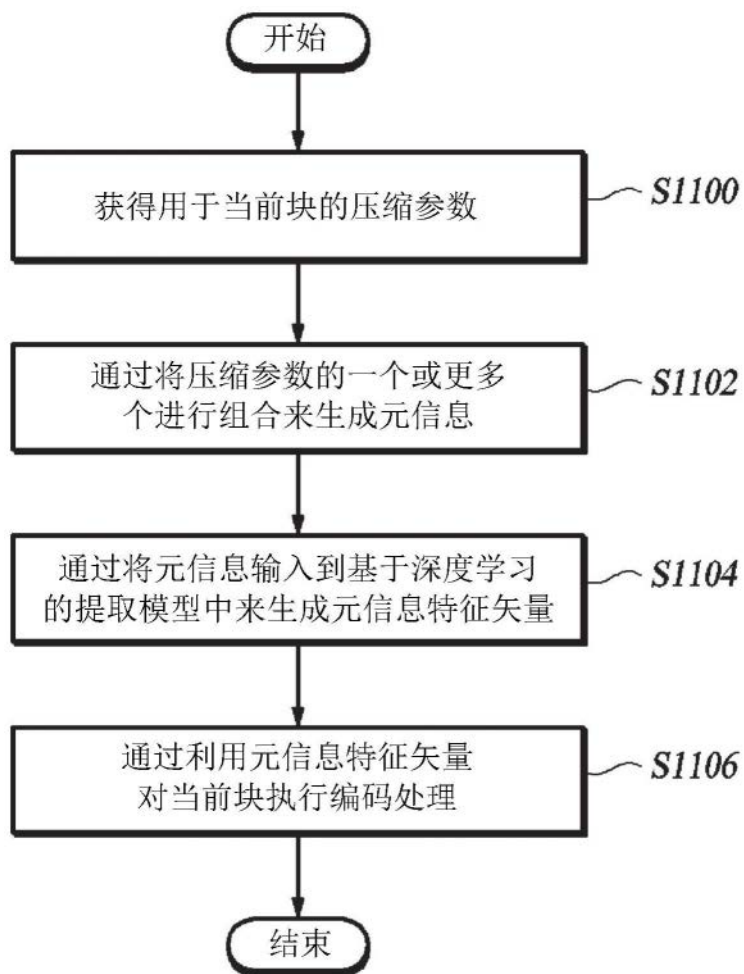


图11

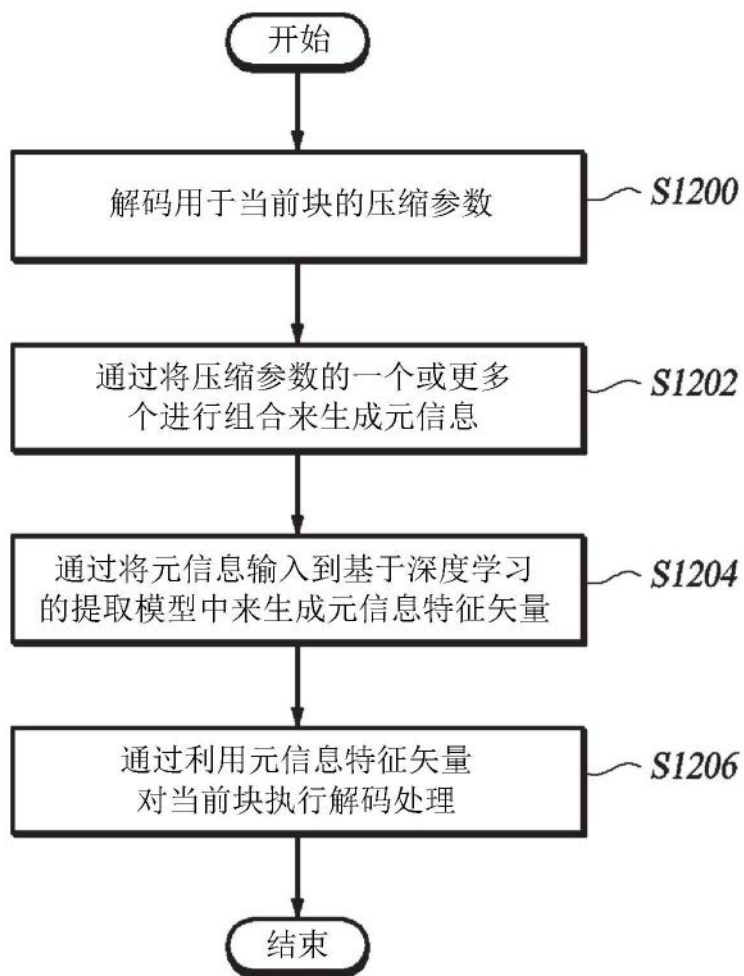


图12