(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116918323 A
(43)申请公布日 2023.10.20

(21)申请号 202280019115.8
(22)申请日 2022.03.03
(30)优先权数据
10-2021-0028794 2021.03.04 KR
10-2022-0026005 2022.02.28 KR
(85)PCT国际申请进入国家阶段日
2023.09.04
(86)PCT国际申请的申请数据
PCT/KR2022/002996 2022.03.03
(87)PCT国际申请的公布数据
WO2022/186620 KO 2022.09.09
(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
梨花女子大学校产学协力团
(72)发明人 姜制远 李订炅 金娜莹 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 19/11(2006.01)

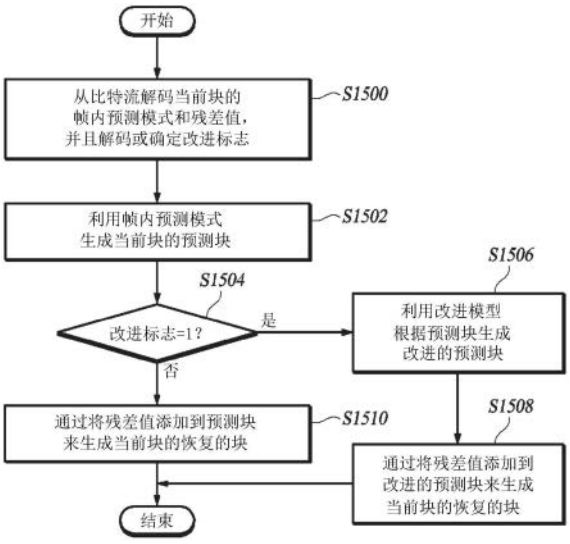
权利要求书2页 说明书17页 附图15页

(54)发明名称

用于改善帧内预测的预测信号的视频编解码方法和装置

(57)摘要

提供了一种视频编解码方法和装置,该编解码方法和装置与在帧内预测中对预测信号进行改进有关,并且为了减少要被编码的残差信号的数据量,本实施方案利用基于可变和固定系数的深度学习模型,根据帧内预测的预测信号来生成邻近于原始视频信号的改进的预测信号。



CN 116918323 A

1.一种视频解码装置,包括:

熵解码器,其配置为从比特流解码当前块的帧内预测模式和残差值,解码改进标志,或者确定改进标志,其中,所述改进标志指示在当前块的帧内预测时是否应用基于深度学习的改进模型;

帧内预测器,其配置为利用帧内预测模式生成当前块的预测块;

信号改进单元,其配置为在改进标志为1时,利用改进模型根据预测块生成改进的预测块;以及

加法器,其配置为在改进标志为1时通过将残差值添加到改进的预测块来生成当前块的恢复的块,或者在改进标志为0时通过将残差值添加到预测块来生成恢复的块。

2.根据权利要求1所述的装置,其中,改进模型利用包括基于固定系数的网络的深度学习模型来实现,并且进行预训练以生成近似于当前块的原始图像的改进的预测块。

3.根据权利要求1所述的装置,其中,改进模型利用包括基于固定系数的网络和基于可变系数的网络的深度学习模型来实现,并且在对当前块的原始图像进行编码时,训练基于可变系数的网络,而基于固定系数的网络是固定的。

4.根据权利要求3所述的装置,其中,所述熵解码器配置为从比特流解码基于可变系数的网络的参数。

5.根据权利要求1所述的装置,其中,所述改进标志从视频编码装置以每个块为基础,或者以每个图像或切片为基础进行发送。

6.根据权利要求1所述的装置,其中,在预定的预测模式用于当前块的帧内预测时,将改进标志确定为1。

7.根据权利要求1所述的装置,其中,当矩阵加权的帧内预测用于当前块的帧内预测时,将改进标志确定为1。

8.根据权利要求1所述的装置,其中,当使用多个参考线时,将改进标志确定为1,但是直接邻近于当前块的行或列的参考样本不用于当前块的帧内预测。

9.根据权利要求1所述的装置,其中,在对于当前块的帧内预测参考样本在邻近的行或列中不能够使用时,将改进标志确定为1。

10.根据权利要求1所述的装置,其中,在从当前块分区的子块用于当前块的帧内预测时,将改进标志确定为1。

11.一种由视频解码装置执行的用于当前块的帧内预测的视频解码方法,所述方法包括:

从比特流解码当前块的帧内预测模式和残差值,解码改进标志,或者确定改进标志,其中,所述改进标志指示在当前块的帧内预测时是否应用基于深度学习的改进模型;

利用帧内预测模式生成当前块的预测块;以及

基于改进标志生成当前块的恢复的块,

其中,在改进标志为1时,生成恢复的块包括:

利用改进模型根据预测块生成改进的预测块;并且

通过将残差值添加到改进的预测块来生成恢复的块,

其中,在改进标志为0时,生成恢复的块包括:

通过将残差值添加到预测块来生成恢复的块。

12.根据权利要求11所述的方法,其中,所述改进标志从视频编码装置以每个块为基础,或者以每个视频序列或切片为基础进行发送。

13.根据权利要求11所述的方法,其中,确定改进标志包括:

在预定的预测模式用于当前块的帧内预测时,将改进标志确定为1。

14.一种由视频编码装置执行的用于当前块的帧内预测的视频编码方法,所述方法包括:

获得当前块的帧内预测模式,获得改进标志,或者确定改进标志,其中,所述改进标志指示在当前块的帧内预测时是否应用基于深度学习的改进模型;

利用帧内预测模式生成当前块的预测块;以及

基于改进标志生成当前块的残差块,

其中,在改进标志为1时,生成残差块包括:

利用改进模型根据预测块生成改进的预测块;并且

通过将当前块减去改进的预测块来生成残差块,

其中,在改进标志为0时,生成残差块包括:

通过将当前块减去预测块来生成残差块。

15.根据权利要求14所述的方法,其中,所述改进模型利用包括基于固定系数的网络的深度学习模型来实现,并且进行预训练以生成近似于当前块的原始图像的改进的预测块。

16.根据权利要求14所述的方法,其中,所述改进模型利用包括基于固定系数的网络和基于可变系数的网络的深度学习模型来实现,并且在对当前块的原始图像进行编码时,训练基于可变系数的网络,而基于固定系数的网络是固定的。

17.根据权利要求16所述的方法,进一步包括:

对基于可变系数的网络的参数进行编码,并且将编码的参数发送到视频解码装置。

18.根据权利要求14所述的方法,进一步包括:

以每个块为基础向视频解码装置发送改进标志,或者以每个图像或切片为基础向视频解码装置发送改进标志。

19.根据权利要求14所述的方法,其中,确定改进标志包括:

当预定的预测模式用于当前块的帧内预测时,将改进标志确定为1。

用于改善帧内预测的预测信号的视频编解码方法和装置

技术领域

[0001] 本发明涉及一种视频编解码方法和装置,用于改善帧内预测中的预测的信号。

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较大的数据量,视频数据需要大量的硬件资源(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,解压接收到的压缩的视频数据,并且播放解压的视频数据。视频压缩技术包括H.264/AVC、高效率视频编解码(High Efficiency Video Coding,HEVC)和多功能视频编解码(Versatile Video Coding,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,要编码的数据量也在增多。相应地,需要一种与现有的压缩技术相比提供更高的编解码效率和改善的影像增强效果的新的压缩技术。

[0006] 近年来,基于深度学习的图像处理技术已经应用于现有的编码基本技术。通过将基于深度学习的图像处理技术应用于现有的编码技术(特别是,诸如帧间预测、帧内预测、环内滤波、变换等的压缩技术),可以提高编解码效率。代表性的应用示例包括基于由深度学习模型生成的虚拟参考帧的帧间预测,并且包括基于去噪模型的环内滤波。因此,需要进一步采用基于深度学习的图像处理技术来提高图像编码/解码的编解码效率。

发明内容

[0007] 技术问题

[0008] 在一些实施方案中,本发明试图提供一种视频编解码方法和装置,用于利用基于可变和固定系数的深度学习模型,根据帧内预测的预测信号来生成近似于原始视频信号的改进的预测信号,以减少要被编码的残差信号的数据量。

[0009] 技术方案

[0010] 本发明的至少一个方面提供了一种视频解码装置。该装置包括熵解码器,所述熵解码器配置为从比特流解码当前块的帧内预测模式和残差值,解码改进标志,或者确定改进标志。改进标志指示在当前块的帧内预测时是否应用基于深度学习的改进模型。该装置还包括帧内预测器,所述帧内预测器配置为利用帧内预测模式生成当前块的预测块。该装置还包括信号改进单元,所述信号改进单元配置为在改进标志为1时,利用改进模型根据预测块生成改进的预测块。该装置还包括加法器,所述加法器配置为在改进标志为1时通过将残差值添加到改进的预测块来生成当前块的恢复的块,或者在改进标志为0时通过将残差值添加到预测块来生成恢复的块。

[0011] 本发明的另一个方面提供了一种由视频解码装置执行的用于当前块的帧内预测

的视频解码方法。该方法包括从比特流解码当前块的帧内预测模式和残差值，解码改进标志，或者确定改进标志。改进标志指示在当前块的帧内预测时是否应用基于深度学习的改进模型。该方法还包括利用帧内预测模式生成当前块的预测块。该方法还包括基于改进标志生成当前块的恢复的块。在改进标志为1时，生成恢复的块包括利用改进模型根据预测块生成改进的预测块，并且通过将残差值添加到改进的预测块来生成恢复的块。当改进标志为0时，生成恢复的块包括通过将残差值添加到预测块来生成恢复的块。

[0012] 本发明的又一个方面提供了一种由视频编码装置执行的用于当前块的帧内预测的视频编码方法。该方法包括获得当前块的帧内预测模式，获得改进标志，或者确定改进标志。改进标志指示在当前块的帧内预测时是否应用基于深度学习的改进模型。该方法还包括利用帧内预测模式生成当前块的预测块。该方法还包括基于改进标志生成当前块的残差块。在改进标志为1时，生成残差块包括利用改进模型根据预测块生成改进的预测块，并且通过将当前块减去改进的预测块来生成残差块。在改进标志为0时，生成残差块包括通过将当前块减去预测块来生成残差块。

[0013] 有益效果

[0014] 如上所述，本发明提供了一种视频编解码方法和装置，用于利用基于可变和固定系数的深度学习模型，根据帧内预测的预测信号生成近似于原始视频信号的改进的预测信号，以减少要被编码的残差信号的数据量，并且提高编解码效率。

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(QTBTTT)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示出帧内预测的编码模式的传输。

[0021] 图7示出根据本发明的一个实施方案的利用帧内预测信号的改进的视频编码装置的框图。

[0022] 图8示出根据本发明的一个实施方案的利用帧内预测信号的改进的视频解码装置的框图。

[0023] 图9示出根据本发明的一个实施方案的包括固定系数网络的改进模型。

[0024] 图10示出根据本发明的另一个实施方案的包括固定系数网络的改进模型。

[0025] 图11示出根据本发明的一个实施方案的改进模型的操作。

[0026] 图12示出根据本发明的一个实施方案的包括可变系数网络的改进模型。

[0027] 图13示出根据本发明的另一个实施方案的包括可变系数网络的改进模型。

[0028] 图14是示出根据本发明的一个实施方案的利用帧内预测信号的改进的视频编码方法的流程图。

[0029] 图15是示出根据本发明的一个实施方案的利用帧内预测信号的改进的视频解码方法的流程图。

具体实施方式

[0030] 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,已省略对相关的已知组件和功能的详细描述。

[0031] 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。

[0032] 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。

[0033] 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。

[0034] 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更多个瓦片(tile)或/和切片(slice)。这里,一个或更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更多个编码树单元(coding tree unit,CTU)。另外,每个CTU通过树结构分割为一个或更多个编码单元(coding unit,CU)。应用于每个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更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,一个或更多个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。

[0035] 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。

[0036] 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。

[0037] 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,可以使用四叉树加二叉树(quadtree plus binarytree,QTBT)结构,或者可以使用四叉树加二叉树三叉树(quadtree plus binarytree ternarytree,QTBTTT)结构。这里,BTTT被添加到树结构以称为多类型树(multiple-type tree,MTT)。

[0038] 图2是用于描述通过利用QTBTTT结构来对块进行分割的方法的示意图。

[0039] 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结构中可以存在多个分割方向。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。

[0040] 替选地,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。

[0041] 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形式的两个块的类型。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。

[0042] CU可以根据从CTU分割的QTBT或QTBTTT而具有各种大小。在下文中,与要编码或解码的CU(即,QTBTTT的叶节点)相对应的块称为“当前块”。当采用QTBTTT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。

[0043] 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。

[0044] 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。

[0045] 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。

[0046] 为了对具有矩形形状的当前块进行高效的方向预测,可以额外地使用图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。

[0047] 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使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。

[0048] 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。

[0049] 帧间预测器124通过利用运动补偿处理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。

[0050] 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。

[0051] 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,使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和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可以由预恢复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可以由预恢复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恢复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恢复图像也可以额外地包括在参考图像列表1中。

[0052] 为了最小化用于编码运动信息所消耗的比特量,可以使用各种方法。

[0053] 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)。

[0054] 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”)。

[0055] 作为用于推导合并候选的相邻块,可以使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。

[0056] 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。

[0057] 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。

[0058] 此后,合并模式和合并跳过模式统称为合并/跳过模式。

[0059] 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。

[0060] 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,可以使用与图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,可以使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。

[0061] 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。

[0062] 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。

[0063] 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。

[0064] 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。

[0065] 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,并且通过利用子块作为变换单元来执行变换。替选地,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。

[0066] 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,可以使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。

[0067] 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。

[0068] 重排单元150可以对量化的残差值执行系数值的重新排列。

[0069] 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使用的扫描方法。

[0070] 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。

[0071] 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。

[0072] 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,以恢复残差块。

[0073] 加法器170将恢复的残差块和由预测器120生成的预测块相加,以恢复当前块。在对下一个块进行帧内预测时,恢复的当前块中的像素可以用作参考像素。

[0074] 环路滤波单元180对恢复的像素执行滤波,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。

[0075] 去块滤波器182对恢复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SAO滤波器184和ALF 186对去块滤波的视频进行额外的滤波。SAO滤波器184和ALF 186是用于补偿由于有损编码(lossy coding)而发生的恢复的像素与原始像素之间的差异的滤波器。SAO滤波器184应用偏移作为CTU单位,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。

[0076] 通过去块滤波器182、SAO滤波器184和ALF 186滤波的恢复的块存储在存储器190中。当一个图像中的所有块被恢复时,恢复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。

[0077] 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。

[0078] 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。

[0079] 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。

[0080] 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,并且提取恢复当前块所需的预测信息和关于残差信号的信息。

[0081] 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。

[0082] 例如,当通过利用QTBTTT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。

[0083] 作为另一个示例,当通过利用QTBTTT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。

[0084] 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。

[0085] 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。

[0086] 此外,熵解码器510提取量化相关的信息并且提取关于量化的当前块的变换系数的信息作为关于残差信号的信息。

[0087] 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)。

[0088] 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。

[0089] 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恢复残差信号,以生成当前块的残差块。

[0090] 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恢复残差信号,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。

[0091] 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。

[0092] 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。

[0093] 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。

[0094] 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。

[0095] 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恢复当前块。在对后续要被解码的块进行帧内预测时，恢复的当前块内的像素用作参考像素。

[0096] 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对恢复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在去块滤波之后对恢复的块执行额外的滤波，以便补偿由于有损编码而发生的恢复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。

[0097] 通过去块滤波器562、SAO滤波器564和ALF 566滤波的恢复的块存储在存储器570中。当一个图像中的所有块被恢复时，恢复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。

[0098] 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，本发明提供了一种视频编解码方法和装置，用于利用基于可变和固定系数的深度学习模型，根据帧内预测的预测信号来生成近似于原始视频信号的改进的预测信号。

[0099] 下面的实施方案可以通常应用于涉及用于视频编码和解码装置的深度学习技术的情况。

[0100] 在下面的描述中，术语待编码/解码的“目标块”可以与如上所述的当前块或编码单元(CU)互换使用，或者术语待编码/解码的“目标块”可以指代编码单元的一些区域。

[0101] I.帧内预测的编码模式

[0102] 如上所述，帧内预测是通过参考位于编码的当前目标块周围的样本来预测当前块的方法。如图3a和图3b所示，多功能视频编解码(VVC)技术可以利用DC/平面模式的非方向预测模式、65种方向预测模式和宽角度帧内预测模式。此外，帧内预测可以利用预测技术，例如多参考线帧内预测(multiple reference line intra prediction,MRLP)、交叉分量线性模型(cross-component linear model,CCLM)、位置决定的帧内预测组合(position dependent intra prediction combination,PDPC)、帧内子分区(intra sub-partitions,ISP)和矩阵加权的帧内预测(matrix-weighted intra prediction,MIP)。

[0103] 在利用MRLP的帧内预测处理中，视频编码/解码装置可以利用多个参考线(multiple reference lines,MRL)以采用更多的参考线。当应用MRL时，视频编码/解码装置可以利用除了原始参考线之外添加在当前块的顶部和左侧的两条线的样本对当前块执行帧内预测。当应用MRL时，指示参考线的索引(mrl_idx)可以用信号通知给视频解码装置以选择参考线。

[0104] CCLM预测是利用代表亮度信号与色度信号之间的相似性的线性模型的帧内预测方法。为了激活CCLM模式，编码装置可以将用于激活CCLM模式的标志用信号通知给视频解码装置。

[0105] CCLM预测首先基于当前色度块来推导相邻参考样本与位于与相邻参考样本相同位置的亮度信号参考样本之间的线性变换函数。此时，可以基于相邻亮度信号的最小值、与

相邻亮度信号同位的色度值、相邻亮度信号的最大值以及与相邻亮度信号同位的色度值来推导线性变换函数。接下来，通过将线性变换函数应用于与色度块同位的亮度样本来执行对色度样本的预测。

[0106] 用于帧内预测的基于规则的预测方法的一种是位置决定的帧内预测组合(Position Dependent Intra Prediction Combination,PDPC)。换句话说，可以基于预定义的运算，利用执行帧内预测的目标块和在空间上与目标块邻近的相邻像素的编码信息来生成预测器(predictor)。

[0107] PDPC修改根据特定的帧内预测模式生成的预测的样本，以生成当前块的帧内预测器。这里，类似于图3a中所示的预测模式，特定的帧内预测模式包括平面模式、DC模式、水平模式(预测模式18)、竖直模式(预测模式50)、左下对角线方向模式(预测模式2)和接近左下对角线方向模式的15种方向模式、以及右上对角线方向模式(预测模式66)和接近右上对角线方向模式的15种方向模式。

[0108] 对于根据特定的帧内预测模式生成的当前块的预测样本，PDPC技术可以利用预定义的权重和相邻像素的位置信息来调整每个像素值以生成预测样本。

[0109] 如上所述，除了非方向模式(即，Planar和DC)之外，亮度块的帧内预测模式还具有细分的方向模式(即，-14到80)，如图3a和图3b所示。在将当前块子分区为相同尺寸的较小块之后，ISP技术跨越整个子块共享当前块的帧内预测，然而，可以对每个子块应用不同的变换。在进行子分区时，块可以在水平或竖直方向上被分区。

[0110] 可以基于预定义的矩阵运算，利用与执行帧内预测的当前块邻近的像素和当前块的编码信息来生成预测器。上述基于规则的预测方法称为矩阵加权的帧内预测(Matrix weighted Intra Prediction,MIP)。

[0111] MIP利用预定义的矩阵运算生成所有或部分帧内预测器。当生成预测器的一部分时，MIP可以通过利用预测器的一部分额外地执行用于上采样或放大的插值(interpolation)来生成与当前块的尺寸相等的最终帧内预测的样本。

[0112] 另一方面，MIP可以选择性地选择在空间上与当前块邻近的像素中的一部分像素，并且使用选择的像素作为当前块的相邻像素。作为另一个实施方案，MIP可以使用根据基于诸如矩阵运算的子采样或缩小的方法的运算所推导的值。

[0113] 根据用于传输帧内预测的编码模式的方法，帧内预测的预测模式以及是否应用预测技术可以从视频编码装置用信号通知给视频解码装置，如图6所示。例如，当用帧内预测对当前块进行编码时，视频编码装置可以将pred_mode_flag用信号通知为0，然后可以利用intra_mip_flag用信号通知是否应用MIP技术。

[0114] 在图6的示例中，当执行当前块的帧内预测时，最可能模式(most probable mode，MPM)技术使用相邻块的帧内预测模式。视频编码装置可以通过传输MPM列表的索引而不是预测模式的索引来提高帧内预测模式的编解码效率。另一方面，图6的示例没有提供用于用信号通知没有应用MIP的编码模式的方法的详细描述。

[0115] 另一方面，当执行帧内预测时，帧内块复制(intra block copy,IBC)技术在与当前块的预测块相同的图像内生成参考块，而不是使用参考样本。此时，块矢量代表指示参考块的位移，并且将其从视频编码装置用信号通知给视频解码装置。

[0116] II.帧内预测的块改进

[0117] 图7示出根据本发明的一个实施方案利用帧内预测信号的改进的视频编码装置的框图。

[0118] 根据本实施方案的视频编码装置额外地包括在作为基本构成元件之一的帧内预测器122之后的信号改进单元710。这里，根据本实施方案的视频编码装置中包括的构成元件不必限于具体示例。例如，视频编码装置可以额外地包括用于训练信号改进单元710中包括的基于深度学习的改进模型的训练单元(未示出)，或者可以实现为与外部训练单元结合操作。

[0119] 帧内预测器122利用预测模式从相邻参考样本生成包括要编码的当前目标块的预测的信号的预测块。

[0120] 视频编码装置可以使用改进标志refinement_flag以指示将改进模型应用于预测块。视频编码装置可以基于每个块向视频解码装置发送改进标志，或者可以在将改进标志包括在SPS中之后基于每个图像或切片发送改进标志。

[0121] 在改进标志为1时，信号改进单元710利用改进模型根据预测块生成改进的预测块。另一方面，训练单元可以训练改进模型，使得改进模型学习信号生成方法，以生成近似于当前块的原始信号的改进的信号。

[0122] 在下面的描述中，术语“改进的预测模型”可以与术语“改进的块”互换使用。

[0123] 在改进标志为1时，视频编码装置通过将当前块减去改进的块来生成残差块，并且在改进标志为0时，视频编码装置通过将当前块减去预测块来生成残差块。视频编码装置可以通过将残差块的残差值输入到变换器140来执行上述编码处理。

[0124] 另一方面，作为示例，视频编码装置可以如下设置改进标志的值。在基于当前块的尺寸确定帧内预测模式的数量N(其中，N是自然数)之后，视频编码装置通过对当前块执行粗略模式决策(Rough Mode Decision，RMD)来确定N个候选预测模式。视频编码装置利用N个候选预测模式和MPM中包括的预测模式来为每个预测模式生成预测块，然后为每个预测块计算率失真代价(rate distortion cost，RD-cost)。此外，视频编码装置通过将改进模型应用于每个预测块来生成改进的预测块，然后为每个改进的预测块计算率失真代价。视频编码装置可以比较改进的块与预测块之间的率失真代价，当利用改进的块的代价变得最小时确定相应的候选预测模式作为当前块的帧内预测模式，并且将改进标志refinement_flag设置为1。

[0125] 图8示出根据本发明的一个实施方案的利用帧内预测信号的改进的视频解码装置的框图。

[0126] 根据本实施方案的视频解码装置额外地包括在作为基本构成元件之一的帧内预测器122之后的信号改进单元810。

[0127] 熵解码器510从比特流对要解码的当前块的帧内预测模式、改进标志refinement_flag以及包括残差值的残差块进行解码。

[0128] 帧内预测器542利用帧内预测模式，从相邻参考样本生成包括要解码的当前块的预测的信号的预测块。

[0129] 在改进标志为1时，信号改进单元810利用改进模型根据预测块生成改进的块。在改进标志为1时，视频解码装置通过将残差值添加到改进的块来生成用于当前块的恢复的块，并且在改进标志为0时，通过将残差值添加到预测块来生成用于当前块的恢复的块。

[0130] 在下面的描述中,对视频编码装置内的改进模型进行描述。以下描述可以同样应用于在视频解码装置中使用的改进模型。

[0131] 在一个实施方案中,改进模型可以是利用仅包括固定系数的卷积层实现的深度学习模型(在下面的描述中,“基于固定系数的网络”)。基于固定系数的网络可以用于改进输入块以近似于原始块。由于训练单元利用各种输入数据预先训练具有相对大量参数的深度学习模型,可以实现包括固定系数网络的改进模型。

[0132] 可以基于噪声预测方法来实现包括基于固定系数的网络的改进模型。例如,如图9所示,改进模型可以通过估计预测块中的噪声并且从预测块移除估计的噪声来生成改进的信号。

[0133] 替选地,改进模型可以使用利用相邻像素值来生成改进的信号的方法。例如,改进模型可以根据IBC的应用使用由块矢量指示的块作为当前块的预测块,并且可以根据预测块生成改进的信号。此外,如图10所示,改进模型可以使用由块矢量找到的预测块作为输入,以基于每个子块或像素生成额外矢量,并且利用这些矢量生成改进的信号。

[0134] 替选地,改进模型可以填充用于当前块的帧内预测的相邻参考样本以用作输入,并且根据输入生成改进的信号。

[0135] 另一方面,如图11所示,改进模型的输入是根据特定模式获得的预测块,并且改进模型的输出对应于改进的预测块。这里,特定模式可以是用于帧内预测的任意预测模式或预测模式的组合。

[0136] 替选地,输入可以由根据用于帧内预测的预测模式的预测块和根据帧间预测的预测块的加权和组成。改进模型可以从输入生成改进的信号。

[0137] 视频编码装置通过比较根据现有编码模式的预测块与改进的块之间的率失真代价来确定是否对预测块进行改进。之后,如上所述,视频解码装置添加指示是否对预测信号进行改进的改进标志refinement_flag,并且将改进标志用信号通知给视频解码装置。当改进的块的率失真代价小于预测块的率失真代价时,视频编码装置可以将改进标志refinement_flag设置为1。

[0138] 率失真代价J_{RD}可以通过以下等式1计算。

[0139] [等式1]

[0140] $J_{RD}=SSD(S,C)+\lambda \cdot R$

[0141] 在等式1中,S代表原始块,并且C代表恢复的块,它们可以是预测块或改进的预测块。平方差之和(The Sum of Squared Difference,SSD)代表原始块与恢复的块之间的相似性。R代表估计的比特率,并且λ代表拉格朗日乘数。

[0142] 另一方面,基于固定系数的网络的参数在视频编码装置与视频解码装置之间共享。

[0143] 在另一个实施方案中,改进模型可以是利用基于固定系数的网络和仅包括可变系数的卷积层实现的深度学习模型(在下面的描述中,“基于可变系数的网络”)。基于可变系数的网络可以用相对少量的参数来实现,这是因为必须传输网络参数。由于训练单元在原始块的编码期间训练深度学习模型,同时固定预训练的基于固定系数的网络,可以实现包括基于可变系数的网络的改进模型。

[0144] 包括基于可变系数的网络的改进模型可以使用掩模映射。掩模映射是通过对基于

固定系数的网络的输出信号x_{fixed}执行适当的操作来生成改进的块$x_{refined}$的矢量或变量。例如，如图12所示，改进模型的基于固定系数的网络生成输出信号x_{fixed}，并且改进网络的基于可变系数的网络生成掩模映射m。改进模型使用掩模映射以执行输出信号x_{fixed}和预测块x的加权和，从而生成改进的块$x_{refined}$。

[0145] 替选地，改进模型可以使用由多个卷积层组成的基于可变系数的网络。例如，如图13所示，改进模型的基于固定系数的网络可以生成输出信号x_{fixed}，并且改进模型的基于可变系数的网络可以根据输出信号x_{fixed}生成改进的块$x_{refined}$。

[0146] 视频编码装置可以通过考虑与基于可变系数的网络参数相关的比特率R'来计算率失真代价J_{RD}，如等式2所示。

[0147] [等式2]

[0148] $J_{RD}=SSD(S,C)+\lambda \cdot (R+R')$

[0149] 例如，由具有3×3核的卷积层组成的基于可变系数的网络总共需要10个参数，包括偏差参数。因此，当利用16位浮点类型(float 16type)传输参数时，额外需要总共160位。

[0150] 另一方面，基于可变系数的网络的参数θ需要从视频编码装置传输到视频解码装置。可以通过考虑参数的尺寸来确定传输周期。例如，参数θ可以在刷新参数的每个帧内帧(I帧)传输。

[0151] 另一方面，如上所述，视频编码装置可以利用比特流显式地传输改进标志refinement_flag。当refinement_flag＝1时，视频解码装置可以通过将改进模型应用于预测块来生成改进的块。另一方面，当refinement_flag＝0时，视频解码装置可以跳过改进模型的应用并且执行传统的帧内预测。如上所述，视频编码装置可以以每个块为基础，或者以每个视频序列或切片为基础传输改进标志。在以每个块为基础的传输的情况下，视频编码装置可以额外地使用指示以每个块为基础的改进标志的存在的更高级别的标志。

[0152] 作为另一个示例，可以隐式地确定改进标志refinement_flag。当确定出refinement_flag＝1时，视频解码装置可以通过将改进模型应用于预测块来生成改进的块。另一方面，当确定出refinement_flag＝0时，视频解码装置可以跳过改进模型的应用并且执行传统的帧内预测。

[0153] 用于隐式地确定改进标志的方法可以采用以下方法的一种。

[0154] 在一种方法中，帧内预测被选择为特定模式。例如，当帧内预测模式是Planar时，改进标志可以被隐式地确定为1。

[0155] 当帧内预测采用MIP模式或PDPC模式而不是传统的帧内预测模式时，改进标志可以被隐式地确定为1。

[0156] 当使用基于MRLP的帧内预测，并且不使用紧密邻近的行或列中的参考样本时，改进标志可以被隐式地确定为1。

[0157] 当对于帧内预测，参考样本在直接邻近的行或列中不可用时，改进标志可以被隐式地确定为1。

[0158] 当对帧内预测应用ISP模式时，改进标志可以被隐式地确定为1。

[0159] 如上所述的改进模型仅接收预测信号并且执行信号改进。作为另一个实施方案，为了实现改善预测信号的相同目标，可以通过利用来自预测之后获得的残差信号、相邻块的信号或通过应用两个或更多不同的预测模式获得的信号的一个或组合的输入来应用改

进模型。

[0160] 另一方面，如上所述，根据本实施方案的预测块的改进可以基于传统的帧内预测模式对预测块进行改进。此外，可以添加预测信号改进模式作为用于视频编码的新预测模式。

[0161] 首先，根据传统的帧内预测模式的预测块的改进可以执行如下。

[0162] 视频编码装置可以通过将根据本实施方案的改进模型应用于通过首先执行可用于编码的所有或部分预测模式而获得的预测块来生成改进的预测块。例如，视频编码装置可以通过将改进模型应用于通过执行DC/Planar非方向预测模式、65种方向预测模式、ISP模式、MIP模式等获得的预测块来生成改进的预测块。

[0163] 替选地，视频编码装置可以在编码时对一个或两个或更多个选择的预测模式执行预测块改进。例如，视频编码装置可以对从65种方向预测模式获得的预测块执行预测块改进。

[0164] 替选地，视频编码装置可以对PU尺寸大于或等于、或者小于或等于预定阈值的块执行预测块改进。

[0165] 根据上述传统的帧内预测模式的预测块的改进可以在视频解码装置中类似地执行。

[0166] 接下来，视频编码装置可以添加预测信号改进模式作为新模式。

[0167] 当添加预测信号改进模式作为新编码模式时，视频编码装置可以根据将率失真代价与现有模式的率失真代价进行比较的结果来选择相应的模式。例如，预测信号改进模式可以添加到现有的帧内预测编码模式，例如非方向预测模式、方向预测模式、ISP和MIP。当应用预测信号改进模式时，视频编码装置可以选择性地使用方向预测模式中的一个方向或任意方向上的参考样本以生成预测块。然后，视频编码装置可以通过应用预测信号改进模型来生成改进的预测块。

[0168] 上述预测信号改进模式可以类似地应用于视频解码装置。

[0169] 在下面的描述中，参考图14和图15对利用帧内预测信号的改进的视频编码方法和视频解码方法进行描述。

[0170] 图14是示出根据本发明的一个实施方案的利用帧内预测信号的改进的视频编码方法的流程图。

[0171] 视频编码装置获得当前块的帧内预测模式，并且获得或确定改进标志(S1400)。这里，改进标志refinement_flag指示当执行当前块的帧内预测时是否应用基于深度学习的改进模型。视频编码装置可以基于每个块向视频解码装置发送改进标志，或者可以在将改进标志包括在SPS中之后基于每个图像或切片发送改进标志。

[0172] 在一个示例中，视频解码装置可以获得如下设置的改进标志的值。视频解码装置比较多个候选帧内预测模式的改进的块与预测块之间的率失真代价。当利用改进的块产生最小代价时，视频编码装置确定相应的候选预测模式作为当前块的帧内预测模式，并且将改进标志设置为1。

[0173] 在另一个示例中，视频编码装置可以如下隐式地确定改进标志的值。在当前块的帧内预测采用预定的预测模式(例如，平面模式)时，视频编码装置可以将改进标志确定为1。在当前块的帧内预测采用MIP或PDPC模式时，视频编码装置可以将改进标志确定为1。此

外,在当前块的帧内预测采用多个参考线但不使用直接邻近于当前块的行或列的参考样本时,视频编码装置可以将改进标志确定为1。在参考样本在相邻行或列中不可用时,视频编码装置可以将改进标志确定为1。此外,在当前块的帧内预测采用ISP模式时,视频编码装置可以将改进标志确定为1。

[0174] 视频编码装置利用帧内预测模式生成当前块的预测块(S1402)。

[0175] 视频编码装置检查改进标志的值(S1404)。

[0176] 在改进标志为1时(S1404中的是),视频编码装置利用改进模型根据预测块生成改进的预测块(S1406),然后通过将当前块减去改进的预测块来生成残差块(S1408)。

[0177] 对改进模型的输入是根据特定模式获得的预测块,并且改进模型的输出对应于改进的预测块。这里,特定模式可以是用于帧内预测的任意预测模式或预测模式的组合。

[0178] 在一个实施方案中,改进模型可以是仅包括基于固定系数的网络的深度学习模型。对包括基于固定系数的模型的改进模型进行预训练,以生成近似于当前块的原始图像的改进的预测块。基于固定系数的网络的参数在视频编码装置与视频解码装置之间共享。

[0179] 在另一个示例中,改进模型可以是包括基于固定系数的网络和基于可变系数的网络的深度学习模型。在包括基于可变系数的网络的改进模型的情况下,在对当前块的原始图像进行编码时,对基于可变系数的网络进行训练,而基于固定系数的网络是固定的。另一方面,视频编码装置可以对基于可变系数的网络的参数进行编码,然后可以将参数发送到视频解码装置。可以通过考虑参数的尺寸来确定传输周期。例如,可以在刷新参数的每个帧内帧(I帧)传输参数。

[0180] 另一方面,在改进标志为0时(S1404中的否),视频编码装置通过将当前块减去预测块来生成残差块(S1410)。

[0181] 之后,视频编码装置可以执行对残差块的残差值进行编码的处理。

[0182] 图15是示出根据本发明的一个实施方案的利用帧内预测信号的改进的视频解码方法的流程图。

[0183] 视频解码装置从比特流解码当前块的帧内预测模式和残差值,并且解码或确定改进标志(S1500)。这里,改进标志refinement_flag指示当执行当前块的帧内预测时是否应用基于深度学习的改进模型。

[0184] 如上所述,改进标志可以从视频编码装置以每个块为基础传输,或者以每个视频序列或切片为基础传输。

[0185] 作为另一个示例,视频解码装置可以类似于视频编码装置的操作而隐式地确定改进标志的值。

[0186] 视频解码装置利用帧内预测模式生成当前块的预测块(S1502)。

[0187] 视频解码装置检查改进标志的值(S1504)。

[0188] 在改进标志为1时(S1504中的是),视频解码装置利用改进模型根据预测块生成改进的预测块(S1506),然后通过将残差值添加到改进的预测块来生成当前块的恢复的块(S1508)。

[0189] 改进模型可以是仅包括基于固定系数的网络的深度学习模型。如上所述,基于固定系数的网络的参数在视频编码装置与视频解码装置之间共享。

[0190] 在另一个示例中,改进可以是包括基于固定系数的网络和基于可变系数的网络的

深度学习模型。在包括基于可变系数的网络的深度学习模型的情况下，视频解码装置从比特流解码基于可变系数的网络的参数。

[0191] 另一方面，在改进标志为0时(S1504中的否)，视频解码装置通过将残差值添加到预测块来生成当前块的恢复的块(S1510)。

[0192] 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。

[0193] 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，本说明书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。

[0194] 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。

[0195] 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，本发明的范围不应当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。

[0196] (附图标记)

[0197] 122:帧内预测器

[0198] 510:熵解码器

[0199] 542:帧内预测器

[0200] 710:信号改进单元

[0201] 810:信号改进单元。

[0202] 相关申请的交叉引用

[0203] 本申请要求于2021年3月4日提交的韩国专利申请No.10-2021-0028794以及于2022年2月28日提交的韩国专利申请No.10-2022-0026005的优先权，其全部内容通过引用结合于本文中。

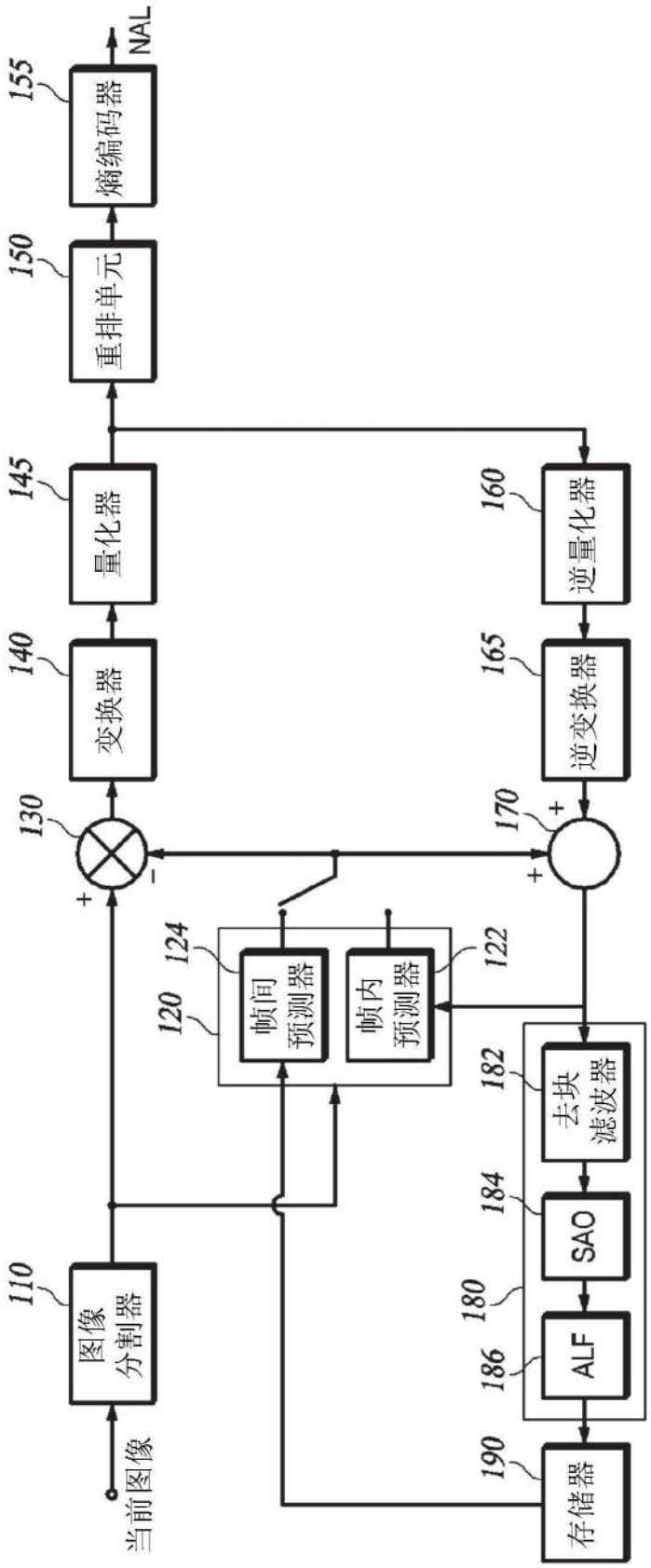

NAL
155
熵编码器
150
重排单元
145
量化器
140
变换器
130
110
图像
分割器
当前图像
160
逆量化器
165
逆变换器
170
120
124
帧间
预测器
122
帧内
预测器
180
182
去块
滤波器
184
SAO
186
ALF
190
存储器

图1

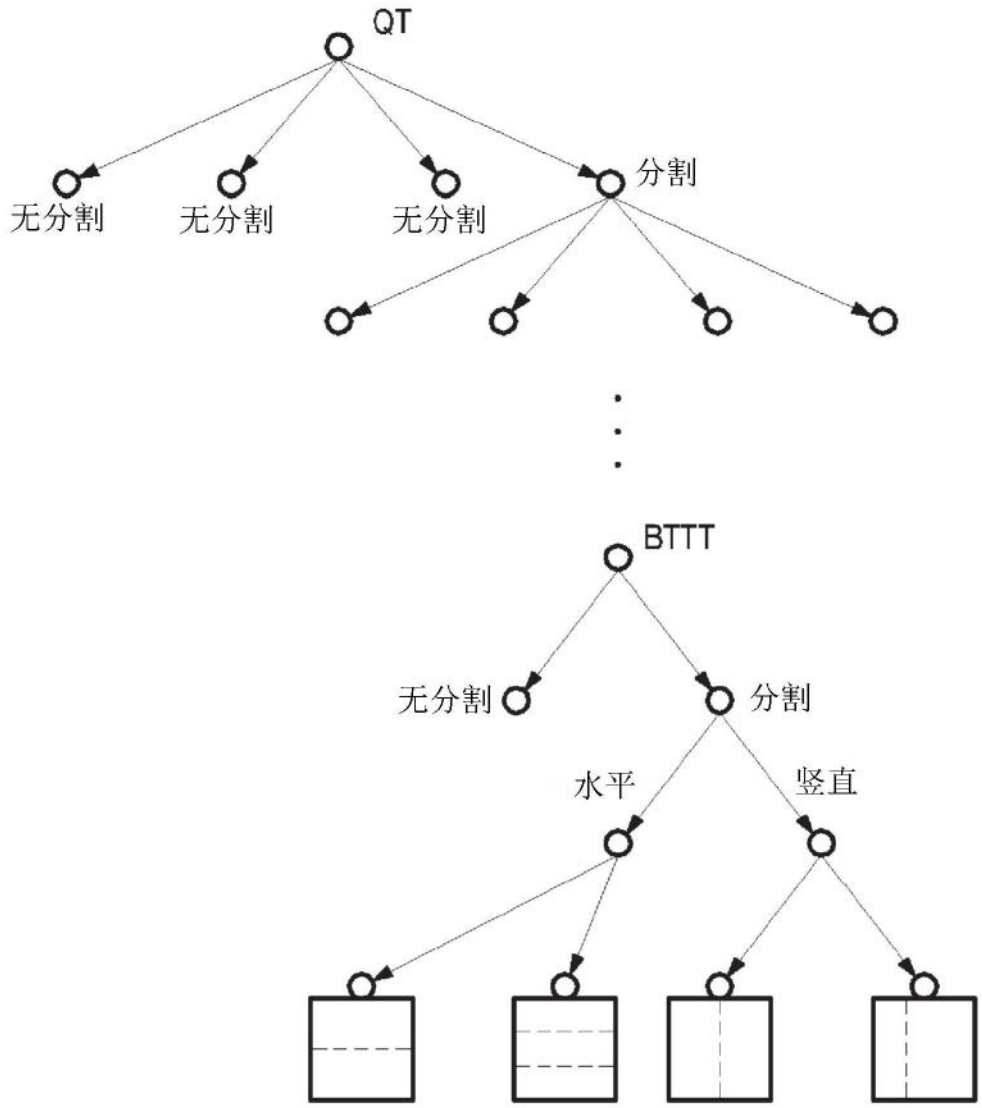


图2

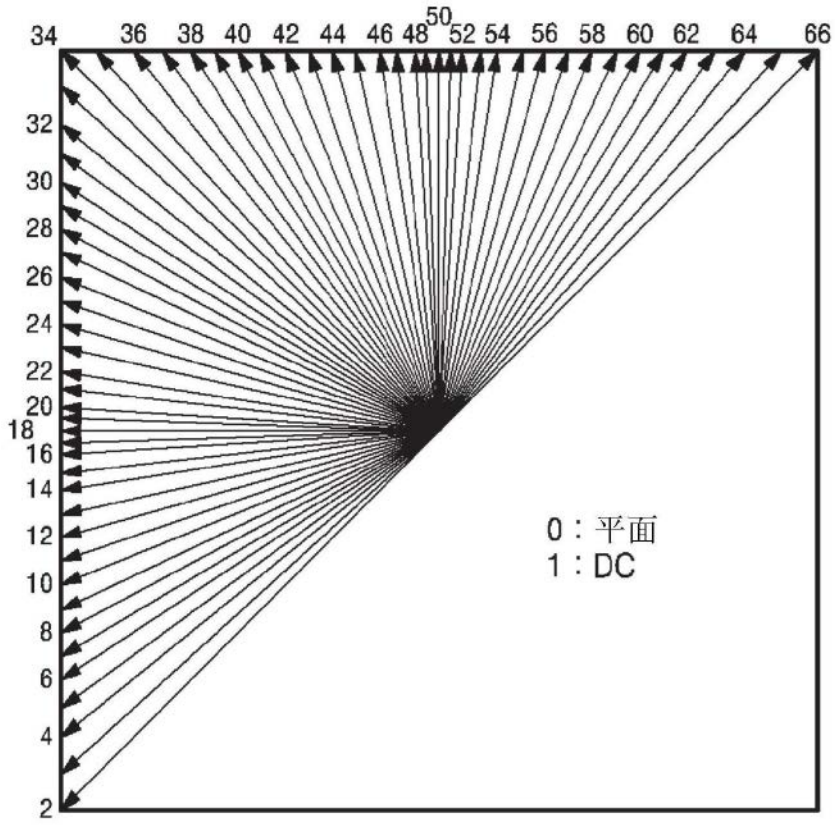

34
36
38
40
42
44
46
48
50
52
54
56
58
60
62
64
66
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0：平面
1：DC

图3a

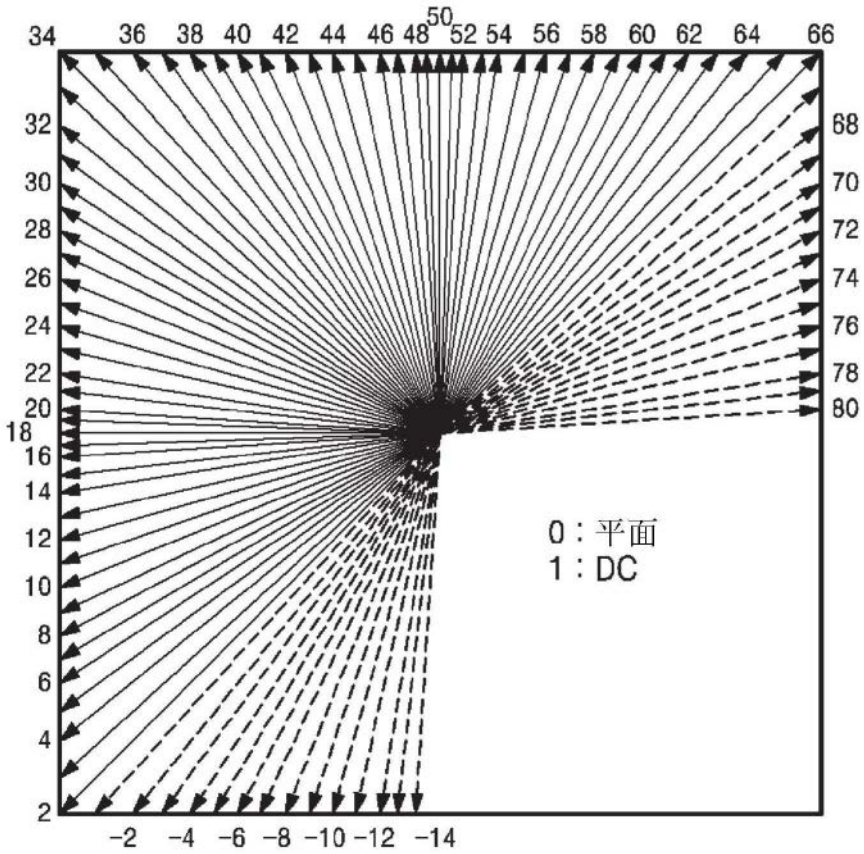

34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66
32 30 28 26 24 22 20 18 16 14 12 10 8 6 4 2
68 70 72 74 76 78 80
-2 -4 -6 -8 -10 -12 -14
0：平面
1：DC

图3b

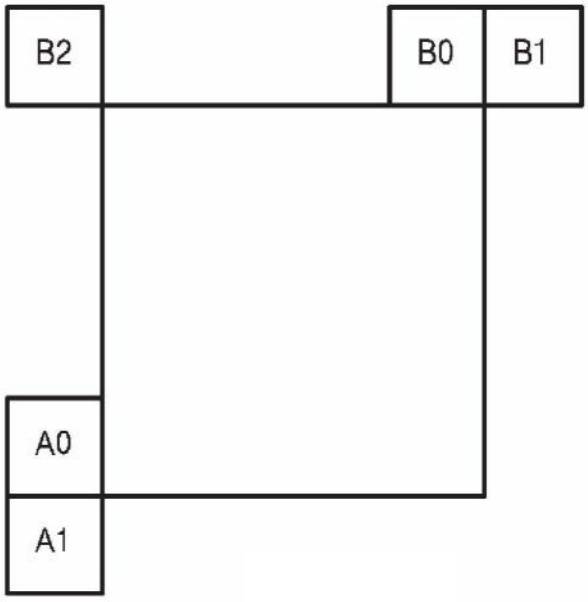

B2
B0
B1
A0
A1

图4

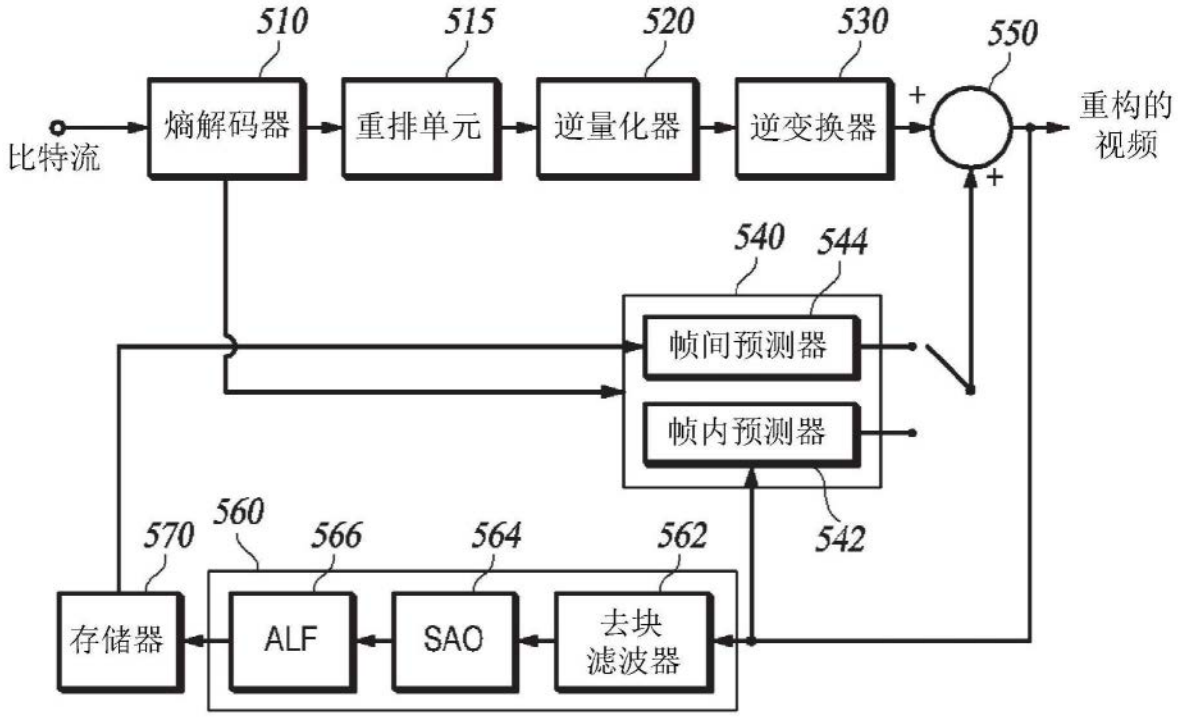

510
515
520
530
550
比特流
熵解码器
重排单元
逆量化器
逆变换器
+
+
重构的
视频
540
544
帧间预测器
帧内预测器
542
570
560
566
564
562
存储器
ALF
SAO
去块
滤波器

图5

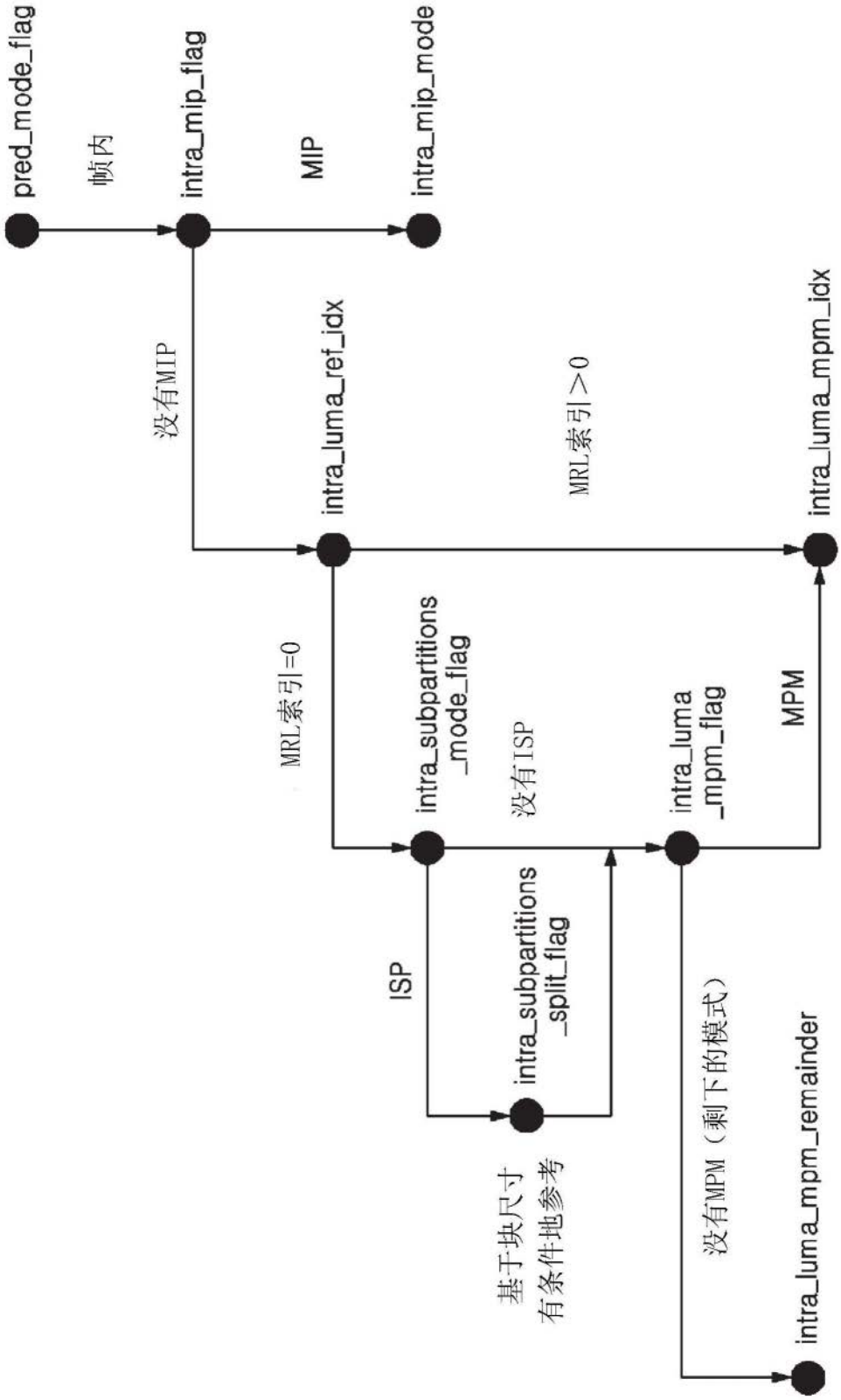

pred_mode_flag
帧内
intra_mip_flag
MIP
intra_mip_mode
没有MIP
intra_luma_ref_idx
MRL索引>0
intra_luma_mpm_idx
MRL索引=0
intra_subpartitions
_mode_flag
没有ISP
intra_luma
_mpm_flag
MPM
ISP
intra_subpartitions
_split_flag
基于块尺寸
有条件地参考
没有MPM（剩下的模式）
intra_luma_mpm_remainder

图6

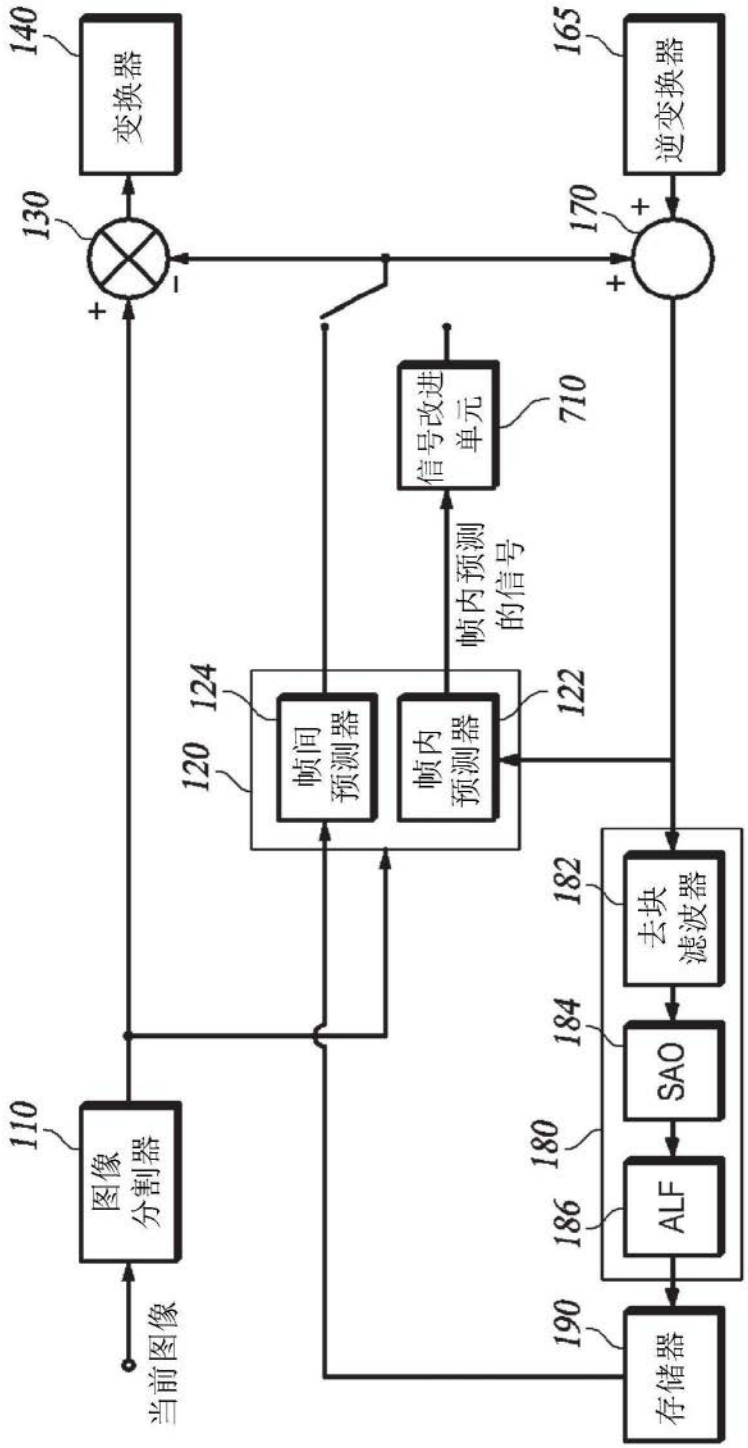

140
变换器
165
逆变换器
130
170
+
-
+
+
信号改进单元
710
帧内预测的信号
124
120
帧间预测器
帧内预测器
122
182
去块滤波器
184
SAO
180
ALF
186
110
图像分割器
当前图像
190
存储器

图7

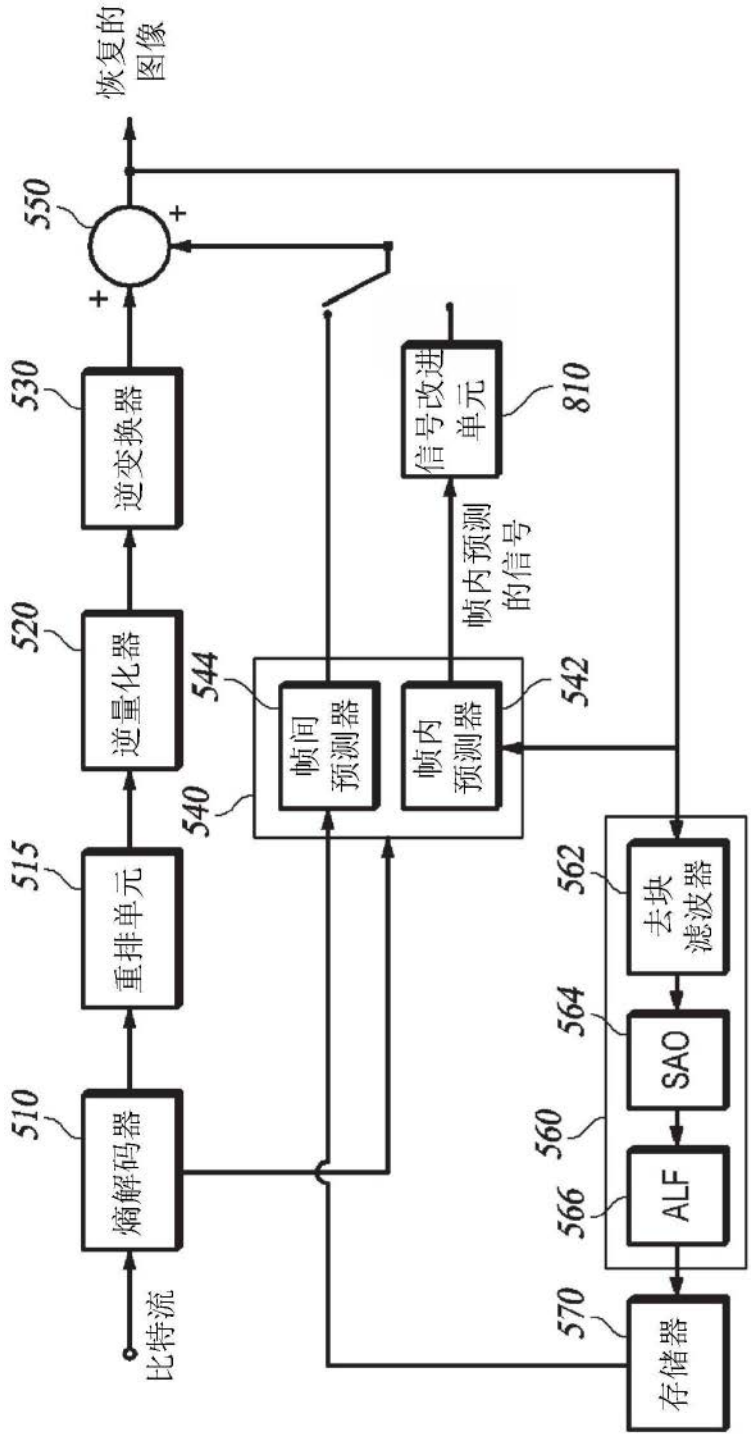

恢复的图像
550
+
+
530
逆变换器
520
逆量化器
515
重排单元
510
熵解码器
比特流
544
帧间预测器
542
帧内预测器
540
帧内预测的信号
信号改进单元
810
562
去块滤波器
564
SAO
560
566
ALF
570
存储器

图8

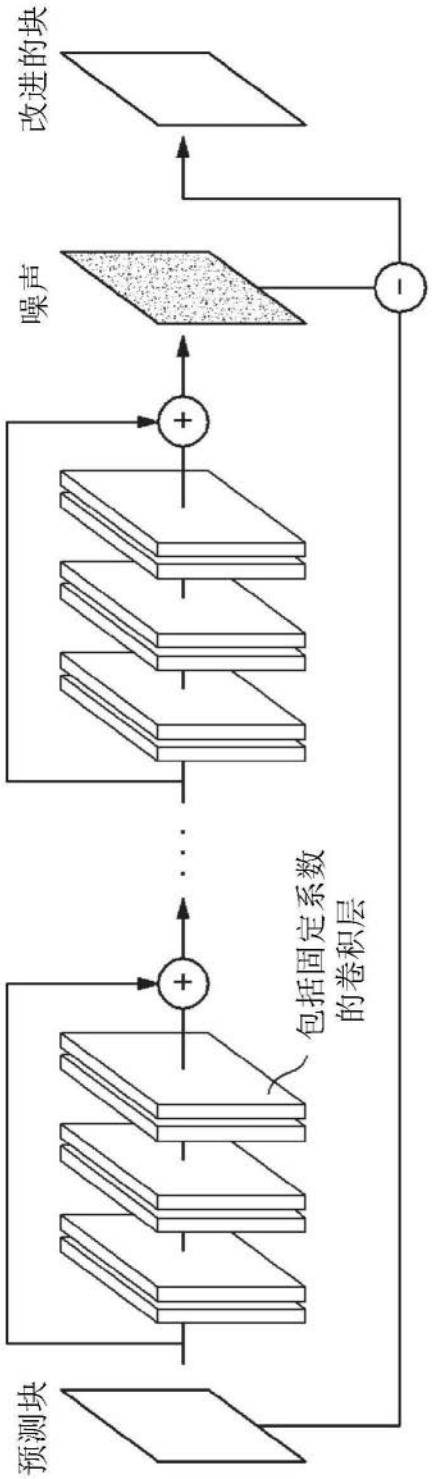


图9

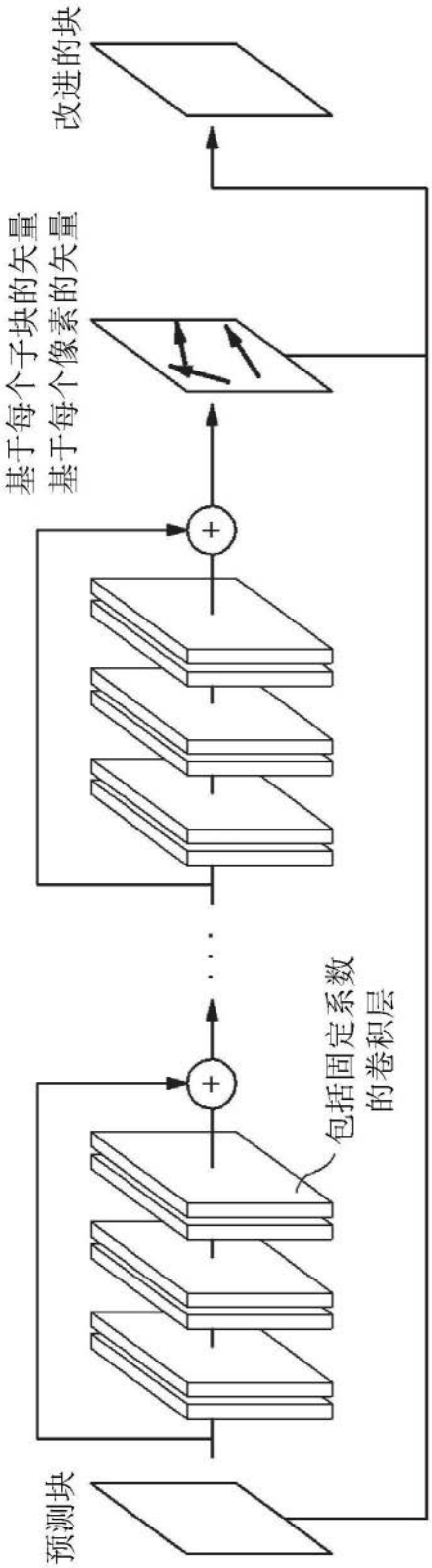


图10

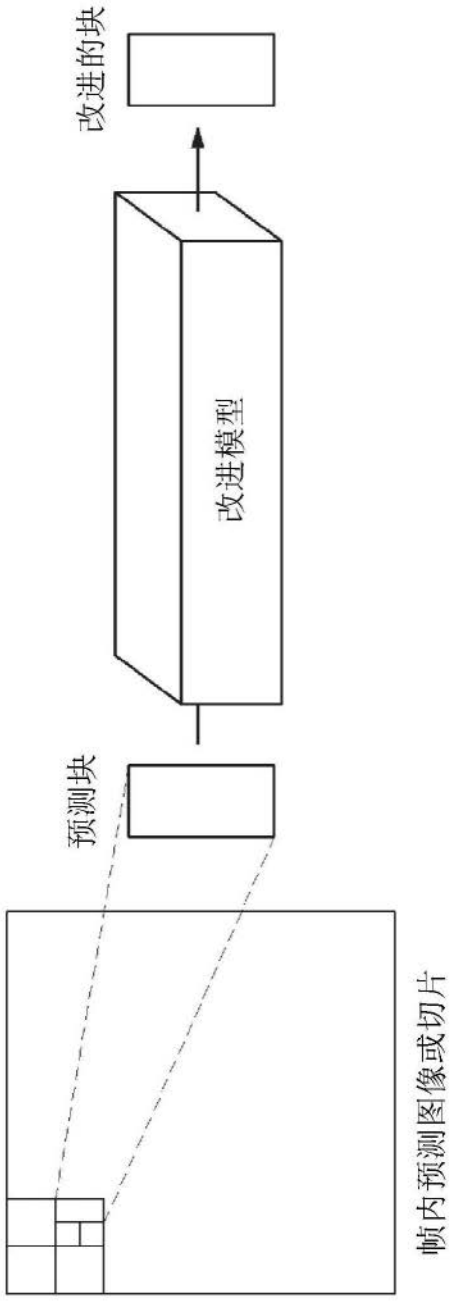


图11

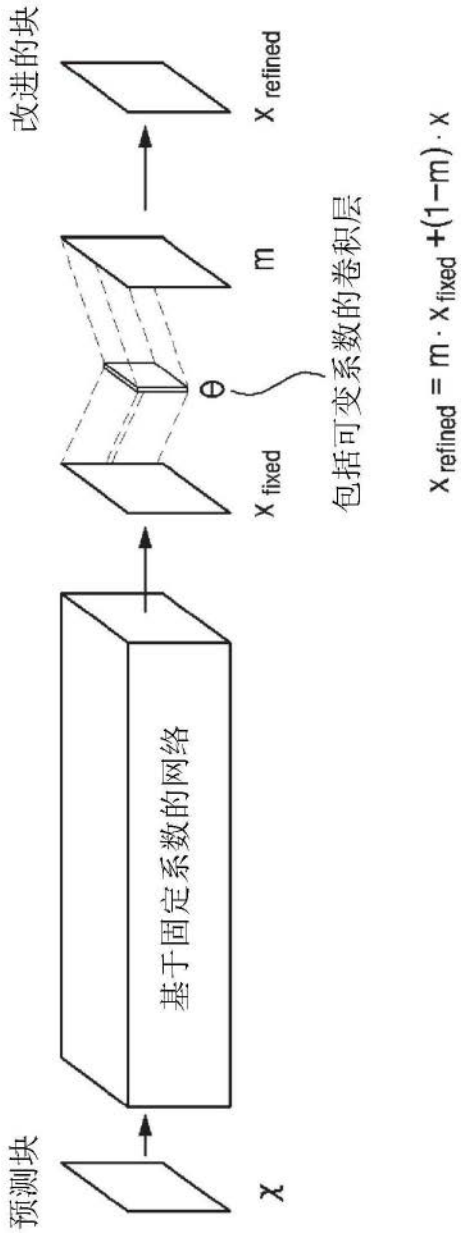

预测块
x
基于固定系数的网络
x fixed
θ
m
x refined
改进的块
包括可变系数的卷积层
x refined = m · x fixed +(1−m) · x

图12

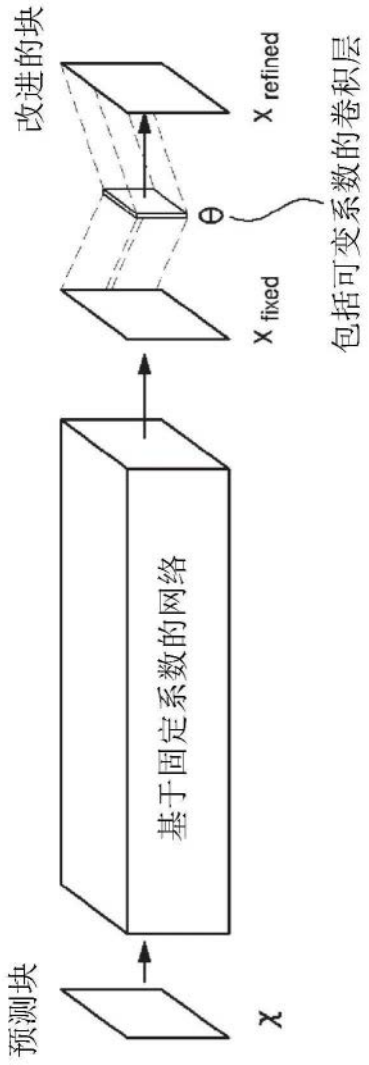


图13

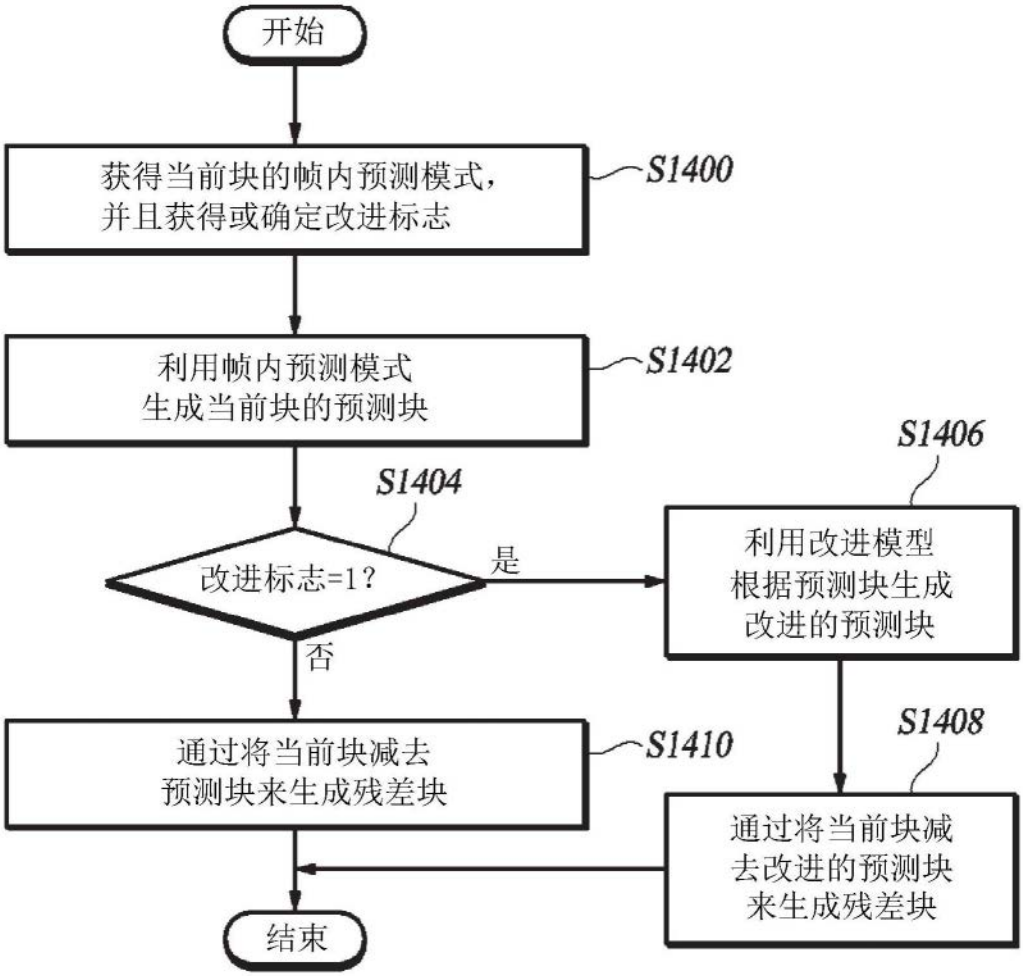

开始
获得当前块的帧内预测模式，
并且获得或确定改进标志
S1400
利用帧内预测模式
生成当前块的预测块
S1402
S1404
改进标志=1？
是
否
S1406
利用改进模型
根据预测块生成
改进的预测块
S1408
通过将当前块减
去改进的预测块
来生成残差块
S1410
通过将当前块减去
预测块来生成残差块
结束

图14

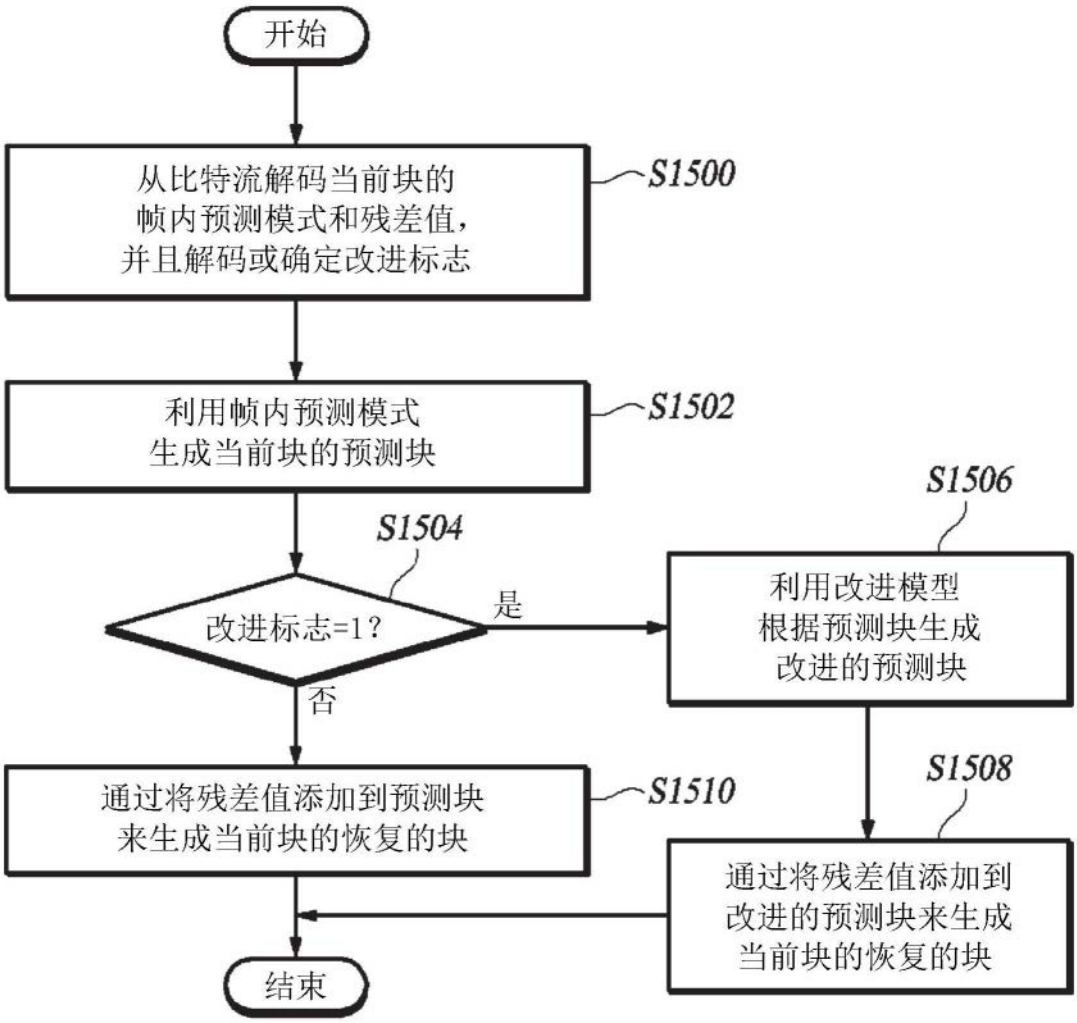

开始
从比特流解码当前块的
帧内预测模式和残差值，
并且解码或确定改进标志
S1500
利用帧内预测模式
生成当前块的预测块
S1502
S1504
改进标志=1？
是
否
S1506
利用改进模型
根据预测块生成
改进的预测块
S1508
通过将残差值添加到
改进的预测块来生成
当前块的恢复的块
通过将残差值添加到预测块
来生成当前块的恢复的块
S1510
结束

图15