(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116194957 A
(43)申请公布日 2023.05.30

(21)申请号 202180064055.7
(22)申请日 2021.09.28
(30)优先权数据
10-2020-0126111 2020.09.28 KR
10-2021-0127795 2021.09.28 KR
(85)PCT国际申请进入国家阶段日
2023.03.17
(86)PCT国际申请的申请数据
PCT/KR2021/013218 2021.09.28
(87)PCT国际申请的公布数据
WO2022/065977 KO 2022.03.31
(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 金挪莹 李订炅 朴胜煜 林和平
(74)专利代理机构 北京路浩知识产权代理有限公司 11002
专利代理师 张晶 崔龙铉
(51)Int.Cl.
*G06T 9/00*(2006.01)

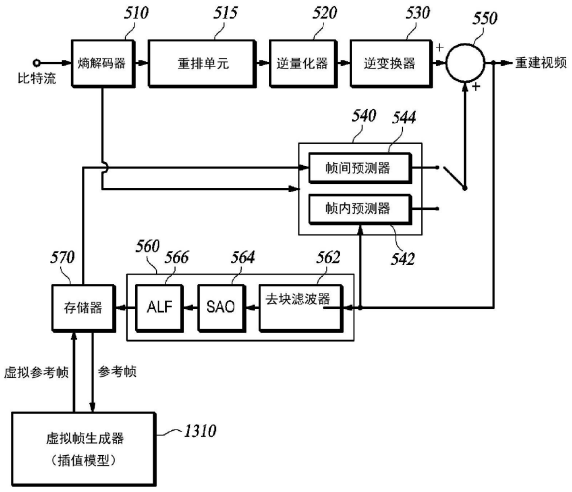
权利要求书2页 说明书21页 附图15页

(54)发明名称

基于可变系数深度学习的帧间预测方法

(57)摘要

本实施例提供一种帧间预测方法，该方法：使可变系数深度学习模型自适应地学习视频的特征；将学习生成的可变系数深度学习模型参数从视频编码装置发送到视频解码装置；以及参考由可变系数深度学习模型生成的虚拟参考帧。




CN 116194957 A

1.一种由视频解码装置执行的视频解码方法,包括:

从比特流中解码可变系数值、仿射预测标志和编码模式,其中,所述仿射预测标志指示是否对当前块应用仿射运动预测,所述编码模式是所述当前块的运动信息的编码模式并指示合并模式或高级运动矢量预测(AMVP)模式;

通过利用插值模型,基于参考帧生成虚拟参考帧,其中,通过所述可变系数值设置包括在所述插值模型中的可变系数网络;以及

当所述编码模式为所述合并模式时,基于所述虚拟参考帧和所述参考帧生成所述当前块的合并候选。

2.根据权利要求1所述的视频解码方法,其中,所述插值模型是下列模型中的一种:

深度学习模型,从所述参考帧直接生成所述虚拟参考帧;

深度学习模型,从所述参考帧生成掩码,然后基于所述掩码对所述参考帧执行加权和以生成所述虚拟参考帧;以及

深度学习模型,从所述参考帧生成光流,然后基于所述光流扭转所述参考帧以生成所述虚拟参考帧。

3.根据权利要求1所述的视频解码方法,其中,所述视频编码装置利用基于由所述插值模型生成的推断帧和原始帧之间的差值的损失函数来更新所述插值模型的所述可变系数网络以生成所述可变系数值。

4.根据权利要求1所述的视频解码方法,其中,所述插值模型进一步包括固定系数网络,并且所述固定系数网络的固定系数值经过基于整个原始参考帧的预训练生成并根据预定的协议设置。

5.根据权利要求4所述的视频解码方法,其中,所述插值模型的所述可变系数网络连接到所述固定系数网络的后端。

6.根据权利要求1所述的视频解码方法,其中,按帧单位或至少一个图片组(GOP)单位对所述可变系数值进行解码。

7.根据权利要求1所述的视频解码方法,其中,生成所述合并候选包括:

搜索空间合并候选;以及

在考虑所述虚拟参考帧的情况下搜索时间合并候选。

8.根据权利要求7所述的视频解码方法,其中,当所述时间合并候选的参考帧和存在与所述当前块相同位置的块的图片都是虚拟参考帧时,将零运动矢量设置为所述合并候选。

9.根据权利要求1所述的视频解码方法,进一步包括:

当所述仿射预测标志为真并且所述编码模式为所述AMVP模式时,生成仿射AMVP候选,

其中,生成所述仿射AMVP候选包括:

在考虑所述虚拟参考帧的情况下添加继承的仿射AMVP候选;以及

在考虑所述虚拟参考帧的情况下添加构建的仿射AMVP候选。

10.根据权利要求9所述的视频解码方法,其中,添加所述继承的仿射AMVP候选包括,当所述继承的仿射AMVP候选的参考图片和所述当前块的参考图片为相同的参考图片时并且当所述相同的参考图片为所述虚拟参考帧时,添加零运动矢量作为所述仿射AMVP候选。

11.根据权利要求9所述的视频解码方法,其中,添加所述构建的仿射AMVP候选包括,当所述构建的仿射AMVP候选的所有控制点运动矢量的参考图片和所述当前块的参考图片为

相同的参考图片时并且当所述相同的参考图片为所述虚拟参考帧时,添加零运动矢量作为参考图片。

12.一种由视频编码装置执行的视频编码方法,包括:

获取先前生成的可变系数值和预设编码模式,其中,所述编码模式是当前块的运动信息的编码模式并指示合并模式或高级运动矢量预测(AMVP)模式;

通过利用插值模型,从参考帧生成虚拟参考帧,其中,通过所述可变系数值设置包括在所述插值模型中的可变系数网络;以及

当所述编码模式为所述合并模式时,基于所述虚拟参考帧和所述参考帧生成所述当前块的合并候选。

13.根据权利要求12所述的视频编码方法,其中,利用基于由所述插值模型生成的推断帧和原始帧之间的差值的损失函数来更新所述插值模型的所述可变系数网络以生成所述可变系数值。

14.根据权利要求12所述的视频编码方法,其中,所述插值模型进一步包括固定系数网络,并且所述固定系数网络的固定系数值经过基于整个原始参考帧的预训练生成并根据预定的协议设置。

15.根据权利要求12所述的视频编码方法,其中,按帧单位或至少一个图片组(GOP)单位对所述可变系数值进行编码。

16.一种视频解码装置,包括:

熵解码器,从比特流中解码可变系数值和编码模式,其中,所述编码模式是当前块的运动信息的编码模式并指示合并模式或高级运动矢量预测(AMVP)模式;

虚拟帧生成器,利用插值模型,从参考帧生成虚拟参考帧,其中,通过所述可变系数值设置包括在所述插值模型中的可变系数网络;以及

帧间预测器,当所述编码模式为所述合并模式时,基于所述虚拟参考帧和所述参考帧生成所述当前块的合并候选。

# 基于可变系数深度学习的帧间预测方法

## 技术领域

[0001] 本公开涉及一种包括在视频编码或解码装置中的基于可变系数深度学习的帧间预测方法。

## 背景技术

[0002] 下面的描述仅提供与本公开相关的背景信息,不构成现有技术。

[0003] 由于视频数据与音频数据或静止图像数据相比具有大的数据量,因此需要包括存储器的大量硬件资源,以存储或发送视频数据而不进行压缩处理。

[0004] 因此,通常在存储或发送视频数据时使用编码器来压缩视频数据并存储或发送,并且解码器接收压缩后的视频数据,对接收的压缩视频数据进行解压,并播放解压缩后的视频数据。视频压缩技术包括H.264/AVC、高效视频编码(HEVC)和多功能视频编码(VVC),与HEVC相比,VVC的编码效率提高了约30%以上。

[0005] 然而,由于图像大小、分辨率和帧速率逐渐增加,因此需要编码的数据量也增加。因此,需要一种提供比现有压缩技术更高的编码效率和改进的图像增强效果的新压缩技术。

[0006] 最近,一种基于深度学习的视频处理技术被应用于现有的编码元件技术。基于深度学习的视频处理技术被应用于现有编码技术中的诸如帧间预测(inter prediction)、帧内预测(intra prediction)、环内滤波或变换等压缩技术。因此可以提高编码效率。代表性的应用示例包括基于根据深度学习模型生成的虚拟参考帧的帧间预测,以及基于图像修复模型的环内滤波(见非专利文献1)。因此,在视频编码或解码中,有必要考虑持续应用基于深度学习的视频处理技术,以提高编码效率。

[0007] [现有技术]

[0008] (非专利文献)

[0009] 非专利文献1:Lee等人,HEVC中基于深度视频预测网络的帧间编码(Deep Video Prediction Network Based Inter-Frame Coding in HEVC),IEEE Access 2020。

## 发明内容

[0010] [技术问题]

[0011] 本公开的目的是提供一种帧间预测方法,用于使可变系数深度学习模型自适应地学习视频的特征,将学习生成的可变系数深度学习模型的参数从视频编码装置发送到视频解码装置,以及参考由可变系数深度学习模型生成的虚拟参考帧。

[0012] [技术方案]

[0013] 根据本公开的实施例,一种由视频解码装置执行的视频解码方法包括:从比特流中解码可变系数值、仿射预测标志和编码模式,所述仿射预测标志指示是否对当前块应用仿射运动预测,所述编码模式是所述当前块的运动信息的编码模式并指示合并模式或高级运动矢量预测(AMVP)模式;通过利用插值模型,基于参考帧生成虚拟参考帧,其中,通过所

述可变系数值设置包括在所述插值模型中的可变系数网络;以及当所述编码模式为所述合并模式时,基于所述虚拟参考帧和所述参考帧生成所述当前块的合并候选。

[0014] 根据本公开的另一实施例,一种由视频编码装置执行的视频编码方法包括:获取先前生成的可变系数值和预设编码模式,其中,所述编码模式是当前块的运动信息的编码模式并指示合并模式或高级运动矢量预测(AMVP)模式;通过利用插值模型,从参考帧生成虚拟参考帧,其中,通过所述可变系数值设置包括在所述插值模型中的可变系数网络;以及当所述编码模式为所述合并模式时,基于所述虚拟参考帧和所述参考帧生成所述当前块的合并候选。

[0015] 根据本公开的另一实施例,一种视频解码装置包括:熵解码器,从比特流中解码可变系数值和编码模式,其中,所述编码模式是当前块的运动信息的编码模式并指示合并模式或高级运动矢量预测(AMVP)模式;虚拟帧生成器,利用插值模型,从参考帧生成虚拟参考帧,其中,通过所述可变系数值设置包括在所述插值模型中的可变系数网络;以及帧间预测器,当所述编码模式为所述合并模式时,基于所述虚拟参考帧和所述参考帧生成所述当前块的合并候选。

[0016] [有益效果]

[0017] 如上所述,根据本实施例,通过提供参考由可变系数深度学习模型生成的虚拟参考帧的帧间预测方法,可以增加帧间预测的效果并提高编码效率。

## 附图说明

[0018] 图1是可以实现本公开技术的视频编码装置的框图。

[0019] 图2是用于说明利用四叉树加二叉树三叉树(QTBTTT)结构划分块的方法的示图。

[0020] 图3A和图3B是示出包括广角帧内预测(intra prediction)模式的多种帧内预测模式的示图。

[0021] 图4是示出当前块的周围块的示图。

[0022] 图5是可以实现本公开技术的视频解码装置的框图。

[0023] 图6是示出根据本公开的实施例的双向OF(BDOF)的概念的示图。

[0024] 图7是示出根据本公开的实施例的在合并/跳过模式中搜索运动矢量候选的过程的流程图。

[0025] 图8是示出根据本公开的实施例的仿射运动预测的示图。

[0026] 图9是示出根据本公开的实施例的以子块为单位的仿射运动预测的示图。

[0027] 图10是示出从周围块的平移运动矢量得出用于仿射运动预测的合并候选的方法的示图。

[0028] 图11是示出根据本公开的实施例的在仿射高级运动矢量预测(AMVP)模式中搜索AMVP候选的过程的流程图。

[0029] 图12是示出根据本公开的实施例的包括插值模型的视频编码装置的框图。

[0030] 图13是示出根据本公开的实施例的包括插值模型的视频解码装置的框图。

[0031] 图14是示出根据本公开的实施例的视频编码装置和视频解码装置中使用的插值模型的操作的示图。

[0032] 图15示出根据本公开的实施例的可变系数插值模型的示例。

[0033] 图16是示出根据本公开的实施例的进一步包括固定系数网络的插值模型的示图。

[0034] 图17是示出根据本公开的实施例的发送可变系数的方法的示图。

[0035] 图18是示出根据本公开的实施例的在帧间预测的合并/跳过模式中使用虚拟参考帧的过程的流程图。

[0036] 图19是示出根据本公开的实施例的在帧间预测的仿射AMVP模式中使用虚拟参考帧的过程的流程图。

## 具体实施方式

[0037] 在下文中,参照附图详细描述本公开的实施例。当附图标记指代各图中的组件时,应当注意的是,尽管在不同的图中示出了相同或等同的组件,但相同或等同的组件可以用相同的附图标记表示。此外,在描述实施例时,省略了对已知相关配置和功能的详细描述,以避免不必要地模糊实施例的主旨。

[0038] 图1是可以实现本公开技术的视频编码装置的框图。在下文中,参照图1的图示,描述了视频编码装置和该装置的子组件。

[0039] 视频编码装置可以包括图片分割器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、循环滤波器单元180和存储器190。

[0040] 视频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。此外,每个组件的功能可以实现为软件,并且微处理器还可以实现为执行与每个组件对应的软件的功能。

[0041] 一个视频由包括多个图片的一个或多个序列构成。每个图片被分割成多个区域,并针对每个区域执行编码。例如,一个图片被分割成一个或多个图块(tile)或/和切片(slice)。此处,一个或多个图块可以被定义为图块组。每个图块或/和切片被分割成一个或多个编码树单元(CTU)。另外,每个CTU通过树结构被分割成一个或多个编码单元(CU)。应用于每个CU的信息被编码为CU的句法(syntax),共同应用于一个CTU中包括的CU的信息被编码为CTU的句法。此外,共同应用于一个切片中的所有块的信息被编码为切片报头(slice header)的句法,应用于构成一个或多个图片的所有块的信息被编码为图片参数集(PPS)或图片报头(picture header)。此外,多个图片共同所指的信息被编码到序列参数集(SPS)中。另外,一个或多个SPS共同涉及的信息被编码为视频参数集(VPS)。此外,共同应用于一个图块或图块组的信息也可以被编码为图块或图块组报头的句法。包括在SPS、PPS、切片报头、图块或图块组报头中的句法可以被称为高级句法。

[0042] 图片分割器110确定编码树单元(CTU)的大小。关于CTU的大小的信息(CTU大小)被编码为SPS或PPS的句法并被发送到视频解码装置。

[0043] 图片分割器110将构成视频的每个图片分割成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(leaf node)成为作为编码的基本单位的编码单元(CU)。

[0044] 树结构可以是四叉树(quadtree,QT)结构,其中上级节点(或父节点)被分割成具有相同大小的四个下级节点(或子节点)。树结构也可以是二叉树(BinaryTree,BT)结构,其中上级节点被分割成两个下级节点。树结构也可以是三叉树(TernaryTree,TT)结构,其中

上级节点以1:2:1的比例被分割成三个下级节点。树结构也可以是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(QTBT)结构,或者可以利用四叉树加二叉树三叉树(QTBTTT)结构。此处,BTTT被添加到树结构中,以被称为多类型树(MTT)。

[0045] 图2是用于描述通过利用QTBTTT结构分割块的方法的示图。

[0046] 如图2所示,CTU可以首先分割成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)。指示QT结构的每个节点是否被分割成下层的四个节点的第一标志(QT_split_flag)由熵编码器155编码并用信号发送到视频解码装置。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,叶节点可被进一步分割成BT结构或TT结构中的至少一个。在BT结构和/或TT结构中可以存在多个分割方向。例如,可以有两个方向,即相应节点的块被水平分割的方向和相应节点的块被垂直分割的方向。如图2所示,当MTT分割开始时,指示节点是否被分割的第二标志(mtt_split_flag)以及如果节点被分割则进一步指示分割方向(垂直或水平)的标志和/或指示分割类型(二元或三元)的标志由熵编码器155编码并用信号发送到视频解码装置。

[0047] 可选地,在对指示每个节点是否被分割成下层的四个节点的第一标志(QT_split_flag)进行编码之前,也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)的值指示每个节点不被分割时,相应节点的块成为分割树结构中的叶节点,并成为作为编码的基本单位的编码单元(CU)。当CU分割标志(split_cu_flag)的值指示每个节点被分割时,视频编码装置通过上述方式开始对第一标志进行编码。

[0048] 当QTBT被用作树结构的另一示例时,可以有两种类型,即,相应节点的块被水平分割成具有相同大小的两个块的类型(即,对称水平分割)以及相应节点的块被垂直分割成具有相同大小的两个块的类型(即,对称垂直分割)。指示BT结构的每个节点是否被分割成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由熵编码器155编码并发送到视频解码装置。另一方面,可以另外存在相应节点的块被分割成两个相互不对称的形式的块的类型。不对称的形式可以包括相应节点的块分割成具有1:3的大小比的两个矩形块的形式,或者还包括相应节点的块沿对角线方向被分割的形式。

[0049] 根据从CTU分割的QTBT或QTBTTT,CU可以具有各种大小。在下文中,对应于待编码或解码的CU(即,QTBTTT的叶节点)的块被称为“当前块”。由于采用了QTBTTT分割,因此当前块的形状除了正方形之外,还可以是矩形形状。

[0050] 预测器120预测当前块以生成预测块。预测器120包括帧内预测器122和帧间预测器(inter-predictor)124。

[0051] 通常,图片中的每个当前块可以被预测性编码。通常,当前块的预测可以通过利用帧内预测技术(利用来自包括当前块的图片的数据)或帧间预测技术(利用来自在包括当前块的图片之前编码的图片的数据)执行。帧间预测包括单向预测和双向预测。

[0052] 帧内预测器122通过利用位于在包括当前块的当前图片中的当前块的周围的像素(参考像素)来预测当前块中的像素。根据预测方向,有多种帧内预测模式。例如,如图3A所示,多种帧内预测模式可以包括具有平面模式和DC模式的2种非定向模式,并且可以包括65种定向模式。根据每个预测模式,待使用的周围像素和计算式被不同地定义。

[0053] 为了对具有矩形形状的当前块进行有效的定向预测,可以另外使用图3B中以虚线箭头表示的定向模式(帧内预测模式#67至#80,#-1至#14)。这些模式可被称为“广角帧内预测模式”。在图3B中,箭头指示用于预测的相应参考样本,不表示预测方向。预测方向与箭头所示的方向相反。在当前块具有矩形形状时,广角帧内预测模式是指在与特定定向模式相反的方向上进行预测而不需要附加的比特发送的模式。在这种情况下,在广角帧内预测模式中,可用于当前块的一些广角帧内预测模式可以由具有矩形形状的当前块的宽度和高度的比率确定。例如,在当前块具有高度小于宽度的矩形形状时,可使用角度小于45度的广角帧内预测模式(帧内预测模式#67至#80)。在当前块具有宽度大于高度的矩形形状时,可使用角度大于-135度的广角帧内预测模式(帧内预测模式#-1至#14)。

[0054] 帧内预测器122可以确定将用于对当前块进行编码的帧内预测模式。在一些示例中,帧内预测器122可以通过利用多种帧内预测模式来编码当前块,并且还从测试的模式中选择待使用的适当的帧内预测模式。例如,帧内预测器122可以通过利用针对多种测试的帧内预测模式的速率失真(rate-distortion)分析来计算速率失真值,并且还在测试的模式中选择具有最佳速率失真特征的帧内预测模式。

[0055] 帧内预测器122在多种帧内预测模式中选择一种帧内预测模式,并通过利用根据选择的帧内预测模式确定的周围像素(参考像素)和计算式来预测当前块。关于选择的帧内预测模式的信息由熵编码器155编码并发送到视频解码装置。

[0056] 帧间预测器124通过利用运动补偿过程生成当前块的预测块。帧间预测器124在比当前图片更早编码和解码的参考图片中搜索与当前块最相似的块,并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),该运动矢量对应于当前图片中的当前块和参考图片中的预测块之间的位移。通常,运动估计是针对亮度(luma)分量执行的,而基于亮度分量计算的运动矢量则用于亮度分量和色度分量两者。包括关于参考图片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由熵编码器155编码并发送到视频解码装置。

[0057] 帧间预测器124还可以对参考图片或参考块执行插值,以增加预测的效果。换言之,通过将滤波器系数应用于包括两个整数样本的多个连续整数样本,对两个连续整数样本之间的子样本进行插值。当针对插值参考图片执行搜索与当前块最相似的块的过程时,运动矢量可以表示到小数单位精度而不是整数样本单位精度。运动矢量的精度或分辨率可以针对每个待编码的目标区域,例如诸如切片、图块、CTU、CU等单位而不同地设置。当应用这种自适应运动矢量分辨率(AMVR)时,应该为每个目标区域用信号发送关于待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,用信号发送关于应用于每个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表示将在下面描述的运动矢量差的精度的信息。

[0058] 另一方面,帧间预测器124可以通过利用双向预测(bi-prediction)执行帧间预测。在双向预测的情况下,利用两个参考图片和表示每张参考图片中与当前块最相似的块位置的两个运动矢量。帧间预测器124分别从参考图片列表0(RefPicList0)和参考图片列表1(RefPicList1)中选择第一参考图片和第二参考图片。帧间预测器124还在各个参考图片中搜索与当前块最相似的块,以生成第一参考块和第二参考块。另外,当前块的预测块是通过对第一参考块和第二参考块进行平均或加权平均而生成的。另外,包括关于用于预测

当前块的两个参考图片的信息和关于两个运动矢量的信息的运动信息被发送到熵编码器155。此处，参考图片列表0可以由预重建图片中的显示顺序在当前图片之前的图片构成，而参考图片列表1可以由预重建图片中的显示顺序在当前图片之后的图片构成。但是，不一定局限于此，显示顺序在当前图片之后的预重建图片可以附加地包括在参考图片列表0中。反之，在当前图片之前的预重建图片也可以附加地包括在参考图片列表1中。

[0059] 为了使用于对运动信息进行编码的比特量最小化，可以利用各种方法。

[0060] 例如，在当前块的参考图片和运动矢量与周围块的参考图片和运动矢量相同时，能够识别周围块的信息被编码以将当前块的运动信息发送到视频解码装置。这样的方法被称为合并模式(merge mode)。

[0061] 在合并模式中，帧间预测器124从当前块的周围块中选择预定数量的合并候选块(以下，称为“合并候选”)。

[0062] 作为用于得出合并候选的周围块，如图4所示，可以使用在当前图片中与当前块相邻的左块A0、底左块A1、顶块B0、顶右块B1和顶左块B2的全部或部分。此外，位于参考图片(可以与用于预测当前块的参考图片相同或不同)而不是当前块所在的当前图片中的块也可以用作合并候选。例如，参考图片中与当前块相同位置的块或与该相同位置的块相邻的块可以附加地用作合并候选。如果通过上述方法选择的合并候选的数量小于预设的数量，则向合并候选添加零矢量。

[0063] 帧间预测器124通过利用周围块配置包括预定数量的合并候选的合并列表。从包括在合并列表中的合并候选中选择待用作当前块的运动信息的合并候选，并且生成用于识别所选择的候选的合并索引信息。所生成的合并索引信息由熵编码器155编码并被发送到视频解码装置。

[0064] 合并跳过(merge skip)模式是合并模式的一个特例。在量化之后，当用于熵编码的所有变换系数接近于零时，仅发送周围块选择信息而不发送残差信号。通过利用合并跳过模式，可以对于有轻微运动的图像、静止图像、屏幕内容图像等实现相对较高的编码效率。

[0065] 在下文中，合并模式和合并跳过模式被统称为合并/跳过模式。

[0066] 对运动信息进行编码的另一种方法是高级运动矢量预测(AMVP)模式。

[0067] 在AMVP模式中，帧间预测器124通过利用当前块的周围块来得出当前块的运动矢量的预测运动矢量候选。作为用于得出预测运动矢量候选的周围块，可以利用图4中示出的当前图片中与当前块相邻的左块A0、顶块B0、顶右块B1、底左块A1和顶左块B2的全部或部分。此外，位于参考图片(可以与用于预测当前块的参考图片相同或不同)而不是当前块所在的当前图片中的块也可以用作用于得出预测运动矢量候选的周围块。例如，可以利用参考图片中与当前块相同位置的块或与该相同位置的块相邻的块。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，则向运动矢量候选添加零矢量。

[0068] 帧间预测器124通过利用周围块的运动矢量得出预测运动矢量候选，并且通过使用预测运动矢量候选确定当前块的运动矢量的预测运动矢量。另外，通过从当前块的运动矢量中减去预测运动矢量来计算运动矢量差。

[0069] 预测运动矢量可以通过对预测运动矢量候选应用预定义的函数(例如，中心值和平均值计算等)来获得。在这种情况下，视频解码装置也知道预定义的函数。此外，由于用于

得出预测运动矢量候选的周围块是已经完成编码和解码的块,因此视频解码装置也可能已经知道周围块的运动矢量。因此,视频编码装置不需要对用于识别预测运动矢量候选的信息进行编码。因此,在这种情况下,关于运动矢量差的信息和关于用于预测当前块的参考图片的信息被编码。

[0070] 另一方面,预测运动矢量也可以通过选择预测运动矢量候选中的任何一个的方式来确定。在这种情况下,用于识别所选择的预测运动矢量候选的信息与关于运动矢量差的信息和关于用于预测当前块的参考图片的信息共同被附加编码。

[0071] 减法器130通过从当前块中减去由帧内预测器122或帧间预测器124生成的预测块来生成残差块。

[0072] 变换器140将残差块中具有空间域的像素值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总大小作为变换单位来变换残差块中的残差信号,或者也可以将残差块分割成多个子块并通过利用子块作为变换单位来执行变换。可选地,将残差块划分为作为变换区域和非变换区域的两个子块,以通过仅利用变换区域子块作为变换单位来变换残差信号。此处,变换区域子块可以是具有基于水平轴(或垂直轴)的大小比为1:1的两个矩形块中的一个。在这种情况下,指示只有子块被变换的标志(cu_sbt_flag)、定向(垂直/水平)信息(cu_sbt_horizontal_flag)和/或位置信息(cu_sbt_pos_flag)由熵编码器155编码并被用信号发送到视频解码装置。此外,变换区域子块的大小可以具有基于水平轴(或垂直轴)的1:3的大小比,在这种情况下,划分相应分割的标志(cu_sbt_quad_flag)另外由熵编码器155编码并被用信号发送到视频解码装置。

[0073] 另一方面,变换器140可以在水平方向和垂直方向上单独执行对残差块的变换。对于变换,可以使用各种类型的变换函数或变换矩阵。例如,用于水平变换和垂直变换的变换函数对可以被定义为多个变换集(MTS)。变换器140可以在MTS中选择一个具有最高变换效率的变换函数对,并在水平方向和垂直方向中的每个方向上变换残差块。关于MTS中的变换函数对的信息(mts_idx)由熵编码器155编码并被用信号发送到视频解码装置。

[0074] 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将量化的变换系数输出到熵编码器155。量化器145也可以在不对任何块或帧进行变换的情况下立即量化相关残差块。量化器145也可以根据变换块中变换系数的位置应用不同的量化系数(缩放值)。应用于以二维量化排列的变换系数的量化矩阵可以被编码并被用信号发送到视频解码装置。

[0075] 重排单元150可以为量化的残差信号执行系数值的重新排列。

[0076] 重排单元150可以通过利用系数扫描将二维系数阵列改变为一维系数序列。例如,重排单元150可以通过利用之字形(zig-zag)扫描或对角线扫描从DC系数到高频域系数进行扫描来输出一维系数序列。根据变换单位的大小和帧内预测模式,也可以用在列方向上扫描二维系数阵列的垂直扫描和在行方向上扫描二维块型系数的水平扫描来代替之字形扫描。换言之,根据变换单位的大小和帧内预测模式,可以在之字形扫描、对角线扫描、垂直扫描和水平扫描中确定待使用的扫描方法。

[0077] 熵编码器155通过利用包括基于上下文的自适应二元算术代码(Context-based Adaptive Binary Arithmetic Code,CABAC)、指数哥伦布(Exponential Golomb)等的各种编码方式对从重排单元150输出的一维量化变换系数序列进行编码而生成比特流。

[0078] 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、MTT分割方向等进行编码，以使视频解码装置与视频编码装置相同地分割块。此外，熵编码器155对关于指示当前块是由帧内预测还是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(运动信息编码模式(合并模式或AMVP模式)，在合并模式的情况下，帧间预测信息是合并索引，在AMVP模式的情况下，帧间预测信息是关于参考图片索引和运动矢量差的信息)进行编码。此外，熵编码器155对与量化有关的信息，即关于量化参数的信息和关于量化矩阵的信息进行编码。

[0079] 逆量化器160对从量化器145输出的量化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间域以重建残差块。

[0080] 加法器170将重建的残差块和由预测器120产生的预测块相加以重建当前块。当对下一个顺序的块进行帧内预测时，重建的当前块中的像素被用作参考像素。

[0081] 循环滤波器单元180对重建的像素执行滤波，以减少由于基于块的预测和变换/量化而产生的块伪影(blocking artifact)、振铃伪影(ringing artifact)、模糊伪影(blurring artifact)等。作为环内滤波器的循环滤波器单元180可以包括去块滤波器182、采样自适应偏移(SAO)滤波器184和自适应循环滤波器(ALF)186的全部或部分。

[0082] 去块滤波器182对重建的块之间的边界进行滤波，以去除由于块单位的编码/解码而产生的块伪影，并且SAO滤波器184和ALF 186对去块滤波的视频执行附加的滤波。SAO滤波器184和ALF 186是用于补偿由于有损编码(lossy coding)而产生的重建的像素和原始像素之间的差异的滤波器。SAO滤波器184应用偏移量作为CTU单位，以提高主观图像质量和编码效率。与此相比，ALF 186执行块单位的滤波，并通过划分相应块的边界和变化量的程度，应用不同的滤波器来补偿失真。关于待用于ALF的滤波器系数的信息可以被编码并被用信号发送到视频解码装置。

[0083] 通过去块滤波器182、SAO滤波器184和ALF 186滤波的重建块被存储在存储器190中。当一个图片中的所有块被重建时，则重建的图片可以被用作用于对之后待编码的图片中的块进行帧间预测的参考图片。

[0084] 图5是可以实现本公开技术的视频解码装置的功能框图。在下文中，参照图5，描述视频解码装置和该装置的子组件。

[0085] 视频解码装置可以被配置为包括熵解码器510、重排单元515、逆量化器520、逆变换器530、预测器540、加法器550、循环滤波器单元560和存储器570。

[0086] 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。此外，每个组件的功能可以实现为软件，并且微处理器还可以实现为执行与每个组件对应的软件的功能。

[0087] 熵解码器510通过解码由视频编码装置生成的比特流提取与块分割有关的信息以确定待解码的当前块，并提取重建当前块所需的预测信息和关于残差信号的信息等。

[0088] 熵解码器510通过从序列参数集(SPS)或图片参数集(PPS)中提取关于CTU大小的信息来确定CTU的大小，并且将图片分割成具有确定的大小的CTU。另外，CTU被确定为树结构的最上层，即根节点，并且提取CTU的分割信息以利用树结构分割CTU。

[0089] 例如，当利用QTBTTT结构分割CTU时，首先提取与QT的分割有关的第一标志(QT_

split_flag)以将每个节点分割成下层的四个节点。另外,针对对应于QT的叶节点的节点,提取与MTT的分割有关的第二标志(MTT_split_flag)、分割方向(垂直/水平)和/或分割类型(二元/三元)信息,以将相应的叶节点分割成MTT结构。因此,QT的叶节点下面的每个节点被递归地分割成BT或TT结构。

[0090] 作为另一示例,当通过利用QTBTTT结构分割CTU时,提取指示CU是否被分割的CU分割标志(split_cu_flag)。当相应的块被分割时,也可以提取第一标志(QT_split_flag)。在分割过程中,针对每个节点,在递归QT分割0次或更多次之后,可能会发生递归MTT分割0次或更多次。例如,针对CTU,MTT分割可以立即发生,或者相反,也可以只发生多次的QT分割。

[0091] 作为另一示例,当通过利用QTBT结构分割CTU时,提取与QT的分割有关的第一标志(QT_split_flag)以将每个节点分割成下层的四个节点。另外,还提取指示与QT的叶节点对应的节点是否被进一步分割为BT的分割标志(split_flag)和分割方向信息。

[0092] 另一方面,当熵解码器510通过利用树结构的分割确定待解码的当前块时,熵解码器510提取关于指示当前块是帧内预测还是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)的句法元素。当预测类型信息指示帧间预测时,熵解码器510提取表示帧间预测信息的句法元素,即运动矢量和运动矢量所指向的参考图片的信息。

[0093] 此外,熵解码器510提取量化相关信息,并提取关于当前块的量化变换系数的信息作为关于残差信号的信息。

[0094] 重排单元515可以再次以与视频编码装置执行的系数扫描顺序相反的顺序将由熵解码器510熵解码的一维量化变换系数的序列改变为二维系数阵列(即,块)。

[0095] 逆量化器520对量化的变换系数进行逆量化并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还可以将不同的量化系数(缩放值)应用于以二维排列的量化变换系数。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)的矩阵应用于量化变换系数的二维阵列来执行逆量化。

[0096] 逆变换器530通过将逆量化的变换系数从频域逆变换到空间域来重建残差信号,从而生成当前块的残差块。

[0097] 此外,当逆变换器530逆变换变换块的部分区域(子块)时,逆变换器530提取指示仅变换变换块的子块的标志(cu_sbt_flag)、子块的定向(垂直/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间域,以重建残差信号,并用“0”值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。

[0098] 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发送的MTS信息(mts_idx)来确定在水平方向和垂直方向中的每个方向上待应用的变换函数或变换矩阵。逆变换器530还通过利用确定的变换函数在水平方向和垂直方向上对变换块中的变换系数执行逆变换。

[0099] 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,帧内预测器542被激活,并且在当前块的预测类型是帧间预测时,帧间预测器544被激活。

[0100] 帧内预测器542从从熵解码器510中提取的用于帧内预测模式的句法元素中确定

多种帧内预测模式中的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周围参考像素来预测当前块。

[0101] 帧间预测器544通过利用从熵解码器510提取的用于帧间预测模式的句法元素来确定当前块的运动矢量和该运动矢量所参考的参考图片。

[0102] 加法器550通过添加从逆变换单元输出的残差块和从帧间预测器或帧内预测单元输出的预测块来重建当前块。重建的当前块中的像素在帧内预测之后待解码的块时被用作参考像素。

[0103] 作为环内滤波器的循环滤波器单元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对重建的块之间的边界执行去块滤波，以消除因块单位的解码而发生的块伪影。SAO滤波器564和ALF 566在去块滤波后对重建的块执行附加的滤波，以补偿因有损编码而产生的重建的像素和原始像素之间的差异。ALF的滤波系数通过利用从比特流解码的滤波系数的信息来确定。

[0104] 通过去块滤波器562、SAO滤波器564和ALF 566滤波的重建块被存储在存储器570中。当一个图片中的所有块被重建时，重建的图片可以被用作用于帧间预测之后待编码的图片中的块的参考图片。

[0105] 本实施例涉及如上所述的视频的编码和解码。更具体地，本实施例提供一种帧间预测方法，其中训练可变系数深度学习模型以自适应地学习视频的特征，将由训练生成的可变系数深度学习模型的参数从视频编码装置发送到视频解码装置，并且参考由可变系数深度学习模型生成的虚拟参考帧。

[0106] 以下实施例可以由视频编码装置的帧间预测器124和视频解码装置的帧间预测器544执行。在下文中，术语“目标块”可以在与如上所述的当前块或编码单元(CU)相同的意义上使用，或者可以指编码单元的部分区域。

[0107] I.双向光流(Bi-directional Optical Flow,BDOF)

[0108] 双向光流(bi-directional OF:BDOF)是在构成视频的样本或物体以恒定速度移动且样本值变化很小的假设下，基于光流(Optical Flow,OF)对使用双向运动预测的样本的运动进行附加补偿的技术。

[0109] 假定指向参考图片$Ref_0$和$Ref_1$中与当前图片的待编码的目标块最相似的相应区域(即参考块)的双向运动矢量$MV_0$和$MV_1$已经通过对目标块的(通常的)双向运动预测如图6中所示被确定。在这种情况下，两个运动矢量是表示整个目标块运动的值。在图6的示例中，由对应于目标块中的样本P的运动矢量$MV_0$指示的参考图片$Ref_0$中的样本为$P_0$，而由对应于目标块中的样本P的运动矢量$MV_1$指示的参考图片$Ref_1$中的样本为$P_1$。

[0110] 在下文中，在图6的示例中假定样本P的运动与目标块的整体运动略有不同。例如，当位于$Ref_0$中的样本A的物体通过当前图片的目标块中的样本P并移动到$Ref_1$中的样本B时，样本A和样本B应该具有非常相似的值。因此，$Ref_0$中与目标块中的样本P最相似的点不是运动矢量$MV_0$所指示的$P_0$，而是通过预定的位移矢量$(v_xt_0, v_yt_0)$移动$P_0$所得到的A。此外，$Ref_1$中与目标块中的样本P最相似的点不是由运动矢量$MV_1$所指示的$P_1$，而是通过预定的位移矢量$(-v_xt_1, -v_yt_1)$移动$P_1$得到的B。此处，$t_0$和$t_1$分别指以当前图片为基准的$Ref_0$和$Ref_1$的时间轴距离，并且是基于图片顺序计数(Picture Order Count,POC)计算的。在这种情况下，$(v_x, v_y)$被称为OF。因此，OF表示以像素为单位的运动矢量，该运动矢量表示图片之间发生的像

素的运动。

[0111] 在BDOF中，当预测当前图片中目标块的样本P的值时，与使用由双向运动矢量$MV_0$和$MV_1$指示的参考样本$P_0$和$P_1$，即预测样本相比，通过使用两个参考样本A和B的值来执行的预测更精确。

[0112] 参考图片$Ref_0$中由对应于目标块中的(亮度)像素(x,y)的运动矢量$MV_0$指示的预测样本的值被定义为$I^{(0)}_{(x,y)}$，并且参考图片$Ref_1$中由对应于目标块中的样本(x,y)的运动矢量$MV_1$指示的预测样本的值被定义为$I^{(1)}_{(x,y)}$。

[0113] 参考图片Ref0中对应于由BDOF运动矢量$(v_x,v_y)$指示的目标块中的样本的样本A的值可以定义为$I^{(0)}(i+v_xt_{0,j}+v_yt_0)$，参考图片$Ref_1$中的样本B的值可以定义为$I^{(1)}(i-v_xt_{1,j}-v_yt_1)$。此外，当假定时间$t_0$和$t_1$相同，并利用“$t_0+t_1$”执行归一化时，样本A的值可以定义为$I^{(0)}(i+1/2v_{x,j}+1/2v_y)$，样本B的值可以定义为$I^{(1)}(i-1/2v_{x,j}-1/2v_y)$。

[0114] 在BDOF中，目标块的像素值可以利用像素的预测样本的值、OF和预测样本的梯度来预测，如等式1所示。

[0115] [等式1]

[0116] $pred(x,y)=(I^{(0)}(x,y)+I^{(1)}(x,y)+b(x,y)+o_{offset})>>shift$

[0117] 此处，像素的预测样本的值是基于用于帧间预测的双向预测的运动信息(即两个运动矢量和两个参考图片)来预测的。“b(x,y)”是利用预测样本中的OF和梯度计算的样本偏移。shift是缩放因子，$o_{offset}$是用于舍入计算的值，是shift的一半。

[0118] II.帧间预测的合并/跳过模式

[0119] 在下文中，参照图7所示的流程图，描述在帧间预测的合并/跳过模式中构建运动矢量的合并候选列表的方法。为了支持合并模式，帧间预测器124可以选择预设数量的(例如，六个)合并候选，并构建合并候选列表。

[0120] 图7是示出根据本公开的实施例的在合并/跳过模式中搜索运动矢量候选的过程的流程图。

[0121] 帧间预测器124搜索空间合并候选(S700)。如图4所示，帧间预测器124从周围块中搜索空间合并候选。最多可以选择四个空间合并候选。

[0122] 帧间预测器124搜索时间合并候选(S702)。帧间预测器124可以添加在参考图片(可以与用于预测当前块的参考图片相同或不同)而不是当前块所在的当前图片中与当前块相同位置的块作为时间合并候选。可以选择一个时间合并候选。

[0123] 帧间预测器124搜索基于历史的预测运动矢量(History-based Motion Vector Predictor，HMVP)候选(S704)。帧间预测器124可以将以前的n(其中n是自然数)个CU的运动矢量存储在表中，然后将这些作为合并候选。该表的大小可以是6，并且以前的CU的运动矢量可以按照先进先出(FIFO)方式来存储。这表明在表中最多存储6个HMVP候选。帧间预测器124可以将存储在表中的HMVP候选中最近的一些运动矢量设置为合并候选。

[0124] 帧间预测器124搜索成对平均MVP(Pairwise Average MVP，PAMVP)候选(S706)。帧间预测器124可以将合并候选列表中的第一候选和第二候选的运动矢量的平均值设置为新的合并候选。

[0125] 当即使执行了所有上述过程(S700至S706)也不能填满合并候选列表时(即，当不能填满预设数量时)，帧间预测器124添加零运动矢量作为剩余的合并候选(S708)。

[0126] III.仿射运动预测

[0127] 如上所述的帧间预测是反映平移运动(translation motion)模型的运动预测。换言之,帧间预测是用于预测水平方向(x轴方向)和垂直方向(y轴方向)的运动的方案。然而,在现实中,除了平移运动外,还可能有各种类型的运动,如旋转、放大或缩小等。根据本实施例的一个方面提供可以反映这种各种类型的运动的仿射运动预测(affine motion prediction)。

[0128] 图8是示出根据本公开的实施例的仿射运动预测的示图。

[0129] 可以有两种用于仿射运动预测的模型。其中一个模型使用当前待编码的目标块的左上角和右上角的两个控制点运动矢量(CPMV),即使用四个参数,如图8的(a)中所示。另一个模型使用目标块的左上角、右上角和左下角的三个控制点运动矢量,即使用六个参数,如图8的(b)所示。

[0130] 4参数仿射模型表示为如等式2所示。在目标块中的样本位置(x,y)的运动可以如等式2所示来计算。此处,假设目标块的左上样本的位置为(0,0)。

[0131] [等式2]

[0132] $$\begin{cases} mv_x = \dfrac{mv_{1x} - mv_{0x}}{W}x - \dfrac{mv_{1y} - mv_{0y}}{W}y + mv_{0x} \\ mv_y = \dfrac{mv_{1y} - mv_{0y}}{W}x + \dfrac{mv_{1x} - mv_{0x}}{W}y + mv_{0y} \end{cases}$$

[0133] 此外,6参数仿射模型表示为如等式(3)所示。目标块中样本位置(x,y)的运动可以如等式3所示来计算。

[0134] [等式3]

[0135] $$\begin{cases} mv_x = \dfrac{mv_{1x} - mv_{0x}}{W}x + \dfrac{mv_{2x} - mv_{0x}}{H}y + mv_{0x} \\ mv_y = \dfrac{mv_{1y} - mv_{0y}}{W}x + \dfrac{mv_{2y} - mv_{0y}}{H}y + mv_{0y} \end{cases}$$

[0136] 此处,$(mv_{0x},mv_{0y})$是左上角控制点的运动矢量,$(mv_{1x},mv_{1y})$是右上角控制点的运动矢量,$(mv_{2x},mv_{2y})$是左下角控制点的运动矢量。W是目标块的水平长度,H是目标块的垂直长度。

[0137] 可以利用根据等式2或等式3计算的运动矢量对目标块中的每个样本执行仿射运动预测。

[0138] 可选地,如图9所示,为了降低计算复杂性,可以以从目标块划分的子块为单位执行仿射运动预测。子块的大小可以是,例如4×4、2×2或8×8。在下文中,描述仿射运动预测是以4×4子块为单位对目标块执行的示例。这一示例只是为了描述的方便,并不限制本公开的范围。

[0139] 在以子块为单位的仿射运动预测中,通过将每个子块的中心位置代入等式2或等式3中的(x,y)来计算每个子块的运动矢量(仿射运动矢量)。此处,中心位置可以是子块的实际中心点,也可以是中心点的右下样本的位置。例如,在左下样本的坐标是(0,0)的4×4子块的情况下,子块的中心位置可以是(1.5,1.5)或(2,2)。利用子块的仿射运动矢量$(mv_x,mv_y)$生成每个子块的预测块。

[0140] 运动矢量$(mv_x,mv_y)$可以被设置为具有1/16的样本精度。在这种情况下,根据等式2

或等式3计算的运动矢量($mv_x$,$mv_y$)可以以1/16样本为单位进行舍入。

[0141] 仿射运动预测不仅可以对亮度(luma)分量执行,而且可以对色度(chroma)分量执行。在4:2:0视频格式的情况下,当以4×4子块为单位对亮度分量执行仿射运动预测时,可以以2×2子块为单位对色度分量执行仿射运动预测。色度分量的每个子块的运动矢量($mv_x$,$mv_y$)可以从相应的亮度分量的运动矢量中得到。可选地,用于色度分量的仿射运动预测的子块的大小可以与亮度分量的子块相同。当以4×4子块为单位对亮度分量执行仿射运动预测时,以4×4子块为单位对色度分量执行仿射运动预测。在这种情况下,由于色度分量的4×4子块对应于亮度分量的四个4×4子块,因此色度分量的子块的运动矢量($mv_x$,$mv_y$)可以从亮度分量的四个相应子块的运动矢量的平均值计算出来。

[0142] 视频编码装置通过执行帧内预测、帧间预测(平移运动预测)、仿射运动预测等,并计算速率-失真(rate-distortion,RD)成本(cost)来选择最佳预测方法。为了执行仿射运动预测,视频编码装置的帧间预测器124确定要使用两种类型的模型中的哪一类,并根据模型的类型确定两个或三个控制点。帧间预测器124利用控制点的运动矢量计算目标块中4×4子块中每一个的运动矢量($mv_x$,$mv_y$)。帧间预测器124通过利用每个子块的运动矢量($mv_x$,$mv_y$)以子块为单位在参考图片中执行运动补偿来生成目标块中每个子块的预测块。

[0143] 视频编码装置的熵编码器155对包括指示仿射运动预测是否应用于目标块的标志、指示仿射模型类型的类型信息以及指示每个控制点的运动矢量的运动信息等的仿射相关的句法元素进行编码并发送到视频解码装置。当执行仿射运动预测时,可以用信号发送类型信息和控制点的运动信息,并且可以用信号发送与根据类型信息确定的数量相同的数量的控制点的运动矢量。

[0144] 视频解码装置利用用信号发送的句法确定仿射模型的类型和控制点运动矢量,并利用等式2或等式3计算目标块中每个4×4子块的运动矢量($mv_x$,$mv_y$)。当用信号发送目标块的仿射运动矢量的运动矢量分辨率信息时,利用诸如舍入的计算修改运动矢量($mv_x$,$mv_y$)以具有由运动矢量分辨率信息识别的精度。

[0145] 视频解码装置通过利用每个子块的运动矢量($mv_x$,$mv_y$)在参考图片中执行运动补偿来生成每个子块的预测块。

[0146] 为了减少对控制点的运动矢量进行编码所需的比特量,可以应用如上所述的一般帧内预测(即平移运动预测)方式。

[0147] 作为示例,在合并模式的情况下,视频编码装置的帧间预测器124构建具有预定数量(例如,五个)的候选的仿射合并候选列表。首先,视频编码装置的帧间预测器124从目标块的周围块中得出继承的(inherited)仿射合并候选。例如,帧间预测器124从图4所示的目标块的周围样本A0、A1、B0、B1和B2中得出预定数量的继承的仿射合并候选,以生成合并候选列表。包括在候选列表中的每个继承的仿射合并候选对应于两个或三个CPMV的一对。

[0148] 帧间预测器124从目标块的周围块中已经以仿射模式预测的周围块的控制点运动矢量中得出继承的仿射合并候选。在一些实施例中,从以仿射模式预测的周围块得出的合并候选的数量可能是有限的。例如,帧间预测器124可以从已经以仿射模式预测的周围块中得出总共两个继承的仿射合并候选,一个来自A0和A1,一个来自B0、B1和B2。优先级可以是A0和A1的顺序,也可以是B0、B1和B2的顺序。

[0149] 另一方面,当合并候选的总数为三个或更多时,帧间预测器124可以从周围块的平

移运动矢量得出与不足数量一样多的“构建的(constructed)仿射合并候选”。

[0150] 图10是示出从周围块的平移运动矢量得出用于仿射运动预测的构建的仿射合并候选的方法的示图。

[0151] 帧间预测器124分别从周围块组{B2、B3和A2}、周围块组{B1和B0}以及周围块组{A1和A0}得出控制点运动矢量CPMV1、CPMV2和CPMV3。作为示例,每个周围块组内的优先级可以是B2、B3和A2的顺序、B1和B0的顺序以及A1和A0的顺序。此外,另一个控制点运动矢量CPMV4从参考图片中的相同位置块(collocated block(并列块))T中得出。帧间预测器124结合四个控制点运动矢量中的两个或三个控制点运动矢量,以生成与不足数量一样多的构建的仿射合并候选。组合的优先级如下。每组中的元素按左上角、右上角和左下角点运动矢量的顺序排列。

[0152] {cpmv1,cpmv2,cpmv3},{cpmv1,cpmv2,cpmv4},{cpmv1,cpmv3,cpmv4},

[0153] {cpmv2,cpmv3,cpmv4},{cpmv1,cpmv2},{cpmv1,cpmv3}

[0154] 当仅利用继承的仿射合并候选和构建的仿射合并候选不能填满合并候选列表时,帧间预测器124可以添加零运动矢量作为候选。

[0155] 帧间预测器124从合并候选列表中选择合并候选,并对目标块执行仿射运动预测。当选择的候选包括两个控制点运动矢量时,帧间预测器124利用4参数模型执行仿射运动预测。另一方面,当选择的候选由三个控制点运动矢量组成时,利用6参数模型执行仿射运动预测。视频编码装置的熵编码器155对指示从合并候选列表中的合并候选中选择的合并候选的索引信息进行编码,并将该索引信息用信号发送到视频解码装置。

[0156] 视频解码装置的熵解码器510对从视频编码装置用信号发送的索引信息进行解码。视频解码装置的帧间预测器544以与视频编码装置相同的方式建合并候选列表,并且利用对应于由索引信息指示的合并候选的控制点运动矢量执行仿射运动预测。

[0157] 作为另一示例,在AMVP模式的情况下,视频编码装置的帧间预测器124确定目标块的仿射模型的类型和控制点运动矢量。帧间预测器124计算作为目标块的实际控制点运动矢量和各控制点的预测运动矢量之间的差的运动矢量差,并对每个控制点的运动矢量差进行编码和发送。为此,视频编码装置的帧间预测器124构建预定数量的(例如,两个)仿射AMVP候选列表。当目标块是4参数类型时,包括在列表中的每个候选由两个控制点运动矢量的一对组成。另一方面,当目标块是6参数类型时,包括在列表中的每个候选由三个控制点运动矢量的一对组成。

[0158] 在下文中,利用图11的示例描述在帧间预测的仿射AMVP模式中构建候选列表的方法。仿射AMVP候选列表可以类似于上述构建仿射合并候选列表的方法得出。

[0159] 图11是示出根据本公开的实施例的在AMVP模式中搜索AMVP候选的过程的流程图。

[0160] 帧间预测器124检查继承的仿射AMVP候选的参考图片是否与当前块的参考图片相同(S1100)。此处,如在上述仿射合并模式中一样,继承的仿射AMVP候选可以是在图4所示的目标块的周围块A0、A1、B0、B1和B2中以仿射模式预测的块。

[0161] 当继承的仿射AMVP候选的参考图片与当前块的参考图片相同时(在S1100中为“是”),帧间预测器124添加继承的仿射AMVP候选(S1102)。

[0162] 当继承的仿射AMVP候选的参考图片与当前块的参考图片不相同时(S1100中为“否”),帧间预测器124检查构建的仿射AMVP候选的所有CPMV的参考图片是否与当前块的参

考图片相同(S1104)。此处,如在上述仿射合并模式中一样,构建的仿射AMVP候选的所有CPMV可以从图10所示的周围样本的运动矢量中得出。

[0163] 当构建的仿射AMVP候选的所有CPMV的参考图片与当前块的参考图片相同时(在S1104中为“是”),帧间预测器124添加构建的仿射AMVP候选(S1106)。

[0164] 在这种情况下,应考虑目标块的仿射模型类型。当目标块的仿射模型类型是4参数类型时,视频编码装置利用周围块的仿射模型得出两个控制点运动矢量(目标块的左上角和右上角点运动矢量)。当目标块的仿射模型类型为6参数类型时,利用周围块的仿射模型得出三个控制点运动矢量(目标块的左上角、右上角和左下角控制点运动矢量)。

[0165] 当周围块是4参数类型时,根据目标块的仿射模型类型,利用周围块的两个控制点运动矢量来预测两个或三个控制点运动矢量。例如,可以利用由等式2表示的周围块的仿射模型。在等式2中,$(mv_{0x},mv_{0y})$和$(mv_{1x},mv_{1y})$分别被替换为周围块的左上角和右上角控制点运动矢量。通过将目标块的控制点的位置与周围块的左上角位置的差值输入(x,y),可以得出目标块的每个控制点的预测运动矢量。

[0166] 当周围块是6参数类型时,根据目标块的仿射模型类型,利用周围块的三个控制点运动矢量预测两个或三个控制点运动矢量。例如,可以利用由等式3表示的周围块的仿射模型。在等式3中,$(mv_{0x},mv_{0y})$、$(mv_{1x},mv_{1y})$和$(mv_{2x},mv_{2y})$分别被替换为周围块的左上角、右上角和左下角控制点运动矢量。通过将目标块的控制点的位置与周围块的左上角位置之间的差值输入(x,y),可以得出目标块的每个控制点的预测运动矢量。

[0167] 当所有CPMV的参考图片与当前块的参考图片不相同时(S1104中为“否”),帧间预测器124添加平移运动矢量作为仿射AMVP候选(S1108)。平移运动矢量可以用于以$mv_0$、$mv_1$和$mv_2$的顺序预测当前块的CPMV。

[0168] 在即使执行所有上述步骤(S1100至S1108)也不能填满候选列表的情况下(即,在不能获得预设的候选数的情况下),帧间预测器124添加零运动矢量作为仿射AMVP候选(S1110)。

[0169] 视频编码装置的帧间预测器124从仿射AMVP列表中选择一个候选,并生成每个实际控制点的运动矢量与选择的候选的相应控制点的预测运动矢量之间的运动矢量差。视频编码装置的熵编码器155对指示目标块的仿射模型类型的类型信息、指示从仿射AMVP列表中的候选中选择的候选的索引信息以及对应于每个控制点的运动矢量差进行编码,并发送到视频解码装置。

[0170] 视频解码装置的帧间预测器544利用从视频编码装置用信号发送的信息确定仿射模型类型,并生成每个控制点的运动矢量差。帧间预测器544以与视频编码装置相同的方式生成仿射AMVP列表,并选择仿射AMVP列表中由用信号发送的索引信息指示的候选。视频解码装置的帧间预测器544将选择的候选的每个控制点的预测运动矢量添加到相应的运动矢量差以计算每个控制点的运动矢量。

[0171] IV.虚拟参考帧的创建和在帧间预测模式中的使用

[0172] 基于可变系数深度学习的视频预测是双环编码(double-loop coding)并包括以下两步过程。该两步包括第一步,在该第一步中,基于可变系数深度学习模型对给定视频进行分析,以获得深度学习模型的参数,并包括第二步,在该第二步中,将获得的深度学习模型的参数应用于视频预测过程,并发送编码的比特流和深度学习模型的参数,用于视频编

码和解码。

[0173] 在第一步中，为了使与视频运动预测相关的深度学习模型的参数适应于各个视频，视频编码装置通过利用各个视频作为训练数据来训练深度学习模型。例如，深度学习模型可以用于实现视频的运动预测信号的细化、运动矢量的精确性的实现、运动矢量的分辨率的改变和参考帧的改进中的一个或多个。

[0174] 深度学习模型可以是包括固定系数的层和包括可变系数的一个或多个卷积层的组合。包括固定系数的层(在下文中称为“固定系数网络”)提供诸如运动预测、运动插值和运动外推(extrapolation)的功能。可选地，只包括可变系数的卷积层(在下文中称为“可变系数网络”)可以作为深度学习模型使用。如上所述，任何类型的深度学习模型都要经过训练以适应各个视频。

[0175] 在第二过程中，视频编码装置和视频解码装置利用经训练的深度学习模型执行运动预测信号的细化、运动矢量的精确性和超分辨率化、参考帧的峰值信噪比(PSNR)的改善等。对执行这些功能的深度学习模型的利用可以以块单位或序列单位控制。例如，可以在视频的序列参数集(SPS)、视频参数集(VPS)、图片参数集(PPS)、或子图片或图块中利用启用深度学习模型的标志。可选地，可以在视频的块单位(即CTU或CU单位)中利用启用深度学习模型的标志。视频编码装置将编码的比特流和深度学习模型的可变系数的值发送到视频解码装置。

[0176] 由于在第一步中训练深度学习模型的可变系数网络可能需要相当多的时间，因此可变系数深度学习难以应用于需要低延迟编码的领域，例如实时视频会议或实时流媒体服务。然而，由于可以利用存储的视频提前执行第一步，因此可变系数深度学习可以容易地应用于使用随机存取编码的领域。

[0177] 以下，描述利用包括可变系数网络作为深度学习模型的视频插值模型的实施例。在下文中，视频插值模型也被称为“可变系数插值模型”或“插值模型”。如上所述，插值模型也可以包括固定系数网络。描述利用插值模型从参考帧生成虚拟参考帧的过程、利用虚拟参考帧对视频进行编码的过程以及通过发送编码的比特流和可变系数深度学习模型对视频进行解码的过程。然而，本公开所适用的领域不限于上述情况，并且可以通过利用可变系数深度学习模型对运动预测信号进行细化并实现运动矢量的精确性来来扩展所适用的领域。

[0178] 通过视频编码装置和视频解码装置，可以通过在参考图片列表中包括由插值模型生成的虚拟参考帧，然后在合并模式和AMVP模式中使用虚拟参考帧进行运动预测来获得编码增益。特别地，众所周知，当虚拟参考帧与原始帧相似时，编码效率会增加(见非专利文献1)。当虚拟参考帧生成时，视频编码装置或视频解码装置可以通过利用自适应地训练为在输入帧上过度拟合(overfitting)的可变系数插值模型更容易地生成与原始帧更相似的帧。因此，在本公开中，描述一种利用如图12、图13和图14所示的可变系数插值模型生成虚拟参考帧并利用该虚拟参考帧提高编码效率的方法。

[0179] 图12是示出根据本公开的实施例的包括插值模型的视频编码装置的框图。

[0180] 如图12所示，视频编码装置可以包括虚拟帧生成器1210，其从存储在存储器190中的已解码的周围参考帧生成虚拟参考帧。此处，虚拟帧生成器1210包括可变系数插值模型。下面描述插值模型的操作。

[0181] 图13是示出根据本公开的实施例的包括插值模型的视频解码装置的框图。

[0182] 如图13所示，视频解码装置可以包括虚拟帧生成器1310，其从存储在存储器570中的已解码的周围参考帧生成虚拟参考帧。此处，虚拟帧生成器1310包括可变系数插值模型。下面描述插值模型的操作。

[0183] 图14是示出根据本公开的实施例的视频编码装置和视频解码装置中使用的插值模型的操作的示图。

[0184] 视频编码装置的虚拟帧生成器1210利用插值模型生成推断帧$X_{t,pred}$，该推断帧$X_{t,pred}$与来自已解码的周围参考帧$X_{t-1}$和$X_{t+1}$的原始帧$X_t$非常相似。推断帧$X_{t,pred}$被用作虚拟参考帧。当使用自适应训练的插值模型来生成与原始帧非常相似的帧$_{Xt,pred}$时，视频编码装置可以生成具有高度相似性的虚拟参考帧。这使得视频编码装置可以对待编码的输入帧使用已被自适应地训练的可变系数插值模型。在这种情况下，视频编码装置将经训练的插值模型的可变系数值发送到视频解码装置。

[0185] 随着包括在插值模型中的可变系数的数量的增加，可变系数插值模型可以被训练成使其更容易将输入帧过度拟合到目标帧。然而，当可变系数的数量增加以利用插值模型生成与原始帧更相似的帧时，这可能导致发送的比特数增加，并降低编码效率。因此，作为可变系数插值模型，应考虑在使用最小数量的可变系数的同时具有良好的编码效率的模型。

[0186] 另一方面，如上所述，视频解码装置的虚拟帧生成器1310利用从视频编码装置发送的可变系数值设置插值模型的可变系数。插值模型从已解码的周围参考帧$X_{t-1}$和$X_{t+1}$推断出与原始帧$X_t$非常相似的帧$X_{t,pred}$。推断帧$X_{t,pred}$被用作虚拟参考帧。

[0187] 在下文中，描述由视频编码装置执行的插值模型的训练。如上所述，插值模型的训练对应于基于可变系数的深度学习的视频预测的第一步。

[0188] 为了训练插值模型以生成虚拟参考帧，视频编码装置可以另外包括训练单元(未示出)，或者可以以链接到外部训练单元的形式实现。

[0189] 如上所述从已解码的周围参考帧$X_{t-1}$和$X_{t+1}$生成推断帧$X_{t,pred}$的插值模型“f”的操作可表示为等式4所示。

[0190] [等式4]

[0191] $X_{t,pred}=f(X_{t-1},X_{t+1},\theta_t)$

[0192] 此处，参数$\theta_t$表示可变系数。

[0193] 训练单元可以将参数$\theta_t$更新为在待编码的第t个帧上过度拟合。在这种情况下，用于训练插值模型的损失函数被定义为如等式5所示。

[0194] [等式5]

[0195] $$L=\frac{1}{H\times W}\sum_{i}^{H\times W}|X_{t,i}-X_{t,i,pred}|^2$$

[0196] 此处，H和W分别表示帧的垂直和水平像素的数量。$X_{t,i,pred}$表示推断帧$X_{t,pred}$的第i个像素的值。$X_{t,i}$表示原始帧的第i个像素的值。如上所述，视频编码装置将为每一帧训练的各自可变系数$\theta_t$的值发送到视频解码装置。视频解码装置可以利用接收到的可变系数$\theta_t$的值设置与包括在视频编码装置中的相同的插值模型的可变系数$\theta_t$，然后利用与等式4中所

示的相同方法生成虚拟参考帧。

[0197] 图15示出根据本公开的实施例的可变系数插值模型的示例。

[0198] 图15中示出的插值模型可以是适合视频处理的基于卷积神经网络(CNN)的深度学习模型。

[0199] 根据本公开的插值模型可以是直接从周围参考帧生成推断帧的深度学习模型，如图15的(a)所示。

[0200] 在根据本公开的另一实施例中，插值模型可以是利用掩码的深度学习模型(见非专利文件1)，如图15的(b)所示。在这种情况下，插值模型可以从参考帧生成掩码，然后利用掩码对已解码的周围参考帧$X_{t-1}$和$X_{t+1}$执行加权和以生成推断帧$X_{t,pred}$。

[0201] 在根据本公开的另一实施例中，插值模型可以是使用光流的深度学习模型，如图15的(c)所示。如关于BDOF所描述的，使用OF的深度学习模型可以在假设时间归一化的基础上生成推断帧$X_{t,pred}$。插值模型可以从参考帧生成OF，然后在OF的基础上扭转(warp)参考帧以生成推断帧$X_{t,pred}$。

[0202] 另一方面，在图15中示出的任何插值模型中，可以基于等式5所示的损失函数更新可变系数。

[0203] 图15中示出的任何深度学习模型可以被用作可变系数插值模型。然而，由于应发送可变系数值，因此只有通过利用最简单的模型才能提高编码效率。另一方面，存在的问题是，当使用较少数量的可变系数时，过度拟合变得困难。

[0204] 如上所述，在插值模型中，包括可变系数的部分被表示为可变系数网络。

[0205] 图16是示出根据本公开的实施例的进一步包括固定系数网络的插值模型的示图。

[0206] 因此，如图16的示例，进一步包括固定系数网络的插值模型从输入帧生成$X_{t-1\rightarrow t}$和$X_{t+1\rightarrow t}$作为中间输出，$X_{t-1\rightarrow t}$和$X_{t+1\rightarrow t}$是类似于待编码的帧的帧。在这种情况下，固定系数网络可以预先在视频编码装置中基于整个数据进行训练。此外，视频编码装置和视频解码装置可以在视频编码装置和视频解码装置中存储相同的固定系数值的状态下根据预定的协议设置固定系数网络。

[0207] 通过利用中间输出$X_{t-1\rightarrow t}$和$X_{t+1\rightarrow t}$，插值模型可以利用较少数量的可变系数提供良好的性能。此外，视频编码装置将应用于可变系数网络的可变系数值发送到视频解码装置。

[0208] 如图16的示例，通过将可变系数网络组合到固定系数网络的后端，可以在待编码的帧上进行过度拟合，但是在另一实施例中，可以只使用可变系数网络而不使用固定系数网络。在这种情况下，可以通过利用基于师生网络(student-teacher network)的训练方法，以包括现有固定系数网络的插值模型的缩小形式实现仅包括可变系数网络的插值模型。

[0209] 在下文中，利用图17的示例描述一种用于减少可变系数的数据量的方案。

[0210] 图17是示出根据本公开的实施例的发送可变系数的方法的示图。

[0211] 视频编码装置可以以帧单位或GOP单位发送可变系数。如图17的(a)所示，可以以每个t为单位，即每一帧，发送可变系数$\theta_t$。然而，视频编码装置可以通过以GOP为单位或以多个GOP为单位发送可变系数来减少可变系数的数据量。

[0212] 例如，当如图17的(b)所示GOP的大小为8时，视频编码装置可以在每GOP的基础上发送可变系数，以发送四个帧中的每个帧的可变系数。此外，视频编码装置可以以2个GOP或

3个GOP为单位延长发送周期。在这种情况下，插值模型可以在输入帧被批量累积的状态下进行训练，并且训练单元可以使用如等式6所示考虑时间“t”的损失函数。

[0213] [等式6]

[0214] $$L = \frac{1}{H \times W \times T} \sum_{t}^{T} \sum_{i}^{H \times W} |X_{t,i} - X_{t,i,pred}|^2$$

[0215] 此处，T表示包含在训练批次中的帧数。

[0216] 另一方面，浮点32位(bit)格式的可变系数$\theta_t$基本上用于训练过程，但是视频编码装置可以将可变系数值改变为16位、8位或更少位的十进制格式，然后将可变系数值发送到视频解码装置。

[0217] 在根据本公开的另一实施例中，视频编码装置可以对可变系数$\theta_t$进行量化，然后将量化的可变系数$\theta_t$发送到视频解码装置。例如，视频编码装置可以将可变系数$\theta_t$的范围划分为均匀的区间以量化可变系数。

[0218] 在根据本公开的另一实施例中，视频编码装置可以发送预设数量的可变系数集，而不是直接发送每个帧或GOP的可变系数值。可选地，在可变系数集通常存储在视频编码装置和视频解码装置中的状态下，视频编码装置可以只发送预定的索引以指示一组可变系数。此处，可变系数集可以在训练插值模型的过程中根据视频序列的信号特征、视频内容的类型、量化参数等提前确定。

[0219] 在下文中，利用图18和图19的流程图描述使用由插值模型生成的虚拟参考帧进行帧间预测的过程。

[0220] 图18是示出根据本公开的实施例的在帧间预测的合并/跳过模式中使用虚拟参考帧的过程的流程图。

[0221] 帧间预测器124搜索空间合并候选(S1800)。如图4所示，帧间预测器124从周围块中搜索空间合并候选。在这种情况下，周围块的参考图片可以是虚拟参考帧或现有参考帧。

[0222] 帧间预测器124在考虑虚拟参考帧的情况下搜索时间合并候选(S1802)。

[0223] 帧间预测器124检查在与当前块相同的位置处存在块的图片和时间合并候选的参考帧是虚拟参考帧还是现有参考帧。

[0224] 当在与当前块相同的位置处存在块的图片和时间合并候选的参考帧都是虚拟参考帧时，帧间预测器124将零运动矢量设置为候选并在跳过或合并模式中使用当前图片的虚拟参考帧。

[0225] 当在与当前块相同的位置处存在块的图片是虚拟参考帧，并且时间合并候选的参考帧是现有参考帧时，帧间预测器124将零运动矢量设置为候选，并且在跳过或合并模式中使用当前图片的虚拟参考帧。可选地，帧间预测器124可以不将候选设置为时间合并候选。

[0226] 另一方面，当在与当前块相同的位置处存在块的图片是现有参考帧，而时间合并候选的参考帧是虚拟参考帧时，帧间预测器124将零运动矢量设置为候选，并在跳过或合并模式中使用当前图片的虚拟参考帧。可选地，帧间预测器124可以不将候选设置为时间合并候选。

[0227] 帧间预测器124搜索HMVP候选(S1804)。当先前n个CU的运动矢量被存储在表中时，帧间预测器124可以包括参考虚拟参考帧的编码CU。如上所述，帧间预测器124可以将存储

在表中的HMVP候选中的最近运动矢量设置为合并候选。

[0228] 帧间预测器124搜索PAMVP候选(S1806)。

[0229] 首先,PAMVP的参考图片被设置为第一合并候选的参考图片。当第一合并候选使用虚拟参考帧时,帧间预测器124可以将对应于零运动矢量和第二候选的运动矢量之间的平均值的运动矢量设置为合并候选。

[0230] 在即使执行所有上述过程(S1800至S1806)也不能填满合并候选列表的情况下(即,在不能填满预设数量的情况下),帧间预测器124添加零运动矢量作为合并候选(S1808)。

[0231] 图19是示出根据本公开的实施例的在帧间预测的仿射AMVP模式中使用虚拟参考帧的过程的流程图。

[0232] 帧间预测器124检查继承的仿射AMVP候选的参考图片是否与当前块的参考图片相同(S1900)。

[0233] 当继承的仿射AMVP候选的参考图片与当前块的参考图片相同时,帧间预测器124检查相同的参考图片是虚拟参考帧还是现有参考帧(S1902)。

[0234] 当相同参考图片是虚拟参考帧时,帧间预测器124添加零运动矢量作为仿射AMVP候选(S1904)。

[0235] 当相同参考图片是现有参考帧时,帧间预测器124添加继承的仿射AMVP候选(S1906)。可选地,帧间预测器124可以不添加候选作为继承的仿射AMVP候选。

[0236] 当继承的仿射AMVP候选的参考图片与当前块的参考图片不相同时,帧间预测器124检查构建的仿射AMVP候选的所有CPMV的参考图片是否与当前块的参考图片相同(S1910)。

[0237] 当构建的仿射AMVP候选的所有CPMV的参考图片与当前块的参考图片相同时,帧间预测器124检查相同的参考图片是虚拟参考帧还是现有参考帧(S1912)。

[0238] 当相同参考图片是虚拟参考帧时,帧间预测器124添加零运动矢量作为仿射AMVP候选(S1914)。

[0239] 当相同参考图片是现有参考帧时,帧间预测器124添加构建的仿射AMVP候选(S1916)。可选地,帧间预测器124可以不添加候选作为构建的仿射AMVP候选。

[0240] 当构建的仿射AMVP候选的所有CPMV的参考图片与当前块的参考图片不相同时,帧间预测器124添加平移运动矢量作为仿射AMVP候选(S1920)。

[0241] 在即使执行所有上述过程(S1900至S1920)也不能填满候选列表的情况下(即,在不能填满预设数量的情况下),帧间预测器124添加零运动矢量作为仿射AMVP候选(S1922)。

[0242] 另一方面,在上述描述中,除了添加平移运动矢量作为候选的步骤之外的步骤(S1900至S1916以及S1922)可以类似地应用于由帧间预测器124搜索仿射合并候选的过程。

[0243] 如上所述,根据本实施例,提供一种帧间预测方法,该方法包括参考由基于可变系数深度学习的插值模型生成的虚拟参考帧,并且该方法增加帧间预测的效果并提高编码效率。

[0244] 根据本实施例的每个流程图描述了各个过程是按顺序执行的,但是本公开不一定限于此。换言之,由于改变和执行流程图中描述的过程或并行执行一个或多个过程是适用的,所以流程图不限于时间序列的顺序。

[0245] 在上述描述中,应当理解的是,实施例可以以许多不同的方式实现。一个或多个示例中描述的功能或方法可以实现为硬件、软件、固件或其任意组合。应该理解的是,本说明书中描述的功能组件被标记为“……单元”,以特别进一步强调其实施的独立性。

[0246] 另一方面,本公开中描述的各种功能或方法也可以通过存储在非暂时性记录介质中的指令来实现,该指令可以由一个或多个处理器读取和执行。非暂时性记录介质包括例如以计算机系统可读的形式存储数据的所有类型的记录装置。例如,非暂时性记录介质包括诸如可擦除可编程只读存储器(EPROM)、闪存驱动器、光学驱动器、磁性硬盘驱动器和固态驱动器(SSD)等存储介质。

[0247] 以上的描述只是示例性地描述了本实施例的技术思想,本领域普通技术人员应该理解的是,在不脱离本实施例的本质特征的范围内,可以进行各种修改和变形。因此,本实施例是用于描述本实施例的技术思想,而不是用于限制本实施例的技术思想,本实施例的技术思想的范围不受这些实施例的限制。本实施例的保护范围应根据所附的权利要求书来解释,所有等同范围内的技术思想也应被解释为包括在本实施例的权利范围。

[0248] (附图标记)

[0249] 190、570:存储器

[0250] 510:熵解码器

[0251] 124、544:帧间预测器

[0252] 1210、1310:虚拟帧生成器

[0253] 相关申请的交叉引用

[0254] 本申请要求于2020年9月28日提交的申请号为10-2020-0126111的韩国专利申请和于2021年9月28日提交的申请号为10-2021-0127795的韩国专利申请的优先权,这两个韩国专利申请的全部公开内容通过引用并入本文。

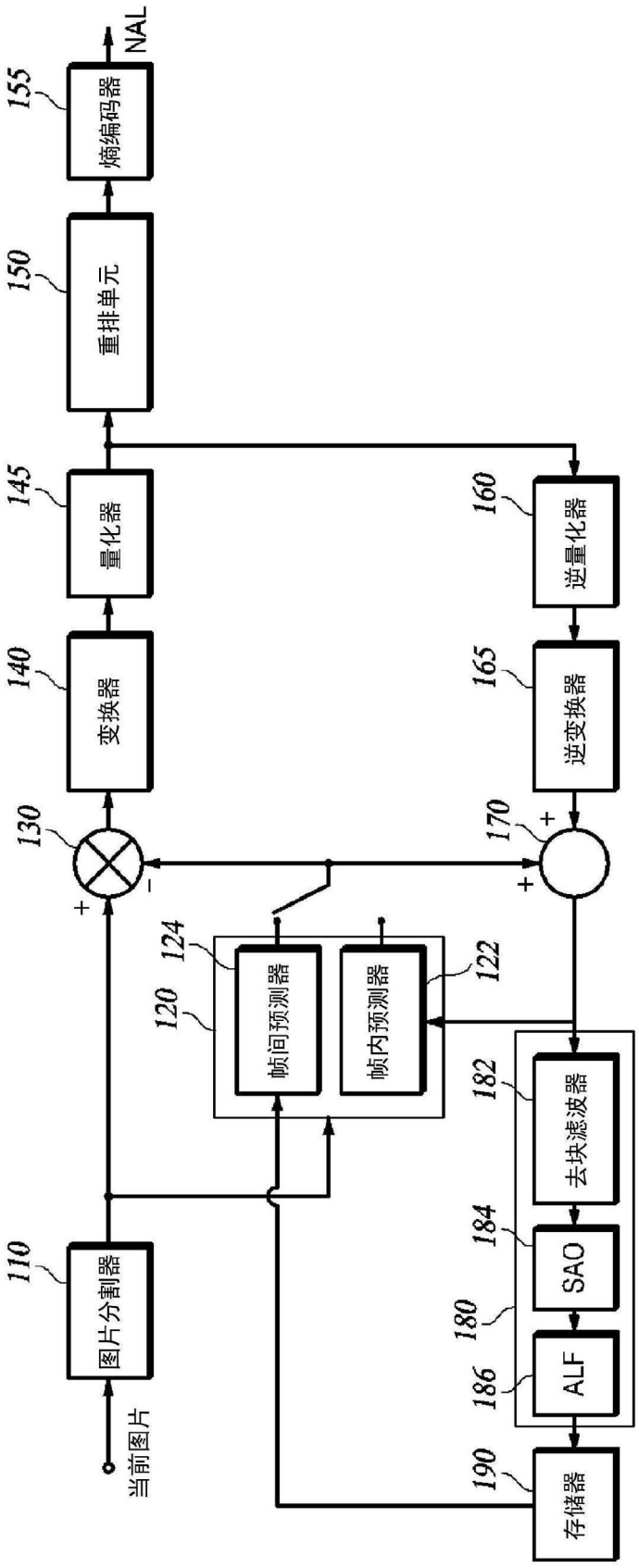

NAL
155
熵编码器
150
重排单元
145
量化器
140
变换器
130
160
逆量化器
165
逆变换器
170
124
帧间预测器
122
帧内预测器
120
182
去块滤波器
184
SAO
180
186
ALF
190
存储器
110
图片分割器
当前图片

图1

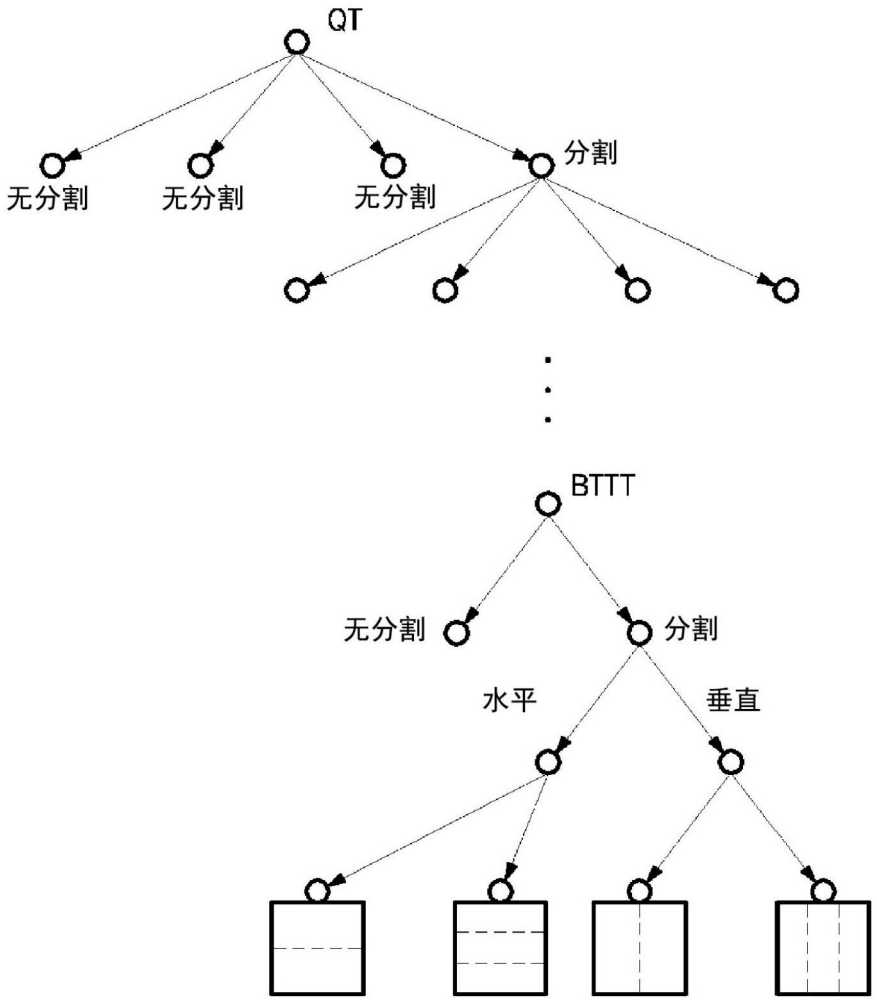



图2

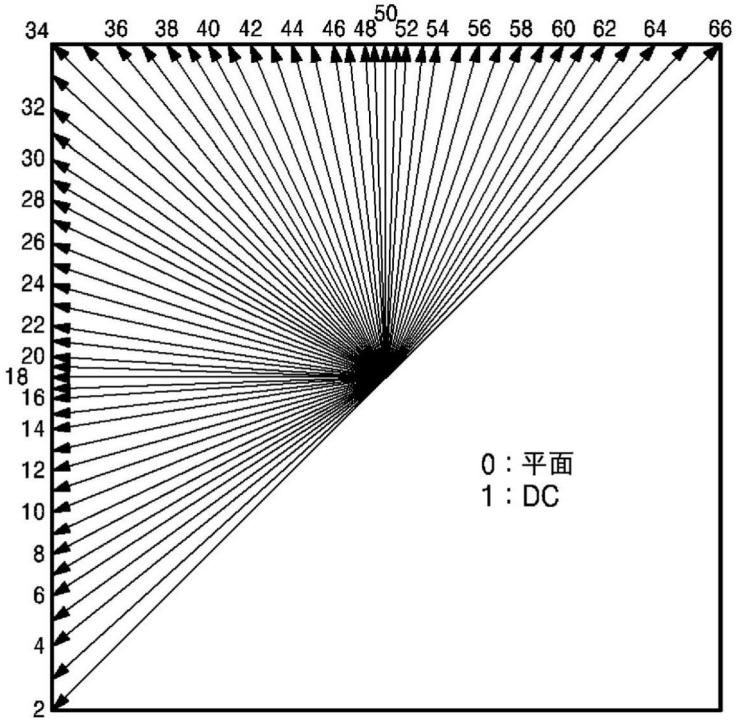

34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66
32 30 28 26 24 22 20 18 16 14 12 10 8 6 4 2
0：平面
1：DC

图3A

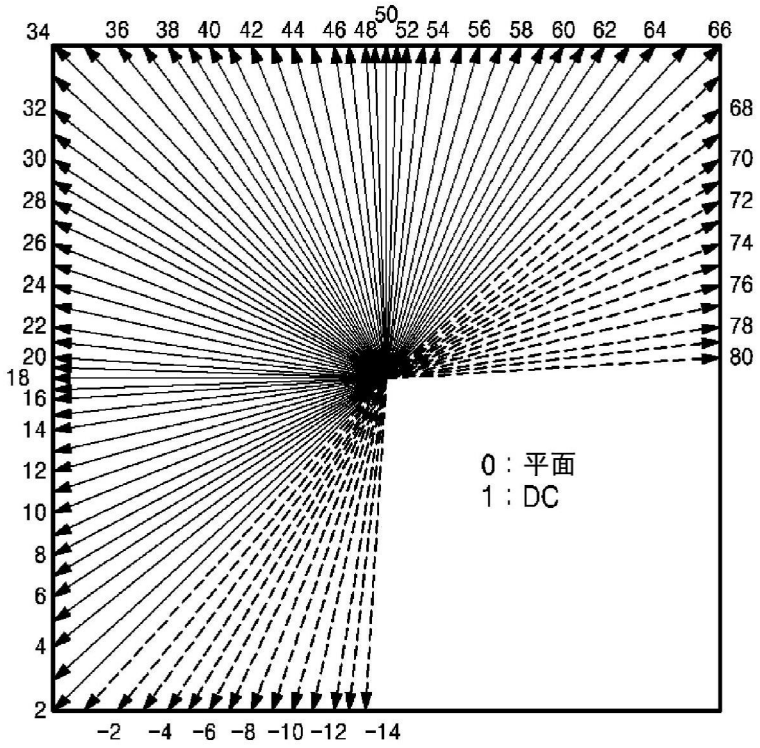

34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66
32 30 28 26 24 22 20 18 16 14 12 10 8 6 4 2
68 70 72 74 76 78 80
-2 -4 -6 -8 -10 -12 -14
0：平面
1：DC

图3B

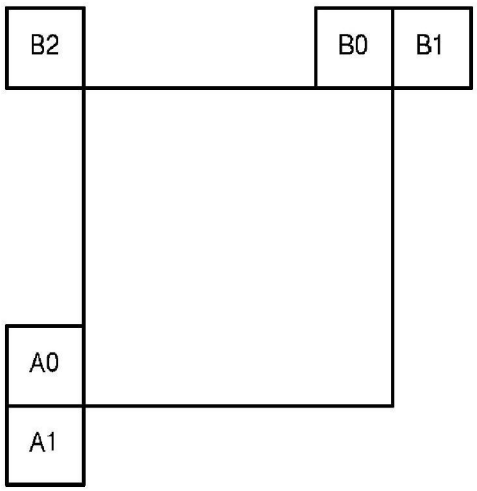

B2
B0
B1
A0
A1

图4

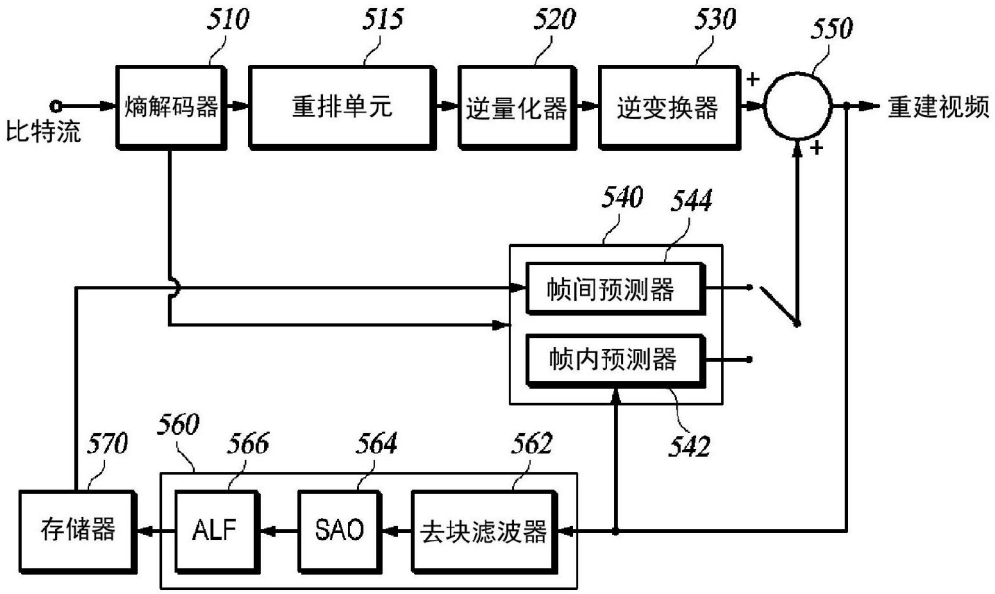

510
515
520
530
550
比特流
熵解码器
重排单元
逆量化器
逆变换器
重建视频
540
544
帧间预测器
帧内预测器
542
570
560
566
564
562
存储器
ALF
SAO
去块滤波器

图5

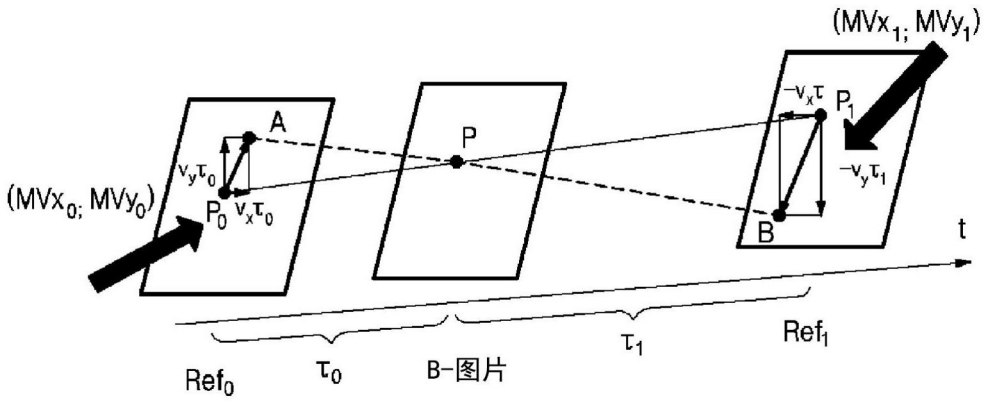

(MVx1; MVy1)
A
P
P1
B
t
(MVx0; MVy0)
P0
Ref0
τ0
B-图片
τ1
Ref1

图6

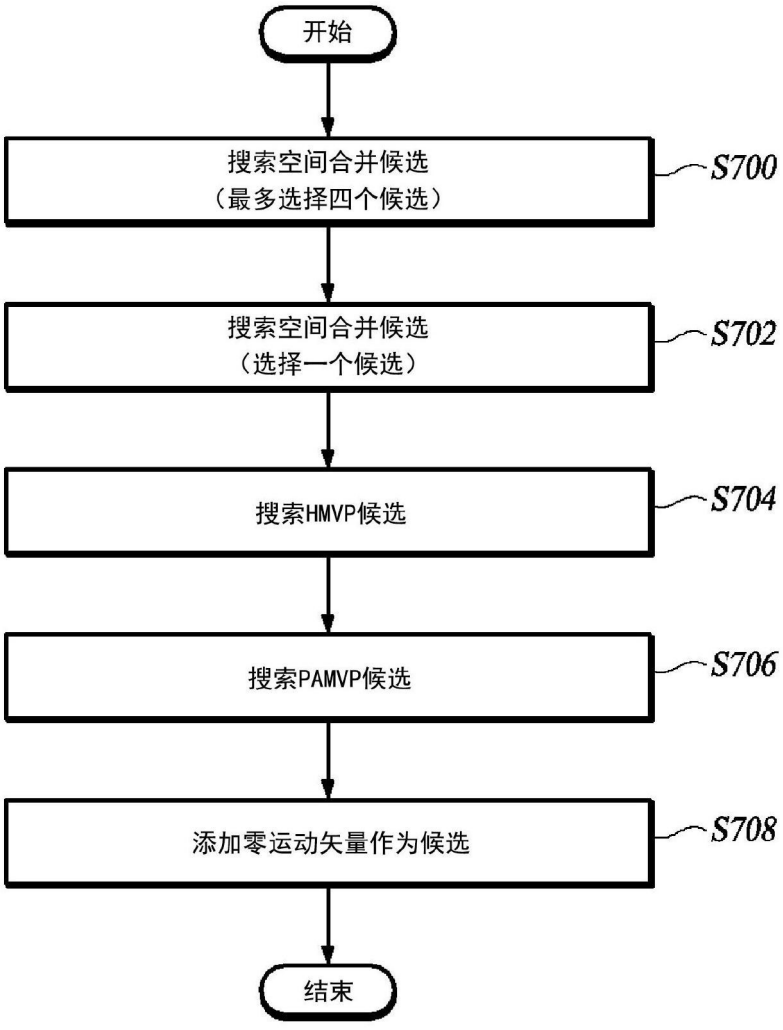

开始
搜索空间合并候选
（最多选择四个候选）
S700
搜索空间合并候选
（选择一个候选）
S702
搜索HMVP候选
S704
搜索PAMVP候选
S706
添加零运动矢量作为候选
S708
结束

图7

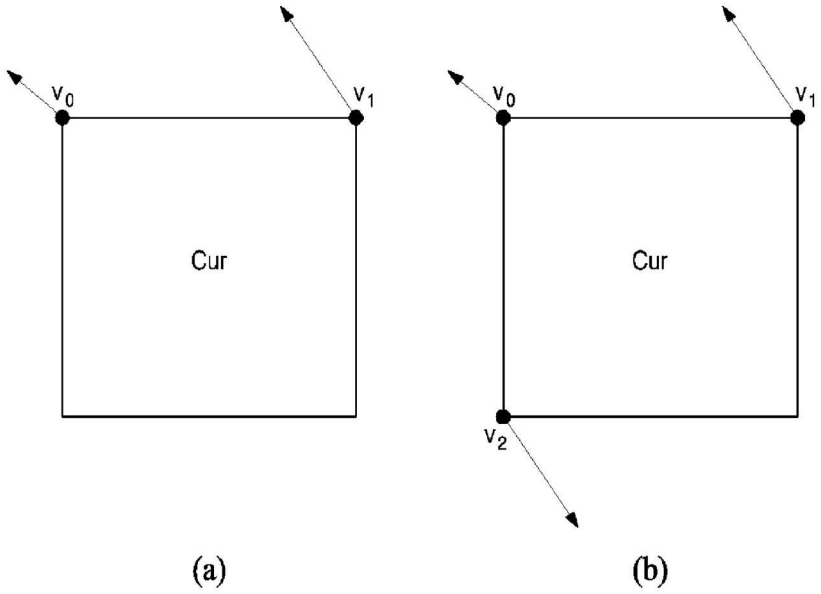



图8

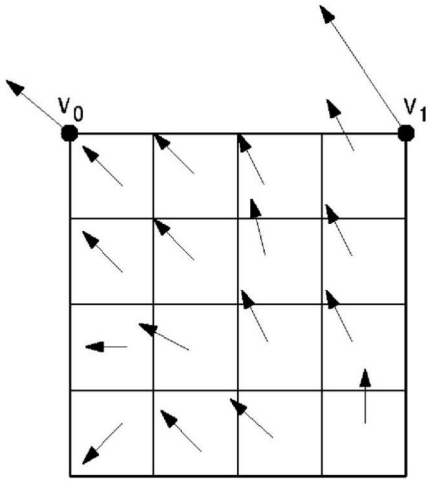



图9

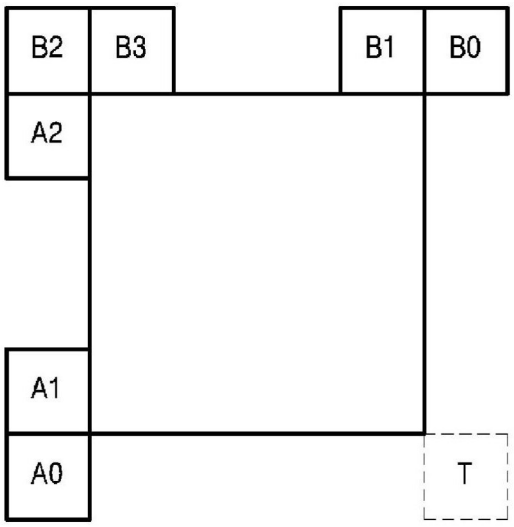



图10

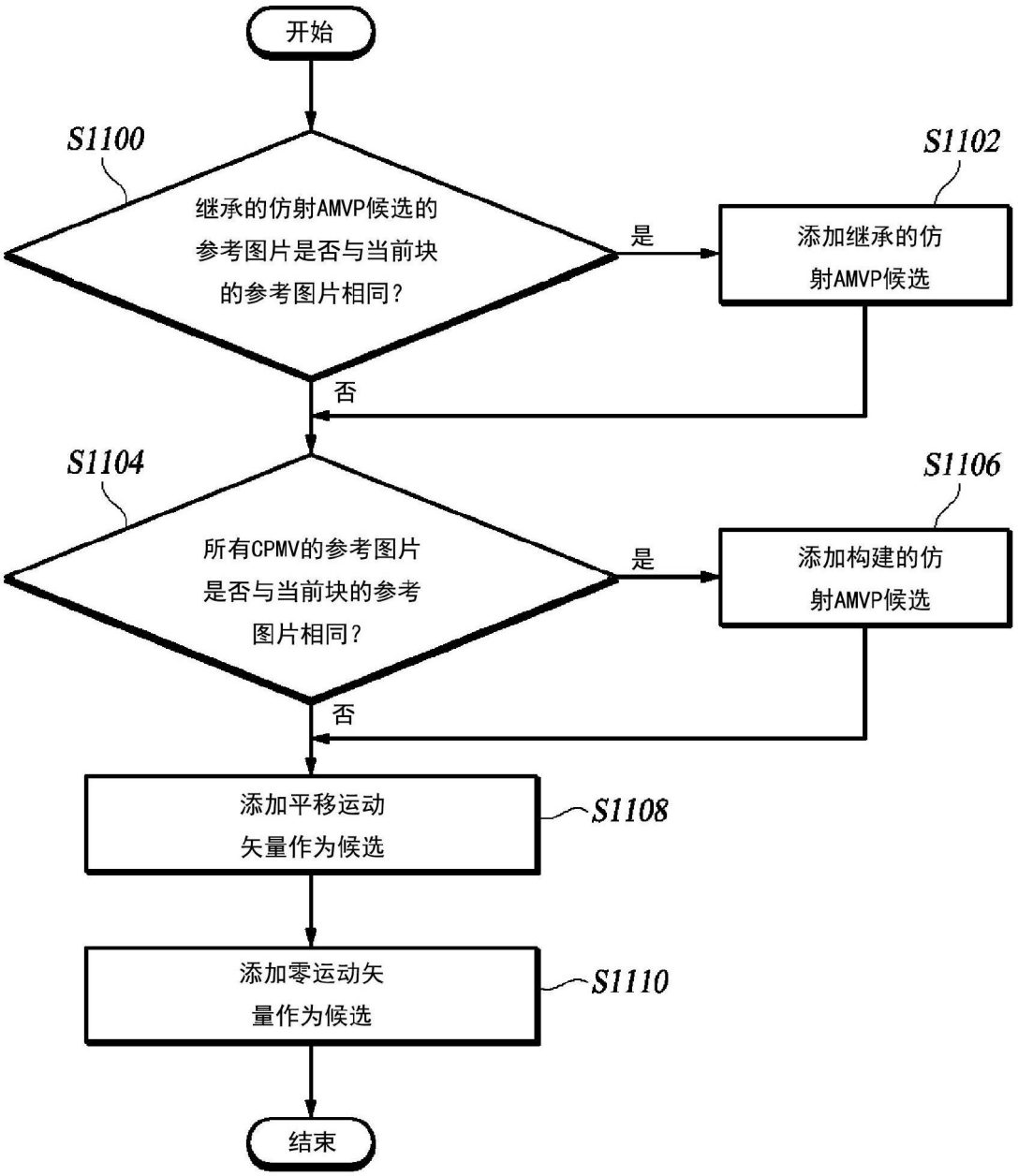

开始
S1100
继承的仿射AMVP候选的参考图片是否与当前块的参考图片相同?
是
S1102
添加继承的仿射AMVP候选
否
S1104
所有CPMV的参考图片是否与当前块的参考图片相同?
是
S1106
添加构建的仿射AMVP候选
否
添加平移运动矢量作为候选
S1108
添加零运动矢量作为候选
S1110
结束

图11

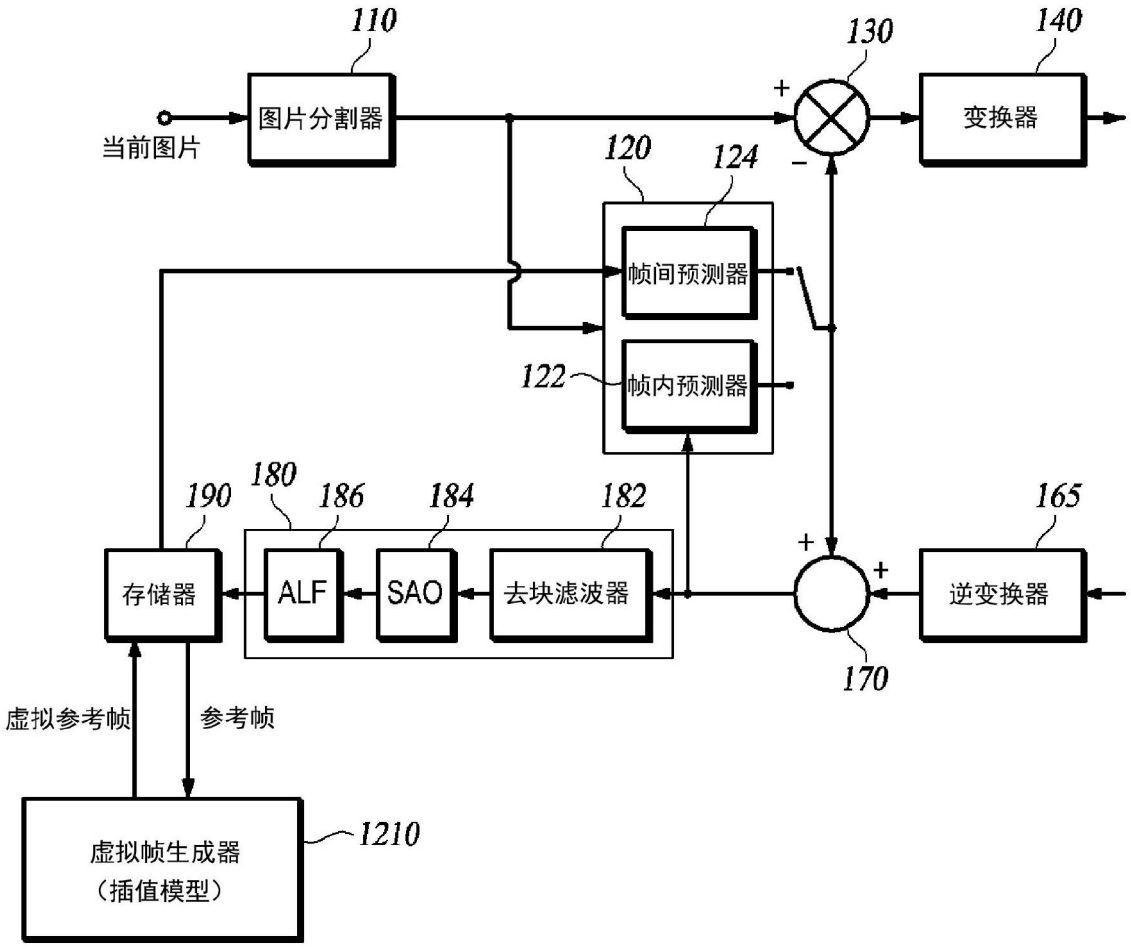

110
图片分割器
当前图片
130
140
变换器
120
124
帧间预测器
122
帧内预测器
190
180
186
184
182
存储器
ALF
SAO
去块滤波器
165
逆变换器
170
虚拟参考帧
参考帧
虚拟帧生成器
（插值模型）
1210

图12

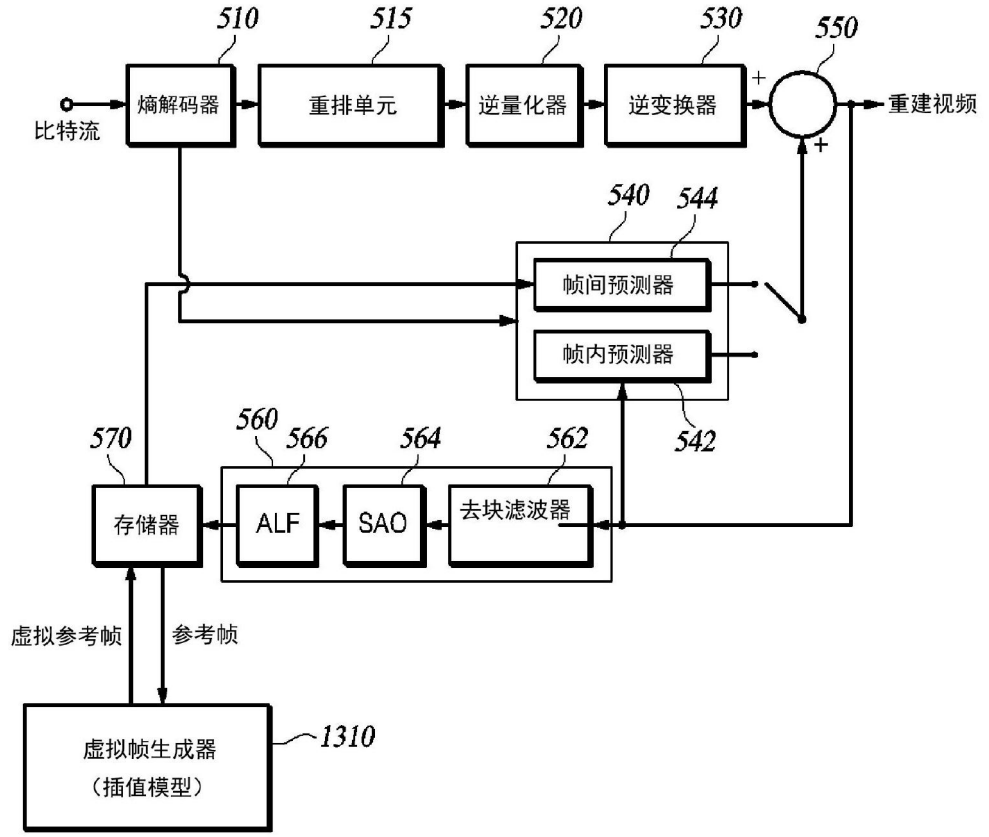

510
515
520
530
550
比特流
熵解码器
重排单元
逆量化器
逆变换器
重建视频
540
544
帧间预测器
帧内预测器
542
570
560
566
564
562
存储器
ALF
SAO
去块滤波器
虚拟参考帧
参考帧
虚拟帧生成器
（插值模型）
1310

图13

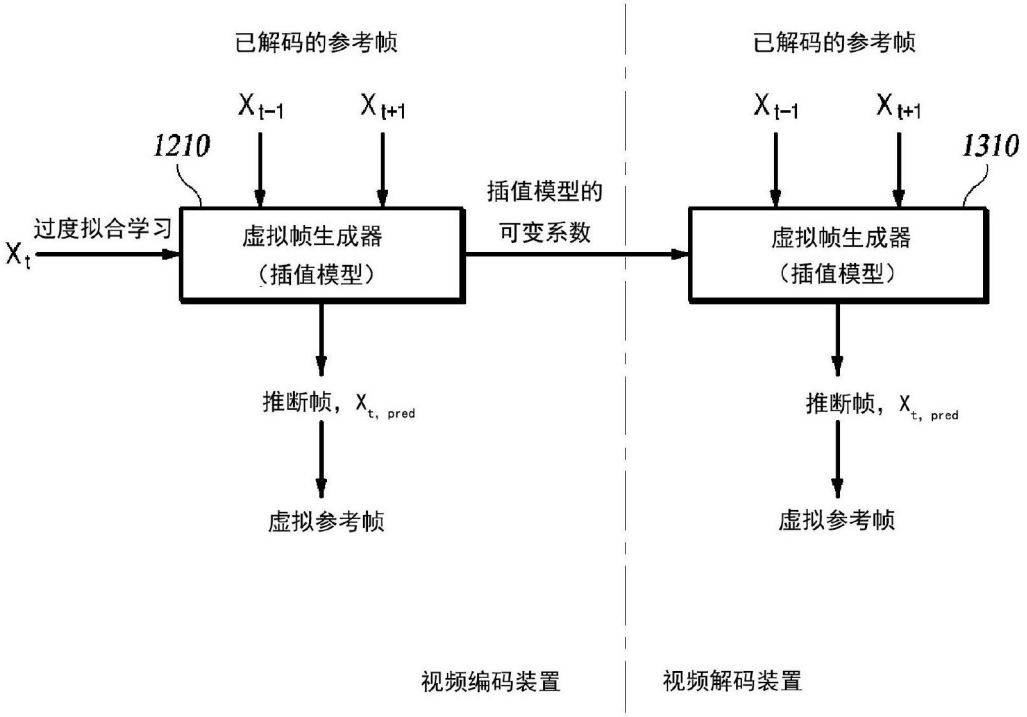
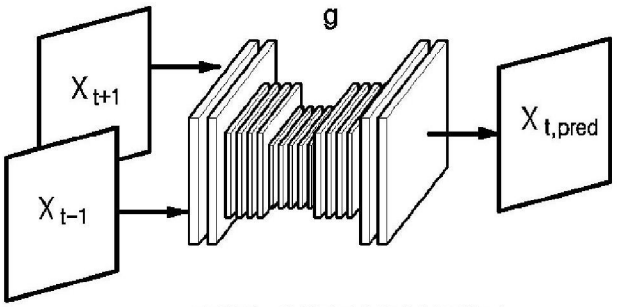

已解码的参考帧
已解码的参考帧
$X_{t-1}$
$X_{t+1}$
$X_{t-1}$
$X_{t+1}$
1210
1310
过度拟合学习
$X_t$
虚拟帧生成器
（插值模型）
插值模型的
可变系数
虚拟帧生成器
（插值模型）
推断帧，$X_{t,\ pred}$
推断帧，$X_{t,\ pred}$
虚拟参考帧
虚拟参考帧
视频编码装置
视频解码装置

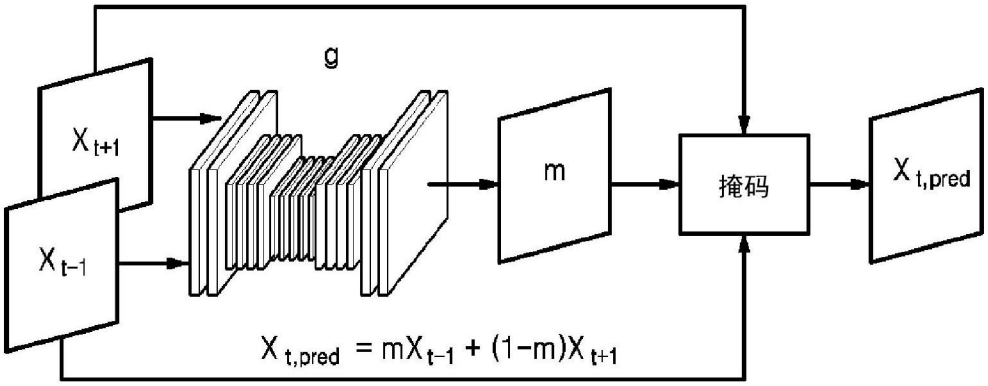
图14




直接生成推断帧的插值模型

$X_{t,pred} = g([X_{t-1}, X_{t+1}])$

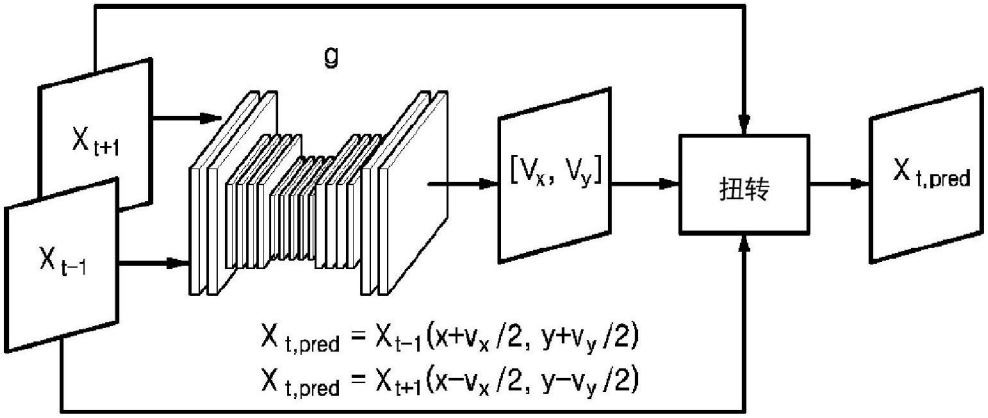
(a)




利用掩码的插值模型

$m = g([X_{t-1}, X_{t+1}])$

(b)




利用OF的插值模型

$[V_x, V_y] = g([X_{t-1}, X_{t+1}])$

(c)

图15

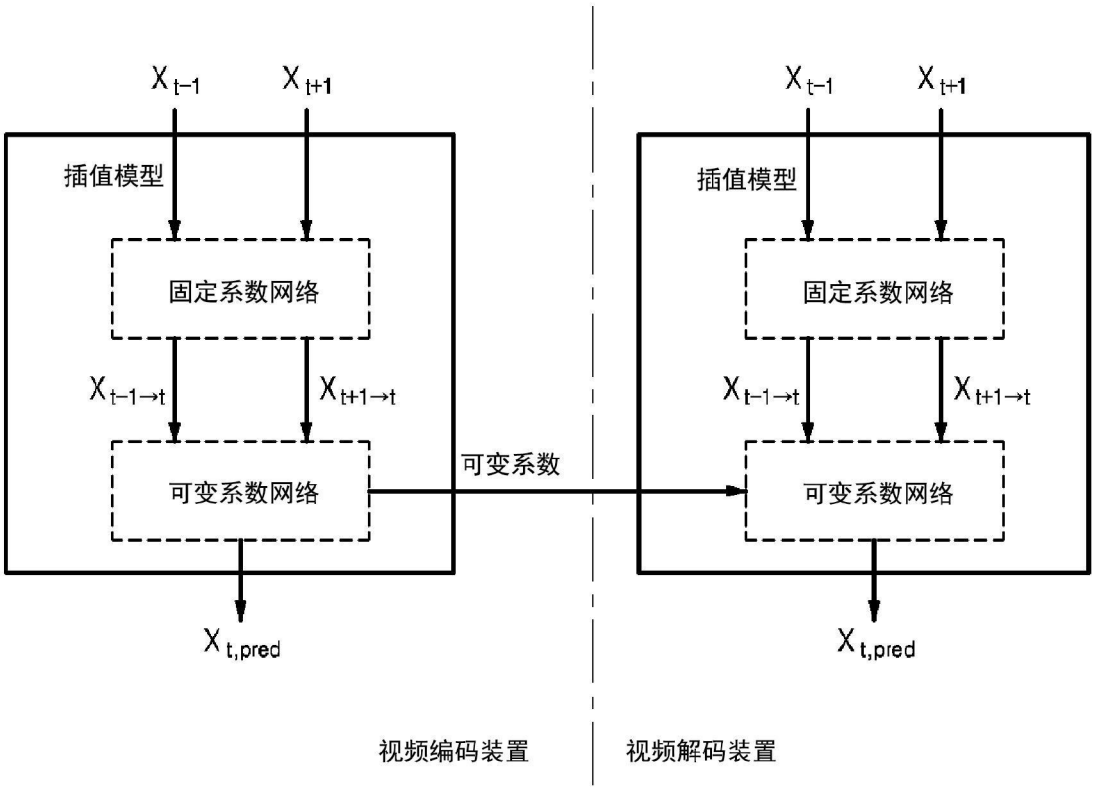

$X_{t-1}$
$X_{t+1}$
插值模型
固定系数网络
$X_{t-1\to t}$
$X_{t+1\to t}$
可变系数网络
$X_{t,pred}$
可变系数
$X_{t-1}$
$X_{t+1}$
插值模型
固定系数网络
$X_{t-1\to t}$
$X_{t+1\to t}$
可变系数网络
$X_{t,pred}$
视频编码装置
视频解码装置

图16

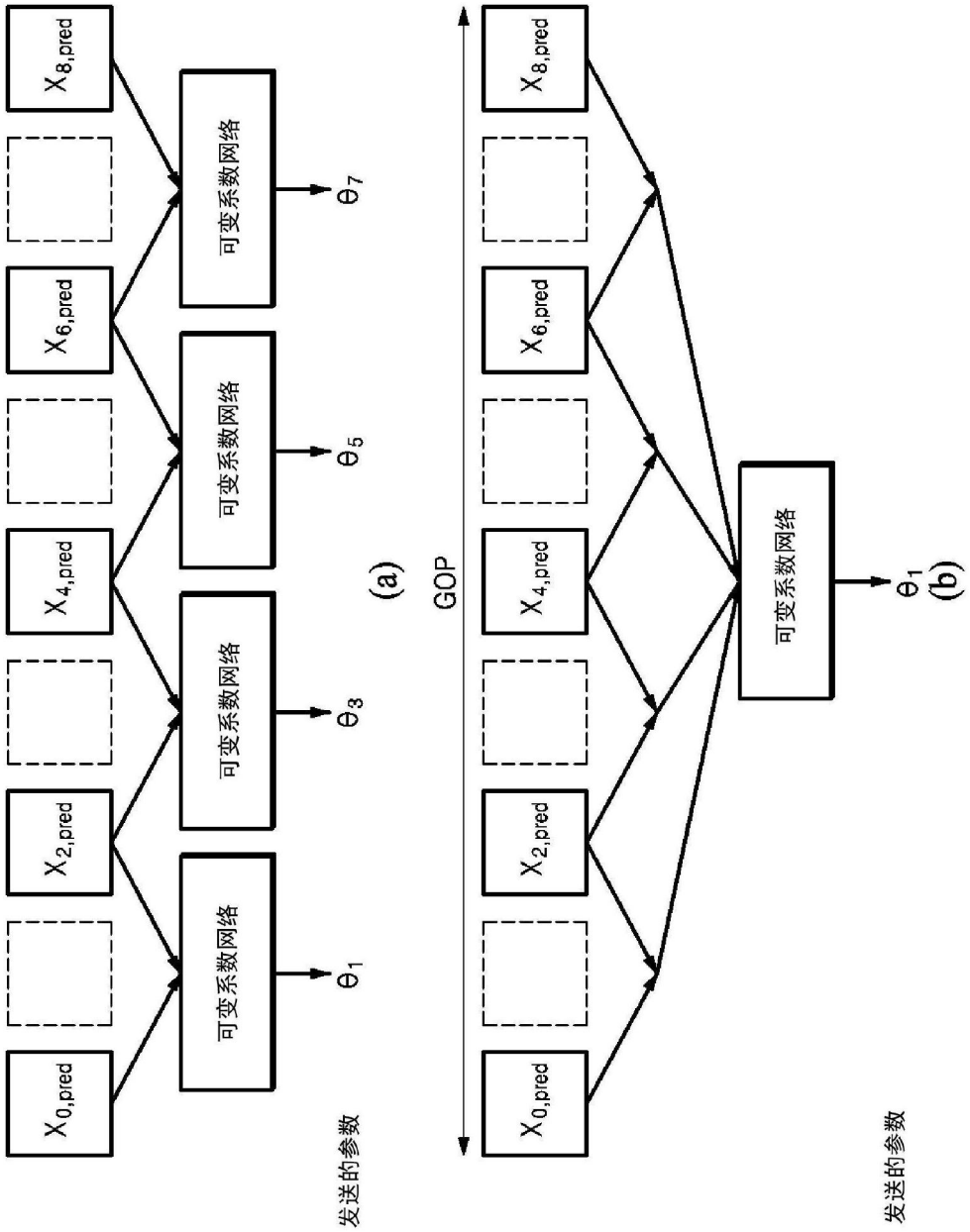

X0,pred
X2,pred
X4,pred
X6,pred
X8,pred
可变系数网络
θ1
θ3
θ5
θ7
发送的参数
(a)
GOP
(b)

图17

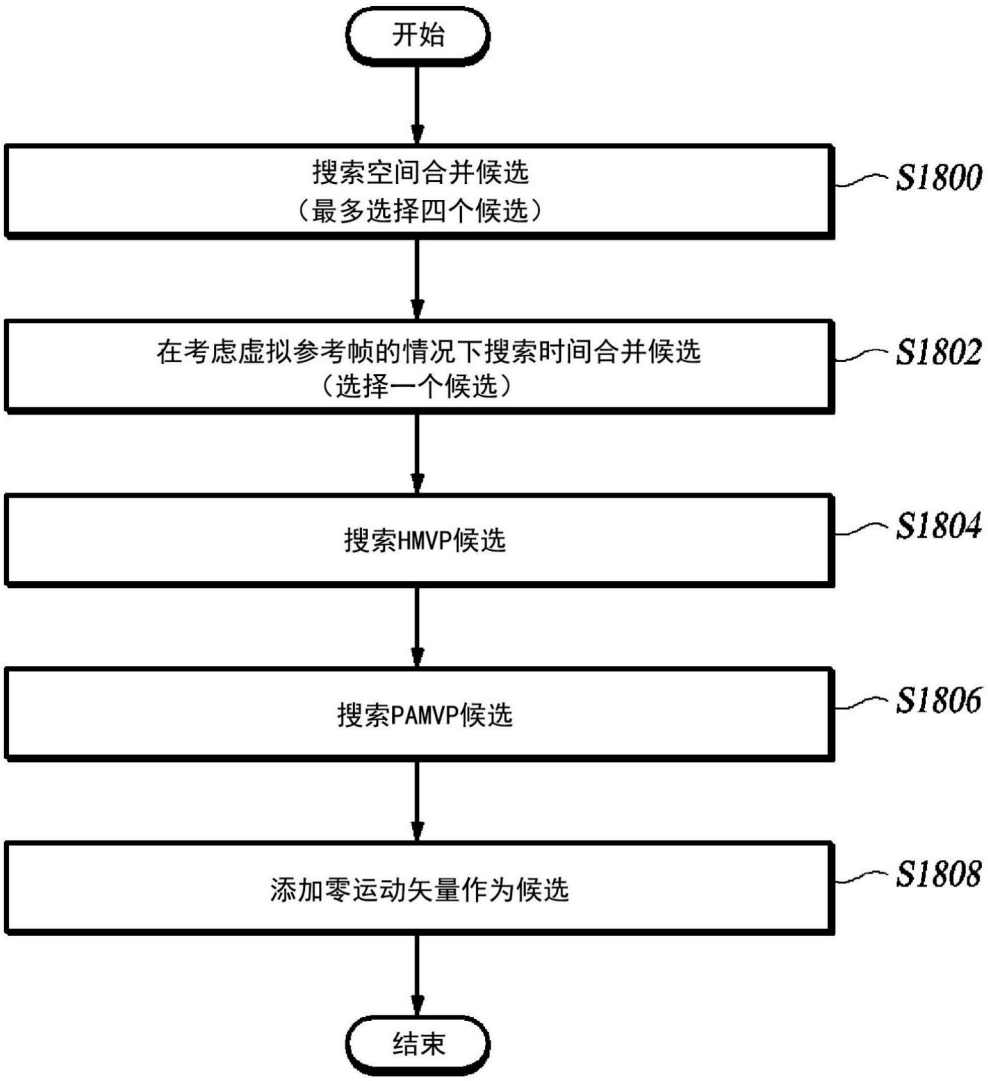

开始
搜索空间合并候选
（最多选择四个候选）
S1800
在考虑虚拟参考帧的情况下搜索时间合并候选
（选择一个候选）
S1802
搜索HMVP候选
S1804
搜索PAMVP候选
S1806
添加零运动矢量作为候选
S1808
结束

图18

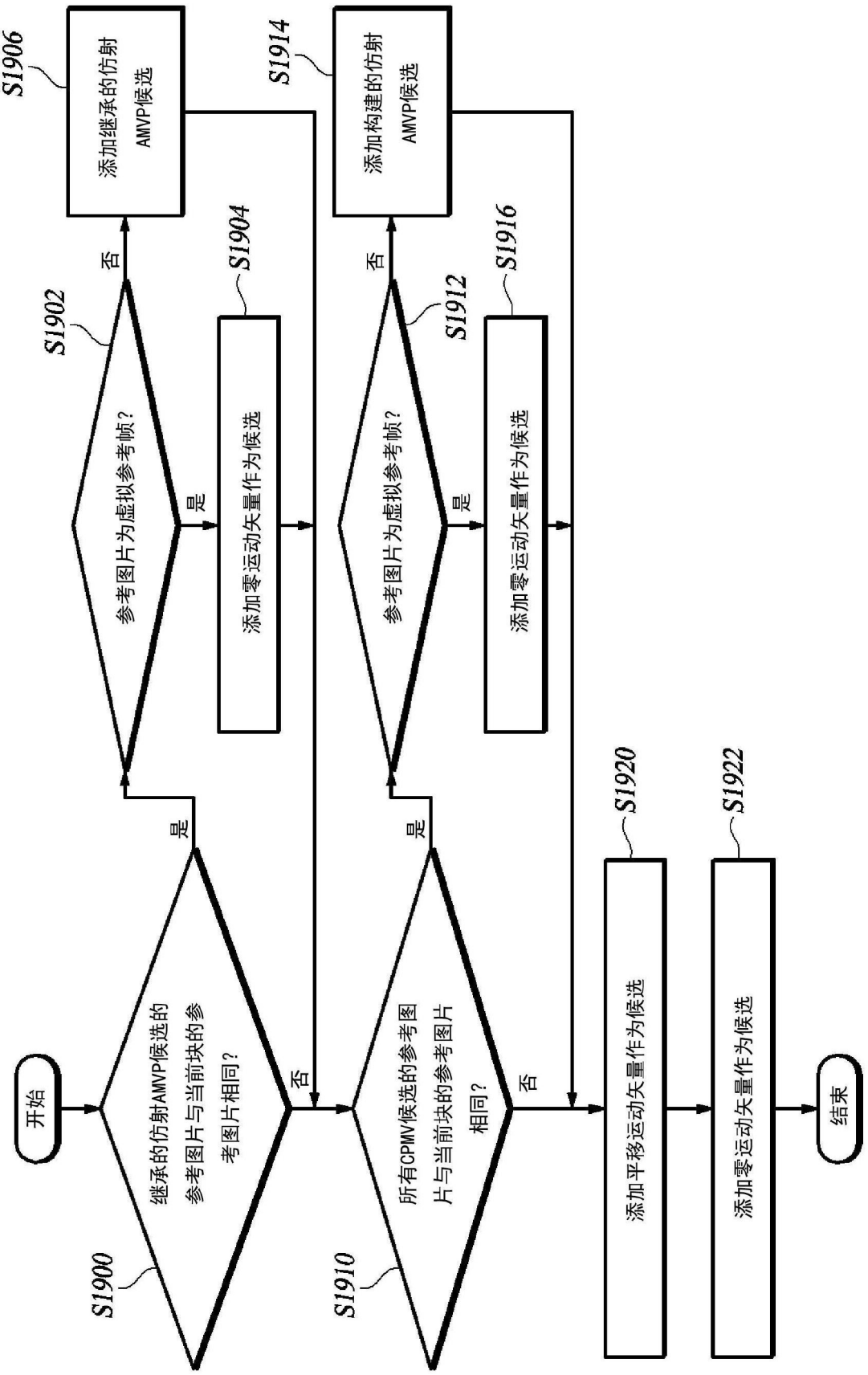

开始
S1900
继承的仿射AMVP候选的参考图片与当前块的参考图片相同?
是
否
S1902
参考图片为虚拟参考帧?
是
否
S1904
添加零运动矢量作为候选
S1906
添加继承的仿射AMVP候选
S1910
所有CPMV候选的参考图片与当前块的参考图片相同?
是
否
S1912
参考图片为虚拟参考帧?
是
否
S1914
添加构建的仿射AMVP候选
S1916
添加零运动矢量作为候选
S1920
添加平移运动矢量作为候选
S1922
添加零运动矢量作为候选
结束

图19