



(12) 发明专利申请

(10) 申请公布号 CN 118805377 A

(43) 申请公布日 2024. 10. 18

(21) 申请号 202380025092.6

(22) 申请日 2023.02.16

(30) 优先权数据

10-2022-0031336 2022.03.14 KR

10-2023-0018329 2023.02.10 KR

(85) PCT国际申请进入国家阶段日

2024.09.02

(86) PCT国际申请的申请数据

PCT/KR2023/002247 2023.02.16

(87) PCT国际申请的公布数据

W02023/177098 KO 2023.09.21

(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 许镇 朴胜煜

(74) 专利代理机构 北京康信知识产权代理有限公司 11240

专利代理师 陈鹏

(51) Int.Cl.

H04N 19/82 (2006.01)

H04N 19/86 (2006.01)

H04N 19/117 (2006.01)

H04N 19/105 (2006.01)

H04N 19/132 (2006.01)

H04N 19/423 (2006.01)

H04N 19/70 (2006.01)

G06T 9/00 (2006.01)

G06N 3/045 (2006.01)

G06N 3/0464 (2006.01)

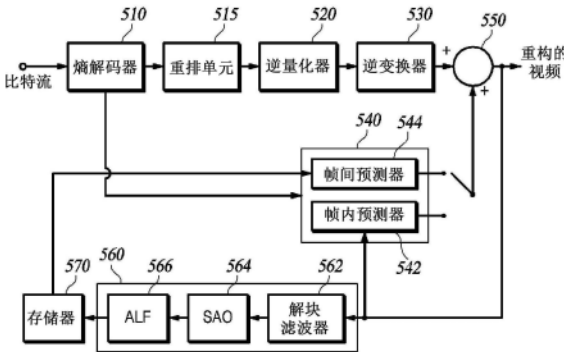
权利要求书2页 说明书19页 附图14页

(54) 发明名称

使用超分辨率环内滤波器的视频编码方法及设备

(57) 摘要

公开了使用超分辨率环内滤波器进行视频编码的方法和设备,本实施方式提供了在通过下采样对视频信号进行压缩并通过上采样恢复视频信号时,基于上采样或超分辨率技术使用环内滤波器进行视频编码的方法和设备。



1. 一种由视频解码设备执行的用于对重构帧进行上采样的方法,所述方法包括:
获得所述重构帧,其中,重构帧是已经由所述视频解码设备根据帧间预测预先对原始帧进行了重构的帧;
获得表示所述重构帧与所述原始帧之间的分辨率差的缩放值;
基于所述缩放值,通过将所述重构帧输入到基于深度学习的重采样环内滤波器中来生成经上采样的重构帧,其中,所述重采样环内滤波器替代包括解块滤波器、采样自适应偏移(SAO)滤波器和自适应环路滤波器的滤波器之一,或者其中,所述重采样环内滤波器插入在所述滤波器之间;以及
将所述滤波器和所述重采样环内滤波器生成的输出帧作为参考帧存储在解码图片缓冲器(DPB)中。
2. 根据权利要求1所述的方法,进一步包括:
从比特流解码指示所述重采样环内滤波器是被启用还是被禁用的标志;以及
当所述标志为真时,执行获得所述缩放值。
3. 根据权利要求2所述的方法,其中,获得所述缩放值包括:
当所述标志为真时,将所述缩放值设定为2。
4. 根据权利要求2所述的方法,其中,获得所述缩放值包括:
当所述标志为真时,从所述比特流解码指示所述缩放值的索引,且基于所述索引设定所述缩放值。
5. 根据权利要求1所述的方法,进一步包括:
从比特流解码指示多个神经网络候选之一的索引;以及
基于所述索引设定所述重采样环内滤波器,
其中,基于所述缩放值对所述神经网络候选中的每一者进行预训练。
6. 根据权利要求1所述的方法,其中,重采样环内滤波器通过使用转置卷积运算执行所述上采样,所述转置卷积运算是提取特征图的卷积运算的逆运算。
7. 根据权利要求1所述的方法,进一步包括:
当所述重采样环内滤波器插入在所述SAO滤波器与所述自适应环路滤波器之间时,将所述经上采样的重构帧输入到所述自适应环路滤波器中;以及
将由所述自适应环路滤波器生成的输出帧作为所述参考帧存储在所述解码图片缓冲器中。
8. 根据权利要求1所述的方法,进一步包括:
当所述重采样环内滤波器应用于当前帧时,在执行所述帧间预测之前,对存储在所述解码图片缓冲器中的所述参考帧进行缩放以在大小上适合所述当前帧。
9. 一种由视频编码设备执行的用于对重构帧进行上采样的方法,所述方法包括:
获得表示原始帧与所述重构帧之间的分辨率差的缩放值,其中,所述重构帧是已经由所述视频编码设备根据帧间预测预先对所述原始帧进行了重构的帧;以及
获得所述重构帧,
其中,当基于所述缩放值的所述重构帧的分辨率小于所述原始帧的分辨率时,所述方法进一步包括:
通过将所述重构帧输入到基于深度学习的重采样环内滤波器中来生成经上采样的重

构帧,其中,所述重采样环内滤波器替代包括解块滤波器、采样自适应偏移(SAO)滤波器和自适应环路滤波器的滤波器之一,或者其中,所述重采样环内滤波器插入在所述滤波器之间;以及

将所述滤波器和所述重采样环内滤波器生成的输出帧作为参考帧存储在解码图片缓冲器(DPB)中。

10.根据权利要求9所述的方法,其中,获得所述缩放值包括:

当对所述原始帧进行下采样时,将所述缩放值设定为2。

11.根据权利要求9所述的方法,其中,获得所述缩放值包括:

在率失真优化方面,确定所述缩放值。

12.根据权利要求9所述的方法,进一步包括:

通过将所述原始帧的水平维度和垂直维度除以所述缩放值来对所述原始帧进行下采样,其中下采样维度被截短或四舍五入到最接近的整数。

13.根据权利要求9所述的方法,进一步包括:

根据基于所述缩放值是否对所述原始帧进行下采样,确定指示是否使用所述重采样环内滤波器的标志;以及

对所述标志进行编码。

14.根据权利要求13所述的方法,其中,确定所述标志包括:

当对所述原始帧进行下采样时,将所述标志设定为真。

15.根据权利要求9所述的视频编码方法,进一步包括:

将指示多个神经网络候选之一的索引确定为所述重采样环内滤波器;以及

对所述索引进行编码,

其中,基于所述缩放值对所述神经网络候选中的每一者进行预训练。

16.一种计算机可读记录介质,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:

获得表示原始帧与重构帧之间的分辨率差的缩放值,其中,所述重构帧是已经由所述视频编码方法根据帧间预测预先对所述原始帧进行了重构的帧;以及

获得所述重构帧,

其中,当基于所述缩放值的所述重构帧的分辨率小于所述原始帧的分辨率时,所述视频编码方法进一步包括:

通过将所述重构帧输入到基于深度学习的重采样环内滤波器中来生成经上采样的重构帧,其中,所述重采样环内滤波器替代包括解块滤波器、采样自适应偏移(SAO)滤波器和自适应环路滤波器的滤波器之一,或者其中,所述重采样环内滤波器插入在所述滤波器之间;以及

将所述滤波器和所述重采样环内滤波器生成的输出帧作为参考帧存储在解码图片缓冲器(DPB)中。

使用超分辨率环内滤波器的视频编码方法及设备

技术领域

[0001] 本公开涉及使用超分辨率环内滤波器的视频编码方法和设备。

背景技术

[0002] 该部分中的陈述仅提供与本公开相关的背景信息,并且不一定构成现有技术。

[0003] 因为视频数据与音频或静止图像数据相比具有大量数据,所以在没有用于压缩的处理的情况下,视频数据需要大量硬件资源(包括存储器)来存储或传输视频数据。

[0004] 因而,编码器通常用于压缩和存储或传输视频数据。解码器接收压缩视频数据,对接收到的压缩视频数据进行解压缩,播放解压缩后的视频数据。视频压缩技术包括H.264/高级视频编码(AVC)、高效率视频编码(HEVC)和多用途视频编码(VVC),其与HEVC相比具有大约30%或更多的改善的编码效率。

[0005] 然而,由于图像大小、分辨率和帧速率逐渐增加,所以要编码的数据量也增加。因而,需要提供比现有压缩技术更高的编码效率和改善的图像增强效果的新压缩技术。最近,已经利用上采样或超分辨率技术来提高视频帧的质量。尤其在环内滤波中,需要考虑有效利用用于上采样或超分辨率增强的这些技术,以增加视频编码效率且增强视频质量。

发明内容

[0006] [技术问题]

[0007] 本发明试图提供一种视频编码方法和设备,所述方法和设备利用基于上采样或超分辨率技术的环内滤波器以通过下采样到压缩和上采样到其重构来处理视频信号来提高视频编码效率并增强视频质量。

[0008] [技术方案]

[0009] 本发明的至少一个方面提供一种由视频解码设备执行的用于对重构帧进行上采样的方法。所述方法包含获得所述重构帧。重构帧是已经由视频解码设备根据帧间预测预先重构的原始帧的帧。所述方法还包含获得表示所述重构帧与所述原始帧之间的分辨率差的缩放值。所述方法还包括:基于所述缩放值,通过将重构帧输入到基于深度学习的重采样环内滤波器中来生成上采样的重构帧。该重采样环内滤波器替代包括解块滤波器、采样自适应偏移(SAO)滤波器和自适应环路滤波器的滤波器之一,或者插入在滤波器之间。所述方法还包含将由所述滤波器和所述重采样环内滤波器生成的输出帧作为参考帧存储在解码图片缓冲器(DPB)中。

[0010] 本发明的另一方面提供一种由视频编码设备执行的用于对重构帧进行上采样的方法。所述方法包含获得表示原始帧与所述重构帧之间的分辨率差的缩放值。重构帧是视频编码设备已经根据帧间预测预先重构的原始帧的帧。所述方法包含获得所述重构帧。当基于所述缩放值的所述重构帧的分辨率小于所述原始帧的分辨率时,所述方法进一步包括:通过将所述重构帧输入到基于深度学习的重采样环内滤波器中,来生成上采样的重构帧。该重采样环内滤波器替代包括解块滤波器、采样自适应偏移(SAO)滤波器和自适应环路

滤波器的滤波器之一,或者插入在滤波器之间。所述方法进一步包括将一旦由所述滤波器和所述重采样环内滤波器生成的输出帧作为参考帧存储在解码图片缓冲器 (DPB) 中。

[0011] 本发明的又一方面提供一种存储由视频编码方法生成的比特流的计算机可读记录介质。视频编码方法包括:获得表示原始帧和重构帧之间的分辨率差的缩放值。重构帧是已经通过视频编码方法根据帧间预测预先重构原始帧的帧。所述视频编码方法还包括获得所述重构帧。当基于所述缩放值的所述重构帧的分辨率小于所述原始帧的分辨率时,所述视频编码方法进一步包括:通过将所述重构帧输入到基于深度学习的重采样环内滤波器中,生成经过上采样的重构帧。该重采样环内滤波器替代包括解块滤波器、采样自适应偏移 (SAO) 滤波器和自适应环路滤波器的滤波器之一,或者插入在滤波器之间。所述视频编码方法还包括:将一旦由所述滤波器和所述重采样环内滤波器生成的输出帧作为参考帧存储在解码图片缓冲器 (DPB) 中。

[0012] [有益效果]

[0013] 如上所述,本发明提供在通过下采样到压缩且上采样到其重构来处理视频信号时基于上采样或超分辨率技术利用环内滤波器的视频编码方法和设备。由此,视频编码方法和设备改善了视频质量并提高了视频编码效率。

附图说明

[0014] 图1是可实施本发明的技术的视频编码设备的框图。

[0015] 图2说明用于使用四叉树加二叉树三叉树 (QTBT) 结构来分割块的方法。

[0016] 图3a和3b示出包括广角帧内预测模式的多个帧内预测模式。

[0017] 图4示出当前块的邻近块。

[0018] 图5是可实施本发明的技术的视频解码设备的框图。

[0019] 图6是示出卷积层的操作的图。

[0020] 图7是示出去卷积层的操作的示图。

[0021] 图8是示出池化层的操作的图。

[0022] 图9是示出单图像超分辨率 (SISR) 网络的示图。

[0023] 图10是示出了在SISR中利用的残差块的示图。

[0024] 图11是示出自适应环路滤波器 (ALF) 的示图。

[0025] 图12是应用了参考图片重采样 (RPR) 技术的视频解码设备的框图。

[0026] 图13是根据本公开的至少一个实施方式的包括重采样环内滤波器 (RIF) 的视频解码设备的框图。

[0027] 图14是示出转置卷积运算的图。

[0028] 图15是示出上池化运算的图。

[0029] 图16是示出根据本公开的至少一个实施方式的利用神经网络超分辨率 (NNSR) 的上采样技术的图。

[0030] 图17是图示根据本公开的至少一个实施方式的RIF的应用的位置的图。

[0031] 图18是根据本公开内容的至少一个实施方式的由视频编码设备执行的用于对重构帧进行上采样的方法的流程图。

[0032] 图19是根据本发明的至少一个实施方式的由视频解码设备执行的用于对重构帧

进行上采样的方法的流程图。

具体实施方式

[0033] 在下文中,参照附图详细描述本公开的一些实施方式。在以下描述中,相同的参考标号表示相同的元件,尽管元件在不同的附图中示出。此外,在一些实施方式的以下描述中,为了清楚和简洁的目的,当认为使本公开的主题模糊时,可以省略相关已知组件和功能的详细描述。

[0034] 图1是可以实现本公开的技术的视频编码设备的框图。在下文中,参考图1的图解,描述视频编码设备和该设备的组件。

[0035] 编码设备可包括图片分离器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、环路滤波器单元180和存储器190。

[0036] 编码设备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,并且还可以实现微处理器以执行与每个组件相对应的软件的功能。

[0037] 一个视频由包括多个图片的一个或多个序列构成。每个画面被划分为多个区域,并对每个区域执行编码。例如,将一个图片分裂成一个或一个以上瓦片或/和切片。在此,一个或多个贴片可以被定义为贴片组。每一瓦片或/和切片分裂成一或多个编码树单元(CTU)。另外,每一CTU通过树结构分裂成一或多个编码单元(CU)。将应用于每一编码单元(CU)的信息编码为CU的语法,且将共同应用于包含于一个CTU中的CU的信息编码为CTU的语法。此外,共同应用于一个片中的所有块的信息被编码为片报头的语法,并且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。此外,将多个图片共同参考的信息编码到序列参数集(SPS)。另外,一或多个SPS共同参考的信息被编码到视频参数集(VPS)。进一步,共同应用于一个瓦片或瓦片组的信息也可编码为瓦片或瓦片组报头的语法。包括在SPS、PPS、片头部、贴片或贴片组头部中的语法可以被称为高级语法。

[0038] 图片分割器110确定编码树单元(CTU)的大小。关于CTU的大小(CTU大小)的信息被编码为SPS或PPS的语法并且被递送到视频解码设备。

[0039] 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),然后通过使用树结构递归分割CTU。树结构中的叶节点变成编码单元(CU),其是编码的基本单元。

[0040] 树结构可以是四叉树(QT),其中,较高节点(或父节点)被分成具有相同大小的四个较低节点(或子节点)。树结构也可以是二进制树(BT),其中较高节点被分成两个较低节点。树结构还可以是三叉树(TT),其中,较高节点以1:2:1的比率被划分为三个较低节点。树结构还可以是QT结构、BT结构和TT结构中的两个或多个结构被混合的结构。例如,可以使用四叉树加二叉树(QTBT)结构或者可以使用四叉树加二叉树三叉树(QTBTTT)结构。这里,将二叉树三叉树(BTTT)添加到树结构以被称为多类型树(MTT)。

[0041] 图2是用于描述通过使用QTBTTT结构分割块的方法的示图。

[0042] 如图2所示,CTU可以首先被拆分成QT结构。四叉树分割可以是递归的,直到分割块的大小达到在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

被划分为下层的四个节点的第一标志 (QT_split_flag) 由熵编码器155编码,并被传输到视频解码设备。当QT的叶节点不大于在BT中允许的根节点的最大块大小 (MaxBTSIZE) 时,叶节点可进一步被划分为BT结构或TT结构中的至少一个。在BT结构和/或TT结构中可存在多个分割方向。例如,可存在两个方向,即,相应节点的块被水平划分的方向和相应节点的块被垂直划分的方向。如图2所示,当MTT划分开始时,指示节点是否被划分的第二标志 (mtt_split_flag) 以及另外指示划分方向 (垂直或水平) 的标志和/或如果节点被划分则指示划分类型 (二进制或三进制) 的标志由熵编码器155编码并用信号通知视频解码设备。

[0043] 可替代地,在编码指示每个节点是否被划分为下层的四个节点的第一标志 (QT_split_flag) 之前,指示节点是否被划分的CU分割标志 (split_cu_flag) 也可被编码。当CU分割标志 (split_cu_flag) 的值指示每一节点未分裂时,对应节点的块变为分裂树结构中的叶节点且变为作为编码的基本单元的CU。当CU分割标志 (split_cu_flag) 的值指示每个节点被分割时,视频编码设备首先通过上述方案开始编码第一标志。

[0044] 当QTBT被用作树结构的另一实例时,可存在两种类型,即,相应节点的块被水平划分为具有相同大小的两个块的类型 (即,对称水平划分) 和相应节点的块被垂直划分为具有相同大小的两个块的类型 (即,对称垂直划分)。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由熵编码器155编码,并被递送到视频解码设备。同时,可以另外呈现将对应节点的块分割为彼此不对称的两个块的类型。非对称形式可包括其中对应节点的块被划分为具有1:3的大小比的两个矩形块的形式,或者还可包括其中对应节点的块被沿对角线方向划分的形式。

[0045] CU可根据来自CTU的QTBT或QTBT分裂而具有不同大小。在下文中,对应于待编码或解码的CU (即,QTBT的叶节点) 的块被称为“当前块”。由于采用QTBT拆分,当前块的形状除了正方形形状外,还可以是矩形形状。

[0046] 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124。

[0047] 一般来说,图片中的当前块中的每一者可经预测性编码。通常,当前块的预测可以通过使用帧内预测技术 (使用来自包括当前块的图片的数据) 或者帧间预测技术 (使用来自在包括当前块的图片之前被编译的图片的数据) 来执行。帧间预测包括单向预测和双向预测两者。

[0048] 帧内预测器122通过使用在包括当前块的当前图片中位于当前块的邻居上的像素 (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中所示,多个帧内预测模式可以包括包含平面模式和DC模式的2个非定向模式,并且可以包括65个定向模式。相邻像素和将被使用的算术等式根据每个预测模式被不同地定义。

[0049] 为了对具有矩形形状的当前块进行有效的定向预测,可以附加地使用如图3B中的虚线箭头图示的定向模式 (#67至#80, 帧内预测模式#-1至#-14)。定向模式可被称作“广角帧内预测模式”。在图3B中,箭头指示用于预测的对应参考样本并且不表示预测方向。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在没有附加比特传输的情况下以与特定方向模式相反的方向执行预测的模式。在这种情况下,在广角帧内预测模式中,可通过具有矩形形状的当前块的宽度和高度的比率确定可用于当前块的一些广角帧内预测模式。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,可使用具有大于-135度的角度的广角帧内预测模式。

[0050] 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。在一些实施例中,帧内预测器122可通过使用多个帧内预测模式对当前块进行编码,且还可从测试模式中选择将使用的适当帧内预测模式。例如,帧内预测器122可通过使用针对多个测试的帧内预测模式的速率-失真分析来计算速率-失真值,并且还可在测试模式中选择具有最佳速率-失真特征的帧内预测模式。

[0051] 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,并通过使用相邻像素(参考像素)和根据选择的帧内预测模式确定的算术等式来预测当前块体。关于所选择的帧内预测模式的信息由熵编码器155编码,并被递送到视频解码设备。

[0052] 帧间预测器124通过使用运动补偿处理来生成用于当前块的预测块。帧间预测器124在比当前画面更早地编码和解码的参考画面中搜索与当前块最相似的块,并且通过使用搜索到的块生成用于当前块的预测块。另外,生成运动矢量(MV),其对应于在当前图片中的当前块和在参考图片中的预测块之间的位移。通常,对亮度分量执行运动估计,并且基于亮度分量计算的运动矢量被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,并被递送到视频解码设备。

[0053] 帧间预测器124还可对参考图片或参考块执行内插以便增加预测的准确度。换句话说,两个连续整数采样之间的子采样通过将滤波器系数应用于包括两个整数采样的多个连续整数采样而被内插。当针对插值的参考画面执行搜索与当前块最相似的块的处理时,可以针对运动矢量表示不是整数采样单位精度,而是十进制单位精度。可针对待编码的每一目标区域(例如,例如切片、瓦片、CTU、CU等单元)不同地设定运动矢量的精度或分辨率。当应用此自适应运动矢量分辨率(AMVR)时,应针对每一目标区域用信号发出关于待应用于每一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表示将在下面描述的运动矢量差的精度的信息。

[0054] 同时,帧间预测器124可通过使用双向预测来执行帧间预测。在双向预测的情况下,使用两个参考图片和表示在每个参考图片中最类似于当前块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)选择第一参考图片和第二参考图片。帧间预测器124还搜索与相应参考图片中的当前块最相似的块以生成第一参考块和第二参考块。另外,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。另外,包括关于用于预测当前块的两个参考图片的信息并且包括关于两个运动矢量的信息的运动信息被递送到熵编码器155。在这里,参考图片列表0可以由在预先重构的图片当中的按照显示顺序的在当前图片之前的图片组成,并且参考图片列表1可以由在预先重构的图片当中的按照显示顺序的在当前图片之后的图片组成。然而,虽然不特别限于此,但是按照显示顺序在当前图片之后的预重构的图片可以附加地包括在参考图片列表0中。相反地,在当前图片之前的预重构的图片也可以被附加地包括在参考图片列表1中。

[0055] 为了最小化用于编码运动信息所消耗的比特数量,可以使用各种方法。

[0056] 例如,当当前块的参考图片和运动矢量与相邻块的参考图片和运动矢量相同时,能够识别相邻块的信息被编码以将当前块的运动信息递送到视频解码设备。这种方法被称为合并模式。

[0057] 在合并模式中,帧间预测器124从当前块的相邻块中选择预定数量的合并候选块(在下文中,称为“合并候选”)。

[0058] 作为用于推导合并候选的相邻块,如在图4中图示的,可以使用在当前图片中相邻于当前块的左块A0、左下块A1、上块B0、右上块B1、和左上块B2中的全部或一些。进一步,定位在参考图片内(可以与用于预测当前块的参考图片相同或不同)的除当前块定位在其处的当前图片之外的块也可以用作合并候选。例如,在参考图片内具有当前块的位于同一地点的块或相邻于位于同一地点的块的块可另外用作合并候选。如果通过上文所描述的方法选择的合并候选的数量小于预设数量,那么将零矢量添加到合并候选。

[0059] 帧间预测器124通过使用相邻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要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。生成的合并索引信息由熵编码器155编码,并被递送到视频解码设备。

[0060] 合并跳跃模式是合并模式的特殊情况。在量化之后,当用于熵编码的所有变换系数接近于零时,仅传输相邻块选择信息而不传输残余信号。通过使用合并跳跃模式,对于具有轻微运动的图片、静止图片、屏幕内容图片等等可以实现相对高的编码效率。

[0061] 在下文中,合并模式和合并跳跃模式被统称为合并/跳跃模式。

[0062] 用于编码运动信息的另一方法是高级运动矢量预测(AMVP)模式。

[0063] 在AMVP模式中,帧间预测器124通过使用当前块的相邻块推导用于当前块的运动矢量的运动矢量预测子候选。作为用于推导运动矢量预测子候选的相邻块,可以使用与在图4中图示的当前图片中的当前块相邻的左块A0、左下块A1、上块B0、右上块B1和左上块B2中的全部或一些。此外,定位在参考图片内(可以与用于预测当前块的参考图片相同或不同)的除当前块定位在其处的当前图片之外的块也可以用作用于推导运动矢量预测子候选的相邻块。例如,可以使用在参考图片之内与当前块的共处一地的块或者相邻于共处一地的块的块。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,则零矢量被添加到运动矢量候选。

[0064] 帧间预测器124通过使用相邻块的运动矢量推导运动矢量预测器候选,并通过使用运动矢量预测器候选确定用于当前块的运动矢量的运动矢量预测器。另外,通过从当前块的运动矢量减去运动矢量预测子来计算运动矢量差。

[0065] 运动矢量预测子可通过将预定义的函数(例如,中心值和平均值计算等)应用于运动矢量预测子候选来获得。在这种情况下,视频解码设备还知道预定义的函数。此外,由于用于推导运动矢量预测子候选的相邻块是编码和解码已经完成的块,所以视频解码设备还可能已经知道相邻块的运动矢量。因此,视频编码设备不需要对用于识别运动矢量预测器候选的信息进行编码。因而,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。

[0066] 同时,还可以通过选择运动矢量预测器候选中的任何一个的方案来确定运动矢量预测器。在这种情况下,用于识别选择的运动矢量预测器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联合地另外编码。

[0067] 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。

[0068] 变换器140将具有空间域的像素值的残差块中的残差信号转换为频域的变换系数。变换器140可通过将残差块的总大小用作变换单元来对残差块中的残差信号进行变换,或者还可将残差块分割成多个子块,并可通过将子块用作变换单元来执行变换。可选地,所述残差块被划分为两个子块,分别为变换区域和非变换区域,以仅使用所述变换区域子块作为变换单元对所述残差信号进行变换。在此,变换区域子块可以是具有基于水平轴(或垂直轴)的1:1的大小比的两个矩形块之一。在这种情况下,标志(cu_sbt_flag)指示仅子块被变换,方向性(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编码并用信号通知视频解码设备。此外,变换区域子块的大小可基于水平轴(或垂直轴)具有1:3的大小比率。在这种情况下,划分相应划分的标志(cu_sbt_quad_flag)另外由熵编码器155编码,并用信号通知视频解码设备。

[0069] 同时,变换器140可在水平方向和垂直方向上单独地对残差块执行变换。对于变换,可使用不同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一对变换函数可定义为多重变换集合(MTS)。变换器140可选择在MTS中具有最高变换效率的一个变换函数对,并可在水平和垂直方向中的每一者上变换残余块。MTS中的变换函数对的信息(mts_idx)由熵编码器155编码并用信号通知视频解码设备。

[0070] 量化器145使用量化参数量化从变换器140输出的变换系数,并将量化的变换系数输出到熵编码器155。量化器145还可以立即量化相关的残差块,而无需用于任何块或帧的变换。量化器145还可根据变换块中的变换系数的位置应用不同的量化系数(缩放值)。应用于以2维布置的量化的变换系数的量化矩阵可被编码并用信号通知给视频解码设备。

[0071] 重排单元150可以执行针对量化的残差值的系数值的重新排列。

[0072] 重排单元150可以通过使用系数扫描将2D系数阵列改变为1D系数序列。例如,重排单元150可以通过使用Z字形扫描或者对角线扫描将DC系数扫描为高频域系数来输出ID系数序列。根据变换单元的大小和帧内预测模式,还可使用沿列方向扫描2D系数阵列的垂直扫描和沿行方向扫描2D块类型系数的水平扫描来代替Z字形扫描。换句话说,根据变换单元的大小和帧内预测模式,可在Z字形扫描、对角扫描、垂直扫描和水平扫描中确定要使用的扫描方法。

[0073] 熵编码器155通过使用包括基于上下文的自适应二进制算术码(CABAC)、指数哥伦布等的各种编码方案对从重排单元150输出的ID量化的变换系数的序列进行编码来生成比特流。

[0074] 此外,熵编码器155对与块划分相关的信息(诸如,CTU大小、CTU分割标志、QT分割标志、MTT划分类型、MTT划分方向等)进行编码,以允许视频解码设备将块等地划分到视频编码设备。此外,熵编码器155对关于指示当前块是通过帧内预测还是帧间预测来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、合并索引的情况下,以及在AMVP模式的情况下,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。

[0075] 逆量化器160对从量化器145输出的量化的变换系数进行去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,以重构残差块。

[0076] 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。当对下一次序块进行帧内预测时,经重构当前块中的像素可用作参考像素。

[0077] 环路滤波器单元180对重构像素执行滤波,以便减少由于基于块的预测和变换/量化而发生的块伪影、振铃伪影、模糊伪影等。作为环路滤波器的环路滤波器单元180可包括解块滤波器182、样本自适应偏移(SAO)滤波器184和自适应环路滤波器(ALF)186的全部或一些。

[0078] 解块滤波器182对重构块之间的边界进行滤波以便去除由于块单位编码/解码而发生的块效应,并且SAO滤波器184和ALF 186执行针对解块的滤波视频的附加滤波。SAO滤波器184及ALF 186是用于补偿经重构像素与原始像素之间的差的滤波器,所述差归因于有损编码而发生。SAO滤波器184应用偏移作为CTU单元以增强主观图片质量和编码效率。另一方面,ALF 186执行块单位滤波,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来补偿失真。关于将用于ALF的滤波器系数的信息可被编码并用信号通知给视频解码设备。

[0079] 通过解块滤波器182、SAO滤波器184和ALF 186滤波的重构块被存储在存储器190中。当一个图片中的所有块被重构时,重构的图片可以用作参考图片,用于对之后要编码的图片内的块进行帧间预测。

[0080] 图5是可实施本发明的技术的视频解码设备的功能框图。在下文中,参见图5,描述视频解码设备和设备的组件。

[0081] 视频解码设备可包括熵解码器510、重排单元515、逆量化器520、逆变换器530、预测器540、加法器550、环路滤波器单元560和存储器570。

[0082] 与图1的视频编码设备相似,视频解码设备的每个部件可实现为硬件或软件或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,并且还可以实现微处理器以执行与每个组件对应的软件的功能。

[0083] 熵解码器510通过对由视频编码设备生成的比特流进行解码以确定将被解码的当前块来提取与块划分相关的信息,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。

[0084] 熵解码器510通过从序列参数集(SPS)或图片参数集(PPS)提取关于CTU大小的信息来确定CTU的大小,并且将图片分割成具有确定大小的CTU。另外,确定CTU为树结构的最高层,即根节点,并且可以提取CTU的划分信息,以使用树结构对CTU进行划分。

[0085] 例如,当使用QTBT结构对CTU进行划分时,首先提取与QT的划分相关的第一标志(QT_split_flag),以将每个节点划分为下层的四个节点。此外,针对与QT的叶节点相应的节点,提取与MTT的划分相关的第二标志(mtt_split_flag)、划分方向(垂直/水平)和/或划分类型(二进制/三元),以将相应的叶节点划分为MTT结构。结果,在QT的叶节点下面的每个节点被递归地分割成BT或TT结构。

[0086] 作为另一实施例,当通过使用QTBT结构来划分CTU时,提取指示CU是否被划分的CU分割标志(split_cu_flag)。当对应的块被划分时,第一标志(QT_split_flag)也可以被提取。在拆分过程期间,对于每个节点,0次或更多次的递归MTT拆分可以在0次或更多次的

递归QT拆分之后发生。例如,对于CTU,MTT分裂可以立即发生,或者相反,仅多次的QT分裂也可以发生。

[0087] 又例如,当使用QTBT结构对CTU进行拆分时,提取与QT的拆分有关的第一标志(QT_split_flag),以将每个节点拆分成下层的四个节点。此外,提取指示与QT的叶节点相对应的节点是否被进一步分割成BT的分割标志(split_flag)以及分割方向信息。

[0088] 同时,当熵解码器510通过使用树结构的划分确定将被解码的当前块时,熵解码器510提取关于指示当前块是帧内预测的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提取表示用于帧间预测信息的语法元素(即,运动矢量和运动矢量参考的参考图片)的信息。

[0089] 此外,熵解码器510提取量化相关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。

[0090] 重新排列单元515可以再次以与由视频编码设备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1D量化的变换系数的序列改变为2D系数阵列(即,块)。

[0091] 逆量化器520去量化量化的变换系数,并且通过使用量化参数去量化量化的变换系数。逆量化器520还可将不同的量化系数(缩放值)应用于以2D布置的量化的变换系数。逆量化器520可通过将来自视频编码设备的量化系数的矩阵(缩放值)应用于量化的变换系数的2D阵列来执行逆量化。

[0092] 逆变换器530通过经由将去量化的变换系数从频域逆变换到空间域中来重构残差信号来生成用于当前块的残差块。

[0093] 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提取仅变换块的子块被变换的标志(cu_sbt_flag)、子块的方向(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中以重构残差信号并且利用值“0”填充未被逆变换的区域作为残差信号以生成用于当前块的最终残差块。

[0094] 此外,当应用MTS时,逆变换器530通过使用从视频编码设备信号通知的MTS信息(mts_idx)来确定在水平方向和垂直方向中的每一个方向上应用的变换索引或变换矩阵。逆变换器530还通过使用所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。

[0095] 预测器540可包括帧内预测器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,并且在当前块的预测类型是帧间预测时激活帧间预测器544。

[0096] 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,在多个帧内预测模式之中确定当前块的帧内预测模式。帧内预测器542还通过根据帧内预测模式使用当前块的相邻参考像素来预测当前块。

[0097] 帧间预测器544通过使用从熵解码器510提取的用于帧间预测模式的语法元素确定当前块的运动矢量和运动矢量参考的参考画面。

[0098] 加法器550通过将从逆变换器530输出的残差块和从帧间预测器544或帧内预测器542输出的预测块相加来重构当前块。在对之后要解码的块进行帧内预测时,重构的当前块内的像素被用作参考像素。

[0099] 作为环路滤波器的环路滤波器单元560可以包括解块滤波器562、SAO滤波器564和ALF 566。解块滤波器562对重构块之间的边界执行解块滤波,以便去除由于块单位解码而发生的块效应。SAO滤波器564和ALF 566在解块滤波之后对重构块执行附加滤波,以便补偿重构像素和原始像素之间的差异,所述差异由于有损编码而发生。通过使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。

[0100] 通过解块滤波器562、SAO滤波器564和ALF 566滤波的重构块被存储在存储器570中。当一个图片中的所有块被重构时,重构的图片可以用作参考图片,用于对之后要编码的图片内的块进行帧间预测。

[0101] 在一些实施方式中,本发明涉及如上文所描述的对视频图像进行编码和解码。更具体来说,本发明提供一种视频编码方法和设备,所述方法和设备在通过下采样到压缩且上采样到其重构来处理视频信号时基于上采样或超分辨率技术利用环内滤波器。

[0102] 下面的实施方式可由视频编码设备中的环路滤波器单元180执行。以下实施方式也可以由视频解码设备中的环路滤波器单元542执行。

[0103] 在重构帧的环内滤波中的视频编码设备可在率失真优化方面生成与本实施例相关联的信令信息。视频编码设备可使用熵编码器155对信令信息进行编码,并将编码的信令信息传输到视频解码设备。视频解码设备可使用熵解码器510从比特流解码与重构帧的环内滤波相关联的信令信息。

[0104] 在以下描述中,术语“目标块”可与当前块或编码单元(CU)互换地使用。术语“目标块”可以指编码单元的一些区域。

[0105] 此外,一个标志的值为真指示何时该标志被设定为1。另外,一个标志的值是假指示何时该标志被设定为0。

[0106] I. 卷积神经网络(CNN)

[0107] CNN指的是由多个卷积层和池化层组成的神经网络,并且是已知最适合图像处理的深度学习技术。卷积层使用多个核或滤波器来提取特征图(也称为“特征”)。构成滤波器的核系数是在学习过程期间确定的参数。

[0108] 在CNN的卷积层中,靠近输入的前层提取响应于简单的低级图像特征(诸如线、点或面)的特征图,而靠近输出的后层提取响应于高级特征(诸如纹理和对象部分)的特征图。

[0109] 图6是示出根据本公开内容的至少一个实施方式的卷积层的操作的示图。

[0110] 卷积层利用卷积运算来从输入图像生成特征图。图6示出了具有核大小为 3×3 的核(或滤波器)。核大小也称为滤波器大小。核具有核参数或滤波器参数,也称为权重。图6中示出的核总共具有九个核参数。核参数可最初设定为任意值,且参数的值可基于训练而更新。

[0111] 卷积层通过使用等于输入图像中的核大小的块来执行卷积运算。在这种情况下,输入图像中的核的大小的块被称为窗口。

[0112] 当以光栅扫描顺序执行对输入图像的滤波时,窗口的位移被称为步幅。在图6的实施例中,步幅是1。如果将步幅设定为2,那么用偏移2个采样的窗口执行卷积运算,从而导致特征图的水平维度和垂直维度为输入图像的水平维度和垂直维度的一半。

[0113] 如上所述,卷积层可以包括多个滤波器。滤波器的数量或核的数量被称作通道。因此,通道的数量等于滤波器的数量。滤波器的数量还确定特征图的维度。

[0114] 填充指的是通过在执行卷积运算之前在输入数据周围填特定值来扩展输入数据的方法。填充主要用于调整输出数据的空间大小。用于填充的值可由超参数来确定,但通常使用零填充。如果不使用填充,则输出数据的空间大小随着每个卷积层而减小,这可能导致边界信息丢失。因此,作为补救,使用填充。因此可以使用填充来均衡卷积层的输出数据与输入数据之间的空间大小。

[0115] 图7是展示了根据本公开的至少一个实施方式的去卷积层的操作的示图。

[0116] 去卷积层执行卷积层的相反操作。去卷积层将特征图作为输入并且生成期望的数据图像作为输出。去卷积层使用与卷积层相同的术语。如图7中的实施例所示,当步幅为1时,特征图的水平维度和垂直维度与输出数据的水平维度和垂直维度相同。当步幅为2时,输出数据的宽度和高度为特征图的宽度和高度的两倍。

[0117] 图8是示出根据本公开的至少一个实施方式的池化层的操作的图。

[0118] 池化层执行池化,其为对由卷积层生成的特征映射进行子采样的过程。池化层利用 2×2 窗口来选择采样,使得输出分别是输入的水平垂直大小的一半。换言之,池化层被用于通过将 2×2 区域聚集成单个采样来减小输入图像或输入特征图的大小。如图8所示,池化方法包括选择 2×2 区域中的最大值的最大池化和生成 2×2 区域的平均值的平均池化。不同于卷积层,池化层不包括需要被训练的变量,并且保持输入中的通道数量与输出中的通道数量相同。

[0119] 池化层的相反被定义为上池化层。上池化层充当池化层的相反,增加维度,并且通常在去卷积层之后使用。

[0120] 卷积编码器-解码器结构是由卷积层和去卷积层对组成的网络结构。卷积编码器由卷积层和池化层组成,以从输入图像输出特征图(或特征向量)。卷积编码器的最终输出向量也被称为潜在向量。卷积解码器由去卷积层和解包层组成,以从特征图或潜在向量生成输出图像。

[0121] 卷积编码器-解码器的输入和输出可以根据应用和网络的目的可变地设定。例如,输入和输出可以是光流图、显著性图、图像帧等。

[0122] 图9是示出SISR网络的示图。

[0123] CNN的应用的实施例是单图像超分辨率(SISR)。SISR网络生成高分辨率图像作为来自低分辨率输入图像的输出。如图9所示,SISR网络可包括多个卷积层。每个卷积层包括激活函数,诸如整流线性单元(ReLU)。可以训练SISR网络的参数,使得所得到的超分辨率(SR)图像接近地面真实值(GT)。

[0124] 使用CNN的SR方法可以通过增加深度(意味着例如通过增加卷积层的数量)来改善SR性能。为了克服学习中可能随着深度增加而发生的过度拟合的问题,可以在SISR网络中使用可以执行跳跃连接和残差学习的残差块。如图10所示,残差块除了用于将卷积运算应用于输入特征 x_1 的路径之外还包括跳跃路径。此外,残差块可以在生成输出 x_{l+1} 时基于学习效率选择应用卷积运算的路径或跳跃路径。在图10的实施例中,残差块包括批次归一化(BN)层。

[0125] 在一个实施例中,用于SISR的增强型深度残差网络(EDSR)通过连续地串联残差块而增加深度来增加网络的性能。使用超深卷积网络(VDSR)的准确图像超分辨率的另一示例是使用残差学习和将残差帧添加到最终输出的基于残差信号的残差学习的视觉几何组或

基于VGG网络的CNN模型。VDSR通过在网络的最末端处添加残差信号以将残差信号添加至输入信号。

[0126] II. 自适应环路滤波器 (ALF)

[0127] VVC的ALF 186、566使用基于维纳-霍普夫方程的自适应线性滤波器来将重构视频帧近似于原始帧。视频编码设备使用SAO 184的输出采样以基于速率-失真优化计算ALF 186的滤波器系数且将其传输到视频解码设备。如图11的实施例所示,ALF 186、566分别被配置为亮度和色度采样的 7×7 菱形形状和 5×5 菱形形状。可通过考虑编码效率与计算复杂性之间的折衷来确定滤波器形状和大小。例如,使用对称FIR滤波器可降低ALF 186、566的计算复杂性。

[0128] 为了导出在图11中示出的滤波器系数 c_i ,在相应位置处利用采样。当前位置 (x, y) 处的经滤波采样 $I(x, y)$ 可如等式1中所示通过7位精度运算来计算。

[0129] 【等式1】

$$[0130] \quad \tilde{I}(x, y) = I(x, y) + \left[\left(\sum_{i=0}^{N-2} c_i r_i + 64 \right) \gg 7 \right]$$

[0131] 在此, r_i 是根据等式2计算的在当前采样和相邻采样之间的差。

[0132] 【等式2】

$$[0133] \quad r_i = \min(b_i, \max(-b_i, I(x+x_i, y+y_i) - I(x, y))) + \min(b_i, \max(-b_i, I(x-x_i, y-y_i) - I(x, y)))$$

[0134] 此处, b_i 是削波参数。

[0135] ALF 186、566针对亮度分量利用最多25组滤波器系数,且将25组滤波器系数应用于 4×4 子块。基于通过使用拉普拉斯滤波器计算的局部块的梯度信息,将 4×4 子块分类为25个类别之一。具体来说,类别分类索引是从表示纹理分量的强度和方向的五个方向属性与子块的五个活动属性的组合导出的。另外,可在滤波之前将例如90度旋转、对角线移位和垂直移位等几何变换应用于滤波器系数。通过使用几何变换来考虑不同方向性,可通过使用滤波器系数的较小集合来处理更多种类的块特性。

[0136] 除了子块等级之外,可以在CTU等级处做出要应用的决定。对于色度分量,在CTU等级处使用多达八个滤波器。只有在亮度ALF在该等级被激活时,色度ALF才被激活。

[0137] 同时,自适应参数集 (APS) 用以传达ALF滤波器参数,包含滤波器系数的集合。如上所述,可针对亮度分量计算多达25组滤波器系数,且可针对色度分量计算多达8组滤波器系数。如果相同ALF系数用于不同切片,那么可用信号通知参考APS的索引,而不是冗余地重新传输ALF系数。

[0138] 在诸如高动态范围 (HDR) 和宽色域 (WCG) 的视频应用中,视频颜色的重构是关键的。交叉分量ALF (CC-ALF) 使用在该位置的当前色度采样和亮度采样之间的相关性,以与ALF并行地修改色度采样。为了从输入的亮度采样生成色度采样的校正,CC-ALF应用线性滤波操作。在这种情况下,没有对称约束的菱形FIR滤波器被用作CC-ALF。

[0139] III. 参考图片重采样 (RPR)

[0140] 在实时视频通信中,自适应分辨率改变 (ARC) 对于诸如自适应视频分辨率和快速流启动的需求是高度需要的。因而,VVC的参考图片重采样 (RPR) 技术可指代具有与当前图

片不同的分辨率的图片。

[0141] 图12是利用RPR技术的视频解码设备的框图。

[0142] 由视频编码设备下采样的视频帧被编码和传输,并且由视频解码设备先前解码的视频帧被存储在存储器570中的解码图片缓冲器(DPB)中。如果当前重构的视频帧具有与参考图片不同的大小,则视频解码设备中的重采样器1210根据大小比率重采样(放大或缩小)参考图片,并且在运动预测和补偿处理中根据大小比率缩放运动矢量。经解码的视频帧存储在DPB中而不进行重采样。当输出时,视频解码设备具有上采样器1220以用于将经下采样的视频帧上采样和重构到其原始大小。然而,上采样可能导致高频分量的丢失,并且因此输出的视频帧可能遭受诸如模糊的问题。

[0143] 在下文中,以视频解码设备为中心描述本公开的实施方式,但是它们可类似地应用于视频编码设备。

[0144] IV. 根据本公开的实施方式

[0145] 在下文中,描述了重采样环内滤波器(RIF)的实现方式。

[0146] 图13是根据本公开的至少一个实施方式的包括重采样环内滤波器的视频解码设备的框图。

[0147] 在根据本公开的至少一个实施方式的视频解码设备中,RIF(重采样环内滤波器)1310可被提供以代替构成常规环内滤波器的解块滤波器562、SAO(采样自适应偏移滤波器)564和ALF(自适应环路滤波器)566中的任一个,或者RIF 1310可被插入在常规环内滤波器的中间。在图13的实例中,RIF 1310位于SAO 564和ALF 566之间。

[0148] 如果输入的压缩图像已经被下采样,则RIF 1310将压缩图像上采样并且重构为原始分辨率。这里,原始分辨率是指原始帧的分辨率。如果输入视频以原始分辨率进入或者用于执行上采样的标志未被启用,则不执行RIF 1310的上采样。

[0149] 在此情况下,适用于VVC RPR技术的技术可用作滤波器来以原始分辨率下采样原始帧。具体地,可以通过使用利用加窗同步函数的低带通滤波器来执行下采样。进一步,在下采样期间,原始帧的水平和垂直维度除以缩放值。在这种情况下,下采样维度被截短或舍入到最近的整数。例如,如果水平方向上的大小是 W 并且缩放值是 S ,则下采样的长度被表示为 $W_d = \text{floor}(W/S)$,其中,floor是递降函数。这里,floor是递降函数。

[0150] 除了简单地除以缩放值之外,可在帧边界处填充像素,使得经下采样的结果为偶数或四的倍数。

[0151] RIF 1310对当前被重构的视频帧进行上采样。上采样缩放可为(但不一定限于)水平 $2x$ 或垂直 $2x$ 。例如,允许十进制增量(诸如 $1.5x$)的上采样,并且还允许多于 $2x$ (诸如 $3x$ 、 $4x$ 等)的缩放。由于上采样,原始分辨率的水平和垂直长度被信令用于重构到原始分辨率。RIF 1310可以利用以下任何技术来在当前重构下对视频帧进行上采样。

[0152] 在一个实施例中,RIF 1310利用有限脉冲响应(FIR)滤波器来对视频帧进行上采样。具体地,应用于在帧间预测期间生成预测块的插值滤波器可用于上采样。

[0153] 作为另一实施例,RIF 1310通过使用基于神经网络执行上采样的神经网络超分辨率(NNSR)技术来对视频帧进行上采样。具体地,可以使用转置卷积,或者可以在上采样之后执行卷积处理。可通过使用卷积中使用的数学运算的逆(即,通过使用转置卷积)来对图像进行上采样,并且可通过使用学习来更新所使用的核。可以基于图7中所示的去卷积层来实

现转置卷积。另外,用于超分辨率增强的网络架构和分量技术(例如上述SISR和EDSR)可用作NNSR。

[0154] 首先,如图14所示,可以实现转置卷积运算。在图14中说明的转置卷积运算中,RIF 1310可利用 3×3 核、步幅2和填充1来生成增加多于输入特征图的水平和垂直维度的输出。如图14的实施例中所示,可将输入特征图的单个像素乘以 3×3 核以生成输出像素的对应集合。

[0155] 另外,RIF 1310可通过使用上述上池化方法来执行上采样。执行图8所示的池化方法的逆过程的上采样方法包括最大上池化和平均上池化。最大上池化记忆最大池化时矩阵中的最大值的位置索引,然后将采样值放置在由上池化时的位置索引指示的位置处。在这种情况下,最大上池化可用相同的值填邻近像素,或者最大上池化可用零填邻近像素。同时,在输入值是输出值的平均值的假设下,平均上池化执行上采样。在图15中示出了平均上池化。

[0156] 另外,RIF 1310可在上采样之后通过使用最近邻或者使用基于双线性或基于双三次内插来执行卷积。在这种情况下,由于输入图像已经上采样至原始分辨率,可以使用去噪神经网络将卷积应用于上采样图像。

[0157] 图16是说明根据本发明的至少一个实施例的利用NNSR的上采样技术的图。

[0158] 在图16的实施例中,对于下采样的输入图像,NNSR通过使用多个卷积层和上池化层来执行上采样。在图16的实施例中, $K_1 \times K_1$ 表示下采样输入图像的分辨率,并且 $K_2 \times K_2$ 表示上采样输出图像的分辨率。此外,上采样层可以实现为如上所述的上池化层。

[0159] 同时,RIF 1310的位置可以是视频解码设备中的环内滤波器560内的任何地方,如图13的实施例中所示。例如,RIF 1310可以位于解块滤波器562、SAO 564和ALF 566中的任何一个处。

[0160] 例如,RIF 1310可以位于SAO 564之后和ALF 566之前,如在图13的实例中。因此,视频解码设备可向ALF 566输入通过将基于FIR的上采样滤波器应用于已经通过SAO 564的下采样输出采样而生成的信号,如图17的实例中。这里,ALF 566的输入信号由 $I(x, y)$ 表示。

[0161] ALF 566的滤波器系数可通过使用维纳-霍普夫等式来计算。具体来说,可导出滤波器系数以在输入信号 $I(x, y)$ 通过ALF 566之后紧密近似原始图像。常规ALF 566移除具有相同分辨率的输入视频帧和输出视频帧之间的量化噪声,但在本实施方式中,输入视频帧和输出视频帧具有不同的分辨率,因此除了移除量化噪声之外,还可导出滤波器系数以减轻由上采样处理引起的模糊效应。当使用与传统ALF 566相同的结构和相同的滤波器系数推导方法时,如上所述,添加RIF 1310可以取代ALF 566的目的。

[0162] 在这种情况下,ALF 566可以是以菱形或方形形状配置的二维线性滤波器。如图11的实施例中所示,ALF 566可被配置为 7×7 菱形和 5×5 菱形以分别用于亮度和色度采样。可通过考虑编码效率与计算复杂性之间的折衷来确定滤波器形状和大小。例如,通过使用对称FIR滤波器可降低ALF 566的计算复杂度。

[0163] 为了导出在图11中示出的滤波器系数 c_i ,使用位置处的采样。当前位置 (x, y) 处的经滤波采样 $I(x, y)$ 可基于7位精度运算如等式1中所示来计算。在等式1中,可根据等式2计算当前采样与相邻采样之间的差 r_i 。在等式2中, b_i 是削波参数。

[0164] ALF 566针对亮度分量利用最多25组滤波器系数且将其应用于 4×4 子块。基于通

过使用拉普拉斯滤波器计算的局部块的梯度信息,将 4×4 子块分类为25个类别之一。具体来说,类别分类索引是从表示纹理分量的强度和方向的五个方向属性与子块的五个活动属性的组合导出。另外,可在滤波之前将例如90度旋转、对角线移位和垂直移位等几何变换应用于滤波器系数。通过使用几何变换来考虑不同方向性,可通过使用滤波器系数的较小集合来处理更多种类的块特性。

[0165] 下面,描述帧间预测处理中RIF 1310的操作。

[0166] 当RIF 1310用于当前帧的帧间预测过程中时,视频解码设备可使用以下技术中的一者用于帧间预测。

[0167] 首先,视频解码设备通过使用固定帧大小执行帧间预测。视频解码设备将重构帧转换为固定分辨率且将其存储在存储器570中的DPB中。视频解码设备接着执行当前帧与存储在DPB中的具有固定分辨率的参考帧之间的运动预测和补偿。视频解码设备可在将其缩放到输出分辨率之后输出固定大小的帧。可替代地,如果固定分辨率和输出分辨率相同,则视频解码设备在没有任何其他转换的情况下输出帧。

[0168] 第二,如果在帧间预测中RIF 1310用于当前帧,则在当前帧和参考帧之间可能存在分辨率差。

[0169] 视频解码设备缩放参考帧以适合当前帧的大小,且接着执行帧间预测。默认地,可基于参考帧与当前帧之间的大小比率来放大或缩小存在于DPB中的帧。

[0170] 或者,视频解码设备在改变当前帧以适合参考帧的大小之后执行帧间预测。在此情况下,视频解码设备缩放所预测运动向量以适合改变的当前帧的大小。

[0171] 如果DPB中的重构帧具有多于一个大小,那么视频解码设备可基于多于一个大小执行帧间预测。为此目的,可以选择当前帧的大小、参考帧的大小和输出大小中的一个或多个。当基于一个或多个大小执行帧间预测时,视频解码设备可确定最终参考分辨率,可相应地缩放预测的运动信息,并存储缩放的运动信息。

[0172] 第三,如果当前帧和参考帧具有不同分辨率,那么视频解码设备可在空间和时间运动候选经缩放以匹配候选块的位置处的分辨率之后执行运动预测。或者,视频解码设备可通过仅考虑空间运动候选而不考虑时间运动候选来执行帧间预测。或者,对于时间运动候选,视频解码设备可将在未缩放位置处的候选块用作运动候选。

[0173] 下面,描述与RIF 1310相关的DPB的操作。

[0174] 已经被RIF 1310进行环内滤波的帧被存储在存储器570中的DPB中。视频解码设备在基于参考帧与当前帧之间的大小比率对DPB中存在的帧进行放大或缩小之后执行帧间预测。为了将帧存储在DPB中,视频解码设备使用以下技术中的一者。

[0175] 视频解码设备将所有帧存储在现有DPB中,而不管帧的大小如何。在该处理中,视频解码设备可在应用RIF 1310以存储具有选定大小的帧或存储具有两个大小的帧之前和之后选择帧的大小。例如,通过使用 $1/4$ 同样多的存储空间,被下采样至 $1/2$ 缩放的视频帧可被存储在常规DPB中。

[0176] 当在参考图片列表中管理参考帧的次序时,视频解码设备可利用常规前向参考图片列表RefPicList[0](下文称作L0)和逆参考图片列表RefPicList[1](下文称作L1)来管理参考帧的次序。

[0177] 或者,如果环内滤波之后的帧具有相对于原始帧改变的分辨率,那么视频解码设

备将改变的帧存储在参考图片列表中的特定索引处。例如,改变的帧可以被存储在第二索引 $\text{RefPictList}[0][1]$ 和 $\text{RefPictList}[1][1]$ 处。

[0178] 如果RIF 1310在帧间预测过程期间用于当前帧,则视频解码设备可利用在当前帧和参考帧之间具有不同分辨率的多个DPB(解码图片缓冲器)。

[0179] 视频解码设备可在应用RIF 1310之前和之后分离帧,并将它们存储在不同的DPB中。因此,一个DPB可以存储缩减的帧,并且另一个DPB可以存储具有输出分辨率的帧。

[0180] 下面描述与RIF 1310相关联的语法的信令。

[0181] 首先,描述与RIF 1310的行为相关的信令。

[0182] RIF 1310被启用还是禁用可被控制在更高等级,诸如视频序列等级、图片等级等。例如,可在SPS上表示处于视频序列等级处的标志以指示RIF 1310是否可用。如果该标志为真,则RIF 1310可用,而如果该标志为假,则RIF 1310不可用。

[0183] 另外,图片等级PPS可用以控制RIF 1310是以与SPS相同的方式启用还是禁用。例如,如果PPS上的对应标志为真,则可以启用RIF 1310。在这种情况下,RIF 1310的输出信号的分辨率相对于输入到环内滤波的信号而改变。另一方面,如果标志为假,则不启用RIF 1310。在这种情况下,在环内滤波器的输入和输出之间的分辨率没有变化。

[0184] 接下来,描述与RIF 1310的缩放值相关联的信令。

[0185] 当在视频序列等级或图片等级用信号通知指示RIF 1310被启用还是被禁用的标志时,改变的分辨率的缩放值被确定为2。此外,可以进一步确定是否在水平方向或垂直方向上应用2或1的缩放值。

[0186] 可从多个可用缩放值中选择除缩放值2之外的缩放值。在这种情况下,可通过使用索引来指示选择的缩放值。例如,如果索引 x_scale_index 是0,则可以保持原始分辨率,如果 x_scale_index 是1,则可以将x轴方向上的缩放值设定为2,并且如果 x_scale_index 是2,则可以将x轴方向上的缩放值设定为4。

[0187] 除了通过使用索引显式地发信号通知缩放值之外,视频解码设备还可通过使用传递到视频解码设备的参数和算术表达式来确定最终缩放。

[0188] 接下来,描述与RIF 1310的上采样滤波器相关联的信令。

[0189] 当神经网络用作RIF 1310时,可用信号通知指示许多候选之一的索引。该指数可指示神经网络的对应的参数集。基于用信号通知的索引,视频解码设备建立用于超分辨率的神经网络。候选神经网络可以根据量化参数、缩放值等来预训练。视频编码设备和视频解码设备可以存储和使用相同结构和对应参数的神经网络。

[0190] 为了对视频内容进行自适应上采样,由视频编码设备使用的神经网络的参数集被存储在比特流中,然后被传输。当参数集被传输时,所述参数集可存储在自适应参数集(APS)中。

[0191] 当FIR滤波器用作RIF 1310且用ALF 566重构帧时,可在存储于APS中时信号通知对应的滤波器系数集。在这种情况下,可基于缩放值使用多个FIR滤波器中的一个。

[0192] 接下来,描述与RIF 1310的输出分辨率相关联的信令。

[0193] 在执行RIF 1310之后获得的输出分辨率被存储在SPS上。

[0194] 通过将下采样的传输图像乘以预定缩放值来计算在执行RIF 1310之后获得的输出分辨率。

[0195] 在执行RIF 1310之后获得的输出分辨率与原始分辨率相同。

[0196] 在一些情况下,在执行RIF 1310之后获得的输出分辨率不同于原始分辨率。在这样的情况下,输出分辨率可以被存储在SPS中并且从SPS传输,或者可以通过乘以预定缩放值来计算。

[0197] 通过设定RIF (1310) 输出的最大分辨率,输出分辨率可被限制为不超过最大分辨率。此外,通过设定RIF (1310) 输入的最小分辨率,可以控制输入分辨率,使得当对原始视频帧进行下采样时输入分辨率不小于最小分辨率。

[0198] 现在参见图18和19,描述了一种基于RIF对重构帧进行上采样的方法。

[0199] 图18是根据本公开的至少一个实施例的由视频编码设备执行的用于对重构帧进行上采样的方法的流程图。

[0200] 视频编码设备获得缩放值 (S1800)。

[0201] 这里,缩放值表示重构帧和原始帧之间的分辨率差。此外,重构帧是已经根据视频编码设备的帧间预测预先重构原始帧的帧。

[0202] 另一方面,如果原始帧被下采样,则缩放值可以被设定为2。

[0203] 或者,在优化速率失真方面,视频编码设备可确定缩放值,编码指示缩放值的索引,且将缩放值传递到视频解码设备。

[0204] 视频编码设备获得重构帧 (S1802)。

[0205] 视频编码设备检查缩放值 (S1804)。

[0206] 如果基于缩放值的重构帧的分辨率小于原始帧的分辨率(在S1804中为是),则视频编码设备执行以下步骤S1806和S1812。

[0207] 视频编码设备将重构帧输入到基于深度学习的重采样环内滤波器中,以生成经上采样的重构帧 (S1806)。这里,重采样环内滤波器替代包括解块滤波器、SAO滤波器和自适应环路滤波器的滤波器中的一个,或者被插入在滤波器之间。经上采样的重构帧具有原始帧的分辨率。

[0208] 在一个实施例中,就针对率失真进行优化而言,视频编码设备确定针对重采样环内滤波器的指示多个神经网络候选之一的索引。随后,视频编码设备可编码所确定的索引。此时,可基于量化参数和缩放值来预训练每个神经网络候选。

[0209] 可替代地,视频编码设备可对从多个神经网络候选中选择为重采样环内滤波器的神经网络的参数集进行编码。在此情况下,可在存储于APS中时信号通知参数集。

[0210] 作为另一实施例,当FIR滤波器用作重采样环内滤波器并且用ALF重构帧时,视频编码设备可基于缩放值利用多个FIR滤波器中的一个。视频编码设备可信号通知存储在APS中的对应滤波器系数集。

[0211] 视频编码设备将由滤波器和重采样环内滤波器生成的输出帧存储在DPB中作为参考帧 (S1808)。

[0212] 视频编码设备确定指示是否使用重采样环内滤波器的标志 (S1810)。所述视频编码设备可以是否基于缩放值对所述原始帧下采样来确定指示是否使用所述重采样环内滤波器的标志。例如,如果原始帧经下采样,那么视频编码设备可将标志设定为真。

[0213] 视频编码设备可对指示是否使用重采样环内滤波器的标志进行编码 (S1812)。

[0214] 另一方面,如果重构帧的分辨率与原始帧的分辨率相同 (S1804中为否),则视频编

码设备执行传统的环内滤波,其中,排除重采样环内滤波器(S1820)。

[0215] 在这种情况下,视频编码设备可将指示是否使用重采样环内滤波器的标志设定为假,然后对所设定的标志进行编码。或者,视频编码设备可不传输指示是否使用重采样环内滤波器的标志。如果没有传输指示是否使用重采样环内滤波器的标志,那么视频解码设备可推断标志的值为假。

[0216] 图19是根据本发明的至少一个实施方式的由视频解码设备执行的用于对重构帧进行上采样的方法的流程图。

[0217] 视频解码设备获得重构帧(S1900)。在这里,重构帧是已经根据视频解码设备的帧间预测预先重构原始帧的帧。

[0218] 视频解码设备从比特流解码指示重采样环内滤波器被启用还是被禁用的标志(S1902),并且检查该标志(S1904)。

[0219] 另一方面,如果没有传输指示重采样环内滤波器被启用还是被禁用的标志,则视频解码设备可推断该标志的值为假。

[0220] 如果指示是否使用重采样环内滤波器的标志为真(在S1904中为是),则视频解码设备从S1906起执行以下步骤。

[0221] 视频解码设备获得缩放值(S1906)。在这里,缩放值表示在重构帧和原始帧之间的分辨率差。

[0222] 如果指示重采样环内滤波器是否被启用的标志为真,则缩放值可被设定为2。

[0223] 或者,视频解码设备可从比特流解码指示缩放值的索引,且基于索引设定缩放值。

[0224] 为了根据缩放值增加重构帧的分辨率,视频解码设备可将重构帧输入到基于深度学习的重采样环内滤波器中,从而生成上采样的重构帧(S1908)。这里,重采样环内滤波器替代包括解块滤波器、SAO滤波器和自适应环路滤波器的滤波器中的一个,或者被插入在滤波器之间。经上采样的重构帧具有原始帧的分辨率。

[0225] 在一个实施例中,视频解码设备可从比特流解码指示多个神经网络候选之一的索引,并随后基于该索引来设定重采样环内滤波器。神经网络候选中的每个可基于量化参数和缩放值来预训练。

[0226] 或者,视频解码设备可从比特流解码重采样环内滤波器的参数集。在此情况下,可在存储于APS中时信号通知参数集。

[0227] 作为另一实施例,当FIR滤波器用作重采样环内滤波器且用ALF重构帧时,视频解码设备可基于缩放值利用多个FIR滤波器中的一者。对应滤波器系数集可存储在APS中且由视频编码设备信号通知。

[0228] 视频解码设备将通过重采样环内滤波器和DPB中的滤波器生成的输出帧存储为参考帧(S1910)。

[0229] 另一方面,如果指示重采样环内滤波器被启用还是被禁用的标志为假(S1904中为否),则视频解码设备执行常规环内滤波,其中排除了重采样环内滤波器(S1920)。

[0230] 尽管各个流程图中的步骤被描述为被顺序地执行,但是这些步骤仅仅例示本公开的一些实施方式的技术构思。因此,本公开所属领域的普通技术人员可通过改变在各个附图中描述的顺序或者通过并行执行两个以上的步骤来执行步骤。所以,各个流程图中的步骤不限于所示的时间序列顺序。

[0231] 应当理解,以上描述呈现了可以以不同其他方式实现的说明性实施例。在一些实施方式中描述的功能可以通过硬件、软件、固件和/或其组合来实现。还应理解,本公开中描述的功能部件用“…单元”标记强烈强调它们独立实现的可能性。

[0232] 同时,在一些实施方式中描述的各种方法或功能可以被实现为存储在能够由一个或多个处理器读取和执行的非暂时性记录介质中的指令。例如,非暂时性记录介质可包括各种类型的记录设备,其中,数据以计算机系统可读的形式存储。例如,非暂时性记录介质可包括存储介质,诸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。

[0233] 尽管出于说明性的目的已经描述了本公开的实施方式,然而,本公开所属领域的普通技术人员应当认识到,在不背离本公开的构思和范围的情况下,各种修改、添加、以及替代是可能的。因此,为了简洁和清晰起见,已经描述了本公开的实施方式。本公开的实施方式的技术构思的范围不受图示的限制。因此,本公开所属领域的普通技术人员应当理解,本公开的范围不应受上述明确描述的実施方式的限制,而是受权利要求及其等同物的限制。

[0234] (参考标号)

[0235] 180:环内滤波器单元

[0236] 155:熵编码器

[0237] 190:存储器

[0238] 510:熵解码器

[0239] 560:环内滤波器单元

[0240] 570:存储器

[0241] 1310:重采样环内滤波器

[0242] 相关申请

[0243] 本申请要求2022年3月14日提交的韩国专利申请第10-2022-0031336号以及2023年2月10日提交的韩国专利申请第10-2023-0018329号的优先权和权益,通过引用将其每一个的全部内容结合在此。

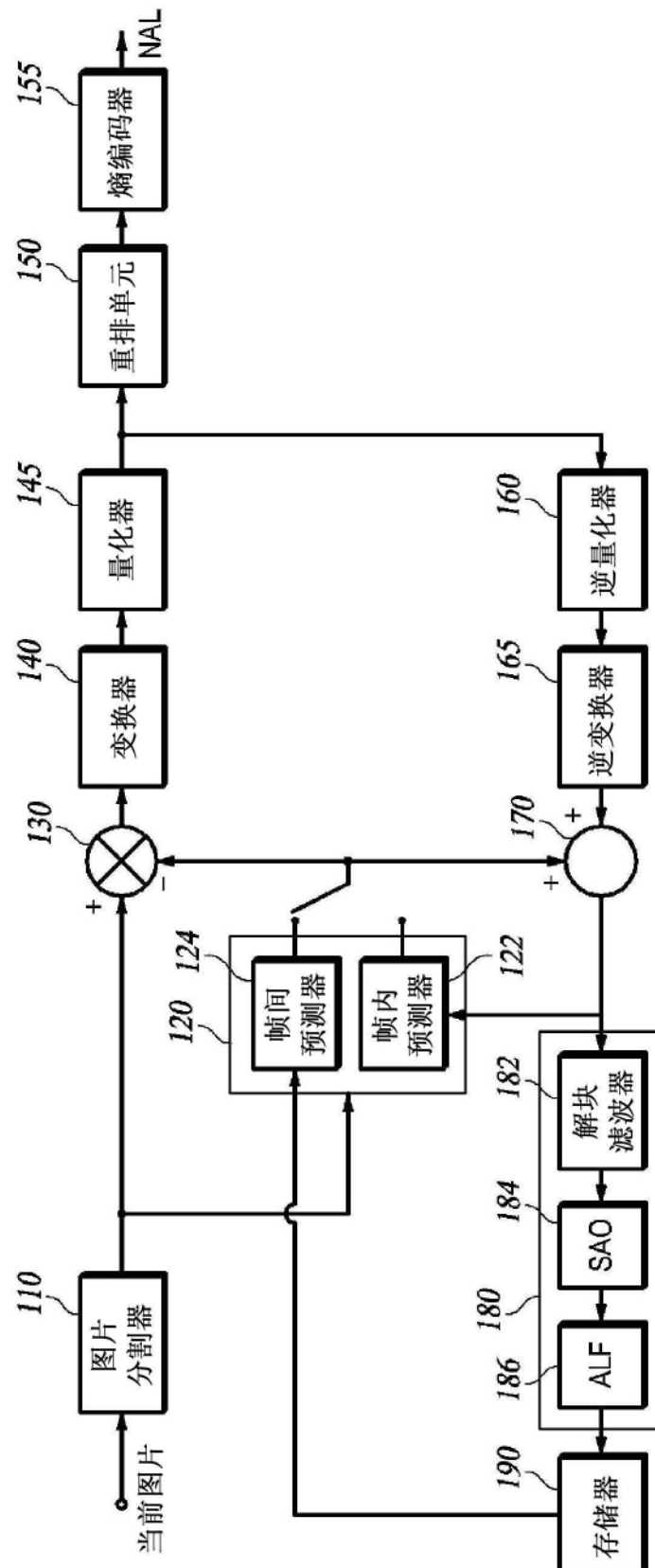


图1

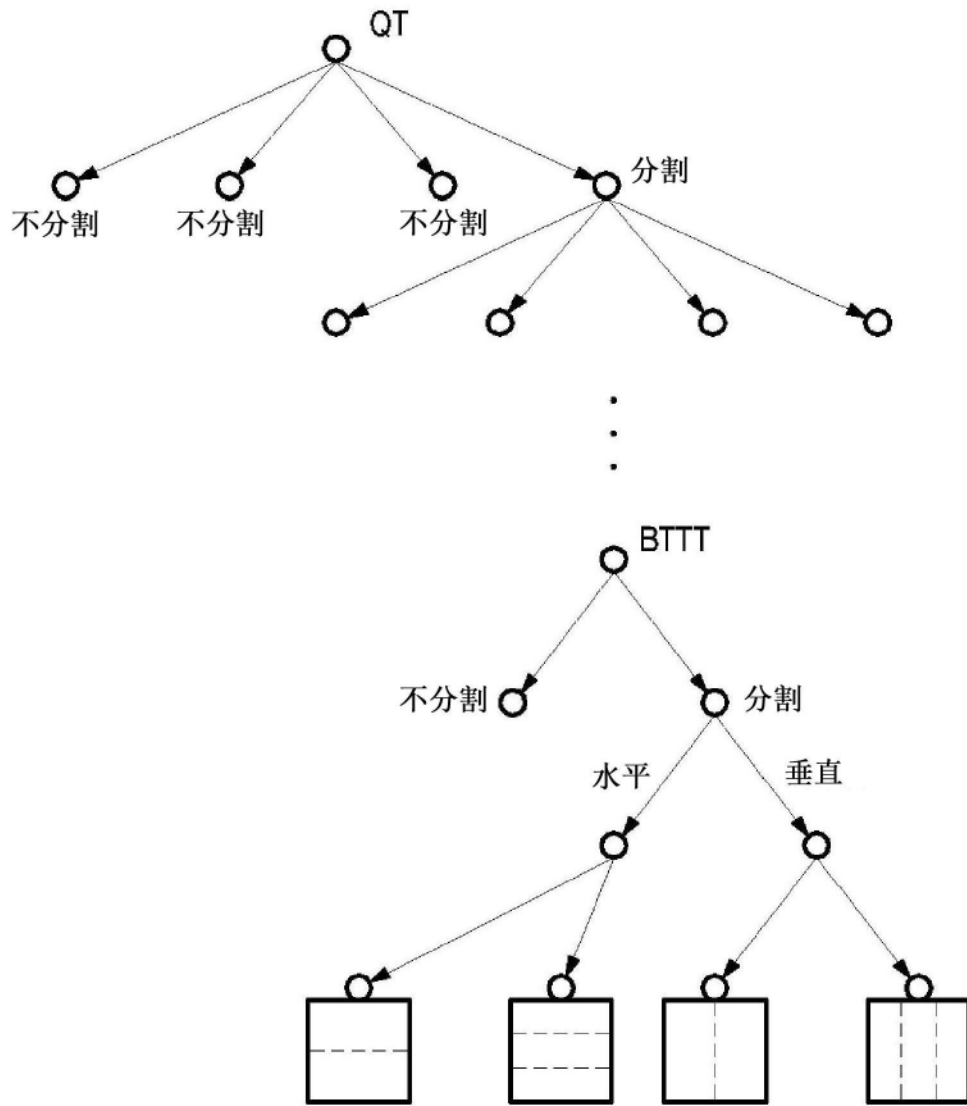


图2

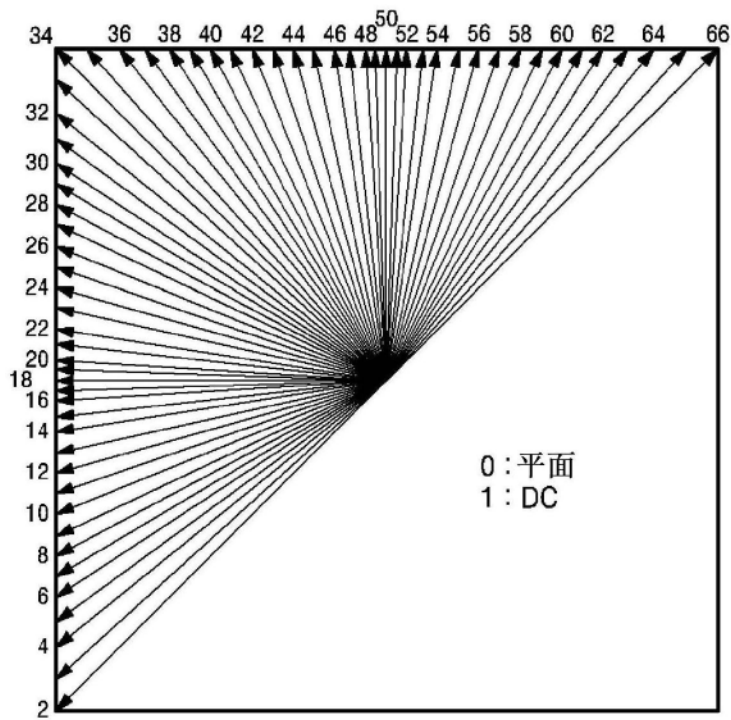


图3a

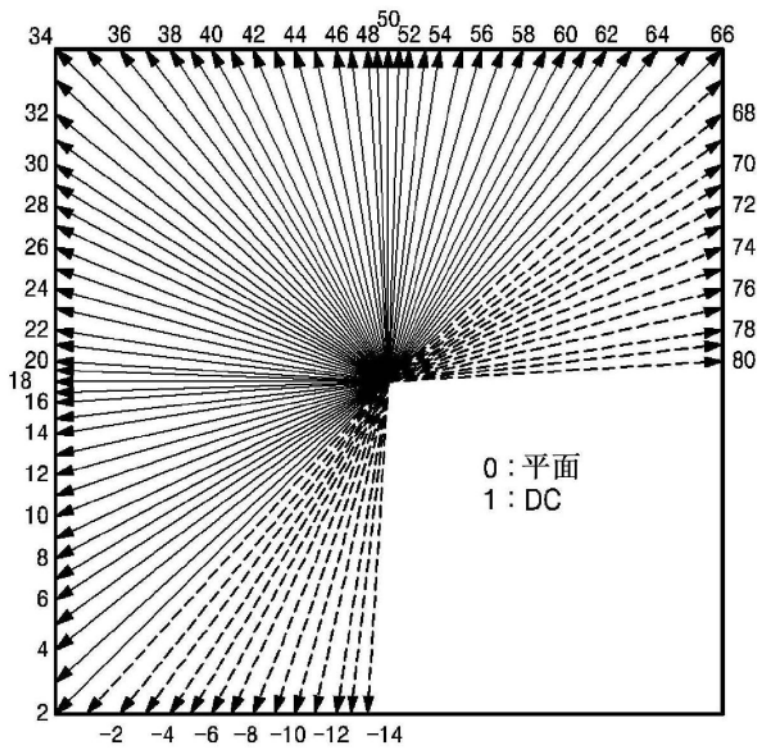


图3b

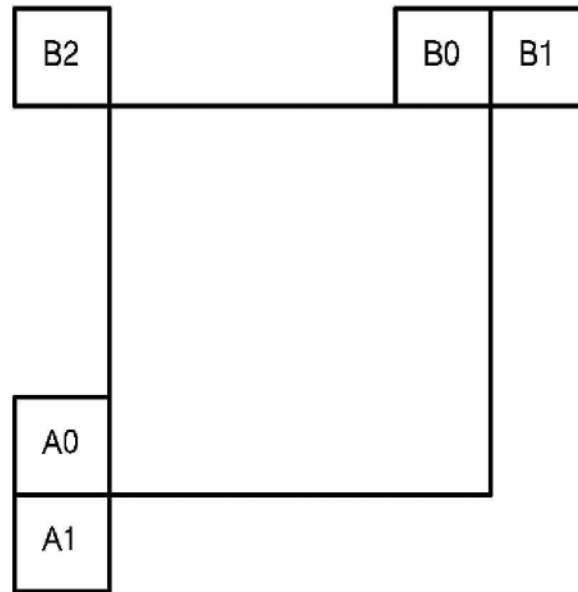


图4

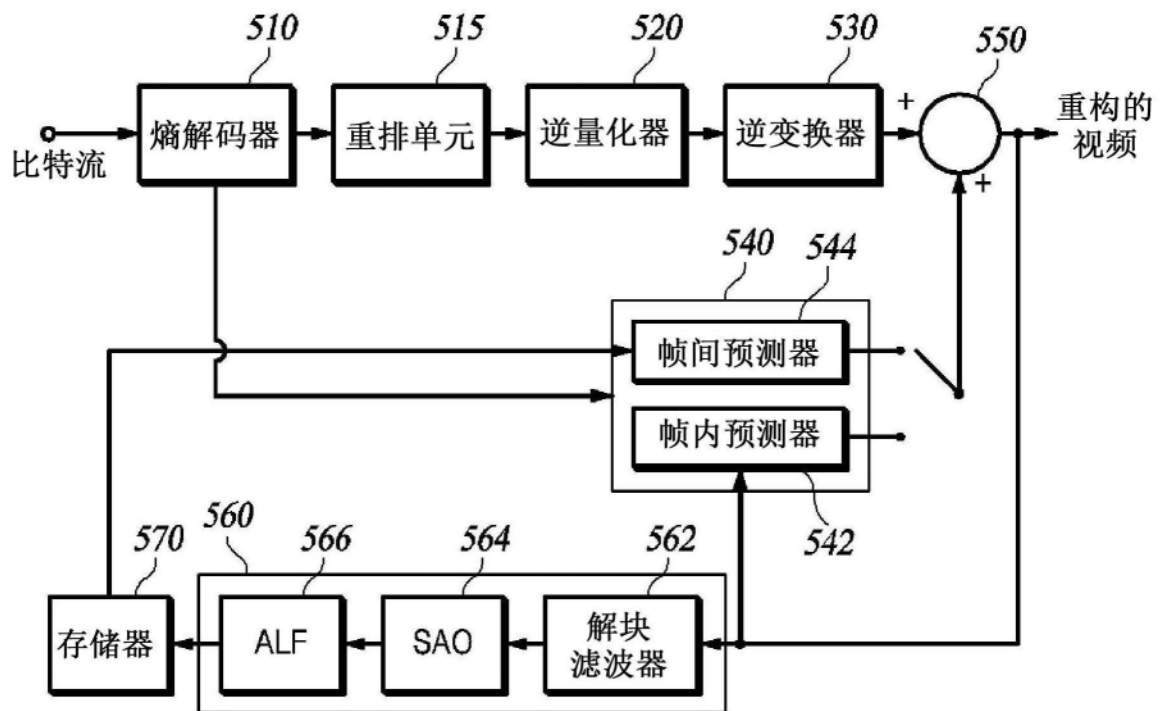


图5

卷积层

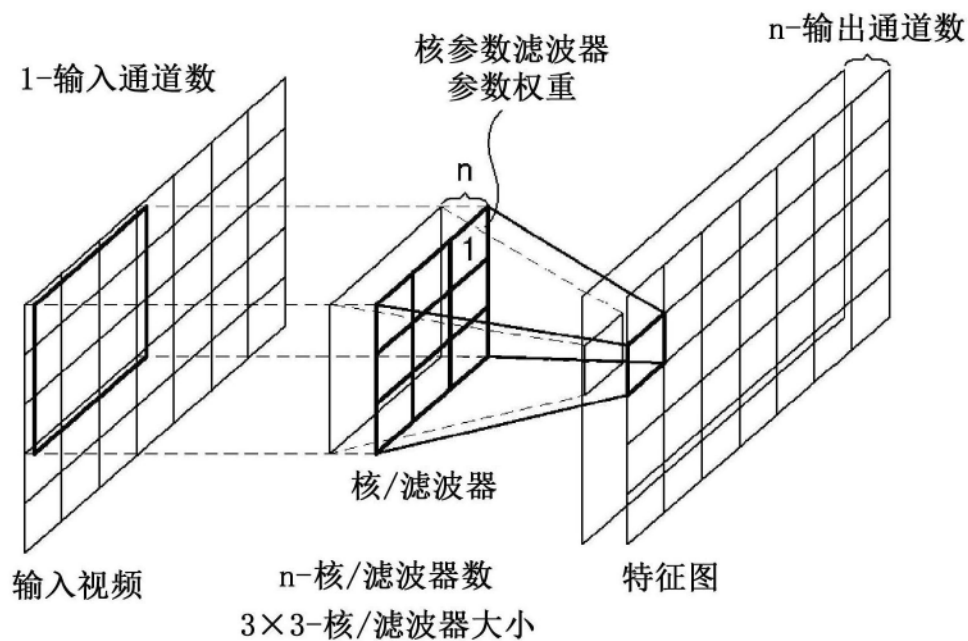


图6

去卷积层

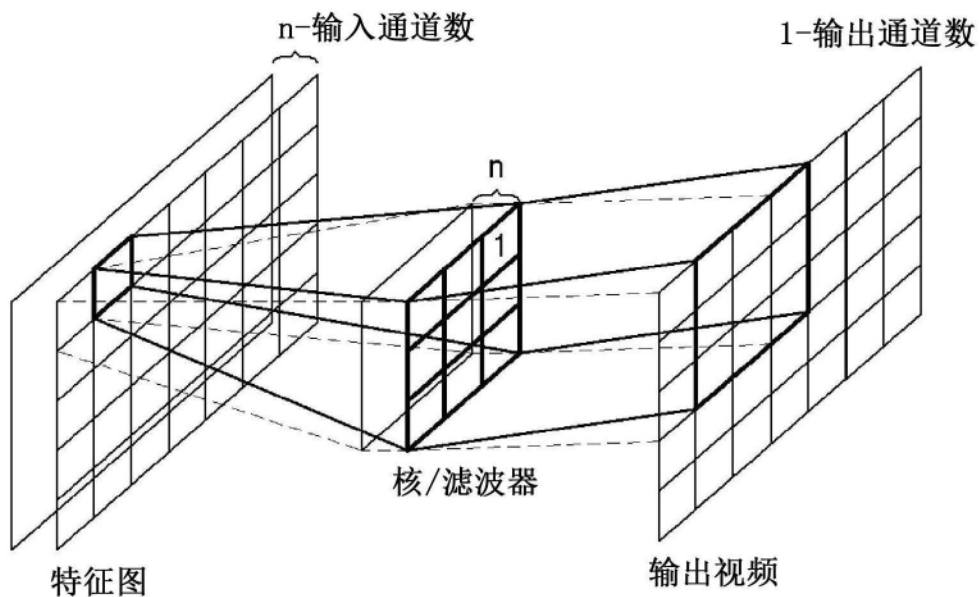


图7

池化层

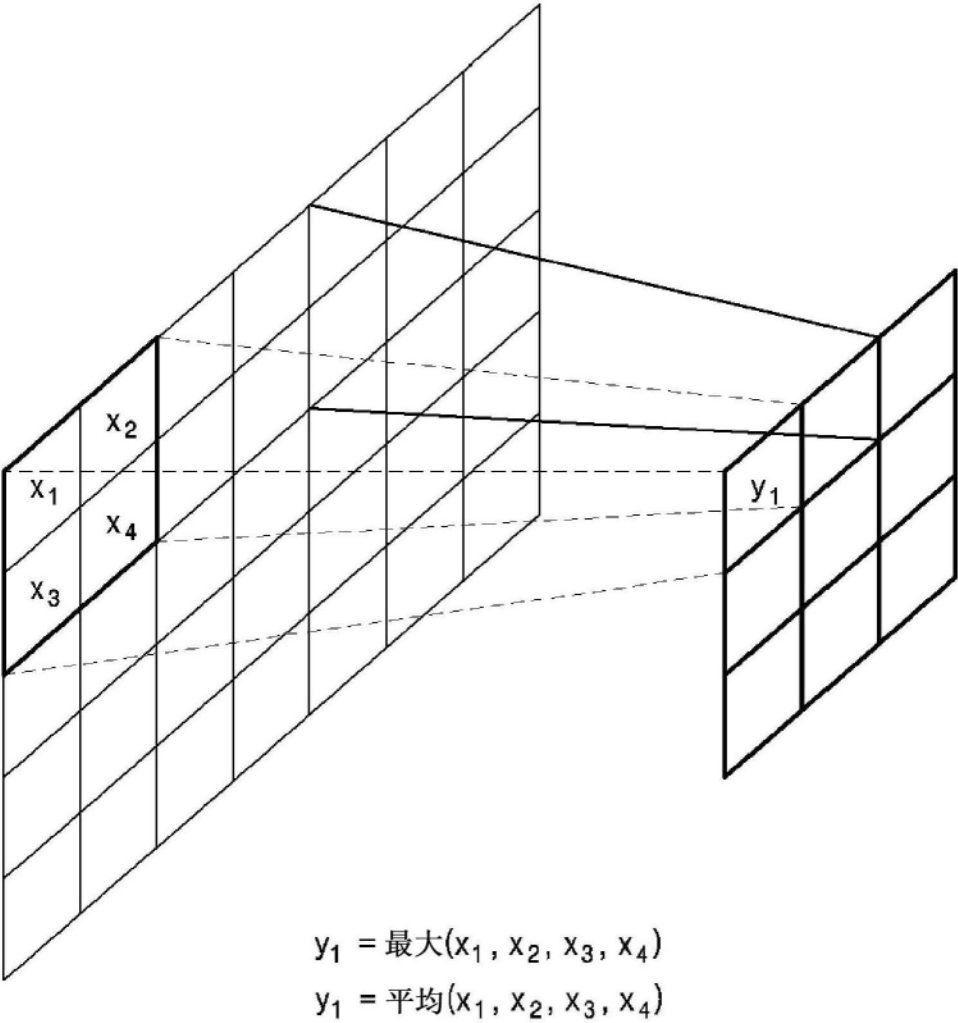


图8

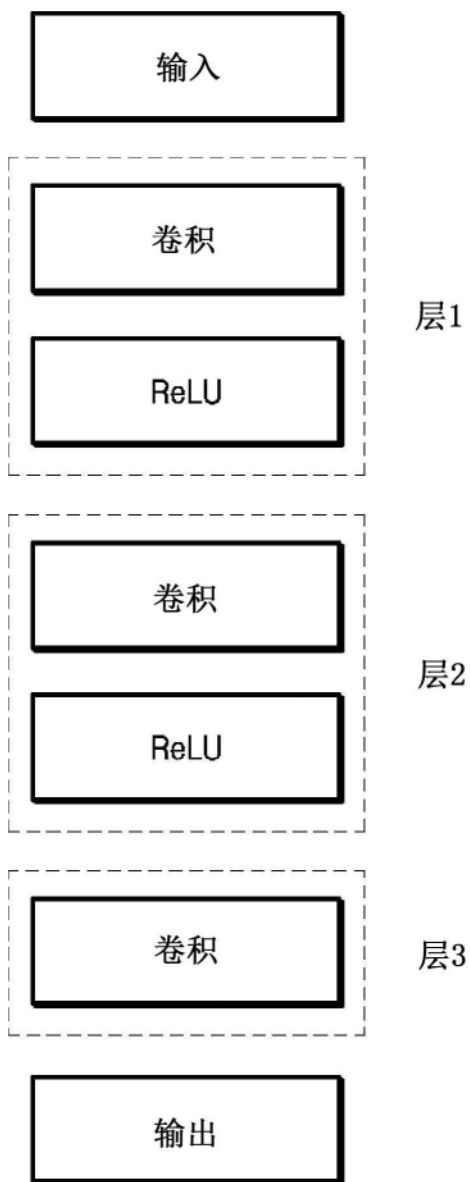


图9

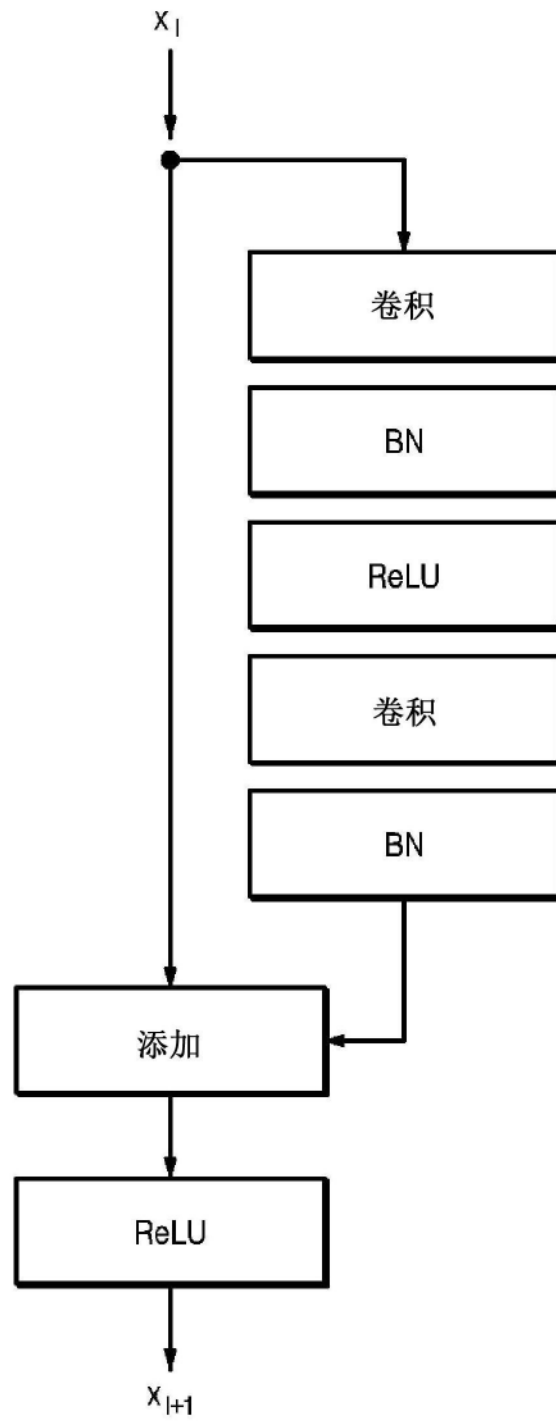


图10

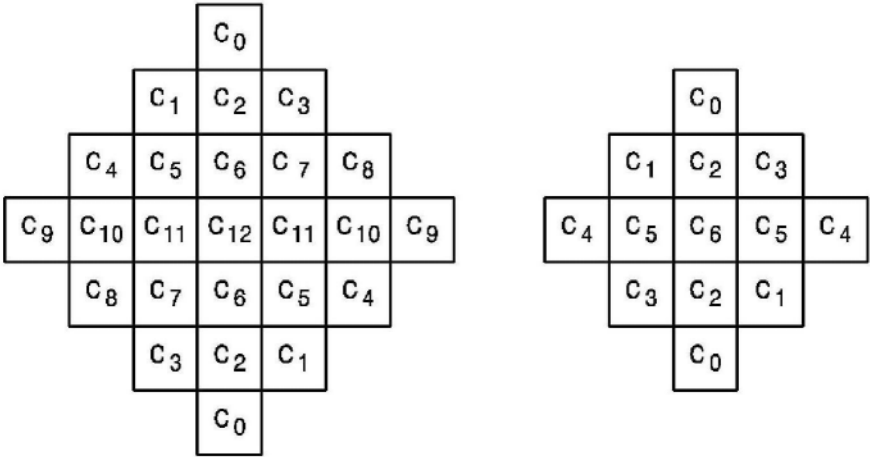


图11

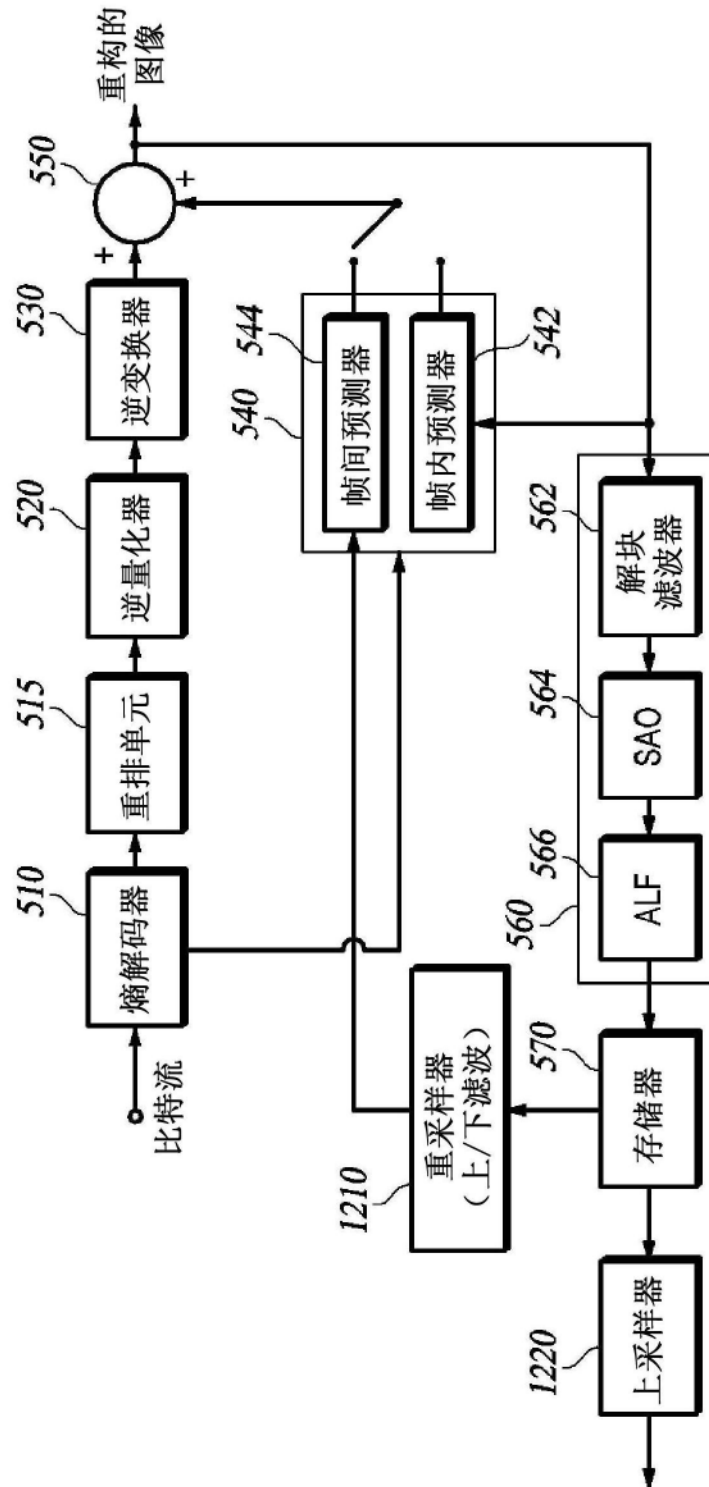


图12

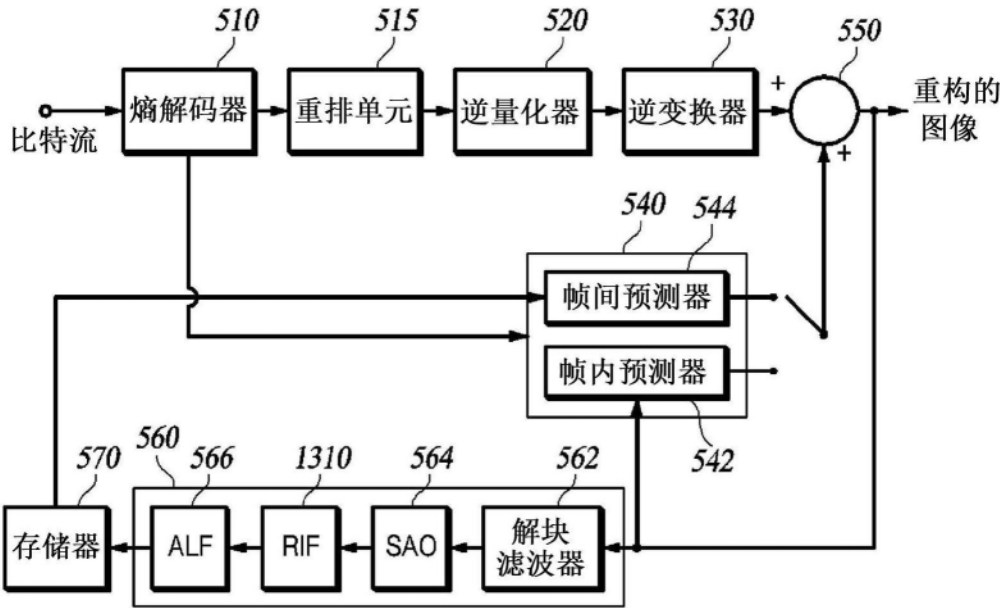


图13

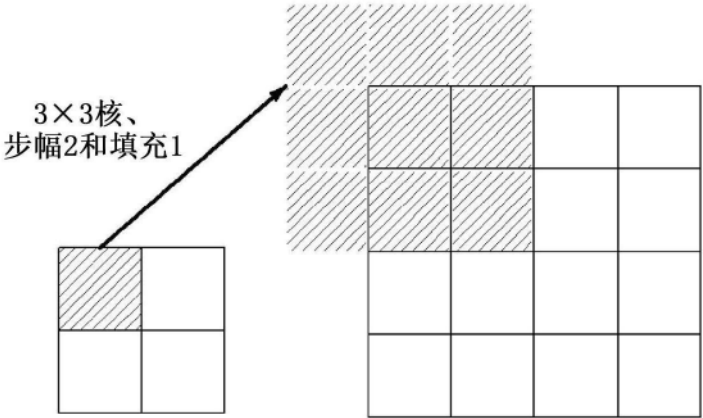


图14

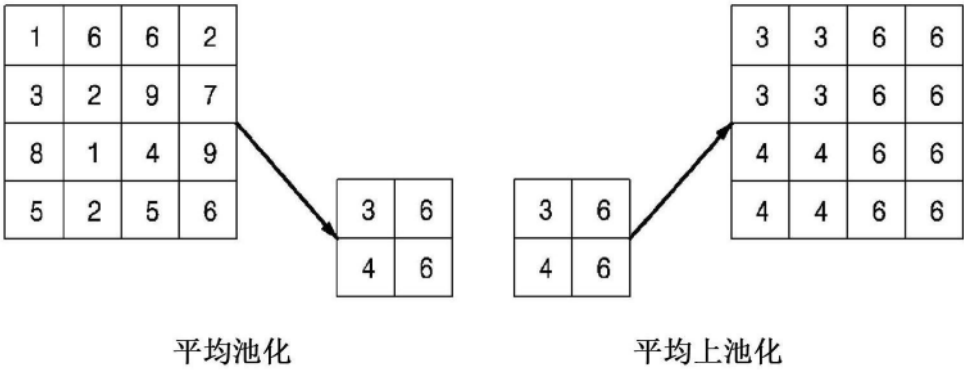


图15

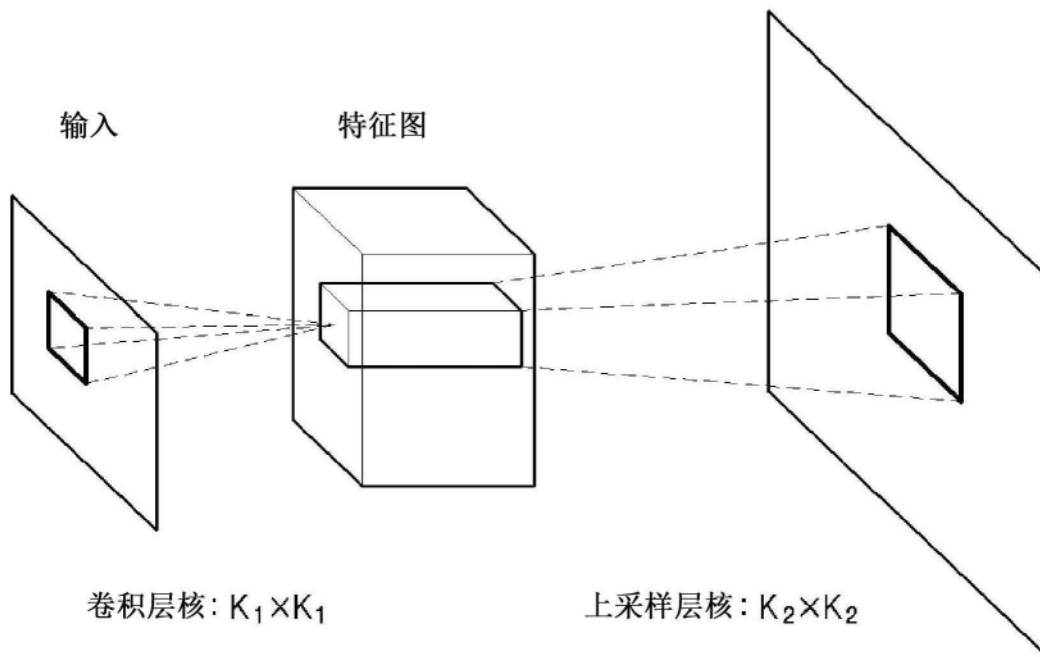


图16

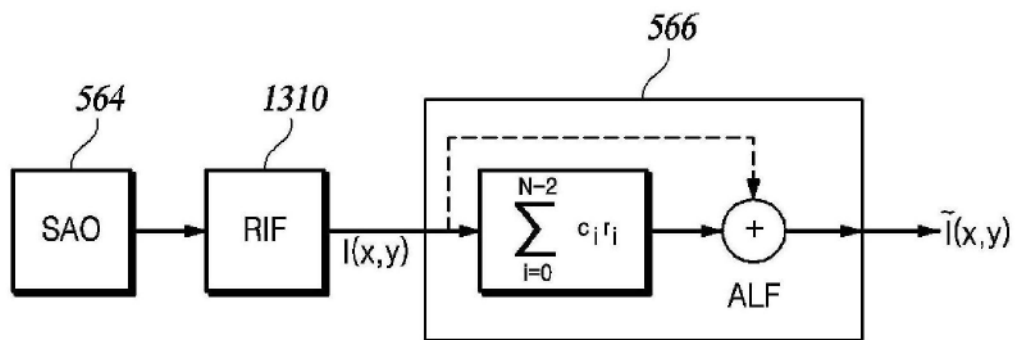


图17

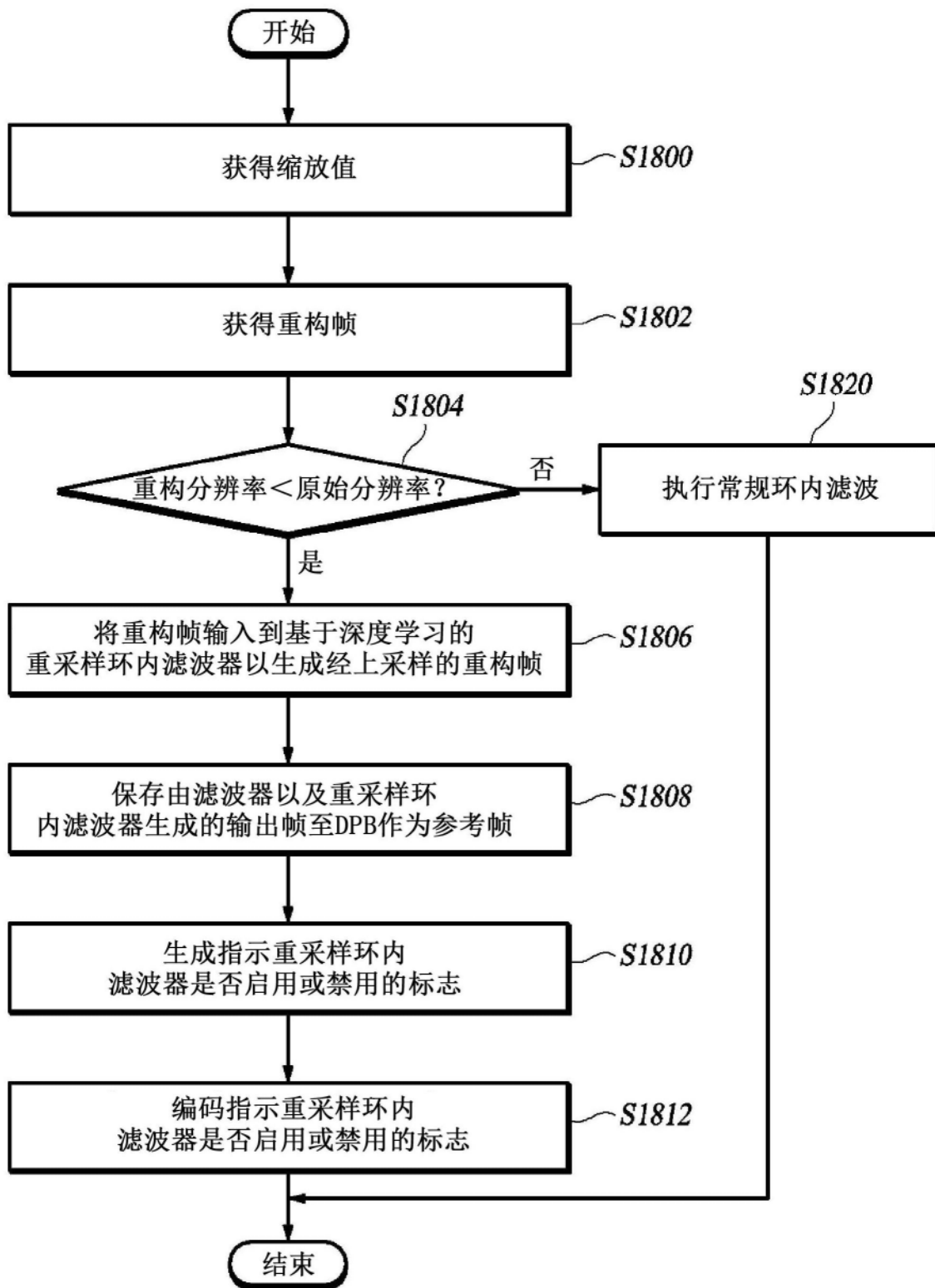


图18

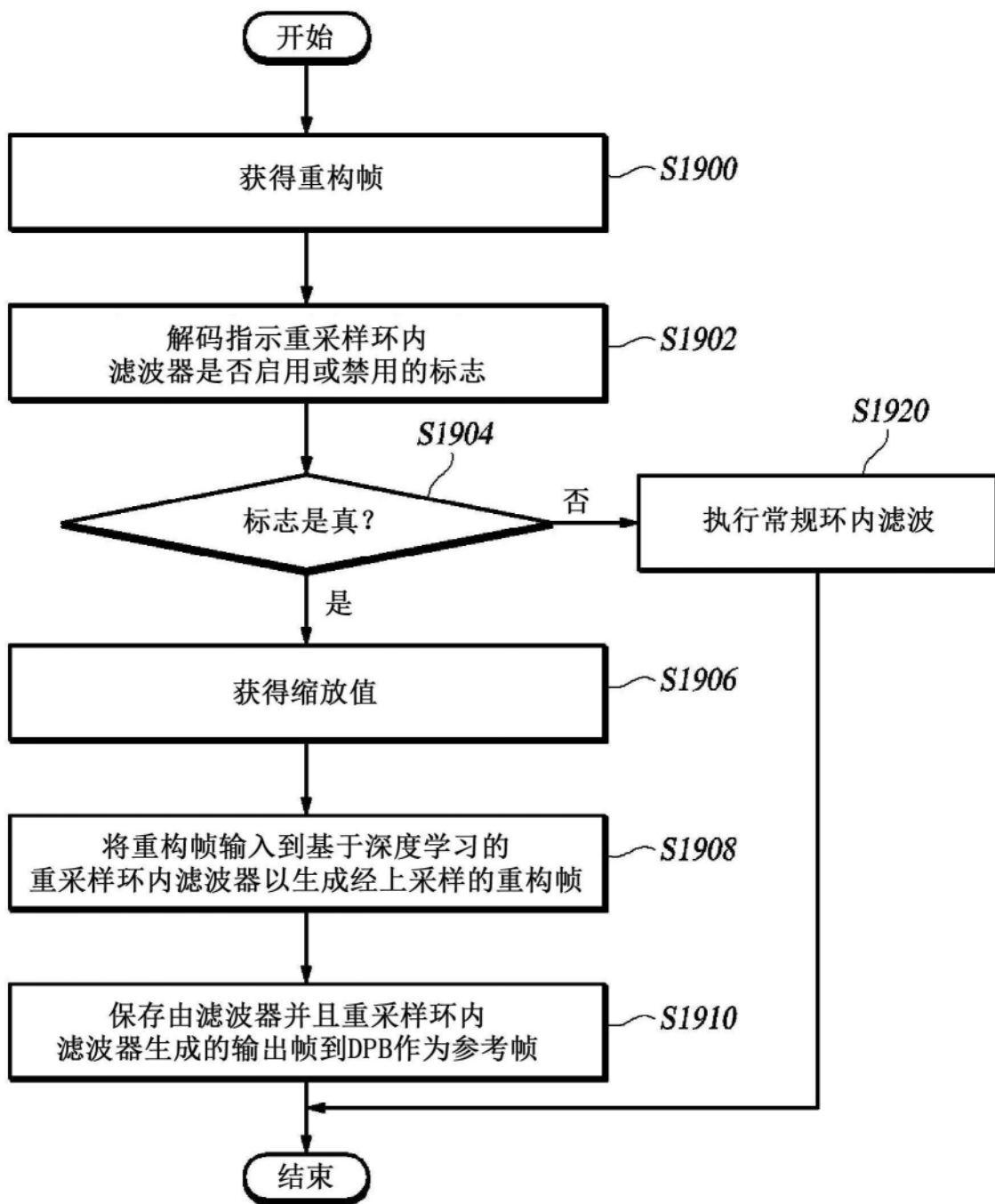


图19