



(12) 发明专利申请

(10) 申请公布号 CN 117693938 A

(43) 申请公布日 2024. 03. 12

(21) 申请号 202280048614.X

(22) 申请日 2022.06.09

(30) 优先权数据

10-2021-0091743 2021.07.13 KR

10-2022-0069466 2022.06.08 KR

(85) PCT国际申请进入国家阶段日

2024.01.09

(86) PCT国际申请的申请数据

PCT/KR2022/008119 2022.06.09

(87) PCT国际申请的公布数据

W02023/287018 KO 2023.01.19

(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 郑惠善 朴胜煜 许镇

(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 19/593 (2014.01)

H04N 19/105 (2014.01)

H04N 19/52 (2014.01)

G06N 3/0464 (2023.01)

G06T 9/00 (2006.01)

H04N 19/132 (2014.01)

H04N 19/124 (2014.01)

H04N 19/176 (2014.01)

H04N 19/13 (2014.01)

H04N 19/11 (2014.01)

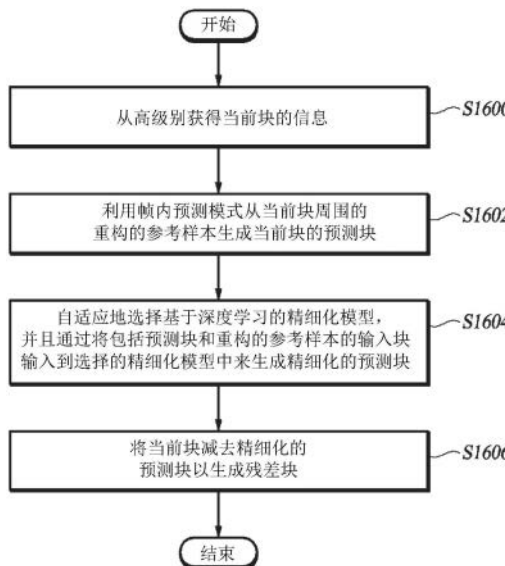
权利要求书3页 说明书19页 附图17页

(54) 发明名称

基于深度学习的精细化帧内预测信号的
视频编解码方法和装置

(57) 摘要

公开了基于深度学习的精细化帧内预测信号的视频编解码方法和装置,并且为了增强视频编码的效率并提高视频质量,本实施方案提供了这样一种视频编解码方法和装置,其通过将相邻重构的参考样本、基于当前块的帧内预测的预测块以及与当前块相关的块信息自适应地输入到深度学习网络来生成精细化的帧内预测信号。



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信息和当前块的残差值,当前块的信息包括当前块的帧内预测模式、当前块的高度、宽度或纵横比、当前块的颜色、包括当前块的图像的类型以及当前块的量化参数;

帧内预测器,其配置为利用帧内预测模式从当前块周围的重构的参考样本生成当前块的预测块;

预测的信号精细化单元,其配置为基于当前块的信息的全部或一些自适应地选择基于深度学习的精细化模型,并且将包括预测块和重构的参考样本的输入块输入到选择的基于深度学习的精细化模型中,以生成精细化的预测块;以及

加法器,其配置为将残差值添加到精细化的预测块,以生成当前块的重构的块。

2. 根据权利要求1所述的视频解码装置,其中,

所述预测的信号精细化单元配置为基于帧内预测模式是否是方向模式、帧内预测模式是否是基于矩阵的预测模式、或者帧内预测模式是否包括在宽角度帧内预测(WAIP)中来自适应地选择基于深度学习的精细化模型。

3. 根据权利要求1所述的视频解码装置,其中,

当帧内预测模式是方向模式时,所述预测的信号精细化单元配置为对帧内预测模式进行聚类,并且基于帧内预测模式包括在竖直模式、水平模式和对角线模式中的哪个聚类来自适应地选择基于深度学习的精细化模型。

4. 根据权利要求1所述的视频解码装置,其中,

所述预测的信号精细化单元配置为基于帧内预测模式是竖直预测还是水平预测来自适应地选择基于深度学习的精细化模型。

5. 根据权利要求1所述的视频解码装置,其中,

所述预测的信号精细化单元配置为基于当前块的高度或宽度是否大于预设阈值来自适应地选择基于深度学习的精细化模型。

6. 根据权利要求1所述的视频解码装置,其中,

预测的信号精细化单元配置为基于当前块的纵横比是否大于预设阈值来自适应地选择基于深度学习的精细化模型。

7. 根据权利要求1所述的视频解码装置,其中,

基于深度学习的精细化模型是包括全连接层的深度学习网络,并且根据学习块的信息来预先进行自适应训练。

8. 根据权利要求7所述的视频解码装置,其中,

所述预测的信号精细化单元配置为将输入块改变为一维矢量,将一维矢量输入到基于深度学习的精细化模型,并且将由基于深度学习的精细化模型生成的一维矢量重塑为预测块的大小。

9. 根据权利要求7所述的视频解码装置,其中,

所述预测的信号精细化单元配置为利用预测块和重构的参考样本来组成大小为 $(h+r) \times (w+r)$ 的输入块,其中, h 是当前块的高度, w 是当前块的宽度,并且 r 是顶部重构的参考样本的高度或左侧重构的参考样本的宽度。

10. 根据权利要求9所述的视频解码装置,其中,

当帧内预测模式是竖直预测时,所述预测的信号精细化单元配置为利用大小为 $(h+r) \times r$ 的左侧重构的参考样本和大小为 $h \times w$ 的预测块来组成输入块,并且当帧内预测模式为水平预测时,所述预测的信号精细化单元配置为利用大小为 $r \times (w+r)$ 的顶部重构的参考样本和大小为 $h \times w$ 的预测块来组成输入块。

11. 根据权利要求7所述的视频解码装置,其中,

所述预测的信号精细化单元配置为利用预测块和重构的参考样本来组成大小为 $(h+r_h) \times (w+r_w)$ 的输入块,其中, h 是当前块的高度, w 是当前块的宽度, r_h 是顶部重构的参考样本的高度,并且 r_w 是左侧重构的参考样本的宽度。

12. 根据权利要求11所述的视频解码装置,其中,

当帧内预测模式是竖直预测时,所述预测的信号精细化单元配置为组成输入块,使得 r_h 大于或等于 r_w ,并且当帧内预测模式为水平预测时,所述预测的信号精细化单元配置为组成输入块,使得 r_w 大于或等于 r_h 。

13. 根据权利要求7所述的视频解码装置,其中,

当帧内预测模式是方向模式时,所述预测的信号精细化单元配置为根据方向模式的斜率将帧内预测模式映射到更接近水平预测或竖直预测的一个。

14. 一种由视频解码装置执行的对当前块进行帧内预测的视频解码方法,所述视频解码方法包括:

从比特流解码当前块的信息和当前块的残差值,当前块的信息包括当前块的帧内预测模式、当前块的高度、宽度或纵横比、当前块的颜色、包括当前块的图像的类型以及当前块的量化参数;

利用帧内预测模式从当前块周围的重构的参考样本生成当前块的预测块;

基于当前块的信息的全部或一些自适应地选择基于深度学习的精细化模型,并且通过将包括预测块和重构的参考样本的输入块输入到选择的基于深度学习的精细化模型中来生成精细化的预测块;以及

通过将残差值添加到精细化的预测块来生成当前块的重构的块。

15. 一种由视频解码装置执行的对当前块进行帧内预测的视频编码方法,所述视频编码方法包括:

从高级别获得当前块的信息,当前块的信息包括当前块的帧内预测模式、当前块的高度、宽度或纵横比、当前块的颜色、包括当前块的图像的类型以及当前块的量化参数;

利用帧内预测模式从当前块周围的重构的参考样本生成当前块的预测块;

基于当前块的信息的全部或一些自适应地选择基于深度学习的精细化模型,并且通过将包括预测块和重构的参考样本的输入块输入到选择的基于深度学习的精细化模型中来生成精细化的预测块;以及

通过将当前块减去精细化的预测块来生成残差块。

16. 根据权利要求15所述的视频编码方法,其中,生成精细化的预测块包括:

根据帧内预测模式是竖直预测还是水平预测来自适应地选择基于深度学习的精细化模型。

17. 根据权利要求15所述的视频编码方法,其中,生成精细化的预测块包括:

根据当前块的高度或宽度是否大于预设阈值来自适应地选择基于深度学习的精细化

模型。

18. 根据权利要求15所述的视频编码方法, 其中, 生成精细化的预测块包括:

根据当前块的纵横比是否大于预设阈值来自适应地选择基于深度学习的精细化模型。

19. 根据权利要求15所述的视频编码方法, 其中, 基于深度学习的精细化模型是包括全连接层的深度学习网络, 并且根据学习块的信息来预先进行自适应训练。

基于深度学习的精细化帧内预测信号的视频编解码方法和装置

技术领域

[0001] 本发明涉及基于深度学习的精细化帧内预测的预测的信号的视频编码方法和装置。

背景技术

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

[0003] 由于视频数据与音频数据或静止影像数据相比具有较大的数据量,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] 对于要编码的视频中的当前块,在现有的帧内预测方法中,基于固定规则从相邻的重新配置的参考样本生成预测块,因此,可能难以生成复杂内容的预测块。此外,由于可以用作参考样本的信息量小于时间相邻信息量,帧内预测方法的编解码性能低于帧间预测方法的编解码性能。因此,可以使用滤波器或深度学习技术,以通过改进预测块来减少与原始块的差异信号。

[0007] 作为利用滤波器的传统技术,存在利用固定规则生成当前块的预测块,然后利用滤波器对预测块进行滤波的方法。利用固定规则的帧内预测方法仅使用有限的信息,因此可能不适合如上所述处理复杂内容。例如,帧内预测的方向模式在选择的方向以外的方向上不使用参考样本。为了解决这个问题,可以对预测块应用滤波。然而,当利用固定滤波器执行滤波时,存在相邻参考样本和整个预测块的信息不能被考虑的问题。

[0008] 作为其他技术,存在通过根据固定规则和滤波器替换帧内预测,利用深度学习技术,在接收相邻参考样本时生成预测块的帧内预测方法。这些技术提出了一种深度学习网络,用于利用相邻参考样本作为输入来生成精确的预测块。因此,基于深度学习的预测方法可以被添加为视频编解码装置中的帧内预测模式的一种,或者替换所有的帧内预测模式。然而,由于基于深度学习的预测方法没有根据现有的帧内预测来利用预测块的信息,存在生成的预测的信号可能不准确的问题。例如,根据现有帧内预测的方向模式生成的预测块是基于方向性来预测的,因此它可以用作当前块的近似结构性信息,但是基于深度学习的预测方法不利用这样的信息。

[0009] 作为另一种技术,存在基于深度学习的预测方法,该方法使用方向模式下预测的

相邻参考样本和块,同时利用深度学习技术。尽管除了参考样本之外,在利用预测块的信息方面存在精细化,但是该方法也具有不适应于与参考样本和预测块相关的信息以及需要额外的信号通知来指示该方法的应用的缺点。

[0010] 因此,为了通过改进预测块来提高视频编解码效率并提供视频质量,利用基于深度学习的帧内预测方法,该方法同时使用相邻参考样本和预测块。此外,还需要考虑自适应地使用与参考样本和预测块相关的信息的方法。

发明内容

[0011] 技术问题

[0012] 本发明旨在提供一种视频编解码方法和装置,用于生成精细化的帧内预测的信号,以提高视频编解码效率和视频质量。视频编解码方法和装置可以根据当前块的帧内预测的预测块、与当前块相关的块信息以及相邻重构的参考样本自适应地输入到深度学习网络中。

[0013] 技术方案

[0014] 本发明的至少一个方面提供了一种视频解码装置。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信息和当前块的残差值。当前块的信息包括当前块的帧内预测模式、当前块的高度、宽度或纵横比、当前块的颜色、包括当前块的图像的类型以及当前块的量化参数。视频解码装置还包括帧内预测器,其配置为利用帧内预测模式从当前块周围的重构的参考样本生成当前块的预测块。视频解码装置还包括预测的信号精细化单元,其配置为基于当前块的信息的全部或一些自适应地选择基于深度学习的精细化模型,并且将包括预测块和重构的参考样本的输入块输入到选择的基于深度学习的精细化模型中,以生成精细化的预测块。视频解码装置还包括加法器,其配置为将残差值添加到精细化的预测块,以生成当前块的重构的块。

[0015] 本发明的另一个方面提供了由视频解码装置执行的对当前块执行帧内预测的视频解码方法。视频解码方法包括从比特流解码当前块的信息和当前块的残差值。当前块的信息包括当前块的帧内预测模式、当前块的高度、宽度或纵横比、当前块的颜色、包括当前块的图像的类型以及当前块的量化参数。视频解码方法还包括利用帧内预测模式从当前块周围的重构的参考样本生成当前块的预测块。视频解码方法还包括基于当前块的信息的全部或一些自适应地选择基于深度学习的精细化模型,并且通过将包括预测块和重构的参考样本的输入块输入到选择的基于深度学习的精细化模型中来生成精细化的预测块。视频解码方法还包括通过将残差值添加到精细化的预测块来生成当前块的重构的块。

[0016] 本发明的又一个方面提供了由视频解码装置执行的对当前块进行帧内预测的视频编码方法。视频编码方法包括从高级别获得当前块的信息。当前块的信息包括当前块的帧内预测模式、当前块的高度、宽度或纵横比、当前块的颜色、包括当前块的图像的类型以及当前块的量化参数。视频编码方法还包括利用帧内预测模式从当前块周围的重构的参考样本生成当前块的预测块。视频编码方法还包括基于当前块的信息的全部或一些自适应地选择基于深度学习的精细化模型,并且通过将包括预测块和重构的参考样本的输入块输入到选择的基于深度学习的精细化模型中来生成精细化的预测块。视频编码方法还包括通过将当前块减去精细化的预测块来生成残差块。

[0017] 有益效果

[0018] 如上所述,本发明提供了一种视频编解码方法和装置,通过将根据当前块的帧内预测的预测块、与当前块相关的块信息以及相邻重构的参考样本自适应地输入到深度学习网络中来生成精细化的帧内预测的信号。因此,视频编解码方法和装置可以提高视频编解码效率和视频质量。

附图说明

[0019] 图1是可以实现本发明技术的视频编码装置的框图。

[0020] 图2示出利用四叉树加二叉树三叉树(QTBT)结构来对块进行分区的方法。

[0021] 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。

[0022] 图4示出当前块的相邻块。

[0023] 图5是可以实现本发明技术的视频解码装置的框图。

[0024] 图6是根据本发明实施方案的利用帧内预测的信号的精细化的视频编码装置的示意性框图。

[0025] 图7是根据本发明实施方案的利用帧内预测的信号的精细化的视频解码装置的示意性框图。

[0026] 图8是示出根据本发明实施方案的包括全连接层的精细化模型的示意图。

[0027] 图9a和图9b是示出根据本发明实施方案的用于帧内预测模式的输入样本的示意图。

[0028] 图10是示出根据本发明另一个实施方案的用于帧内预测模式的输入样本的示意图。

[0029] 图11是示出根据本发明实施方案的利用预处理过程的精细化模型的示意图。

[0030] 图12是示出根据本发明实施方案的用于输入的参考样本的确定的示意图。

[0031] 图13是示出根据本发明另一个实施方案的扩展的精细化模型的示意图。

[0032] 图14是示出根据本发明另一个实施方案的在帧内预测的方向模式下自适应地操作的自适应精细化模型的示意图。

[0033] 图15是示出根据本发明实施方案的基于卷积神经网络(convolutional neural network, CNN)的精细化模型的示意图。

[0034] 图16是示出根据本发明实施方案的利用帧内预测的信号的精细化的视频编码方法的流程图。

[0035] 图17是示出根据本发明实施方案的利用帧内预测的信号的精细化的视频解码方法的流程图。

具体实施方式

[0036] 在下文中,将参考所附说明性附图对本发明的一些实施方案进行详细描述。在下面的描述中,相同的附图标记表示相同的元件,尽管元件在不同的附图中示出。此外,在一些实施方案的以下描述中,当认为相关的已知组件和功能的详细描述模糊了本发明的主题时,为了清楚和简洁起见,可以省略对相关的已知组件和功能的详细描述。

[0037] 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。

[0038] 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。

[0039] 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。

[0040] 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)。应用于每个编码单元(CU)的信息被编码为CU的语法,并且共同应用于包括在一个CTU中的CU的信息被编码为CTU的语法。另外,共同应用于一个切片中的所有块的信息被编码为切片头的语法,而应用于构成一个或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。

[0041] 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。

[0042] 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。

[0043] 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,QTBTBT)结构。这里,二叉树三叉树(BTBT)被添加到树结构以称为多类型树(multiple-type tree,MTT)。

[0044] 图2是用于描述通过利用QTBTBT结构来对块进行分割的方法的示意图。

[0045] 如图2所示,CTU可以首先分割为QT结构。四叉树分割可以是递归的,直到分割块的大小达到QT中允许的叶节点的最小块大小(MinQTSize)。由熵编码器155对指示QT结构的每个节点是否被分割为下层的四个节点的第一标志(QT_split_flag)进行编码,并将其用信号通知视频解码装置。当QT的叶节点不大于BT中允许的根节点的最大块大小(MaxBTSize)时,可以将叶节点进一步分割为BT结构或TT结构的至少一个。在BT结构和/或TT结构中可能存在多个分割方向。例如,可以存在两个方向,即,水平地分割相应节点的块的方向以及竖

直地分割相应节点的块的方向。如图2所示,当MTT分割开始时,由熵编码器155对指示节点是否被分割的第二标志(mtt_split_flag)、以及在节点被分割的情况下额外指示分割方向(竖直或水平)的标志和/或指示分割类型(二叉或三叉)的标志进行编码,并将其用信号通知视频解码装置。

[0046] 替代地,在对指示每个节点是否被分割为下层的四个节点的第一标志(QT_split_flag)进行编码之前,还可以对指示节点是否被分割的CU分割标志(split_cu_flag)进行编码。当CU分割标志(split_cu_flag)的值指示出每个节点没有被分割时,相应节点的块成为分割树结构中的叶节点并成为CU,CU是编码的基本单元。当CU分割标志(split_cu_flag)的值指示出每个节点被分割时,视频编码装置以上述方案对第一标志首先开始编码。

[0047] 当QTBT用作树结构的另一个示例时,可以存在两种类型,即将相应节点的块水平地分割为具有相同大小的两个块的类型(即,对称水平分割)以及将相应节点的块竖直地分割为具有相同大小的两个块的类型(即,对称竖直分割)。由熵编码器155对指示BT结构的每个节点是否被分割为下层的块的分割标志(split_flag)和指示分割类型的分割类型信息进行编码,并将其传送至视频解码装置。另一方面,可以另外存在相应节点的块被分割为彼此非对称的两个块的类型。非对称形式可以包括相应节点的块被分割为大小比率为1:3的两个矩形块的形式,或者还可以包括相应节点的块在对角线方向上被分割的形式。

[0048] CU可以根据从CTU分割的QTBT或QTBTBT而具有各种大小。在下文中,与要编码或解码的CU(即,QTBTBT的叶节点)相对应的块称为“当前块”。当采用QTBTBT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。

[0049] 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。

[0050] 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。

[0051] 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。

[0052] 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。

[0053] 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。

[0054] 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。

[0055] 帧间预测器124通过利用运动补偿处理来生成当前块的预测块。帧间预测器124在比当前图像更早已被编码和解码的参考图像中搜索与当前块最相似的块,并且通过利用搜索到的块来生成当前块的预测块。另外,生成运动矢量(motion vector,MV),所述运动矢量与当前图像中的当前块和参考图像中的预测块之间的位移(displacement)相对应。通常,对亮度(luma)分量执行运动估计,并且基于亮度分量计算的运动矢量用于亮度分量和色度分量两者。由熵编码器155对包括参考图像的信息和关于用于预测当前块的运动矢量的信息的运动信息进行编码,并将其传送至视频解码装置。

[0056] 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。

[0057] 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包括关于两个运动矢量的信息的运动信息被传送至熵编码器155。这里,参考图像列表0可以由预重构图像中按显示顺序在当前图像之前的图像构成,并且参考图像列表1可以由预重构图像中按显示顺序在当前图像之后的图像构成。然而,尽管不特别限于此,但可以将显示顺序中在当前图像之后的预重构图像额外地包括在参考图像列表0中。相反,在当前图像之前的预重构图像也可以额外地包括在参考图像列表1中。

[0058] 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。

[0059] 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)。

[0060] 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”)。

[0061] 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。

[0062] 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。

[0063] 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。

[0064] 此后,合并模式和合并跳过模式统称为合并/跳过模式。

[0065] 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。

[0066] 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。

[0067] 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。

[0068] 可以通过将预定义函数(例如,中值和平均值计算等)应用于运动矢量预测候选来获取运动矢量预测。在这种情况下,视频解码装置还知道预定义功能。此外,由于用于推导运动矢量预测候选的相邻块是已经完成编码和解码的块,所以视频解码装置也可能已经知道相邻块的运动矢量。因此,视频编码装置不需要对用于识别运动矢量预测候选的信息进行编码。相应地,在这种情况下,对于关于运动矢量差的信息和关于用于预测当前块的参考图像的信息进行编码。

[0069] 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。

[0070] 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。

[0071] 变换器140将具有空域的像素值的残差块中的残差信号变换为频域的变换系数。变换器140可以通过利用残差块的整个大小作为变换单元来变换残差块中的残差信号,或者也可以将残差块分割为多个子块,并且可以通过利用子块作为变换单元来执行变换。替代地,将残差块划分成两个子块,即变换区域和非变换区域,以通过仅利用变换区域子块作为变换单元来变换残差信号。这里,变换区域子块可以是基于水平轴线(或垂直轴线)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(或垂直轴线)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。

[0072] 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,并且可以在水平方向和竖直方向的每个上变换残差块。由熵编码器155对关于MTS中的变换函数对的信息(mts_idx)进行编码,并将其用信号通知视频解码装置。

[0073] 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。

[0074] 重排单元150可以对量化的残差值执行系数值的重新排列。

[0075] 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。

[0076] 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。

[0077] 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。

[0078] 逆量化器160对从量化器145输出的量化的变换系数进行逆量化,以生成变换系数。逆变换器165将从逆量化器160输出的变换系数从频域变换到空域,以重构残差块。

[0079] 加法器170将重构的残差块和由预测器120生成的预测块相加,以重构当前块。在对下一个块进行帧内预测时,重构的当前块中的像素用作参考像素。

[0080] 环路滤波单元180对重构的像素执行滤波,以减少由于基于块的预测和变换/量化而发生的块伪影(blocking artifacts)、振铃伪影(ringing artifacts)、模糊伪影(blurring artifacts)等。作为环内滤波器的环路滤波单元180可以包括去块滤波器182、样本自适应偏移(sample adaptive offset, SAO)滤波器184和自适应环路滤波器(adaptive loop filter, ALF)186的全部或一些。

[0081] 去块滤波器182对重构的块之间的边界进行滤波,以去除由于块单元编码/解码而发生的块伪影(blocking artifacts),并且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)而发生的重构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,以增强主观图像质量和编码效率。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。

[0082] 通过去块滤波器182、SAO滤波器184和ALF 186滤波的重构的块存储在存储器190中。当一个图像中的所有块被重构时,重构的图像可以用作用于对随后要被编码的图像内的块进行帧间预测的参考图像。

[0083] 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。

[0084] 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。

[0085] 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。

[0086] 熵解码器510通过解码由视频编码装置生成的比特流来提取与块分割相关的信息,以确定要解码的当前块,并且提取重构当前块所需的预测信息和关于残差信号的信息。

[0087] 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),并且提取CTU的分割信息以通过利用树结构来分割CTU。

[0088] 例如,当通过利用QTBT结构来分割CTU时,首先提取与QT的分割相关的第一标志(QT_split_flag),以将每个节点分割为下层的四个节点。另外,相对于与QT的叶节点相对应的节点来提取与MTT的分割相关的第二标志(mtt_split_flag)、分割方向(竖直/水平)和/或分割类型(二叉/三叉),以将相应的叶节点分割为MTT结构。结果,QT的叶节点下方的每个节点被递归地分割为BT或TT结构。

[0089] 作为另一个示例,当通过利用QTBT结构来分割CTU时,提取指示是否分割CU的CU分割标志(split_cu_flag)。当分割相应块时,还可以提取第一标志(QT_split_flag)。在分割过程期间,对于每个节点,在递归QT分割0次或更多次之后可能发生0次或更多次的递归

MTT分割。例如,对于CTU,MTT分割可以立即发生,或者相反,也可以仅发生多次QT分割。

[0090] 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。

[0091] 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。

[0092] 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。

[0093] 重排单元515可以以由视频编码装置执行的系数扫描顺序的相反顺序,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)。

[0094] 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。

[0095] 逆变换器530通过将逆量化的变换系数从频域逆变换到空域来重构残差信号,以生成当前块的残差块。

[0096] 此外,当逆变换器530对变换块的部分区域(子块)进行逆变换时,逆变换器530提取仅对变换块的子块进行变换的标志(cu_sbt_flag)、子块的方向(竖直/水平)信息(cu_sbt_horizontal_flag)和/或子块的位置信息(cu_sbt_pos_flag)。逆变换器530还将相应子块的变换系数从频域逆变换到空域以重构残差信号,并且用值“0”填充未被逆变换的区域作为残差信号,以生成当前块的最终残差块。

[0097] 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。

[0098] 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。

[0099] 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。

[0100] 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。

[0101] 加法器550通过将从逆变换器530输出的残差块与从帧间预测器544或帧内预测器542输出的预测块相加来重构当前块。在对后续要被解码的块进行帧内预测时,重构的当前块内的像素用作参考像素。

[0102] 作为环内滤波器的环路滤波单元560可以包括去块滤波器562、SAO滤波器564和ALF 566。去块滤波器562对重构的块之间的边界执行去块滤波,以便去除由于块单元解码而发生的块伪影。SAO滤波器564和ALF 566在去块滤波之后对重构的块执行额外的滤波,以便补偿由于有损编码而发生的重构的像素与原始像素之间的差异。通过利用关于从比特流解码的滤波器系数的信息来确定ALF的滤波器系数。

[0103] 通过去块滤波器562、SAO滤波器564和ALF 566滤波的重构的块存储在存储器570中。当一个图像中的所有块被重构时,重构的图像可以用作用于对随后要被编码的图像内的块进行帧间预测的参考图像。

[0104] 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,通过将根据当前块的帧内预测的预测块、与当前块相关的块信息以及相邻重构的参考样本自适应地输入到深度学习网络中来生成精细化的帧内预测的信号。

[0105] 以下实施方案可以普遍应用于视频编码装置和使用视频编码装置的深度学习技术的部件。

[0106] 在下面的描述中,术语待编码/解码的“目标块”可以以与如上所述的当前块或编码单元(CU)相同的意义使用,或者术语“目标块”可以意味着编码单元的部分区域。

[0107] 此外,块的纵横比被定义为通过将块的水平长度除以竖直长度而获得的值。

[0108] I. 帧内预测模式

[0109] 如上所述,帧内预测是通过参考当前要编码的块周围存在的样本来预测当前块的方法。在多功能视频编解码(VVC)技术中,帧内预测模式除了非方向模式(即,平面和DC)之外,还具有细分的方向模式(即,2至66),如图3a所示。此外,如添加到图3b的示例,亮度块的帧内预测模式具有根据宽角度帧内预测(wide-angle intra prediction,WAIP)的方向模式(-14到-1以及67到80)。在这种帧内预测中,根据方向模式的预测块可以表示当前块的近似结构性信息。

[0110] 此外,帧内预测可以使用预测技术,例如多参考行帧内预测(multiple reference line intra prediction,MRLP)、交叉分量线性模型(cross-component linear model,CCLM)、位置相关帧内预测组合(position dependent intra prediction combination,PDPC)、帧内子分区(intra sub-partitions,ISP)、基于矩阵的帧内预测(matrix-based intra prediction,MIP)等。

[0111] 在利用MRLP的帧内预测过程中,视频编码/解码装置可以通过利用多参考行(multiple reference line,MRL)来使用更多的参考行。当应用MRL时,视频编码/解码装置可以利用除了原始参考行之外添加到当前块的顶部和左侧的两条线的样本来对当前块执行帧内预测。为了在应用MRL时选择参考行,可以将指示参考行的索引(mrl_idx)用信号通知视频解码装置。

[0112] CCLM预测是帧内预测方法,其使用代表亮度信号与色度信号之间的相似性的线性模型。为了激活CCLM模式,视频编码装置可以将激活CCLM模式的标志用信号通知视频解码装置。

[0113] 在CCLM预测中,首先,推导当前色度块的相邻参考样本与存在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。这里,可以基于相邻亮度信号的最小

值、在与周边亮度信号相同位置的色度值、相邻亮度信号的最大值以及在与周边亮度信号相同位置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在与色度块相同位置的亮度样本来执行色度样本的预测。

[0114] 在帧内预测中,基于规则的预测方法的一种是位置相关帧内预测组合(dependent intra prediction combination,PDPC)。换句话说,可以通过利用执行帧内预测的目标块和与目标块在空间上邻近的相邻像素的编码信息,基于预定义的操作来生成预测因子。

[0115] PDPC修改根据特定帧内预测模式生成的预测样本,以生成当前块的帧内预测因子。这里,在图3a所示的预测模式中,特定的帧内预测模式包括平面、DC、水平(预测模式#18)、竖直(预测模式#50)、左下对角线方向模式(预测模式#2)及其附近的15个方向模式、以及右上对角线方向模式(预测模式#66)及其附近的15个方向模式。

[0116] 在PDPC中,对于根据特定帧内预测模式生成的当前块的预测样本,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的值,以生成预测样本。

[0117] ISP技术将当前块细分为相同大小的小块,然后在所有子块之间共享当前块的帧内预测模式。然而,ISP技术可以对每个子块应用不同的变换。在细分的情况下,可以水平地或竖直地划分块。

[0118] 可以基于利用执行帧内预测的当前块的相邻像素和当前块的编码信息的预定义的矩阵运算来生成预测因子。这种基于规则的预测方法称为MIP。

[0119] MIP通过利用预定义的矩阵运算来生成帧内预测因子的全部或一部分。当生成预测因子的一部分时,MIP可以通过利用预测因子的部分额外地执行用于升尺度的上采样或插置来生成等于当前块的大小的最终帧内预测样本。

[0120] 另一方面,MIP可以选择性地选择与当前块在空间上邻近的一些像素,并且可以将它们用作当前块的相邻像素。在另一个实施方案中,MIP可以使用根据基于用于矩阵运算的方法(例如,子采样、降尺度等)的运算而推导出的值。

[0121] 最可能模式(Most probable mode,MPM)技术在当前块的帧内预测期间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。

[0122] 另一方面,帧内块复制(intra block copy,IBC)是在执行帧内预测时,在与当前块的预测块相同的图像中生成参考块而不是利用参考样本的技术。这里,块矢量是指示参考块的位移,并且将块矢量从视频编码装置用信号通知视频解码装置。

[0123] II. 帧内预测的预测的信号精细化

[0124] 图6是根据本发明实施方案的利用帧内预测的信号的精细化的视频编码装置的示意性框图。

[0125] 根据本实施方案的视频编码装置额外地包括位于帧内预测器122(其是现有组件)的后方的预测的信号精细化单元610。这里,包括在根据本实施方案的视频编码装置中的组件不一定限于此。例如,视频编码装置可以额外地包括用于训练包括在预测的信号精细化单元610中的基于深度学习的精细化模型的训练单元(未示出),或者可以以链接到外部训练单元的形式实现。

[0126] 帧内预测器122利用要编码的当前块的帧内预测模式,从相邻参考样本生成包括当前块的预测的信号的预测块。

[0127] 预测的信号精细化单元610通过利用精细化模型,从预测块和相邻重构的参考样本生成精细化的预测块。另一方面,训练单元可以训练精细化模型,使得精细化模型可以学习精细化的信号生成方法,以生成接近当前块的原始信号的精细化的信号。

[0128] 在下文中,术语精细化的预测块可以与术语“精细化的块”互换使用。此外,重构的参考样本可以与“重构的参考样本”互换使用。

[0129] 视频编码装置通过将当前块减去精细化的块来生成残差块。视频编码装置可以将残差块的残差值输入到变换器140,并且可以执行如上所述的编码处理。

[0130] 图7是根据本发明实施方案的利用帧内预测的信号的精细化的视频解码装置的示意性框图。

[0131] 根据本实施方案的视频解码装置额外地包括位于帧内预测器542(其是现有组件)的后方的预测的信号精细化单元710。

[0132] 熵解码器510从比特流解码残差块,所述残差块包括要解码的当前块的帧内预测模式和残差值。

[0133] 帧内预测器542通过利用要解码的当前块的帧内预测模式,从相邻参考样本生成包括当前块的预测的信号的预测块。

[0134] 预测的信号精细化单元710利用精细化模型,从预测块和相邻重构的参考样本生成精细化的块。视频解码装置通过将残差值添加到精细化的块来生成当前块的重构的块。

[0135] 在下文中,对视频编码装置内的精细化模型及其学习方法进行描述。以下描述可以同样应用于视频解码装置内的精细化模型。

[0136] 作为精细化预测块的深度学习网络,精细化模型f可以由等式1表示。

[0137] [等式1]

$$[0138] \quad x_{\text{hat}_p} = f(x_p, x_r)$$

[0139] 精细化模型f接收由帧内预测和相邻重构的参考样本 x_r 生成的预测块 x_p 作为输入,并且生成精细化的预测块 x_{hat_p} 。

[0140] 另一方面,训练单元更新精细化模型的参数,使得精细化模型可以学习如何基于如等式2所示的损失函数L生成接近原始块 y_p 的输出,因此可以训练精细化模型。

[0141] [等式2]

$$[0142] \quad L = \sum |y_p - x_{\text{hat}_p}|$$

[0143] 这里,等式2中所示的损失函数使用预测块与原始块之间的L1损失,但不一定限于此。换句话说,可以使用可以表达预测块与原始块之间的差异的任何损失函数。

[0144] 在本实施方案中,在如等式3中定义块中的自适应精细化模型f之后,训练单元可以根据用于学习的块的信息,利用等式2中所示的损失函数来训练不同的精细化模型。

[0145] [等式3]

$$[0146] \quad x_{\text{hat}_p} = f(x_p, x_r, m, b, l, p, q)$$

[0147] 这里,m、b、l、p和q表示块的信息。M表示块的帧内预测模式,并且b表示块的大小或纵横比(长与宽之比)。l表示块是亮度块还是色度块,并且p表示包括块的图像类型。q表示块的量化参数(quantization parameter, QP)。

[0148] 此后,视频编码装置可以通过基于从高级别传输的当前块的信息应用不同训练的精细化模型来生成精细化的块。

[0149] 在以下实施方案中,应用不同的精细化模型 f 意味着利用不同的滤波器输出具有不同值的块。还包括应用或不应用 f 的含义。这里,不应用 f 与应用 f 具有相同的含义, f 的操作类似于单位矩阵,在不改变输入信号的情况下按原样输出输入信号。

[0150] 在下文中,对根据块信息 m 、 b 、 l 、 p 和 q 的实施方案进行描述。另一方面,可以组合一个或更多个实施方案。换句话说,可以使用块信息 m 、 b 、 l 、 p 和 q 的全部或一些。这里,块可以代表训练过程中用于学习的块或帧内预测过程中的当前块。

[0151] 视频编码装置可以根据预测模式 m 是否是方向的来应用不同的 f 。视频编码装置可以根据预测模式 m 是否是基于矩阵的预测(例如,MIP)来应用不同的 f 。视频编码装置可以根据预测模式 m 是否包括在WAIP中来应用不同的 f 。视频编码装置可以对方向预测模式 m 进行聚类,并且根据方向预测模式 m 包括在竖直模式、水平模式和对角线模式中的哪个聚类来应用不同的 f 。此外,视频编码装置可以基于预测模式 m 是竖直预测还是水平预测来应用不同的 f 。

[0152] 如果 b 是块的水平长度或竖直长度,则视频编码装置可以将 b 与预设阈值进行比较,并且根据结果来应用不同的 f 。如果 b 是块的纵横比,则视频编码装置可以将其与预设阈值进行比较,并且根据结果来应用不同的 f 。

[0153] 视频编码装置可以根据值 l (即,块是亮度块还是色度块)来应用不同的 f 。

[0154] 视频编码装置可以根据 p 的类型,即,图像是I(帧内)、B(双向预测)还是P(预测)图像来应用不同的 f 。

[0155] 视频编码装置可以根据值 q ,即,QP值或其范围来应用不同的 f 。

[0156] 在下文中,对用于精细化预测块的精细化模型的结构和操作进行描述。

[0157] 图8是示出根据本发明实施方案的包括全连接层的精细化模型的示意图。

[0158] 如图8所示,精细化模型可以是包括三个全连接(fully-connected,FC)层的深度学习网络。在每个全连接层之后,存在修正线性单元(rectified linear unit,ReLU)层。作为精细化模型的输入,使用大小 $(h+r) \times (w+r)$ 的输入块,该输入块包括根据帧内预测模式确定的预测块和相邻重构的参考样本。这里, h 是当前块的高度(竖直长度),并且 w 是当前块的宽度(水平长度)。另外, r 代表左侧重构的参考样本的宽度或顶部重构的参考样本的高度。这些参考样本可以位于当前块的左侧和顶部。

[0159] 输入块被展平成大小为 $(h+r) \times (w+r) \times 1$ 的一维矢量,然后被输入到精细化模型中。精细化模型输出大小为 $hw \times 1$ 的一维矢量,并且可以通过最终重塑一维矢量来生成精细化的块。这里,当基于根据特定帧内预测模式(例如,竖直预测、水平预测或类似MIP的基于矩阵的预测)生成的预测块来训练精细化模型时,预测块被精细化以适合该预测模式的特征。

[0160] 另一方面,在利用具有图8所示的结构精细化模型的情况下,不按原样输入大小为 $(h+r) \times (w+r)$ 的块,并且可以根据帧内预测模式不同地选择输入样本。

[0161] 图9a和图9b是示出根据本发明实施方案的用于帧内预测模式的输入样本的示意图。

[0162] 如图9a所示,可以根据帧内预测模式来选择精细化模型的输入样本。

[0163] 在竖直预测的情况下,利用顶部参考样本生成预测块。因此,在生成预测块时没有使用的大小为 $(h+r) \times r$ 的左侧参考样本以及大小为 $h \times w$ 的预测块可以用作精细化模型的输入。

[0164] 另一方面,在水平预测的情况下,利用左侧参考样本来生成预测块。因此,在生成预测块时没有使用的大小为 $r \times (w+r)$ 的顶部参考样本以及大小为 $h \times w$ 的预测块可以用作精细化模型的输入。

[0165] 替代地,如在图9b的示例中,可以相反地选择输入样本。

[0166] 在竖直预测的情况下,利用顶部参考样本来生成预测块。相应地,大小为 $r \times (w+r)$ 的顶部参考样本和大小为 $h \times w$ 的预测块可以用作精细化模型的输入。

[0167] 另一方面,在水平预测的情况下,利用左侧参考样本来生成预测块。相应地,大小为 $(h+r) \times r$ 的左侧参考样本和大小为 $h \times w$ 的预测块可以用作精细化模型的输入。

[0168] 作为另一个示例,对于竖直预测和水平预测,邻近当前块的大小为 $(h+r) \times r$ 的左侧参考样本、邻近当前块的大小为 $(w \times r)$ 的顶部参考样本以及大小为 $h \times w$ 的预测块可以用作精细化模型的输入。

[0169] 基于上述方法,如在图10的示例中,左侧和顶部参考样本的数量可以设置为对于竖直方向模式或水平方向模式是不同的。在图10的示例中, r_w 代表基于预测块的左侧参考样本线的数量,并且 r_h 代表顶部参考样本的线的数量。

[0170] 在竖直预测的情况下,可以组合精细化模型的输入,使得当生成输入块时使用的顶部参考样本线等于或大于左侧参考样本线($r_h \geq r_w$)。换句话说,大小为 $(h+r_h) \times r_w$ 的左侧参考样本、大小为 $r_h \times w$ 的顶部参考样本、以及大小为 $h \times w$ 的预测块可以用作精细化模型的输入。

[0171] 另一方面,在水平预测的情况下,可以组合精细化模型的输入,使得当生成输入块时使用的左侧参考样本线等于或大于顶部参考样本线($r_h \leq r_w$)。换句话说,大小为 $r_h \times (w+r_w)$ 的顶部参考样本、大小为 $h \times r_w$ 的左侧参考样本以及大小为 $h \times w$ 的预测块可以用作精细化模型的输入。

[0172] 作为另一个示例,当利用图8所示的精细化模型时,大小为 $(h+r) \times (w+r)$ 的块可以不按原样输入,但是可以根据预测模式来应用诸如旋转的预处理过程。例如,如图11的示例所示,可以通过在帧内预测期间以方向模式对竖直预测块应用竖直镜像和90度顺时针旋转来生成水平预测块。利用预处理过程和针对水平预测块自适应训练的精细化模型,视频编码装置可以执行竖直预测块的滤波。

[0173] 作为另一个示例,当任意方向预测模式应用于图8所示的精细化模型时,在帧内预测模式中使用的同时包括在参考样本线中的样本包括在精细化模型的输入块中。例如,在MRLP中,从预测块的顶部和左侧开始的第一行、第二行和第三行可以是当前块的参考样本线。

[0174] 作为另一个示例,如图12所示,利用确定的左侧样本和方向模式来确定在向上方向上的参考样本,然后确定的样本包括在精细化模型的输入块中。类似地,在利用顶部样本和方向模式来确定在左侧方向上的参考样本之后,确定的样本包括在精细化模型的输入块中。

[0175] 作为另一个示例,当任意方向预测模式应用于图8所示的精细化模型时,根据方向性的斜率,预测模式可以映射到水平方向或竖直方向的更接近的一个。此后,可以以图9a、图9b和图10所示的方式来确定参考样本。

[0176] 作为另一个示例,在利用图8所示的精细化模型的情况下,可以不同地选择输入样

本,而不是按原样输入大小为 $(h+r) \times (w+r)$ 的块,但是一些样本可以被填补,然后被输入。例如,如果参考样本不能包括在CTU的块边界处,则可以替代地包括其他可用的参考样本,或者构成预测块的一些样本可以包括在精细化模型的输入中。

[0177] 作为另一个示例,在利用图8所示的精细化模型的情况下,视频编码装置可以将指示是否使用先前训练的精细化模型的参数的标志用信号通知视频解码装置。这里,视频解码装置可以通过组合预定的精细化模型的参数和标志来对预测块执行滤波。

[0178] 作为另一个示例,在利用图8所示的精细化模型的情况下,代替从视频编码装置接收指示是否使用先前训练的精细化模型的参数的标志,视频解码装置可以根据先前接收到的编码信息来确定是否使用精细化模型。换句话说,视频解码装置通过组合根据当前块的信息 m 、 b 、 l 、 p 和 q 的全部或一些确定的预定的参数和条件来执行滤波。例如,如果当前块的预测模式 m 是竖直模式或水平模式,则视频解码装置可以应用精细化模型。

[0179] 作为另一个示例,在利用图8所示的精细化模型的情况下,视频编码装置可以使用各种精细化模型的一个,而不是一个精细化模型。因此,视频编码装置可以传输指示多个精细化模型的一个的索引,并且视频解码装置可以组合接收到的索引和指示是否使用精细化模型的参数的标志,并且对预测块执行滤波。

[0180] 作为另一个示例,在利用图8所示的精细化模型的情况下,视频编码装置可以选择各种精细化模型的一个,而不是利用一个精细化模型。这里,视频编码装置可以根据先前从高级别传输的当前块的编码信息来确定在多个精细化模型中使用哪个精细化模型。类似地,视频解码装置可以根据当前块的信息 m 、 b 、 l 、 p 和 q 的全部或一些来确定多个精细化模型中的一个精细化模型,并且对预测块执行滤波。

[0181] 作为另一个示例,在利用图8所示的精细化模型的情况下,视频编码装置可以将针对每个序列单独训练的精细化模型的参数的全部或一些发送到视频解码装置。这里,一些参数可以是构成精细化模型的一些层的参数,或者可以代表滤波器中的诸如偏移的参数。这里,代替利用预定的参数,视频解码装置可以用通过比特流接收到的参数来配置精细化模型,然后对预测块执行滤波。

[0182] 以上给出的描述可以同样地应用于视频解码装置内的精细化模型。

[0183] 图13是示出根据本发明另一个实施方案的扩展的精细化模型的示意图。

[0184] 可以通过向如图8所示的精细化模型添加卷积层和空间注意力来扩展精细化模型。扩展的精细化模型可以解决在利用现有的方向模式执行帧内预测时没有使用其他方向的参考样本的信息的问题。例如,如图13所示,扩展的精细化模型可以应用于根据方向模式中的水平方向预测而确定的 8×8 预测块。通过将预测块输入到卷积层而获得的特征图、从预测块的顶部的参考样本生成的特征矢量以及由图8所示的精细化模型生成的块相乘,可以生成考虑了其他方向的参考样本的信息的精细化的块。在图13的示例中,生成特征矢量的分量包括卷积层、池化层和激活层。扩展的精细化模型可以利用基于左侧参考样本在水平方向上预测的预测块的注意力结构来反映顶部参考样本的特征。最后,图13所示的扩展的精细化模型可以生成精细化的预测块。

[0185] 图14是示出根据本发明另一个实施方案的在帧内预测的方向模式下自适应地操作的自适应精细化模型的示意图。

[0186] 图14所示的自适应精细化模型包括以下四个组件。第一组件①从多个边界样本提

取特征。该组件是从预测块周围的参考样本线提取特征的 1×1 卷积层。第二组件②提取预测块的特征。如上所述,在方向模式中预测的角度预测区域可以表示当前块的近似结构性信息。因此,对于在方向模式下预测的预测块,利用卷积层来提取对应于结构性信息的特征。

[0187] 第三组件③利用注意力结构来融合从第一组件和第二组件提取的特征。这里,注意力结构通过在精细化预测块的每个像素时控制相邻参考样本的作用来允许高效地使用空间信息。最后,第四组件④是输出层,它利用卷积层从第三组件的输出生成精细化的块。

[0188] 图15是示出根据本发明实施方案的基于CNN的精细化模型的示意图。

[0189] 基于卷积神经网络(CNN)的精细化模型可以通过利用预测块和相邻重构的参考样本作为输入来执行基于CNN的图像质量精细化,从而提高当前块的图像质量。在图15的示例中,输入块包括相邻重构的参考样本和预测样本。这里,相邻参考样本可以包括编解码误差,并且预测样本可以包括预测误差和编解码误差。此外,基于CNN的精细化模型生成具有与输入块相同大小的输出块。因此,在CNN网络的训练期间,训练单元可以利用诸如等式4所示的总损失函数(Total Loss Function, TL)。

[0190] [等式4]

[0191] $TL = \alpha \cdot L(\text{输出}_{\text{blk}}, GT_{\text{blk}}) + (1 - \alpha) \cdot L(\text{输出}_{\text{blk}} \cdot \text{掩码}, GT_{\text{blk}} \cdot \text{掩码})$

[0192] 这里,掩码是矩阵,其中掩码的大小与输入块的大小相同,并且用于预测样本的掩码值为1,而其余样本的掩码值为0。基准(Ground truth, GT)代表用于学习的目标块。训练单元可以计算相邻参考样本和预测样本的每个的损失,然后可以使用利用权重的加权和作为总损失。通过将权重 α 设置为与相邻参考样本的面积和预测样本的面积之和成反比,可以精细化等式4中第二项的重要性。这种基于CNN的精细化模型的优点是,即使输入块的大小改变,也能够运行。

[0193] 在下文中,利用图16和图17的图示对基于深度学习的精细化帧内预测的信号的视频编码方法和视频解码方法进行描述。

[0194] 图16是示出根据本发明实施方案的利用帧内预测的信号的精细化的视频编码方法的流程图。

[0195] 视频编码装置从高级别获得当前块的信息(S1600)。这里,当前块的信息包括:当前块的帧内预测模式 m 、关于当前块的高度、宽度或纵横比的信息 b 、当前块的颜色 l 、包括当前块的当前图像的类型 p 以及当前块的量化参数 q 。

[0196] 视频编码装置利用帧内预测模式从当前块周围的重构的参考样本生成当前块的预测块(S1602)。

[0197] 视频编码装置基于当前块的信息的全部或部分来自适应地选择基于深度学习的精细化模型,并且通过将包括预测块和重构的参考样本的输入块输入到选择的精细化模型中来生成精细化的预测块(S1604)。

[0198] 对于块自适应精细化模型 f ,如等式3所示,训练单元可以根据用于学习的块的信息来预训练不同的精细化模型。此后,基于当前块的输入信息,视频编码装置可以通过将输入块应用于不同训练的精细化模型来生成精细化的预测块。

[0199] 在下文中,对根据当前块的信息 m 、 b 、 l 、 p 和 q 的实施方案进行描述。另一方面,可以组合一个或更多个实施方案。换句话说,可以使用当前块的信息 m 、 b 、 l 、 p 和 q 的全部或一些。

[0200] 例如,视频编码装置可以根据预测模式 m 是否是方向的来应用不同的精细化模型。视频编码装置可以基于预测模式 m 是否是基于矩阵的预测(例如,MIP)来应用不同的精细化模型。视频编码装置可以基于预测模式 m 是否包括在WAIP中来应用不同的精细化模型。对于方向的预测模式 m ,视频编码装置可以对 m 进行聚类,并且可以基于方向的预测模式 m 包括在竖直模式、水平模式和对角线模式中的哪个聚类来应用不同的精细化模型。此外,视频编码装置可以基于预测模式 m 是竖直预测还是水平预测来应用不同的精细化模型。

[0201] 如果 b 是当前块的水平长度或竖直长度,则视频编码装置可以将 b 与预设阈值进行比较,并且可以根据结果来应用不同的精细化模型。如果 b 是当前块的纵横比,则视频编码装置可以将 b 与预设阈值进行比较,并且根据结果来应用不同的精细化模型。

[0202] 视频编码装置可以根据值 l (即,当前块是亮度块还是色度块)来应用不同的精细化模型。

[0203] 视频编码装置可以根据 p 的类型(即,当前图像是图像I、B还是P)来应用不同的精细化模型。

[0204] 视频编码装置可以基于值 q (即,当前块的QP值或其范围)来应用不同的精细化模型。

[0205] 另一方面,视频编码装置可以将指示是否使用先前训练的精细化模型的参数的标志用信号通知视频解码装置。这里,视频解码装置可以通过组合预定的精细化模型的参数和标志来对预测块执行滤波。

[0206] 如上所述,视频编码装置可以不使用单个精细化模型,而是可以使用各种精细化模型的一个。因此,视频编码装置可以传输指示多个精细化模型的一个的索引,并且视频解码装置可以组合接收到的索引和指示是否使用精细化模型的参数的标志,并且可以对预测块执行滤波。

[0207] 作为另一个示例,视频编码装置可以将针对每个序列单独训练的精细化模型的参数的全部或一些发送到视频解码装置。这里,一些参数可以是构成精细化模型的一些的参数,或者可以代表滤波器中的诸如偏移的参数。这里,代替利用预定的参数,视频解码装置可以用通过比特流接收到的参数来配置精细化模型,然后对预测块执行滤波。

[0208] 视频编码装置将当前块减去精细化的预测块以生成残差块(S1606)。

[0209] 图17是示出根据本发明实施方案的利用帧内预测的信号的精细化的视频解码方法的流程图。

[0210] 视频解码装置从比特流解码当前块的信息和当前块的残差值(S1700)。当前块的信息包括:当前块的帧内预测模式 m 、关于当前块的高度、宽度或纵横比的信息 b 、当前块的颜色 l 、包括当前块的当前图像的类型 p 以及当前块的量化参数 q 。

[0211] 视频解码装置通过利用帧内预测模式从当前块周围的重构的参考样本生成当前块的预测块(S1702)。

[0212] 视频解码装置基于当前块的信息的全部或部分来自适应地选择基于深度学习的精细化模型,并且通过将包括预测块和重构的参考样本的输入块输入到选择的精细化模型中来生成精细化的预测块(S1704)。

[0213] 由训练单元根据块的信息来预先自适应地训练精细化模型,从而进行学习。

[0214] 另一方面,视频解码装置根据当前块的信息 m 、 b 、 l 、 p 和 q 自适应地选择精细化模型

的实施方案与视频编码装置的实施方案相同,因此省略了进一步的详细描述。

[0215] 视频解码装置将残差值添加到精细化的预测块,以生成当前块的重构的块(S1706)。

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

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

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

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

[0220] (附图标记)

[0221] 122:帧内预测器

[0222] 510:熵解码器

[0223] 542:帧内预测器

[0224] 610:预测的信号精细化单元

[0225] 710:预测的信号精细化单元

[0226] 相关申请的交叉引用

[0227] 本申请要求于2021年7月13日提交的韩国专利申请No.10-2021-0091743以及于2022年6月8日提交的韩国专利申请No.10-2022-0069466的优先权和权益,其每个的全部内容通过引用结合于本文中。

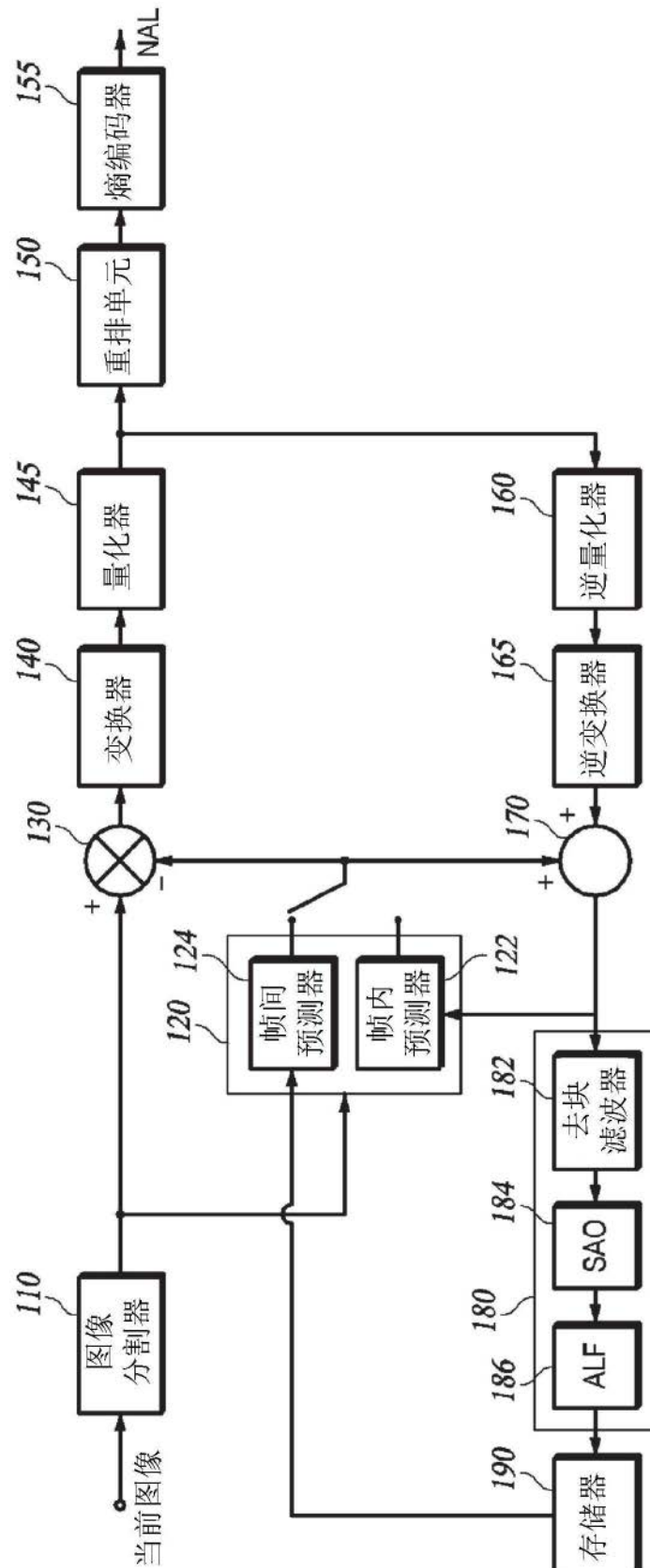


图1

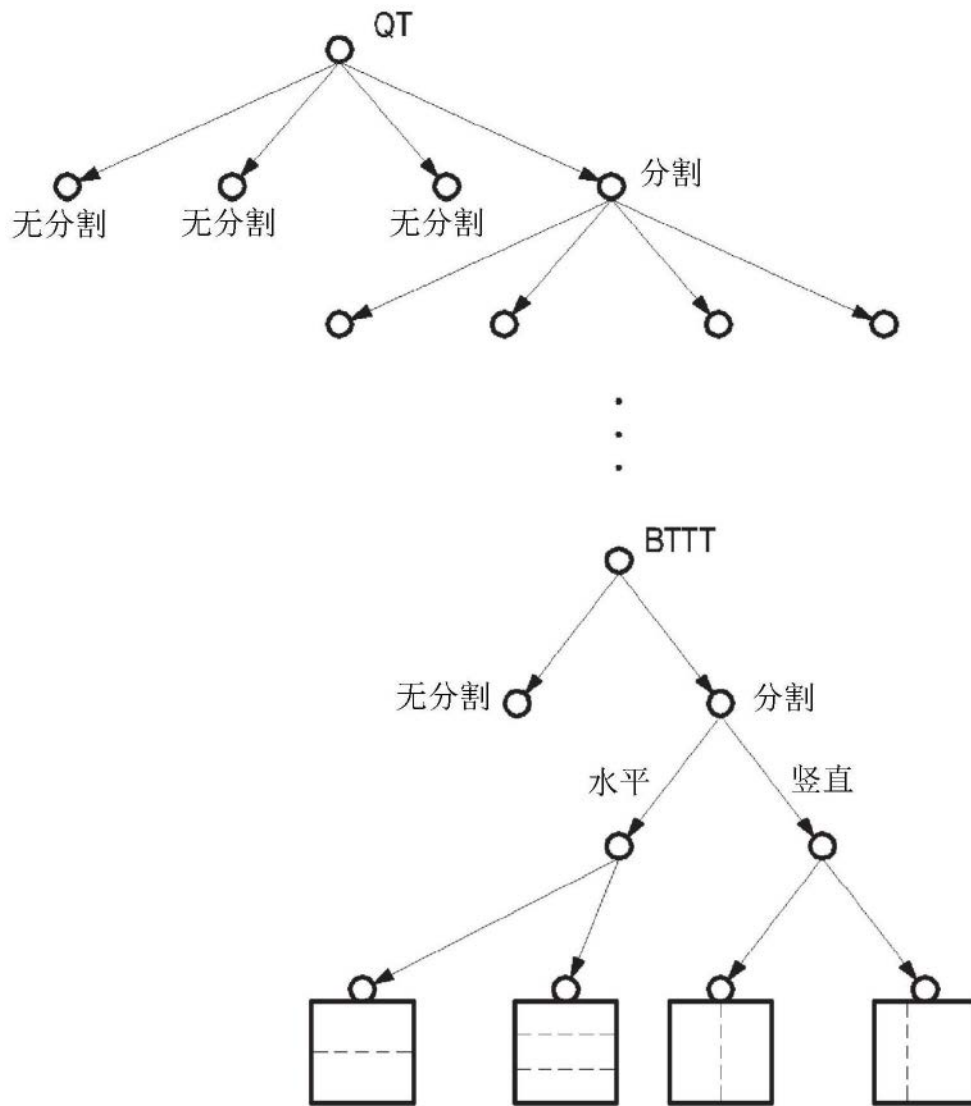


图2

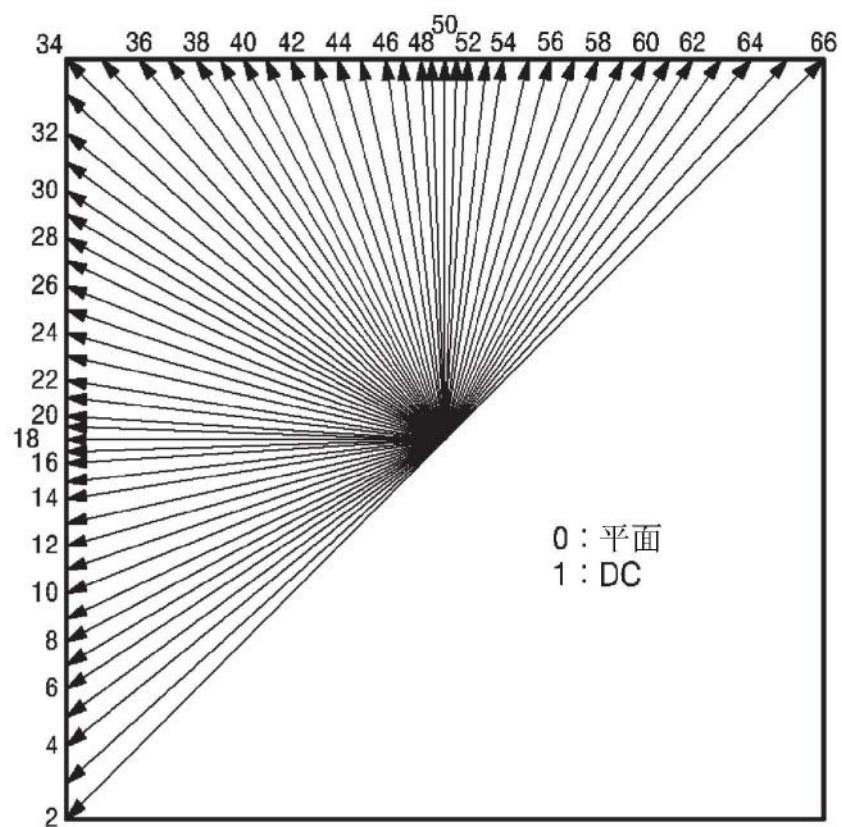


图3a

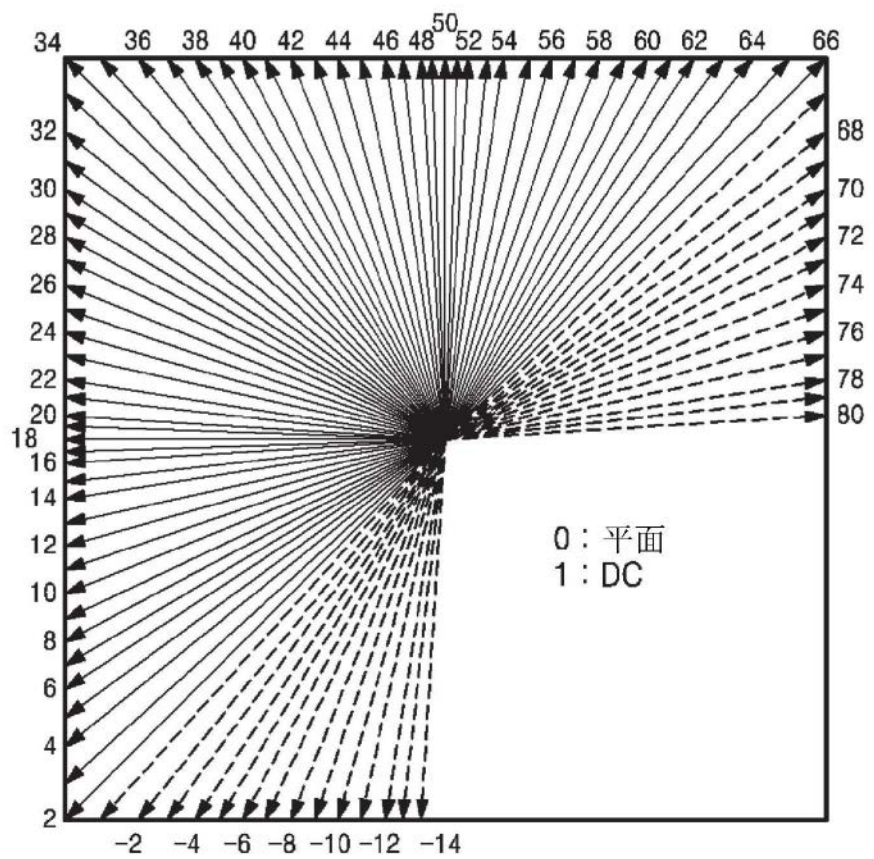


图3b

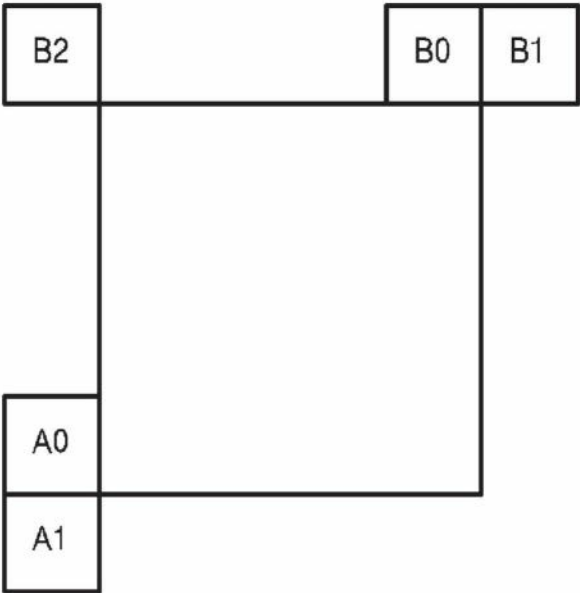


图4

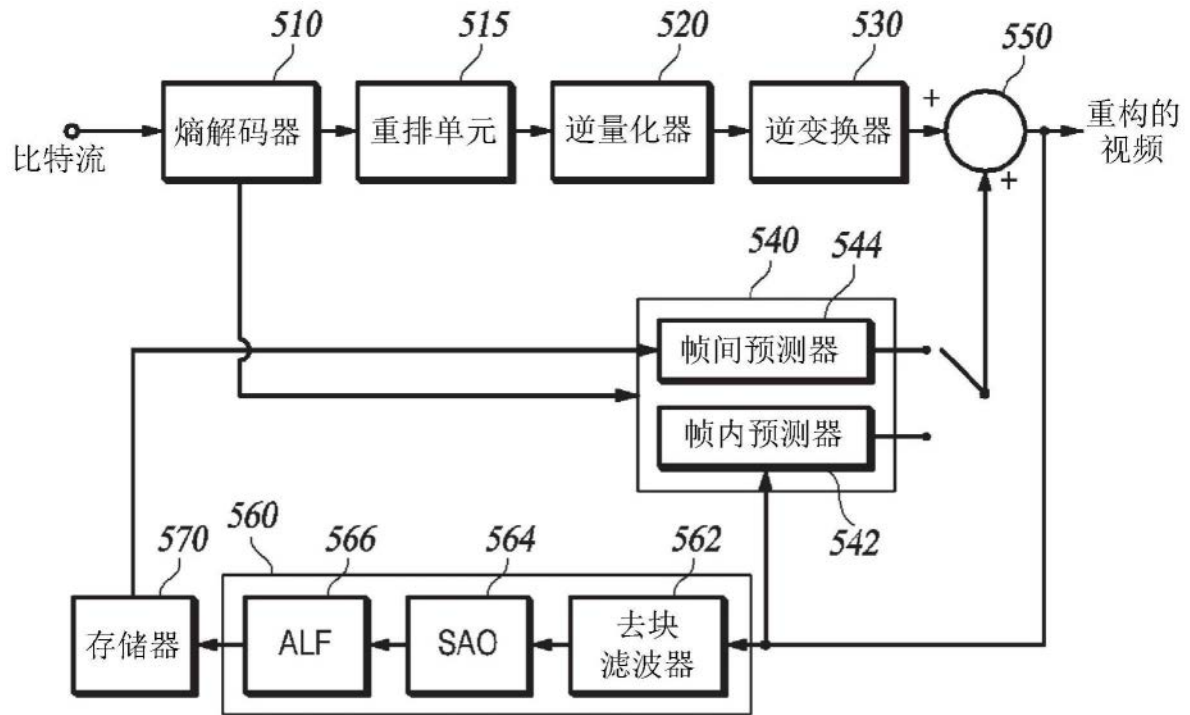


图5

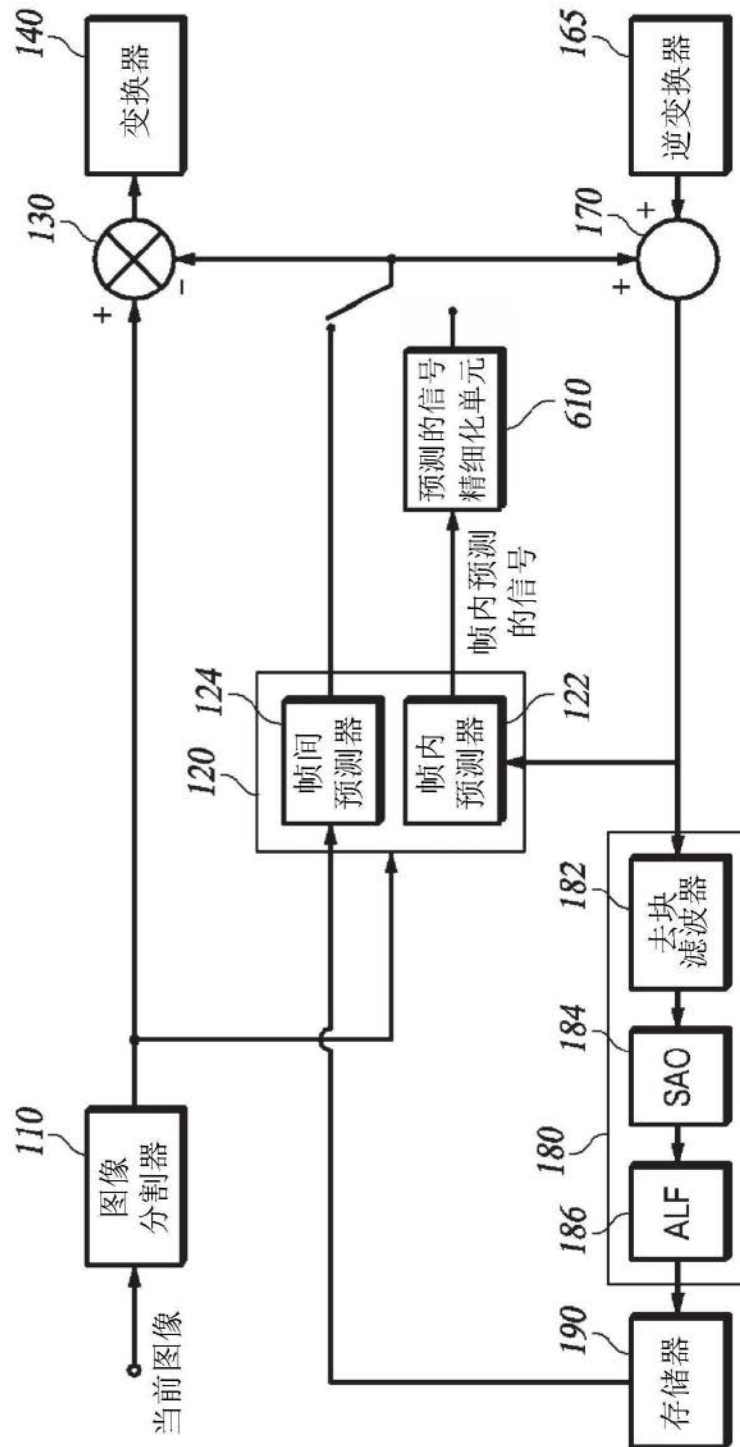


图6

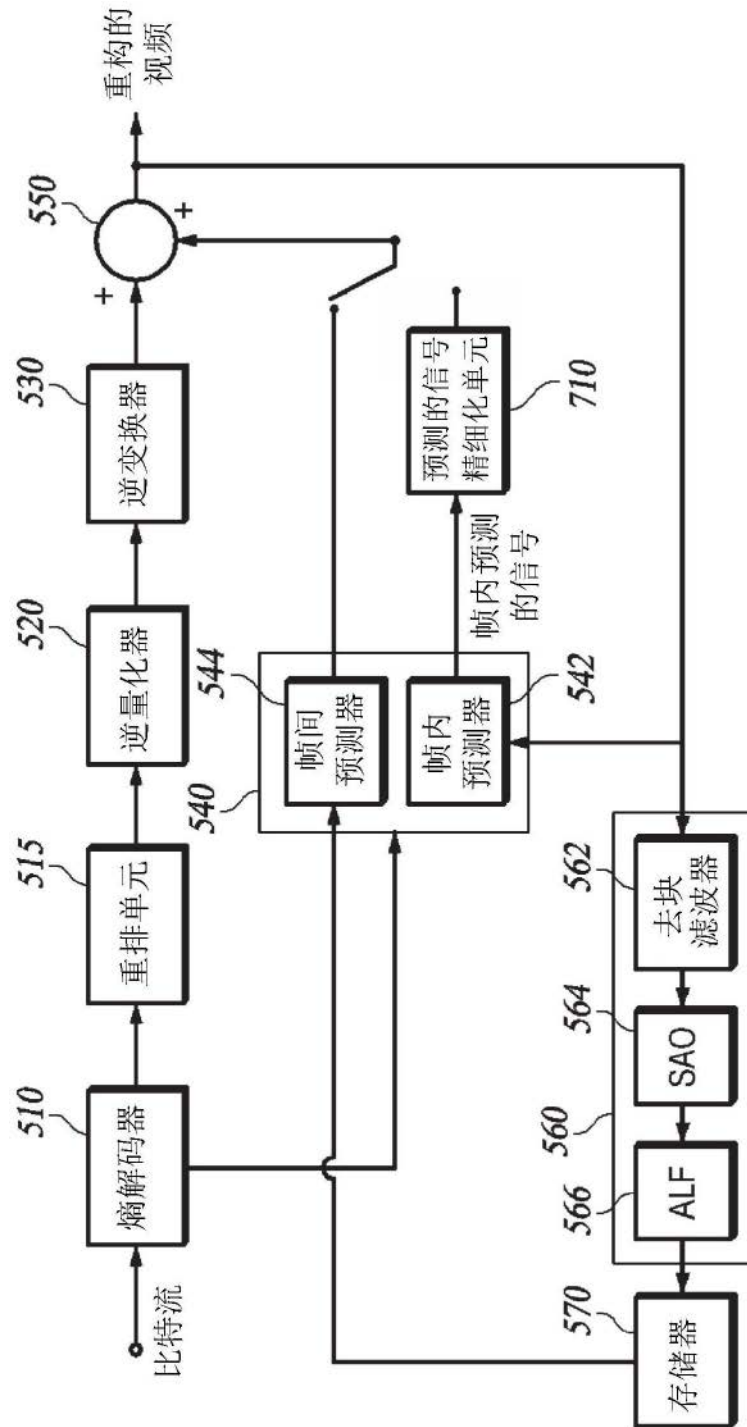
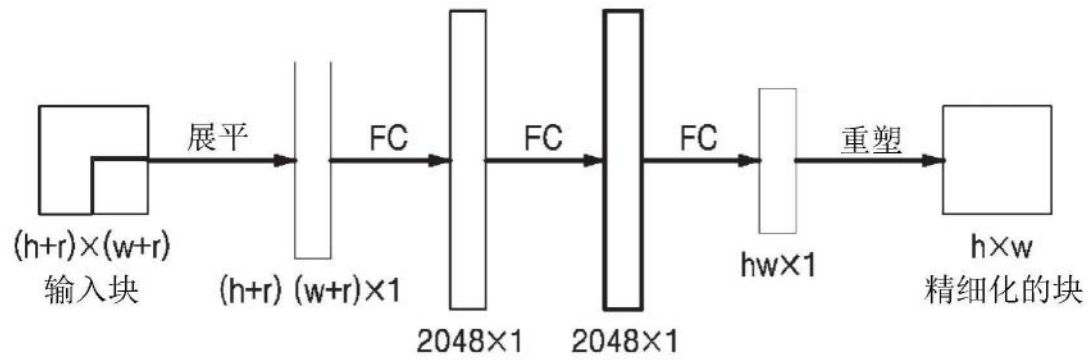


图7



h : 预测块的高度

w : 预测块的宽度

r : 参考样本的长度

图8

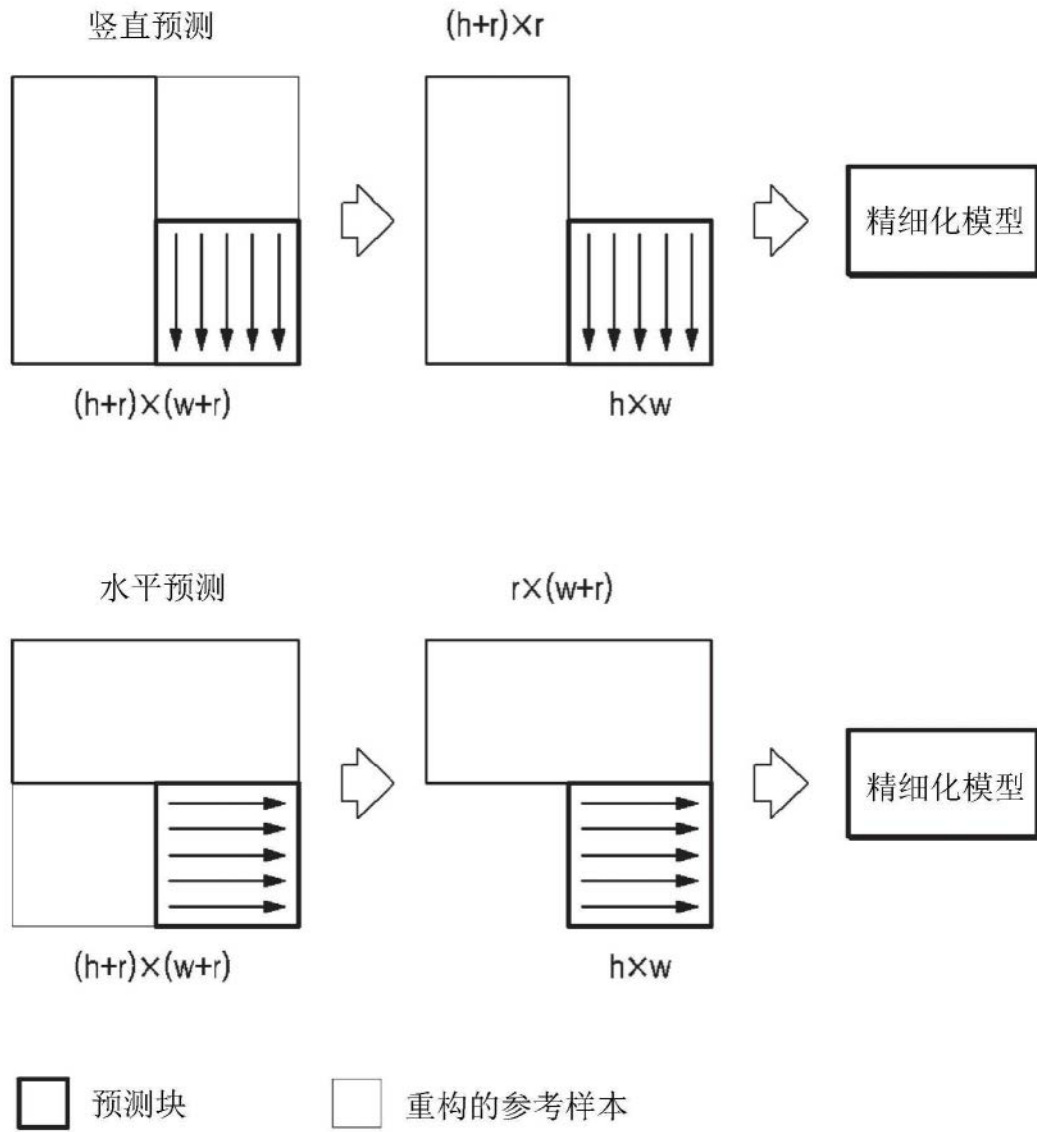


图9a

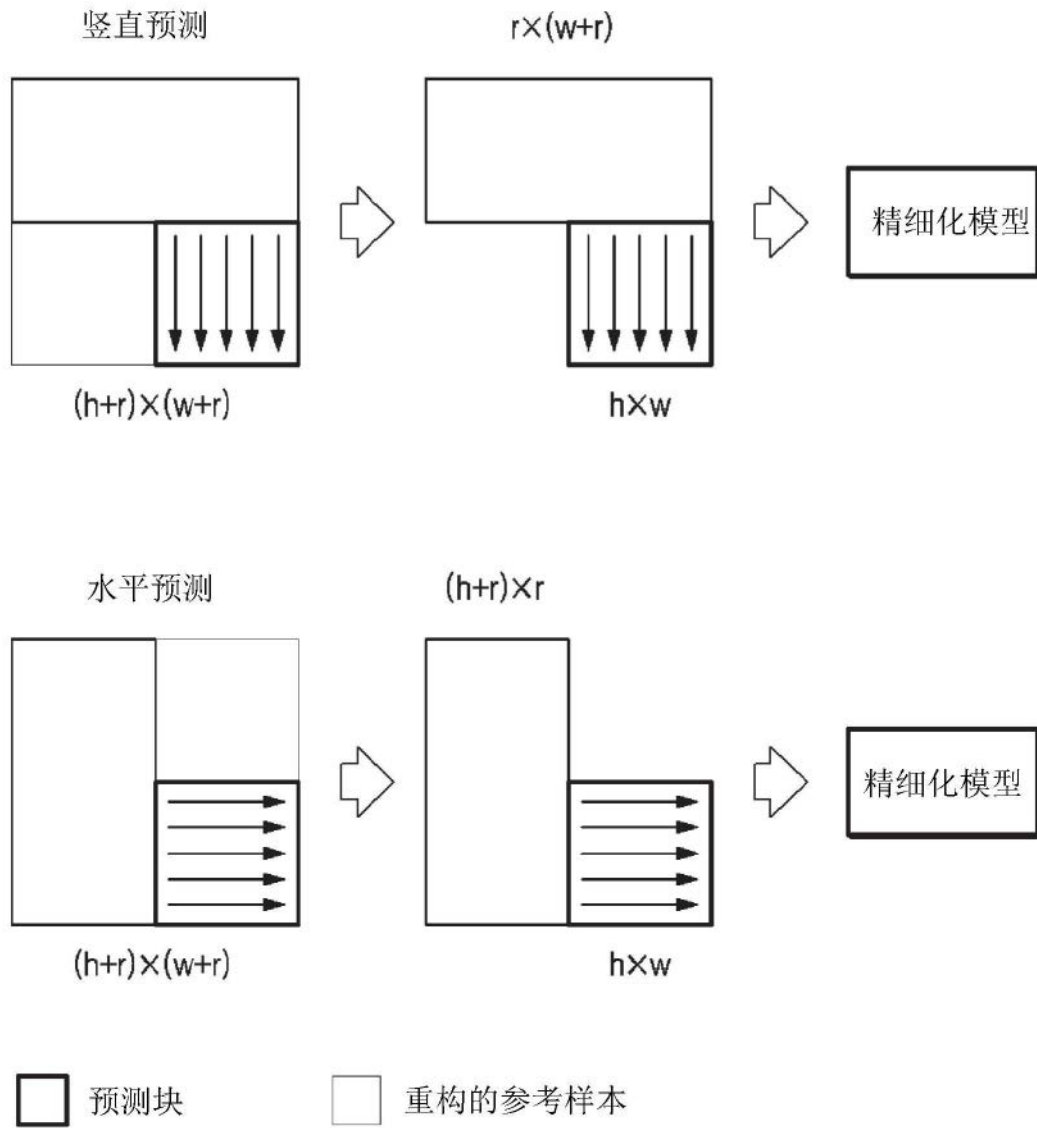


图9b

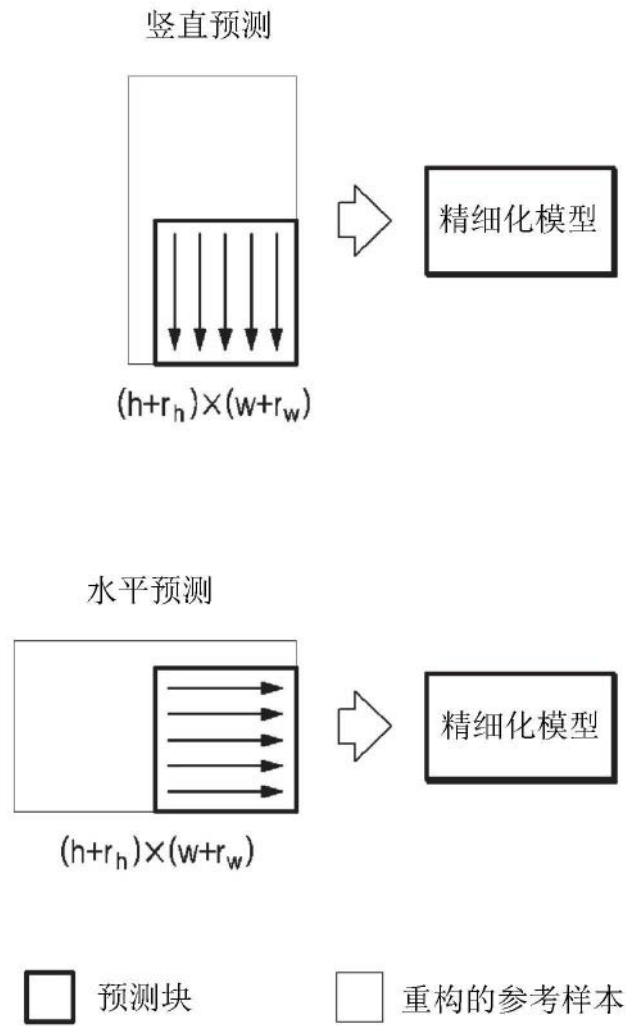


图10

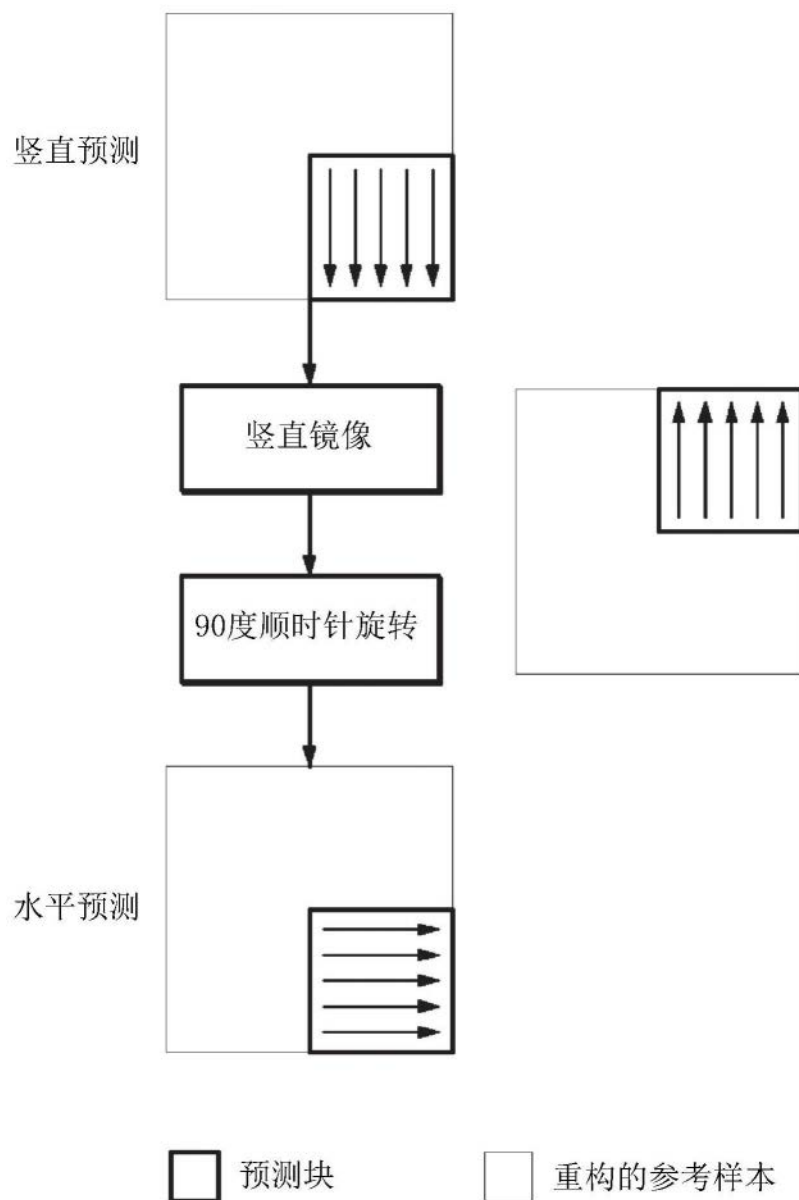


图11

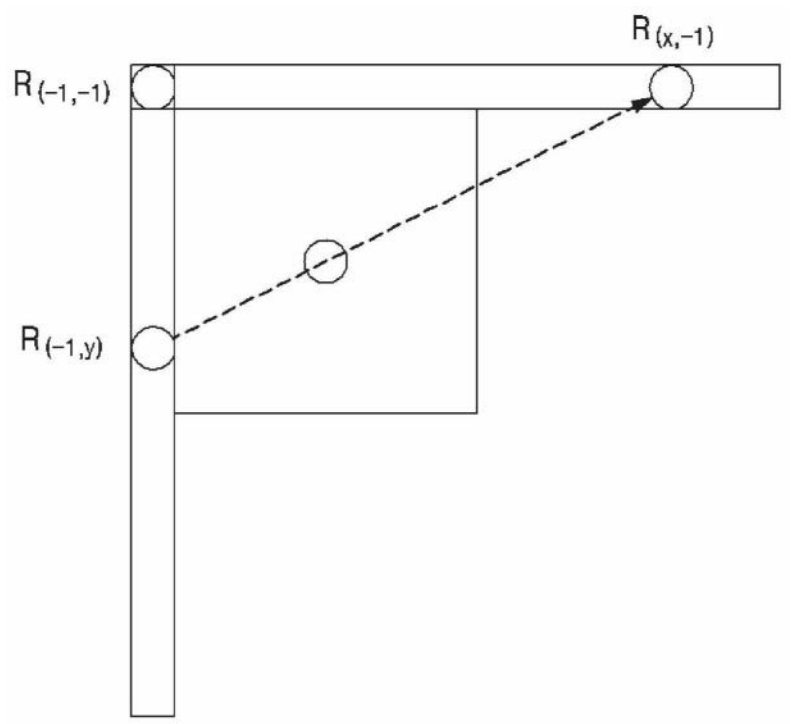


图12

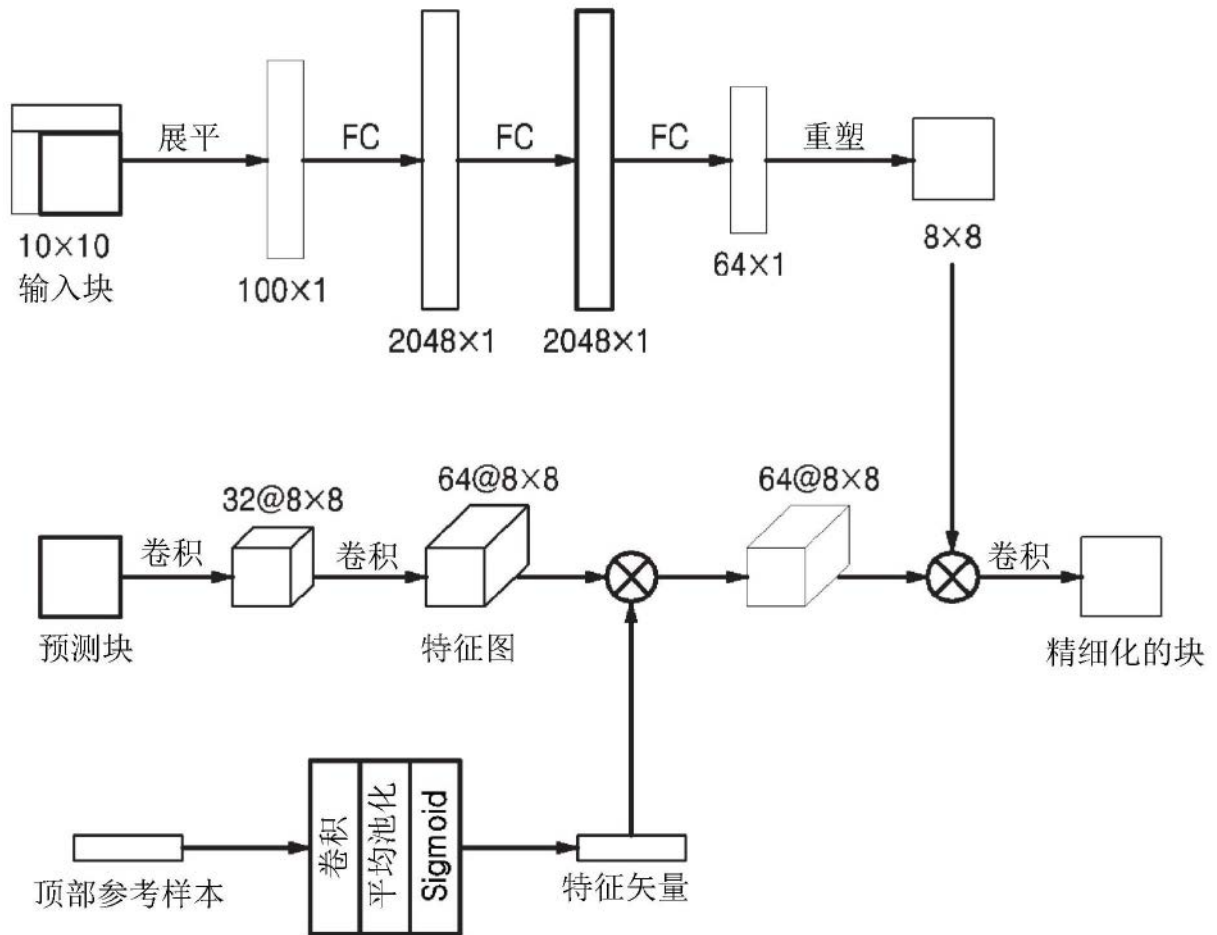


图13

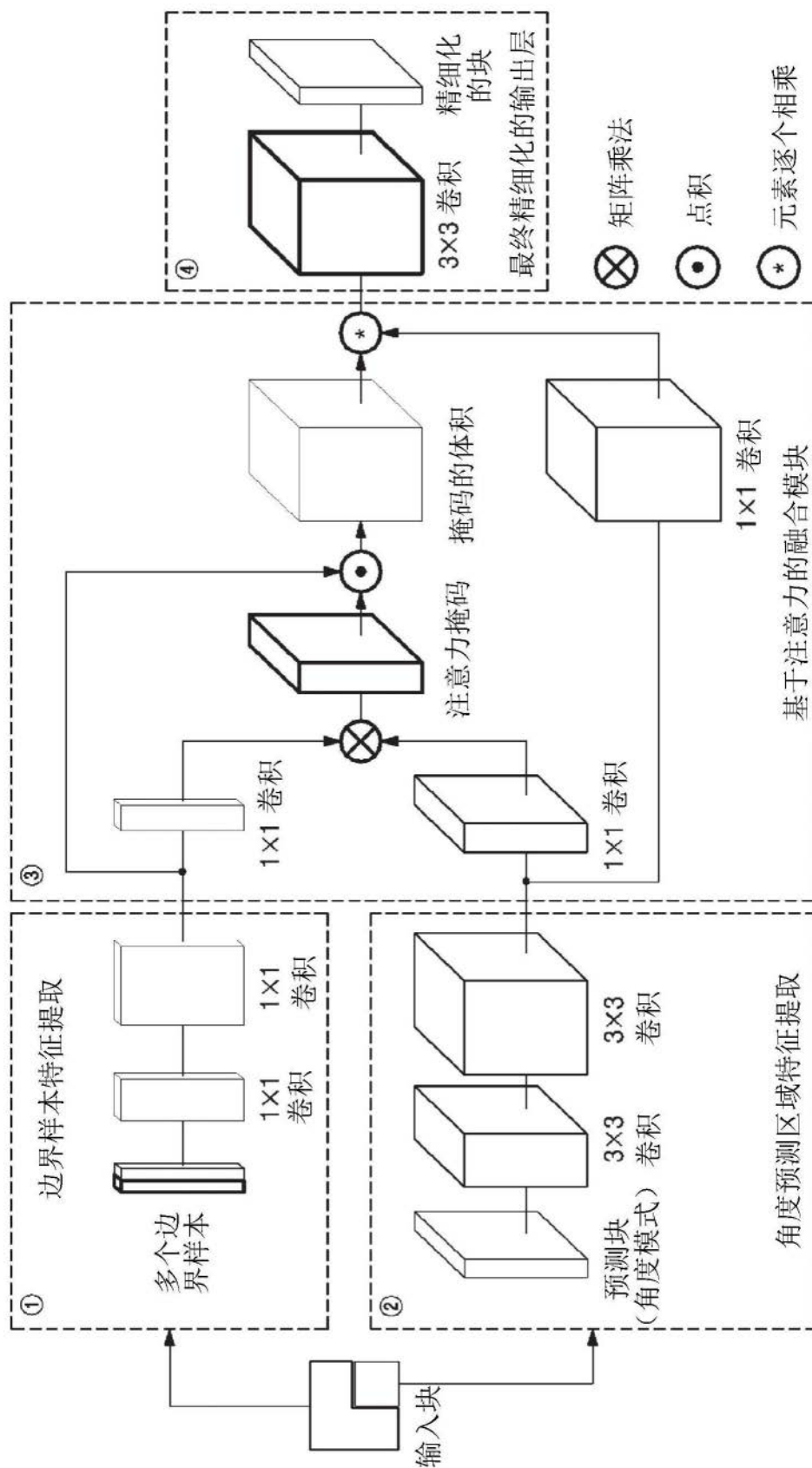


图14

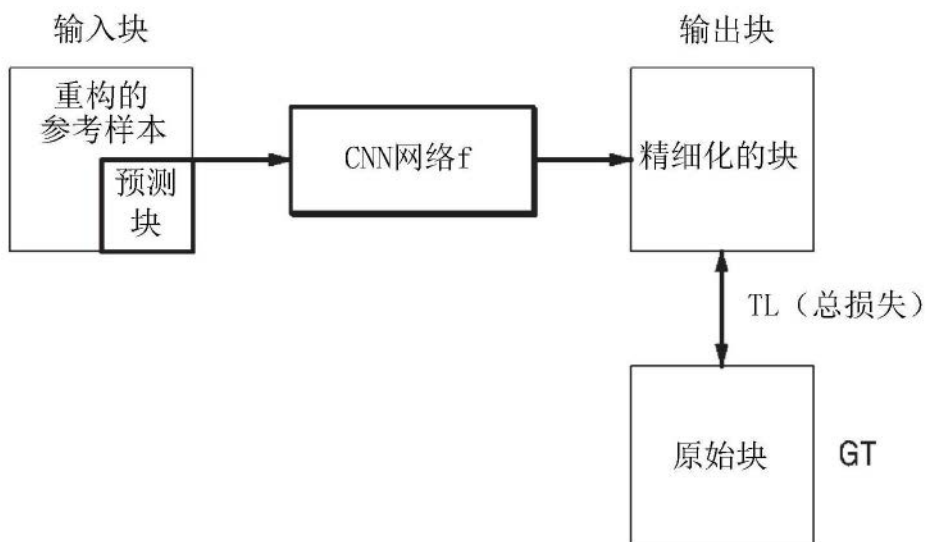


图15

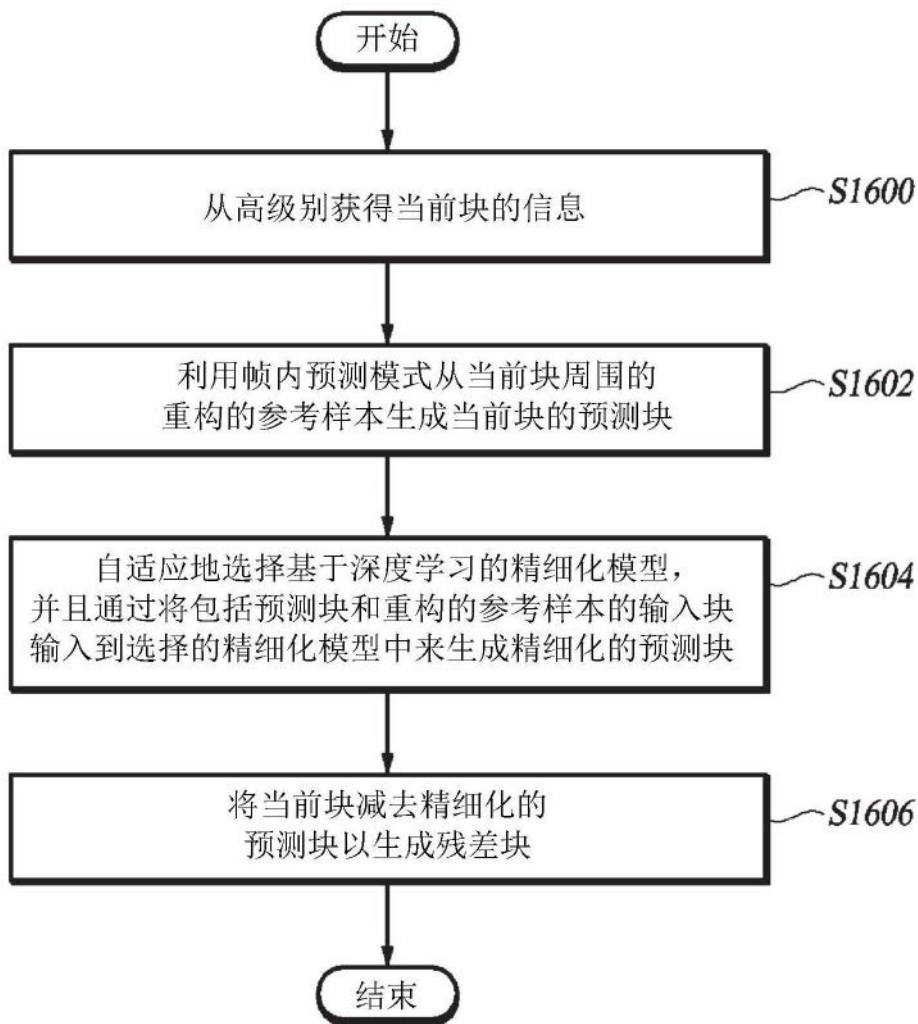


图16

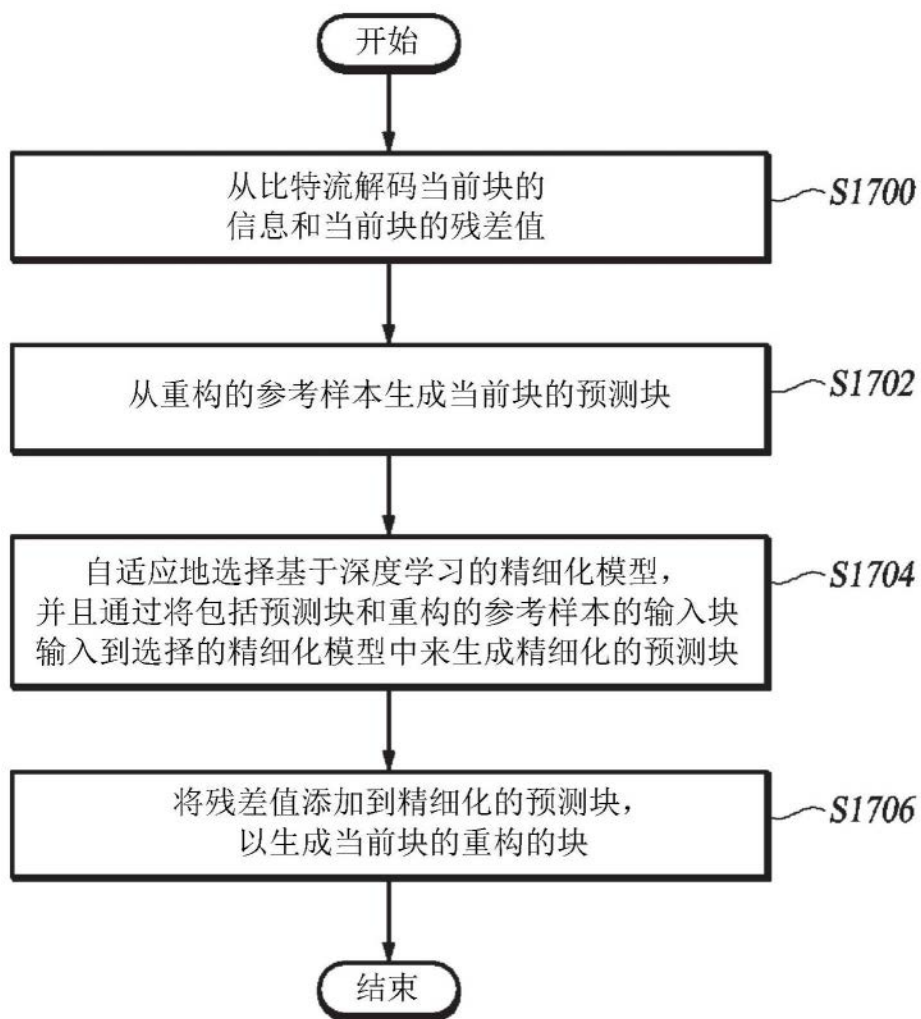


图17