



(12) 发明专利申请

(10) 申请公布号 CN 119856493 A

(43) 申请公布日 2025. 04. 18

(21) 申请号 202380064513.6

(22) 申请日 2023.09.07

(30) 优先权数据

10-2022-0114925 2022.09.13 KR

(85) PCT国际申请进入国家阶段日

2025.03.07

(86) PCT国际申请的申请数据

PCT/KR2023/013396 2023.09.07

(87) PCT国际申请的公布数据

W02024/058494 KO 2024.03.21

(71) 申请人 LG电子株式会社

地址 韩国首尔

(72) 发明人 郭东圭 林宰显 金哲瑾

(74) 专利代理机构 北京三友知识产权代理有限公司 11127

专利代理师 刘久亮

(51) Int.Cl.

H04N 19/593 (2006.01)

H04N 19/105 (2006.01)

H04N 19/132 (2006.01)

H04N 19/176 (2006.01)

H04N 19/13 (2006.01)

H04N 19/70 (2006.01)

G06T 9/00 (2006.01)

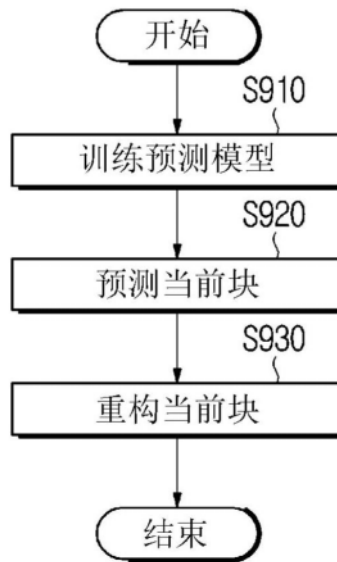
权利要求书1页 说明书24页 附图14页

(54) 发明名称

基于预测模型训练的特征编码/解码方法和设备、以及其中存储有比特流的记录介质

(57) 摘要

提供了特征编码/解码方法和设备、以及由该特征编码方法生成的计算机可读记录介质。根据本公开,由特征解码设备执行的解码方法包括以下步骤:基于重构的邻居样本对预测模型进行训练;通过使用经训练的预测模型来预测当前块;以及基于当前块的残差信息和所预测的当前块来重构当前块。



1. 一种由特征解码设备执行的特征解码方法,所述特征解码方法包括以下步骤:  
基于重构的邻居样本对预测模型进行训练;  
通过使用经训练的预测模型来预测当前块;以及  
基于所述当前块的残差信息和所预测的当前块来重构所述当前块。
2. 根据权利要求1所述的特征解码方法,其中,所述训练的步骤包括以下步骤:  
基于所述重构的邻居样本和第一预测模型来预测所述当前块;以及  
对所述第一预测模型进行训练,以使得基于所述所预测的当前块和所述重构的邻居样本之间的差异导出的值减小。
3. 根据权利要求2所述的特征解码方法,其中,所述第一预测模型被训练基于关于次数的信息导出的次数一样多次,并且  
其中,关于所述次数的所述信息是从比特流获得的。
4. 根据权利要求3所述的特征解码方法,其中,关于所述次数的所述信息指示从所述第一预测模型被训练的次数中减去预定的默认次数而得到的次数。
5. 根据权利要求1所述的特征解码方法,其中,基于预测模式信息来执行训练,所述预测模式信息是从比特流获得的并且指示所述当前块的预测模式为基于训练的预测。
6. 根据权利要求2所述的特征解码方法,其中,所述第一预测模型是至少一个或更多个候选预测模型当中的、由从比特流获得的预测模型信息指示的预测模型。
7. 根据权利要求2所述的特征解码方法,其中,所述第一预测模型是基于预测模型候选列表或预定预测模型中的一者确定的。
8. 根据权利要求7所述的特征解码方法,其中,基于从比特流获得的第一标志的值为第一值,所述第一预测模型是基于所述预测模型候选列表确定的,并且  
其中,基于所述第一标志的所述值为第二值,所述第一预测模型是基于所述预定预测模型确定的。
9. 根据权利要求7所述的特征解码方法,  
其中,所述预测模型候选列表包括用于预测至少一个邻居块的至少一个预测模型候选,并且  
其中,将所述第一预测模型确定为所述至少一个预测模型候选当中的、由从比特流获得的选择信息指示的预测模型候选。
10. 一种由特征编码设备执行的特征编码方法,所述特征编码方法包括以下步骤:  
基于重构的邻居样本对预测模型进行训练;  
通过使用经训练的所述预测模型来预测当前块;以及  
对基于所述当前块和所预测的当前块导出的残差信息进行编码。
11. 一种计算机可读记录介质,所述计算机可读记录介质用于存储由根据权利要求10所述的特征编码方法生成的比特流。
12. 一种用于发送由特征编码方法生成的比特流的方法,  
其中,所述特征编码方法包括以下步骤:  
基于重构的邻居样本对预测模型进行训练;  
通过使用经训练的预测模型来预测当前块;以及  
对基于所述当前块和所预测的当前块导出的残差信息进行编码。

## 基于预测模型训练的特征编码/解码方法和设备、以及其中存储有比特流的记录介质

### 技术领域

[0001] 本公开涉及特征编码/解码方法及装置,并且具体地,涉及基于预测模型训练的特征编码/解码方法和设备、以及存储由本公开的特征编码方法/设备生成的比特流的记录介质。

### 背景技术

[0002] 随着机器学习技术的发展,对基于图像处理的人工智能服务的需求日益增加。为了在有限的资源内有效处理人工智能服务所需的大量图像数据,对机器任务性能优化的图像压缩技术至关重要。然而,现有的图像压缩技术已经以针对人类视觉的高分辨率、高质量图像处理为目标而开发,并且存在不适合人工智能服务的问题。相应地,适合人工智能服务的面向机器的新型图像压缩技术的研发正在积极进行。

### 发明内容

[0003] 技术问题

[0004] 本公开的目的是提供一种具有改进的编码/解码效率的特征编码/解码方法和设备。

[0005] 本公开的另一目的是提供一种特征编码/解码方法和设备,其训练预测模型并使用该预测模型进行预测。

[0006] 本公开的另一目的是提供一种特征编码/解码方法和设备,其配置完全训练的预测模型的列表,并利用列表中的预测模型之一进行预测或作为用于训练的初始值。

[0007] 本公开的另一目的是提供一种发送由根据本公开的特征编码方法或设备生成的比特流的方法。

[0008] 本公开的另一目的是提供一种存储由根据本公开的特征编码方法或设备生成的比特流的记录介质。

[0009] 本公开的另一目的是提供一种存储由根据本公开的特征解码设备接收、解码并用于重构的特征的比特流的记录介质。

[0010] 本公开所解决的技术问题不限于上述技术问题,并且本文中未描述的其他技术问题将从以下描述中对于本领域技术人员来说将变得显而易见。

[0011] 技术方案

[0012] 根据本公开的一方面的由特征解码设备执行的特征解码方法可以包括:基于重构的邻居样本对预测模型进行训练,通过使用经训练的预测模型来预测当前块,以及基于当前块的残差信息和所预测的当前块重构当前块。

[0013] 根据本公开的另一方面的由特征编码设备执行的特征编码方法可以包括:基于重构的邻居样本对预测模型进行训练,使用经训练的预测模型来预测当前块,以及对基于当前块和所预测的当前块导出的残差信息进行编码。

[0014] 根据本公开的另一方面的另一种记录介质可以存储由本公开的特征编码方法或特征编码设备生成的比特流。

[0015] 根据本公开的另一方面的比特流传输方法可以向特征解码设备发送由本公开的所述特征编码方法或所述特征编码设备生成的比特流。

[0016] 上面关于本公开简要概括的特征仅仅是本公开下面详细描述的例子性方面,并且不限制本公开的范围。

[0017] 有益效果

[0018] 根据本公开,能够提供一种具有改进的编码/解码效率的特征编码/解码方法和设备。

[0019] 此外,根据本公开,由于可以通过训练实际增加可用于预测的预测模型的数量,因此能够应用适合于当前块的特征的预测模型。

[0020] 此外,根据本公开,与可变预测模型相比,可以减少比特流中的比特量。

[0021] 本领域的技术人员将会领会到,通过本公开能够实现的效果不限于上文具体描述的内容,并且从详细描述中将更清楚地理解本公开的其他优点。

## 附图说明

[0022] 图1是示意性地示出可适用本公开的实施方式的VCM系统的视图。

[0023] 图2是示意性地示出可适用本公开的实施方式的VCM管线结构的图。

[0024] 图3是示意性地示出可适用本公开的实施方式的图像/视频编码器的图。

[0025] 图4是示意性地示出可适用本公开的实施方式的图像/视频解码器的图。

[0026] 图5是示意性地图示可适用本公开的实施方式的特征/特征图编码过程的流程图。

[0027] 图6是示意性地图示可适用本公开的实施方式的特征/特征图解码过程的流程图。

[0028] 图7是示出可用于训练预测模型的重构的邻居样本和当前编码/解码目标样本的示例的视图。

[0029] 图8是示出根据本公开的实施方式的特征编码方法的流程图。

[0030] 图9是示出根据本公开的实施方式的特征解码方法的流程图。

[0031] 图10是示出根据本公开的另一实施方式的特征编码方法的流程图。

[0032] 图11是示出根据本公开的另一实施方式的特征解码方法的流程图。

[0033] 图12a和图12b是示出根据本公开的又一实施方式的特征编码方法的流程图。

[0034] 图13a和图13b是示出根据本公开的又一实施方式的特征解码方法的流程图。

[0035] 图14是示出可适用本公开的实施方式的内容流传输系统的示例的视图。

[0036] 图15是示出可适用本公开的实施方式的内容流传输系统的另一示例的视图。

## 具体实施方式

[0037] 下面,将参考附图对本公开的实施方案进行详细描述,使得本领域的技术人员能够容易地实现。然而,本公开能够以各种不同的形式来实现,并且不限于本文描述的实施方案。

[0038] 在描述本公开时,在确定相关已知功能或构造的详细描述使本公开的范围不必要地含糊不清,则将省略其详细描述。在附图中,省略了与本公开的描述无关的部分,并且相

似的附图标记被附上相似的部分。

[0039] 在本公开中,当一个组件“连接”、“联接”或“链接”到另一组件时,它不仅可以包括直接连接关系,还可以包括中间组件存在的间接连接关系。另外,当一个组件“包括”或“具有”其他组件时,除非另有说明,否则意味着可以进一步包括其他组件,而不是排除其他组件。

[0040] 在本公开中,术语第一、第二等可以仅用于将一个组件与其他组件区分开的目的,并且不限制组件的顺序或重要性,除非另有说明。因此,在本公开的范围,一个实施方式中的第一组件在另一个实施方式中可以被称为第二组件,并且类似地,一个实施方式中的第二组件在另一个实施方式中可以被称为第一组件。

[0041] 在本公开中,彼此区分的组件旨在清楚地描述每个特征,并不意味着这些组件有必要被分离。即,多个组件可以集成并实现在一个硬件或软件单元中,或者一个组件可以分布并实现在多个硬件或软件单元中。因此,即使没有另外说明,这些组件被集成或组件被分布的实施方式也包括在本公开的范围。

[0042] 在本公开中,各种实施方式中描述的组件不一定意指必要组件,并且一些组件可以是可选组件。因此,由实施方式中描述的组件的子集组成的实施方式也包括在本公开的范围。另外,除了各种实施方式中描述的组件之外,包括其他组件的实施方式也包括在本公开的范围。

[0043] 本公开涉及图像的编码和解码,并且本公开中使用的术语可以具有本公开所属的技术领域中常用的一般含义,除非在本公开中新定义。

[0044] 本公开可以被应用于多功能视频编码 (VVC) 标准和/或机器视频编码 (VCM) 标准中公开的方法。另外,本公开可以被应用于基本视频编码 (EVC) 标准、AOMedia 视频1 (AV1) 标准、第二代音频视频编码标准 (AVS2) 或下一代视频/图像编码标准 (例如,H.267或H.268等) 中公开的方法。

[0045] 本公开提供与视频/图像编译相关的各种实施方式,并且除非另有说明,否则这些实施例可以彼此组合地执行。在本公开中,“视频”指代根据时间的流逝的一系列图像的集合。“图像”可以由人工智能 (AI) 生成的信息。在通过AI执行一系列任务的过程中使用的输入信息、在信息处理过程中生成的信息以及输出信息可以被用作图像。在本公开中,“图片”通常指代表示特定时间段内的一个图像的单元,并且切片/图块是编码时构成图片的一部分的编译单元。一个图片可以由一个或多个切片/图块组成。另外,切片/图块可以包括一个或多个编译树单元 (CTU)。CTU可以被分割为一个或多个CU。图块是存在于图片中的特定图块行和特定图块列中的矩形区域,并且可以由多个CTU组成。图块列可以被定义为CTU的矩形区域,可以具有与图片相同的高度,并且可以具有由从诸如图片参数集的比特流部分用信号发送的句法元素指定的宽度。图块行可以被定义为CTU的矩形区域,可以具有与图片相同的宽度,并且可以具有由从诸如图片参数集的比特流部分用信号发送的句法元素指定的高度。图块扫描是对图片进行分割的CTU的某种连续排序方法。这里,CTU可以根据图块内的CTU光栅扫描顺序地排序,并且图片中的图块可以根据图片的图块的光栅扫描顺序来顺序地排序。切片可以包含整数个完整的图块,或者可以包含一个图片的一个图块内的连续整数个完整的CTU行。切片可以专门包含在单个NAL单元中。一个图片可以由一个或多个图块组组成。一个图块组可以包括一个或多个图块。拼块可以指示图片中的图块内的CTU行的

矩形区域。一个图块可以包括一个或多个拼块。拼块可以指代图块中的CTU行的矩形区域。一个图块可以被拆分成多个拼块,并且每个拼块可以包括属于一个图块的一个或多个CTU行。未拆分成多个拼块的图块也可以被视为拼块。

[0046] 在本公开中,“像素”或“像元”可以意指构成一个图片(或图像)的最小单元。另外,“样本”可以用作与像素相对应的术语。样本通常可以表示像素或像素的值,并且可以仅表示亮度分量的像素/像素值或仅表示色度分量的像素/像素值。

[0047] 在实施方式中,特别是当应用于VCM时,当存在由一组具有不同特性和含义的分量的集合组成的图片时,像素/像素值可以表示通过独立信息或各个分量的组合、合成和分析生成的分量的像素/像素值。例如,在RGB输入中,可以仅表示R的像素/像素值,可以仅表示G的像素/像素值,或者可以仅表示B的像素/像素值。例如,可以仅表示使用R、G和B分量合成的亮度分量的像素/像素值。例如,可以仅表示通过从分量中分析R、G和B分量而提取的图像和信息的像素/像素值。

[0048] 在本公开中,“单元”可以表示图像处理的基本单元。该单元可以包括图片的特定区域和与该区域相关的信息中的至少之一。一个单元可以包括一个亮度块和两个色度(例如,Cb和Cr)块。在一些情况下,该单元可以与诸如“样本阵列”、“块”或“区域”的术语互换使用。在一般情况下, $M \times N$ 块可以包括M列和N行的样本(或样本阵列)或变换系数的集合(或阵列)。在实施方式中,特别地,特别是当应用于VCM时,该单元可以表示包含用于执行特定任务的信息的基本单元。

[0049] 在本公开中,“当前块”可以意指“当前编译块”、“当前编译单元”、“编译目标块”、“解码目标块”或“处理目标块”之一。当执行预测时,“当前块”可以意指“当前预测块”或“预测目标块”。当执行变换(逆变换)/量化(解量化)时,“当前块”可以意指“当前变换块”或“变换目标块”。当执行滤波时,“当前块”可以意指“滤波目标块”。

[0050] 另外,在本公开中,“当前块”可以意指“当前块的亮度块”,除非明确地陈述为色度块。“当前块的色度块”可以通过包括诸如“色度块”或“当前色度块”的色度块的显式描述来表达。

[0051] 在本公开中,术语“/”和“,”应解释为指示“和/或”。例如,表述“A/B”和“A,B”可以意指“A和/或B”。此外,“A/B/C”和“A,B,C”可以意指“A、B和/或C中的至少一个”。

[0052] 在本公开中,术语“或”应解释为指示“和/或”。例如,表述“A或B”可以包括1)仅“A”,2)仅“B”,和/或3)“A和B”这两者。换句话说,在本公开中,术语“或”应当被解释为指示“附加地或另选地”。

[0053] 本公开涉及机器视频/图像编码(VCM)。

[0054] VCM指代出于机器视觉的目的,对源图像/视频的部分或从源图像/视频获得的信息进行编码/解码的压缩技术。在VCM中,编码/解码目标可以被称为特征。该特征可以指代基于任务目的、要求、周围环境等从源图像/视频中提取的信息。该特征可以具有与源图像/视频不同的信息形式,并且相应地,该特征的压缩方法和表达格式也可能与视频源的不同。

[0055] VCM可以被应用于多种应用领域。例如,在识别和跟踪物体或人的监控系统中,VCM可以被用于存储或发送物体识别信息。此外,在智能交通或智能交通系统中,VCM可以被用于将从GPS收集的车辆位置信息、从LIDAR、雷达等收集的感测信息以及各种车辆控制信息发送到其他车辆或基础设施。此外,在智慧城市领域,VCM可以被用于执行互连传感器节点

或设备的单独任务。

[0056] 本公开提供特征/特征图编译的各种实施方式。除非另有说明,本公开的实施方式可以单独实施,或者可以两个或更多个组合实施。

[0057] VCM系统的概述

[0058] 图1是示意性地示出可适用本公开的实施方式的VCM系统的图。

[0059] 参照图1,VCM系统可以包括编码设备10和解码设备20。

[0060] 编码设备10可通过压缩/编码从源图像/视频中提取的特征/特征图来生成比特流,并通过存储介质或网络将生成的比特流传输到解码设备20。编码设备10也可称为特征编码设备。在VCM系统中,可在神经网络的每个隐藏层处生成特征/特征图。生成的特征图的大小和通道数可根据神经网络的类型或隐藏层的位置而变化。在本公开中,特征图可称为特征集,特征或特征图可称为“特征信息”。

[0061] 编码设备10可以包括特征获取单元11、编码单元12和传输单元13。

[0062] 特征获取单元11可以获取针对源图像/视频的特征/特征图。取决于实施方式,特征获取单元11可以从外部设备,例如,特征提取网络,获取特征/特征图。在这种情况下,特征获取单元11执行特征接收接口功能。另选地,特征获取单元11可以通过使用源图像/视频作为输入执行神经网络(例如,CNN、DNN等)来获取特征/特征图。在这种情况下,特征获取单元11执行特征提取网络功能。

[0063] 根据实施方式,编码设备10可以进一步包括用于获取源图像/视频的源图像生成器(未示出)。源图像生成器可以利用图像传感器、相机模块等来实现,并且可以通过图像/视频捕获、合成或生成过程来获取源图像/视频。在这种情况下,生成的源图像/视频可以被发送到特征提取网络并用作用于提取特征/特征图的输入数据。

[0064] 编码单元12可以对由特征获取单元11获取的特征/特征图进行编码。编码单元12可以执行诸如预测、变换和量化之类的一系列过程以增加编码效率。编码的数据(编码的特征/特征图信息)能够以比特流的形式输出。包含编码的特征/特征图信息的比特流可以被称为VCM比特流。

[0065] 传输单元13可获得以比特流形式输出的特征/特征图信息或数据,并通过数字存储介质或网络以文件或流的形式将特征/特征图信息或数据转发到解码设备20或另一外部对象。这里,数字存储介质可以包括各种存储介质,例如USB、SD、CD、DVD、蓝光、HDD和SSD。传输单元13可以包括用于生成具有预定文件格式的媒体文件的元件或用于通过广播/通信网络传输数据的元件。传输单元13可以作为与编码单元12分开的传输装置提供,并且在这种情况下,传输装置可以包括至少一个用于获得以比特流形式输出的特征/特征图信息或数据的处理器和用于以文件或流的形式转发特征/特征图信息或数据的传输单元。

[0066] 解码设备20可以从编码设备10获得特征/特征图信息,并且基于获得后的信息重构特征/特征图。

[0067] 解码设备20可以包括接收单元21和解码单元22。

[0068] 接收单元21可以从编码设备10接收比特流,从接收到的比特流获得特征/特征图信息,并将其发送到解码单元22。

[0069] 解码单元22可以基于获取的特征/特征图信息对特征/特征图进行解码。解码单元22可以执行与编码单元12的操作相对应的一系列过程,诸如解量化、逆变换和预测,以增加

解码效率。

[0070] 取决于实施方式,解码设备20可以进一步包括任务分析/渲染单元23。

[0071] 任务分析/渲染单元23可以基于解码的特征/特征图来执行任务分析。另外,任务分析/渲染单元23可以将解码的特征/特征图渲染成适合任务执行的形式。可以基于任务分析结果和渲染的特征/特征图来执行各种机器(面向的)任务。

[0072] 如上所述,VCM系统可以根据用户和/或机器请求、任务目的和周围环境对从源图像/视频提取的特征进行编码/解码,并且基于解码的特征执行各种机器(面向的)任务。VCM系统可以通过扩展/重新设计视频/图像编译系统来实现,并且可以执行VCM标准中定义的各种编码/解码方法。

[0073] VCM管线

[0074] 图2是示意性地示出可适用本公开的实施方式的VCM管线结构的图。

[0075] 参照图2,VCM管线200可以包括用于对图像/视频进行编码/解码的第一管线210和用于对特征/特征图进行编码/解码的第二管线220。在本公开中,第一管线210可以被称为视频编解码器管线,并且第二管线220可以被称为特征编解码器管线。

[0076] 第一管线210可以包括用于对输入图像/视频进行编码的第一阶段211和用于对编码的图像/视频进行解码以生成重构的图像/视频的第二阶段212。重构的图像/视频可以被用于人类观看,即,人类视觉。

[0077] 第二管线220可以包括用于从输入图像/视频提取特征/特征图的第三阶段221、用于对提取的特征/特征图进行编码的第四阶段222、以及用于对编码的特征/特征图进行解码以生成重构的特征/特征图的第五阶段223。重构的特征/特征图可以被用于机器(视觉)任务。这里,机器(视觉)任务可以指代其中图像/视频被机器消费的任务。机器(视觉)任务可以被应用于服务场景,诸如例如监控、智能交通、智能城市、智能工业、智能内容等。取决于实施方式,重构的特征/特征图可以被用于人类视觉。

[0078] 取决于实施方式,在第四阶段222中编码的特征/特征图可以被传送到第一阶段221并被用于对图像/视频进行编码。在这种情况下,可以基于编码的特征/特征图来生成附加比特流,并且所生成的附加比特流可以被传送到第二阶段222并用于解码图像/视频。

[0079] 取决于实施方式,在第五阶段223中解码的特征/特征图可以被传送到第二阶段222并用于解码图像/视频。

[0080] 图2示出VCM管线200包括第一管线210和第二管线220的情况,但这仅仅是示例并且本公开的实施方式不限于此。例如,VCM管线200可以仅包括第二管线220,或者第二管线220可以被扩展为多个特征编解码器管线。

[0081] 另外,在第一管线210中,第一阶段211可以由图像/视频编码器执行,并且第二阶段212可以由图像/视频解码器执行。另外,在第二管线220中,第三阶段221可以由VCM编码器(或特征/特征图编码器)执行,并且第四阶段222可以由VCM解码器(或特征/特征图解码器)执行。在下文中,将详细描述编码器/解码器结构。

[0082] 编码器

[0083] 图3是示意性地示出可适用本公开的实施方式的图像/视频编码器的图。

[0084] 参照图3,图像/视频编码器300可以进一步包括图像分割器310、预测器320、残差处理器330、熵编码器340、和加法器350、滤波器360以及存储器370。预测器320可以包括帧

间预测器321和帧内预测器322。残差处理器330可以包括变换器332、量化器333、解量化器334和逆变换器335。残差处理器330可以进一步包括减法器331。加法器350可以被称为重构器或重构块生成器。取决于实施方式,图像分割器310、预测器320、残差处理器330、熵编码器340、加法器350和滤波器360可以由一个或多个硬件组件(例如,编码器芯片组或处理器)配置。另外,存储器370可以包括解码图片缓冲器(DPB)并且可以由数字存储介质配置。上述硬件组件可以进一步包括作为内部/外部组件的存储器370。

[0085] 图像分割器310可以将输入到图像/视频编码器300的输入图像(或图片、帧)分割成一个或多个处理单元。作为示例,处理单元可以被称为编译单元(CU)。编译单元可以根据二叉树三元树(QTBT)结构从编译树单元(CTU)或最大编译单元(LCU)被递归地分割。例如,一个编译单元可以基于二叉树结构、三元树结构和/或三元结构被分割为多个更深深度的编译单元。在这种情况下,例如,可以首先应用二叉树结构,并且可以稍后应用三元树结构和/或三元树结构。另选地,可以首先应用三元树结构。根据本公开的图像/视频编码过程可以基于不再被分割的最终编译单元来执行。在这种情况下,可以基于根据图像特性的编译效率来使用最大编译单元作为最终编译单元,或者如有必要,可以将编译单元递归地分割为更深深度的编译单元以使用作为最终编译单元的具有最佳大小的编译单元。这里,编译过程可以包括诸如预测、变换和重构的过程,这将在后面描述。作为另一示例,处理单元可以进一步包括预测单元(PU)或变换单元(TU)。在这种情况下,预测单元和变换单元均可以从上述最终编译单元划分或分割。预测单元可以是样本预测的单元,并且变换单元可以是用于导出变换系数的单元和/或用于从变换系数导出残差信号的单元。

[0086] 在一些情况下,该单元可以与诸如块或区域的术语互换使用。在一般情况下, $M \times N$ 块可以表示由M列和N行组成的样本或变换系数的集合。样本通常可以表示像素或像素值,并且可以仅表示亮度分量的像素/像素值,或者仅表示色度分量的像素/像素值。样本可以被用作与像素或像元相对应的术语。

[0087] 图像/视频编码器300可以通过从输入图像信号(原始块、原始样本阵列)减去从帧间预测器321或帧内预测器322输出的预测信号(预测块、预测样本阵列)来生成残差信号(残差块、残差样本阵列)并将生成的残差信号发送到变换器332。在这种情况下,如所示的,从图像/视频编码器300内的输入图像信号(原始块、预测样本阵列)中减去预测信号(预测块、预测样本阵列)的单元可以被称为减法器331。预测器可以对处理目标块(在下文中称为当前块)执行预测,并生成包括用于该当前块的预测样本的预测块。预测器可以确定在当前块或CU单元中是否应用帧内预测或者帧间预测。预测器可以生成诸如预测模式信息的与预测相关的各种信息,并且将其传送到熵编码器340。关于预测的信息可以在熵编码器340中被编码并且以比特流的形式输出。

[0088] 帧内预测器322可以通过参考当前图片中的样本来预测当前块。此时,取决于预测模式,所参考的样本可以位于当前块的邻居中或者可以位于远离当前块的位置。在帧内预测中,预测模式可以包括多个非定向模式和多个定向模式。非定向模式可以包括例如DC模式和平面模式。根据预测方向的详细程度,定向模式可以包括例如33个定向预测模式或65个定向预测模式。然而,这仅是示例,可以取决于设置使用更多或更少的定向预测模式。帧内预测器322可以通过使用应用于邻近块的预测模式来确定应用于当前块的预测模式。

[0089] 帧间预测器321可以基于由参照图片上的运动矢量指定的参考块(参考样本阵列)

来导出用于当前块的预测块。在这种情况下,为了减少在帧间预测模式中发送的运动信息量,可以基于邻近块和当前块之间的运动信息的相关性以块、子块或样本为单位来预测运动信息。运动信息可以包括运动矢量和参照图片索引。运动信息可以进一步包括帧间预测方向(L0预测、L1预测、Bi预测等)信息。在帧间预测的情况下,邻近块可以包括当前图片中存在的空间邻近块和参照图片中存在的时间邻近块。包括参考块的参照图片和包括时间邻近块的参照图片可以相同或不同。时间邻近块可以被称为并置参考块、并置CU(colCU)等,并且包括时间邻近块的参照图片可以被称为并置图片(colPic)。例如,帧间预测器321可以基于邻近块构造运动信息候选列表,并生成指示哪个候选被用于导出当前块的运动矢量和/或参照图片索引的信息。可以基于各种预测模式来执行帧间预测,并且例如,在跳过模式和合并模式的情况下,帧间预测器321可以使用邻近块的运动信息作为当前块的运动信息。在跳过模式的情况下,与合并模式不同,可以不发送残差信号。在运动矢量预测(MVP)模式的情况下,邻近块的运动矢量可以被用作运动矢量预测器,并且可以用信号发送运动矢量差以指示当前块的运动矢量。

[0090] 预测器320可以基于各种预测方法来生成预测信号。例如,对于一个块的预测,预测器不仅可以应用帧内预测或帧间预测,而且可以同时应用帧内预测和帧间预测这两者。这可以称为组合的帧间和帧内预测(CIIP)。另外,预测器可以基于用于块的预测的帧内块复制(IBC)预测模式或调色板模式。IBC预测模式或调色板模式可以被用于游戏等的内容图像/视频编译,例如,屏幕内容编译(SCC)。IBC基本上在当前图片内执行预测,但是可以与帧间预测类似地执行,因为在当前图片内导出参考块。也就是说,IBC可以使用本公开中描述的帧间预测技术中的至少一种。调色板模式可以被视为帧内编译或帧内预测的示例。当应用调色板模式时,可以基于关于调色板表和调色板索引的信息来用信号发送图片内的样本值。

[0091] 由预测器320生成的预测信号可以被用于生成重构信号或生成残差信号。变换器332可以通过将变换技术应用于残差信号来生成变换系数。例如,变换技术可以包括离散余弦变换(DCT)、离散正弦变换(DST)、karhunen-loève变换(KLT)、基于图形的变换(GBT)或条件非线性变换(CNT)中的至少一种。这里,GBT指代当像素之间的关系信息由图形来表示时从图形获得的变换。CNT指代基于使用所有先前重构的像素生成的预测信号获得的变换。另外,变换处理可以被应用于具有相同大小的正方形像素块,或者可以应用于具有可变大小的非正方形块。

[0092] 量化器130可以对变换系数进行量化并将它们发送到熵编码器190。熵编码器190可以对量化的信号(关于量化的变换系数的信息)进行编码并输出比特流。关于量化的变换系数的信息可以被称为残差信息。量化器130可以基于系数扫描顺序将以块形式的量化变换系数重新排序为一维矢量形式,并且基于以一维矢量形式的量化的变换系数生成关于量化的变换系数的信息。熵编码器340可以执行各种编码方法,诸如例如指数哥伦布、上下文自适应可变长度编码(CAVLC)、上下文自适应二进制算术编码(CABAC)等。熵编码器340可以一起或单独地对除了量化的变换系数(例如,句法元素的值等)之外的视频/图像重构所必要的信息进行编码。编码的信息(例如,编码的视频/图像信息)能够以比特流的形式以网络抽象层(NAL)为单位来发送或存储。视频/图像信息可以进一步包括关于诸如自适应参数集(APS)、图片参数集(PPS)、序列参数集(SPS)或视频参数集(VPS)的各种参数集的信息。另

外,视频/图像信息可以进一步包括一般约束信息。另外,视频/图像信息可以进一步包括生成和使用编码的信息的方法、目的等。在本公开中,从图像/视频编码器传送/用信号发送到图像/视频解码器的信息和/或句法元素可以包括在图像/视频信息中。图像/视频信息可以通过上述编码过程被编码并且被包括在比特流中。比特流可以通过网络发送或者可以存储在数字存储介质中。网络可以包括广播网络和/或通信网络,并且数字存储介质可以包括诸如USB、SD、CD、DVD、蓝光、HDD、SSD等的各种存储介质。发送从熵编码器340输出的信号的发送器(未示出)和/或存储该信号的存储单元(未示出)可以被配置为图像/视频编码器300的内部/外部元件,或者发送器可以被包括在熵编码器340中。

[0093] 从量化器130输出的量化的变换系数可以被用于生成预测信号。例如,可以通过解量化器334和逆变换器335对量化的变换系数应用解量化和逆变换来重构残差信号(残差块或残差样本)。加法器350将重构的残差信号与从帧间预测器321或帧内预测器322输出的预测信号相加以生成重构信号(重构图片、重构块、重构样本阵列)。在针对处理目标块不存在残差的情况下,诸如应用跳跃模式的情况,预测块可以被用作重构块。加法器350可以被称为重构器或重构块生成器。所生成的重构信号可以被用于当前图片中的下一个处理目标块的帧内预测,并且可以被用于通过如下所述的滤波来进行下一个图片的帧间预测。

[0094] 另外,具有色度缩放的亮度映射(LMCS)在图片编码和/或重构过程中适用。

[0095] 滤波器360可以通过对重构信号应用滤波来改进主观/客观图像质量。例如,滤波器360可以通过对重构图片应用各种滤波方法来生成修改的重构图片,并将修改的重构图片存储在存储器370中,具体地,存储在存储器370的DPB中。各种滤波方法可包括:例如,去块滤波、样本自适应偏移、自适应环路滤波器、双边滤波器等。滤波器360可以生成与滤波有关的各种信息,并将所生成的信息发送到熵编码器190。与滤波有关的信息可以由熵编码器340编码并以比特流的形式输出。

[0096] 发送到存储器370的修改的重构图片可以被用作帧间预测器321中的参照图片。通过此,可以避免编码器和解码器之间的预测失配并且可以改进编译效率。

[0097] 存储器370的DPB可以存储修改的重构图片以用作帧间预测器321中的参照图片。存储器370可以存储从其中导出(或编码)当前图片中的运动信息的块的运动信息和/或已经重构的图片中的块的运动信息。存储的运动信息可以被传送到帧间预测器321以用作空间邻近块的运动信息或时间邻近块的运动信息。存储器370可以存储当前图片中的重构块的重构样本,并且可以将存储的重构样本传送到帧内预测器322。

[0098] 另外,VCM编码器(或特征/特征图编码器)基本上执行一系列过程,诸如预测、变换和量化来对特征/特征图进行编码,并且因此可以基本上具有与参照图3描述的图像/视频编码器300相同/相似的结构。然而,VCM编码器与图像/视频编码器300的不同之处在于特征/特征图是编码目标,并且因此在每个单元(或组件)的名称(例如,图像分割器310等)及其具体操作内容中可能与图像/视频编码器300不同。稍后将详细描述VCM编码器的具体操作。

[0099] 解码器

[0100] 图4是示意性地示出可适用本公开的实施方式的图像/视频解码器的图。

[0101] 参照图4,图像/视频解码器400可以包括熵解码器410、残差处理器420、预测器430、加法器440、滤波器450以及存储器460。预测器430可以包括帧间预测器431和帧内预测

器432。残差处理器420可以包括解量化器421和逆变换器422。取决于实施方式,熵解码器410、残差处理器420、预测器430、加法器440和滤波器450可以由一个硬件组件(例如,解码器)配置。芯片组或处理器)配置。另外,存储器460可以包括解码图片缓冲器(DPB)并且可以由数字存储介质配置。硬件组件可以进一步包括存储器460作为内部/外部组件。

[0102] 当输入包含视频/图像信息的比特流时,图像/视频解码器400可以与在图3的图像/视频编码器300中处理图像/视频信息的过程相对应地重构图像/视频。例如,图像/视频解码器400可以基于从比特流获得的块分割相关信息来导出单元/块。图像/视频解码器400可以使用在图像/视频编码器中应用的处理单元来执行解码。相应地,解码的处理单元例如可以是编译单元,并且编译单元可以根据二叉树结构、二叉树结构和/或三叉树结构从编译树单元或最大编译单元进行分割。可以从编译单元导出一个或多个变换单元。另外,通过图像/视频解码器400解码和输出的重构的图像信号可以通过回放设备来播放。

[0103] 图像/视频解码器400可以接收以比特流的形式从图3的编码器输出的信号,并且通过熵解码器410解码接收到的信号。例如,熵解码器410可以解析比特流以导出对于图像重构(或图片重构)所必要的信息(例如,图像/视频信息)。该图像/视频信息可以进一步包括关于各种参数集的信息,诸如自适应参数集(APS)、图片参数集(PPS)、序列参数集(SPS)或视频参数集(VPS)。另外,图像/视频信息可以进一步包括一般约束信息。另外,图像/视频信息可以包括生成和使用解码信息的方法、目的等。图像/视频解码器400可以进一步基于关于参数集的信息和/或一般约束信息来对图片进行解码。可以通过解码过程从比特流中解码和获得用信号发送的/接收到的信息和/或句法元素。例如,熵解码器410可以基于诸如指数哥伦布编译、CAVLC或CABAC的编译方法对比特流中的信息进行解码,并且输出对于图像重构所必要的句法元素的值和与残差相关的变换系数的量化值。更具体地,在CABAC熵解码方法中,可以在比特流中接收与每个句法元素相对应的bin,可以使用解码目标句法元素信息以及邻近块和解码目标块的解码信息或者关于先前步骤中解码的符号/bin的信息来确定上下文模型,可以根据所确定的上下文模型来预测bin的出现概率,并可以对bin进行算术解码以生成与每个句法元素的值相对应的符号。此时,CABAC熵解码方法可以在确定上下文模型之后将关于解码的符号/bin的信息用于下一个符号/bin的上下文模型来更新上下文模型。在熵解码器410中解码的信息当中的关于预测的信息被提供给预测器(帧间预测器432和帧内预测器431),并且通过在熵解码器410中执行熵解码获得的残差值,即,量化的变换系数和相关的参数信息可以被输入到残差处理器420。残差处理器420可以导出残差信号(残差块、残差样本、残差样本阵列)。另外,由熵解码器410解码的信息当中的关于滤波的信息可以被提供给滤波器450。同时,接收从图像/视频编码器输出的信号的接收器(未示出)可以进一步被配置为图像/视频解码器400的内部/外部元件,或者接收器可以是熵解码器410的组件。同时,根据本公开的图像/视频解码器可以被称为图像/视频解码装置,并且图像/视频解码器可以被划分为信息解码器(图像/视频信息解码器)和样本解码器(图像/视频样本解码器)。在这种情况下,信息解码器可以包括熵解码器410,并且样本解码器可以包括解量化器321、逆变换器322、加法器440、滤波器450和存储器460、帧间预测器432或帧内预测器431中的至少一个。

[0104] 解量化器421可以对量化的变换系数进行解量化并输出变换系数。解量化器421可以将量化的变换系数重新排列成二维块形式。在这种情况下,可以基于在图像/视频编码器

中执行的系数扫描顺序来执行重新排列。解量化器321可以使用量化参数(例如,量化步长信息)对量化的变换系数执行解量化并获得变换系数。

[0105] 逆变换器422对变换系数进行逆变换以获得残差信号(残差块、残差样本阵列)。

[0106] 预测器430可以对当前块执行预测并且生成包括用于当前块的预测样本的预测块。预测器可以基于关于从熵解码器410输出的预测的信息来确定是否将帧内预测或者帧间预测应用于当前块,并且可以确定特定的帧内/帧间预测模式。

[0107] 预测器420可以基于各种预测方法生成预测信号。例如,预测器不仅可以对一个块的预测应用帧内预测或帧间预测,而且可以同时应用帧内预测和帧间预测。这可以称为组合的帧间和帧内预测(CIIP)。另外,预测器可以基于用于块的预测的帧内块复制(IBC)预测模式或调色板模式。IBC预测模式或调色板模式可以例如用于诸如游戏的内容的图像/视频编译,诸如屏幕内容编译(SCC)。在IBC中,预测基本上在当前图片内执行,但是可以与帧间预测类似地执行,因为在当前图片内导出参考块。也就是说,IBC可以使用此文档中描述的帧间预测技术中的至少一种。调色板模式可以被视为帧内编译或帧内预测的示例。当应用调色板模式时,关于调色板表和调色板索引的信息可以被包括在图像/视频信息中并用信号发送。

[0108] 帧内预测器431可以通过参考当前图片中的样本来预测当前块。取决于预测模式,参考的样本可以位于当前块的邻居中,或者可以位于远离当前块的位置。在帧内预测中,预测模式可以包括多个非定向模式和多个定向模式。帧内预测器431可以使用应用于邻近块的预测模式来确定应用于当前块的预测模式。

[0109] 帧间预测器432可以基于由参照图片中的运动矢量指定的参考块(参考样本阵列)来导出用于当前块的预测块。此时,为了减少在帧间预测模式中发送的运动信息量,可以基于邻近块和当前块之间的运动信息的相关性以块、子块或样本为单位来预测运动信息。运动信息可以包括运动矢量和参照图片索引。运动信息可以进一步包括帧间预测方向(L0预测、L1预测、Bi预测等)信息。在帧间预测的情况下,邻近块可以包括当前图片中存在的空间邻近块和参照图片中存在的的时间邻近块。例如,帧间预测器432可以基于邻近块构造运动信息候选列表,并且基于接收到的候选选择信息导出当前块的运动矢量和/或参照图片索引。可以基于各种预测模式来执行帧间预测,并且关于预测的信息可以包括指示用于当前块的帧间预测的模式的信息。

[0110] 加法器440可以通过将获得的残差信号与从预测器(包括帧间预测器432和/或帧内预测器431)输出的预测信号(预测块、预测样本阵列)相加来生成重构信号(重构图片、重构块、重构样本阵列)。如果针对处理目标块不存在残差,诸如当应用跳过模式时,则预测块可以被用作重构块。

[0111] 加法器440可以被称为重构器或重构块生成器。所生成的重构信号可以被用于当前图片中的下一个处理目标块的帧内预测,可以在如稍后描述的滤波之后被输出,或者可以被用于下一个图片的帧间预测。

[0112] 另外,具有色度缩放的亮度映射(LMCS)在图片解码过程中可适用。

[0113] 滤波器450能够通过通过对重构信号应用滤波来改进主观/客观图像质量。例如,滤波器450可以通过将各种滤波方法应用于重构图片来生成修改的重构图片,并将修改的重构图片发送到存储器460中,具体地发送到存储器460的DPB。各种滤波方法可以包括:例如,去

块滤波、样本自适应偏移、自适应环路滤波器、双边滤波器等。

[0114] 存储在存储器460的DPB中的(修改的)重构图片可以被用作帧间预测器432中的参照图片。存储器460可以存储从其中导出当前图片中的运动信息的块的运动信息和/或已经重构的图片中的块的运动信息。存储的运动信息可以被传送到帧间预测器432以用作空间邻近块的运动信息或时间邻近块的运动信息。存储器460可以存储当前图片中的重构块的重构样本并将它们传送到帧内预测器431。

[0115] 另外,VCM解码器(或特征/特征图解码器)执行诸如预测、逆变换和解量化的一系列过程来解码特征/特征图,并且可以基本上具有与上面参照图4描述的图像/视频解码器400相同/相似的结构。然而,VCM解码器与图像/视频解码器400的不同之处在于特征/特征图是解码目标,并且在每个单元(或组件)的名称(例如,DPB等)及其具体操作上可能与图像/视频解码器400不同。VCM解码器的操作可以对应于VCM编码器的操作,并且具体操作将在后面详细描述。

[0116] 特征/特征图编码过程

[0117] 图5是示意性地图示可适用本公开的実施方式的特征/特征图编码过程的流程图。

[0118] 参照图5,特征/特征图编码过程可以包括预测过程(S510)、残差处理过程(S520)和信息编码过程(S530)。

[0119] 预测过程(S510)可以由上面参照图3描述的预测器320执行。

[0120] 具体地,帧内预测器322可以通过参考当前特征/特征图中的特征元素来预测当前块(即,当前编码目标特征元素的集合)。可以基于构成特征/特征图的特征元素的空间相似性来执行帧内预测。例如,图像/视频内的相同兴趣区(RoI)中包括的特征元素可以被估计为具有相似的数据分布特性。因此,帧内预测器322可以通过参考包括当前块的兴趣区内已经重构的特征元素来预测当前块。此时,取决于预测模式,所参考的特征元素可以位于与当前块相邻的位置或者可以位于远离当前块的位置。用于特征/特征图编码的帧内预测模式可以包括多个非定向预测模式和多个定向预测模式。非定向预测模式可以包括例如与图像/视频编码过程的DC模式和平面模式相对应的预测模式。另外,定向模式可以包括与例如图像/视频编码过程的33个定向模式或65个定向模式相对应的预测模式。然而,这是示例,并且可以取决于实施方式以各种方式设置/改变帧内预测模式的类型和数量。

[0121] 帧间预测器321可以基于由关于参考特征/特征图的运动信息指定的参考块(即,参考特征元素的集合)来预测当前块。可以基于构成特征/特征图的特征元素的时间相似性来执行帧间预测。例如,时间上连续的特征可能具有相似的数据分布特性。因此,帧间预测器321可以通过参考时间上与当前特征相邻的特征的已经重构的特征元素来预测当前块。此时,用于指定参考的特征元素的运动信息可以包括运动矢量和参考特征/特征图索引。运动信息可以进一步包括关于帧间预测方向的信息(例如,L0预测、L1预测、Bi预测等)。在帧间预测的情况下,邻近块可以包括存在于当前特征/特征图内的空间邻近块和存在于参考特征/特征图内的时间邻近块。包括参考块的参考特征/特征图和包括时间邻近块的参考特征/特征图可以相同或不同。时间邻近块可以被称为并置参考块等,并且包括时间邻近块的参考特征/特征图可以被称为并置特征/特征图。帧间预测器321可以基于邻近块构造运动信息候选列表,并生成指示哪个候选被用于导出当前块的运动矢量和/或参考特征/特征图索引的信息。可以基于各种预测模式来执行帧间预测。例如,在跳过模式和合并模式的情况

下,帧间预测器321可以使用邻近块的运动信息作为当前块的运动信息。在跳过模式的情况下,与合并模式不同,可以不发送残差信号。在运动矢量预测(MVP)模式的情况下,邻近块的运动矢量被用作运动矢量预测器,并且当前块的运动矢量可以通过用信号发送运动矢量差来指示。除了上述的帧内预测和帧间预测之外,预测器320还可以基于各种预测方法来生成预测信号。

[0122] 由预测器320生成的预测信号可以被用于生成残差信号(残差块、残差特征元素)(S520)。残差处理过程(S520)可以由上面参照图3描述的残差处理器330执行。另外,可以通过针对残差信号的变换和/或量化过程来生成(量化的)变换系数,并且熵编码器340可以将关于比特流中的(量化的)变换系数的信息编码为残差信息(S530)。另外,除了残差信息之外,熵编码器340还可以在比特流中编码对于特征/特征图重构所必要的信息,诸如预测信息(例如,预测模式信息、运动信息等)。

[0123] 另外,特征/特征图编码过程可以进一步不仅包括用于对用于特征/特征图重构的信息(例如,预测信息、残差信息、分割信息等)进行编码并将其以比特流的形式输出的过程(S530)、用于为当前特征/特征图生成重构的特征/特征图的过程,以及用于对重构的特征/特征图应用环路滤波的过程(可选的)。

[0124] VCM编码器可以通过解量化和逆变换从量化的变换系数导出(修改的)残差特征,并且基于步骤S510的输出的预测的特征和(修改的)残差特征生成重构的特征/特征图。以这种方式生成的重构的特征/特征图可以与VCM解码器中生成的重构的特征/特征图相同。当对重构的特征/特征图执行环路滤波过程时,可以通过对重构的特征/特征图进行环路滤波过程来生成修改的重构的特征/特征图。修改后的重构的特征/特征图可以存储在解码特征缓冲器(DFB)或存储器中,并在稍后的特征/特征图预测过程中用作参考特征/特征图。另外,(环路)滤波相关信息(参数)可以被编码并以比特流的形式输出。通过环路滤波过程,可以去除在特征/特征图编译期间可能出现的噪声,并且可以改进基于特征/特征图的任务性能。另外,通过在编码器阶段和解码器阶段都执行环路滤波过程,能够保证预测结果的同一性,能够改进特征/特征图编译的可靠性,并且能够减少用于特征/特征图编译的数据传输量。

[0125] 特征/特征图解码过程

[0126] 图6是示意性地图示可适用本公开的実施方式的特征/特征图解码过程的流程图。

[0127] 参照图6,特征/特征图解码过程可以包括图像/视频信息获取过程(S610)、特征/特征图重构过程(S620至S640)、以及用于重构的特征/特征图的环路滤波过程(S650)。可以对通过本公开中描述的帧间/帧内预测(S620)和残差处理(S630)、用于对在本公开中描述的量化的变换系数的解量化和逆变换过程获取的预测信号和残差信号执行特征/特征图重构过程。可以通过针对重构的特征/特征图的环路滤波过程来生成修改的重构的特征/特征图,并且修改的重构的特征/特征图可以被输出作为解码的特征/特征图。解码的特征/特征图可以存储在解码特征缓冲器(DFB)或存储器中,并且当解码特征/特征图时用作帧间预测过程中的参考特征/特征图。在一些情况下,可以省略上述环路滤波过程。在这种情况下,重构的特征/特征图可以在没有变化的情况下作为解码的特征/特征图被输出,并存储在解码特征缓冲器(DFB)或存储器中,并且然后在解码特征/特征图时被用作帧间预测过程中的参考特征/特征图。

## [0128] 实施方式

[0129] 在生成了图像的压缩比特流以执行机器任务的情况下,在现有技术中,选择任一个预定的预测模型来执行帧内预测。

[0130] 由于可选择的预测模型数量是有限的,现有技术可能不足以应用适合于当前块的特征的预测模型,从而可能增加比特流的比特量。

[0131] 在应用可变预测模型来执行适合于当前块的特征的预测时,需要在特征编码装置10和特征解码装置20上使用相同的参数。因此,需要将在特征编码装置10中使用的预测模型的参数发送到特征解码装置20,并且预测模型可变性的增加也可能增加对应参数的数量,从而可能增加比特流的比特量。

[0132] 本公开提出了用于对当前块执行帧内预测的方法,该方法不仅使用已经完全训练并因而具有预定参数的预测模型,而且通过使用新学习的参数的预测模型,新学习的参数是通过使用重构的邻居参考样本(重构的邻居样本)来学习的。

[0133] 此外,本公开提出了用于将完全学习的参数存储在列表中,然后使用列表中的参数之一作为用于当前块预测的参数的方法。此外,本公开还提出了用于将列表中的参数之一作为用于预测当前块的训练的初始值的方法。

[0134] 因此,通过本公开,如果生成了图像的压缩比特流来执行机器任务,则与现有方法相比,具有较小大小(少量比特)的比特流可以使得机器任务能够以相似的准确程度执行。

[0135] 本公开中描述的实施方式可以单独或以其组合进行操作。

[0136] 图8是示出根据本公开的実施方式的特征编码方法的流程图。

[0137] 参照图8,特征编码装置10可以基于重构的邻居样本来训练预测模型(S810)。“预测模型”可以是“预测模型的参数”,并且“预测模型的训练”可以是“对预测模型的参数的训练”。对预测模型的训练可以进行n次或更多次(n是等于或大于0的整数)。

[0138] 重构的邻居样本可以通过基于训练的帧内预测来解码的样本,也可以是通过使用现有的帧内预测方法解码的样本。此外,重构的邻居样本可以通过使用包括帧间预测在内的其他预测方法解码的样本。

[0139] 在图7中示出了可用于训练预测模型的重构的邻居样本和当前编码/解码目标样本(当前样本或当前块)的示例。参照图7,在标记有向量x的区域中的样本表示当前编码/解码目标样本,并且在标记有 $r$ 、 $r_0$ 、 $r_1$ 、 $r_2$ 、 $r_3$ 、 $x_0$ 、 $x_1$ 、 $x_2$ 和 $x_3$ 的区域中的样本表示与当前块左上方相邻的重构的邻居样本。

[0140] 在图7中用作训练数据的邻居样本的位置分布、形状和数量仅是为方便描述的一个示例,而实际的邻居样本可以具有与图7的示例不同的分布、形状或数量。例如,用作训练数据以确保帧内预测模式的编码和解码的独立性的重构的邻居样本可以被限制为仅在帧内编码和解码模式中使用,或者可以被限制为仅在预定义的相同编码和解码单元(例如图块、切片等)中使用。

[0141] 在使用重构的邻居样本 $r$ 的值来预测当前块的样本(当前样本) $x$ 的值的情况下,由具有参数 $\theta_t$ 的预测模型 $f$ 预测的预测样本的值 $x_{\text{pred}}$ 可以如等式1所示地表示。

[0142] [等式1]

$$[0143] \quad x_{\text{pred}} = f(r; \theta_t)$$

[0144] 在等式1中, $\theta_t$ 表示根据预定的训练方法经过t次训练过程的预测模型的参数值。

[0145] 训练 $\theta_t$ 的过程可以被描述为重复搜索使得损失函数 $L_{tr}$ 最小化的 $\theta_t$ 的优化过程, 损失函数 $L_{tr}$ 可以由L2范数之和来表示, 如等式2所示。

[0146] [等式2]

$$L_{tr} = \sum_n ||x_n - f(r_n; \theta_t)||_2$$

[0148] 如等式2所示, 训练 $\theta_t$ 的过程可以从初始值开始, 并以N对重构的参考样本 $\{r_0, x_0\}$ ,  $\{r_1, x_1\}$ ,  $\{r_2, x_2\}$ ,  $\{r_3, x_3\}$ ,  $\dots$ ,  $\{r_N, x_N\}$ 作为训练数据, 并且以重构的 $x_N$ 作为输入。等式2的损失函数仅是为方便描述的一个示例, 可以使用不同的等式。

[0149] 特征编码装置10可以使用经训练的预测模型来预测当前块(S820)。特征编码装置10可以基于所预测的当前块(预测块)和当前块生成针对当前块的残差信息, 并且对所生成的残差信息进行编码(S830)。

[0150] 图9是示出根据本公开的実施方式的特征解码方法的流程图。

[0151] 参照图9, 特征解码装置20可以基于重构的邻居样本训练预测模型(S910)。

[0152] 预测模型可以是预测模型的参数, 并且预测模型的训练可以是对预测模型的参数的训练。对预测模型的训练可以执行n次或更多次(n为大于或等于0的整数)。

[0153] 重构的邻居样本可以通过基于训练的帧内预测而解码的样本, 或者是通过使用现有的帧内预测方法而解码的样本。此外, 重构的邻居样本可以通过使用包括帧间预测在内的其他预测方法而解码的样本。用于训练预测模型的重构的邻居样本和当前编码/解码目标样本的示例可与图7的示例相同。

[0154] 特征解码装置20可以通过使用经训练的预测模型对当前块进行预测(S920)。特征解码装置20可以基于针对当前块的预测块和残差块重构当前块(S930)。残差块可以基于从特征编码装置10用信号通知的残差信息导出。

[0155] 实施方式1

[0156] 实施方式1对应于用于通过使用重构的邻居样本训练预测模型, 然后通过经训练的预测模型执行预测的方法。

[0157] 图10是示出根据实施方式1的特征编码方法的流程图。图10的特征编码方法示出了训练预测模型的参数 $\theta_t$ 、预测当前样本并对其差异进行编码的过程。

[0158] 参照图10, 可以配置可用于训练预测模型的多种变量(S1002), 并且可以基于重构的邻居样本训练预测模型(S1004)。预测模型的参数 $\theta_t$ 的训练(S1004)可以从初始值 $\theta_0$ 开始, 并继续进行以降低损失函数 $L_{tr}$ , 如等式2所例示的。

[0159] 具有初始值 $\theta_0$ 的预测模型可以是“第一预测模型”。根据实施方式, 第一预测模型可以是具有初始值 $\theta_0$ 的预测, 或者是在重复训练预测模型时作为先前训练的结果而导出的预测模型。因此, 预测模型的训练可以包括基于重构的邻居样本和第一预测模型预测当前块的过程。

[0160] 如等式2中所例示的损失函数 $L_{tr}$ 可以对应于基于预测样本和重构的邻居样本之间的差异而导出的值。因此, 预测模型的训练可以包括训练第一模型以使得基于预测样本和重构的邻居样本之间的差异而导出的值减小的过程。

[0161] 训练结果可以存储为 $\theta_{t+1}$ , 并且可以重复执行这样的过程( $t=t+1$ )(S1006)。上述过程可重复执行预定义的次数 $T_{max}$ (S1012), 并且可以通过使用应用了重构的邻居样本r和训练结果 $\theta_t$ 的预测模型来导出等式1的预测值 $x_{pred}$ (S1006)。此外, 如等式3中所示, 可以以L2

范数的形式导出相对于当前样本的值 $x$ 的损失函数 $L(x, x_{\text{pred}})$  (S1008)。

[0162] [等式3]

$$[0163] \quad L(x, x_{\text{pred}}) = \|x - x_{\text{pred}}\|_2$$

[0164] 等式3的损失函数仅是为方便描述的一个示例,可以使用不同的等式。

[0165] 损失函数 $L(x, x_{\text{pred}})$ 的最小值可以是 $L_{\text{min}}$ ,对应时间的值 $\theta_t$ 可以设置为 $\theta_{\text{opt}}$ ,并且对应时间的 $t$ 可以设置为 $t_{\text{opt}}$ (关于次数的信息) (S1016)。另选地,在损失函数 $L(x, x_{\text{pred}})$ 的值小于预定义阈值 $L_{\text{thr}}$  (S1008)的情况下,无法执行训练达到最大次数 $T_{\text{max}}$  (S1012),那么可以将对应时间的值 $\theta_t$ 设置为 $\theta_{\text{opt}}$ ,可以将对应时间的 $t$ 设置为 $t_{\text{opt}}$  (S1014)。

[0166] 可以如等式4所示地导出用于编码的帧内预测结果 $x_{\text{pred}}$ ,并且可以如等式5所示地导出当前样本的值 $x$ 与预测值 $x_{\text{pred}}$ 之间的差 $x_{\text{diff}}$ 。

[0167] [等式4]

$$[0168] \quad x_{\text{pred}} = f(r; \theta_{\text{opt}})$$

[0169] [等式5]

$$[0170] \quad x_{\text{diff}} = x - x_{\text{pred}}$$

[0171] 可以对差 $x_{\text{diff}}$ 进行编码 (S1024),并从特征编码装置10向特征解码装置20用信号通知。另外,如果在解码时,训练过程重复了与编码时的相同的次数,则可以导出与编码相同的训练结果 $\theta_t$ ,并且由于这可以应用于预测,因此也可以对 $t_{\text{opt}}$ 进行编码 (S1020),并且可以从特征编码装置10用信号通知给特征解码装置20。

[0172] 如果 $t_{\text{opt}} = 0$  (S1018),则可以对 $f_{\text{skip}}$ (第一标志)进行编码 (S1022),并且可以从特征编码装置10用信号通知给特征解码装置20,该第一标志指示不需要执行预测模型的训练过程。在这种情况下,可以跳过对 $t_{\text{opt}}$ 的编码。

[0173] 根据实施方式,可以在特征编码装置10和特征解码装置20中预定义 $t_{\text{opt}}$ 的默认值。在这种情况下,可以不对 $t_{\text{opt}}$ 进行编码。另外,虽然图10描述了仅定义一个预测模型的情况,但是可以定义多个预测模型,并且这种情况可以具有索引,该索引被编码以从多个预测模型中选择任一个预测模型。

[0174] 图11是示出根据实施方式1的特征解码方法的流程图。图11的特征解码方法示出了训练预测模型的参数 $\theta_t$ 、预测当前样本以及通过使用解码的差异来重构当前样本的过程。

[0175] 参照图11,可以配置可用于训练预测模型的多种变量 (S1100),并且可以基于重构的邻居样本来训练预测模型 (S1030)。与特征编码方法类似,在训练预测模型的参数 $\theta_t$ 时,能够重复执行以下过程 (S1140):从初始值 $\theta_0$ 开始,继续进行以降低等式2所例示的损失函数 $L_{\text{tr}}$ ,并且将训练结果存储为 $\theta_{t+1}$ 。

[0176] 具有初始值 $\theta_0$ 的预测模型可以是“第一预测模型”。根据实施方式,第一预测模型可以是具有初始值 $\theta_0$ 的预测或者是当预测模型的训练被重复时作为先前训练的结果而导出的预测模型。因此,预测模型的训练可以包括基于重构的邻居样本和第一预测模型来预测当前块的过程。

[0177] 等式2中所例示的损失函数 $L_{\text{tr}}$ 可以对应于基于预测样本和重构的邻居样本之间的差异而导出的值。因此,预测模型的训练可以包括训练第一模型以使得基于预测样本和重构的邻居样本之间的差异而导出的值减小的过程。

[0178] 当训练过程重复与解码的 $t_{opt}$  (S1120) 指示的数量相同的次数时 (S1150), 特征解码装置20可以因此获得与特征编码装置10用于预测的 $\theta_{t_{opt}}$  相同的预测模型参数。随着训练过程重复与 $t_{opt}$  所指示数量相同的次数, 并且使用结果进行预测, 可以如等式6所示地导出与特征编码装置10中相同的预测值 $x_{pred}$  (S1160)。

[0179] [等式6]

$$[0180] \quad x_{pred} = f(r; \theta_{t_{opt}})$$

[0181] 另选地, 在作为指示不需要单独的预测模型参数训练过程的标志的 $f_{skip}$  (第一标志) 的值为1时 (S1110), 可以跳过对 $t_{opt}$  的解码 (S1120) 和 $\theta_{t_{opt}}$  的训练过程 (S1130、S1140、S1150), 使用预测模型的初始默认值执行预测, 从而可以导出预测值 $x_{pred}$  (S1160)。

[0182] 如等式7所示, 通过使用预测值 $x_{pred}$  和解码的差 $\hat{x}_{diff}$ , 可以导出当前样本的重构值 $\hat{x}$ 。

[0183] [等式7]

$$[0184] \quad \hat{x} = x_{pred} + \hat{x}_{diff}$$

[0185] 根据实施方式, 可以在特征编码装置10和特征解码装置20中预定义 $t_{opt}$  的默认值。在这种情况下, 可以不对 $t_{opt}$  进行解码。另外, 虽然图11描述了仅定义一个预测模型的情况, 但是可以定义多个预测模型, 并且这种情况可以具有索引, 该索引被解码以从多个预测模型中选择任一个预测模型。

[0186] 表1和表2示出了用信号通知以执行根据实施方式1的特征编码方法和特征解码方法的句法元素的示例。具体地, 表1示出了指示是否在视频序列内使用基于训练的帧内预测的句法元素。表1示出了指示是否使用基于训练的帧内预测以及可用块的大小的SPS级别的句法元素, 但这仅是一个示例, 并且可以以诸如PPS、图片头、切片头之类的其他级别来定义句法元素。表2示出了指示以下项的句法元素: 个体CU的帧内预测模式是否是基于训练的帧内预测、所使用的预测模型的索引、是否训练附加预测模型参数、以及训练次数。

[0187] [表1]

	seq_parameter_set_rbsp() {	描述符
	...	
	<b>sps_lip_enabled_flag</b>	<b>u(1)</b>
	if(sps_lip_enabled_flag) {	
[0188]	<b>sps_lip_enabled_greater_flag</b>	<b>u(1)</b>
	...	
	}	
	...	
	}	

[0189] [表2]

	<code>coding_unit(x0, y0, cbWidth, cbHeight, ...) {</code>	描述符
	<code>...</code>	
	<code>if(sps_lip_enabled_flag &amp;&amp; (sps_lip_enabled_greater_flag    (cbWidth &lt;= 4 &amp;&amp; cbHeight &lt;=4))) {</code>	
	<code>coding_unit(x0, y0, ...) {</code>	描述符
	<code>...</code>	
	<code>if(sps_lip_enabled_flag) {</code>	
	<code>intra_lip_flag</code>	ae(v)
[0190]	<code>if(intra_lip_flag) {</code>	
	<code>intra_lip_model_idx[ x0 ][ y0 ]</code>	ae(v)
	<code>intra_lip_skip_learning_flag</code>	ae(v)
	<code>if(intra_lip_skip_learning_flag == 0) {</code>	
	<code>intra_lip_num_itr[ x0 ][ y0 ]</code>	ue(v)
	<code>}</code>	
	<code>}</code>	
	<code>}</code>	
	<code>...</code>	
	<code>}</code>	

[0191] sps\_lip\_enabled\_flag指示是否在视频序列中使用基于训练的帧内预测。sps\_lip\_enabled\_greater\_flag指示是否在视频序列内大于预定义大小的块中可以使用基于训练的帧内预测。表2描述了预定义大小为4且sps\_lip\_enabled\_greater\_flag的数量为1,但这仅是为了方便描述的一个示例。也就是说,预定义大小可以定义为另一值,并且对于各种块大小可以存在多个sps\_lip\_enabled\_greater\_flag。

[0192] intra\_lip\_flag (帧内预测模式信息) 指示对应CU的预测模式是否是基于训练的帧内预测。intra\_lip\_model\_idx (预测模型信息) 指示在向对应CU应用基于训练的帧内预测时将使用的预定义预测模型的索引。intra\_lip\_model\_idx可以指示至少一个或更多个候选预测模型当中将要用作第一预测模型的预测模型。intra\_lip\_skip\_learning\_flag (第一标志) 指示在解码对应CU时是否需要针对预测模型参数的附加训练过程,这是与f\_skip相对应的句法元素。当intra\_lip\_skip\_learning\_flag具有0值时,它指示应当执行附加训练,并且当intra\_lip\_skip\_learning\_flag具有值1时,它指示在不进行进一步训练的情况下通过使用对应预测模型参数的初始默认值执行预测。intra\_lip\_num\_itr指示使用在解码CU时重构的CU样本和确定的预测模型执行训练的次数(次数信息)。intra\_lip\_num\_itr是与t\_opt相对应的句法元素。intra\_lip\_num\_itr可以以各种方式进行编码,例如通过按原样编码t\_opt的值或示出与默认值(默认次数)的差,如t\_opt = 默认值 + intra\_lip\_num\_itr。

[0193] 实施方式2

[0194] 实施方式2提出了将经训练的参数(或经训练的预测模型)存储在列表中的方法、通过使用列表中存储的任一个预测模型作为训练的初始值来执行训练的方法、以及在不执行训练过程的情况下通过使用列表中存储的任一个预测模型来执行预测的方法。

[0195] 其中存储了经训练的预测模型或未训练的预测模型的列表可以称为“预测模型候选列表”或“参数列表T”,从预测模型候选列表中的预测模型中选择的预测模型可以称为“第一预测模型”。

[0196] 图12a和图12b是示出根据实施方式2的特征编码方法的流程图。图12的特征编码

方法示出了选择和训练预测模型的参数 $\theta_t$ 、预测当前样本并对其差异进行编码的过程。

[0197] 参照图12a和图12b,可以将用于预测重构块的预测模型存储在预测模型候选列表(或参数列表T)中(S1228)。也就是说,预测模型候选列表可以包括用于预测至少一个相邻块的至少一个预测模型候选。例如,如果块n被完全编码,并且在编码该块时使用了参数 $\theta_{opt}^{(n)}$ 进行预测,则在编码块n时,预测模型候选列表中存储有 $T = \{\theta_{opt}^{(0)}, \theta_{opt}^{(1)}, \theta_{opt}^{(2)}, \dots, \theta_{opt}^{(N-1)}\}$ 的值。为了便于描述,图12a和图12b描述了每个 $\theta_{opt}^{(n)}$ 都存储在预测模型候选列表中,但可以仅存储 $\theta_{opt}^{(n)}$ 的一部分或者可以根据预定规则删除旧项。

[0198] 像等式2中一样,可以执行对预测模型的参数 $\theta_t$ 的训练,以减小损失函数 $L_{tr}$ (S1214),并且可以将训练结果存储为 $\theta_{t+1}$ (S1216)。此外,可以重复执行这些过程。在这种情况下,训练的初始值 $\theta_0$ 可以像实施方式1一样从默认初始值开始,或者从预测模型候选列表中存储的参数当中的第 $n\_idx$ 个参数 $\theta_{opt}^{(n\_idx)}$ 开始。也就是说,可以基于预测模型候选列表和预定预测模型中的一者来确定第一预测模型。具体而言,与训练的初始值相对应的第一预测模型可以被确定为预定预测模型(默认初始值),或者可以被确定为预测模型候选列表中存储的预测模型当中的第 $n\_index$ 个预测模型。

[0199] 对应过程可重复执行预定义的次数 $T_{max}$ (S1222),并且可以通过使用应用了重构的邻居样本 $r$ 和训练结果 $\theta_t$ 的预测模型如等式1中那样导出预测值 $x_{pred}$ (S1216)。此外,如等式3所例示的,可以导出相对于当前样本的值 $x$ 的损失函数 $L(x, x_{pred})$ (S1218)。

[0200] 损失函数 $L(x, x_{pred})$ 的最小值可以是 $L_{min}$ ,对应时间的值 $\theta_t$ 可以设置为 $\theta_{opt}$ ,对应时间的 $t$ 可以设置为 $t_{opt}$ (关于次数的信息)(S1226)。另选地,在损失函数 $L(x, x_{pred})$ 的值小于预定义阈值 $L_{thr}$ 的情况下(S1218),无法执行训练达最大次数 $T_{max}$ (S1222),然后可将对应时间的值 $\theta_t$ 设置为 $\theta_{opt}$ ,并且可以将对应时间的 $t$ 设置为 $t_{opt}$ (S1224)。

[0201] 根据实施方式,在不进行预测模型的训练过程的情况下,可以原样使用预定义的默认初始值执行预测,或者可使用作为预测模型候选列表当中的第 $n\_idx$ 个预测模型的 $\theta_{opt}^{(n\_idx)}$ 执行预测(S1204)。

[0202] 如等式4和等式5所例示的,可以导出预测结果 $x_{pred}$ 以及当前样本的值 $x$ 与 $x_{pred}$ 之间的差 $x_{diff}$ 。可以对差 $x_{diff}$ 进行编码(S1232),并且从特征编码装置10发信号通知给特征解码装置20。另外,如果在解码时,训练过程重复与在编码时相同的次数,则可以导出与编码相同的训练结果 $\theta_t$ ,并且由于这可以应用于预测,因此也可以对 $t_{opt}$ 进行编码(S1230),并且可以从特征编码装置10发信号通知给特征解码装置20。如果 $t_{opt} = 0$ (S1208),则可以对作为指示不需要执行预测模型的训练过程的标志的 $f\_skip$ (第一标志)进行编码(S1212),并从特征编码装置10发信号通知给特征解码装置20。在这种情况下,可以跳过 $t_{opt}$ 的编码。

[0203] 如图12a和图12b所示,在训练过程之前,可以确定预测模型训练的初始值 $\theta_0$ 是从默认初始值开始,还是从预测模型候选列表中的预测模型中的一个预测模型开始,或者是否在跳过训练过程时原样使用预测模型候选列表中的预测模型(S1202和S1203)。但是,这仅是一个示例,可以部分或全部执行编码过程(或训练过程),然后选择并执行更有效的方法。在这样的示例中,在用于预测的 $\theta_{opt}$ 是经历了附加训练过程的结果值的情况下,可以将

$\theta_{opt}$  添加到预测模型候选列表中。

[0204] 虽然在图12a和图12b中没有表达,但是可以存在用于将 $\theta_{opt}$ 添加到预测模型候选列表的条件。此外,可以存在用于从预测模型候选列表中删除在预测模型候选列表中存储的 $\theta_{opt}^{(n)}$ 中的一些的条件。

[0205] 可以对值f\_list进行编码(S1203和S1204),并且从特征编码装置10发信号通知给特征解码装置20,该值指示是使用存储在预测模型候选列表中的预测模型执行训练和预测,还是使用默认初始值执行预测。在这种情况下,可以对指示用于预测的预测模型的索引n\_idx(选择信息)进行编码(S1204),并且从特征编码装置10发信号通知给特征解码装置20。

[0206] 根据实施方式,可以在特征编码装置10和特征解码装置20中预定义t\_opt的默认值。在这种情况下,可以不对t\_opt进行编码。此外,虽然图12a和图12b描述了仅定义一个预测模型的情况,但可以定义多个预测模型,并且这种情况可以具有被编码为从多个预测模型中选择任一个预测模型的索引。此外,可以存在多个预测模型候选列表。

[0207] 图13a和图13b是示出根据实施方式2的特征解码方法的流程图。图13a和图13b的特征解码方法示出了选择和训练预测模型的参数 $\theta_t$ 、预测当前样本以及通过使用解码的差来重构当前样本的过程。

[0208] 参照图13a和图13b,可以对 $x_{diff}$ 进行解码,并且可以设置 $t=0$ (S1302)。依据作为指示是否使用存储在预测模型候选列表中的预测模型来执行训练和预测的标志f\_list的值(S1304),可以确定预测模型参数的初始值 $\theta_0$ 。具体地,当 $f\_list=1$ 时,可以对索引n\_idx(选择信息)进行解码,并且可以将预测模型参数的初始值 $\theta_0$ 确定为预测模型候选列表中的由n\_idx指示的预测模型 $\theta_{opt}^{(n\_idx)}$ 。当 $f\_list=0$ 时,可以将预测模型参数的初始值 $\theta_0$ 确定为预定义的默认值(S1308)。

[0209] 类似于特征编码方法,如等式3所例示的,可以重复执行预测模型的参数 $\theta_t$ 的训练(S1314),以降低损失函数 $L_{tr}$ ,并将训练结果存储为 $\theta_{t+1}$ (S1316)。当训练过程重复解码的t\_opt(S1312)所指示的数量一样多次时(S1318),特征解码装置20可因此获得与特征编码装置10用于预测的 $\theta_{opt}$ 相同的预测模型参数。另选地,当作为指示不需要单独的训练过程的标志的f\_skip(第一标志)被解码为值1时(S1310),可以依据f\_list的值不执行t\_opt的解码和预测模型参数的训练(S1304),并且可以使用预测模型的默认值或预测模型候选列表中的第n\_idx个预测模型 $\theta_{opt}^{(n\_idx)}$ 作为 $\theta_{opt}$ 。

[0210] 由于训练过程重复t\_opt所指示的数量一样多次并且使用结果执行预测,因此可以如等式6所示地导出与特征编码装置10中相同的预测值 $x_{pred}$ (S1322)。此外,如等式7所示,通过使用预测值 $x_{pred}$ 和解码的差 $\hat{x}_{diff}$ ,可以导出当前样本的重构值 $\hat{x}$ (S1322)。

[0211] 根据实施方式,可以在特征编码装置10和特征解码装置20中预定义t\_opt的默认值。在这种情况下,可以不解码t\_opt。另外,虽然图13a和图13b描述了仅定义一个预测模型的情况,但是可以定义多个预测模型,并且这种情况可以具有索引,该索引被解码以从多个预测模型中选择任一个预测模型。另外,可以存在多个预测模型候选列表。

[0212] 当 $f\_skip=0$ (S1310)时,可以将用于预测的 $\theta_{opt}$ 添加到预测模型候选列表中(S1320)。虽然图13a和图13b中没有表达,但可以存在用于将 $\theta_{opt}$ 添加到预测模型候选列表

中的条件。此外,可以存在用于从预测模型候选列表中删除在预测模型候选列表中存储的  $\theta_{opt}^{(n)}$  中的一些的条件。

[0213] 表3和表4示出了被用信号通知以执行根据实施方式2的特征编码方法和特征解码方法的句法元素的示例。具体地,表3示出了指示是否在视频序列内使用基于训练的帧内预测以及基于训练的帧内预测适用的块的大小的句法元素。表3示出了SPS级别的句法元素,但这仅是一个示例,并且可以以诸如PPS、图片头、切片头之类的其他级别来定义句法元素。表4示出了指示以下项的句法元素:个体CU的帧内预测模式是否是基于训练的帧内预测、所使用的预测模型的索引、是否使用存储在预测模型候选列表中的预测模型、是否训练附加预测模型参数、以及训练次数。

[0214] [表3]

	描述符
seq_parameter_set_rbsp() {	
...	
<b>sps_lip_enabled_flag</b>	<b>u(1)</b>
if(sps_lip_enabled_flag) {	
<b>sps_lip_enabled_grater_flag</b>	<b>u(1)</b>
<b>sps_lip_parm_list_flag</b>	<b>u(1)</b>
}	
...	
}	

[0215]

[0216] [表4]

	描述符
coding_unit(x0, y0, ...) {	
...	
if(sps_lip_enabled_flag && (sps_lip_enabled_grater_flag    (cbWidth <= 4 && cbHeight <=4))) {	
if(sps_lip_enabled_flag) {	
<b>intra_lip_flag</b>	<b>ae(v)</b>
if(intra_lip_flag) {	
<b>intra_lip_model_idx[ x0 ][ y0 ]</b>	<b>ae(v)</b>
if(sps_lip_parm_list_flag) {	
<b>intra_lip_parm_from_list_flag</b>	<b>ae(v)</b>
if(intra_lip_parm_from_list_flag) {	
<b>intra_lip_parm_idx[ x0 ][ y0 ]</b>	<b>ae(v)</b>
}	
}	
<b>intra_lip_skip_learning_flag</b>	<b>ae(v)</b>
if(intra_lip_skip_learning_flag == 0) {	
<b>intra_lip_num_itr[ x0 ][ y0 ]</b>	<b>ue(v)</b>
}	
}	
}	
...	
}	

[0217]

[0218] sps\_lip\_enabled\_flag指示是否在视频序列中使用基于训练的帧内预测。sps\_lip\_enabled\_grater\_flag指示是否在视频序列内大于预定义大小的块中可以使用基于

训练的帧内预测。表3描述了预定义大小为4且sps\_lip\_enabled\_greater\_flag的数量为1,但这仅是为了方便描述的一个示例。也就是说,预定义大小可以定义为另一值,并且针对各种块大小可以存在多个sps\_lip\_enabled\_greater\_flag。sps\_lip\_parm\_list\_flag(第一标志)指示是否在视频序列中使用预测模型候选列表中存储的预测模型。

[0219] intra\_lip\_flag(帧内预测模式信息)指示对应CU的预测模式是否是基于训练的帧内预测。intra\_lip\_model\_idx(预测模型信息)指示在向对应CU应用基于训练的帧内预测时将使用的预定义预测模型的索引。intra\_lip\_model\_idx可以指示至少一个或更多个候选预测模型当中将要用作第一预测模型的预测模型。intra\_lip\_parm\_from\_list\_flag指示在解码对应CU时存储在预测模型候选列表中的值是否用作预测模型参数的初始值。intra\_lip\_parm\_from\_list\_flag是与f\_lst对应的句法元素。当intra\_lip\_parm\_from\_list\_flag==0时,使用默认初始值,并且当intra\_lip\_parm\_from\_list\_flag==1时,附加地解码指示存储在预测模型候选列表中的值的索引。intra\_lip\_parm\_idx是指示预测模型候选列表中存储的预测模型当中的将用作用于解码的当前CU的初始值的预测模型的句法元素。intra\_lip\_parm\_idx是与n\_idx(选择信息)对应的句法元素。intra\_lip\_skip\_learning\_flag(第一标志)指示在解码对应CU时是否需要针对预测模型参数的附加训练过程,这是与f\_skip对应的句法元素。当intra\_lip\_skip\_learning\_flag具有值0时,它指示应当执行附加训练,并且当intra\_lip\_skip\_learning\_flag具有值1时,它指示在不进行进一步训练的情况下通过使用对应预测模型参数的初始默认值来执行预测。intra\_lip\_num\_itr指示使用当CU被解码时重构的CU的样本和确定的预测模型来执行训练的次数(次数信息)。intra\_lip\_num\_itr是与t\_opt对应的句法元素。可以以各种方式对intra\_lip\_num\_itr进行编码,例如,按原样对t\_opt的值进行编码,或者示出与默认值(默认次数)的差,如 $t\_opt = \text{默认值} + \text{intra\_lip\_num\_itr}$ 。

### [0220] 实施方式3

[0221] 上述实施方式已描述了f\_skip(第一标志)的编码,该第一标志指示如果 $t\_opt = 0$ ,则不需要在解码过程中执行预测模型的训练过程。

[0222] 根据实施方式,t\_opt==0指示执行零次训练,并且这可以指示不执行训练。因此,当不执行训练时,特征编码装置10可以将t\_opt的值编码为0,并且当t\_opt的值为0时,特征解码装置20可以不执行训练。此外,当t\_opt的值为0时,可以不对f\_skip进行编码和解码。

[0223] 虽然为了描述清楚起见,上述本公开的示例性方法被表示为一系列操作,但是并不旨在限制其中步骤被执行的顺序,并且必要时这些步骤可以被同时执行或以不同的顺序被执行。为了实现根据本公开的方法,所描述的步骤可以进一步包括其他步骤,可以包括除了一些步骤之外的剩余步骤,或者可以包括除了一些步骤之外的其他附加步骤。

[0224] 本公开中,执行预先确定的操作(步骤)的图像编码设备或图像解码设备可以执行确认相应操作(步骤)的执行条件或情形的操作(步骤)。例如,如果描述当满足预先确定的条件时执行预先确定的操作,则图像编码设备或图像解码设备可以在确定是否满足预先确定的条件之后执行预先确定的操作。

[0225] 本公开的各种实施方式并不是所有可能组合的列表,而是旨在描述本公开的代表性方面,并且各种实施方式中描述的内容可以独立地应用或者以两个或更多个的组合应

用。

[0226] 本公开中描述的实施方式可以在处理器、微处理器、控制器或芯片上实现和执行。例如,各图中所示的功能单元可以在计算机、处理器、微处理器、控制器或芯片上实现和执行。在这种情况下,用于实现的信息(例如,关于指令的信息)或算法可以存储在数字存储介质中。

[0227] 另外,应用本公开的实施方式的解码器(解码装置)和编码器(编码装置)可以被包括在多媒体广播发送和接收设备、移动通信终端、家庭影院视频设备、数字影院视频设备、监控摄像机、视频聊天设备、诸如视频通信的实时通信设备、移动流设备、存储介质、摄像机、视频点播(VoD)服务提供设备、OTT视频(顶部视频)设备、互联网流服务提供设备、三维(3D)视频设备、增强现实(AR)设备、可视电话视频设备、交通终端(例如,车辆(包括自动驾驶车辆)终端、机器人终端、飞机终端、船舶终端等)和医疗视频设备等中,并且可以被用于处理视频信号或数据信号。例如,OTT视频设备可以包括游戏控制台、蓝光播放器、互联网接入电视、家庭影院系统、智能电话、平板PC、数字录像机(DVR)等。

[0228] 另外,应用本公开的实施方式的处理方法能够以由计算机执行并存储在计算机可读记录介质中的程序的形式产生。具有根据此文档的实施方式的数据结构的多媒体数据也可以存储在计算机可读记录介质中。计算机可读记录介质包括存储计算机可读数据的所有类型的存储设备和分布式存储设备。计算机可读记录介质包括例如蓝光盘(BD)、通用串行总线(USB)、ROM、PROM、EPROM、EEPROM、RAM、CD-ROM、磁带、软盘和光学数据存储设备。另外,计算机可读记录介质包括以载波形式实现的介质(例如,经由互联网传输)。另外,通过编码方法产生的比特流可以存储在计算机可读记录介质中或者通过有线或无线通信网络发送。

[0229] 另外,本公开的实施方式可以通过程序代码实现为计算机程序产品,并且程序代码可以由本公开的实施方式在计算机上执行。程序代码可以存储在计算机可读的载体上。

[0230] 图14是图示可适用本公开的实施方式的内容流传输系统的示例的视图。

[0231] 参照图14,应用本公开的实施方式的内容流传输系统可以主要包括编码服务器、流媒体服务器、网络服务器、web存储器、用户设备和多媒体输入设备。

[0232] 编码服务器将从诸如智能电话、相机、摄像机等的多媒体输入设备输入的内容压缩为数字数据以生成比特流并将该比特流发送到流媒体服务器。作为另一示例,当诸如智能电话、相机、摄像机等的多媒体输入设备直接生成比特流时,可以省略编码服务器。

[0233] 比特流可以由应用本公开的实施方式的图像编码方法或图像编码设备生成,并且流媒体服务器可以在发送或接收比特流的过程中临时存储比特流。

[0234] 流媒体服务器基于用户的请求通过网络服务器将多媒体数据发送到用户设备,并且网络服务器用作用于向用户通知服务的媒介。当用户向网络服务器请求期望的服务时,网络服务器可以将其传送到流媒体服务器,并且流媒体服务器可以向用户发送多媒体数据。在这种情况下,内容流传输系统可以包括单独的控制服务器。在这种情况下,控制服务器用作控制内容流传输系统中的设备之间的命令/响应。

[0235] 流传输服务器可以从媒体存储和/或编码服务器接收内容。例如,当从编码服务器接收内容时,可以实时接收内容。在这种情况下,为了提供流畅的流服务,流媒体服务器可以在预先确定的时间内存储比特流。

[0236] 用户设备的示例可以包括移动电话、智能电话、膝上型计算机、数字广播终端、个

人数字助理 (PDA)、便携式多媒体播放器 (PMP)、导航、板式PC、平板PC、超级本、可穿戴设备 (例如,智能手表、智能眼镜、头戴式显示器)、数字电视、台式电脑、数字标牌等。

[0237] 内容流传输系统中的每个服务器可以作为分布式服务器来操作,在这种情况下,从每个服务器接收到的数据可以被分发。

[0238] 图15是示出可适用本公开的实施方式的内容流传输系统的另一示例的图。

[0239] 参照图15,在诸如VCM的实施方式中,根据设备的性能、用户的请求、要执行的任务的特性等,可以在用户终端中执行任务或者可以在外部设备 (例如,流媒体服务器、分析服务器等) 中执行任务。以这种方式,为了将执行任务所必要的信息发送到外部设备,用户终端可以生成包括直接或通过编码服务器执行任务所必要的信息 (例如,诸如任务、神经网络和/或用法的信息) 的比特流。

[0240] 在实施方式中,分析服务器可以在对从用户终端 (或从编码服务器) 接收到的编码的信息进行解码之后执行用户终端请求的任务。此时,分析服务器可以将通过任务执行获得的结果发送回到用户终端,或者可以将其发送到另一链接的服务服务器 (例如,网络服务器)。例如,分析服务器可以将通过执行确定火灾的任务获得的结果发送到火灾相关服务器。在这种情况下,分析服务器可以包括单独的控制服务器。在这种情况下,控制服务器可以用作控制与分析服务器和服务器相关联的每个设备之间的命令/响应。另外,分析服务器可以基于要由用户设备执行的任务和可以执行的任务信息来从网络服务器请求期望的信息。当分析服务器向网络服务器请求期望的服务时,网络服务器将其发送到分析服务器,并且分析服务器可以将数据发送到用户终端。在这种情况下,内容流传输系统的控制服务器可以用作控制流传输系统中的设备之间的命令/响应。本公开的范围包括用于使能够在装置或计算机上执行根据各种实施方式的方法的操作的软件或机器可执行命令 (例如,操作系统、应用、固件、程序等)、其上存储有这样的软件或命令并且可在装置或计算机上执行的非暂时性计算机可读介质。

[0241] 工业实用性

[0242] 本公开的实施方式可以用于对特征/特征图进行编码或解码。

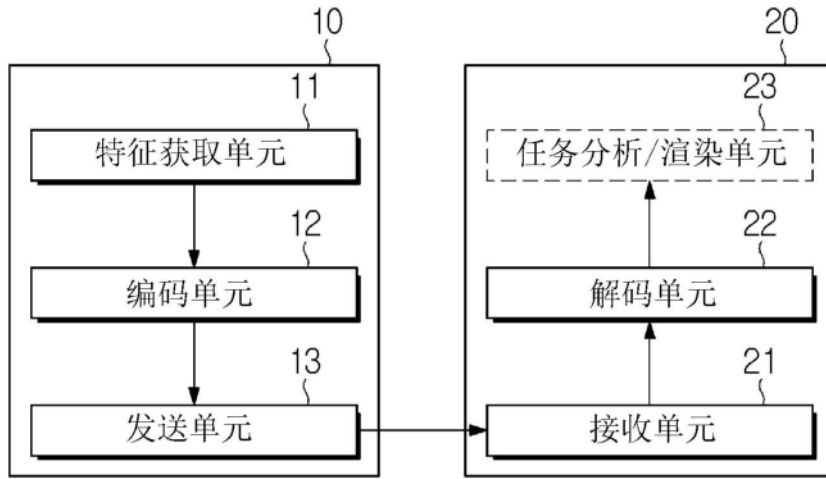


图1

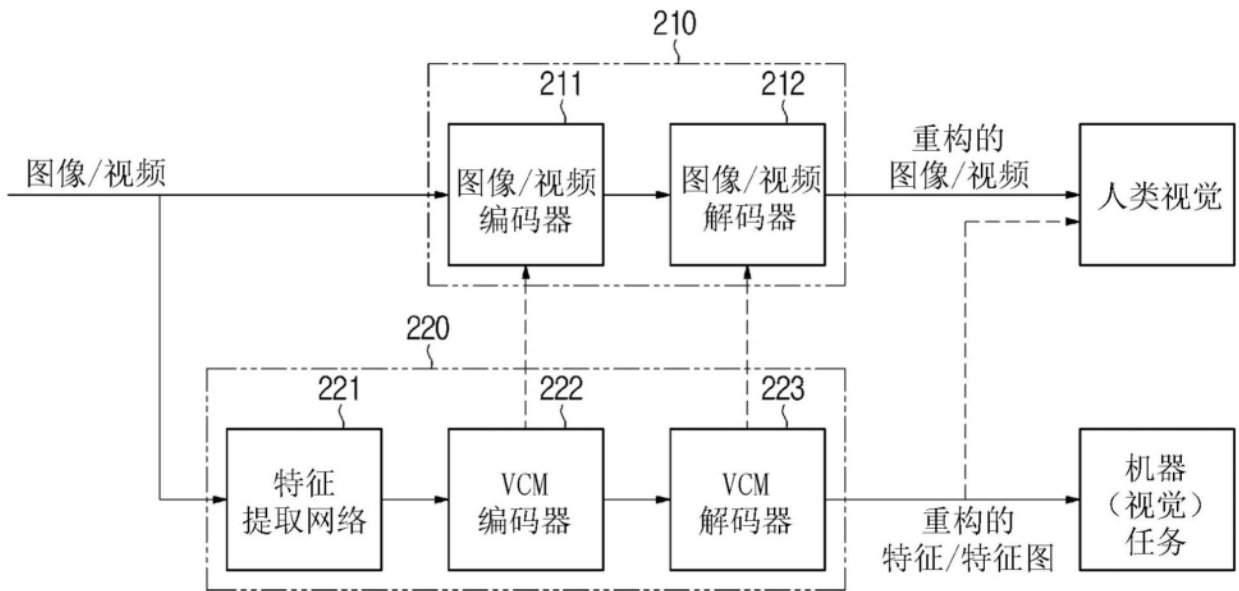


图2

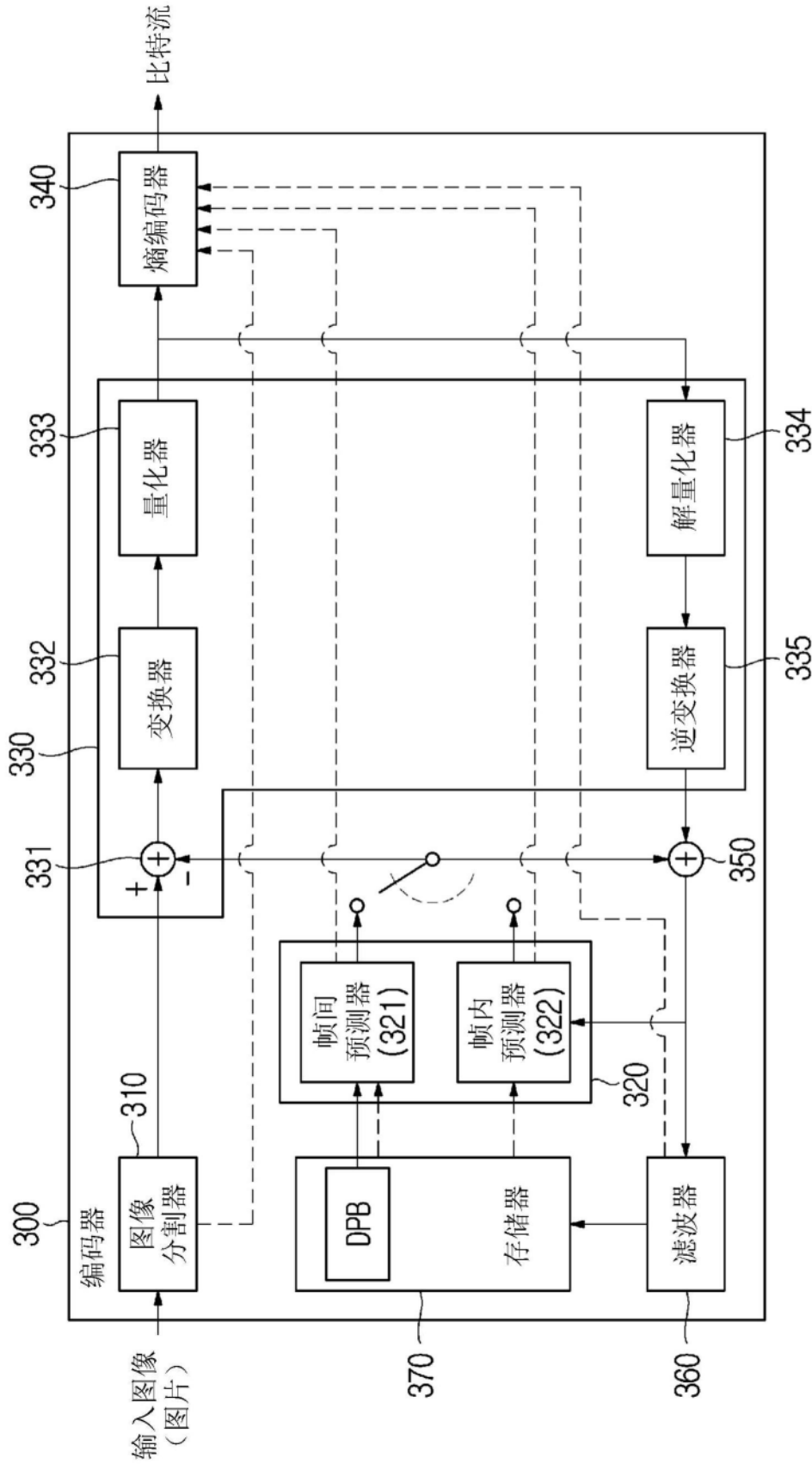


图3

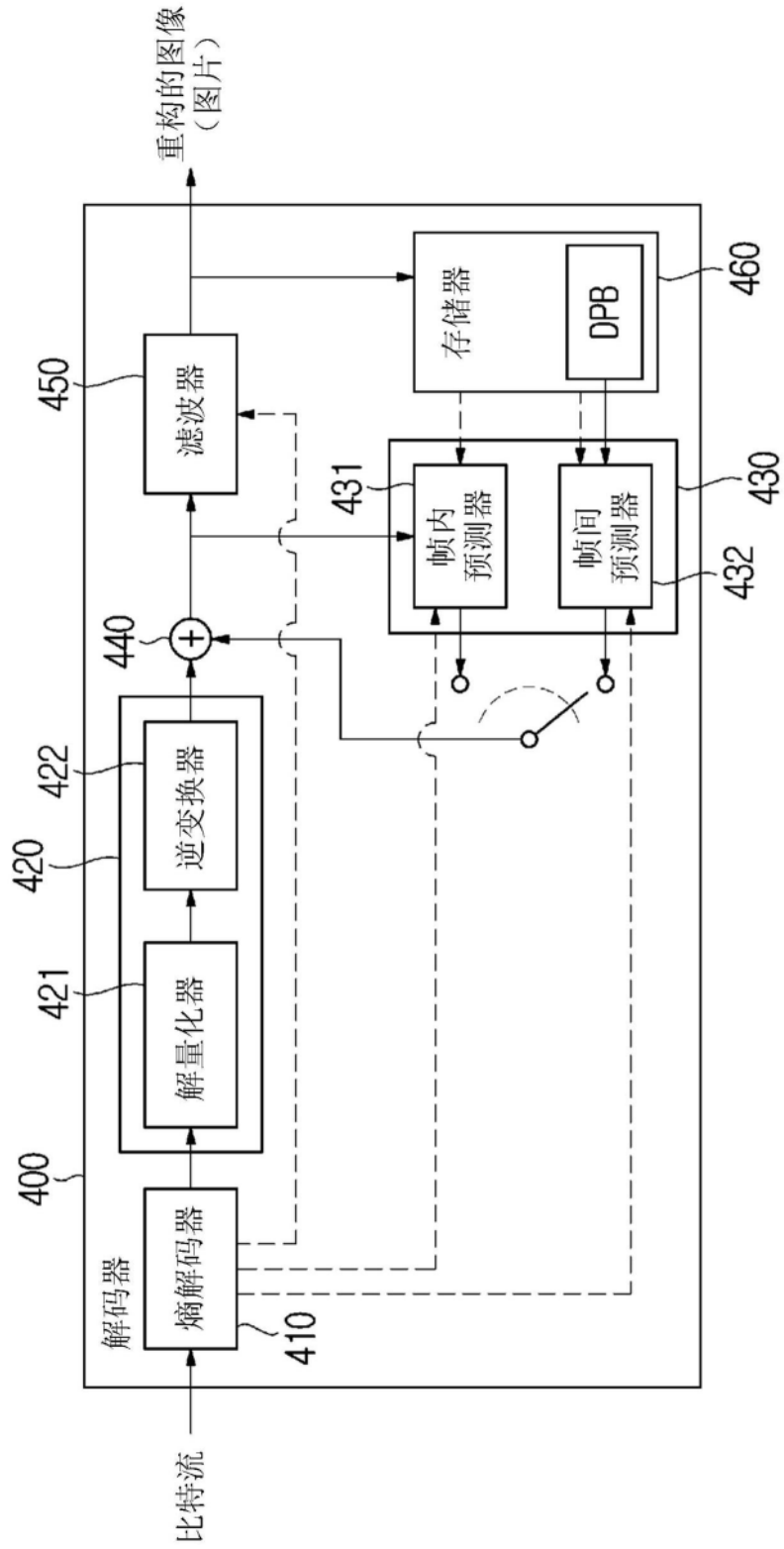


图4

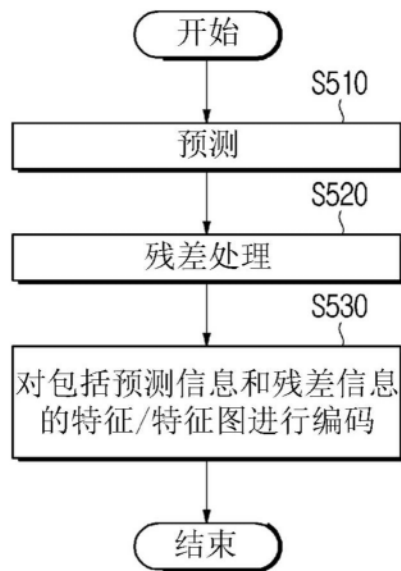


图5

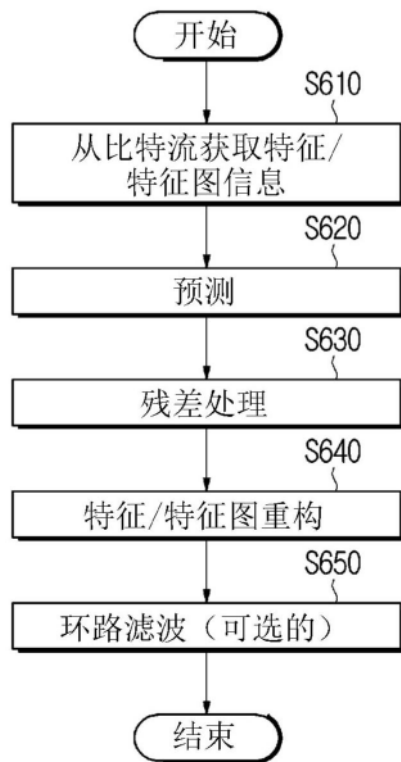


图6

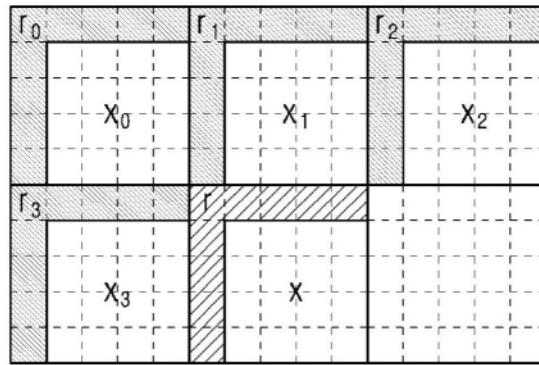


图7

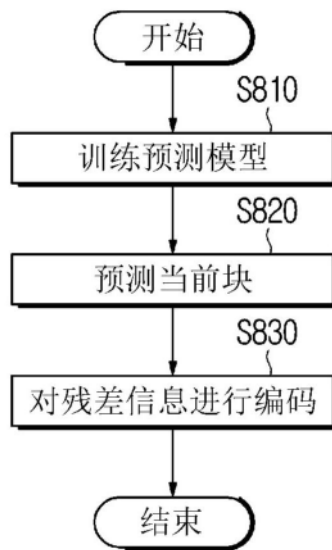


图8

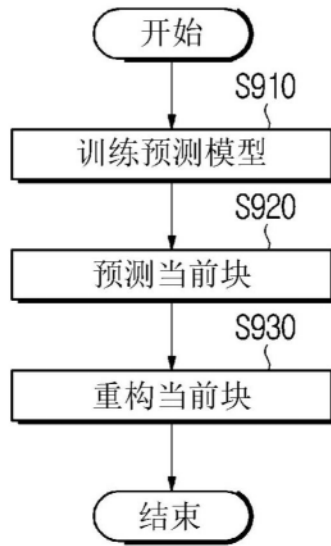


图9

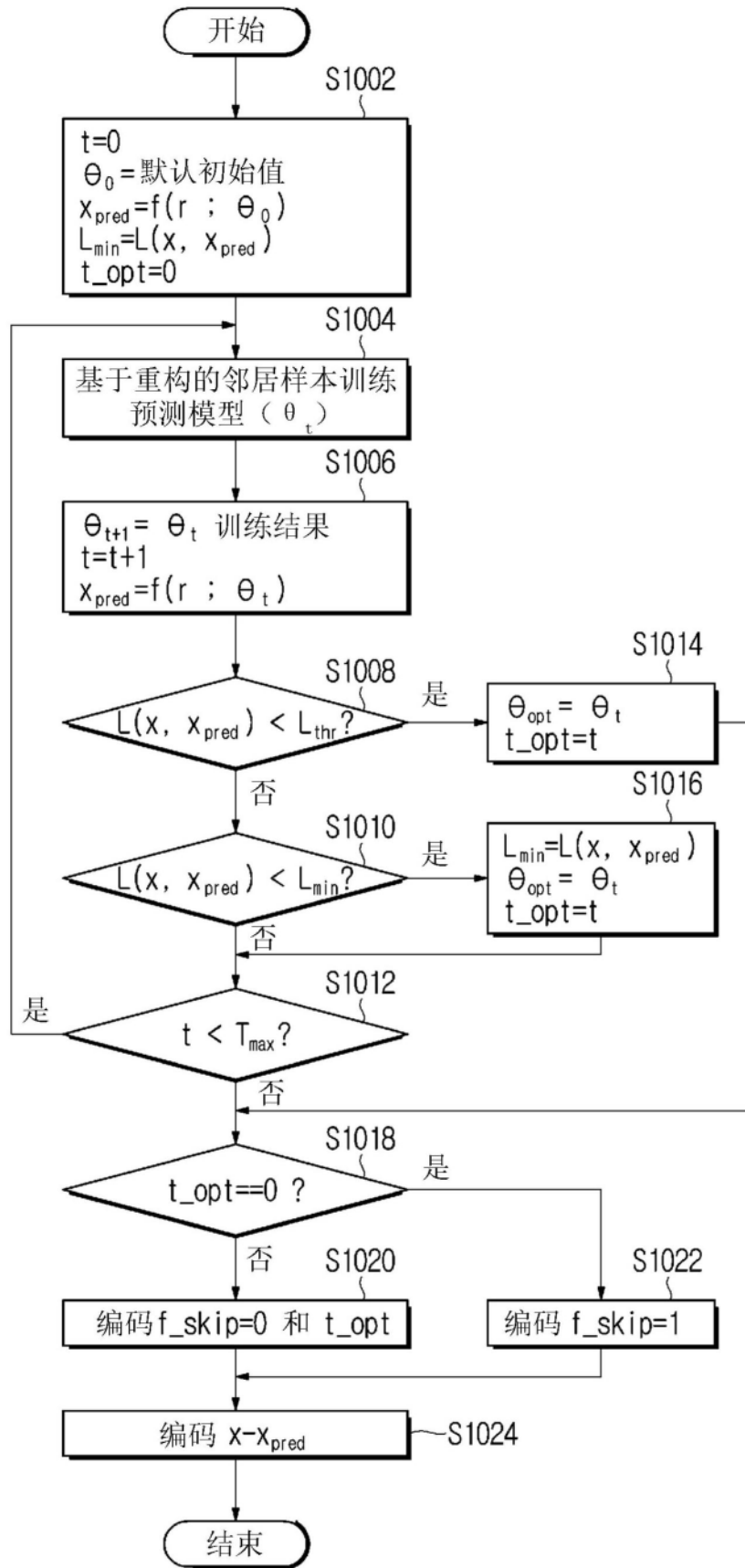


图10

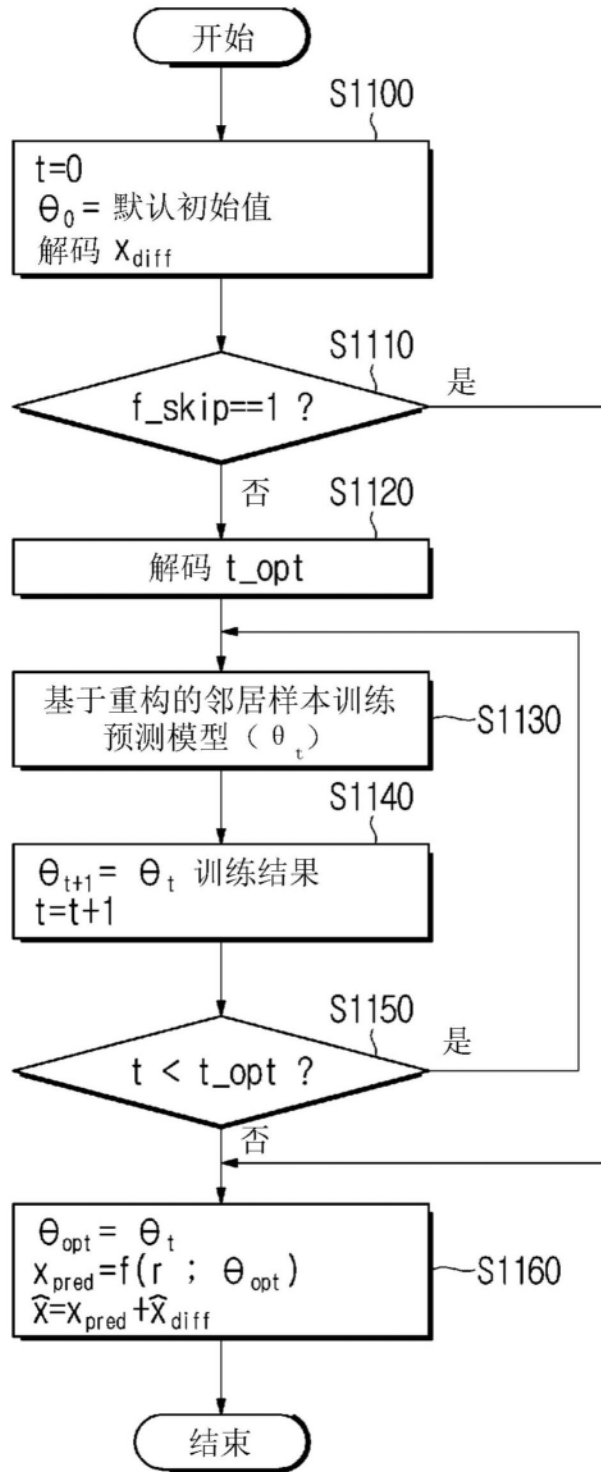


图11

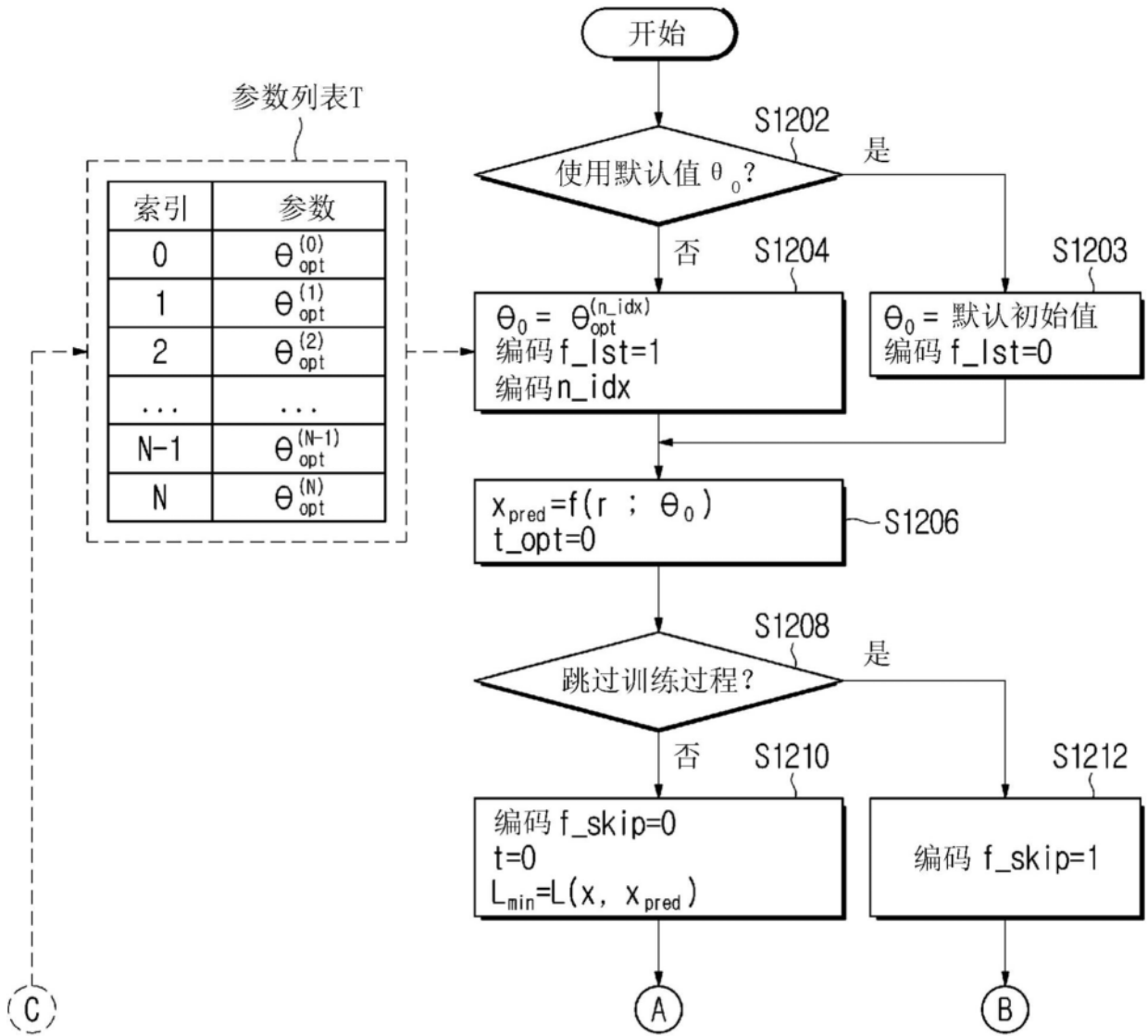


图12a

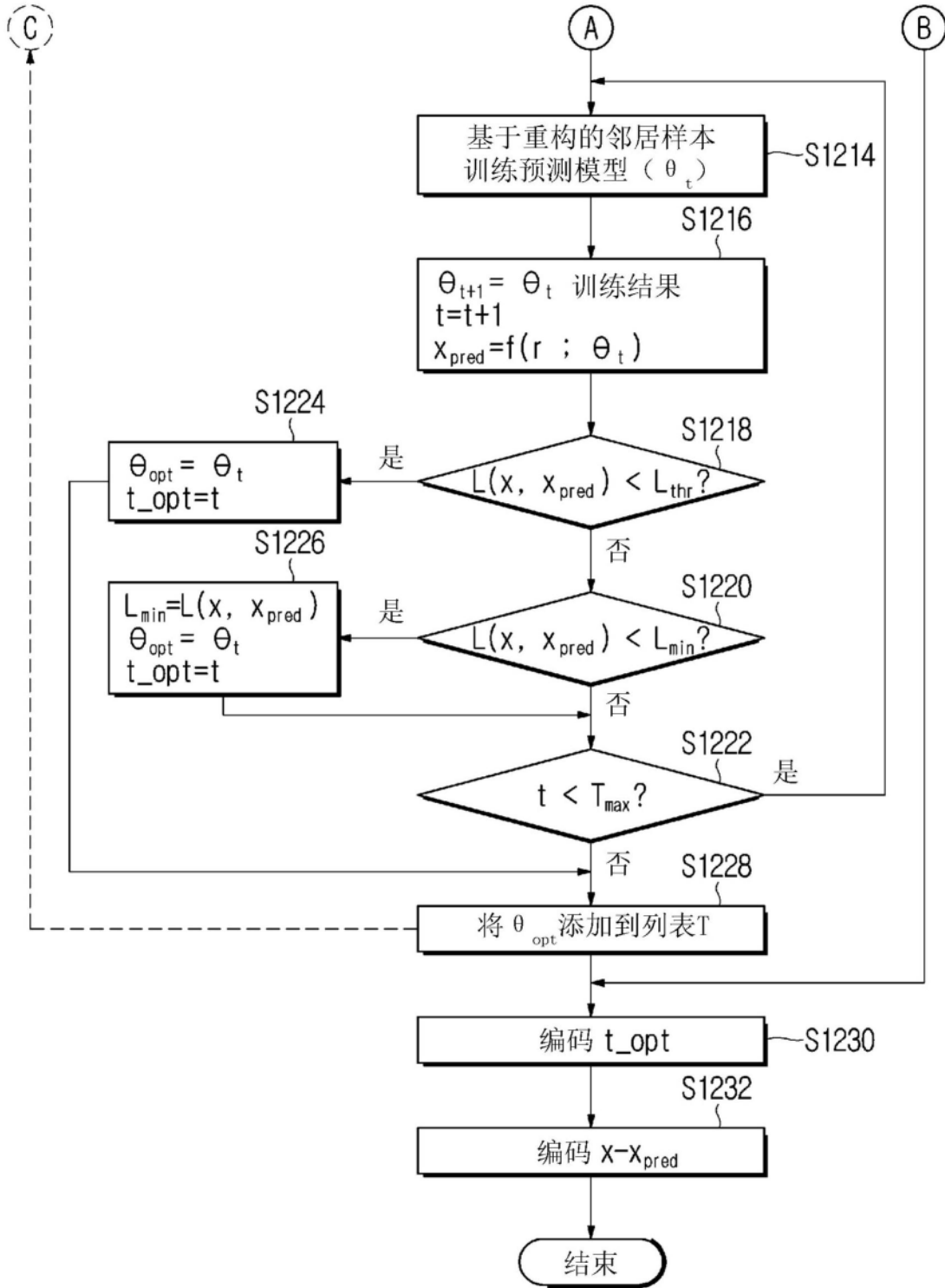


图12b

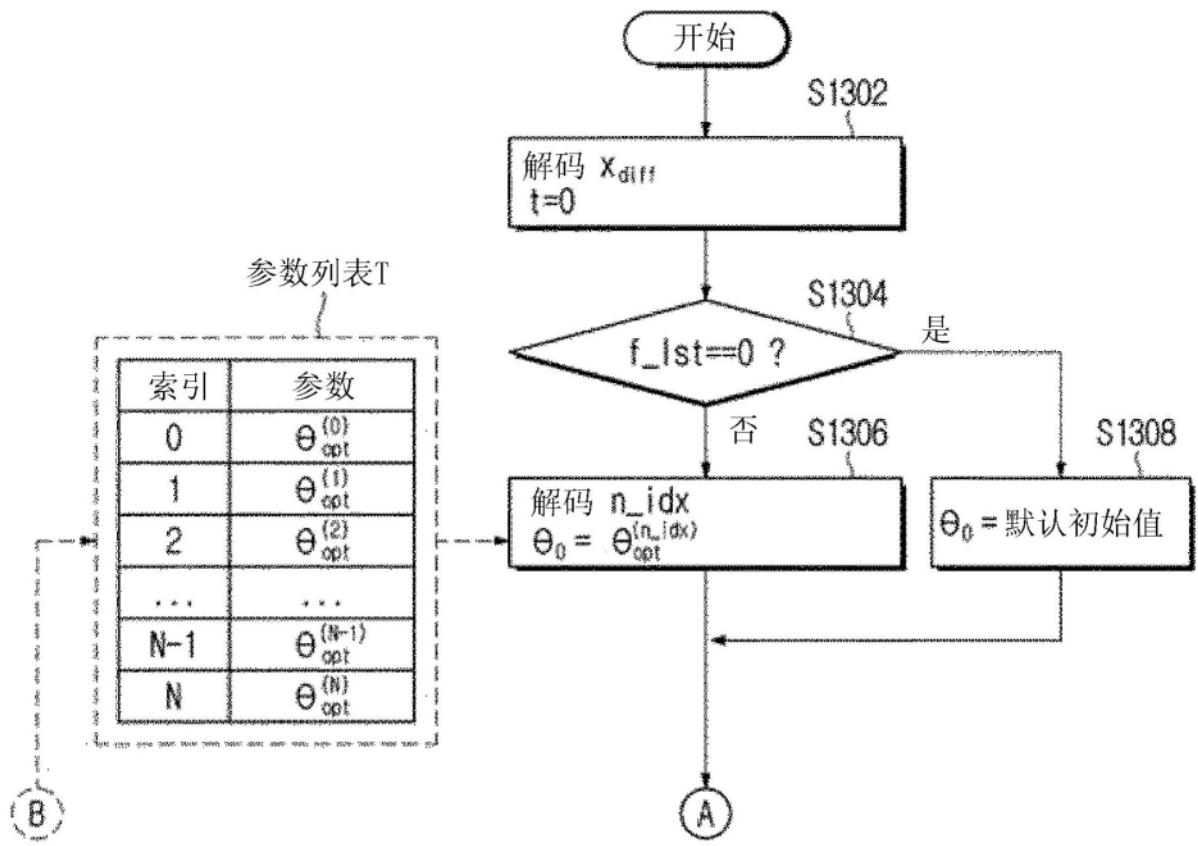


图13a

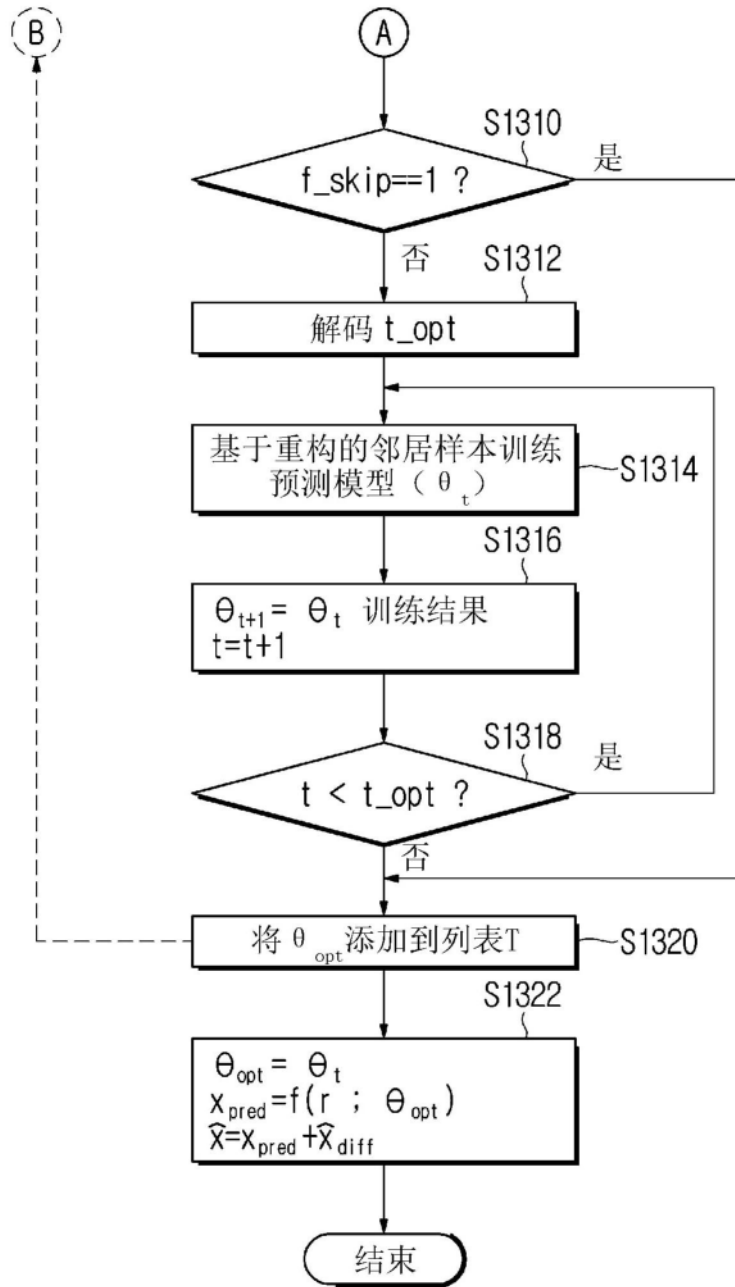


图13b

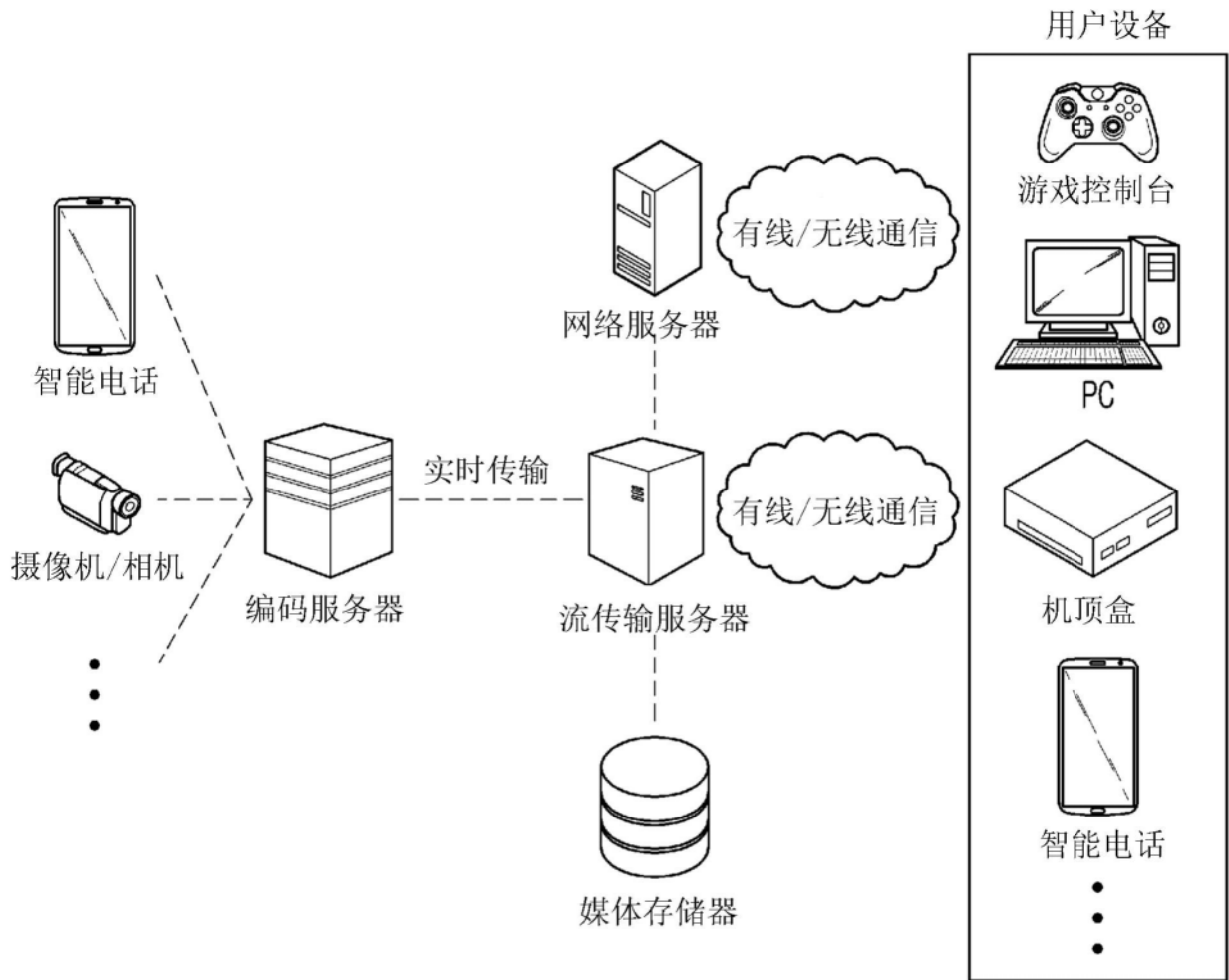


图14

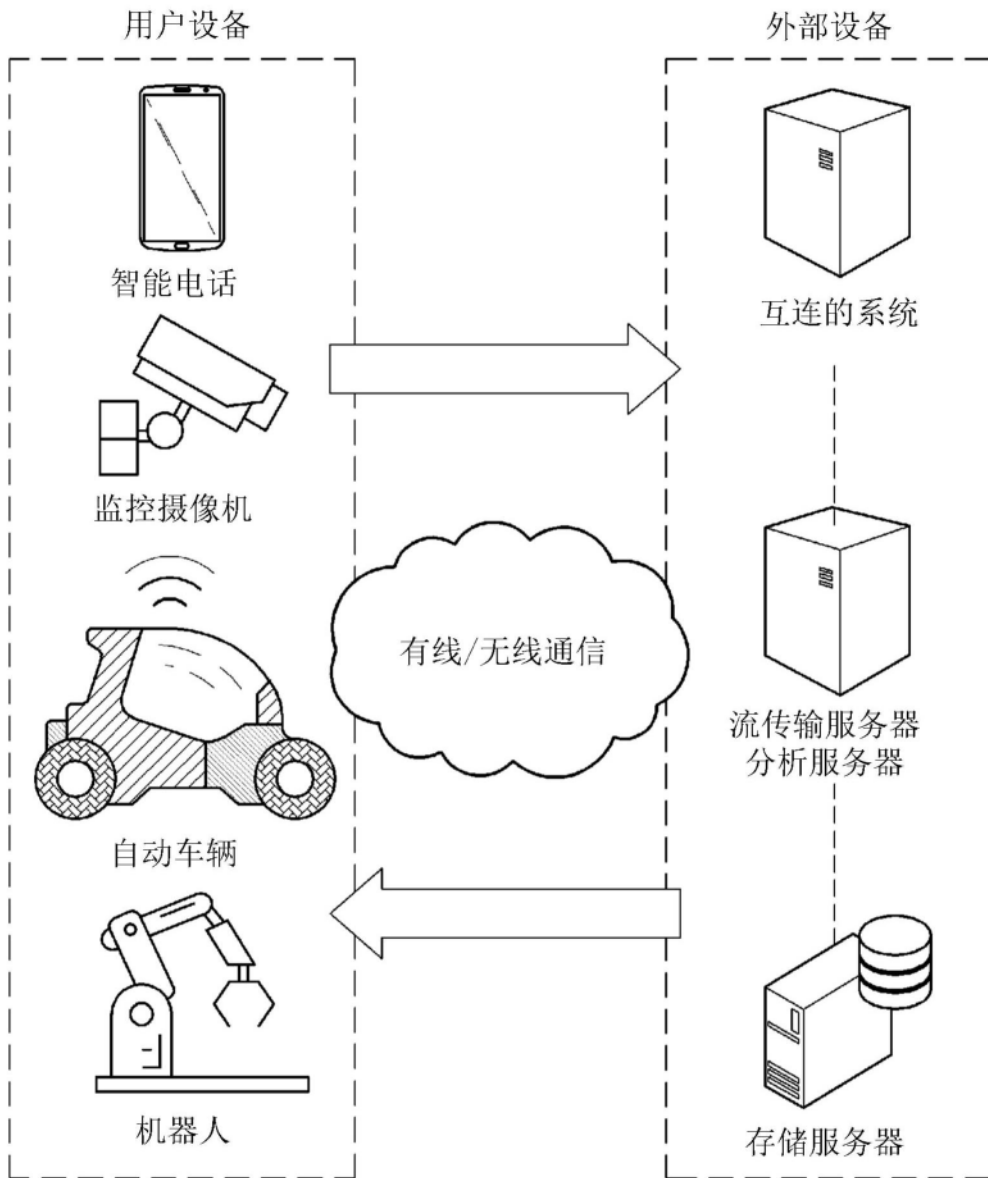


图15