(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 115769575 B
(45)授权公告日 2025.02.07

(21)申请号 202180015923.2
(22)申请日 2021.10.05
(65)同一申请的已公布的文献号
申请公布号 CN 115769575 A
(43)申请公布日 2023.03.07
(30)优先权数据
63/130,249 2020.12.23 US
17/490,967 2021.09.30 US
(85)PCT国际申请进入国家阶段日
2022.08.22
(86)PCT国际申请的申请数据
PCT/US2021/053591 2021.10.05
(87)PCT国际申请的公布数据
WO2022/139935 EN 2022.06.30

(73)专利权人 腾讯美国有限责任公司
地址 美国加利福尼亚州帕洛阿尔托公园大道2747号
(72)发明人 马杜·柏林加色克里什南
山姆如迪亚士万狄·卡胡 赵欣
刘杉
(74)专利代理机构 深圳市深佳知识产权代理事务所(普通合伙) 44285
专利代理师 王曙聘

(51)Int.Cl.
H04N 19/132(2006.01)
(56)对比文件
US 2018091817 A1,2018.03.29
WO 2019185883 A1,2019.10.03
审查员 平蕾

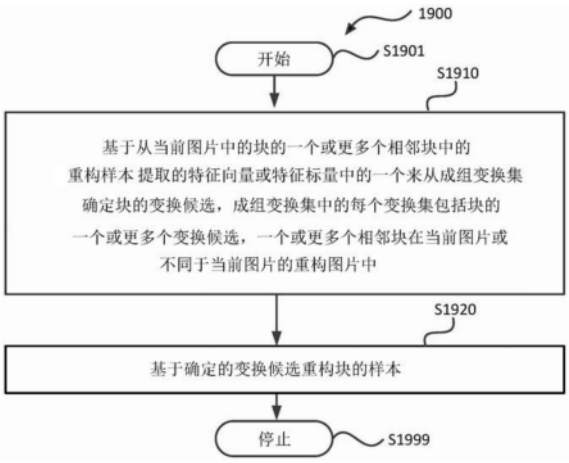
权利要求书2页 说明书33页 附图22页

(54)发明名称

视频解码的方法、计算机装置、设备及存储介质

(57)摘要

本公开内容的各方面提供包括用于视频解码的处理电路的设备和方法。处理电路基于从当前图片的块的一个或更多个相邻块中的重建样本提取的特征向量$\vec{F}$或特征标量S的一个,从成组变换集确定块的变换候选。成组变换集中的每个变换集可以包括块的一个或更多个变换候选。一个或更多个相邻块可以在当前图片中或与当前图片不同的重建图片中。处理电路基于确定的变换候选重建块的样本。处理电路可以基于块的经编码信息中指示的块的预测模式从成组变换集中选择变换集子组,并且基于重建样本从变换集子组确定变换候选。



CN 115769575 B

1.一种用于解码器中的视频解码的方法,其特征在于,所述方法包括:

基于从当前图片中的块的一个或更多个相邻块中的重建样本提取的特征向量$\vec{F}$或特征标量S中的一个,从成组变换集确定所述块的变换候选,所述成组变换集中的每个变换集包括所述块的一个或更多个变换候选,一个或更多个相邻块在所述当前图片中或不同于所述当前图片的重建图片中;以及

基于所确定的变换候选重建所述块的样本;

所述从成组变换集确定所述块的变换候选包括:

基于对所述块的一个或更多个相邻块中的所述重建样本的统计分析,确定所述特征向量$\vec{F}$或所述特征标量S中的一个;

基于所述块的经编码信息中指示的所述块的预测模式,从所述成组变换集中选择变换集子组;以及

执行以下操作中的一个

(i)基于所述特征向量$\vec{F}$或所述特征标量S中的一个,从所述变换集子组确定变换集,

(ii)基于所述特征向量$\vec{F}$或所述特征标量S中的一个,从所述变换集子组确定所述变换候选,或者

(iii)基于所述经编码信息中的索引从所述变换集子组中选择所述变换集,并且基于所述特征向量$\vec{F}$或所述特征标量S中的一个从所选择的变换集确定所述变换候选;

所述特征向量$\vec{F}$或所述特征标量S中的一个是所述特征向量$\vec{F}$;并且确定所述特征向量$\vec{F}$或所述特征标量S中的一个还包括:所述特征向量$\vec{F}$包括分别指示所述块的左侧的相邻列的重建样本的样本值和所述块的上侧的相邻行的重建样本的样本值的变量的联合变化性,所述联合变化性由分别指示所述块的左侧的相邻列的重建样本的样本值和所述块的上侧的相邻行的重建样本的样本值的变量的协方差或二阶矩确定。

2.根据权利要求1所述的方法,其特征在于,所述从所述变换集子组确定所述变换候选包括:

基于从所述块的一个或更多个相邻块中的所述重建样本提取的所述特征向量$\vec{F}$或所述特征标量S中的一个,选择所述变换集子组的一个变换集;以及

基于在所述经编码信息中用信号指示的索引,从所述变换集子组中的所选择的一个变换集确定所述块的变换候选。

3.根据权利要求1所述的方法,其特征在于,所述从所述变换集子组确定所述变换候选包括:

基于在所述经编码信息中用信号指示的索引,选择所述变换集子组中的一个变换集;以及

基于从所述块的一个或更多个相邻块中的所述重建样本提取的所述特征向量$\vec{F}$或所述特征标量S中的一个,从所述变换集子组中的所选择的一个变换集确定所述块的变换候选。

4.根据权利要求1所述的方法,其特征在于,所述从所述变换集子组确定所述变换候选

包括：

基于从所述块的一个或更多个相邻块中的所述重建样本提取的所述特征向量$\vec{F}$或所述特征标量S中的一个，从所述变换集子组隐式地确定所述变换候选。

5.根据权利要求1所述的方法，其特征在于，预先定义分类向量集$\vec{C}$和与所述分类向量集$\vec{C}$相关联的阈值集K_v，

所述方法还包括计算所述变量的联合变化性与从被包括在所述分类向量集$\vec{C}$中的分类向量子集$\vec{C'}$中选择的分类向量之间的距离，并且

所述执行包括执行以下操作中的一个

(i)基于所述距离与从被包括在所述阈值集K_v中的阈值子集K_v'中选择的阈值的比较，从所述变换集子组确定所述变换集，

(ii)基于所述距离与所述阈值的比较，从所述变换集子组确定所述变换候选，或者

(iii)基于所述经编码信息中的索引从所述变换集子组中选择所述变换集，并且基于所述距离与所述阈值的比较从所选择的变换集确定所述变换候选。

6.一种计算机装置，其特征在于，所述计算机装置包括：

一个或更多个计算机可读非暂态存储介质，其被配置成存储计算机程序代码；以及

一个或更多个计算机处理器，其被配置成访问所述计算机程序代码并且按所述计算机程序代码的指示来执行根据权利要求1至5中任一项所述的方法。

7.一种视频解码设备，包括处理电路，其被配置成执行根据权利要求1至5中任一项所述的方法。

8.一种非暂态计算机可读存储介质，其特征在于，所述非暂态计算机可读存储介质存储计算机程序代码，当由至少一个处理器执行所述计算机程序代码时，所述至少一个处理器执行根据权利要求1至5中任一项所述的方法。

9.一种处理视频比特流的方法，其特征在于，所述视频比特流根据权利要求1至5任一项所述的方法进行解码。

视频解码的方法、计算机装置、设备及存储介质

[0001] 引用合并

[0002] 本申请要求2021年9月30日提交的美国专利申请第17/490,967号“METHOD AND APPARATUS FOR VIDEO CODING”的优先权权益,该美国专利申请要求2020年12月23日提交的美国临时申请第63/130,249号的“Feature based transform selection”的优先权权益。在此通过引用将在先申请的全部公开内容全部并入。

技术领域

[0003] 本公开内容描述总体上涉及视频编码/解码的实施方式,并且更特别地涉及一种视频解码的方法、计算机装置、视频解码设备以及非暂态计算机可读存储介质。

背景技术

[0004] 本文提供的背景技术描述是为了总体上呈现本公开内容的上下文的目的。当前署名的发明人的工作,就在该背景技术部分中描述该工作的程度而言,以及在提交时可能不构成现有技术的说明书的各方面,既未明示地也未隐含地被承认为是本公开内容的现有技术。

[0005] 可以使用具有运动补偿的图片间预测来执行视频编码和解码。未压缩数字视频可以包括一系列图片,每个图片具有例如1920×1080个亮度样本和相关色度样本的空间维度。该系列图片可以具有例如每秒60幅图片或60Hz的固定或可变的图片速率(非正式地也称为帧速率)。未压缩视频具有特定的比特率要求。例如,每样本8比特的1080p60 4:2:0视频(60Hz帧速率的1920×1080亮度样本分辨率)需要接近1.5Gbit/s带宽。一小时的这样的视频需要超过600千兆字节(GByte)的存储空间。

[0006] 视频编码和解码的一个目的可以是通过压缩减少输入视频信号中的冗余。压缩可以帮助减少前述带宽或存储空间要求,在一些情况下减少两个数量级或更多。可以采用无损压缩和有损压缩两者及其组合。无损压缩是指可以由压缩的原始信号重建原始信号的精确副本的技术。当使用有损压缩时,重建信号可能与原始信号不相同,但是原始信号与重建信号之间的失真足够小,以使得重建信号对于预期应用有用。在视频的情况下,广泛采用有损压缩。容许的失真量取决于应用;例如,某些消费者流式传输应用的用户可以比电视分发应用的用户容忍更高的失真。可实现的压缩比可以反映:较高的可允许/可容忍的失真可以产生较高的压缩比。

[0007] 视频编码器和解码器可以利用来自若干宽泛类别的技术,包括例如运动补偿、变换、量化和熵编码。

[0008] 视频编解码技术可以包括被称为帧内编码的技术。在帧内编码中,在不参考来自先前重建的参考图片的样本或其他数据的情况下表示样本值。在一些视频编解码器中,将图片在空间上细分为样本块。当所有的样本块都以帧内模式编码时,该图片可以是帧内图片。帧内图片及其派生(例如独立解码器刷新图片)可以用于重置解码器状态,并且因此可以用作经编码视频比特流和视频会话中的第一图片,或用作静止图片。可以使帧内块的样

本经受变换,并且可以在熵编码之前量化变换系数。帧内预测可以是使预变换域的样本值最小的技术。在一些情况下,变换之后的DC值越小,且AC系数越小,在给定量化步长尺寸下表示熵编码之后的块所需的比特越少。

[0009] 传统的帧内编码,例如从例如MPEG-2代编码技术中已知的帧内编码,不使用帧内预测。然而,一些较新的视频压缩技术包括从例如在空间相邻的并且按解码顺序在前的数据块的编码/解码期间获得的周围样本数据和/或元数据进行尝试的技术。这样的技术此后被称为"帧内预测"技术。注意,在至少一些情况下,帧内预测仅使用来自重建中的当前图片而不是来自参考图片的参考数据。

[0010] 可以有许多不同形式的帧内预测。当在给定视频编码技术中可以使用多于一种的这样的技术时,可以以帧内预测模式对所使用的技术进行编码。在某些情况下,模式可以具有子模式和/或参数,并且这些子模式和/或参数可以被单独地编码或被包括在模式码字中。对于给定模式/子模式/参数组合要使用哪些码字可能对通过帧内预测的编码效率增益有影响,并且用于将码字转换成比特流的熵编码技术也是如此。现有技术缺乏与基于块的一个或更多个相邻块的重建样本(例如,特征指示符(例如,块的一个或更多个相邻块的重建样本的特征向量或特征标量))的变换集或变换内核选择方案相关的视频编码技术。

发明内容

[0011] 根据实施方式,提供了一种用于解码器中的视频解码的方法。所述方法包括:基于从当前图片中的块的一个或更多个相邻块中的重建样本提取的特征向量或特征标量S中的一个,从成组变换集确定所述块的变换候选,所述成组变换集中的每个变换集包括所述块的一个或更多个变换候选,一个或更多个相邻块在所述当前图片中或不同于所述当前图片的重建图片中;以及基于所确定的变换候选重建所述块的样本。

[0012] 根据实施方式,提供了一种计算机装置,所述计算机装置包括:一个或更多个计算机可读非暂态存储介质,其被配置成存储计算机程序代码;以及一个或更多个计算机处理器,其被配置成访问所述计算机程序代码并且按所述计算机程序代码的指示来执行上述用于解码器中的视频解码的方法。

[0013] 根据实施方式,提供了一种视频解码设备,包括处理电路,其被配置成执行上述用于解码器中的视频解码的方法。

[0014] 根据实施方式,提供了一种非暂态计算机可读存储介质,所述非暂态计算机可读存储介质存储计算机程序代码,当由至少一个处理器执行所述计算机程序代码时,所述至少一个处理器执行上述用于解码器中的视频解码的方法。

[0015] 本发明的视频解码的方法、计算机装置、视频解码设备以及非暂态计算机可读存储介质提供了一种与基于块的一个或更多个相邻块的重建样本(例如,特征指示符(例如,块的一个或更多个相邻块的重建样本的特征向量或特征标量))的变换集或变换内核选择方案相关的视频编码技术。

附图说明

[0016] 根据以下详细描述和附图,所公开的主题的其他特征、性质和各种优点将变得更加明显,在附图中:

[0017] 图1A是帧内预测模式的示例性子集的示意性图示;

[0018] 图1B是示例性帧内预测方向的图示;

[0019] 图2是一个示例中的当前块及其周围空间合并候选的示意性图示;

[0020] 图3是根据实施方式的通信系统的简化框图的示意性图示;

[0021] 图4是根据另一实施方式的通信系统的简化框图的示意性图示;

[0022] 图5是根据实施方式的解码器的简化框图的示意性图示;

[0023] 图6是根据实施方式的编码器的简化框图的示意性图示;

[0024] 图7示出了根据另一实施方式的编码器的框图;

[0025] 图8示出了根据另一实施方式的解码器的框图;

[0026] 图9示出了根据本公开内容的实施方式的用于编码块的标称模式的示例。

[0027] 图10示出了根据本公开内容的方面的非定向平滑帧内预测的示例。

[0028] 图11示出了根据本公开内容的实施方式的基于递归滤波的帧内预测器的示例。

[0029] 图12示出了根据本公开内容的实施方式的用于编码块的多个参考线的示例。

[0030] 图13示出了根据本公开内容的实施方式的初级变换基函数的示例。

[0031] 图14A示出了根据本公开内容的实施方式的基于变换块大小和预测模式的各种变换内核的可用性的示例性相关性。

[0032] 图14B示出了根据本公开内容的实施方式的基于帧内预测模式的示例性变换类型选择。

[0033] 图15至图16示出了根据本公开内容的实施方式的分别使用16×64变换和16×48变换的两个变换编码过程的示例。

[0034] 图17A至图17D示出了根据本公开内容的实施方式的针对帧内预测模式观察到的(以灰度表示的)示例性残差图案。

[0035] 图18示出了根据本公开内容的实施方式的块的示例性空间相邻样本。

[0036] 图19示出了概述根据本公开内容的实施方式的过程的流程图。

[0037] 图20是根据实施方式的计算机系统的示意性图示。

具体实施方式

[0038] 帧内预测的特定模式被引入H.264,在H.265中被改进,并且在诸如联合探索模型(joint exploration model,JEM)、通用视频编码(versatile video coding,VVC)和基准集(benchmark set,BMS)的较新的编码技术中被进一步改进。可以使用属于已经可用的样本的相邻样本值来形成预测器块。根据方向将相邻样本的样本值复制到预测器块中。对使用方向的参考可以在比特流中编码或可以自身预测。

[0039] 参照图1A,在右下方描绘了从H.265的33个可能的预测器方向(对应于35个帧内模式的33个角模式)中已知的九个预测器方向的子集。箭头会聚的点(101)表示被预测的样本。箭头表示样本被预测的方向。例如,箭头(102)指示样本(101)是从与水平成45度角的右上的一个或更多个样本预测的。类似地,箭头(103)指示样本(101)是从样本(101)的与水平成22.5度角的左下方的一个或更多个样本来预测的。

[0040] 仍然参照图1A,在左上方,描绘了4×4个样本的方块(104)(由虚粗体线指示)。方块(104)包括16个样本,每个样本用“S”、其在Y维度上的位置(例如,行索引)和其在X维度上

的位置(例如,列索引)来标记。例如,样本S21是Y维度上的第二样本(从顶部)和X维度上的第一样本(从左侧)。类似地,样本S44是块(104)中在Y维度和X维度上的第四样本。由于块的大小是4×4个样本,所以S44在右下。进一步显示了遵循类似编号方案的参考样本。参考样本用R、其相对于块(104)的Y位置(例如,行索引)和X位置(列索引)来标记。在H.264和H.265两者中,预测样本与重建中的块相邻;因此不需要使用负值。

[0041] 通过复制来自由用信号指示的预测方向所划定的相邻样本的参考样本值,可以进行帧内图片预测。例如,假设经编码视频比特流包括以下信令:对于该块,该信令指示与箭头(102)一致的预测方向,也就是说,从与水平成45度角的右上方的一个或更多个预测样本来对样本进行预测。在这种情况下,样本S41、样本S32、样本S23和样本S14是从相同的参考样本R05预测的。然后从参考样本R08预测样本S44。

[0042] 在某些情况下,多个参考样本的值可以例如通过插值来组合,以便计算参考样本;尤其是当方向不能以45度均匀地划分时。

[0043] 随着视频编码技术的发展,可能的方向的数目也在增加。在H.264(2003年)中,可以表示九个不同的方向。在H.265(2013年)中增加到33,并且在公开时JEM/VVC/BMS可以支持多达65个方向。已经进行了实验来识别最可能的方向,并且熵编码中的某些技术被用于以少量的比特表示那些可能的方向,从而对于较不可能的方向接受某些惩罚。此外,有时可以从在相邻的已解码的块中使用的相邻方向来预测方向本身。

[0044] 图1B示出了描绘根据JEM的65个帧内预测方向以示出预测方向的数目随时间的增加的示意图(180)。

[0045] 经编码视频比特流中表示方向的帧内预测方向比特的映射可以随视频编码技术不同而不同;并且范围可以例如从预测方向到帧内预测模式的简单直接映射、到码字、到涉及最可能模式的复杂自适应方案以及类似技术。然而,在所有情况下,可以存在统计上与某些其他方向相比在视频内容中更不可能出现的某些方向。由于视频压缩的目标是减少冗余,所以在良好工作的视频编码技术中,与更可能的方向相比,那些较不可能的方向将由更多的比特来表示。

[0046] 运动补偿可以是有损压缩技术,并且可以涉及这样的技术:来自先前重建的图片或其部分(参考图片)的样本数据块在由运动向量(此后称为MV)指示的方向上进行空间移位之后,被用于预测新重建的图片或图片部片。在一些情况下,参考图片可以与当前正在重建的图片相同。MV可以具有两个维度X和Y,或三个维度,第三个维度是使用中的参考图片的指示(后者可以间接地是时间维度)。

[0047] 在一些视频压缩技术中,可应用于样本数据的某个区域的MV可以从其他MV预测,例如从与空间上邻近重建中的区域并且在解码顺序中在该MV之前的样本数据的另一区域相关的MV。这样做可以显著减少编码MV所需的数据量,从而消除冗余并增加压缩。MV预测可以有效地工作,例如,因为当对从摄像机得到的输入视频信号(称为自然视频)进行编码时,存在比单个MV所适用的区域更大的区域沿相似方向移动的统计可能性,并且因此在某些情况下可以使用从相邻区域的MV得到的相似运动向量来预测。这导致针对给定区域所发现的MV与从周围MV预测的MV相似或相同,并且转而可以在熵编码之后,用比直接编码MV将使用的比特数更少的比特数来表示。在一些情况下,MV预测可以是从原始信号(即,样本流)得出的信号(即,MV)的无损压缩的示例。在其他情况下,MV预测本身可能是有损的,例如,因为当

从若干周围MV计算预测器时的舍入误差。

[0048] H.265/HEVC(ITU-T Rec.H.265,“High Efficiency Video Coding”,2016年12月)中描述了各种MV预测机制。在H.265提供的许多MV预测机制中,本文描述的是一种此后被称为“空间合并”的技术。

[0049] 参照图2,当前块(201)包括在运动搜索过程期间由编码器发现能够从已经空间移位的相同大小的先前块预测的样本。代替直接编码MV,可以使用与表示为A0、A1和B0、B1、B2(分别为202至206)的五个周围样本中的任一个相关联的MV,从与一个或更多个参考图片相关联的元数据,例如从最近的(按解码顺序)参考图片来得出MV。在H.265中,MV预测可以使用来自相邻块正在使用的相同参考图片的预测器。

[0050] 图3示出了根据本公开内容的实施方式的通信系统(300)的简化框图。通信系统(300)包括多个终端设备,这些终端设备可以经由例如网络(350)彼此通信。例如,通信系统(300)包括经由网络(350)互连的第一对终端设备(310)和(320)。在图3的示例中,第一对终端设备(310)和(320)执行数据的单向传输。例如,终端设备(310)可以对视频数据(例如,由终端设备(310)捕获的视频图片的流)进行编码,以经由网络(350)传输到另一终端设备(320)。编码的视频数据可以以一个或更多个经编码视频比特流的形式被发送。终端设备(320)可以从网络(350)接收经编码视频数据,解码该经编码视频数据以恢复视频图片,并且根据所恢复的视频数据显示视频图片。单向数据传输在媒体服务应用等中可能是常见的。

[0051] 在另一示例中,通信系统(300)包括第二对终端设备(330)和(340),其执行可能例如在视频会议期间发生的经编码视频数据的双向传输。对于数据的双向传输,在示例中,终端设备(330)和(340)中的每个终端设备可以对视频数据(例如,由终端设备捕获的视频图片的流)进行编码以经由网络(350)传输到终端设备(330)和(340)中的另一终端设备。终端设备(330)和(340)中的每个终端设备还可以接收由终端设备(330)和(340)中的另一终端设备传输的经编码视频数据,并且可以解码经编码视频数据以恢复视频图片,并且可以根据所恢复的视频数据在可访问显示设备处显示视频图片。

[0052] 在图3的示例中,终端设备(310)、(320)、(330)和(40)可以被图示为服务器、个人计算机和智能电话,但是本公开内容的原理可以不限于此。本公开内容的实施方式可应用于膝上型计算机、平板计算机、媒体播放器和/或专用视频会议设备。网络(350)表示在终端设备(310)、(320)、(330)和(340)之间传送经编码视频数据的任意数量的网络,包括例如有线(连线)和/或无线通信网络。通信网络(350)可以在电路交换和/或分组交换信道中交换数据。代表性网络包括电信网络、局域网、广域网和/或因特网。出于当前讨论的目的,除非以下在本文中说明,否则网络(350)的架构和拓扑结构对于本公开内容的操作可能是不重要的。

[0053] 图4示出作为所公开的主题的应用的示例的视频编码器和视频解码器在流式传输环境中的布置。所公开的主题可以同样适用于其他具有视频功能的应用,包括例如视频会议、数字TV、将经压缩视频存储在包括CD、DVD、记忆棒等的数字媒介上。

[0054] 流式传输系统可以包括捕获子系统(413),该捕获子系统可以包括视频源(401),例如数字摄像机,该视频源创建例如未压缩的视频图片流(402)。在示例中,视频图片流(402)包括由数字摄像机拍摄的样本。被描绘为粗线以强调与编码的视频数据(404)(或经

编码视频比特流)相比时的高数据量的视频图片流(402)可以由包括耦合至视频源(401)的视频编码器(403)的电子设备(420)来处理。视频编码器(403)可以包括硬件、软件或其组合,以实施或实现如下面更详细描述的所公开的主题的各方面。被描绘为细线以强调与视频图片流(402)相比时的较低的数据量的编码的视频数据(404)(或编码的视频比特流(404))可以存储在流式传输服务器(405)上以供将来使用。一个或更多个流式传输客户端子系统,例如图4中的客户端子系统(406)和(408),可以访问流式传输服务器(405)以检索编码的视频数据(404)的副本(407)和(409)。客户端子系统(406)可以包括例如电子设备(430)中的视频解码器(410)。视频解码器(410)对编码的视频数据的输入副本(407)进行解码,并且创建可以在显示器(412)(例如,显示屏)或其他呈现设备(未描绘)上呈现的视频图片(411)的输出流。在一些流式传输系统中,可以根据某些视频编码/压缩标准对编码的视频数据(404)、(407)以及(409)(例如,视频比特流)进行编码。这些标准的示例包括ITU-T建议H.265。在示例中,正在开发中的视频编码标准被非正式地称为通用视频编码(Versatile Video Coding,VVC)。所公开的主题可以在VVC的上下文中使用。

[0055] 注意,电子设备(420)和(430)可以包括其他部件(未示出)。例如,电子设备(420)可以包括视频解码器(未示出),并且电子设备(430)也可以包括视频编码器(未示出)。

[0056] 图5示出了根据本公开内容的实施方式的视频解码器(510)的框图。视频解码器(510)可以被包括在电子设备(530)中。电子设备(530)可以包括接收器(531)(例如,接收电路)。视频解码器(510)可以用来代替图4示例中的视频解码器(410)。

[0057] 接收器(531)可以接收要由视频解码器(510)解码的一个或更多个经编码视频序列;在相同或另一实施方式中,每次一个经编码视频序列,其中每个经编码视频序列的解码独立于其他经编码视频序列。可以从信道(501)接收经编码视频序列,该信道可以是到存储编码的视频数据的存储装置的硬件/软件链路。接收器(531)可以接收具有其他数据的编码的视频数据,例如经编码音频数据和/或辅助数据流,其可以被转发到它们各自的使用实体(未描绘)。接收器(531)可以将经编码视频序列与其他数据分开。为了对抗网络抖动,缓冲存储器(515)可以耦合在接收器(531)与熵解码器/解析器(520)(此后称为“解析器(520)”)之间。在某些应用中,缓冲存储器(515)是视频解码器(510)的一部分。在其他实施方式中,缓冲存储器可以在视频解码器(510)外部(未描绘)。在其他实施方式中,在视频解码器(510)外部可以有缓冲存储器(未描绘),例如以对抗网络抖动,并且另外在视频解码器(510)内部可以有另一缓冲存储器(515),例如以处理播放定时。当接收器(531)从具有足够带宽和可控性的存储/转发设备或从等同步网络接收数据时,可能不需要缓冲存储器(515),或者缓冲存储器可以很小。为了在诸如因特网的尽力而为分组网络上使用,可能需要缓冲存储器(515),其可以相对较大,并且可以有利地具有自适应大小,并且可以至少部分地在视频解码器(510)外部的操作系统或类似元件(未描绘)中实现。

[0058] 视频解码器(510)可以包括解析器(520)以从经编码视频序列重建符号(521)。这些符号的类别包括用于管理视频解码器(510)的操作的信息,以及潜在地包括用于控制呈现设备的信息,所述呈现设备是诸如不是电子设备(530)的组成部分但是可以耦合至电子设备(530)的呈现设备(512)(例如,显示屏),如图5所示。用于呈现设备的控制信息可以是补充增强信息(SEI消息)或视频可用性信息(Video Usability Information,VUI)参数集片段(未描绘)的形式。解析器(520)可以解析/熵解码所接收的经编码视频序列。经编码视

频序列的编码可以根据视频编码技术或标准,并且可以遵循各种原理,包括可变长度编码、霍夫曼编码、具有或不具有上下文敏感的算术编码等。解析器(520)可以基于与组相对应的至少一个参数、从经编码视频序列中提取用于视频解码器中的像素的子组中的至少一个的子组参数的集合。子组可以包括图片组(Groups of Picture,GOP)、图片、图块、切片、宏块、编码单元(Coding Unit,CU)、块、变换单元(Transform Unit,TU)、预测单元(Prediction Unit,PU)等。解析器(520)还可以从经编码视频序列中提取诸如变换系数、量化器参数值、运动向量等的信息。

[0059] 解析器(520)可以对从缓冲存储器(515)接收的视频序列执行熵解码/解析操作,以便创建符号(521)。

[0060] 根据经编码视频图片或其部分的类型(例如:帧间图片和帧内图片、帧间块和帧内块)和其他因素,符号(521)的重建可以涉及多个不同的单元。涉及到哪些单元以及如何涉及这些单元能够由被解析器(520)从经编码视频序列解析的子组控制信息来控制。为了清楚起见,没有描绘在解析器(520)与下面的多个单元之间的这样的子组控制信息的流。

[0061] 除了已经提到的功能块之外,视频解码器(510)可以在概念上细分为多个功能单元,如下所述。在商业约束下操作的实际实现中,这些单元中的许多彼此紧密地交互,并且可以至少部分地集成到彼此中。然而,为了描述所公开的主题的目的,以下概念上细分成功能单元是合适的。

[0062] 第一单元是缩放器/逆变换单元(551)。缩放器/逆变换单元(551)从解析器(520)接收量化变换系数以及控制信息作为符号(521),该控制信息包括要使用何种变换、块大小、量化因子、量化缩放矩阵等。缩放器/逆变换单元(551)可以输出包括可以被输入到聚合器(555)中的样本值的块。

[0063] 在一些情况下,缩放器/逆变换(551)的输出样本可以涉及帧内编码块;即:不使用来自先前重建的图片的预测性信息,而是可以使用来自当前图片的先前重建部分的预测性信息的块。这样的预测性信息可以由帧内图片预测单元(552)提供。在一些情况下,帧内图片预测单元(552)使用从当前图片缓冲器(558)获取的周围的已经重建的信息来生成与重建中的块相同大小和形状的块。当前图片缓冲器(558)缓冲例如部分重建的当前图片和/或完全重建的当前图片。在一些情况下,聚合器(555)基于每个样本将帧内预测单元(552)已经生成的预测信息添加到由缩放器/逆变换单元(551)提供的输出样本信息。

[0064] 在其他情况下,缩放器/逆变换单元(551)的输出样本可以与经帧间编码的并且可能经运动补偿的块有关。在这样的情况下,运动补偿预测单元(553)可以访问参考图片存储器(557)以获取用于预测的样本。在根据与块相关的符号(521)对所获取的样本进行运动补偿之后,这些样本可以由聚合器(555)添加到缩放器/逆变换单元(551)的输出(在这种情况下称为残差样本或残差信号),以便生成输出样本信息。运动补偿预测单元(553)从其获取预测样本的参考图片存储器(557)内的地址可以由运动向量控制,所述运动向量可以以可以具有例如X、Y和参考图片分量的符号(521)的形式被运动补偿预测单元(553)使用。运动补偿还可以包括当子样本精确运动向量在使用中时从参考图片存储器(557)获取的样本值的插值、运动向量预测机制等。

[0065] 聚合器(555)的输出样本可以在环路滤波器单元(556)中经受各种环路滤波技术。视频压缩技术可以包括环路内滤波器技术,其由被包括在经编码视频序列(也称为经编码

视频比特流)中的参数控制,并且作为来自解析器(520)的符号(521)可用于环路滤波器单元(556),但是也可以响应于在对经编码图片或经编码视频序列的先前(按解码顺序)部分进行解码期间获得的元信息,以及响应于先前重建和环路滤波的样本值。

[0066] 环路滤波器单元(556)的输出可以是样本流,其可以被输出到呈现设备(512)以及被存储在参考图片存储器(557)中以供在将来的帧间图片预测中使用。

[0067] 某些经编码图片一旦被完全重建,就可以被用作用于未来预测的参考图片。例如,一旦对应于当前图片的经编码图片被完全重建并且经编码图片已经被识别为参考图片(例如,通过解析器(520)),则当前图片缓冲器(558)可以变为参考图片存储器(557)的一部分,并且新的当前图片缓冲器可以在开始重建随后的经编码图片之前被重新分配。

[0068] 视频解码器(510)可以根据标准(例如ITU-T建议H.265)中的预定视频压缩技术来执行解码操作。在经编码视频序列遵守视频压缩技术或标准的语法及视频压缩技术或标准中所记载的配置文件两者的意义上,经编码视频序列可以符合由所使用的视频压缩技术或标准指定的语法。具体地,配置文件可以从视频压缩技术或标准中可用的所有工具中选择某些工具作为按照该配置文件可用的仅有工具。合规性所必需的还有,经编码视频序列的复杂度在由视频压缩技术或标准的等级所限定的范围内。在一些情况下,等级限制最大图片大小、最大帧速率、最大重建采样率(例如以每秒百万样本来测量)、最大参考图片大小等。在一些情况下,由等级设置的限制可以通过假设参考解码器(Hypothetical Reference Decoder,HRD)规范和在经编码视频序列中用信号指示的用于HRD缓冲器管理的元数据来进一步限制。

[0069] 在实施方式中,接收器(531)可以接收具有编码的视频的附加(冗余)数据。附加数据可以被包括作为经编码视频序列的一部分。附加数据可以由视频解码器(510)使用以适当地解码数据和/或更准确地重建原始视频数据。附加数据可以是例如时间、空间或信噪比(signal noise ratio,SNR)增强层、冗余切片、冗余图片、前向纠错码等的形式。

[0070] 图6示出了根据本公开内容的实施方式的视频编码器(603)的框图。视频编码器(603)被包括在电子设备(620)中。电子设备(620)包括发送器(640)(例如,发送电路)。视频编码器(603)可以用来代替图4的示例中的视频编码器(403)。

[0071] 视频编码器(603)可以从视频源(601)(其不是图6的示例中的电子设备(620)的一部分)接收视频样本,该视频源可以捕获要由视频编码器(603)编码的视频图像。在另一示例中,视频源(601)是电子设备(620)的一部分。

[0072] 视频源(601)可以以数字视频样本流的形式提供要由视频编码器(603)编码的源视频序列,该数字视频样本流可以具有任何适当的位深度(例如:8位、10位、12位、……)、任何颜色空间(例如,BT.601Y CrCB、RGB、……)以及任何适当的采样结构(例如Y CrCb 4:2:0、Y CrCb 4:4:4)。在媒体服务系统中,视频源(601)可以是存储预先准备的视频的存储设备。在视频会议系统中,视频源(601)可以是捕获本地图像信息作为视频序列的摄像机。视频数据可以作为在顺序观看时赋予运动的多个单独图片来提供。图片本身可以被组织为像素的空间阵列,其中,取决于使用中的采样结构、颜色空间等,每个像素可以包括一个或更多个样本。本领域技术人员可以容易地理解像素与样本之间的关系。下面的描述集中在样本上。

[0073] 根据实施方式,视频编码器(603)可以实时地或在应用所需的任何其他时间约束

下将源视频序列的图片编码和压缩为经编码视频序列(643)。强制实行适当的编码速度是控制器(650)的一个功能。在一些实施方式中,控制器(650)控制如下所述的其他功能单元,并且在功能上耦合至其他功能单元。为了清楚起见,没有描绘该耦合。由控制器(650)设置的参数可以包括速率控制相关参数(图片跳过、量化器、速率失真优化技术的λ值……)、图片大小、图片组(group of picture,GOP)布局、最大运动向量搜索范围等。控制器(650)可以被配置成具有与针对特定系统设计而优化的视频编码器(603)有关的其他适当功能。

[0074] 在一些实施方式中,视频编码器(603)被配置成在编码环路中操作。作为过于简化的描述,在示例中,编码环路可以包括源编码器(630)(例如,负责基于要编码的输入图片和参考图片创建诸如符号流的符号)和嵌入在视频编码器(603)中的(本地)解码器(633)。解码器(633)以与(远程)解码器也将创建的类似的方式(因为符号与经编码视频比特流之间的任何压缩在所公开的主题中考虑的视频压缩技术中是无损的)来重建符号以创建样本数据。将重建的样本流(样本数据)输入到参考图片存储器(634)。由于符号流的解码导致与解码器位置(本地或远程)无关的位精确(bit-exact)结果,所以参考图片存储器(634)中的内容在本地编码器与远程编码器之间也是位精确的。换句话说,编码器的预测部分“看到”当解码器在解码期间使用预测时将“看到”的完全相同的样本值作为参考图片样本。参考图片同步性(以及如果例如由于信道误差而不能保持同步性而导致的漂移)的这种基本原理也被用在一些相关技术中。

[0075] “本地”解码器(633)的操作可以与“远程”解码器相同,“远程”解码器是例如已经结合图5在上面详细描述的视频解码器(510)。然而,简要地参照图5,由于符号是可用的并且通过熵编码器(645)和解析器(520)将符号编码/解码为经编码视频序列可以是无损的,所以视频解码器(510)的熵解码部分(包括缓冲存储器(515)和解析器(520))可能不能完全在本地解码器(633)中实现。

[0076] 在这一点上可以观察到,除了存在于解码器中的解析/熵解码之外的任何解码器技术也必须以基本相同的功能形式存在于相应的编码器中。为此,所公开的主题集中于解码器操作。编码器技术的描述可以被简化,因为它们是全面描述的解码器技术的相反技术。下面仅在某些领域需要并提供更详细的描述。

[0077] 在操作期间,在一些示例中,源编码器(630)可以执行运动补偿预测性编码,其参考来自视频序列的被指定为“参考图片”的一个或更多个先前经编码图片预测性地编码输入图片。以该方式,编码引擎(632)对输入图片的像素块与可以被选为输入图片的预测参考的参考图片的像素块之间的差异进行编码。

[0078] 本地视频解码器(633)可以基于由源编码器(630)创建的符号对可以被指定为参考图片的图片的经编码视频数据进行解码。编码引擎(632)的操作可以有利地是有损过程。当可以在视频解码器(图6中未示出)处解码经编码视频数据时,重建的视频序列通常可以是具有一些错误的源视频序列的副本。本地视频解码器(633)复制可以由视频解码器对参考图片执行的解码过程,并且可以使得重建的参考图片被存储在参考图片缓存(634)中。以此方式,视频编码器(603)可以本地存储重建的参考图片的副本,所述副本具有与将由远端视频解码器获得的重建的参考图片共同的内容(不存在传输错误)。

[0079] 预测器(635)可以执行对编码引擎(632)的预测搜索。也就是说,对于要编码的新图片,预测器(635)可以从参考图片存储器(634)搜索可以充当用于新图片的适当预测参考

的样本数据(作为候选参考像素块)或诸如参考图片运动向量、块形状等的特定元数据。预测器(635)可以样本块逐像素块地进行操作以找到适当的预测参考。在一些情况下,如由预测器(635)获得的搜索结果所确定的,输入图片可以具有从存储在参考图片存储器(634)中的多个参考图片获得的预测参考。

[0080] 控制器(650)可以管理源编码器(630)的编码操作,包括例如用于编码视频数据的参数和子组参数的设置。

[0081] 所有前述功能单元的输出可以在熵编码器(645)中经历熵编码。熵编码器(645)通过根据诸如霍夫曼编码、可变长度编码、算术编码等的技术对符号无损压缩,将由各种功能单元生成的符号转换为经编码视频序列。

[0082] 发送器(640)可以缓冲由熵编码器(645)创建的经编码视频序列,以准备经由通信信道(660)进行传输,该通信信道可以是到将存储编码的视频数据的存储设备的硬件/软件链路。发送器(640)可以将来自视频编码器(603)的经编码视频数据与要发送的其他数据(例如,经编码音频数据和/或辅助数据流(未示出源))合并。

[0083] 控制器(650)可以管理视频编码器(603)的操作。在编码期间,控制器(650)可以向每个经编码图片指定某个经编码图片类型,这可以影响可以应用于相应图片的编码技术。例如,图片通常可以被指定为以下图片类型中的一个:

[0084] 帧内图片(I图片)可以是在不使用序列中的任何其他图片作为预测源的情况下可以被编码和解码的图片。一些视频编解码器允许不同类型的帧内图片,包括例如独立解码器刷新(Independent Decoder Refresh,“IDR”)图片。本领域技术人员知道I图片的那些变型及其各自的应用和特征。

[0085] 预测性图片(P图片)可以是使用帧内预测或帧间预测来编码和解码的图片,其中帧内预测或帧间预测使用至多一个运动向量和参考索引来预测每个块的样本值。

[0086] 双向预测性图片(B图片)可以是使用帧内预测或帧间预测来编码和解码的图片,其中帧内预测或帧间预测使用至多两个运动向量和参考索引来预测每个块的样本值。类似地,多个预测性图片可以使用多于两个参考图片和相关联的元数据以用于重建单个块。

[0087] 源图片通常可以在空间上细分为多个样本块(例如,每个4×4、8×8、4×8或16×16个样本的块)并且被逐块地编码。可以参考通过应用于块的相应图片的编码分配所确定其他(已经编码)块来预测性地编码块。例如,I图片的块可以被非预测地编码,或者它们可以参考同一图片的已经编码的块被预测地编码(空间预测或帧内预测)。P图片的像素块可以参考一个先前经编码参考图片经由空间预测或经由时间预测被预测性地编码。B图片的块可以参考一个或两个先前经编码参考图片经由空间预测或经由时间预测被预测性地编码。

[0088] 视频编码器(603)可以根据预定视频编码技术或标准(例如ITU-T建议H.265)执行编码操作。在其操作中,视频编码器(603)可以执行各种压缩操作,包括利用输入视频序列中的时间和空间冗余的预测性编码操作。因此,经编码视频数据可以符合由所使用的视频编码技术或标准指定的语法。

[0089] 在实施方式中,发送器(640)可以与编码的视频一起发送附加数据。源编码器(630)可以包括这样的数据作为经编码视频序列的一部分。附加数据可以包括时间/空间/SNR增强层、诸如冗余图片和切片的其他形式的冗余数据、SEI消息、VUI参数集片段等。

[0090] 视频可以被捕获为时间序列中的多个源图片(视频图片)。帧内图片预测(通常缩写为帧内预测)利用给定图片中的空间相关性,并且图片间预测利用图片之间的(时间或其他)相关性。在示例中,将被称为当前图片的、正在被编码/解码的特定图片划分为块。在当前图片中的块类似于视频中的先前经编码且仍缓冲的参考图片中的参考块时,可以通过被称为运动向量的向量来编码当前图片中的块。运动向量指向参考图片中的参考块,并且在使用多个参考图片的情况下,可以具有标识参考图片的第三维。

[0091] 在一些实施方式中,双向预测技术可以用于帧间图片预测。根据双向预测技术,使用两个参考图片,例如第一参考图片和第二参考图片,其按照解码顺序均在视频中的当前图片之前(但可能在显示顺序上分别在过去和未来)。可以通过指向第一参考图片中的第一参考块的第一运动向量和指向第二参考图片中的第二参考块的第二运动向量对当前图片中的块进行编码。可以通过第一参考块和第二参考块的组合来对块进行预测。

[0092] 此外,可以在帧间图片预测中使用合并模式技术以提高编码效率。

[0093] 根据本公开内容的一些实施方式,以块为单位执行诸如帧间图片预测和帧内图片预测的预测。例如,根据HEVC标准,将视频图片序列中的图片划分成编码树单元(coding tree unit,CTU)以供压缩,图片中的CTU具有相同大小,例如64×64像素、32×32像素或16×16像素。通常,CTU包括三个编码树块(coding tree block,CTB),它们是一个亮度CTB和两个色度CTB。可以将每个CTU递归地以四叉树分割成一个或多个编码单元(coding unit,CU)。例如,64×64像素的CTU可以被分割成64×64像素的一个CU,或32×32像素的4个CU,或16×16像素的16个CU。在示例中,分析每个CU以确定CU的预测类型,例如帧间预测类型或帧内预测类型。取决于时间和/或空间可预测性,将CU分割成一个或更多个预测单元(prediction unit,PU)。通常,每个PU包括亮度预测块(prediction block,PB)和两个色度PB。在实施方式中,以预测块为单位执行编码(编码/解码)中的预测操作。使用亮度预测块作为预测块的示例,预测块包括像素(例如,8×8像素、16×16像素、8×16像素、16×8像素等)的值(例如,亮度值)的矩阵。

[0094] 图7示出了根据本公开内容的另一实施方式的视频编码器(703)的图。视频编码器(703)被配置成接收视频图片序列中的当前视频图片内的样本值的处理块(例如,预测块),并且将处理块编码为作为经编码视频序列的部分的经编码图片。在示例中,使用视频编码器(703)代替图4的示例中的视频编码器(403)。

[0095] 在HEVC示例中,视频编码器(703)接收处理块(例如,8×8样本的预测块等)的样本值的矩阵。视频编码器(703)例如使用率失真优化来确定是否使用帧内模式、帧间模式或双向预测模式最佳地编码处理块。当处理块将以帧内模式被编码时,视频编码器(703)可以使用帧内预测技术来将处理块编码成经编码图片;并且当处理块将以帧间模式或双向预测模式编码时,视频编码器(703)可以分别使用帧间预测或双向预测技术将处理块编码成经编码图片。在某些视频编码技术中,合并模式可以是帧间图片预测子模式,其中,在不借助预测器外部的已编码运动矢量分量的情况下,从一个或更多个运动向量预测器得到运动向量。在某些其他视频编码技术中,可能存在可适用于主题块的运动向量分量。在示例中,视频编码器(703)包括其他部件,例如用于确定处理块的模式的模式决策模块(未示出)。

[0096] 在图7的示例中,视频编码器(703)包括如图7所示耦合在一起的帧间编码器(730)、帧内编码器(722)、残差计算器(723)、开关(726)、残差编码器(724)、通用控制器

(721)以及熵编码器(725)。

[0097] 帧间编码器(730)被配置成接收当前块(例如,处理块)的样本,将块与参考图片中的一个或更多个参考块(例如,先前图片和后续图片中的块)进行比较,生成帧间预测信息(例如,根据帧间编码技术的冗余信息的描述、运动向量、合并模式信息),并且使用任何合适的技术基于帧间预测信息来计算帧间预测结果(例如,已预测块)。在一些示例中,参考图片是基于编码的视频信息解码的经解码参考图片。

[0098] 帧内编码器(722)被配置成接收当前块(例如,处理块)的样本,在一些情况下将该块与同一图片中已经被编码的块进行比较,在变换之后生成量化系数,并且在一些情况下还生成帧内预测信息(例如,根据一个或更多个帧内编码技术的帧内预测方向信息)。在示例中,帧内编码器(722)还基于同一图片中的参考块和帧内预测信息来计算帧内预测结果(例如,已预测块)。

[0099] 通用控制器(721)被配置成确定通用控制数据,并且基于通用控制数据来控制视频编码器(703)的其他部件。在示例中,通用控制器(721)确定块的模式,并且基于该模式向开关(726)提供控制信号。例如,当模式是帧内模式时,通用控制器(721)控制开关(726)以选择供残差计算器(723)使用的帧内模式结果,并且控制熵编码器(725)以选择帧内预测信息并且将帧内预测信息包括在比特流中;并且当模式是帧间模式时,通用控制器(721)控制开关(726)以选择供残差计算器(723)使用的帧间预测结果,并且控制熵编码器(725)以选择帧间预测信息并且将帧间预测信息包括在比特流中。

[0100] 残差计算器(723)被配置成计算接收到的块与从帧内编码器(722)或帧间编码器(730)选择的预测结果之间的差(残差数据)。残差编码器(724)被配置成基于残差数据进行操作以对残差数据进行编码以生成变换系数。在示例中,残差编码器(724)被配置成将残差数据从空间域转换到频域,并且生成变换系数。然后,对变换系数进行量化处理,以获得量化的变换系数。在各种实施方式中,视频编码器(703)还包括残差解码器(728)。残差解码器(728)被配置成执行逆变换,并且生成解码的残差数据。解码的残差数据可以由帧内编码器(722)和帧间编码器(730)适当地使用。例如,帧间编码器(730)可以基于解码残差数据和帧间预测信息来生成解码块,并且帧内编码器(722)可以基于解码的残差数据和帧间预测信息来生成解码块。在一些示例中,解码块被适当地处理以生成解码的图片,并且解码的图片可以在存储器电路(未示出)中被缓冲并且被用作参考图片。

[0101] 熵编码器(725)被配置成格式化比特流以包括编码的块。熵编码器(725)根据诸如HEVC标准的合适标准被配置成包括各种信息。在示例中,熵编码器(725)被配置成包括通用控制数据、所选择的预测信息(例如,帧内预测信息或帧间预测信息)、残差信息和在比特流中的其他合适的信息。注意,根据所公开的主题,当在帧间模式或双向预测模式的合并子模式下对块进行编码时,不存在残差信息。

[0102] 图8示出了根据本公开内容的另一实施方式的视频解码器(810)的图。视频解码器(810)被配置成接收作为经编码视频序列的一部分的经编码图片,并且对经编码图片进行解码以生成重建图片。在示例中,使用视频解码器(810)代替图4的示例中的视频解码器(410)。

[0103] 在图8的示例中,视频解码器(810)包括如图8所示耦合在一起的熵解码器(871)、帧间解码器(880)、残差解码器(873)、重建模块(874)以及帧内解码器(872)。

[0104] 熵解码器(871)可以被配置成根据经编码图片重建表示构成经编码图片的语法元素的某些符号。这样的符号可以包括例如对块进行编码的模式(诸如,帧内模式、帧间模式、双向预测模式,后两者的合并子模式或另一子模式)、可以识别分别用于由帧内解码器(872)或帧间解码器(880)进行的预测的某些样本或元数据的预测信息(诸如,帧内预测信息或帧间预测信息)、例如呈量化变换系数的形式的残差信息等。在示例中,当预测模式是帧间模式或双向预测模式时,帧间预测信息被提供给帧间解码器(880);并且当预测类型是帧内预测类型时,帧内预测信息被提供给帧内解码器(872)。残差信息可以经过逆量化,并且被提供给残差解码器(873)。

[0105] 帧间解码器(880)被配置成接收帧间预测信息,并且基于帧间预测信息生成帧间预测结果。

[0106] 帧内解码器(872)被配置成接收帧内预测信息,并且基于帧内预测信息生成预测结果。

[0107] 残差解码器(873)被配置成执行逆量化以提取解量化的变换系数,并且处理解量化的变换系数以将残差从频域转换到空间域。残差解码器(873)也可能需要某些控制信息(以包括量化器参数(Quantizer Parameter,QP)),并且该信息可以由熵解码器(871)提供(未标示数据路径,因为这仅仅是低量控制信息)。

[0108] 重建模块(874)被配置成在空间域中组合由残差解码器(873)输出的残差和预测结果(视情况而由帧间预测模块或帧内预测模块输出)以形成重建的块,该重建的块可以是重建的图片的一部分,该重建的图片又可以是重建的视频的一部分。应当注意,可以执行其他合适的操作,例如解块操作等,以改善视觉质量。

[0109] 注意,可以使用任何合适的技术来实现视频编码器(403)、(603)和(703)以及视频解码器(410)、(510)和(810)。在实施方式中,视频编码器(403)、(603)和(703)以及视频解码器(410)、(510)和(810)可以使用一个或更多个集成电路来实现。在另一实施方式中,视频编码器(403)、(603)和(703)以及视频解码器(410)、(510)和(810)可以使用执行软件指令的一个或更多个处理器来实现。

[0110] 公开了与基于块的一个或更多个相邻块的重建样本(例如,特征指示符(例如,块的一个或更多个相邻块的重建样本的特征向量或特征标量))的变换集或变换内核选择方案相关的视频编码技术。视频编码格式可以包括为通过因特网进行视频传输而设计的开放视频编码格式,例如AOMedia Video 1(AV1)或超越AV1的下一代AOMedia Video格式。视频编码标准可以包括高效率视频编码(High Efficiency Video Coding,HEVC)标准,超越HEVC的下一代视频编码(例如,通用视频编码(Versatile Video Coding,VVC))等。

[0111] 可以在帧内预测(例如AV1、VVC等)中使用各种帧内预测模式。在实施方式中,例如在AV1中,使用定向帧内预测。在定向帧内预测中,可以通过根据沿一方向的相邻重建样本进行外推来生成块的预测样本。方向对应于角度。用于定向帧内预测以预测块的预测样本的模式可以称为定向模式(也称为定向预测模式、定向帧内模式、定向帧内预测模式、角度模式)。每个定向模式可以对应于不同的角度或不同的方向。在示例中,例如在开放视频编码格式VP9中,八个定向模式对应于从45°到207°的八个角度,如图9所示。八个定向模式也可以称为标称模式(例如,V_PRED、H_PRED、D45_PRED、D135_PRED、D113_PRED、D157_PRED、D203_PRED和D67_PRED)。为了在定向纹理中(例如在AV1中)利用更多种类的空间冗余,可以

将定向模式扩展到例如超过八个标称模式以达到具有更精细粒度和更多角度(或方向)的角度集合,如图9所示。

[0112] 图9示出了根据本公开内容的实施方式的编码块(coding block,CB)(910)的标称模式的示例。某些角度(称为标称角度)可以对应于标称模式。在示例中,八个标称角度(或标称帧内角度)(901)至(908)分别对应于八个标称模式(例如,V_PRED、H_PRED、D45_PRED、D135_PRED、D113_PRE、D157_PRED、D203_PRED和D67_PRED)。八个标称角度(901)至(908)以及八个标称模式可以分别被称为V_PRED、H_PRED、D45_PRED、D135_PRED、D113_PRED、D157_PRED、D203_PRED和D67_PRED。标称模式索引可以指示标称模式(例如,八个标称模式中的一个)。在示例中,用信号指示标称模式索引。

[0113] 此外,每个标称角度可以对应于多个更精细角度(例如,七个更精细角度),并且因此可以例如在AV1中使用56个角度(或预测角度)或56个定向模式(或角度模式、定向帧内预测模式)。每个预测角度可以由标称角度和角度偏移(或角度增量(δ))表示。角度偏移可以通过将偏移整数I(例如,-3、-2、-1、0、1、2或3)与步长(例如,3°)相乘来获得。在示例中,预测角度等于标称角度与角度偏移之和。在示例中,例如在AV1中,可以用信号指示标称模式(例如,八个标称模式(901)至(908))以及某些非角度平滑模式(例如,如下所述的五个非角度平滑模式,例如DC模式、PAETH模式、SMOOTH模式、竖直SMOOTH模式和水平SMOOTH模式)。随后,如果当前预测模式是定向模式(或角度模式),可以进一步用信号发送索引以指示对应于标称角度的角度偏移(例如,偏移整数I)。在示例中,可以基于标称模式索引和指示相对于标称模式的角度偏移的索引来确定定向模式(例如,56个定向模式中的一个)。在示例中,为了经由通用方式实现定向预测模式,用统一定向预测器(predictor)来实现例如在AV1中使用的56个定向模式,统一定向预测器可以将每个像素投影到参考子像素位置并且通过2抽头双线性滤波器来内插参考像素。

[0114] 可以在块(例如CB)的帧内预测中使用非定向平滑帧内预测器(也称为非定向平滑帧内预测模式、非定向平滑模式、非角度平滑模式)。在一些示例中(例如,在AV1中),五种非定向平滑帧内预测模式包括DC模式或DC预测器(例如,DC)、PAETH模式或PAETH预测器(例如,PAETH)、SMOOTH模式或SMOOTH预测器(例如,SMOOTH)、竖直SMOOTH模式(称为SMOOTH_V模式、SMOOTH_V预测器、SMOOTH_V)和水平SMOOTH模式(称为SMOOTH_H模式、SMOOTH_H预测器或SMOOTH_H)。

[0115] 图10示出了根据本公开内容的方面的非定向平滑帧内预测模式(例如,DC模式、PAETH模式、SMOOTH模式、SMOOTH_V模式和SMOOTH_H模式)的示例。为了基于DC预测器预测CB(1000)中的样本(1001),可以使用左相邻样本(1012)的第一值和上相邻样本(或顶部相邻样本)(1011)的第二值的平均值作为预测器。

[0116] 为了基于PAETH预测器来预测样本(1001),可以获得左相邻样本(1012)的第一值、顶部相邻样本(1011)的第二值和左上相邻样本(1013)的第三值。然后,使用式1获得参考值。

[0117] 参考值=第一值+第二值-第三值(式1)

[0118] 可以将最接近参考值的第一值、第二值和第三值中的一个设置为样本(1001)的预测值。

[0119] SMOOTH_V模式、SMOOTH_H模式和SMOOTH模式可以分别使用竖直方向、水平方向以

及竖直方向和水平方向的平均值的二次插值来预测CB(1000)。为了基于SMOOTH预测器来预测样本(1001),可以使用第一值、第二值、右样本(1014)的值和底部样本(1016)的值的平均(例如,加权组合)。在各种示例中,右样本(1014)和底部样本(1016)未被重建,并且因此,右上相邻样本(1015)的值和左下相邻样本(1017)的值可以分别替换右样本(1014)和底部样本(1016)的值。因此,可以使用第一值、第二值、右上相邻样本(1015)的值和左下相邻样本(1017)的值的平均值(例如,加权组合)作为SMOOTH预测器。为了基于SMOOTH_V预测器来预测样本(1001),可以使用顶部相邻样本(1011)的第二值与左下相邻样本(1017)的值的平均值(例如,加权组合)。为了基于SMOOTH_H预测器来预测样本(1001),可以使用左相邻样本(1012)的第一值与右上相邻样本(1015)的值的平均值(例如,加权组合)。

[0120] 图11示出了根据本公开内容的实施方式的基于递归滤波的帧内预测器(也称为滤波器帧内模式或递归滤波模式)的示例。为了捕获与边缘上的参考对象的衰减空间相关性,可以将滤波器帧内模式用于诸如CB

[0121] (1100)的块。在示例中,CB(1100)是亮度块。可以将亮度块(1100)划分为多个分片(例如,八(4×2)个分片B0-B7)。分片B0-B7中的每一个可以具有多个相邻样本。例如,分片B0具有七个相邻样本(或七个相邻元素)R00-R06,包括四个顶部相邻样本R01-R04、两个左相邻样本R05-R06和左上相邻样本R00。类似地,分片B7具有七个相邻样本R70-R76,包括四个顶部相邻样本R71-R74、两个左相邻样本R75-R76和左上相邻样本R70。

[0122] 在一些示例中,例如针对AV1预先设计多个(例如,五个)滤波器帧内模式(或多个递归滤波模式)。每个滤波器帧内模式可以由反映对应4×2分片(例如,B0)中的样本(或像素)与邻近4×2分片B0的七个相邻元素(例如,R00-R06)之间的相关性的一组八个7抽头滤波器来表示。7抽头滤波器的加权因子可以是位置相关的。对于分片B0-B7中的每一个,七个相邻元素(例如,对于B0为R00-R06,对于B7为R70-R76)可以用于预测对应分片中的样本。在示例中,相邻元素R00-R06用于预测分片B0中的样本。在示例中,相邻元素R70-R76用于预测分片B7中的样本。对于CB(1100)中的某些分片例如分片B0,已经重建了所有七个相邻元素(例如,R00-R06)。对于CB(1100)中的其他分片,七个相邻元素中的至少一个未被重建,并且因此可以使用直接相邻元素的预测值(或直接相邻元素的预测样本)作为参考。例如,分片B7的七个相邻元素R70-R76未被重建,因此可以使用直接相邻元素的预测样本。

[0123] 可以从亮度样本预测色度样本。在实施方式中,亮度到色度模式(例如,CfL模式、CfL预测器)是仅色度帧内预测器,仅色度帧内预测器可以将色度样本(或像素)建模为重合的重建亮度样本(或像素)的线性函数。例如,CfL预测可以使用如下的式(2)来表示。

[0124] $CfL(\alpha)=\alpha L^{A}+D$ (式2)

[0125] 其中,L^{A}表示亮度分量的AC贡献,α表示线性模型的缩放参数,并且D表示色度分量的DC贡献。在示例中,基于色度分辨率对重建亮度像素进行子采样,并且减去平均值以形成AC贡献(例如,L^{A})。为了从AC贡献近似色度AC分量,不需要解码器计算缩放参数α,而是在一些示例中,例如在AV1中,CfL模式基于原始色度像素来确定缩放参数α并且在比特流中用信号指示缩放参数α,从而降低解码器复杂度并产生更精确的预测。色度分量的DC贡献可以使用帧内DC模式来计算。帧内DC模式对于大多数色度内容可以是足够的,并且具有成熟的快速实现方式。

[0126] 多线帧内预测可以使用更多的参考线进行帧内预测。参考线可以包括图片中的多

个样本。在示例中,参考线包括成行的样本和成列的样本。在示例中,编码器可以确定并用信号指示用于生成帧内预测器的参考线。可在帧内预测模式之前用信号发送指示参考线的索引(也称为参考线索引)。在示例中,当用信号指示非零参考线索引时,仅允许MPM。图12示出了用于CB(1210)的四条参考线的示例。参照图12,参考线可以包括多达六个段例如段A至段F以及左上参考样本。例如,参考线0包括段B和段E以及左上参考样本。例如,参考线3包括段A至段F以及左上参考样本。段A和段F可以分别用来自段B和段E的最接近的样本填充。在一些示例中,例如在HEVC中,仅将一条参考线(例如,邻近CB(1210)的参考线0)用于帧内预测。在一些示例中,例如在VVC中,将多条参考线(例如,参考线0、参考线1和参考线3)用于帧内预测。

[0127] 通常,可以使用例如以上参照图9至图12描述的各种帧内预测模式中的一个或合适的组合来预测块。

[0128] 下面描述例如在AOMedia Video 1(AV1)中使用的那些初级变换的实施方式。可以对包括残差(例如,空间域中的残差)的变换块(transform block,TB)执行前向变换(例如,在编码器中),以获得包括频域(或空间频率域)中的变换系数的TB。包括空间域中的残差的TB可以被称为残差TB,并且包括频域中的变换系数的TB可以被称为系数TB。在示例中,前向变换包括可以将残差TB变换为系数TB的前向初级变换。在示例中,前向变换包括前向初级变换和前向次级变换,其中前向初级变换可以将残差TB变换为中间系数TB,并且前向次级变换可以将中间系数TB变换为系数TB。

[0129] 可以对频域中的系数TB执行逆变换(例如,在编码器或解码器中)以获得空间域中的残差TB。在示例中,逆变换包括可以将系数TB变换为残差TB的逆初级变换。在示例中,逆变换包括逆初级变换和逆次级变换,其中逆次级变换可以将系数TB变换为中间系数TB,并且逆初级变换可以将中间系数TB变换为残差TB。

[0130] 通常,初级变换可以指前向初级变换或逆初级变换,其中在残差TB与系数TB之间执行初级变换。在一些实施方式中,初级变换可以是可分离变换,其中2D初级变换可以包括水平初级变换(也称为水平变换)和竖直初级变换(也称为竖直变换)。次级变换可以指前向次级变换或逆次级变换,其中在中间系数TB与系数TB之间执行次级变换。

[0131] 为了支持扩展编码块分区,如在本公开内容中所描述的,可以例如在AV1中使用多个变换大小(例如,对于每个维度,范围从4点到64点)和变换形状(例如,具有宽度与高度的比为2:1、1:2、4:1或1:4的矩形,正方形)。

[0132] 2D变换过程可以使用混合变换内核,混合变换内核可以包括用于经编码残差块的每个维度的不同1D变换。初级1D变换可以包括(a)4点、8点、16点、32点、64点DCT-2;(b)4点、8点、16点非对称DST(asymmetric DST,ADST)(例如,DST-4、DST-7)和对应的翻转版本(例如,ADST的翻转版本或FlipADST可以以反向顺序应用ADST);和/或(c)4点、8点、16点、32点恒等变换(identity transform,IDTX)。图13示出了根据本公开内容的实施方式的初级变换基函数的示例。图13的示例中的初级变换基函数包括具有N点输入的DCT-2和非对称DST(DST-4和DST-7)的基函数。图13所示的初级变换基函数可以用于AV1。

[0133] 混合变换内核的可用性可以取决于变换块大小和预测模式。图14A示出了基于变换块大小(例如,第三列中所示的大小)和预测模式(例如,第三列中示出的帧内预测和帧间预测)的各种变换内核(例如,第一列中示出和第二列中描述的变换类型)的可用性的示例

性相关性。可以在AV1中使用基于预测模式和变换块大小的的示例性混合变换内核和可用性。参照图14A,符号“→”和“↓”分别表示水平维度(也称为水平方向)和竖直维度(也称为竖直方向)。符号“√”和“×”表示变换内核对于对应块大小和预测模式的可用性。例如,符号“√”表示变换核是可用的,并且符号“×”表示变换内核是不可用的。

[0134] 在示例中,变换类型(1410)由图14A的第一列中所示的ADST_DCT表示。如图14A的第二列所示,变换类型(1410)包括竖直方向上的ADST和水平方向上的DCT。根据图14A的第三列,当块大小小于或等于16×16

[0135] (例如,16×16样本、16×16亮度样本)时,变换类型(1410)可以用于帧内预测和帧间预测。

[0136] 在示例中,如图14A的第一列所示,变换类型(1420)由V_ADST表示。如图14A的第二列所示,变换类型(1420)包括竖直方向上的ADST和水平方向上的IDTX(即单位矩阵(identity matrix))。因此,变换类型(1420)(例如,V_ADST)在竖直方向上执行而不在水平方向上执行。根据图14A的第三列,不论块大小如何,变换类型(1420)均不可用于帧内预测。当块大小小于16×16(例如,16×16样本、16×16亮度样本)时,变换类型(1420)可用于帧间预测。

[0137] 在示例中,图14A适用于亮度分量。对于色度分量,可以隐式地执行变换类型(或变换内核)选择。在示例中,如图14B所示,对于帧内预测残差,可以根据帧内预测模式来选择变换类型。在示例中,图14B所示的变换类型选择可适用于色度分量。对于帧间预测残差,可以根据共位亮度块的变换类型选择来选择变换类型。因此,在示例中,不在比特流中用信号指示用于色度分量的变换类型。

[0138] 可以将诸如初级变换、次级变换的变换应用于诸如CB的块。在示例中,变换包括初级变换、次级变换的组合。变换可以是非可分离变换、可分离变换、或非可分离变换和可分离变换的组合。

[0139] 可以例如在VVC中执行次级变换。在一些示例中,如图15至图16中所示,例如在VVC中,可以在编码器侧处的前向初级变换与量化之间以及解码器侧处的解量化与逆初级变换之间应用低频非可分离变换(low-frequency non-separable transform,LFNST)(其也称为简化次级变换(reduced secondary transform,RST)),以进一步使初级变换系数解相关。

[0140] 使用4×4输入块(或输入矩阵)X作为示例(如式3所示),可以如下描述可以在LFNST中使用的非可分离变换的应用。如式3至式4中所示,为了应用4×4非可分离变换(例如,LFNST),4×4输入块X可以由向量$\vec{X}$表示。

[0141] $$X=\begin{bmatrix} X_{00} & X_{01} & X_{02} & X_{03} \\ X_{10} & X_{11} & X_{12} & X_{13} \\ X_{20} & X_{21} & X_{22} & X_{23} \\ X_{30} & X_{31} & X_{32} & X_{33} \end{bmatrix} \quad (式3)$$

[0142] $$\vec{X}=[X_{00}\ X_{01}\ X_{02}\ X_{03}\ X_{10}\ X_{11}\ X_{12}\ X_{13}\ X_{20}\ X_{21}\ X_{22}\ X_{23}\ X_{30}\ X_{31}\ X_{32}\ X_{33}]^T \quad (式4)$$

[0143] 可以将非可分离变换计算为$\vec{F}=T\cdot\vec{X}$,其中$\vec{F}$表示变换系数向量,并且T是16×16变换矩阵。随后可以使用4×4输入块的扫描顺序(例如,水平扫描顺序、竖直扫描顺序、Z

字形扫描顺序或对角线扫描顺序)将16×1系数向量 $\vec{F}$ 重新组织成4×4输出块(或输出矩阵、系数块)。具有较小索引的变换系数可以以较小的扫描索引放置在4×4系数块中。

[0144] 非可分离次级变换可以应用于块(例如,CB)。在一些示例中,如图15至图16所示,例如在VVC中,在前向初级变换与量化之间(例如,在编码器侧处)以及在解量化与逆初级变换之间(例如,在解码器侧处)应用LFNST。

[0145] 图15至图16分别示出了使用16×64变换(或64×16变换,这取决于变换是前向还是逆次级变换)和16×48变换(或48×16变换,这取决于变换是前向还是逆次级变换)的两个变换编码过程(1700)和(1800)的示例。参照图15,在过程(1700)中,在编码器侧处,可以首先对块(例如,残差块)执行前向初级变换(1710)以获得系数块(1713)。随后,可以将前向次级变换(或前向LFNST)(1712)应用于系数块(1713)。在前向次级变换(1712)中,系数块(1713)的左上角处的4×4子块A-D的64个系数可以由长度为64的向量表示,并且可以将长度为64的向量与64×16的变换矩阵(即,宽度为64且高度为16)相乘,从而得到长度为16的向量。将长度为16的向量中的元素填充回系数块(1713)的左上4×4子块A中。子块B-D中的系数可以是零。然后,在量化步骤(1714)处对前向次级变换(1712)之后产生的系数进行量化,并且对其进行熵编码以生成比特流(1716)中的经编码比特。

[0146] 经编码比特可以在解码器侧处被接收,并且被熵解码,随后进行解量化步骤(1724)以生成系数块(1723)。可以执行逆次级变换(或逆LFNST)

[0147] (1722),例如逆RST8×8,以例如从左上4×4子块E处的16个系数获得64个系数。可以将64个系数填充回4×4子块E-H。此外,可以用逆初级变换(1720)来处理逆次级变换(1722)之后的系数块(1723)中的系数,以获得恢复的残差块。

[0148] 图16的示例的过程(1800)类似于过程(1700),除了在前向次级变换(1712)期间处理更少的(即,48个)系数之外。具体地,用大小为48×16的较小变换矩阵来处理子块A-C中的48个系数。使用48×16的较小变换矩阵可以减小用于存储变换矩阵的存储器大小和计算次数(例如,乘法、加法、减法等),并且因此可以降低计算复杂度。

[0149] 在示例中,根据块(例如,CB)的块大小来应用4×4非可分离变换(例如,4×4LFNST)或8×8非可分离变换(例如,8×8LFNST)。块的块大小可以包括宽度、高度等。例如,在宽度和高度的最小值小于阈值(例如8)(例如,min(宽度,高度)<8)的情况下,将4×4LFNST应用于块。例如,在宽度和高度的最小值大于阈值(例如4)(例如,min(宽度,高度)>4))的情况下,将8×8LFNST应用于块。

[0150] 非可分离变换(例如,LFNST)可以基于直接矩阵相乘方法,并且因此可以单遍实现而无需迭代。为了减小非可分离变换矩阵的维度并且使存储变换系数的存储空间和计算复杂度最小化,可以在LFNST中使用简化的非可分离变换方法(或RST)。因此,在简化的非可分离变换中,可以将N(例如,对于8×8非可分离次级变换(non-separable secondary transform,NSST),N是64)维向量映射到不同空间中的R维向量,其中N/R(R<N)是简化因子。因此,代替N×N矩阵,RST矩阵是如式5中描述的R×N矩阵。

[0151] $$T_{RxN}=\begin{bmatrix} t_{11} & t_{12} & t_{13} & \cdots & t_{1N} \\ t_{21} & t_{22} & t_{23} & & t_{2N} \\ & \vdots & & \ddots & \vdots \\ t_{R1} & t_{R2} & t_{R3} & \cdots & t_{RN} \end{bmatrix} \quad (式5)$$

[0152] 在式5中,R×N变换矩阵的R行是N维空间的R个基。逆变换矩阵可以是前向变换中使用的变换矩阵(例如,T_{RxN})的转置。对于8×8LFNST,可以应用简化因子4,并且在8×8非可分离变换中使用的64×64直接矩阵可以简化为16×64直接矩阵,如图15所示。替选地,可以应用大于4的简化因子,并且在8×8非可分离变换中使用的64×64直接矩阵可以简化为16×48直接矩阵,如图16所示。因此,可以在解码器侧处使用48×16逆RST矩阵以在8×8左上区域中生成核心(初级)变换系数。

[0153] 参照图16,当应用16×48矩阵代替具有相同变换集配置的16×64矩阵时,16×48矩阵的输入包括来自左上8×8块中的除右下4×4块D之外的三个4×4块A、B和C的48个输入数据。通过降低维度,可以在性能下降最小的情况下减少用于存储LFNST矩阵的存储器使用量,例如从10KB减小到8KB。

[0154] 为了降低复杂度,可以将LFNST限制为在第一系数子组之外的系数是非有效的情况下是适用的。在示例中,可以将LFNST限制为仅在第一系数子组之外的所有系数均为非有效的情况下才适用。参照图15至图16,第一系数子组对应于左上块E,并且因此块E之外的系数是非有效的。

[0155] 在示例中,当应用LFNST时,仅初级变换系数是非有效的(例如,零)。在示例中,当应用LFNST时,所有仅初级变换系数均为零。仅初级变换系数可以指在没有次级变换的情况下从初级变换获得的变换系数。因此,LFNST索引信令可以以最后有效位置为条件,从而避免LFNST中的额外系数扫描。在一些示例中,额外系数扫描用于检查特定位置处的有效变换系数。在示例中,在每像素乘法方面,LFNST的最坏情况处理例如将4×4块和8×8块的非可分离变换分别限制为8×16变换和8×48变换。在上述情况下,当应用LFNST时,最后有效扫描位置可以小于8。对于其他大小,当应用LFNST时,最后有效扫描位置可以小于16。对于4×N和N×4的CB,并且N大于8,该限制可以意味着LFNST应用于CB中的左上4×4区域。在示例中,该限制意味着LFNST对仅在CB中的左上4×4区域仅应用一次。在示例中,当应用LFNST时,所有仅初级系数是非有效的(例如,零),用于初级变换的操作的数目减少。从编码器的角度来看,当测试LFNST变换时,可以显著地简化变换系数的量化。对于例如按照扫描顺序的前16个系数,可以最大程度地进行率失真优化量化,可以将剩余系数设置为零。

[0156] 可以如下所述选择LFNST变换(例如,变换内核、变换核心或变换矩阵)。在实施方式中,可以使用多个变换集,并且在LFNST中的多个变换集中的每一个中可以包括一个或更多个非可分离变换矩阵(或内核)。根据本公开内容的各方面,可以从多个变换集中选择变换集,并且可以从变换集中的一个或更多个非可分离变换矩阵中选择非可分离变换矩阵。

[0157] 表1示出了根据本公开内容的实施方式的从帧内预测模式到多个变换集的示例性映射。该映射指示帧内预测模式与多个变换集之间的关系。诸如表1指示的关系可以是预定义的,并且可以存储在编码器和解码器中。

[0158] 表1:变换集选择表

[0159]

IntraPredMode	Tr.set index
IntraPredMode＜c	1
0＜＝IntraPredMode＜＝1	0
2＜＝IntraPredMode＜＝12	1
13＜＝IntraPredMode＜＝23	2

24<＝IntraPredMode<＝44	3
45<＝IntraPredMode<＝55	2
56<＝IntraPredMode<＝80	1
81<＝IntraPredMode<＝83	0

[0160] 参考表1,多个变换集包括四个变换集,例如,由从0至3的变换集索引(例如,Tr.set index)分别表示的变换集0至变换集3。索引(例如,IntraPredMode)可以指示帧内预测模式,并且可以基于索引和表1来获得变换集索引。因此,可以基于帧内预测模式来确定变换集。在示例中,如果三个跨分量线性模型(cross component linear model,CCLM)模式(例如,INTRA_LT_CCLM、INTRA_T_CCLM或INTRA_L_CCLM)中的一个用于CB(例如,81≤IntraPredMode≤83),则针对该CB选择变换集0。

[0161] 如上所述,每个变换集可以包括一个或更多个非可分离变换矩阵。一个或更多个非可分离变换矩阵之一可以通过例如显式地用信号指示的LFNST索引来选择。LFNST索引可以例如在用信号指示变换系数之后在每个帧内编码CU(例如,CB)的比特流中用信号指示一次。在实施方式中,每个变换集包括两个非可分离变换矩阵(内核),并且所选择的非可分离次级变换候选可以是两个非可分离变换矩阵之一。在一些示例中,不将LFNST应用于CB(例如,用变换跳过模式编码的CB或CB的非零系数的数目小于阈值)。在示例中,当LFNST不应用于CB时,针对该CB不用信号指示LFNST索引。LFNST索引的默认值可以是零并且不用信号指示,从而指示LFNST不应用于CB。

[0162] 在实施方式中,将LFNST限制为仅在第一系数子组之外的所有系数均为非有效时适用,LFNST索引的编码可以取决于最后有效系数的位置。LFNST索引可以是上下文编码的。在示例中,LFNST索引的上下文编码不取决于帧内预测模式,并且仅第一二进制位被上下文编码。LFNST可以应用于帧内切片或帧间切片中的帧内编码CU,并且可以应用于亮度分量和色度分量两者。如果启用双树,则可以分别用信号指示用于亮度分量和色度分量的LFNST索引。对于帧间切片(例如,禁用双树)而言,可以用信号指示单个LFNST索引并且LFNST索引可以用于亮度分量和色度分量两者。

[0163] 可以使用帧内子分区(intra sub-partition,ISP)编码模式。在ISP编码模式中,可以根据块大小将亮度帧内预测块竖直或水平地划分为2或4个子分区。在一些示例中,当将RST应用于每个可行子分区时,性能提升是微不足道的。因此,在一些示例中,当选择ISP模式时,禁用LFNST并且不用信号指示LFNST索引(或RST索引)。对ISP预测的残差禁用RST或LFNST可以降低编码复杂度。在一些示例中,当选择基于矩阵的帧内预测模式(matrix-based intra predication mode,MIP)时,禁用LFNST并且不用信号指示LFNST索引。

[0164] 在一些示例中,由于最大变换大小限制(例如,64×64)而隐式地分割(TU平铺(tiling))大于64×64的CU,LFNST索引搜索可以针对某一数目的解码流水线级将数据缓冲提高四倍。因此,可以将LFNST所被允许的最大大小限制为64×64。在示例中,仅通过离散余弦变换(discrete cosine transform,DCT)类型2(DCT-2)变换来启用LFNST。

[0165] 在一些示例中,可分离变换方案对于捕获定向纹理图案(例如,沿45°或135°方向的边缘)可能不是高效的。例如,在上述情形中,非可分离变换方案可以提高编码效率。为了降低计算复杂度和存储器使用量,可以使用非可分离变换方案作为次级变换,该次级变换应用于从初级变换获得的低频变换系数。

[0166] 在某些实现方式中,例如从成组变换内核中选择待使用的变换内核是基于预测模式信息。预测模式信息可以指示预测模式。

[0167] 在一些示例中,预测模式信息单独仅提供针对预测模式观察到的残差图案的整个空间的粗略表示。图17A至图17D示出了根据本公开内容的实施方式的针对作为AV1中的帧内预测模式的定向模式D45_PRED(45°)观察到的示例性残差图案(以灰度表示)。根据本公开内容的各方面,相邻经重建样本可以提供用于更高效地表示残差图案的附加信息。因此,公开了除了基于预测模式信息之外还基于相邻重建样本的变换集选择和/或变换内核选择的方法。例如,除了预测模式信息之外,还使用相邻重建样本的特征指示符(例如,特征向量$\vec{F}$或特征标量S)作为选择过程的输入。

[0168] 在本公开内容中,术语块可以指PB、CB、经编码块、编码单元(coding unit,CU)、变换块(transform block,TB)、变换单元(transform unit,TU)、亮度块(例如,亮度CB)、色度块(例如,色度CB)等。

[0169] 在本公开内容中,块的大小可以指块宽度、块高度、块纵横比(例如,块宽度与块高度的比、块高度与块宽度的比)、块面积大小或块面积(例如,块宽度×块高度)、块宽度和块高度的最小值、块宽度和块高度的最大值等。

[0170] 在本公开内容中,块的变换内核可以用于初级变换、次级变换、第三级变换或超越第三级变换的变换方案。变换内核可以用于可分离变换或非可分离变换。变换内核可以用于亮度块、色度块、帧间预测、帧内预测等。变换内核也可以称为变换核心、变换候选、变换内核选项等。在示例中,变换内核是变换矩阵。可以至少基于变换内核来对块进行变换。因此,本公开内容中的方法可以应用于初级变换、次级变换、第三级变换、超越第三级变换的任何变换方案、可分离变换、非可分离变换、亮度块、色度块、帧间预测、帧内预测等。

[0171] 在本公开内容中,变换集可以指成组变换内核或变换内核选项。变换集可以包括一个或更多个变换内核或变换内核选项。在示例中,可以从变换集中选择块的变换内核。

[0172] 根据本公开内容的各方面,相邻重建样本可以用于从成组变换集中确定变换内核。当前图片中正在重建的块的相邻重建样本可以用于从成组变换集中确定块的变换内核。成组变换集可以是预定的。在示例中,成组变换集预存储在编码器和/或解码器中。

[0173] 块的相邻重建样本(例如,相邻重建样本集)可以指来自当前图片中的先前解码的相邻块的重建样本(例如,成组重建样本)或先前解码的图片中的重建样本。先前解码的相邻块可以包括该块的一个或更多个相邻块,并且相邻重建样本也被称为该块的一个或更多个相邻块中的重建样本。该块的一个或更多个相邻块可以在当前图片或与当前图片不同的重建图片(例如,参考图片)中。该块的相邻重建样本可以包括在当前图片中的块的空间相邻样本和/或在与当前图片不同的另一图片(例如,先前解码的图片)中的块的时间相邻样本。图18示出了根据本公开内容的实施方式的块的示例性空间相邻样本。该块是当前图片中的当前块(1851)。样本1-4和A-X是当前块(1851)的空间相邻样本并且已经被重建。在示例中,样本1-4和A-X在当前块(1851)的一个或更多个重建相邻块中。根据本公开内容的各方面,样本1-4和A-X中的一个或更多个可以用于确定当前块(1851)的变换内核。在示例中,当前块(1851)的相邻重建样本包括左上相邻样本4、包括样本A-F的左相邻列(1852)以及包括样本G-L的上相邻行(1854),并且用于确定当前块(1851)的变换内核。在示例中,当前块(1851)的相邻重建样本包括左上相邻样本1-4、包括样本A-F和M-R的左相邻列(1852)至

(1853)、以及包括样本G-L和S-X的上相邻行(1854)-(1855),并且用于确定当前块(1851)的变换内核。

[0174] 成组变换集可以包括一个或更多个变换集。一个或更多个变换集中的每一个可以包括任何合适的一个或更多个变换内核。因此,成组变换集可以包括在初级变换、次级变换、第三级变换或超越三级变换的变换方案中使用的变换内核。成组变换集可以包括作为可分离变换和/或非可分离变换的变换内核。成组变换集可以包括用于亮度块、色度块、帧间预测、帧内预测等的变换内核。

[0175] 根据本公开内容的各方面,可以基于当前图片中的块的一个或更多个相邻块中的重建样本从成组变换集确定该块的变换内核(或变换候选)。可以基于从当前图片中的块的一个或更多个相邻块中的重建样本中提取的特征指示符(例如,特征向量$\vec{F}$或特征标量S)从成组变换集中确定该块的变换内核(或变换候选)。成组变换集中的每个集合可以包括该块的一个或更多个变换候选。一个或更多个相邻块可以在当前图片或与当前图片不同的重建图片中。此外,可以基于所确定的变换候选来重建块的样本。

[0176] 在实施方式中,可以从相邻重建样本中提取相邻重建样本的特征指示符(例如,特征向量$\vec{F}$或特征标量S)。因此,可以基于特征指示符来确定块(例如,当前块(1851))的变换内核。特征指示符可以是向量(称为特征向量$\vec{F}$)或标量(称为特征标量S)。可以基于相邻重建样本的统计信息或统计分析,例如基于与相邻重建样本相关联的一个或更多个变量,来提取特征指示符。一个或更多个变量可以指示相邻重建样本的样本值的一个或更多个分布。

[0177] 在实施方式中,可以基于块的相邻重建样本(例如,特征指示符)和预测模式信息来确定块(例如,当前块(1851))的变换内核。块的预测模式信息可以指示用于对块进行预测(例如块的帧间预测或帧内预测)的信息。根据本公开内容的各方面,块的预测模式信息可以指示块的预测模式(例如,帧内预测模式、帧间预测模式)。在示例中,基于块的相邻重建样本(例如,特征指示符)和预测模式来确定块的变换内核。

[0178] 在示例中,块被帧内编码或帧内预测,并且预测模式信息被称为帧内预测模式信息。块的预测模式信息(例如,帧内预测模式信息)可以指示用于块的帧内预测模式。帧内预测模式可以指用于块的帧内预测的预测模式,例如图9中描述的定向模式(或定向预测模式)、图10中描述的非定向预测模式(例如,DC模式、PAETH模式、SMOOTH模式、SMOOTH_V模式或SMOOTH_H模式)或图11中描述的递归滤波模式。帧内预测模式还可以指本公开内容中描述的预测模式、本公开内容中描述的预测模式的合适变型或本公开内容中描述的预测模式的合适组合。例如,可以将帧内预测模式与图12中描述的多线帧内预测组合。

[0179] 更具体地,例如,可以基于来自经编码视频比特流的块的经编码信息从成组变换集中选择变换集子组。块的经编码信息可以包括指示块的预测模式(例如,帧内预测模式或帧间预测模式)的预测模式信息。在实施方式中,可以基于预测模式信息(例如,预测模式)从成组变换集中选择变换集子组。

[0180] 根据本公开内容的各方面,可以基于在块的经编码信息中指示的块的预测模式从成组变换集中选择变换集子组。此外,可以基于块的一个或更多个相邻块中的重建样本来确定来自变换集子组的变换候选。例如,可以基于从块的一个或更多个相邻块中的重建样

本提取的特征指示符(例如,特征向量$\vec{F}$或特征标量S)来确定来自变换集子组的变换候选。

[0181] 在从成组变换集中选择变换集子组之后,可以使用如下所述的任何合适的方法至少基于相邻重建样本(例如,特征指示符)来进一步确定块的变换内核。

[0182] 在实施方式中,使用块的相邻重建样本(或块的一个或更多个相邻块中的重建样本)从所选择的变换集子组中识别或选择所选择的变换集子组中的变换集。在示例中,使用从块的一个或更多个相邻块中的重建样本提取的特征指示符(例如,特征向量$\vec{F}$或特征标量S)来识别或选择所选择的变换集子组中的变换集。此外,可以基于第一索引从所识别的(或所选择的)变换集确定(例如,选择)块的变换内核(或变换候选)。经编码信息可以指示(例如,包括)第一索引。第一索引可以在经编码视频比特流中被用信号指示。

[0183] 在实施方式中,基于第二索引从所选择的变换集子组中选择所选择的变换集子组中的变换集。经编码信息可以指示(例如,包括)第二索引。第二索引可以在经编码视频比特流中被用信号指示。此外,可以基于块的相邻重建样本(或块的一个或更多个相邻块中的重建样本)(例如,特征指示符)来确定(例如,选择)来自所选择的变换集的块的变换内核(或变换候选)。在示例中,可以使用块的特征指示符(例如,特征向量$\vec{F}$或特征标量S)来确定所选择的变换集中的变换内核(或变换候选)。

[0184] 在实施方式中,基于块的相邻重建样本从所选择的变换集子组中隐式地确定(或识别)块的变换内核(或变换候选)。在示例中,使用来自所选择的变换集子组的块的特征指示符(例如,特征向量$\vec{F}$或特征标量S)并且不使用经编码视频比特流中用信号指示的索引来隐式地确定(或识别)变换内核(或变换候选)。

[0185] 可以基于相邻的重建样本来确定(例如,得到)块的特征指示符(例如,特征向量$\vec{F}$或特征标量S)。根据本公开内容的各方面,可以根据对块的相邻重建样本(或块的一个或更多个相邻块中的重建样本)的统计分析确定或得到块的特征指示符(例如,特征向量$\vec{F}$或特征标量S)。如上所述,可以执行以下操作中的一个:(i)基于特征指示符(例如,特征向量$\vec{F}$或特征标量S)从变换集子组确定变换集,(ii)基于特征指示符(例如,特征向量$\vec{F}$或特征标量S)从变换集子组确定变换候选,或者(iii)可以基于经编码信息中的索引从变换集子组中选择变换集,并且基于特征指示符(例如,特征向量$\vec{F}$或特征标量S)从所选择的变换集确定变换候选。

[0186] 在实施方式中,单个变量X可以用于表示相邻重建样本,并且块的特征指示符是块的特征标量S。变量X可以指示相邻重建样本的样本值。在示例中,变量X是包括相邻重建样本的样本值的数组,并且反映了样本值在相邻重建样本上的分布。特征标量S可以指示样本值的统计信息。特征标量S可以包括但不限于可以从相邻重建样本获得的变量X的标量定量测量,例如变量X的均值(或一阶矩)、变量X的方差(或二阶矩)、变量X的偏度(或三阶矩)。在示例中,变量X被称为随机变量X。

[0187] 在示例中,特征指示符是特征标量S。特征标量S被确定为变量X的矩(例如,一阶矩、二阶矩、三阶矩等),其指示块的一个或更多个相邻块中的重建样本的样本值。

[0188] 变量X的示例如图18所示。在示例中,变量X包括与当前块(1851)相邻的重建相邻样本,包括样本4、左相邻列(1852)和上相邻行(1854)。在示例中,变量X包括重建相邻样本1至4和A至X。

[0189] 在实施方式中,可以使用多个变量(例如,两个变量Y和Z)来分别表示相邻重建样本(如果可用)的不同集合(例如,对应Y的第一集合和对应Z的第二集合),并且块的特征指示符是该块的特征向量$\vec{F}$。多个变量可以分别指示相邻重建样本的不同集合的样本值。在示例中,变量Y是包括相邻重建样本的第一集合的第一样本值的数组,并且反映了相邻重建样本的第一集合的第一样本值的分布。变量Z是包括相邻重建样本的第二集合的第二样本值的数组,并且反映了相邻重建样本的第二集合的第二样本值的分布。特征向量$\vec{F}$可以指示多个变量的统计信息。

[0190] 特征向量$\vec{F}$可以包括但不限于可以从相邻重建样本获得的多个变量的统计分析,例如多个变量的联合变化性测量(例如,多个变量(例如,两个变量Y和Z)的协方差或二阶矩)。在示例中,该多个变量(例如,两个变量Y和Z)被称为随机变量。

[0191] 图18中示出了多个变量Y和Z的示例。在示例中,变量Y包括左相邻列(1852),并且变量Z包括上相邻行(1854)。在示例子中,变量Y和变量Z中的一个包括左上样本4。

[0192] 在示例中,特征指示符是特征向量$\vec{F}$。特征向量$\vec{F}$被确定为变量(例如,变量Y和变量Z)的联合变化性,这些变量分别指示块(例如,当前块(1851))的左侧的相邻列(例如,左相邻列(1852))的重建样本的样本值和块(例如,当前块(1851))的上侧的相邻行(例如,上相邻行(1854))的重建样本的样本值。

[0193] 在示例中,变量Y包括左相邻列(1852)至(1853),并且变量Z包括上相邻行(1854)至(1855)。在示例中,变量Y和变量Z中的一个包括左上样本1至4。

[0194] 根据本公开内容的各方面,当将特征指示符用于确定块的变换内核或者用于确定包括块的变换内核的变换集时,可以基于特征指示符和阈值来确定变换内核或变换集。可以从预定义的阈值集中选择阈值。可以使用任何合适的方法基于特征指示符和阈值来确定变换内核或变换集。下面描述一些示例性方法。如上所述,根据本公开内容的各方面,例如基于预测模式信息从成组变换集中选择变换集子组。随后,在示例中,使用块的特征指示符和阈值从所选择的变换集子组中识别变换集。在另一示例中,使用经编码视频比特流中用信号指示的索引(例如,第二索引)来选择所选择的变换集子组中的变换集,并且可以使用块的特征指示符和阈值来确定(例如,选择)来自所选择的变换集的块的变换内核。在另一示例中,使用来自所选择的变换集子组的特征指示符和阈值来隐式地识别变换内核。

[0195] 例如,为了分类的目的,可以预定义阈值集(表示为K_s)。在示例中,可以执行以下操作中的一个:(i)基于变量X的矩和从预定义的阈值集K_s中选择的阈值从变换集子组确定变换集,(ii)基于变量X的矩和阈值从变换集子组确定变换候选,或者(iii)基于经编码信息中的索引从变换集子组中选择变换集,并且基于变量X的矩和阈值从所选择的变换集确定变换候选。

[0196] 在实施方式中,特征标量S用于例如从如上所述的所选择的变换集子组识别的块的变换内核或者识别包括块的变换内核的变换集。阈值集K_s可以包括一个或更多个第一阈值(或第一阈值)。

[0197] 在示例中,特征标量S是变量X的定量测量,例如变量X的均值、方差、偏度等。例如,特征标量S是变量X的矩。变量的矩可以是变量X的一阶矩(或均值)、变量X的二阶矩(或方差)或变量X的三阶矩(例如偏度)中的一个。

[0198] 每个预测模式(例如,帧内预测模式或帧间预测模式)可以对应于阈值集K_s中的唯一阈值子集K_s',例如,这指示多个预测模式与阈值集K_s中的多个阈值子集之间的单射映射(injective mapping)。块的预测模式可以是多个预测模式中的一个。每个预测模式与对应的阈值子集K_s'之间的映射是单射映射。例如,阈值集K_s包括与第一预测模式相对应的阈值子集K_{s1}'和与第二预测模式相对应的阈值子集K_{s2}',并且阈值子集K_{s1}'中没有元素或没有阈值与阈值子集K_{s2}'中的元素或阈值相同。

[0199] 在示例中,特征标量S是变量X的定量测量,例如变量X的均值、方差、偏度等。每个预测模式可以对应于阈值集K_s中的任何阈值子集K_s'。每个预测模式与阈值集K_s中的对应阈值子集K_s'之间的映射是非单射映射。在示例中,多个预测模式(例如,第一预测模式和第二预测模式)可以对应于阈值集K_s中的相同阈值子集K_s'。在示例中,阈值集K_s包括与第一预测模式相对应的阈值子集K_{s3}'和与第二预测模式相对应的阈值子集K_{s4}',并且阈值子集K_{s3}'中的元素(或阈值)与阈值子集K_{s4}'中的元素(或阈值)相同。在示例中,阈值集K_s是单个元素集合。

[0200] 在实施方式中,阈值集K_s中的阈值子集K_s'的元素是取决于量化索引(或与量化索引相关联的对应量化步长)的阈值。通常,每个量化索引可以对应于唯一量化步长。在每个预测模式与阈值集K_s中的对应阈值子集K_s'之间的映射是非单射或单射的情况下,以上描述是适用的。

[0201] 在实施方式中,仅定义阈值集K_s中的阈值子集K_s'的与特定量化索引(或对应的量化步长)相对应的一个元素。可以使用映射函数得出阈值子集K_s'中的剩余元素(如果有)。映射函数可以是线性的或非线性的。映射函数可以由编码器和/或解码器预先定义和使用。阈值子集K_s'中的剩余元素可以对应于不同的量化索引或(或对应的量化步长)。在每个预测模式与阈值集K_s中的对应阈值子集K_s'之间的映射是非单射或单射的情况下,以上描述是适用的。

[0202] 在实施方式中,可以使用查找表来确定阈值子集K_s'。例如,使用查找表来基于预测模式和/或量化索引来为阈值子集K_s'选择元素。在示例中,查找表包括预测模式、量化索引(或对应的量化步长)与阈值的关系。可以遍历查找表以使用预测模式(例如,帧内预测模式或帧间预测模式)和/或量化索引(或对应的量化步长)来确定阈值子集K_s'(例如,为阈值子集K_s'选择元素)。在每个预测模式与阈值集K_s中的对应阈值子集K_s'之间的映射是非单射或单射的情况下,以上描述是适用的。

[0203] 在实施方式中,从量化索引(或对应的量化步长)到阈值的映射是线性映射。线性映射中使用的参数,例如斜率和截距,可以是预先定义的或使用块的经编码信息得出的。经编码信息可以包括但不限于块大小、量化索引(或对应的量化步长)、预测模式(例如,帧内预测模式或帧间预测模式)。线性映射(例如,线性映射中所使用的参数)可以是预先定义的或基于块大小、量化索引(或对应的量化步长)、预测模式中的一个或组合来得出的。如果从量化索引(或对应的量化步长)到阈值的映射是非线性映射,则可以适当地调整该描述。在每个预测模式与阈值集K_s中的对应阈值子集K_s'之间的映射是非单射或单射的情况下,以上

描述是适用的。

[0204] 如上所述,特征标量S可以用于识别包括块的变换内核的变换集或者用于识别块的变换内核,并且阈值集K_s可以包括一个或更多个第一阈值。在实施方式中,取决于块的块大小来例如从阈值集K_s中选择阈值。在实施方式中,例如从阈值集K_s中对阈值的选择取决于量化中使用的量化参数,例如量化索引(或对应的量化步长)。在实施方式中,例如从阈值集K_s中对阈值的选择取决于预测模式(例如,帧内预测模式和/或帧间预测模式)。

[0205] 在实施方式中,特征向量$\vec{F}$用于识别块的变换内核或者用于识别包括块的变换内核的变换集。为了分类的目的,可以预先定义分类向量集$\vec{C}$和与分类向量集$\vec{C}$相关联的阈值集(表示为K_v)。阈值集K_v可以包括一个或更多个第二阈值或者多个第二阈值。

[0206] 在示例中,对于特征向量$\vec{F}$,每个预测模式(例如,帧内预测模式或帧间预测模式)可以对应于分类向量集$\vec{C}$中的唯一分类向量子集$\vec{C'}$和阈值集K_v中的唯一阈值子集K_v'。唯一分类向量子集$\vec{C'}$对应于唯一阈值子集K_v'。每个预测模式与对应的分类向量子集$\vec{C'}$(连同相关联的阈值子集K_v')之间的映射是单射映射。例如,阈值集K_v包括阈值子集K_{v1}'和阈值子集K_{v2}'。分类向量集$\vec{C}$包括分类向量子集$\overrightarrow{C1'}$和分类向量子集$\overrightarrow{C2'}$。一种预测模式对应于阈值子集K_{v1}'和分类向量子集$\overrightarrow{C1'}$,并且另一预测模式对应于阈值子集K_{v2}'和分类向量子集$\overrightarrow{C2'}$。阈值子集K_{v1}'中没有阈值与阈值子集K_{v2}'中的阈值相同,并且分类向量子集$\overrightarrow{C1'}$中没有分类向量与分类向量子集$\overrightarrow{C2'}$中的分类向量相同。

[0207] 在示例中,每个预测模式(例如,帧内预测模式或帧间预测模式)可以对应于阈值集K_v中的任何阈值子集K_v'以及分类向量集$\vec{C}$中的任何对应的分类向量子集$\vec{C'}$。每个预测模式与阈值集K_v中的对应的阈值子集K_v'之间的映射是非单射映射。每个预测模式与分类向量集$\vec{C}$中的对应的分类向量子集$\vec{C'}$之间的映射是非单射映射。在示例中,每个预测模式与对应的阈值子集K_v'以及对应的分类向量子集$\vec{C'}$之间的映射是非单射映射。在示例中,多个预测模式可以对应于同一阈值子集K_v'(以及与阈值子集K_v'相关联的对应的分类向量子集$\vec{C'}$)。

[0208] 在实施方式中,阈值集K_v中的阈值子集K_v'的元素是取决于量化索引(或对应的量化步长)的阈值。同样,分类向量集$\vec{C}$中的对应分类向量子集$\vec{C'}$的元素是取决于量化索引(或对应的量化步长)的分类向量。在(i)每个预测模式与(ii)对应的阈值子集K_v'以及与阈值子集K_v'相对应的分类向量子集$\vec{C'}$之间的映射是非单射或单射的情况下,以上描述是适用的。

[0209] 在实施方式中,仅定义阈值集K_s中的阈值子集K_v'的一个元素以及分类向量集$\vec{C}$中的对应的分类向量子集$\vec{C'}$的对应于特定量化索引(或与特定量化索引相对应的量化步长)一个元素。可以使用映射函数来得出阈值子集K_v'中的剩余元素(如果有)和分类向量子集

$\vec{C}$ 中的剩余元素(如果有)。映射函数可以是线性的或非线性的。映射函数可以由编码器和/或解码器预先定义和使用。阈值子集K_v'中的剩余元素和对应的分类向量子集$\vec{C}$中的剩余元素可以对应于不同的量化索引(或对应于量化索引的量化步长)。在(i)每个预测模式与(ii)对应的阈值子集K_v'以及与阈值子集K_v'相对应的分类向量子集$\vec{C}$之间的映射是非单射或单射的情况下,以上描述是适用的。

[0210] 在实施方式中,可以使用查找表以通过为阈值子集K_v'选择元素来确定阈值子集K_v',并且通过为分类向量子集$\vec{C}$选择元素来确定对应的分类向量子集$\vec{C}$。在示例中,查找表包括预测模式、量化索引(或对应于量化索引的量化步长)、阈值与分类向量的关系。可以遍历查找表以使用预测模式(例如,帧内预测模式或帧间预测模式)和/或量化索引(或对应于量化索引的量化步长)来确定阈值子集K_v'(例如,为阈值子集K_v'选择元素)和对应的分类向量子集$\vec{C}$(例如,为分类向量子集$\vec{C}$选择元素)。在(i)每个预测模式与(ii)对应的阈值子集K_v'以及与阈值子集K_v'相对应的分类向量子集$\vec{C}$之间的映射是非单射或单射的情况下,以上描述是适用的。

[0211] 在实施方式中,从块的相邻重建样本获得特征向量$\vec{F}$。此外,使用预测模式信息和量化索引(或对应的量化步长),从分类向量子集$\vec{C}$中获得分类向量(例如,子集$\vec{C}$中的元素$\overrightarrow{\{C'\}}$),并且从阈值子集K_v'中获得与该分类向量相对应的阈值(例如,阈值子集K_v'中的元素{Kv′})。注意,{}

[0212] (例如,$\overrightarrow{\{C'\}}$)表示对应集合或子集(例如,分类向量子集$\vec{C}$)的元素。预测模式信息可以指示预测模式(例如,帧内预测模式或帧间预测模式)。可以计算两个向量(即,特征向量$\vec{F}$和分类向量$\overrightarrow{\{C'\}}$)之间的距离,并且然后将该距离与阈值{Kv′}进行比较。因此,可以基于阈值{Kv′}与作为特征向量$\vec{F}$与分类向量$\overrightarrow{\{C'\}}$之间的距离的比较来识别变换集或变换候选。

[0213] 在示例中,可以执行以下操作中的一个:(i)基于距离与从被包括在阈值集K_v中的阈值子集K_v'中选择的阈值(例如,{Kv′})的比较,从变换集子组确定变换集,(ii)基于距离与阈值(例如,{Kv′})的比较,从变换集子组确定变换候选,或者(iii)基于经编码信息中的索引从变换集子组选择变换集,并且基于距离与阈值(例如,{Kv′})的比较从所选择的变换集确定变换候选。可以计算特征向量$\vec{F}$(例如,变量的联合变化性)与从预定义的分类向量集$\vec{C}$中选择的分类向量$\overrightarrow{\{C'\}}$之间的距离。在示例中,计算特征向量$\vec{F}$(例如,变量的联合变化性)与从被包括在分类向量集$\vec{C}$中的分类向量子集$\vec{C}$中选择的分类向量$\overrightarrow{\{C'\}}$之间的距离。

[0214] 在示例中,该距离可以被确定为但不限于两个向量(即,特征向量$\vec{F}$和分类向量

$\overrightarrow{\{C'\}}$)之间的点积。替选地,该距离可以被确定为$\sum_{i=0}^{N-1}\left|\vec{F}(i)-\overrightarrow{C'(i)}\right|^{\alpha}$,其中,参数α是预先定义的常数并且参数N指示特征向量$\vec{F}$和分类向量$\overrightarrow{\{C'\}}$的维度(例如,大小)。

[0215] 在一个实施方式中,该比较可以确定但不限于(i)距离是否小于或等于阈值(例如,$\vec{F}.\overrightarrow{\{C'\}} \leq \{Kv'\}$)或(ii)距离是否大于或等于阈值(例如,$\vec{F}.\overrightarrow{\{C'\}} \geq \{Kv'\}$)其中{}表示对应集合的元素。

[0216] 在一些实施方式中,可以限制使用相邻重建样本(例如相邻重建样本的特征指示符)来从变换集子组中选择变换候选。对于使用经编码信息——例如块的预测模式(例如,帧内预测模式、帧间预测模式)——选择的某些变换集子组,变换候选或变换内核识别过程不使用相邻重建样本的特征指示符(例如,特征向量$\vec{F}$或特征标量S)。替代地,变换候选可以由在经编码视频比特流中用信号指示的一个或更多个相关联的索引来标识。在示例中,某些变换集子组中的每一个具有减少数目的可用变换集和/或变换内核。在示例中,上述限制适用但不限于某些预测模式,例如竖直SMOOTH模式SMOOTH_V模式和水平SMOOTH模式SMOOTH_H。因此,对于基于某些预测模式(例如,SMOOTH_V和SMOOTH_H)而选择的某些变换集子组,块的变换候选可以通过在经编码视频比特流中用信号指示的一个或更多个相关联的索引来标识并且不基于块的特征指示符来确定。

[0217] 图19示出了概述根据本公开内容的实施方式的过程(1900)的流程图。过程(1900)可以用于重建块,例如CB、CU、PB、TB、TU、亮度块(例如,亮度CB或亮度TB)、色度块(例如,色度CB或色度TB)等。在各种实施方式中,过程(1900)由处理电路执行,处理电路例如是终端设备(310)、(320)、(330)和(340)中的处理电路、执行视频编码器(403)的功能的处理电路、执行视频解码器(410)的功能的处理电路、执行视频解码器(510)的功能的处理电路、执行视频编码器(603)的功能的处理电路等。在一些实施方式中,过程(1900)以软件指令实现,因此当处理电路执行软件指令时,处理电路执行过程(1900)。处理在(S1901)处开始并且进行至(S1910)。

[0218] 在(S1910)处,可以基于当前图片的块的一个或更多个相邻块中的重建样本,例如从块的一个或更多个相邻块中的重建样本提取的特征指示符(例如,特征向量$\vec{F}$或特征标量S中的一个),从成组变换集确定块的变换候选。成组变换集中的每个变换集可以包括块的一个或更多个变换候选。一个或更多个相邻块可以在当前图片或与当前图片不同的重建图片中。

[0219] 在实施方式中,可以基于块的经编码信息中指示的块的预测模式(例如,帧内预测模式、帧间预测模式)从成组变换集中选择变换集子组。可以基于从块的一个或更多个相邻块中的重建样本提取的特征指示符(例如,特征向量$\vec{F}$或特征标量S中的一个)来从变换集子组确定变换候选。

[0220] 在示例中,可以基于从块的一个或更多个相邻块中的重建样本提取的特征指示符(例如,特征向量$\vec{F}$或特征标量S中的一个)来选择变换集子组中的一个变换集。可以基于在经编码信息中用信号指示的索引从变换集子组中的所选择的一个变换集确定块的变换候

选。

[0221] 在示例中,可以基于经编码信息中用信号指示的第二索引来选择变换集子组的一个变换集。可以基于从块的一个或更多个相邻块中的重建样本提取的特征指示符(例如,特征向量$\vec{F}$或特征标量S中的一个)来从变换集子组的所选择的一个变换集确定块的变换候选。

[0222] 在示例中,可以基于从块的一个或更多个相邻块中的重建样本提取的特征指示符(例如,特征向量$\vec{F}$或特征标量S中的一个)从变换集子组隐式地确定变换候选。

[0223] 在示例中,基于对块的一个或更多个相邻块中的重建样本的统计分析来确定特征指示符(例如,特征向量$\vec{F}$或特征标量S)。此外,可以执行以下操作中的一个:(i)基于特征指示符(例如,特征向量$\vec{F}$或特征标量S)从变换集子组确定变换集,(ii)基于特征指示符(例如,特征向量$\vec{F}$或特征标量S)从变换集子组确定变换候选,(iii)基于经编码信息中的第二索引从变换集子组中选择变换集,并且基于特征指示符(例如,特征向量$\vec{F}$或特征标量S)从所选择的变换集确定变换候选。

[0224] 在示例中,特征指示符是特征标量S并且被确定为指示块的一个或更多个相邻块中的重建样本的样本值的变量的矩。可以预先定义阈值集K_s。因此,可以执行以下操作中的一个:(i)基于变量的矩和从阈值集K_s选择的阈值从变换集子组确定变换集,(ii)基于变量的矩和阈值从变换集子组确定变换候选,或者(iii)基于经编码信息中的第二索引从变换集子组中选择变换集,并且基于变量的矩和阈值从所选择的变换集确定变换候选。

[0225] 在示例中,变量的矩是变量的一阶矩、变量的二阶矩或变量的三阶矩中的一个。块的预测模式是多个预测模式中的一个。多个预测模式中的每一个对应于阈值集K_s中的唯一阈值子集K_s',其指示多个预测模式与阈值集K_s中的多个阈值子集之间的单射映射。

[0226] 基于(i)块的块大小、(ii)量化参数或(iii)块的预测模式中的一个从阈值集K_s中选择阈值。

[0227] 在示例中,特征指示符是特征向量$\vec{F}$并且被确定为分别指示块的左侧的相邻列的重建样本的样本值和块的上侧的相邻行的重建样本的样本值的变量的联合变化性。预先定义分类向量集$\vec{C}$和与分类向量集$\vec{C}$相关联的阈值集K_v。计算变量的联合变化性与从被包括在分类向量集$\vec{C}$中的分类向量子集$\vec{C'}$选择的分类向量之间的距离。因此,可以执行以下操作中的一个:

[0228] (i)基于距离与从被包括在阈值集K_v中的子集K_v'中选择的阈值的比较,从变换集子组确定变换集,(ii)基于距离与阈值K_v的比较,从变换集子组确定变换候选,或者(iii)基于经编码信息中的第二索引从变换集子组中选择变换集,并且基于距离与阈值的比较从所选择的变换集确定变换候选。

[0229] 在(S1920)处,可以基于确定的变换候选重建块的样本。

[0230] 可以适当地调整过程(1900)。可以修改和/或省略过程(1900)中的(一个或多个)步骤。可以添加(一个或多个)附加步骤。可以使用任何合适的实现顺序。

[0231] 本公开内容中的实施方式可以单独使用或以任何顺序组合使用。此外,方法(或实施方式)、编码器和解码器中的每一个均可以通过处理电路(例如,一个或更多个处理器或者一个或更多个集成电路)来实现。在一个示例中,一个或更多个处理器执行存储在非暂态计算机可读介质中的程序。本公开内容中的实施方式可以应用于亮度块或色度块。

[0232] 可以使用计算机可读指令将上述技术实现为计算机软件,并且将其物理地存储在一个或更多个计算机可读介质中。例如,图20示出了适于实现所公开的主题的某些实施方式的计算机系统(2000)。

[0233] 计算机软件可以使用任何合适的机器代码或计算机语言来编码,其可以经受汇编、编译、链接等机制以创建包括指令的代码,所述指令可以由一个或更多个计算机中央处理单元(central processing unit,CPU)、图形处理单元(Graphics Processing Unit,GPU)等直接执行,或者通过解释、微代码执行等来执行。

[0234] 指令可以在各种类型的计算机或其部件上执行,包括例如个人计算机、平板计算机、服务器、智能电话、游戏设备、物联网设备等。

[0235] 图20中所示的计算机系统(2000)的部件本质上是示例性的,并且不旨在对实现本公开内容的实施方式的计算机软件的使用范围或功能提出任何限制。也不应该将部件的配置解释为具有与计算机系统(2000)的示例性实施方式中示出的任一部件或其组合相关的任何依赖性或要求。

[0236] 计算机系统(2000)可以包括某些人机接口输入设备。这样的人机接口输入设备可以响应于一个或更多个人类用户通过例如触觉输入(诸如:键击、滑动、数据手套移动)、音频输入(诸如:语音、拍手)、视觉输入(诸如:手势)、嗅觉输入(未描绘)的输入。人机接口设备还可以用于捕获不一定与人的有意识输入直接相关的某些媒体,诸如音频(诸如:语音、音乐、环境声音)、图像(诸如:扫描图像、从静止图像摄像机获得的摄影图像)、视频(诸如:二维视频、包括立体视频的三维视频)。

[0237] 输入人机接口设备可以包括以下中的一个或更多个(每个仅描绘一个):键盘(2001)、鼠标(2002)、轨迹板(2003)、触摸屏(2010)、数据手套(未示出)、操纵杆(2005)、麦克风(2006)、扫描仪(2007)、摄像机(2008)。

[0238] 计算机系统(2000)还可以包括某些人机接口输出设备。这样的人机接口输出设备可以通过例如触觉输出、声音、光和气味/味道来刺激一个或更多个人类用户的感觉。这样的人机接口输出设备可以包括触觉输出设备(例如,通过触摸屏(2010)、数据手套(未示出)或操纵杆(2005)的触觉反馈,但是还可以存在不用作输入设备的触觉反馈设备)、音频输出设备(诸如:扬声器(2009)、耳机(未描绘))、视觉输出设备(诸如屏幕(2010),以包括CRT屏幕、LCD屏幕、等离子屏幕、OLED屏幕,每一者具有或不具有触摸屏输入能力,每一者具有或不具有触觉反馈能力——其中的一些可以通过诸如立体图形输出、虚拟现实眼镜(未描绘)、全息显示器和烟箱(未描绘)的装置来输出二维视觉输出或多于三维输出),以及打印机(未描绘)。

[0239] 计算机系统(2000)还可以包括人类可访问的存储设备及其相关联的介质,诸如包括具有CD/DVD等介质(2021)的CD/DVD ROM/RW(2020)的光学介质、拇指驱动器(2022)、可移动硬盘驱动器或固态驱动器(2023)、诸如磁带和软盘的传统磁介质(未描绘)、诸如安全软件狗的基于专用ROM/ASIC/PLD的设备(未描绘)等。

[0240] 本领域技术人员还应当理解,结合本发明所公开的主题使用的术语“计算机可读介质”不包括传输介质、载波或其他瞬态信号。

[0241] 计算机系统(2000)还可以包括至一个或更多个通信网络(2055)的接口(2054)。网络可以例如是无线的、有线的、光学的。网络还可以是本地的、广域的、城市的、车载的和工业的、实时的、延迟容忍的等。网络的示例包括:诸如以太网的局域网;无线LAN;包括GSM(Global System for Mobile Communication)、3G(Third Generation)、4G(Fourth Generation)、5G(Fifth Generation)、LTE(Long Term Evolution)等的蜂窝网络;包括有线电视、卫星电视和地面广播电视的电视有线连接或无线广域数字网络;包括CAN总线的车载和工业网络等。某些网络通常需要外部网络接口适配器,其附接至某些通用数据端口或外围总线(2049)(例如,计算机系统(2000)的USB端口);其他网络通常通过附接至如下所述的系统总线而集成到计算机系统(2000)的核(例如以太网接口集成到PC计算机系统或蜂窝网络接口集成到智能电话计算机系统)中。使用这些网络中的任何一个,计算机系统(2000)可以与其他实体通信。这样的通信可以是单向仅接收的(例如,广播电视)、单向仅发送的(例如,至某些CAN总线设备的CAN总线)、或双向的,例如,到使用局域或广域数字网络的其他计算机系统。可以在如上所述的这些网络和网络接口中的每个网络和网络接口上使用某些协议和协议栈。

[0242] 上述人机接口设备、人可访问的存储设备以及网络接口可以附接至计算机系统(2000)的核(2040)。

[0243] 核(2040)可以包括一个或更多个中央处理单元(Central Processing Unit,CPU)(2041)、图形处理单元(Graphics Processing Unit,GPU)(2042)、现场可编程门区(Field Programmable Gate Area,FPGA)形式的专用可编程处理单元(2043)、用于某些任务的硬件加速器(2044)等。这些设备,连同只读存储器(Read-only memory,ROM)(2045)、随机存取存储器

[0244] (2046)、诸如内部非用户可访问硬盘驱动器的内部大容量存储装置、SSD

[0245] (2047)等,可以通过系统总线(2048)连接。在一些计算机系统中,系统总线(2048)可以以一个或更个物理插头的形式访问,以使得能够通过附加的CPU、GPU等来扩展。外围设备可以直接或通过外围总线(2049)附接至核的系统总线(2048)。在示例中,屏幕(2010)可以连接至图形适配器(2050)。外围总线的架构包括PCI、USB等。

[0246] CPU(2041)、GPU(2042)、FPGA(2043)和加速器(2044)可以执行某些指令,这些指令组合起来可以构成上述计算机代码。该计算机代码可以存储在ROM(2045)或RAM(2046)中。过渡数据也可以存储在RAM(2046)中,而永久数据可以存储在例如内部大容量存储装置(2047)中。可以通过使用高速缓冲存储器来实现对任何存储器设备的快速存储和检索,所述高速缓冲存储器可以与一个或更多个CPU(2041)、GPU(2042)、大容量存储装置(2047)、ROM(2045)、RAM(2046)等紧密关联。

[0247] 计算机可读介质上可以具有用于执行各种计算机实现的操作的计算机代码。介质和计算机代码可以是为本公开的目的而专门设计和构造的介质和计算机代码,或者它们可以是计算机软件领域的技术人员公知和可用的类型的。

[0248] 作为示例而非限制,具有架构(2000)的计算机系统,并且尤其是核(2040)可以由于处理器(包括CPU、GPU、FPGA、加速器等)执行包含在一个或更多个有形计算机可读介质中

的软件而提供功能。这样的计算机可读介质可以是与如上所述的用户可访问大容量存储装置以及诸如核内部大容量存储装置(2047)或ROM(2045)的具有非暂态性的核(2040)的特定存储装置相关联的介质。实现本公开内容的各种实施方式的软件可以存储在这样的设备中并且由核(2040)执行。根据特定需要,计算机可读介质可以包括一个或更多个存储器设备或芯片。软件可以使核(2040)并且具体地使其中的处理器(包括CPU、GPU、FPGA等)执行本文描述的特定过程或特定过程的特定部分,包括定义存储在RAM(2046)中的数据结构以及根据由软件定义的过程修改这样的数据结构。另外或作为替选,计算机系统可以提供作为硬连线或以其他方式包含在电路(例如:加速器(2044))中的逻辑的结果的功能,该逻辑可以代替软件或与软件一起操作以执行本文描述的特定过程或特定过程的特定部分。在适当的情况下,对软件的引用可以包含逻辑,反之亦然。在适当的情况下,对计算机可读介质的引用可以包括存储用于执行的软件的电路(例如集成电路(integrated circuit,IC))、包含用于执行的逻辑的电路、或存储用于执行的软件的电路以及包含用于执行的逻辑的电路两者。本公开内容包含硬件和软件的任何合适的组合。

[0249] 附录A:首字母缩写

[0250] JEM:联合探索模型

[0251] VVC:通用视频编码

[0252] BMS:基准集

[0253] MV:运动向量

[0254] HEVC:高效视频编码

[0255] SEI:补充增强信息

[0256] VUI:视频可用性信息

[0257] GOP:图片组

[0258] TU:变换单元

[0259] PU:预测单元

[0260] CTU:编码树单元

[0261] CTB:编码树块

[0262] PB:预测块

[0263] HRD:假设参考解码器

[0264] SNR:信噪比

[0265] CPU:中央处理单元

[0266] GPU:图形处理单元

[0267] CRT:阴极射线管

[0268] LCD:液晶显示器

[0269] OLED:有机发光二极管

[0270] CD:致密盘

[0271] DVD:数字视频光盘

[0272] ROM:只读存储器

[0273] RAM:随机存取存储器

[0274] ASIC:专用集成电路

[0275] PLD:可编程逻辑设备

[0276] LAN:局域网

[0277] GSM:全球移动通信系统

[0278] LTE:长期演进

[0279] CANBus:控制器局域网总线

[0280] USB:通用串行总线

[0281] PCI:外围部件互连

[0282] FPGA:现场可编程门区

[0283] SSD:固态驱动器

[0284] IC:集成电路

[0285] CU:编码单元

[0286] 虽然本公开内容已经描述了若干示例性实施方式,但是存在落入本公开内容的范围内的改变、置换和各种替代等同物。因此,应当理解,本领域技术人员将能够设计出虽然未在本文明确示出或描述但体现本公开内容的原理并且因此在本公开内容的精神和范围内的许多系统和方法。

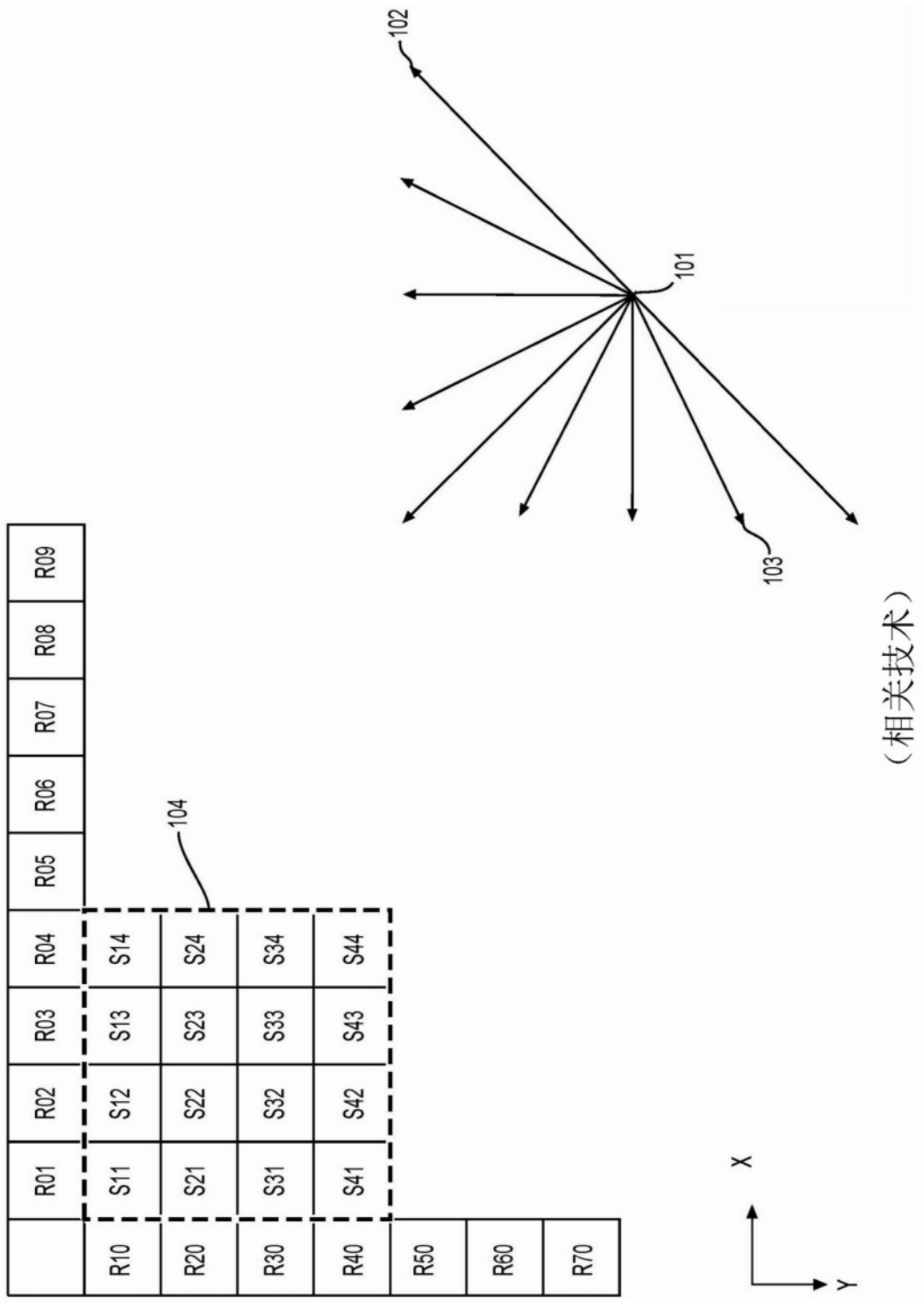

102
101
103
104
R09
R08
R07
R06
R05
R04
R03
R02
R01
S14
S24
S34
S44
S13
S23
S33
S43
S12
S22
S32
S42
S11
S21
S31
S41
R10
R20
R30
R40
R50
R60
R70
X
Y
（相关技术）

图1A

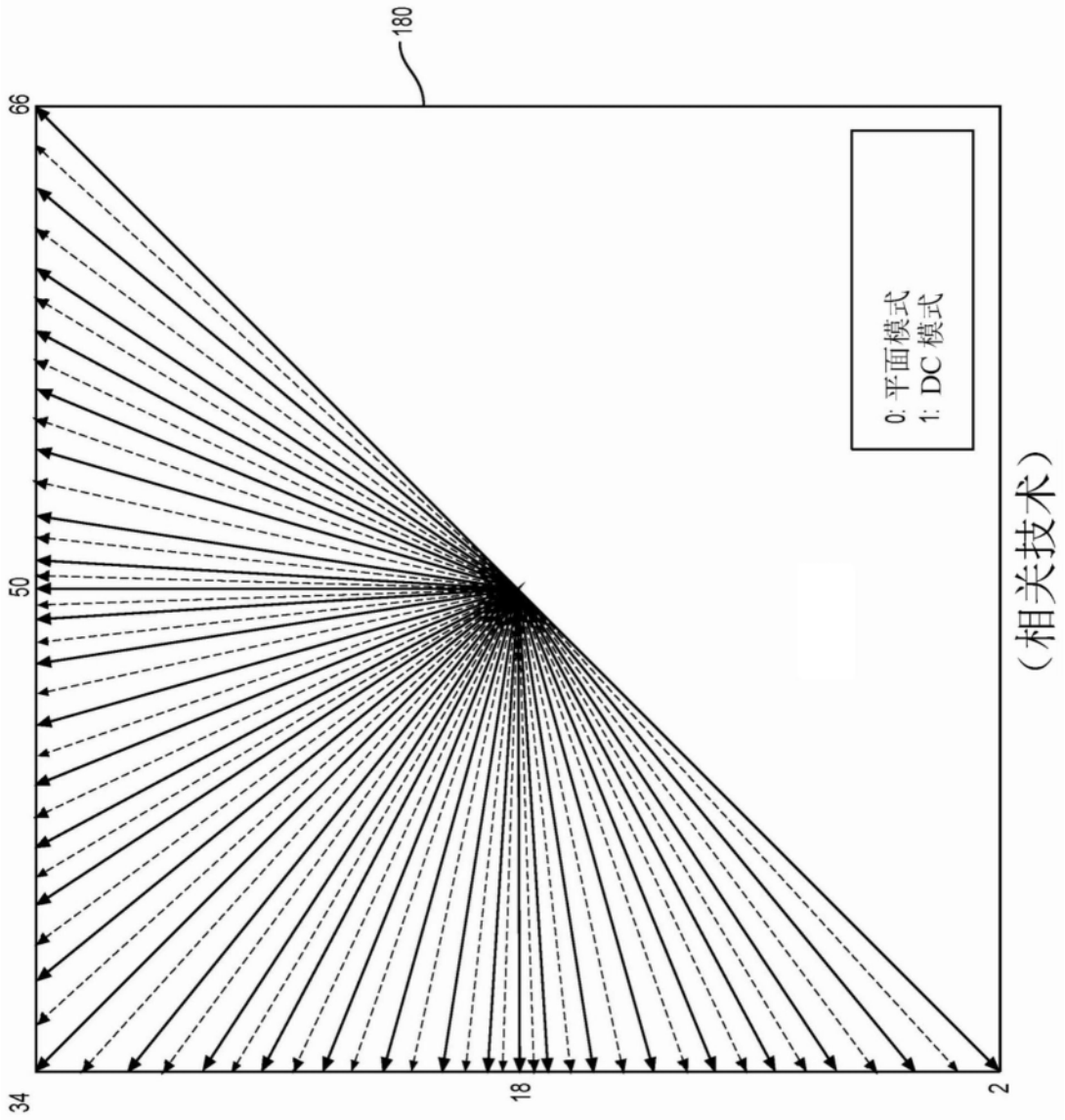


图1B

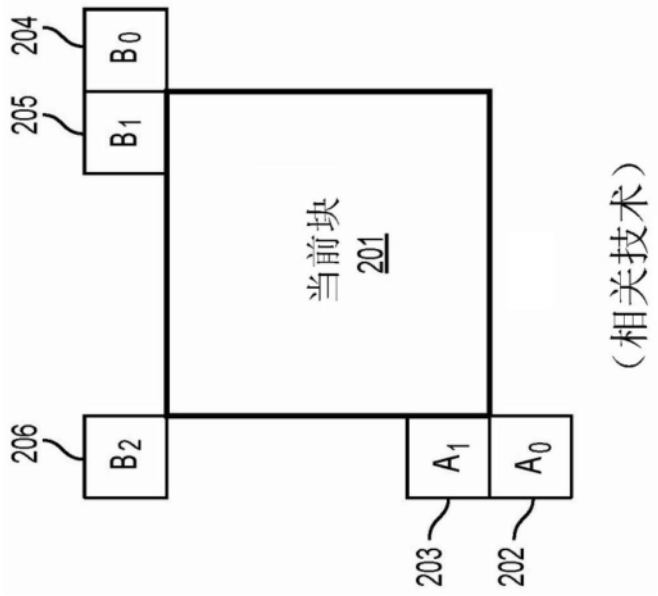


图2

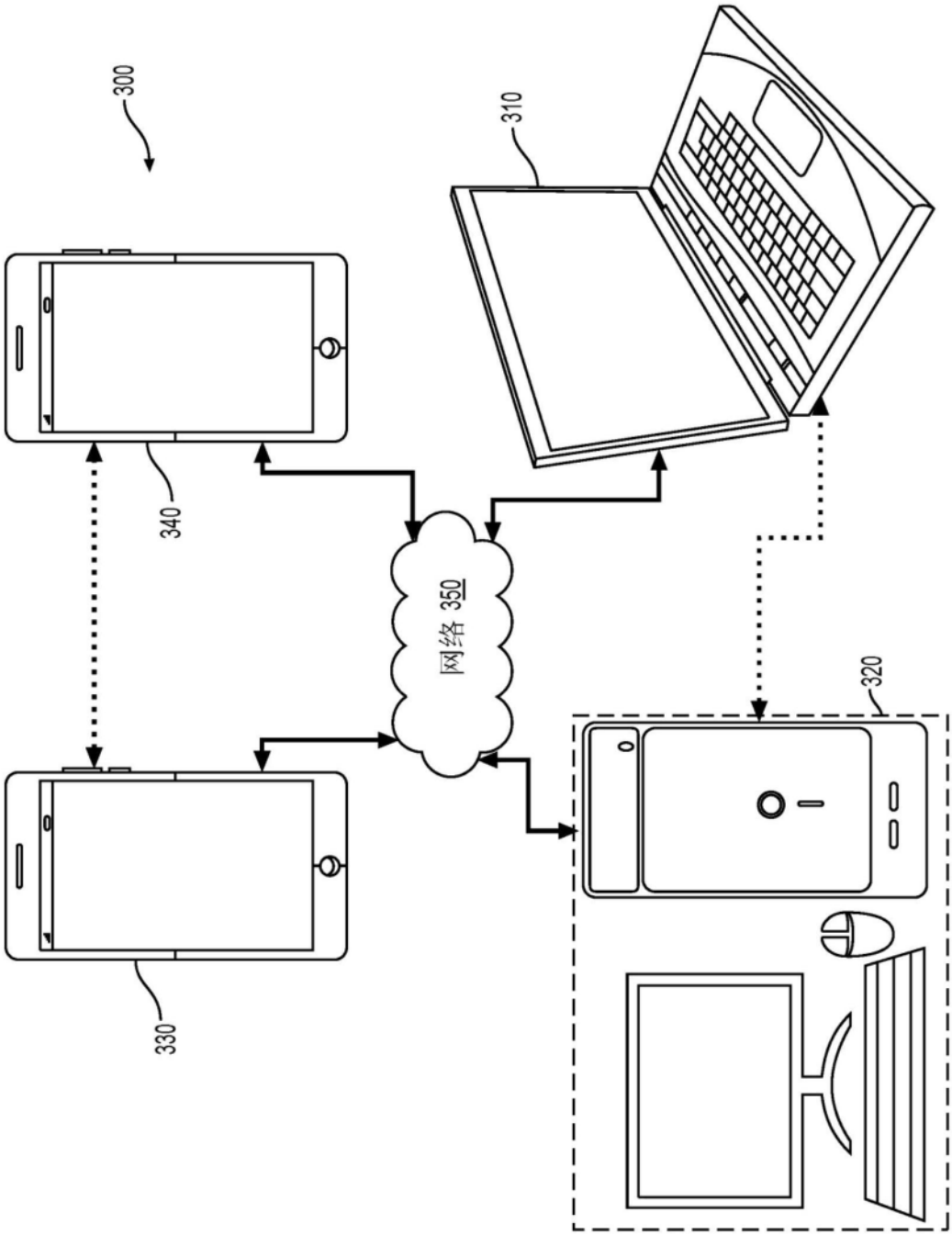


图3

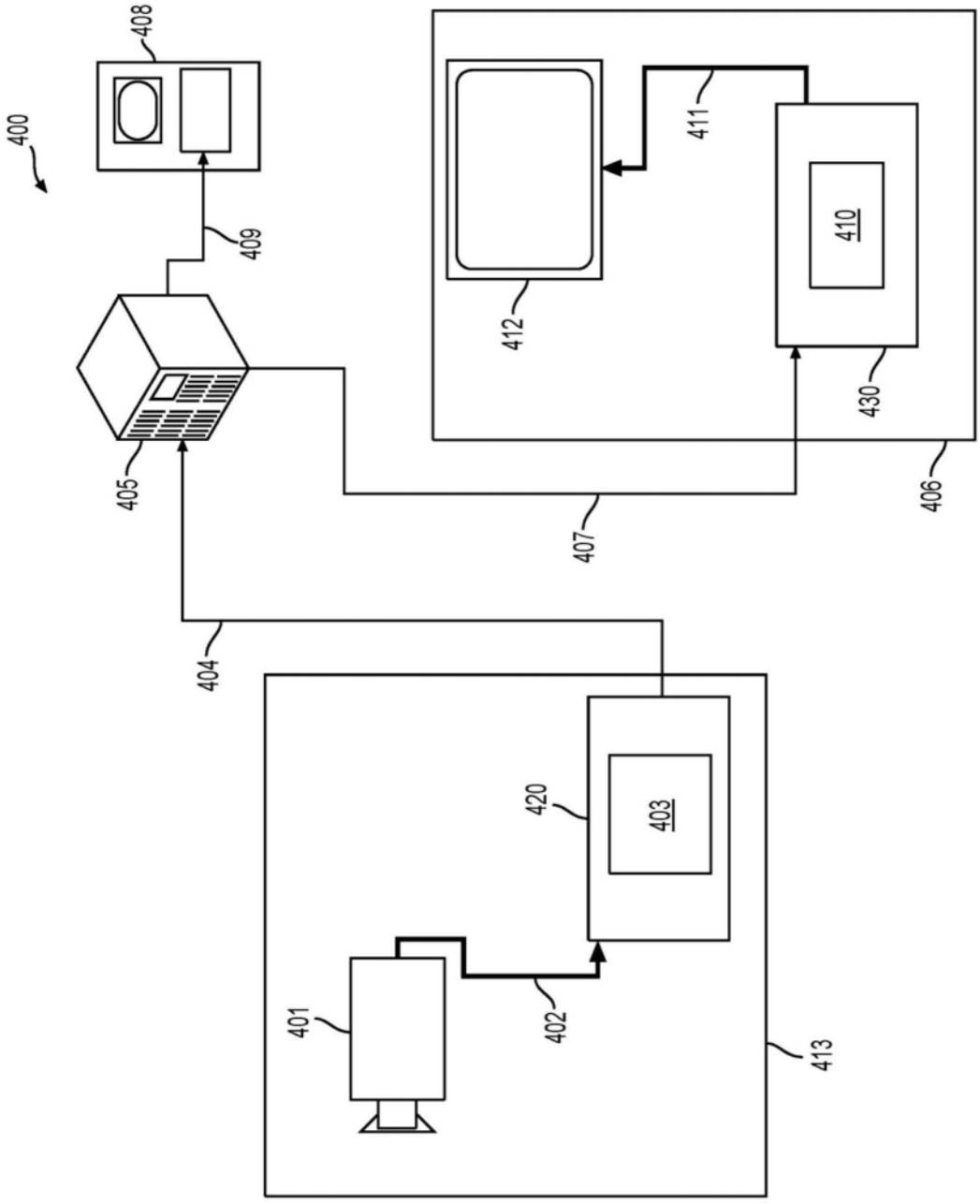


图4

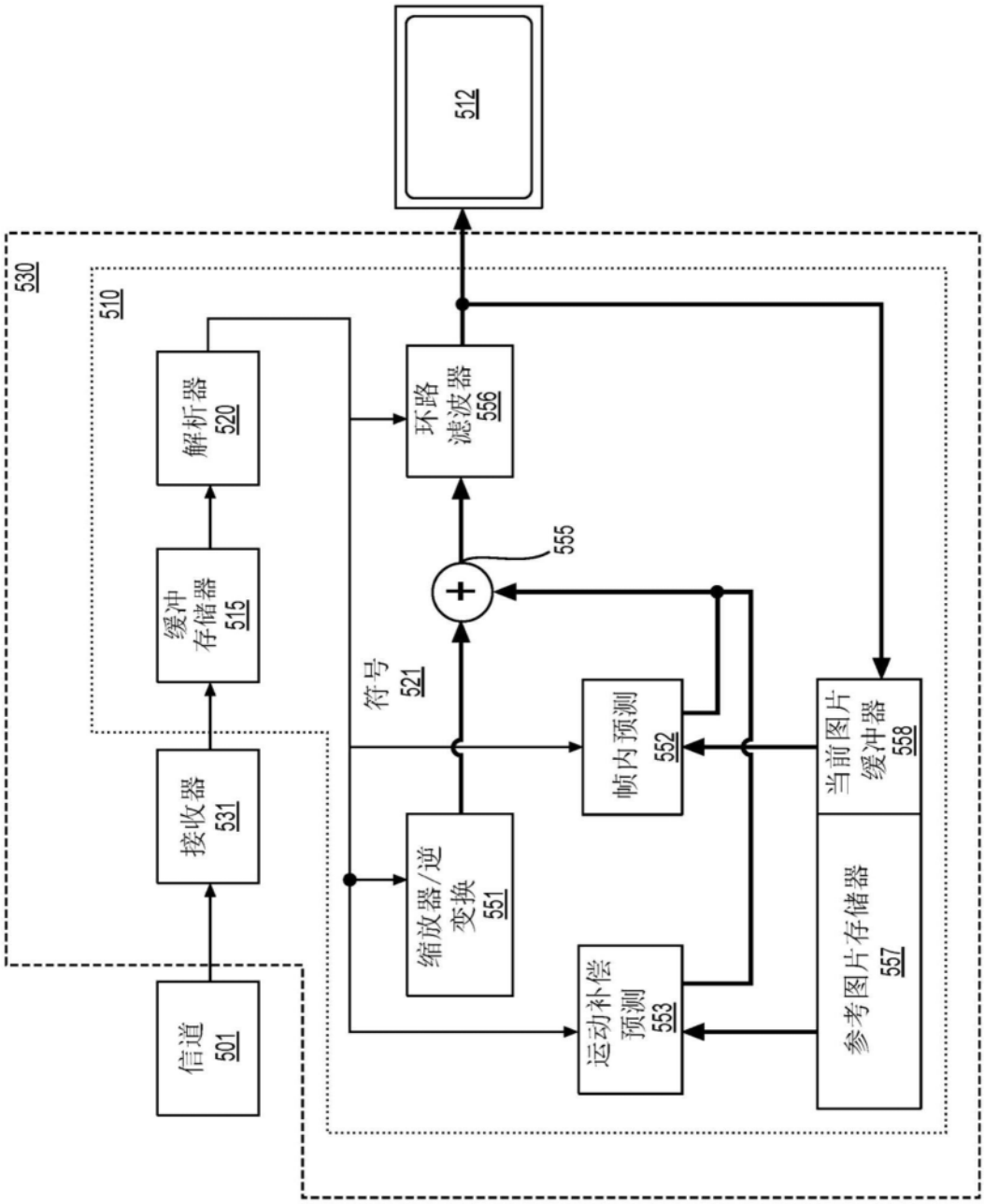

512
530
510
解析器
520
环路
滤波器
556
555
缓冲
存储器
515
符号
521
帧内预测
552
当前图片
缓冲器
558
接收器
531
缩放器/逆
变换
551
运动补偿
预测
553
参考图片存储器
557
信道
501

图5

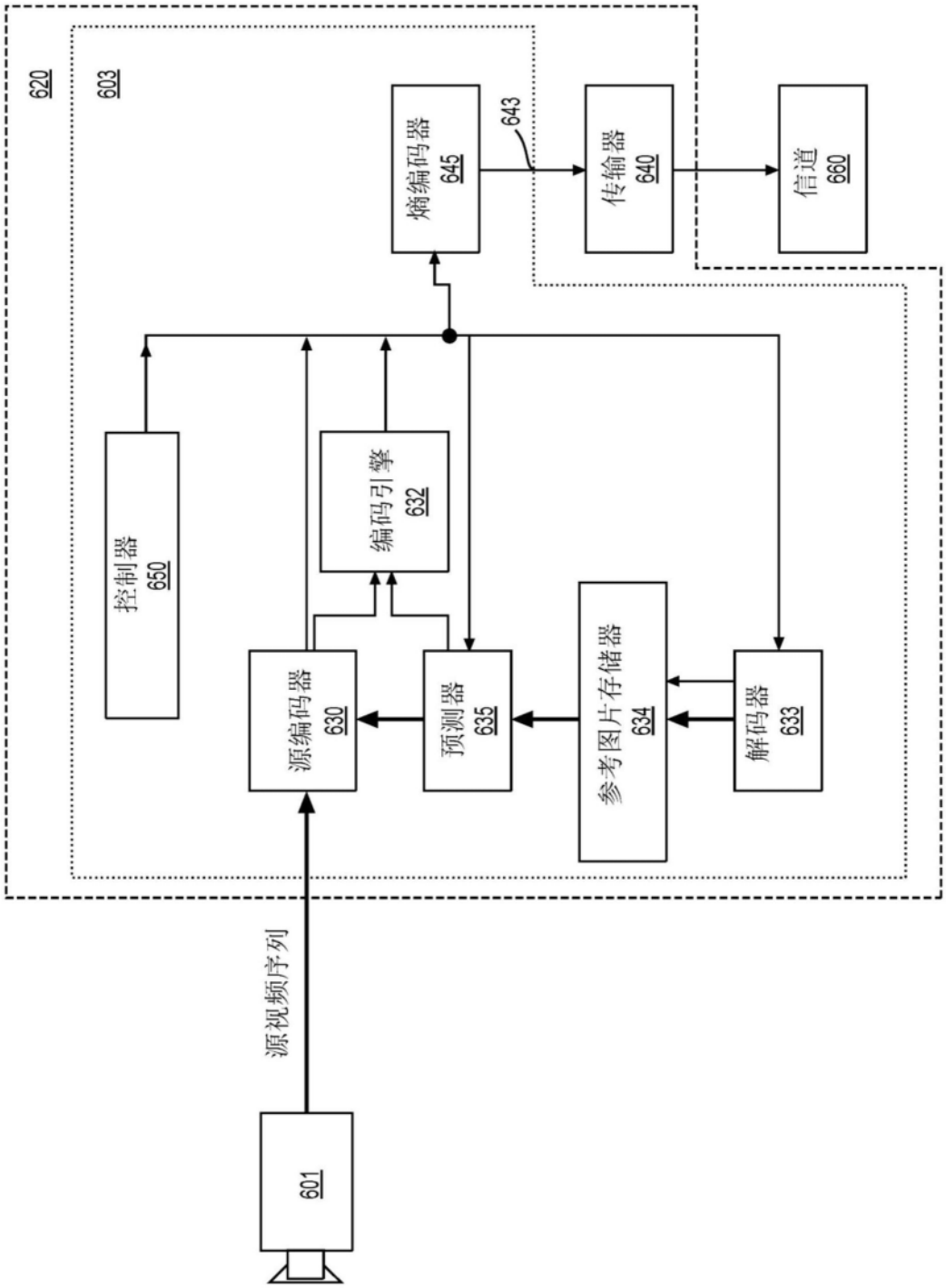

620
603
熵编码器
645
643
传输器
640
信道
660
控制器
650
编码引擎
632
源编码器
630
预测器
635
参考图片存储器
634
解码器
633
源视频序列
601

图6

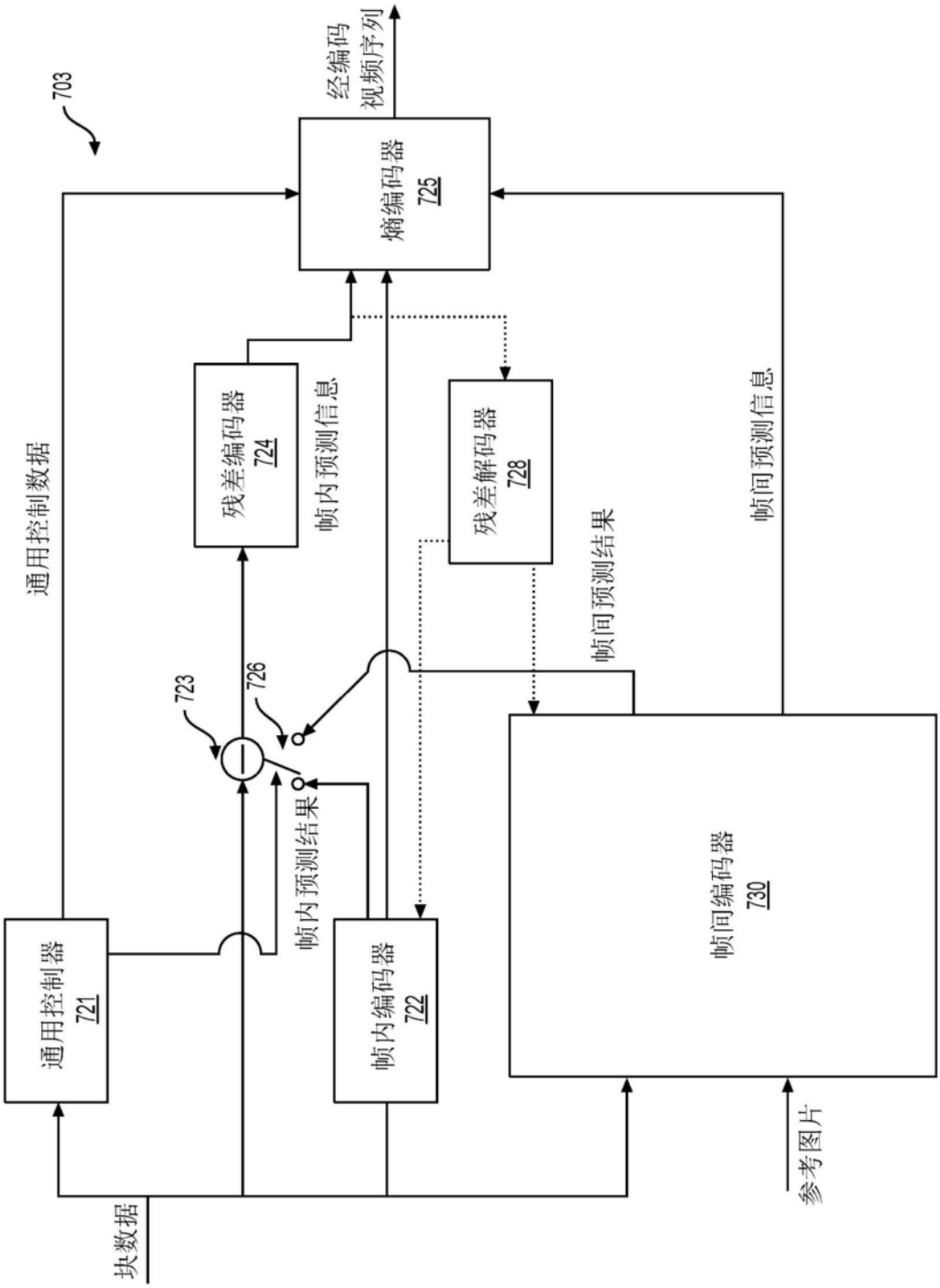

703
经编码
视频序列
熵编码器
725
通用控制数据
残差编码器
724
帧内预测信息
残差解码器
728
帧间预测结果
帧间预测信息
723
726
帧内预测结果
通用控制器
721
帧内编码器
722
帧间编码器
730
块数据
参考图片

图7

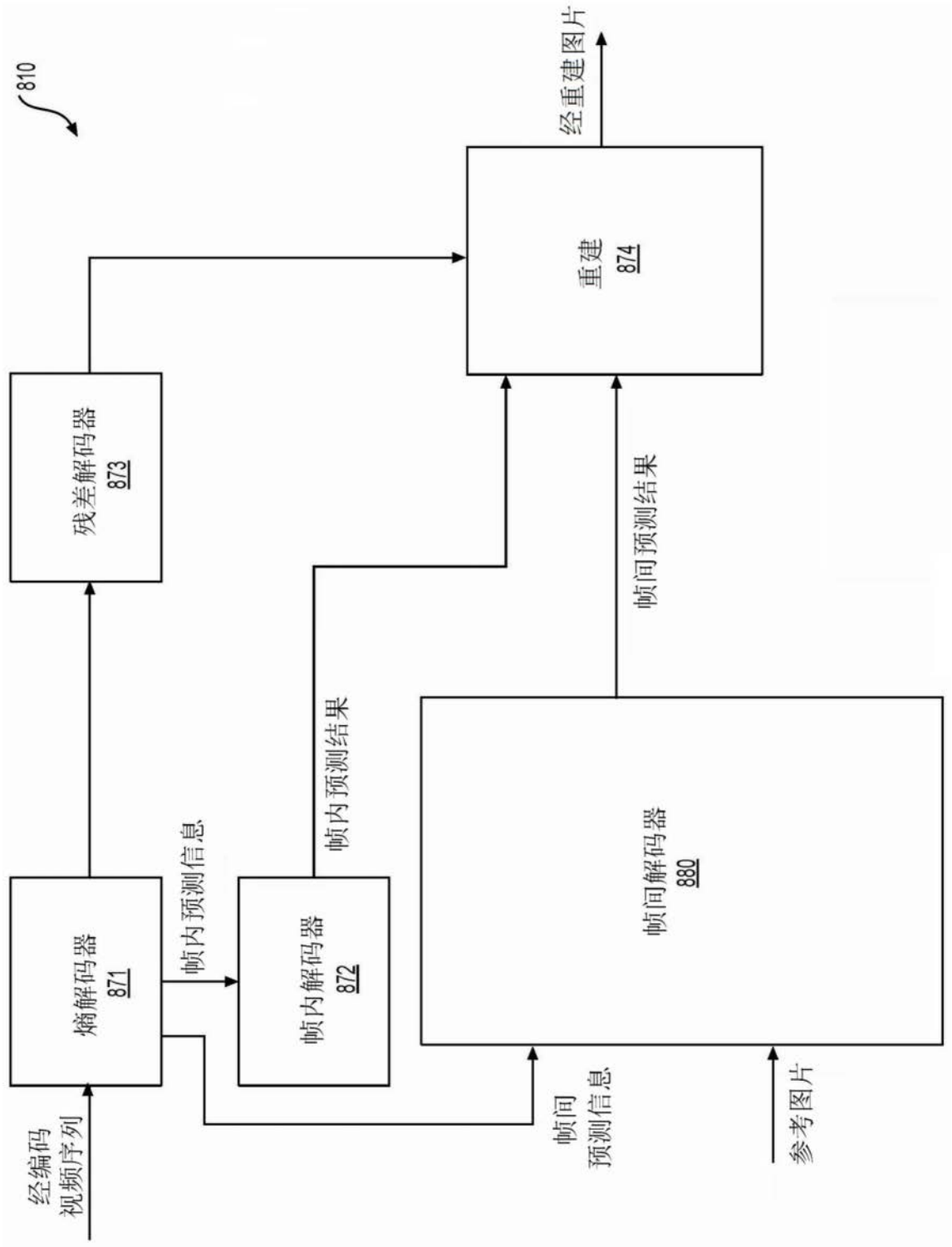

810
经重建图片
重建
874
残差解码器
873
帧间预测结果
帧内预测结果
帧间解码器
880
熵解码器
871
帧内预测信息
帧内解码器
872
经编码
视频序列
帧间
预测信息
参考图片

图8

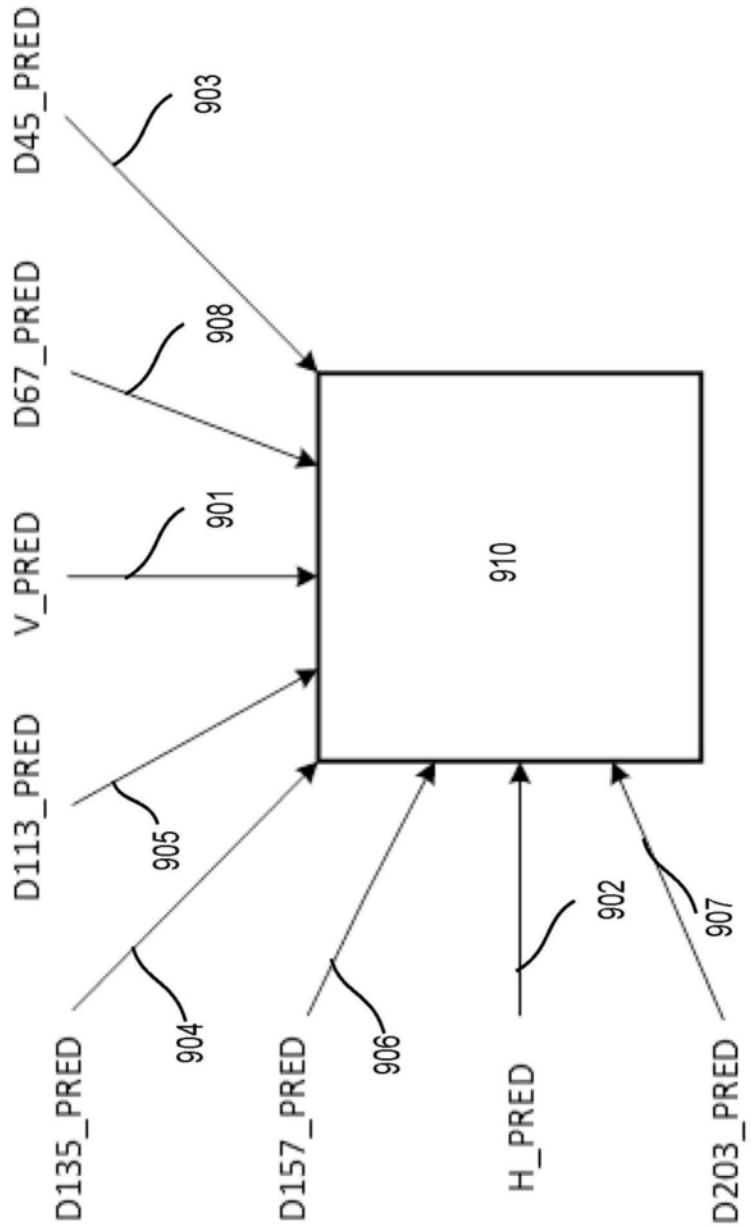

D45_PRED
903
D67_PRED
908
V_PRED
901
D113_PRED
905
D135_PRED
904
D157_PRED
906
H_PRED
902
D203_PRED
907
910

图9

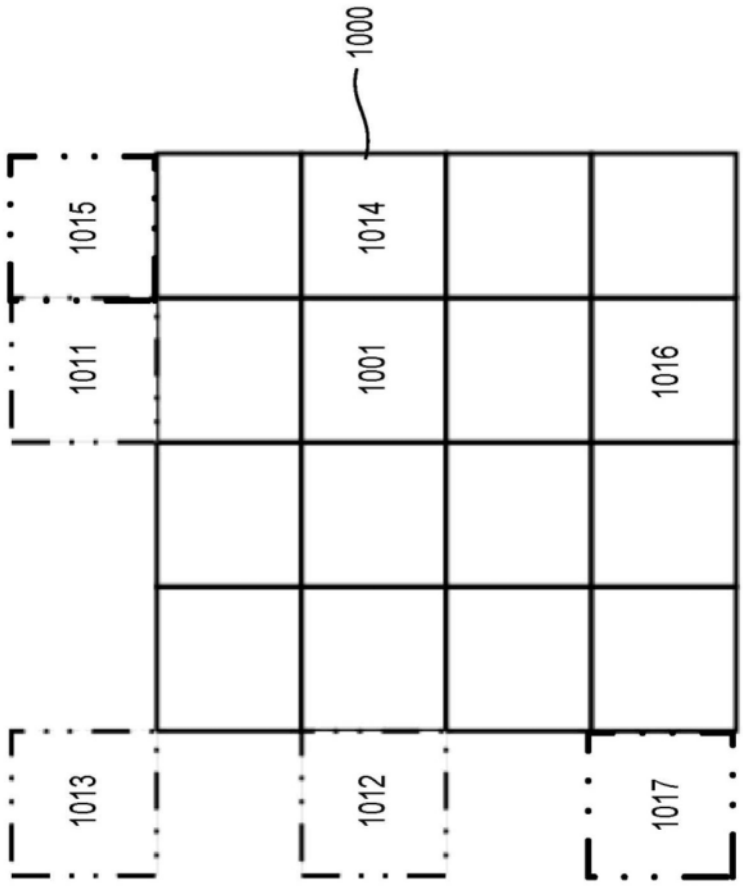


图10

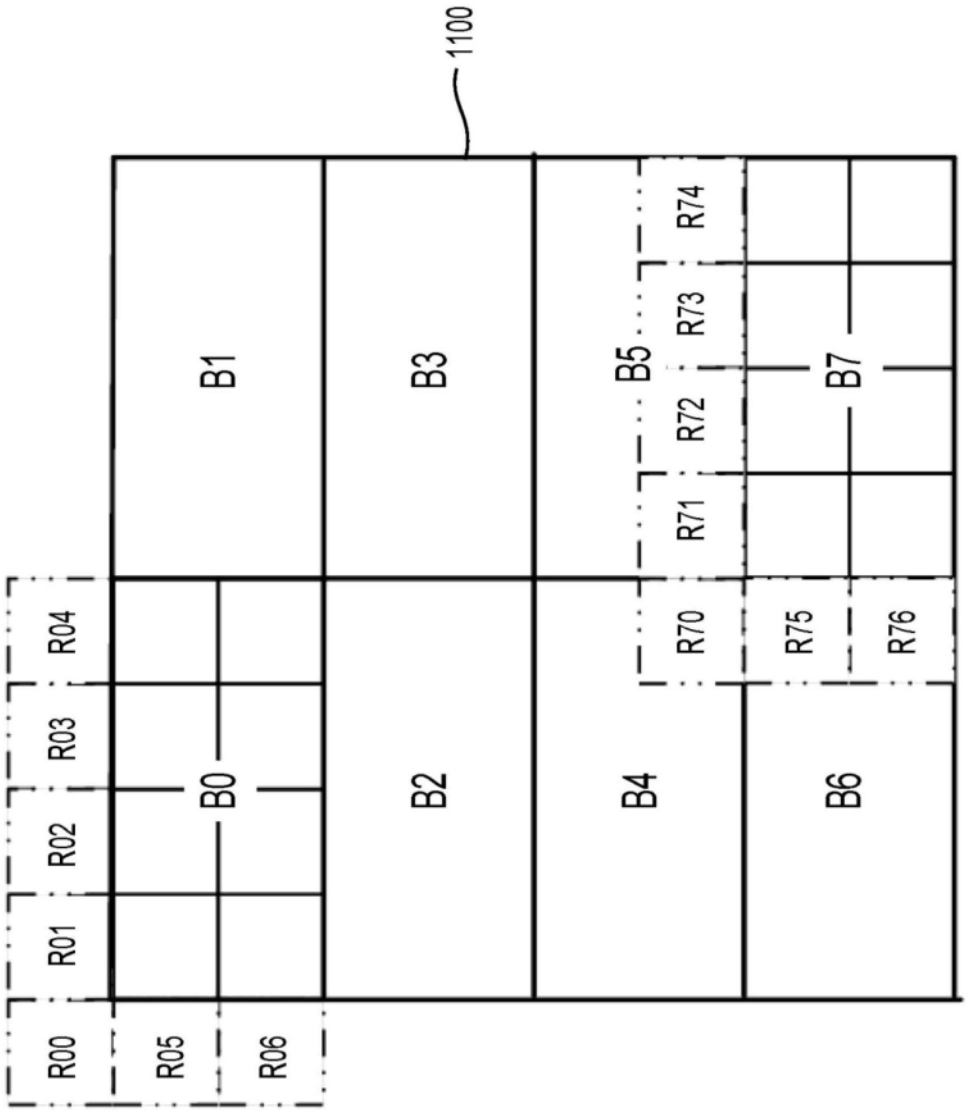

1100
B1
B3
B5
B7
B0
B2
B4
B6
R00
R01
R02
R03
R04
R05
R06
R70
R71
R72
R73
R74
R75
R76

图11

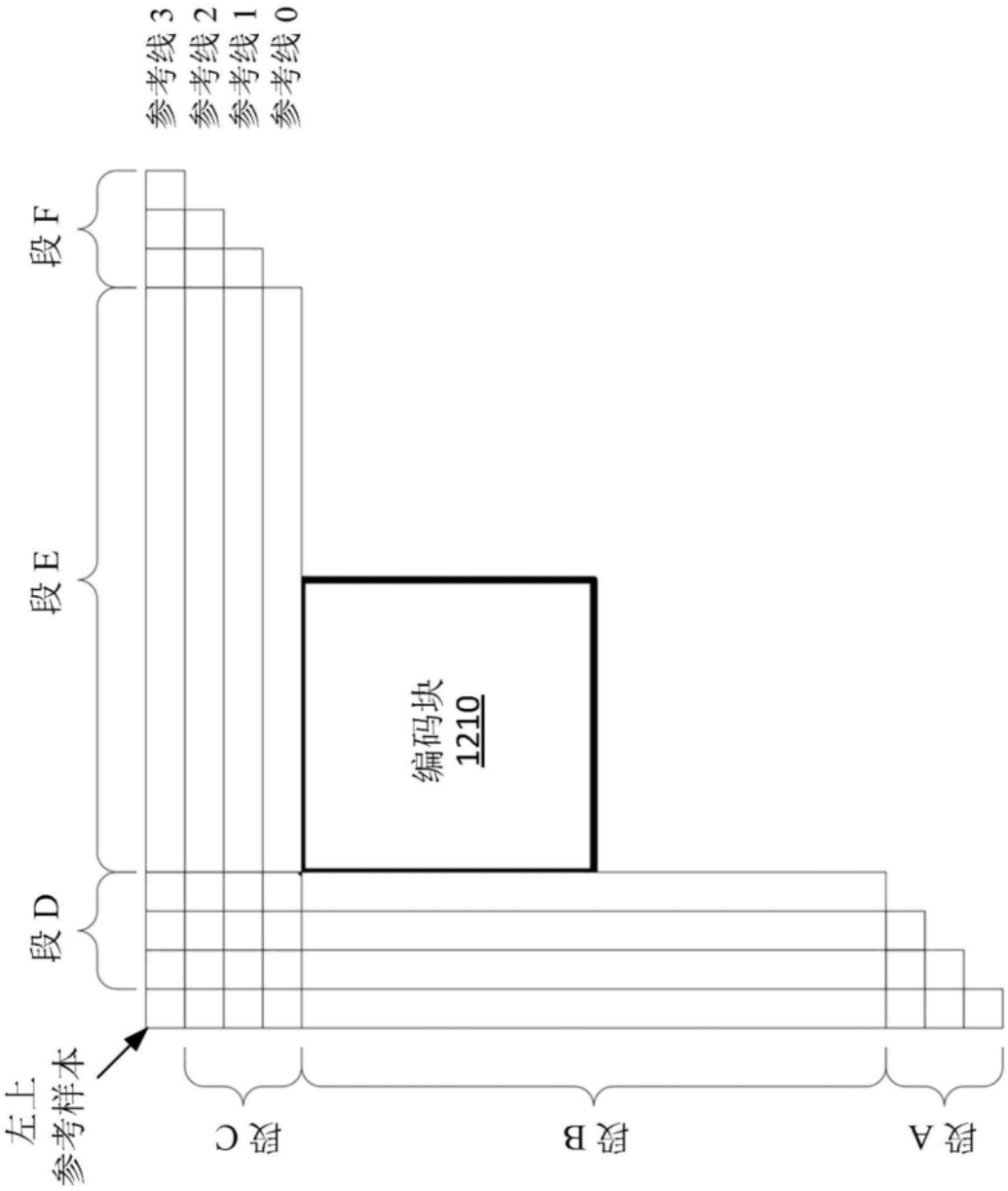


图12

变换类型	基函数 $T_i(j)$, $i,j=0,1,\ldots,N-1$
DCT-2	$T_i(j)=\omega_0\cdot\sqrt{\frac{2}{N}}\cdot\cos\left(\frac{\pi\cdot i\cdot(2j+1)}{2N}\right)$ 其中 $\omega_0=\begin{cases}\sqrt{\frac{2}{N}} & i=0\\ 1 & i\neq 0\end{cases}$
DST-4	$T_i(j)=\sqrt{\frac{2}{N}}\cdot\sin\left(\frac{\pi\cdot(2i+1)\cdot(2j+1)}{4N}\right)$
DST-7	$T_i(j)=\sqrt{\frac{4}{2N+1}}\cdot\sin\left(\frac{\pi\cdot(2i+1)\cdot(j+1)}{2N+1}\right)$

图13

变换类型	描述	预测模式	
		帧内	帧间
DCT_DCT	DCT ↓ 与 →	✓(所有块大小)	✓(所有块大小)
ADST_DCT	ADST ↓; DCT →	✓（块大小 ≤ 16×16）	✓（块大小 ≤ 16×16）
DCT_ADST	DCT ↓; ADST →		
ADST_ADST	ADST ↓ 与 →		
FLIPADST_DCT	FLIPADST ↓; DCT →	×	✓（块大小 ≤ 16×16）
DCT_FLIPADST	DCT ↓; FLIPADST →		
FLIPADST_FLIPADST	FLIPADST ↓ 与 →		
ADST_FLIPADST	ADST ↓; FLIPADST →		
FLIPADST_ADST	FLIPADST ↓; ADST →		
IDTX	IDTX ↓ 与 →	✓（块大小 ≤ 16×16）	✓（块大小 ≤ 32×32）
V_DCT	DCT ↓; IDTX →	✓（块大小 < 16×16）	✓（块大小 ≤ 16×16）
H_DCT	IDTX ↓; DCT →		
V_ADST	ADST ↓; IDTX →	×	✓（块大小 < 16×16）
H_ADST	IDTX ↓; ADST →		
V_FLIPADST	FLIPADST ↓; IDTX →	×	✓（块大小 < 16×16）
H_FLIPADST	IDTX ↓; FLIPADST →		

1410

1420

图14A

帧内预测	竖直变换	水平变换
DC_PRED	DCT	DCT
V_PRED	ADST	DCT
H_PRED	DCT	ADST
D45_PRED	DCT	DCT
D135_PRED	ADST	ADST
D113_PRED	ADST	DCT
D157_PRED	DCT	ADST
D203_PRED	DCT	ADST
D67_PRED	ADST	DCT
SMOOTH_PRED	ADST	ADST
SMOOTH_V_PRED	ADST	DCT
SMOOTH_H_PRED	DCT	ADST
PAETH_PRED	ADST	ADST

图14B

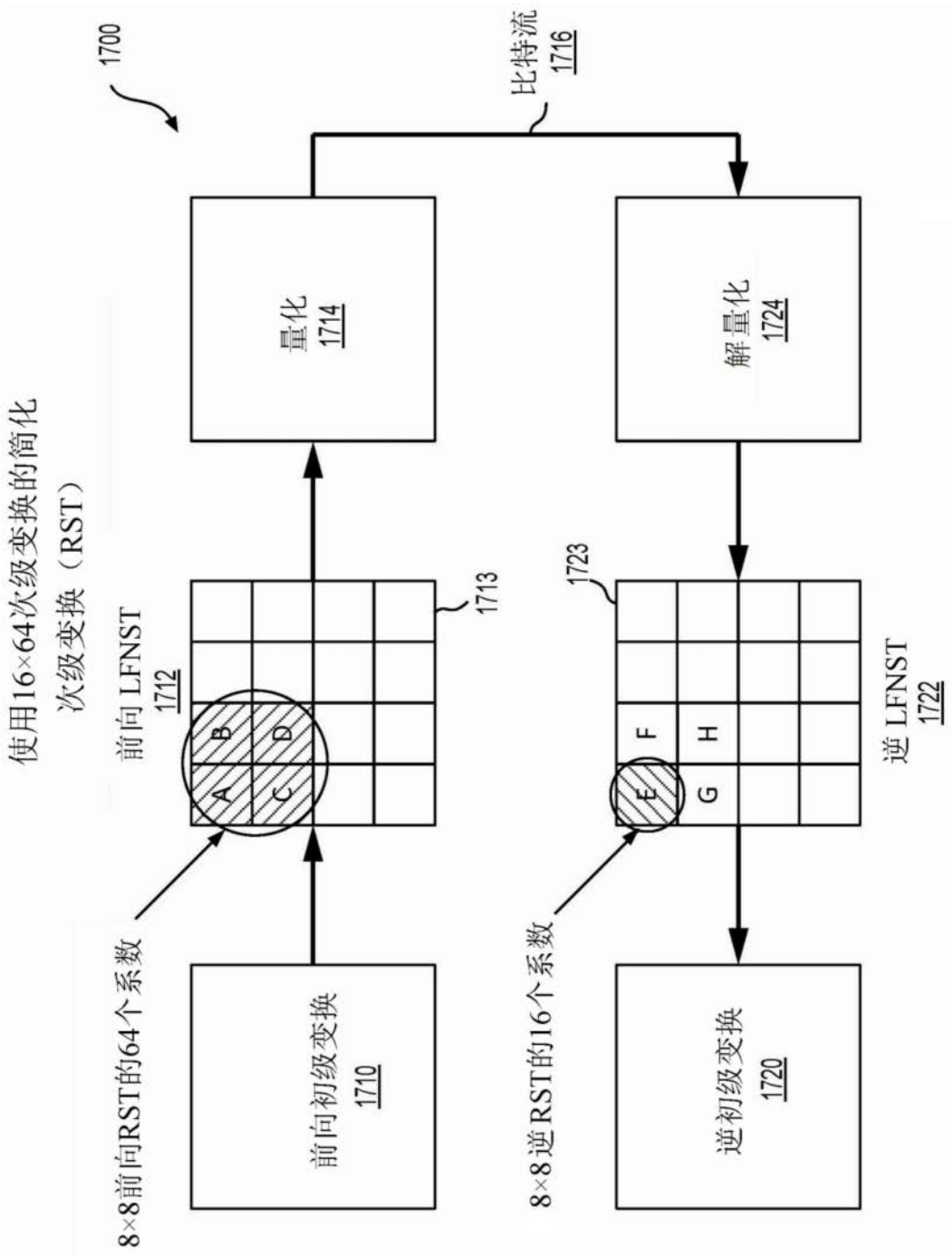

使用16×64次级变换的简化次级变换（RST）
1700
比特流
1716
量化
1714
解量化
1724
前向 LFNST
1712
1713
1723
逆 LFNST
1722
A
B
C
D
E
F
G
H
8×8前向RST的64个系数
8×8逆RST的16个系数
前向初级变换
1710
逆初级变换
1720

图15

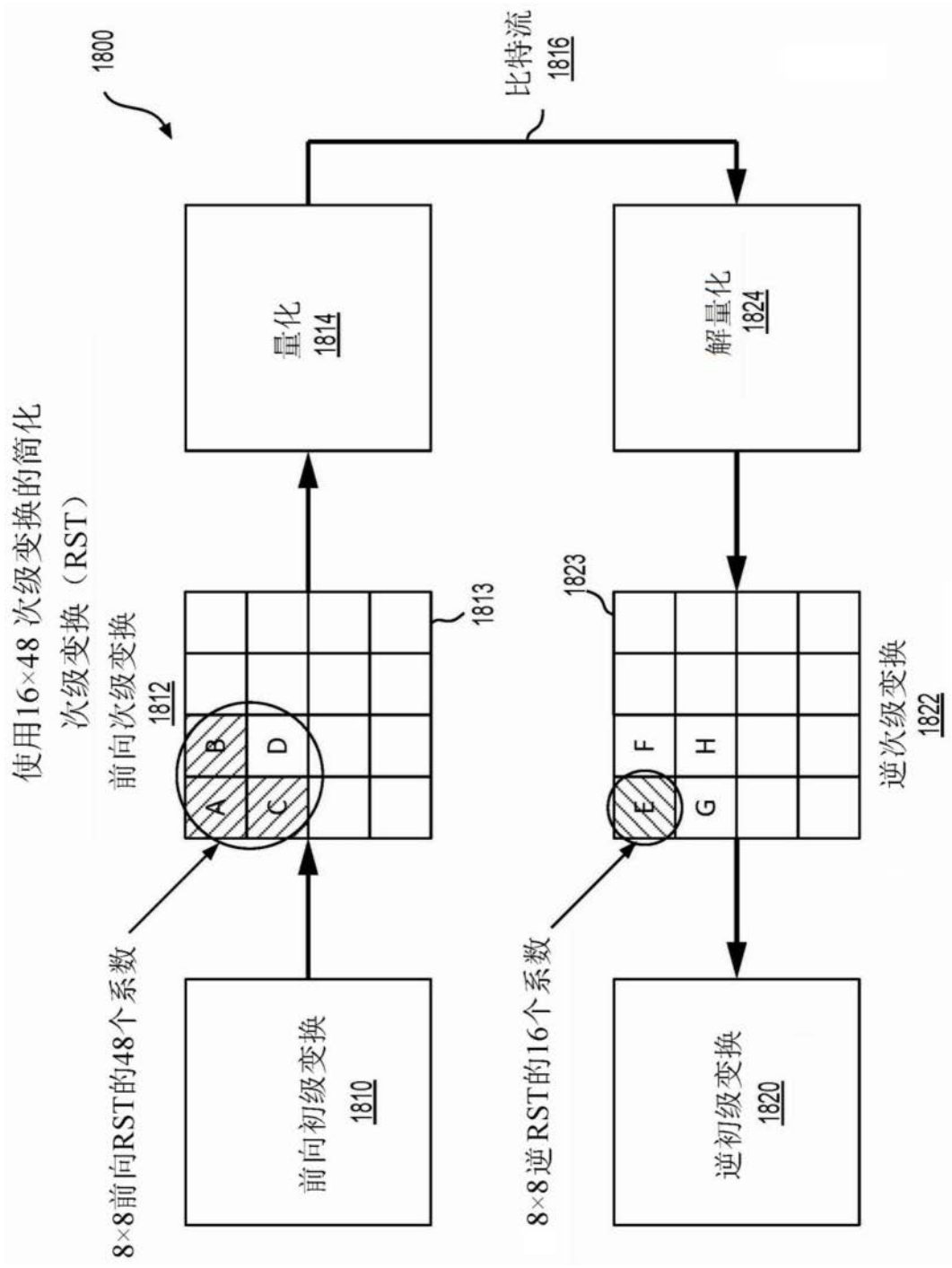

1800
比特流
1816
量化
1814
解量化
1824
使用16×48次级变换的简化次级变换（RST）
前向次级变换
1812
1813
1823
逆次级变换
1822
A
B
C
D
E
F
G
H
8×8前向RST的48个系数
8×8逆RST的16个系数
前向初级变换
1810
逆初级变换
1820

图16

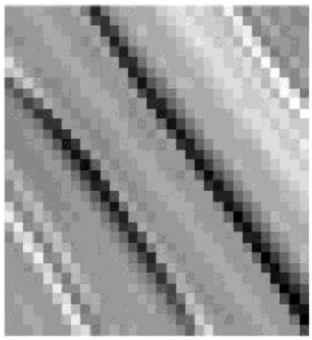


图17A

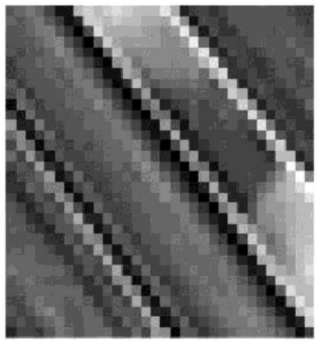


图17B

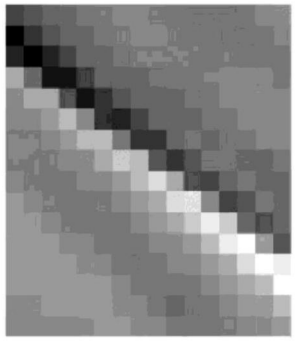


图17C

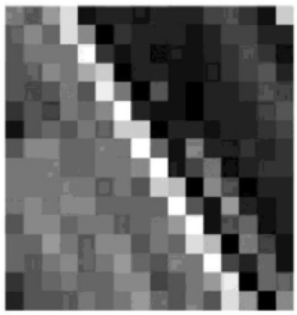


图17D

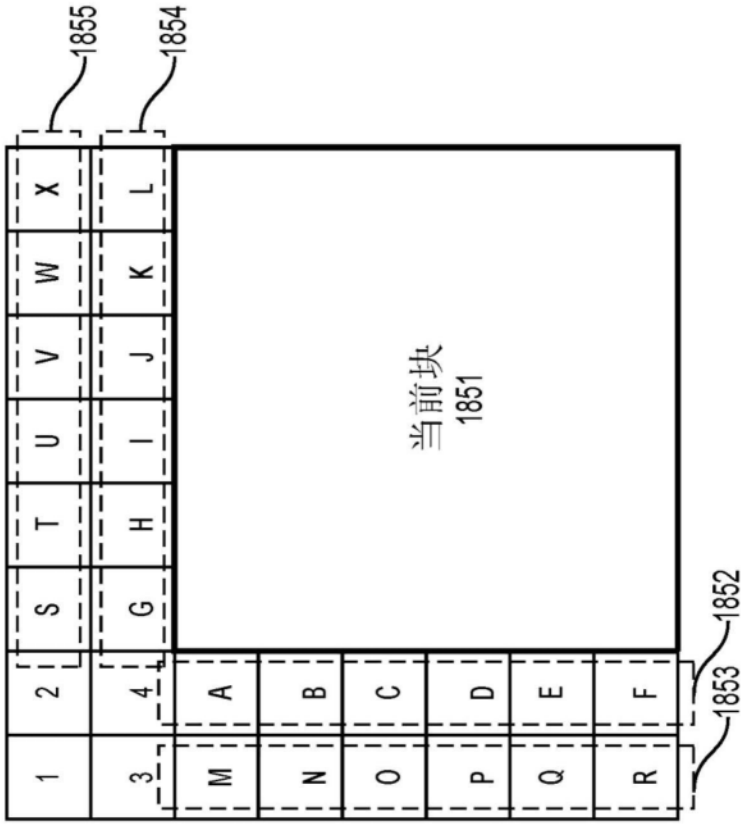

1855
1854
X
W
V
U
T
S
L
K
J
I
H
G
当前块
1851
1852
1853
2
4
1
3
A
B
C
D
E
F
M
N
O
P
Q
R

图18

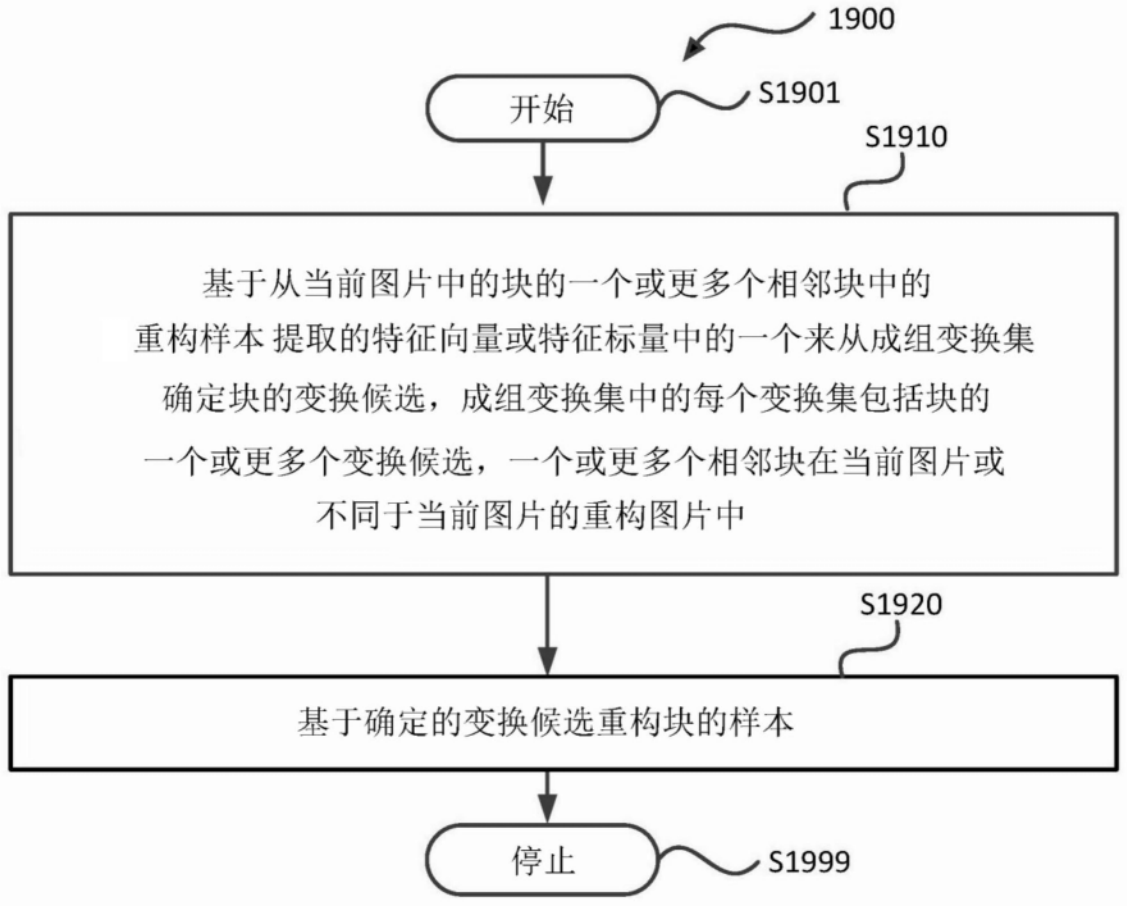

1900
开始
S1901
S1910
基于从当前图片中的块的一个或更多个相邻块中的
重构样本 提取的特征向量或特征标量中的一个来从成组变换集
确定块的变换候选，成组变换集中的每个变换集包括块的
一个或更多个变换候选，一个或更多个相邻块在当前图片或
不同于当前图片的重构图片中
S1920
基于确定的变换候选重构块的样本
停止
S1999

图19

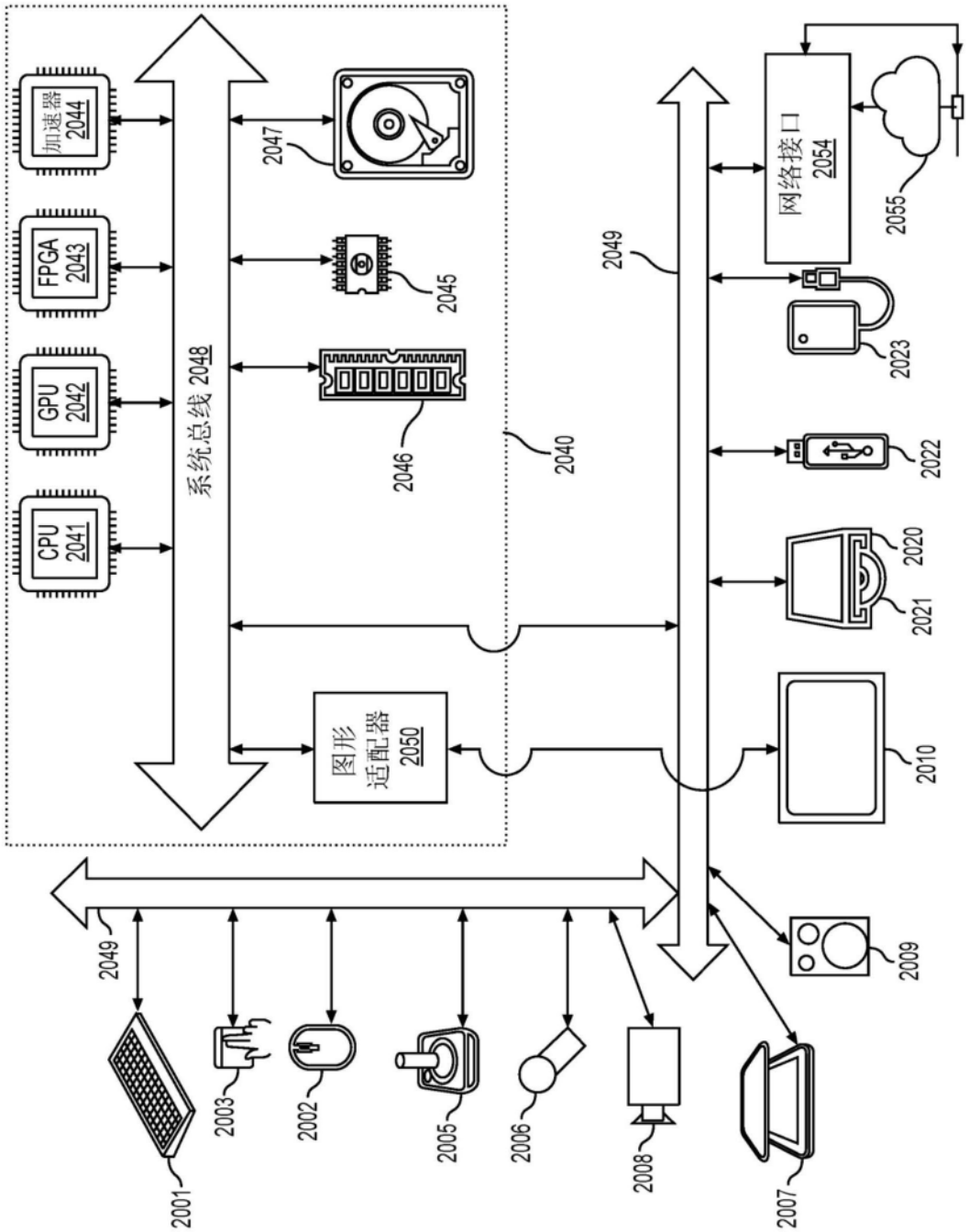

加速器 2044
FPGA 2043
GPU 2042
CPU 2041
系统总线 2048
2047
2045
2046
2040
图形适配器 2050
2049
网络接口 2054
2055
2023
2022
2020
2021
2010
2009
2001
2003
2002
2005
2006
2008
2007

图20