(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 116325751 B
(45)授权公告日 2024.10.15

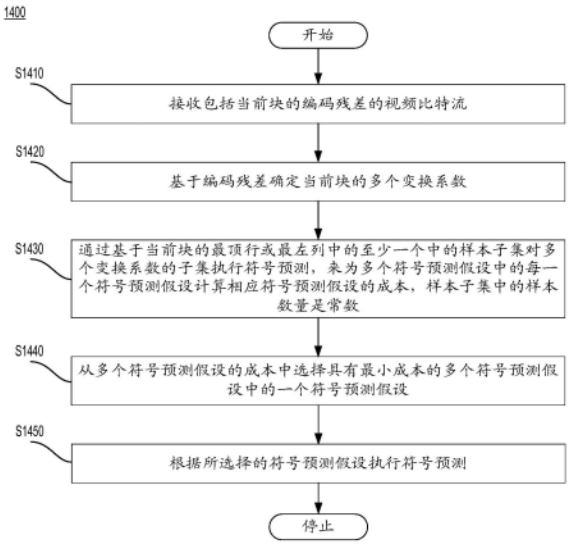
(21)申请号 202180062088.8
(22)申请日 2021.10.04
(65)同一申请的已公布的文献号
申请公布号 CN 116325751 A
(43)申请公布日 2023.06.23
(30)优先权数据
63/217,021 2021.06.30 US
17/490,562 2021.09.30 US
(85)PCT国际申请进入国家阶段日
2023.03.09
(86)PCT国际申请的申请数据
PCT/US2021/053399 2021.10.04
(87)PCT国际申请的公布数据
WO2023/277938 EN 2023.01.05
(73)专利权人 腾讯美国有限责任公司
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 欧阳祥 李翔 刘杉
(74)专利代理机构 北京康信知识产权代理有限责任公司 11240
专利代理师 周婷婷
(51)Int.Cl.
H04N 19/61(2006.01)
(56)对比文件
WO 2019172802 A1,2019.09.12
Felix Henry (Orange,B-Com).《Residual Coefficient Sign Prediction》.《Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 : JVET-D0031》.2016,正文第1-5页.
Cheung Auyeung, Xiang Li, Shan Liu.《EE2-related: Low Complexity Sign Prediction》.《Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29 : JVET-W0108-v4》.2021,正文第1-3页.
审查员 师佳彬
权利要求书1页 说明书23页 附图16页

(54)发明名称

视频解码的方法、装置及存储介质

(57)摘要

本公开的各方面包括用于视频解码的方法、装置和非暂时性计算机可读存储介质。一种装置包括处理电路,该处理电路接收包括当前块的编码残差的视频比特流。处理电路基于编码残差确定当前块的多个变换系数。处理电路通过基于当前块的最顶行或最左列中的至少一个中的样本子集对多个变换系数的子集执行符号预测,来为多个符号预测假设中的每一个符号预测假设计算相应符号预测假设的成本。样本子集中的样本数量是常数。处理电路选择具有最小成本的一个符号预测假设。处理电路根据所选择的符号预测假设执行符号预测。



CN 116325751 B

1.一种视频解码的方法,其特征在于,所述方法包括:

接收视频比特流;

确定所述视频比特流中当前块的多个变换系数;

从所述当前块的最顶行中的重构样本中确定出第一样本子集,并从所述当前块的最左列中的重构样本中确定出第二样本子集,其中,所述第一样本子集中的第一样本在所述最顶行中的位置是均匀间隔的,或所述第一样本子集中的至少两个第一样本在所述最顶行中的位置是相邻的,所述第二样本子集中的第二样本在所述最左列中的位置是均匀间隔的,或所述第二样本子集中的至少两个第二样本在所述最左列中的位置是相邻的;

通过所述第一样本子集和所述第二样本子集中的至少之一对所述多个变换系数的子集执行符号预测,来为多个符号预测假设中的每一个符号预测假设计算相应符号预测假设的成本,所述第一样本子集中的样本数量是小于所述当前块的最顶行中样本总数的第一常数,所述第二样本子集中的样本数量是小于所述当前块的最左列中样本总数的第二常数;

从所述多个符号预测假设的成本中选择具有最小成本的所述多个符号预测假设中的一个符号预测假设;以及

根据所选择的符号预测假设执行所述符号预测。

2.根据权利要求1所述的方法,其特征在于,所述第一常数和所述第二常数是预先确定的。

3.根据权利要求1所述的方法,其特征在于,基于所述多个变换系数的子集的数量来确定所述第一常数和所述第二常数。

4.根据权利要求1所述的方法,其特征在于,基于所述当前块的宽度来确定所述第一样本子集中第一样本的位置,并基于所述当前块的高度来确定所述第二样本子集中第二样本的位置。

5.根据权利要求1所述的方法,其特征在于,所述计算包括:

基于(i)所述当前块的所述第一样本子集和所述第二样本子集中的至少之一,以及(ii)所述当前块上方的行或所述当前块左侧的列中的至少一个中的相邻重构样本,来为所述多个符号预测假设中的每一个符号预测假设计算所述相应符号预测假设的成本。

6.一种视频解码装置,其特征在于,包括:

存储器,用于存储指令;

处理器,用于调用所述存储器存储的指令实现权利要求1-5任一项所述的方法。

7.一种存储指令的非暂时性计算机可读存储介质,其特征在于,当由至少一个处理器执行时,所述指令使得所述至少一个处理器执行权利要求1-5任一项所述的方法。

视频解码的方法、装置及存储介质

[0001] 相关申请

[0002] 本申请要求2021年9月30日提交的题为“用于视频编码的方法和装置”的美国专利申请号17/490,562的权益,该申请要求2021年6月30日提交的题为“低复杂信号预测”的美国临时申请号63/217,021的优先权。在先申请的公开内容在此全部通过引用结合于此。

技术领域

[0003] 本公开描述了总体上涉及视频编码的实施例。

背景技术

[0004] 这里提供的背景描述是为了总体呈现本公开的上下文。在该背景技术部分中所描述的工作的程度上,目前署名的发明人的工作以及在申请时可能不符合现有技术的描述的方面,既不明确也不暗示地承认是本公开的现有技术。

[0005] 可以使用具有运动补偿的帧间图片预测来执行视频编码和解码。未压缩的视频可以包括一系列图片,每个图片具有例如1920×1080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固定或可变的图像速率(非正式地也称为帧速率),例如,每秒60幅图片或60Hz。未压缩的视频有重要的比特率要求。例如,每样本8比特的1080p60 4:2:0视频(60Hz帧速率的1920x1080亮度样本分辨率)需要接近1.5Gbit/s的带宽。一小时这样的视频需要超过600GB的存储空间。

[0006] 视频编码和解码的一个目的是通过压缩来减少输入视频信号中的冗余。压缩有助于降低上述带宽或存储空间需求,在某些情况下可降低两个数量级或更多。可以采用无损压缩和有损压缩及其组合。无损压缩是指可以从压缩的原始信号中重构出原始信号的精确副本的技术。当使用有损压缩时,重构信号可能与原始信号不相同,但是原始信号和重构信号之间的失真足够小,使得重构信号对预期应用有用。对于视频来说,广泛采用的是有损压缩。被容许的失真量取决于应用;例如,某些消费者流应用的用户可能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出:更高的容许/可容忍失真可以产生更高的压缩比。

[0007] 视频编码器和解码器可以利用几大类技术,包括例如运动补偿、变换、量化和熵编码。

[0008] 视频编解码器技术可以包括称为帧内编码的技术。在帧内编码中,样本值是在不参考来自先前重构的参考图片的样本或其他数据的情况下表示的。在某些视频编解码器中,图片在空间上被细分为样本块。当所有样本块都以帧内模式编码时,该图片可以是帧内图片。帧内图片及其派生图片(例如,独立的解码器刷新图片)可以用于重置解码器状态,因此可以用作编码视频比特流和视频会话中的第一图片,或者用作静止图像。可以对帧内块的样本可以被变换,并且在熵编码之前可以对变换系数进行量化。帧内预测可以是一种在预变换域中最小化样本值的技术。在一些情况下,变换后的DC值越小,AC系数越小,在给定量化步长下表示熵编码后的块所需的比特就越少。

[0009] 例如,从MPEG-2代编码技术中已知的传统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例如尝试从相邻的空间中且在解码顺序上在前的数据块的编码和/或解码期间获得的周围样本数据和/或元数据的技术。这种技术此后被称为“帧内预测”技术。注意,至少在一些情况下,帧内预测所使用的参考数据仅来自重构中的当前图片,而不是来自参考图片。

[0010] 可以有许多不同形式的帧内预测。当在给定的视频编码技术中可以使用多于一种这样的技术时,所使用的技术可以在帧内预测模式中被编码。在某些情况下,模式可以有子模式和/或参数,这些子模式和/或参数可以被单独编码或包含在模式码字中。对于给定的模式、子模式和/或参数组合,使用哪个码字会对通过帧内预测的编码效率增益产生影响,因此,对于将码字转换成比特流的熵编码技术也会产生影响。

[0011] H.264引入了特定的帧内预测模式,在H.265中进行了改进,并在更新的编码技术中进一步改进,例如,共同探索模型(Joint Exploration Model,简称JEM)、通用视频编码(Versatile Video Coding,简称VVC)和基准集(Benchmark Set,简称BMS)。使用属于已经是可用样本的相邻样本值来形成预测块。根据方向将相邻样本的样本值复制到预测块中。对使用中的方向的参考可以被编码进比特流中,或者可以被其本身预测出。

[0012] 参考图1A,右下方描绘了从H.265的33个可能的预测方向(对应于35个帧内模式的33个角度模式)中已知的9个预测方向的子集。箭头会聚的点(101)表示被预测的样本。箭头表示的是来自哪个被预测的样本的方向。例如,箭头(102)指示样本(101)是从与水平线成45°角的右上方的一个或多个样本预测到的。类似地,箭头(103)指示样本(101)是从与水平线成22.5°角的样本(101)左下方的一个或多个样本预测到的。

[0013] 仍然参考图1A,在左上方描绘了4×4样本的正方形块(104)(由粗虚线表示)。正方形块(104)包括16个样本,每个样本标有“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)。例如,样本S21是Y维度中的第二个样本(从顶部算起)和X维度中的第一个样本(从左侧算起)。类似地,样本S44在Y和X维度上都是块(104)中的第四个样本。由于该块的大小为4×4个样本,所以S44位于右下角。还示出了遵循类似编号方案的参考样本。用R标记的参考样本、其相对于块的Y位置(例如,行索引)和X位置(列索引)来标记(104)。在H.264和H.265中,预测样本与重构中的块相邻;因此,不需要使用负值。

[0014] 帧内图片预测可以通过从合适的信号预测方向的相邻样本中复制参考样本值来工作。例如,假设已编码的视频比特流包含针对此块的信号,其指示与箭头(102)一致的预测方向——即,从与水平成45°角的右上方的一个或多个预测样本来预测出的样本。在这种情况下,从相同的参考样本R05预测样本S41、S32、S23和S14。然后从参考样本R08预测样本S44。

[0015] 在某些情况下,多个参考样本的值可以组合,例如,通过插值,以便计算参考样本;尤其是当方向不能被45°整除时。

[0016] 随着视频编码技术的发展,可能的方向的数量已经增加。在H.264(2003年)中,可以表示九个不同方向。这在H.265(2013年)中增加到33个,而JEM/VVC/BMS在披露时可以支持多达65个方向。已经通过实验来识别最可能的方向,并且熵编码中的某些技术用于以少量比特来表示那些可能的方向,对于不太可能的方向接受一定的惩罚。此外,方向本身有时可以从相邻的已经解码的块中使用的相邻方向来预测。

[0017] 图1B示出了示意图(105),描绘了根据JEM的65个帧内预测方向,以说明预测方向的数量随着时间的推移而增加。

[0018] 编码视频比特流中帧内预测方向比特的映射所代表的方向可以根据视频编码技术的不同而不同;例如从预测方向到帧内预测模式的简单直接映射,其范围可以覆盖到码字,到涉及最可能模式的复杂自适应方案,以及类似的技术。然而,在所有情况下,与某些其他方向相比,某些方向在统计上不太可能出现在视频内容中。由于视频压缩的目标是减少冗余,在较好的视频编码技术中,那些不太可能的方向将由更多的比特来表示,相对那更可能的方向。

[0019] 运动补偿可以是有损压缩技术,并且可以涉及这样的技术,其中,来自先前重构的图片或其一部分(参考图片)的样本数据块在由运动矢量(此后称为MV)指示的方向上进行空间移位之后,可以用于预测新重构的图片或图片部分。在某些情况下,参考图片可以与当前正在重构的图片相同。MV可以具有两个维度X和Y,或者三个维度,第三个维度是使用中的参考图片的指示(参考图片间接可以是时间维度)。

[0020] 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中预测,例如,从与空间上邻近正在重构的区域的并且在解码顺序上在该MV之前的样本数据的另一个区域相关的MV中预测。这样做可以大大减少编码MV所需的数据量,从而消除冗余并提高压缩率。MV预测可以有效地工作,例如,因为当对从相机导出的输入视频信号(称为自然视频)进行编码时,存在统计可能性,即比单个MV可应用的区域大的区域在相似的方向上移动,因此,在某些情况下,可以使用从相邻区域的MV导出的相似MV来预测。这使得对于给定区域发现的MV与从周围MV预测出的MV相似或相同,并且在熵编码之后,这又可以用比直接编码MV时所使用的更少的比特数来表示。在某些情况下,MV预测可以是从原始信号(即:样本流)导出的信号(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耗的,例如,当从几个周围的MV计算预测值时,会出现舍入误差。

[0021] 在H.265/HEVC(ITU-T Rec.H.265,“High Efficiency Video Coding”,2016年12月)中描述各种MV预测机制。在H.265提供的许多MV预测机制中,本文描述了一种此后被称为“空间合并”的技术。

[0022] 参考图1C,当前块(111)可以包括由编码器在运动搜索过程中发现的样本,以便从已经被空间移位的相同大小的先前块中预测。并非直接编码该MV,可以使用与五个周围样本(表示为A0、A1和B0、B1、B2(分别为112至116))中的任一个相关联的MV,从与一个或多个参考图片相关联的元数据中导出该MV,例如,从最近的(按照解码顺序)参考图片中导出。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。

发明内容

[0023] 本公开的各方面包括用于视频解码的装置。一种装置包括处理电路,该处理电路接收包括当前块的编码残差的视频比特流。处理电路基于编码残差确定当前块的多个变换系数。处理电路通过基于当前块的最顶行或最左列中的至少一个中的样本子集对多个变换系数的子集执行符号预测,来为多个符号预测假设中的每一个符号预测假设计算相应符号预测假设的成本。样本子集中的样本数量是常数。处理电路选择具有最小成本的一个符号预测假设。处理电路根据所选择的符号预测假设执行符号预测。

[0024] 在一个实施例中,常数是预先确定的。例如,常数可以是4或8。

[0025] 在一个实施例中,基于多个变换系数的子集的数量确定常数。例如,可以预先确定多个变换系数的子集的数量(例如,变换块中的4×4低频变换系数),并且可以确定常数小于多个变换系数的子集的预定数量。当多个变换系数的子集被预先确定为4×4低频变换系数时,对于最顶行或最左列,常数可以是4。

[0026] 在一个实施例中,基于块的一个维度与阈值之间的比较来确定常数。在一个示例中,当块的宽度(或高度)大于阈值时,最顶行(或最左列)中的常数可以是第一数量。否则,最顶行(或最左列)中的常数可以是小于第一数量的第二数量。

[0027] 在一个实施例中,基于块的一个维度与一个以上阈值之间的比较来确定常数。在一个示例中,块的尺寸可以与两个阈值进行比较。当块的宽度(或高度)不大于第一阈值时,最顶行(或最左列)中的常数可以是第一数量。当块的宽度(或高度)大于第一阈值并小于第二阈值时,最顶行(或最左列)中的常数可以是第二数量。当块的宽度(或高度)大于第二阈值时,最顶行(或最左列)中的常数可以是第三数量。在另一个示例中,块的维度可以与不同数量的阈值进行比较。

[0028] 在一个实施例中,基于包括常数的多个值的查找表来确定常数值。多个值中的每一个都与块的一个维度的不同值相关联。例如,当块的宽度(或高度)在第一范围内时,根据查找表,最顶行(或最左列)中的常数可以是第一数量。当块的宽度(或高度)在第二范围内时,根据查找表,最顶行(或最左列)中的常数可以是第二数量,依此类推。

[0029] 在一个实施例中,基于块的一个维度来确定当前块的最顶行或最左列中的至少一个中的样本子集的位置。例如,基于块的宽度(或高度)来确定块的最顶行(或最左列)中的样本子集的位置。

[0030] 在一个实施例中,当前块的最顶行或最左列中的至少一个中的样本子集的至少两个位置是相邻的。例如,最顶行(或最左列)中的两个样本的位置彼此相邻。

[0031] 在一个实施例中,当前块的最顶行或最左列中的至少一个中的样本子集的位置是均匀间隔的。

[0032] 在一个实施例中,处理电路基于(i)当前块的最顶行或最左列中的至少一个中的样本子集,以及(ii)当前块上方的行或当前块左侧的列中的至少一个中的相邻重构样本,来为多个符号预测假设中的每一个符号预测假设计算相应符号预测假设的成本。

[0033] 本公开的各方面提供了用于视频解码的方法。这些方法可以执行由用于视频编码/解码的装置执行的过程中的任何一个或组合。在该方法中,接收包括当前块的编码残差的视频比特流。基于编码残差确定当前块的多个变换系数。通过基于当前块的最顶行或最左列中的至少一个中的样本子集对多个变换系数的子集执行符号预测,来为多个符号预测假设中的每一个符号预测假设计算相应符号预测假设的成本。样本子集中的样本数量是常数。选择具有最小成本的一个符号预测假设。根据所选择的符号预测假设执行符号预测。

[0034] 本公开的各方面还提供了存储指令的非暂时性计算机可读介质,当由至少一个处理器执行时,指令使得至少一个处理器执行用于视频解码的方法中的任何一种或组合。

附图说明

[0035] 从以下详细描述和附图中,所公开主题的进一步特征、性质和各种优点将变得更

加明显,其中:

[0036] 图1A是帧内预测模式的示示例集的示意图;

[0037] 图1B是示例性帧内预测方向的图示;

[0038] 图1C是一个示例中当前块及其周围空间合并候选的示意图;

[0039] 图2是根据一个实施例的通信系统的简化框图的示意图;

[0040] 图3是根据一个实施例的通信系统的简化框图的示意图;

[0041] 图4是根据一个实施例的解码器的简化框图的示意图;

[0042] 图5是根据一个实施例的编码器的简化框图的示意图;

[0043] 图6示出了根据另一实施例的编码器的框图;

[0044] 图7示出了根据另一实施例的解码器的框图;

[0045] 图8示出了根据本公开的实施例的示例性符号预测过程;

[0046] 图9示出了根据本公开的实施例的假设重构边界的成本的示例性计算;

[0047] 图10示出了根据本公开的另一实施例的假设重构边界的成本的示例性计算;

[0048] 图11示出了根据本公开的另一实施例的假设重构边界的成本的示例性计算;

[0049] 图12示出了根据本公开的另一实施例的假设重构边界的成本的示例性计算;

[0050] 图13示出了根据本公开的另一实施例的示例性符号预测过程;

[0051] 图14示出了根据一个实施例的示例性流程图;以及

[0052] 图15是根据一个实施例的计算机系统的示意图。

具体实施方式

[0053] I.视频解码器和编码器系统

[0054] 图2示出了根据本公开实施例的通信系统(200)的简化框图。通信系统(200)包括能够经由例如网络(250)相互通信的多个终端设备。例如,通信系统(200)包括经由网络(250)互连的第一对终端设备(210)和(220)。在图2的示例中,第一对终端设备(210)和(220)执行数据的单向传输。例如,终端设备(210)可以对视频数据(例如,由终端设备(210)捕捉的视频图片流)进行编码,以便经由网络(250)传输到另一个终端设备(220)。编码视频数据可以以一个或多个编码视频比特流的形式传输。终端设备(220)可以从网络(250)接收已编码视频数据,解码上述已编码的视频数据,以恢复视频图片,并根据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场景中是常见的。

[0055] 在另一示例中,通信系统(200)包括第二对终端设备(230)和(240),其执行对已编码视频数据的双向传输,例如对于在视频会议期间产生的编码视频数据。对于数据的双向传输,在一个示例中,终端设备(230)和(240)的每个终端设备可以对视频数据(例如,由终端设备捕捉的视频图片流)进行编码,以便经由网络(250)传输到终端设备(230)和(240)的另一个终端设备。终端设备(230)和(240)的每个终端设备还可以接收由终端设备(230)和(240)的另一个终端设备传输的编码视频数据,并且可以解码已编码视频数据,以恢复视频图片,并且可以根据恢复的视频数据在可访问的显示设备上显示视频图片。

[0056] 在图2的示例中,终端设备(210)、终端装置(220)、终端装置(230)和终端装置(240)可以被示为服务器、个人计算机和智能电话,但是本公开的原理可以不限于此。本公开的实施例适用于膝上型计算机、平板计算机、媒体播放器和/或专用视频会议设备。网络

(250)表示在终端设备(210)、终端装置(220)、终端装置(230)和终端装置(240)之间传送编码视频数据的任意数量的网络,包括例如线路(有线)和/或无线通信网络。通信网络(250)可以在电路交换和/或分组交换信道中交换数据。代表性的网络包括电信网络、局域网、广域网和/或互联网。出于当前讨论的目的,网络(250)的架构和拓扑对于本公开的操作可能是不重要的,除非在下文中解释。

[0057] 作为所公开的主题的应用的示例,图3示出了视频编码器和视频解码器在流环境中的放置。所公开的主题同样可应用于其他支持视频的应用,包括例如视频会议、数字电视、在包括CD、DVD、记忆棒等的数字媒体上存储压缩视频等。

[0058] 流式传输系统可以包括:捕捉子系统(313),其可以包括视频源(301),例如,数码相机;创建例如未压缩的视频图片流(302)。在一个示例中,视频图片流(302)包括由数码相机拍摄的样本。当与编码视频数据(304)(或编码视频比特流)相比时,视频图片流(302)被描绘为强调高数据量的粗线,可以由包括耦合到视频源(301)的视频编码器(303)的电子设备(320)来处理。视频编码器(303)可以包括硬件、软件或其组合,以实现或实施如下面更详细描述的所公开的主题的各方面。编码视频数据(304)(或编码视频比特流(304))被描绘为细线,以强调当与视频图片流(302)相比时其为较低的数据量,可以存储在流服务器(305)上,以供将来使用。一个或多个流客户端子系统(例如,图3中的客户端子系统(306)和(308))可以访问流服务器(305),以检索编码视频数据(304)的副本(307)和(309)。客户端子系统(306)可以包括例如电子设备(330)中的视频解码器(310)。视频解码器(310)对编码视频数据的副本(307)进行解码,并创建可以在显示器(312)(例如,显示屏)或其他渲染设备(未示出)上渲染的输出视频图片流(311)。在一些流系统中,编码的视频数据(304)、编码视频数据的副本(307)和(309)(例如,编码视频比特流)可以根据某些视频编码/压缩标准来编码。这些标准的示例包括ITU-T建议H.265。在一个示例中,正在开发的视频编码标准被非正式地称为通用视频编码(VVC)。所公开的主题可以在VVC的环境中使用。

[0059] 需要注意的是,电子设备(320)和(330)可以包括其他组件(未示出)。例如,电子设备(320)可以包括视频解码器(未示出),并且电子设备(330)也可以包括视频编码器(未示出)。

[0060] 图4示出了根据本公开实施例的视频解码器(410)的框图。视频解码器(410)可以包括在电子设备(430)中。电子设备(430)可以包括接收机(431)(例如,接收电路)。视频解码器(410)可以用来代替图3示例中的视频解码器(310)。

[0061] 接收机(431)可以接收将由视频解码器(410)解码的一个或多个编码视频序列;在同一个或另一个实施例中,一次一个编码视频序列,其中,每个编码视频序列的解码独立于其他编码视频序列。可以从信道(401)接收编码视频序列,该信道可以是到存储编码视频数据的存储设备的硬件/软件链接。接收机(431)可以接收编码视频数据和其他数据,例如,编码的音频数据和/或辅助数据流,这些数据可以被转发到其相应的使用实体(未示出)。接收机(431)可以将编码视频序列与其他数据分离。为了对抗网络抖动,缓冲存储器(415)可以耦合在接收机(431)和熵解码器/解析器(420)(以下称为“解析器(420)”)之间。在某些应用中,缓冲存储器(415)是视频解码器(410)的一部分。在其他情况下,可以在视频解码器(410)之外(未示出)。在其他情况下,在视频解码器(410)之外可以有缓冲存储器(未示出),例如,用于对抗网络抖动,此外,在视频解码器(410)内部可以有另一个缓冲存储器(415),

例如,用于处理播放定时。当接收机(431)从具有足够带宽和可控性的存储/转发设备或者从等同步网络中接收数据时,缓冲存储器(415)可以是不需要的,或者可以是小的。为了在诸如因特网之类的尽力而为的分组网络上使用,可能需要缓冲存储器(415),该缓冲存储器可以相对较大,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(410)外部的操作系统或类似元件(未示出)中实现。

[0062] 视频解码器(410)可以包括解析器(420),以从编码的视频序列中重构符号(421)。这些符号的类别包括用于管理视频解码器(410)的操作的信息以及潜在地控制诸如渲染设备(412)(例如,显示屏)等渲染设备的信息,该渲染设备不是电子设备(430)的组成部分,但是可以耦合到电子设备(430),如图4所示。用于渲染设备的控制信息可以是补充增强信息(Supplemental Enhancement Information,简称SEI消息)或视频可用性信息(Video Usability Information,简称VUI)参数集片段(未示出)的形式。解析器(420)可以对接收到的编码视频序列进行解析/熵解码。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性的算术编码等。解析器(420)可以基于对应于该组的至少一个参数,从编码视频序列中提取视频解码器中的至少一个像素子组的一组子组参数。子组可以包括图片组(Groups of Pictures,简称GOP)、图片、图块、切片、宏块、编码单元(Coding Units,简称CU)、块、变换单元(Transform Units,简称TU)、预测单元(Prediction Units,简称PU)等。解析器(420)还可以从编码视频序列中提取信息,例如,变换系数、量化器参数值、MV等。

[0063] 解析器(420)可以对从缓冲存储器(415)接收的视频序列执行熵解码/解析操作,以便创建符号(421)。

[0064] 根据编码视频图片或其部分的类型(例如:帧间和帧内图片、帧间和帧内块)以及其他因素,符号(421)的重构可以涉及多个不同的单元。可以通过由解析器(420)从编码视频序列中解析的子组控制信息来控制涉及哪些单元以及如何涉及。为了清楚起见,没有描述解析器(420)和下面的多个单元之间的这种子组控制信息流。

[0065] 除了已经提到的功能块之外,视频解码器(410)可以在概念上细分成如下所述的多个功能单元。在商业限制下操作的实际实现中,许多这些单元彼此紧密交互,并且可以至少部分地彼此集成。然而,为了描述所公开的主题,在概念上细分成以下功能单元是合适的。

[0066] 第一单元是定标器/逆变换单元(451)。定标器/逆变换单元(451)接收量化的变换系数以及控制信息,包括使用哪个变换、块大小、量化因子、量化缩放矩阵等,作为来自解析器(420)的符号(421)。定标器/逆变换单元(451)可以输出包括样本值的块,这些块可以被输入到聚集器(455)中。

[0067] 在一些情况下,定标器/逆变换单元(451)的输出样本可以属于帧内编码块;即:没有使用来自先前重构图片的预测信息,但是可以使用来自当前图片的先前重构部分的预测信息的块。这种预测信息可以由帧内图片预测单元(452)提供。在一些情况下,帧内图片预测单元(452)使用从当前图片缓冲器(458)获取的周围已经重构的信息,来生成与重构中的块具有相同大小和形状的块。当前图片缓冲器(458)缓冲例如部分重构的当前图片和/或完全重构的当前图片。在一些情况下,聚集器(455)基于每个样本将帧内预测单元(452)已经生成的预测信息添加到由定标器/逆变换单元(451)提供的输出样本信息。

[0068] 在其他情况下,定标器/逆变换单元(451)的输出样本可以属于帧间编码的并且可能是运动补偿的块。在这种情况下,运动补偿预测单元(453)可以访问参考图片存储器(457),以获取用于预测的样本。在根据与该块有关的符号(421)对获取的样本进行运动补偿之后,这些样本可以由聚集器(455)添加到定标器/逆变换单元(451)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(453)从中获取预测样本的参考图片存储器(457)内的地址可以由运动矢量来控制,对于运动补偿预测单元(453)可行的是,以符号(421)的形式获得这些地址,这些符号可以具有例如X、Y和参考图片组件。当使用子采样精确运动向量(Motion Vector,简称MV)时,运动补偿还可以包括从参考图片存储器(457)获取的样本值的插值、MV预测机制等。

[0069] 聚集器(455)的输出样本可以在环路滤波器单元(456)中经受各种环路滤波技术。视频压缩技术可以包括环路滤波技术,该技术由编码视频序列(也称为编码视频比特流)中包含的参数控制,并且该参数可以作为来自解析器(420)的符号(421)用于环路滤波单元(456)中,而且也可以对在编码图像或编码视频序列的先前(按照解码顺序)部分的解码期间获得的元信息进行响应,以及对先前的重构和环路滤波的样本值进行响应。

[0070] 环路滤波器单元(456)的输出可以是样本流,该样本流可以输出到渲染设备(412)以及存储在参考图片存储器(457)中,以用于将来的帧间图片预测。

[0071] 一旦完全重构,某些编码图片可以用作未来预测的参考图片。例如,一旦完全重构与当前图片对应的编码图像,并且该编码的图片已经被识别为参考图片(例如,通过解析器(420)),则当前图片缓冲器(458)可以成为参考图片存储器(457)的一部分,并且在开始对接下来的编码图片的重构之前,可以重新分配新的当前图片缓冲器。

[0072] 视频解码器(410)可以根据诸如ITU-T Rec.H.265等标准中的预定视频压缩技术来执行解码操作。在编码视频序列符合视频压缩技术或标准的语法以及视频压缩技术或标准中记载的简介的意义上,编码视频序列可以符合由所使用的视频压缩技术或标准规定的语法。具体地,简介可以从视频压缩技术或标准中所有可用的工具中选择某些工具,作为在该简介下唯一可用的工具。符合标准还需要编码视频序列的复杂度在视频压缩技术或标准的级别所定义的范围内。在某些情况下,级别限制了最大图片大小、最大帧速率、最大重构采样率(例如,以每秒兆样本为单位测量)、最大参考图片大小等。在某些情况下,由级别设置的限制可以通过假设参考解码器(Hypothetical Reference Decoder,简称HRD)规范和用于编码视频序列中信令的HRD缓冲管理的元数据来进一步限制。

[0073] 在一个实施例中,接收机(431)可以接收具有编码视频的额外(冗余)数据。该额外数据可以作为一部分包括在编码视频序列中。视频解码器(410)可以使用额外数据来正确解码数据和/或更准确地重构原始视频数据。额外数据可以是例如时间、空间或信噪比(SNR)增强层、冗余切片、冗余图片、前向纠错码等形式。

[0074] 图5示出了根据本公开实施例的视频编码器(503)的框图。视频编码器(503)包括在电子设备(520)中。电子设备(520)包括发射机(540)(例如,发射电路)。视频编码器(503)可以用来代替图3示例中的视频编码器(303)。

[0075] 视频编码器(503)可以从视频源(501)(其不是图5示例中的电子设备(520)的一部分)接收视频样本,该视频源可以捕捉要由视频编码器(503)编码的视频图片。在另一示例中,视频源(501)是电子设备(520)的一部分。

[0076] 视频源(501)可以以数字视频样本流的形式提供要由视频编码器(503)编码的源视频序列,该数字视频样本流可以具有任何合适的比特深度(例如:8比特、10比特、12比特、...)、任何颜色空间(例如,BT.501Y CrCB、RGB、...)和任何合适的采样结构(例如,Y CrCb 4:2:0、Y CrCb4:4:4)。在媒体服务系统中,视频源(501)可以是存储先前准备的视频的存储设备。在视频会议系统中,视频源(501)可以是捕捉本地图像信息作为视频序列的相机。视频数据可以以多个单独的图片的形式提供,当按顺序观看时,这些图片赋予运动。这些图片本身可以被组织为像素的空间阵列,其中,每个像素可以包括一个或多个样本,这取决于使用中的采样结构、颜色空间等。本领域技术人员可以容易地理解像素和样本之间的关系。下面的描述集中在样本上。

[0077] 根据一个实施例,视频编码器(503)可以实时地或者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成编码视频序列(543)。实施适当的编码速度是控制器(550)的一个功能。在一些实施例中,控制器(550)控制如下所述的其他功能单元,并且在功能上耦合到其他功能单元。为了清楚起见,没有描述耦合。控制器(550)设置的参数可以包括速率控制相关参数(图片跳过、量化器、率失真优化技术的λ值、…)、图片大小、图片组(GOP)布局、最大MV允许的参考区域等。控制器(550)可以被配置为具有针对特定系统设计而优化的与视频编码器(503)相关的其他合适的功能。

[0078] 在一些实施例中,视频编码器(503)被配置为在编码循环中操作。作为简化的描述,在一个示例中,编码环路可以包括源编码器(530)(例如,用于基于待编码的输入图片和参考图片来创建符号,例如,符号流)以及嵌入在视频编码器(503)中的(本地)解码器(533)。解码器(533)以类似于(远程)解码器也将创建的方式重构符号,以创建样本数据(因为在所公开的主题中考虑的视频压缩技术中,符号和编码视频比特流之间的任何压缩都是无损的)。重构的样本流(样本数据)被输入到参考图片存储器(534)。由于符号流的解码导致位精确结果独立于解码器位置(本地或远程),所以参考图片存储器(534)中的内容在本地编码器和远程编码器之间也是位精确的。换言之,当在解码期间使用预测时,编码器的预测部分作为参考图片样本"看到"与解码器"看到"的样本值完全相同的样本值。也在一些相关技术中使用参考图片同步性(以及如果不能保持同步性而由此产生的漂移,例如,由于信道误差)的该基本原理。

[0079] "本地"解码器(533)的操作可以与诸如视频解码器(410)之类的"远程"解码器的操作相同,上面已经结合图4对其进行了详细描述。然而,简要地参考图4,由于符号是可用的,并且熵编码器(545)和解析器(420)对编码视频序列的符号编码/解码可以是无损的,所以可以不完全在本地解码器(533)中实现视频解码器(410)的熵解码部分,包括缓冲存储器(415)和解析器(420)。

[0080] 在这一点上可以观察到,除了解码器中存在的解析/熵解码之外,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简化,因为这些技术是全面描述的解码器技术的逆。仅在某些领域,下面需要并且提供了更详细的描述。

[0081] 在操作期间,在一些示例中,源编码器(530)可以执行运动补偿预测编码,其参考来自视频序列的被指定为"参考图片"的一个或多个先前已编码的图片来对输入图片进行预测性地编码。以这种方式,编码引擎(532)对输入图片的像素块和可被选为对输入图片进

行预测参考的参考图片的像素块之间的差异进行编码。

[0082] 本地视频解码器(533)可以基于由源编码器(530)创建的符号,对可以被指定为参考图片的图片的编码视频数据进行解码。编码引擎(532)的操作可以是有利地有损过程。当已编码的视频数据可以在视频解码器(图6中未示出)处被解码时,重构的视频序列通常可以是具有一些误差的源视频序列的副本。本地视频解码器(533)可以复制由视频解码器对参考图片执行的解码过程,并且可以使得重构的参考图片存储在参考图片缓存(534)中。以这种方式,视频编码器(503)可以本地存储重构的参考图片的副本,这些副本具有与将由远端视频解码器获得的重构的参考图片相同的内容(不存在传输误差)。

[0083] 预测器(535)可以对编码引擎(532)执行预测搜索。也就是说,对于待编码的新图片,预测器(535)可以在参考图片存储器(534)中搜索样本数据(作为候选参考像素块)或某些元数据,例如,参考图片运动矢量、块形状等,其可以用作新图片的适当预测参考。预测器(535)可以在逐个样本块-像素块的基础上操作,以找到合适的预测参考。在一些情况下,如由预测器(535)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(534)中的多个参考图片中提取的预测参考。

[0084] 控制器(550)可以管理源编码器(530)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。

[0085] 所有前述功能单元的输出可以在熵编码器中经历熵编码(545)。熵编码器(545)通过根据诸如Huffman编码、可变长度编码、算术编码等技术的无损压缩符号,将由各种功能单元生成的符号转换成编码的视频序列。

[0086] 发射机(540)可以缓冲由熵编码器(545)创建的编码视频序列,以准备经由通信信道(560)传输,通信信道可以是到将存储编码视频数据的存储设备的硬件/软件链接。发射机(540)可以将来自视频编码器(503)的编码视频数据与待传输的其他数据合并,例如,编码音频数据和/或辅助数据流(源未示出)。

[0087] 控制器(550)可以管理视频编码器(503)的操作。在编码期间,控制器(550)可以向每个编码图片分配特定的编码图片类型,这可以影响应用于相应图片的编码技术。例如,图片通常可以被分配为以下图片类型之一:

[0088] 帧内图片(I图片)可以是不使用序列中的任何其他图片作为预测源而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(“IDR”)图片。本领域技术人员知道I图片的那些变体以及其相应的应用和特征。

[0089] 预测图片(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,该预测使用了最多一个MV和参考索引来预测每个块的样本值。

[0090] 双向预测图片(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,该预测使用了最多两个MV和参考索引来预测每个块的样本值。类似地,多预测图片可以使用两个以上的参考图片和相关元数据来重构单个块。

[0091] 源图片通常可以在空间上被细分成多个样本块(例如,每个样本块为4×4、8×8、4×8或16×16个样本块),并且在逐块的基础上编码。参考其他(已经编码的)块来对块进行预测性地编码,这里其他块是由应用于块的相应图片的编码分配来确定的。例如,I图片的块可以被非预测性地编码,或者可以参考同一图片的已经编码的块被预测性地编码(空间预测或帧内预测)。参考一个先前编码的参考图片,经由空间预测或经由时间预测,对P图片

的像素块进行预测性地编码。参考一个或两个先前编码的参考图片,经由空间预测或经由时间预测,可以对B图片的块进行预测性地编码。

[0092] 视频编码器(503)可以根据预定的视频编码技术或标准(例如,ITU-TRec.H.265)来执行编码操作。在其操作中,视频编码器(503)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编码操作。因此,编码的视频数据可以符合由正在使用的视频编码技术或标准指定的语法。

[0093] 在一个实施例中,发射机(540)可以与编码视频一起传输额外数据。源编码器(530)可以包括作为编码视频序列的一部分的数据,。额外数据可以包括时间/空间/SNR增强层、其他形式的冗余数据(例如,冗余图片和切片)、SEI消息、VUI参数集片段等。

[0094] 从视频中捕捉作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)是利用给定图片中的空间相关性,而帧间图片预测是利用图片之间的(时间或其他)相关性。在一个示例中,一个在编码/解码中的特定图片,被称为当前图片,被分割成块。在当前图片中的块类似于视频中先前编码且仍被缓冲的参考图片中的参考块时,当前图片中的块可以通过称为MV的矢量来编码。在使用多个参考图片的情况下,MV指向参考图片中的参考块,并且可以具有识别参考图片的第三维。

[0095] 在一些实施例中,可以在帧间图片预测中使用双向预测技术。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,这两个参考图片在解码顺序上都在视频中的当前图片之前(但是在显示顺序上可以分别在过去和未来)。当前图片中的块可由指向第一参考图片中的第一参考块的第一MV和指向第二参考图片中的第二参考块的第二MV来编码。可以通过第一参考块和第二参考块的组合来预测该块。

[0096] 此外,可以在帧间图片预测中使用合并模式技术来提高编码效率。

[0097] 根据本公开的一些实施例,例如,帧间图片预测和帧内图片预测以块为单位执行预测。例如,根据HEVC标准,视频图片序列中的图片被分割成编码树单元(Coding Tree Units,简称CTU),用于压缩,图片中的CTU具有相同的大小,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 Blocks,简称CTB),即一个亮度CTB和两个色度CTB。每个CTU可以被递归地四叉树分割成一个或多个编码单元(CU)。例如,64×64像素的CTU可以被分割成一个64×64像素的CU,或者4个32×32像素的CU,或者16个16×16像素的CU。在一个示例中,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。根据时间和/或空间可预测性,CU被分成一个或多个预测单元(Prediction Units,简称PU)。通常,每个PU包括一个亮度预测块(Prediction Block,简称PB)和两个色度PB。在实施例中,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,预测块包含像素的值(例如,亮度值)的矩阵,例如,8×8像素、16×16像素、8×16像素、16×8像素等。

[0098] 图6示出了根据本公开的另一实施例的视频编码器(603)的示意图。视频编码器(603)被配置为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),并将该处理块编码成作为编码视频序列的一部分的编码图像。在一个示例中,视频编码器(603)用于代替图3示例中的视频编码器(303)。

[0099] 在HEVC示例中,视频编码器(603)接收处理块的样本值矩阵,例如,8×8样本的预测块等。视频编码器(603)确定使用帧内模式、帧间模式还是使用例如率失真优化的双向预

测模式,来对处理块进行最佳编码。当要以帧内模式对处理块进行编码时,视频编码器(603)可以使用帧内预测技术将处理块编码成编码图片;并且当要以帧间模式或双向预测模式对处理块进行编码时,视频编码器(603)可以分别使用帧间预测或双向预测技术来将处理块编码成编码图片。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,从一个或多个MV预测器中导出MV,而不受益于预测器之外的编码MV组件。在某些其他视频编码技术中,可能存在适用于对象块的MV组件。在一个示例中,视频编码器(603)包括其他组件,例如,用于确定处理块的模式的模式判定模块(未示出)。

[0100] 在图6的示例中,视频编码器(603)包括如图6所示耦合在一起的帧间编码器(630)、帧内编码器(622)、残差计算器(623)、交换机(626)、残差编码器(624)、通用控制器(621)和熵编码器(625)。

[0101] 帧间编码器(630)被配置为接收当前块(例如,处理块)的样本,将该块与参考图片中的一个或多个参考块(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、运动矢量、合并模式信息),并且基于使用了任何合适的技术得到的帧间预测信息来计算帧间预测结果(例如,预测块)。在一些示例中,参考图片是基于编码视频信息解码的解码参考图片。

[0102] 帧内编码器(622)被配置为接收当前块(例如,处理块)的样本,在一些情况下,将该块与已经在同一图片中编码的块进行比较,在变换后生成量化系数,并且在一些情况下,还生成帧内预测信息(例如,根据一种或多种帧内编码技术的帧内预测方向信息)。在一个示例中,帧内编码器(622)还基于同一图片中的帧内预测信息和参考块来计算帧内预测结果(例如,预测块)。

[0103] 通用控制器(621)被配置为确定通用控制数据并基于通用控制数据控制视频编码器(603)的其他组件。在一个示例中,通用控制器(621)确定块的模式,并基于该模式向交换机(626)提供控制信号。例如,当模式是帧内模式时,通用控制器(621)控制交换机(626)选择帧内模式结果供残差计算器(623)使用,并控制熵编码器(625)选择帧内预测信息并将帧内预测信息包括在比特流中;当模式是帧间模式时,通用控制器(621)控制交换机(626)选择帧间预测结果供残差计算器(623)使用,并控制熵编码器(625)选择帧间预测信息并将帧间预测信息包括在比特流中。

[0104] 残差计算器(623)被配置为计算接收的块和从帧内编码器(622)或帧间编码器(630)选择的预测结果之间的差(残差数据)。残差编码器(624)被配置为基于残差数据进行操作,以编码残差数据,来生成变换系数。在一个示例中,残差编码器(624)被配置为将残差数据从空间域转换到频域,并生成变换系数。然后对变换系数进行量化处理,以获得量化的变换系数。在各种实施例中,视频编码器(603)还包括残差解码器(728)。残差解码器(728)被配置为执行逆变换,并生成解码后的残差数据。解码的残差数据可以被帧内编码器(622)和帧间编码器(630)适当地使用。例如,帧间编码器(630)可以基于解码的残差数据和帧间预测信息生成解码的块,帧内编码器(622)可以基于解码的残差数据和帧内预测信息生成解码的块。适当地处理被解码的块,以生成已解码的图片,并且已解码的图片可以缓存在存储器电路(未示出)中,并且在一些示例中用作参考图片。

[0105] 熵编码器(625)被配置为格式化比特流,以包括编码块。熵编码器(625)被配置为根据合适的标准,例如,HEVC,包括各种信息。在一个示例中,熵编码器(625)被配置为包括

通用控制数据、所选预测信息(例如,帧内预测信息或帧间预测信息)、残差信息和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中对块进行编码时,没有残差信息。

[0106] 图7示出了根据本公开的另一实施例的视频解码器(710)的示意图。视频解码器(710)被配置为接收作为编码视频序列的一部分的编码图片,并对编码图片进行解码,以生成重构图片。在一个示例中,视频解码器(710)用于代替图3示例中的视频解码器(310)。

[0107] 在图7的示例中,视频解码器(710)包括如图7所示耦合在一起的熵解码器(771)、帧间解码器(780)、残差解码器(773)、重构模块(774)和帧内解码器(772)。

[0108] 熵解码器(771)可以被配置为从编码图像中重构某些符号,这些符号表示组成编码图像的语法元素。这种符号可以包括例如对块进行编码的模式(例如,帧内模式、帧间模式、双向预测模式、合并子模式或另一子模式中的后两种模式)、可以识别出分别由帧内解码器(772)或帧间解码器(780)进行预测的某些样本或元数据的预测信息(例如,帧内预测信息或帧间预测信息)、例如量化变换系数形式的残差信息等。在一个示例中,当预测模式是帧间或双向预测模式时,帧间预测信息被提供给帧间解码器(780);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(772)。残差信息可以经历逆量化,并被提供给残差解码器(773)。

[0109] 帧间解码器(780)被配置为接收帧间预测信息,并基于帧间预测信息生成帧间预测结果。

[0110] 帧内解码器(772)被配置为接收帧内预测信息,并基于帧内预测信息生成预测结果。

[0111] 残差解码器(773)被配置为执行逆量化,以提取去量化的变换系数,并处理去量化的变换系数,以将残差从频域转换到空间域。残差解码器(773)还可能需要某些控制信息(以包括量化参数(QP)),并且该信息可以由熵解码器(771)提供(数据路径未示出,因为这可能只是少量控制信息)。

[0112] 重构模块(774)被配置为在空间域中组合由残差解码器(773)输出的残差和预测结果(由帧间或帧内预测模块输出,视情况而定),以形成重构块,该重构块可以是重构图片的一部分,该重构图片又可以是重构视频的一部分。注意,可以执行诸如去块操作等其他合适的操作来提高视觉质量。

[0113] 注意,视频编码器(303)、视频编码器(503)和视频编码器(603)以及视频解码器(310)、视频解码器(410)和视频解码器(710)可以使用任何合适的技术来实现。在一个实施例中,视频编码器(303)、视频编码器(503)和视频编码器(603)以及视频解码器(310)、视频解码器(410)和视频解码器(710)可以使用一个或多个集成电路来实现。在另一实施例中,视频编码器(303)、视频编码器(503)和视频编码器(603)以及视频解码器(310)、视频解码器(410)和视频解码器(710)可以使用执行软件指令的一个或多个处理器来实现。

[0114] II.变换系数的符号预测

[0115] 在诸如探索性压缩模型(Explorative Compression Model,简称ECM)1.0的一些相关示例中,可以实现残差编码中的变换系数的符号预测过程。在一些情况下,变换块(Transform Block,简称TB)的一部分中的非零变换系数可以用于符号预测过程。TB的部分可以包括TB的低频变换系数。低频变换系数是靠近TB左上角的变换系数。在EMC-1.0测试条

件下,可以预测包括TB的最低4x4频率分量的TB的一部分中的非零变换系数的最多8个符号。例如,TB的左上部分(例如,8x8 TB)可以包括TB的最低4×4频率分量。

[0116] 当预测TB中的变换系数的多个符号(例如,n个符号)时,编码器和/或解码器可以以预定的搜索顺序搜索n个符号,以便最小化空间域中TB的边界处的不连续性的成本。

[0117] 图8示出了根据本公开的实施例的最小化TB的边界处的不连续性成本的示例性符号预测过程(800)。在诸如ECM-1.0之类的一些相关示例中,符号预测过程的搜索顺序可以基于n个符号位,其顺序例如是格雷码的编码顺序,格雷码是二进制数字系统的顺序,使得两个连续值仅在一位上不同。每个符号位可以表示不同变换系数的符号。在本公开中,n个符号位的每个搜索点可以被称为假设。也就是说,n个不同符号的组合可以称为假设,其中,每个符号可以是正的或负的。对于n个符号位,总共有2^n个假设。2^n个假设是n个不同符号的2^n个不同组合,其中,每个符号可以是正的或负的。

[0118] 在图8中,符号预测过程(800)从n个符号位的2^n个预测假设中以预定顺序选择符号预测假设,如步骤(S810)所示。然后,符号预测过程(800)使用TB的顶边界和/或左边界中的所有采样位置来计算所选择的符号预测假设的成本,将所选择的符号预测假设的成本与一个或多个先前计算的其他符号预测假设的成本进行比较,并且选择具有最低成本的符号预测假设,作为在现有计算的成本中具有最小成本的符号预测候选,如步骤(S820)所示。在步骤(S830),符号预测过程(800)确定是否测试了所有假设。如果测试了所有假设,则具有最小成本的符号预测候选中的符号,可以用作预测TB的变换系数的符号的符号预测器。否则,符号预测过程(800)可以重复步骤(S810)、(S820)和(S830)。

[0119] 需要注意的是,图9-12中的每个当前块(900)-(1200)的最顶行和/或最左列中的重构像素可以被称为假设重构边界。在一些情况下,最顶行和/或最左列中的重构像素可以通过使用与这些像素对应的变换系数的预测值和绝对值来重构,并且可能不具有最终的重构值。

[0120] 图9示出了根据本公开的实施例的假设重构边界的成本的示例性计算。在图9中,当前块(900)的左侧和顶部重构相邻像素都是可用的。对于当前块(900)的最左列(或左手侧(Left Hand Side,简称LHS))的每个重构像素$p_{0,y}$,当可用时,使用两个左侧重构相邻像素执行简单的线性预测,以获得其预测$pred_{0,y}=(2\cdot p_{-1,y}-p_{-2,y})$。预测$pred_{0,y}$和重构像素$p_{0,y}$之间的绝对差被添加到假设的成本中。类似地,对于当前块(900)的最顶行中的每个重构像素$p_{x,0}$,当可用时,使用两个顶部重构相邻像素执行简单的线性预测,以获得其预测$pred_{x,0}=(2\cdot p_{x,-1}-p_{x,-2})$。预测$pred_{x,0}$和重构像素$p_{x,0}$之间的绝对差被添加到假设的成本中。因此,假设的成本可以用(等式1)来表示。

[0121] $$ost=\sum_{x=0}^{w-1}\left|\left(2p_{x,-1}-p_{x,-2}\right)-p_{x,0}\right|+\sum_{y=0}^{h-1}\left|\left(2p_{-1,y}-p_{-2,y}\right)-p_{0,y}\right| \quad \text{(等式1)}$$

[0122] 图10示出了根据本公开的另一实施例的假设重构边界的成本的示例性计算。在图10中,只有当前块(1000)的左重构相邻像素是可用的。类似于(等式1)中的计算,假设的成本计算可以用(等式2)表示,其中,仅考虑基于左侧重构相邻像素的绝对差。

[0123] $$cost=\sum_{y=0}^{h-1}\left|\left(2p_{-1,y}-p_{-2,y}\right)-p_{0,y}\right| \quad \text{(等式2)}$$

[0124] 图11示出了根据本公开的另一实施例的假设重构边界的成本的示例性计算。在图

11中,只有当前块(1100)的顶部重构的相邻像素是可用的。类似于(等式1)中的计算,假设的成本计算可以用(等式3)表示,其中,仅考虑基于顶部重构的相邻像素的绝对差。

[0125] $$cost = \sum_{x=0}^{w-1} \left| \left(2p_{x,-1} - p_{x,-2}\right) - p_{x,0} \right| \quad \text{(等式 3)}$$

[0126] III.低复杂度符号预测

[0127] 在诸如ECM-1.0之类的一些相关示例中,符号预测过程的成本的计算可以与块的宽度和/或高度成比例,如(等式1)-(等式3)所示。此外,为了预测n个符号位,该计算可以执行2^n次。可以发现,符号预测过程的成本计算是ECM-1.0解码器中的热点。因此,有必要减少这种计算。

[0128] 本公开包括减少符号预测过程的成本计算的方法。

[0129] 根据本公开的一些方面,成本的计算可以独立于块的宽度和高度。例如,当在假设成本的计算中使用块的顶部和/或左边界的预定数量的像素时,成本的计算可以独立于块的维度。

[0130] 图12示出了根据本公开的实施例的假设重构边界的成本的示例性计算。在图12中,符号预测可适用于TB的一部分,其包含(例如)具有最低频率分量(例如,最低频率变换系数)的TB的4×4子块。在当前块(1200)的顶部重构的相邻像素是可用的时,预定数量的像素(例如,来自当前块(1200)的顶边界的4个样本)可以用于类似于(等式3)的线性预测中。类似地,在当前块(1200)的左重构相邻像素是可用的时,预定数量的像素(例如,来自当前块的左边界的4个样本)可以用于类似于(等式2)的线性预测中。

[0131] 根据本公开的各方面,用于成本计算的线性预测中的预定数量的样本,可以基于符号预测的最低频率分量的数量。样本的预定数量可以独立于块的大小。

[0132] 在图12中,当前块(1200)具有宽度w和高度h,它们是4的倍数。线性预测中使用的四个预定顶边界样本$p_{1,0}$、$p_{3,0}$、$p_{5,0}$和$p_{7,0}$可以位于如图12所示的以下水平位置:

[0133] $x_1 = w/8$ (等式4)

[0134] $x_2 = x_1 + w/4$ (等式5)

[0135] $x_3 = x2 + w/4$ (等式6)

[0136] $x_4 = x_3 + w/4$ (等式7)

[0137] 其中,/表示硬件实现或编程语言(例如,C编程语言)中的整数除法。

[0138] 用于线性预测的四个预定左边界样本$p_{0,0}$、$p_{0,2}$、$p_{0,4}$、和$p_{0,6}$可以位于如图12所示的以下垂直位置:

[0139] $y_1 = h - 1 - h/8$ (等式8)

[0140] $y_2 = y1 - h/4$ (等式9)

[0141] $y_3 = y_2 - h/4$ (等式10)

[0142] $y_4 = y_3 - h/4$ (等式11)

[0143] 其中,/表示硬件实现或编程语言(例如,C编程语言)中的整数除法。

[0144] 图13示出了根据本公开的实施例的最小化TB的边界处的不连续性成本的示例性符号预测过程(1300)。在图13中,符号预测过程(1300)基于TB的边界处的像素的子采样,从而可以选择预定数量的像素来计算成本。

[0145] 在图13中,符号预测过程(1300)确定TB的边界处的像素的子采样位置,如步骤

(S1310)所示。然后,如步骤(S1320)所示,符号预测过程(1300)从n个符号位的2^n个预测假设中以预定顺序选择符号预测假设。然后,符号预测过程(1300)使用TB的顶边界和/或左边界中的子采样位置来计算所选择的符号预测假设的成本,将所选择的符号预测假设的成本与一个或多个先前计算的其他符号预测假设的成本进行比较,并且选择具有最低成本的符号预测假设,作为在现有计算的成本中具有最小成本的符号预测候选,如步骤(S1330)所示。在步骤(S1340),符号预测过程(1300)确定是否测试了所有假设。如果测试了所有假设,则具有最小成本的符号预测候选中的符号可以用作一组符号预测器,用于预测TB的变换系数的符号。否则,符号预测过程(1300)可以重复步骤(S1310)-(S1340)。

[0146] 在一个实施例中,对于宽度w和高度h是4的倍数的块,该块的顶边界样本的子采样位置可以表示为以下水平位置:

[0147] $x_1 = x_{偏移}$ (等式12)

[0148] $x_2 = x_1 + \delta x$ (等式13)

[0149] $x_3 = x_2 + \delta x$ (等式14)

[0150] $x_4 = x_3 + \delta x$ (等式15)

[0151] 其中,/表示C中的整数除法,$x_{偏移}$是预先确定的非负偏移,δx是两个相邻样本之间的距离。$x_{偏移}$的示例值可以包括0和w/8。δx的示例值可以包括1和w/4。

[0152] 块的左边界样本的子采样位置可以表示为以下垂直位置:

[0153] $y_1 = h-1-y_{偏移}$ (等式16)

[0154] $y_2 = y_1 - \delta y$ (等式17)

[0155] $y_3 = y_2 - \delta y$ (等式18)

[0156] $y_4 = y_3 - \delta y$ (等式19)

[0157] 其中,/表示硬件实现或编程语言(例如,C编程语言)中的整数除法,$y_{偏移}$是预先确定的非负偏移,δy是两个相邻样本之间的距离。$y_{偏移}$的示例值可以包括0和h/8。δy的示例值可以包括1和h/4。

[0158] 在一个实施例中,对于宽度w和高度h是4的倍数的块,该块的顶边界样本的子采样位置可以表示为以下水平位置:

[0159] $x_1 = x_{偏移}$ (等式20)

[0160] $x_2 = x_1 + \delta x_1$ (等式21)

[0161] $x_3 = x_2 + \delta x_2$ (等式22)

[0162] $x_4 = x_3 + \delta x_3$ (等式23)

[0163] 其中,/表示硬件实现或编程语言(例如,C编程语言)中的整数除法,$x_{偏移}$是预先确定的非负偏移,δx_1、δx_2、δx_3是$x_4 < w$的两个相邻样本之间的正距离。

[0164] 块的左边界样本的子采样位置可以表示为以下垂直位置:

[0165] $y_1 = h-1-y_{偏移}$ (等式24)

[0166] $y_2 = y_1 - \delta y_1$ (等式25)

[0167] $y_3 = y_2 - \delta y_2$ (等式26)

[0168] $y_4 = y_3 - \delta y_3$ (等式27)

[0169] 其中,/表示C中的整数除法,$y_{偏移}$是预先确定的非负偏移量,δy_1、δy_2、δy_3是$y_4 \geq 4$的两个相邻样本之间的正距离。

[0170] 本公开包括可应用于具有符号预测过程的视频编码系统的方法,其中,可通过选择使当前块(例如,TB)的边界处的不连续性的成本最小化的符号来预测低频变换系数的符号。例如,本公开包括通过在成本计算中对像素进行子采样来减少成本计算的方法。

[0171] 根据本公开的各方面,子采样过程中的样本数量可以基于TB中符合符号预测过程的预定数量的低频分量。子采样过程中的样本数量可以独立于TB的大小。例如,适合符号预测过程的低频分量可以是TB的4×4最低频率变换系数。

[0172] 在一个实施例中,可以预先确定用于符号预测过程中的成本计算的TB的顶边界处的样本数量。例如,对于单指令多数据(Single Instruction Multiple Data,简称SIMD)实现,样本的预定数量可以是4或8。

[0173] 在一个实施例中,可以预先确定用于符号预测过程中的成本计算的TB左边界的样本数量。例如,对于SIMD实现,样本的预定数量可以是4或8。

[0174] 在一个实施例中,用于符号预测过程中的成本计算的TB的顶边界处的样本可以是均匀间隔的。

[0175] 在一个实施例中,用于符号预测过程中的成本计算的TB的左边界处的样本可以是均匀间隔的。

[0176] 在一些实施例中,子采样过程只能应用于某些大小的TB。例如,如果子采样过程中的样本数量是8,则子采样过程可以应用于16×16的TB,而且不能应用于8×8的TB。

[0177] 在一个实施例中,可以基于宽度w来预先确定TB的顶边界处的样本数量。例如,当宽度w大于阈值时,样本数量可以被预先确定为第一数量。否则,样本数量可以被预先确定为第二数量。对于w的至少一个值,预定数量可以小于w。例如,当样本的预定数量为4时,4x4 TB可以不被子采样,但是8x8 TB可以被子采样。

[0178] 在一个实施例中,可以基于高度h来预先确定TB的左边界处的样本数量。例如,当高度h大于阈值时,样本数量可以被预先确定为第一数量。否则,样本数量可以被预先确定为第二数量。对于至少一个h值,预定数量可以小于h。例如,当样本的预定数量为4时,4x4 TB可以不被子采样,但是8x8 TB可以被子采样。

[0179] 在一个实施例中,TB的顶边界处的采样位置可以基于宽度w预先确定,并且是均匀间隔的。例如,可以基于(等式12)-(等式15)来预先确定TB的顶边界处的采样位置。

[0180] 在一个实施例中,TB的顶边界处的采样位置可以基于宽度w预先确定,并且不是均匀间隔的。例如,可以基于(等式20)-(等式23)来预先确定TB的顶边界处的采样位置。

[0181] 在一个实施例中,可以基于宽度w预先确定TB的顶边界处的采样位置,并且至少两个采样位置彼此相邻。

[0182] 在一个实施例中,TB的左边界处的采样位置可以基于高度h来预先确定,并且是均匀间隔的。例如,可以基于(等式16)-(等式19)来预先确定TB的左边界处的采样位置。

[0183] 在一个实施例中,TB的左边界处的采样位置可以基于高度h来预先确定,并且不是均匀间隔的。例如,可以基于(等式24)-(等式27)来预先确定TB的左边界处的采样位置。

[0184] 在一个实施例中,可以基于高度h来预先确定TB的左边界处的采样位置,并且至少两个采样位置彼此相邻。

[0185] IV.流程图

[0186] 图14示出了概述根据本公开的实施例的示例性过程(1400)的流程图。在各种实施

例中,过程(1400)由处理电路执行,例如,终端设备(210)、(220)、(230)和(240)中的处理电路、执行视频编码器(303)的功能的处理电路、执行视频解码器(310)的功能的处理电路、执行视频编码器(503)的功能的处理电路等。在一些实施例中,过程(1400)在软件指令中实现,因此当处理电路执行软件指令时,处理电路执行过程(1400)。

[0187] 过程(1400)通常可以在步骤(S1410)开始,其中,过程(1400)接收包括当前块的编码残差的视频比特流。然后,过程(1400)进行到步骤(S1420)。

[0188] 在步骤(S1420),过程(1400)基于编码残差确定当前块的多个变换系数。然后,过程(1400)进行到步骤(S1430)。

[0189] 在步骤(S1430),过程(1400)通过基于当前块的最顶行或最左列中的至少一个中的样本子集对多个变换系数的子集执行符号预测,来为多个符号预测假设中的每一个符号预测假设计算相应符号预测假设的成本。样本子集中的样本数量是常数。然后,过程(1400)进行到步骤(S1440)。

[0190] 在步骤(S1440),过程(1400)从多个符号预测假设的成本中选择具有最小成本的多个符号预测假设中的一个符号预测假设。然后,过程(1400)进行到步骤(S1450)。

[0191] 在步骤(S1450),过程(1400)根据所选择的符号预测假设执行符号预测。然后,过程(1400)终止。

[0192] 在一个实施例中,常数是预先确定的。例如,常数可以是4或8。

[0193] 在一个实施例中,基于多个变换系数的子集的数量来确定常数。例如,可以预先确定多个变换系数的子集的数量(例如,TB中的4×4低频变换系数),并且可以确定常数小于多个变换系数的子集的预定数量。当多个变换系数的子集被预先确定为4×4低频变换系数时,对于最顶行或最左列,常数可以是4。

[0194] 在一个实施例中,基于块的一个维度和阈值之间的比较来确定常数。在一个示例中,当块的宽度(或高度)大于阈值时,最顶行(或最左列)中的常数可以是第一数量。否则,最顶行(或最左列)中的常数可以是小于第一数量的第二数量。

[0195] 在一个实施例中,基于块的一个维度与一个以上阈值之间的比较来确定常数。在一个示例中,块的尺寸可以与两个阈值进行比较。当块的宽度(或高度)不大于第一阈值时,最顶行(或最左列)中的常数可以是第一数量。当块的宽度(或高度)大于第一阈值且小于第二阈值时,最顶行(或最左列)中的常数可以是第二数量。当块的宽度(或高度)大于第二阈值时,最顶行(或最左列)中的常数可以是第三数量。在另一个示例中,块的维度可以与不同数量的阈值进行比较。

[0196] 在一个实施例中,基于包括常数的多个值的查找表来确定上述常数。多个值中的每一个值都与块的一个维度的不同值相关联。例如,当块的宽度(或高度)在第一范围内时,根据查找表,最顶行(或最左列)中的常数可以是第一数量。当块的宽度(或高度)在第二范围内时,根据查找表,最顶行(或最左列)中的常数可以是第二数量,依此类推。

[0197] 在一个实施例中,基于块的一个维度来确定当前块的最顶行或最左列中的至少一个中的样本子集的位置。例如,基于块的宽度(或高度)来确定块的最顶行(或最左列)中的样本子集的位置。

[0198] 在一个实施例中,当前块的最顶行或最左列中的至少一个中的样本子集的至少两个位置是相邻的。例如,最顶行(或最左列)中的两个样本的位置彼此相邻。

[0199] 在一个实施例中,当前块的最顶行或最左列中的至少一个中的样本子集的位置是均匀间隔的。例如,在图12中,最顶行中的四个样本$p_{1,0}$、$p_{3,0}$、$p_{5,0}$和$p_{7,0}$(或最左列中的$p_{0,0}$、$p_{0,2}$、$p_{0,4}$和$p_{0,6}$)的位置是均匀间隔的。

[0200] 在一个实施例中,过程(1400)基于(i)当前块的最顶行或最左列中的至少一个中的样本子集,以及(ii)当前块上方的行或当前块左侧的列中的至少一个中的相邻重构样本,来为多个符号预测假设中的每一个符号预测假设计算相应符号预测假设的成本。

[0201] V.计算机系统

[0202] 上述技术可以被实现为使用计算机可读指令的计算机软件,并且物理地存储在一个或多个计算机可读介质中。例如,图15示出了适于实现所公开主题的某些实施例的计算机系统(1500)。

[0203] 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、链接或类似机制来创建包括指令的代码,这些指令可以由一个或多个计算机中央处理单元(Central Processing Units,简称CPU)、图形处理单元(Graphics Processing Units,简称GPU)等直接执行,或者通过解释、微代码执行等来执行。

[0204] 指令可以在各种类型的计算机或其组件上执行,包括例如个人计算机、平板计算机、服务器、智能手机、游戏设备、物联网设备等。

[0205] 图15中所示的计算机系统(1500)的组件本质上是示例性的,并且不旨在对实现本公开的实施例的计算机软件的使用范围或功能提出任何限制。组件的配置也不应被解释为对计算机系统(1500)的示例性实施例中所示的任何一个组件或组件组合有任何依赖性或要求。

[0206] 计算机系统(1500)可以包括某些人机接口输入设备。这种人机接口输入设备可以响应一个或多个人类用户的输入,例如触觉输入(例如:击键、滑动、数据手套移动)、音频输入(例如:语音、鼓掌)、视觉输入(例如:手势)、嗅觉输入(未示出)进行的输入。人机接口设备还可以用于捕捉不一定与人的有意识输入直接相关的某些媒体,例如,音频(例如:语音、音乐、环境声音)、图像(例如:扫描图像、从静止图像相机获得的照片图像)、视频(例如,二维视频、包括立体视频的三维视频)。

[0207] 输入人机接口设备可以包括以下一个或多个(每个仅描绘了一个):键盘(1501)、鼠标(1502)、轨迹板(1503)、触摸屏(1510)、数据手套(未示出)、操纵杆(1505)、麦克风(1506)、扫描仪(1507)、相机(1508)。

[0208] 计算机系统(1500)还可以包括某些人机接口输出设备。这种人机接口输出设备可以刺激一个或多个人类用户的感觉,例如通过触觉输出、声音、光和气味/味道。这种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(1510)、数据手套(未示出)或操纵杆(1505)的触觉反馈,但是也可以有不用作输入设备的触觉反馈设备)、音频输出设备(例如:扬声器(1509)、耳机(未示出))、视觉输出设备(例如,屏幕(1510),包括CRT屏幕、LCD屏幕、等离子屏幕、OLED屏幕,每个都具有或不具有触摸屏输入能力,每个都具有或不具有触觉反馈能力——其中一些能够输出二维视觉输出或多于三维的输出,通过诸如立体输出之类的方式;虚拟现实眼镜(未示出)、全息显示器和烟雾箱(未示出))以及打印机(未示出)。这些视觉输出设备(例如,屏幕(1510))可以通过图形适配器(1550)连接到系统总线(1548)。

[0209] 计算机系统(1500)还可以包括人类可访问的存储设备及其相关联的介质,例如,

包括具有CD/DVD或类似介质(1521)的CD/DVD ROM/RW(1520)的光学介质、拇指驱动器(1522)、可移动硬盘驱动器或固态驱动器(1523)、诸如磁带和软盘(未示出)之类的传统磁介质、诸如安全加密狗(未示出)之类的专用ROM/ASIC/PLD设备等。

[0210] 本领域技术人员还应该理解,结合当前公开的主题使用的术语“计算机可读介质”不包括传输介质、载波或其他瞬时信号。

[0211] 计算机系统(1500)还可以包括到一个或多个通信网络(1555)的通信网络接口(1554)。一个或多个通信网络(1555)例如可以是无线的、有线的、光学的。一个或多个通信网络(1555)还可以是局域的、广域的、大都市的、车辆的和工业的、实时的、延迟容忍的等。一个或多个通信网络(1555)的示例包括诸如以太网、无线LAN之类的局域网、包括GSM、3G、4G、5G、LTE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或无线广域数字网络、包括CANBus的车辆和工业网络等。某些网络通常需要连接到某些通用数据端口或外围总线(1549)(例如,计算机系统(1500)的USB端口)的外部网络接口适配器;其他的通常通过连接到如下所述的系统总线而集成到计算机系统(1500)的核心中(例如,PC计算机系统中的以太网接口或智能电话计算机系统中的蜂窝网络接口)。使用这些网络中的任何一个,计算机系统(1500)可以与其他实体通信。这种通信可以是单向的、只接收的(例如,广播电视)、单向的、只发送的(例如,到某些CANbus设备的CANbus),或者是双向的,例如,到使用局域或广域数字网络的其他计算机系统。如上所述,某些协议和协议栈可以用在这些网络和网络接口的每一个上。

[0212] 前述人机接口设备、人类可访问的存储设备和网络接口可以附接到计算机系统(1500)的核心(1540)。

[0213] 核心(1540)可以包括一个或多个中央处理单元(Central Processing Units,简称CPU)(1541)、图形处理单元(Graphics Processing Units,简称GPU)(1542)、现场可编程门区域(Field Programmable Gate Areas,简称FPGA)(1543)形式的专用可编程处理单元、用于特定任务的硬件加速器(1544)、图形适配器(1550)等。这些设备连同只读存储器(Read-only memory,简称ROM)(1545)、随机存取存储器(1546)、诸如内部非用户可访问硬盘驱动器、SSD之类的内部大容量存储器(1547)可以通过系统总线(1548)连接。在一些计算机系统中,系统总线(1548)可以以一个或多个物理插头的形式访问,以允许额外CPU、GPU等的扩展。外围设备可以直接或者通过外围总线(1549)连接到核心的系统总线(1548)。在一个示例中,屏幕(1510)可以连接到图形适配器(1550)。外围总线的架构包括PCI、USB等。

[0214] CPU(1541)、GPU(1542)、FPGA(1543)和加速器(1544)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(1545)或RAM(1546)中。过渡数据也可以存储在RAM(1546)中,而永久数据可以存储在例如内部大容量存储器(1547)中。可以通过使用高速缓冲存储器来实现对任何存储设备的快速存储和检索,高速缓冲存储器可以与一个或多个CPU(1541)、GPU(1542)、大容量存储器(1547)、ROM(1545)、RAM(1546)等紧密关联。

[0215] 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了本公开的目的而专门设计和构造的,或者可以是计算机软件领域的技术人员公知和可获得的类型。

[0216] 作为示例而非限制,具有架构的计算机系统(1500),特别是核心(1540)可以提供

作为处理器(包括CPU、GPU、FPGA、加速器等)的结果的功能,其执行嵌入在一个或多个有形的计算机可读介质中的软件。这种计算机可读介质可以是与如上所述的用户可访问的大容量存储器相关联的介质以及具有非暂时性的核心(1540)的某些存储器,例如,核心内部大容量存储器(1547)或ROM(1545)。实现本公开的各种实施例的软件可以存储在这样的设备中并由核心(1540)执行。根据特定需要,计算机可读介质可以包括一个或多个存储设备或芯片。该软件可以使核心(1540)并且特别是其中的处理器(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(1546)中的数据结构并且根据软件定义的过程修改这样的数据结构。此外或作为替代,计算机系统提供作为硬连线或以其他方式包含在电路(例如:加速器(1144))中的逻辑的结果的功能,其可以代替软件或与软件一起操作来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括逻辑,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有所要执行的软件的电路(例如,集成电路(Integrated Circuit,简称IC))、包含用于执行的逻辑的电路或者这两者。本公开包含硬件和软件的任何合适的组合。

[0217] 虽然本公开已经描述了几个示例性实施例,但是存在落入本公开范围内的变更、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许多系统和方法,尽管在此没有明确示出或描述,但是这些系统和方法体现了本公开的原理,并且因此在本公开的精神和范围内。

[0218] 附录A:缩略语

[0219] ALF:自适应环路滤波器(Adaptive Loop Filter)

[0220] AMVP:高级运动矢量预测(Advanced Motion Vector Prediction)

[0221] APS:自适应参数集(Adaptation Parameter Set)

[0222] ASIC:专用集成电路(Application-Specific Integrated Circuit)

[0223] ATMVP:替代/高级时间运动矢量预测(Alternative/Advanced Temporal Motion Vector Prediction)

[0224] AV1:AOMedia视频1

[0225] AV2:AOMedia视频2

[0226] BMS:基准集(Benchmark Set)

[0227] BV:块矢量(Block Vector)

[0228] CANBus:控制器区域网络总线(Controller Area Network Bus)

[0229] CB:编码块(Coding Block)

[0230] CC-ALF:交叉分量自适应环路滤波器(Cross-Component Adaptive Loop Filter)

[0231] CD:光盘(Compact Disc)

[0232] CDEF:约束方向增强滤波器(Constrained Directional Enhancement Filter)

[0233] CPR:当前图片参考(Current Picture Referencing)

[0234] CPU:中央处理器(Central Processing Unit)

[0235] CRT:阴极射线管(Cathode Ray Tube)

[0236] CTB:编码树块(Coding Tree Block)

[0237] CTU:编码树单元(Coding Tree Unit)

[0238] CU:编码单元(Coding Unit)

[0239] DPB:解码器图片缓冲器(Decoder Picture Buffer)

[0240] DPCM:差分脉冲编码调制(Differential Pulse-Code Modulation)

[0241] DPS:解码参数集(Decoding Parameter Set)

[0242] DVD:数字视频光盘(Digital Video Disc)

[0243] FPGA:现场可编程门区域(Field Programmable Gate Area)

[0244] JCCR:共同CbCr残差编码(Joint CbCr Residual Coding)

[0245] JVET:共同视频探索小组(Joint Video Exploration Team)

[0246] GOP:图片组(Groups of Pictures)

[0247] GPU:图形处理单元(Graphics Processing Unit)

[0248] GSM:全球移动通信系统(Global System for Mobile communications)

[0249] HDR:高动态范围(High Dynamic Range)

[0250] HEVC:高效视频编码(High Efficiency Video Coding)

[0251] HRD:假设参考解码器(Hypothetical Reference Decoder)

[0252] IBC:帧内块复制(Intra Block Copy)

[0253] IC:集成电路(Integrated Circuit)

[0254] ISP:帧内子分区(Intra Sub-Partitions)

[0255] JEM:共同探索模式(Joint Exploration Model)

[0256] LAN:局域网(Local Area Network)

[0257] LCD:液晶显示器(Liquid-Crystal Display)

[0258] LR:环路恢复滤波器(Loop Restoration Filter)

[0259] LRU:环路恢复单元(Loop Restoration Unit)

[0260] LTE:长期演进(Long-Term Evolution)

[0261] MPM:最可能模式(Most Probable Mode)

[0262] MV:运动矢量(Motion Vector)

[0263] OLED:有机发光二极管(Organic Light-Emitting Diode)

[0264] PB:预测块(Prediction Blocks)

[0265] PCI:外围组件互连(Peripheral Component Interconnect)

[0266] PDPC:位置相关预测组合(Position Dependent Prediction Combination)

[0267] PLD:可编程逻辑器件(Programmable Logic Device)

[0268] PPS:图片参数集(Picture Parameter Set)

[0269] PU:预测单元(Prediction Unit)

[0270] RAM:随机存取存储器(Random Access Memory)

[0271] ROM:只读存储器(Read-Only Memory)

[0272] SAO:样本自适应偏移(Sample Adaptive Offset)

[0273] SCC:屏幕内容编码(Screen Content Coding)

[0274] SDR:标准动态范围(Standard Dynamic Range)

[0275] SEI:补充增强信息(Supplementary Enhancement Information)

[0276] SNR:信噪比(Signal Noise Ratio)

[0277] SPS:序列参数集(Sequence Parameter Set)

[0278] SSD:固态硬盘(Solid-state Drive)

[0279] TU:变换单元(Transform Unit)

[0280] USB:通用串行总线(Universal Serial Bus)

[0281] VPS:视频参数集(Video Parameter Set)

[0282] VUI:视频可用性信息(Video Usability Information)

[0283] VVC:多功能视频编码(Versatile Video Coding)

[0284] WAIP:广角帧内预测(Wide-Angle Intra Prediction)

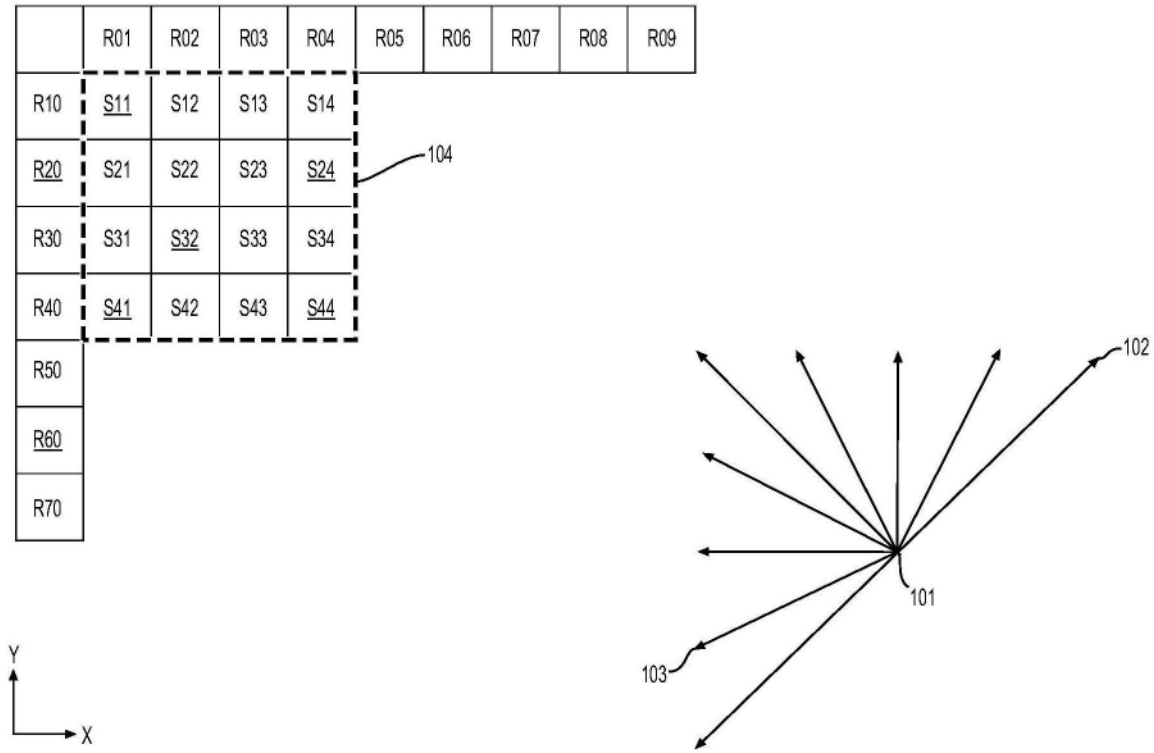

R01
R02
R03
R04
R05
R06
R07
R08
R09
R10
S11
S12
S13
S14
R20
S21
S22
S23
S24
R30
S31
S32
S33
S34
R40
S41
S42
S43
S44
R50
R60
R70
104
102
101
103
Y
X

图1A(相关技术)

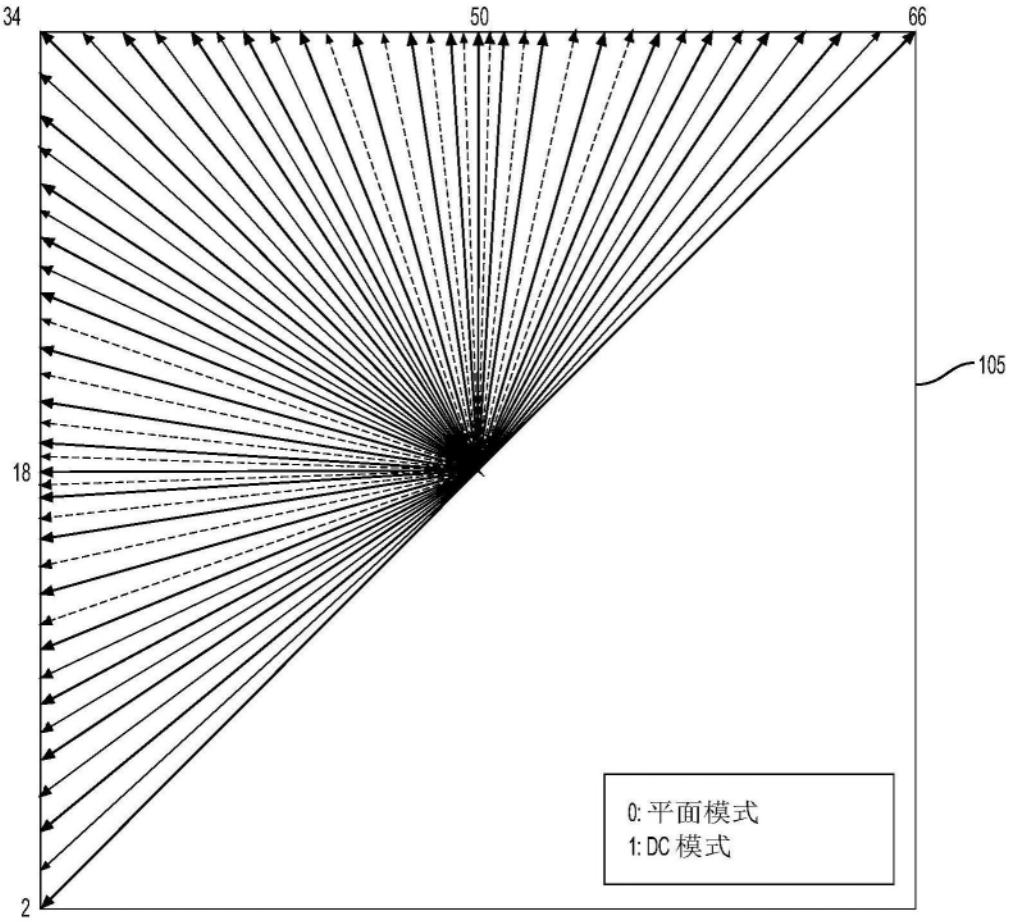


图1B(相关技术)

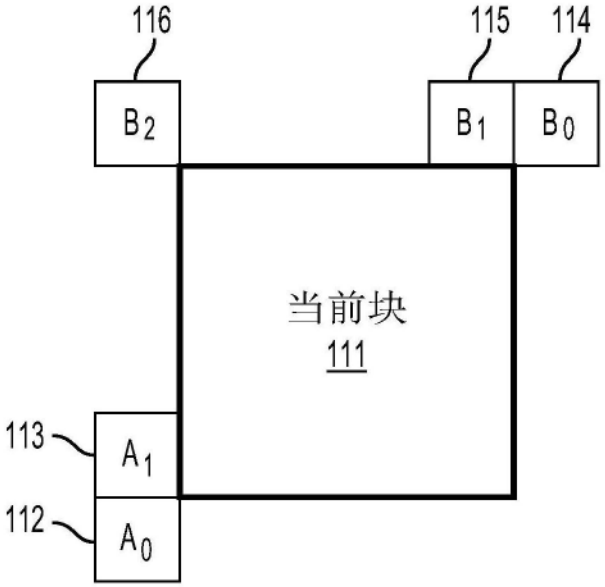


图1C(相关技术)

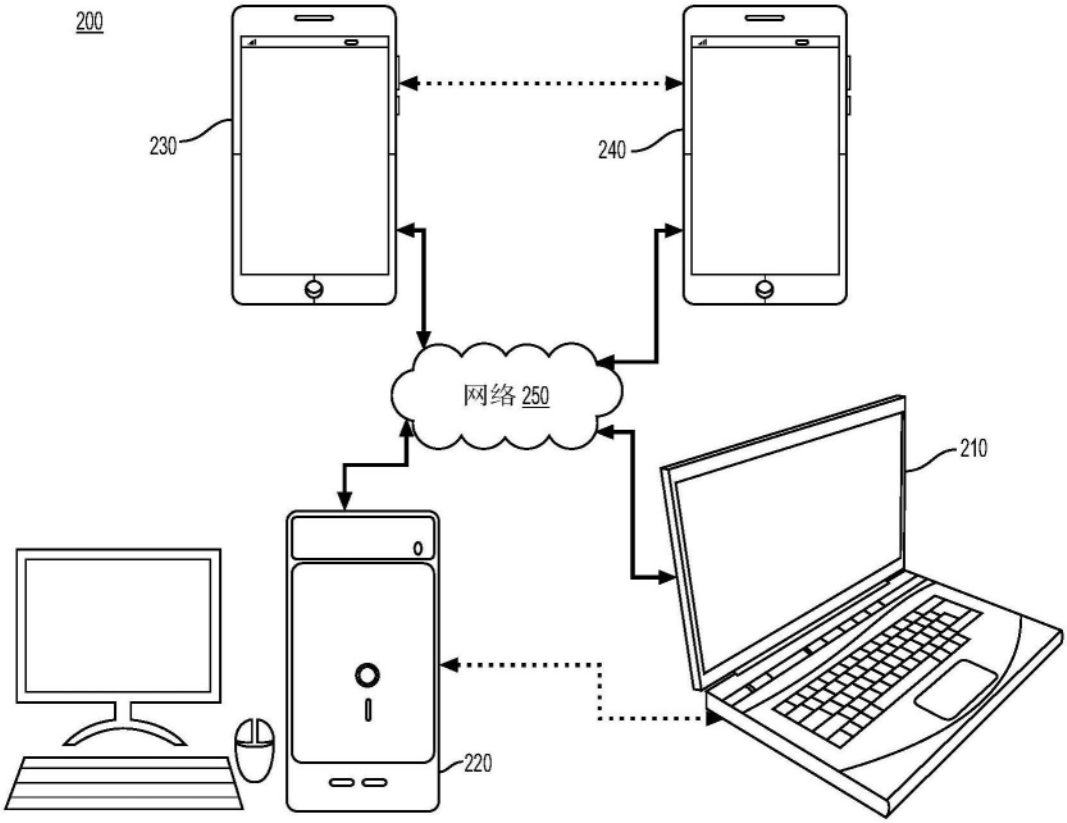


图2

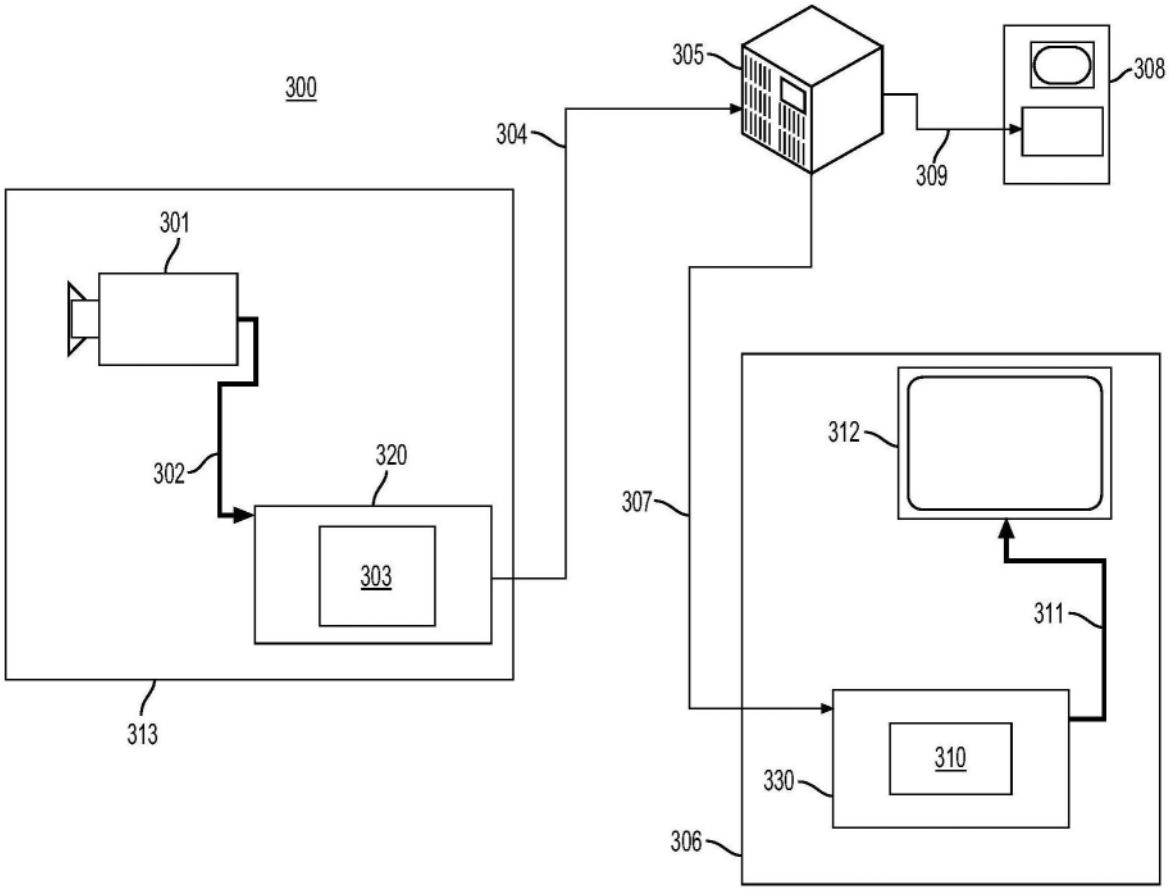


图3

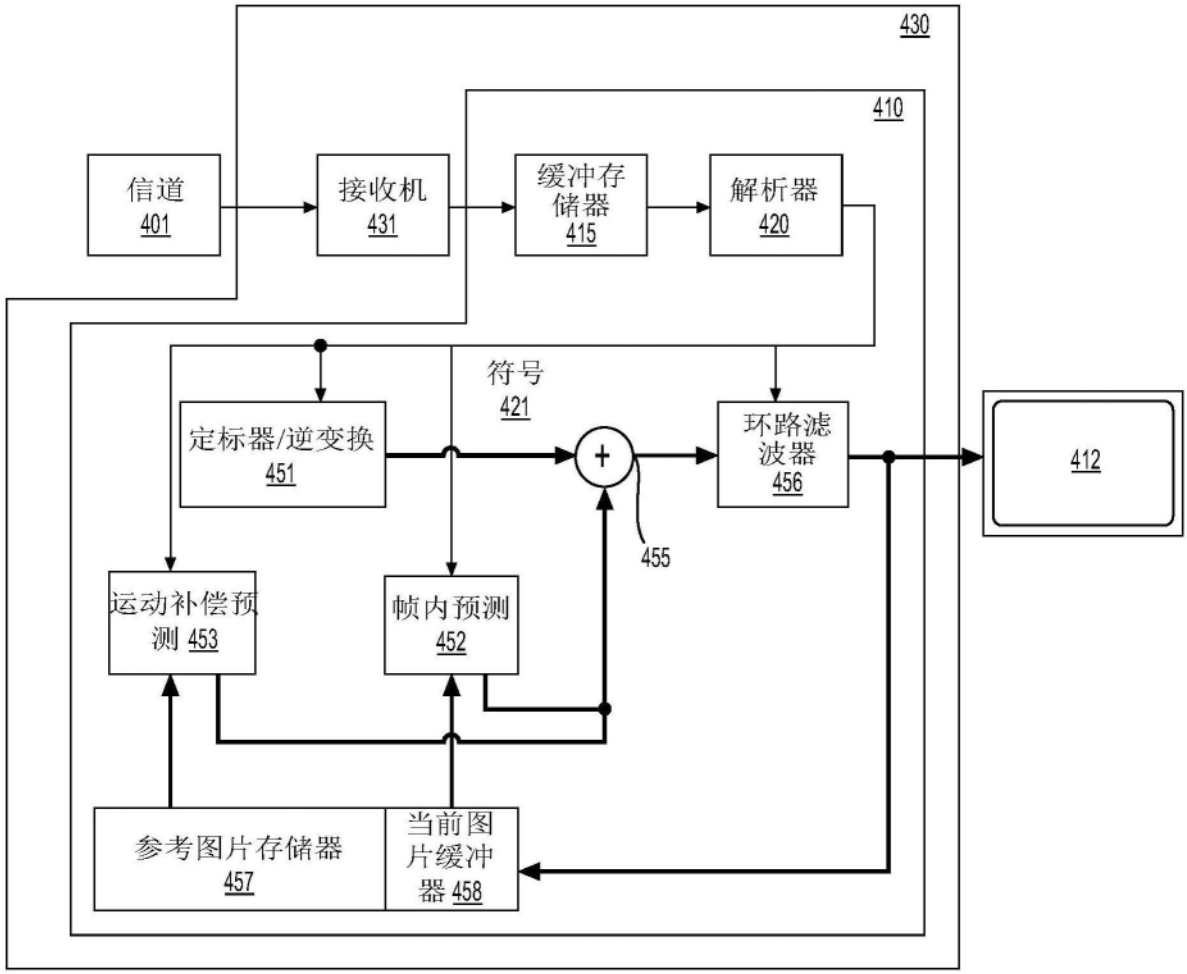

430
410
信道
401
接收机
431
缓冲存储器
415
解析器
420
符号
421
定标器/逆变换
451
环路滤波器
456
412
455
运动补偿预测 453
帧内预测
452
参考图片存储器
457
当前图片缓冲器 458

图4

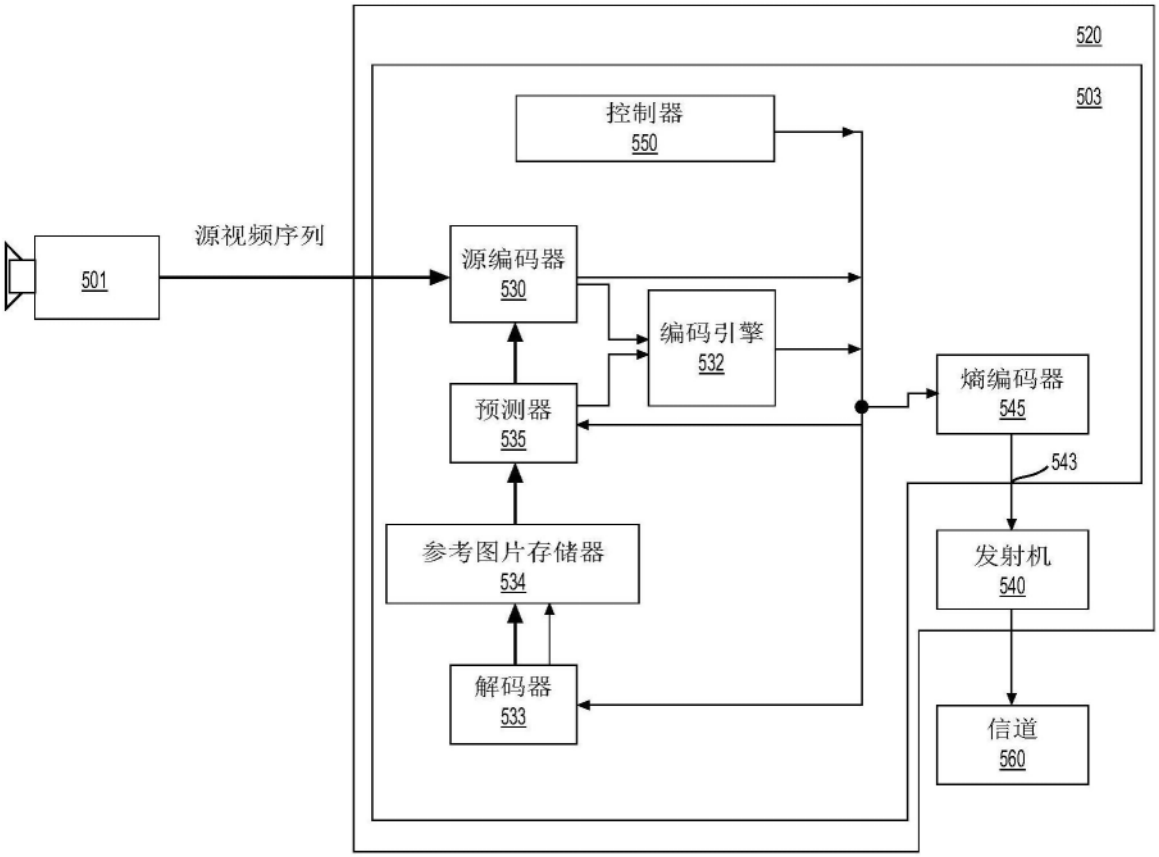

520
503
控制器
550
源视频序列
501
源编码器
530
编码引擎
532
预测器
535
熵编码器
545
543
参考图片存储器
534
发射机
540
解码器
533
信道
560

图5

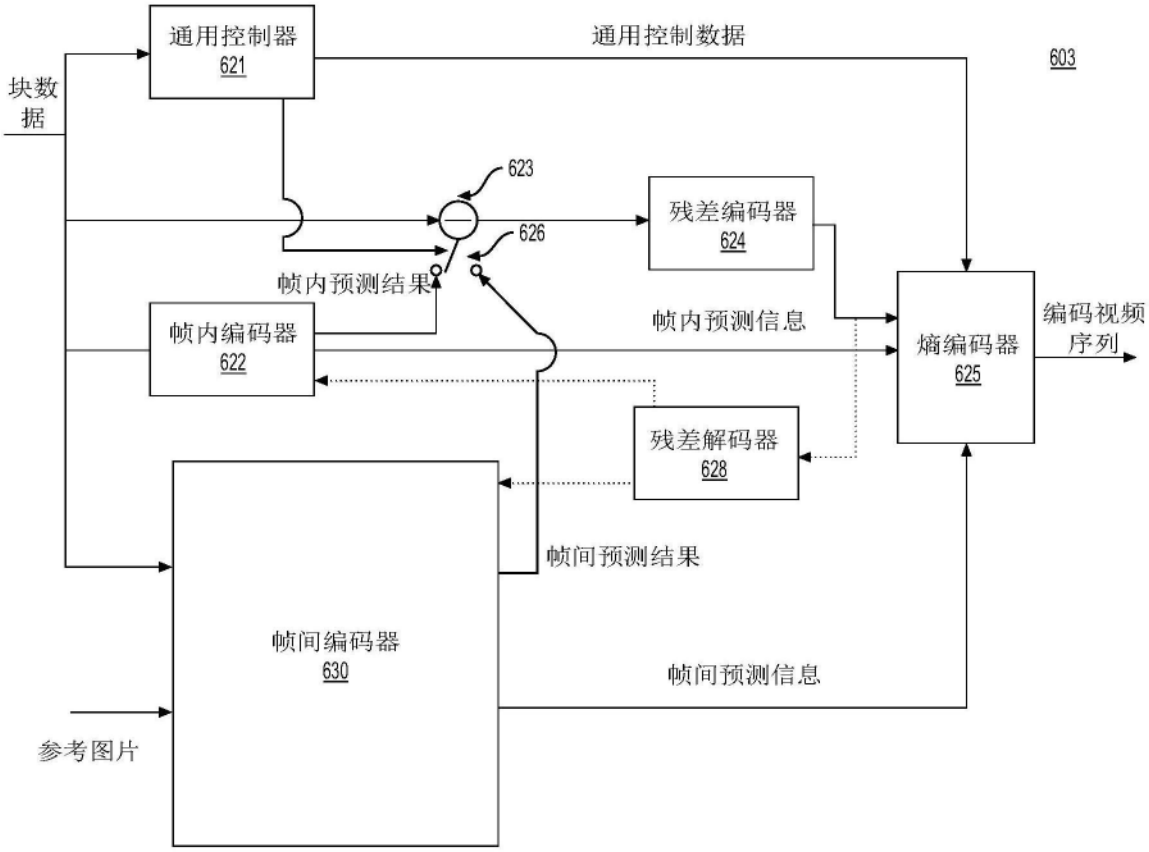

通用控制器
621
通用控制数据
603
块数据
623
残差编码器
624
626
帧内预测结果
帧内预测信息
帧内编码器
622
熵编码器
625
编码视频序列
残差解码器
628
帧间预测结果
帧间编码器
630
帧间预测信息
参考图片

图6

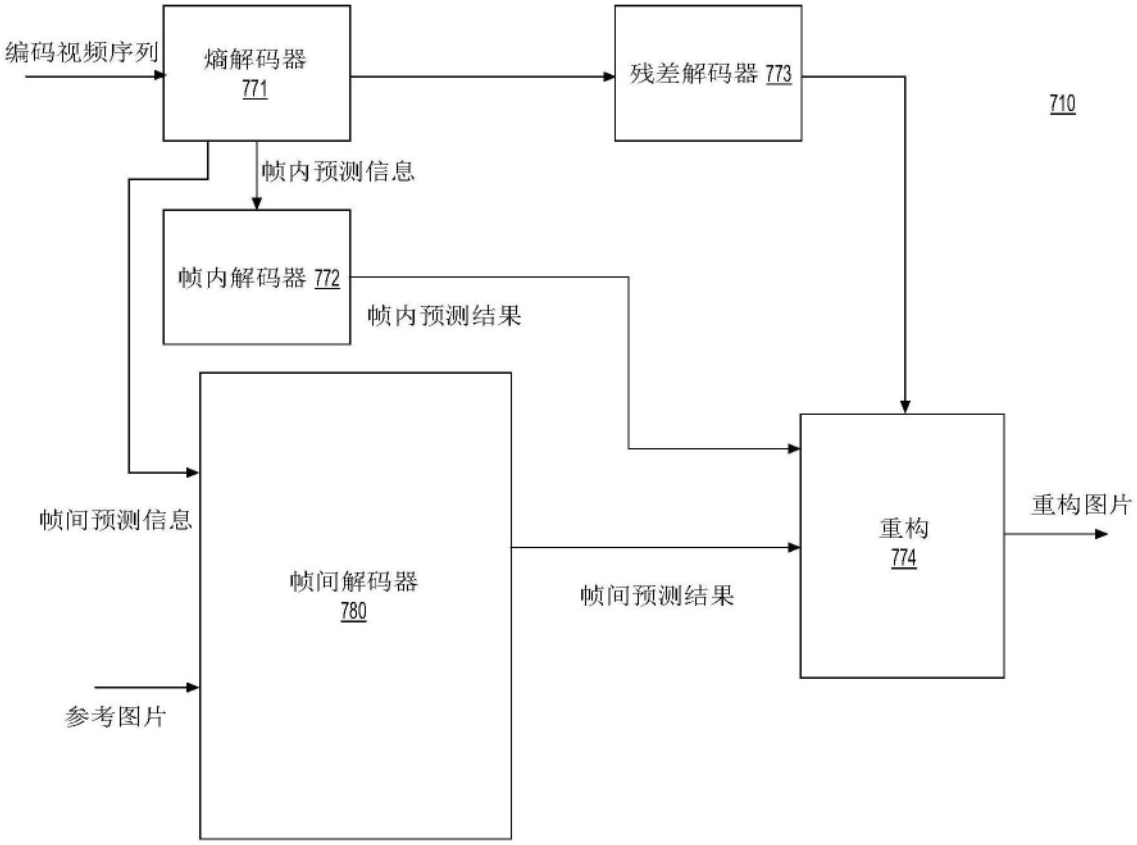

编码视频序列
熵解码器
771
残差解码器 773
710
帧内预测信息
帧内解码器 772
帧内预测结果
帧间预测信息
帧间解码器
780
参考图片
帧间预测结果
重构
774
重构图片

图7

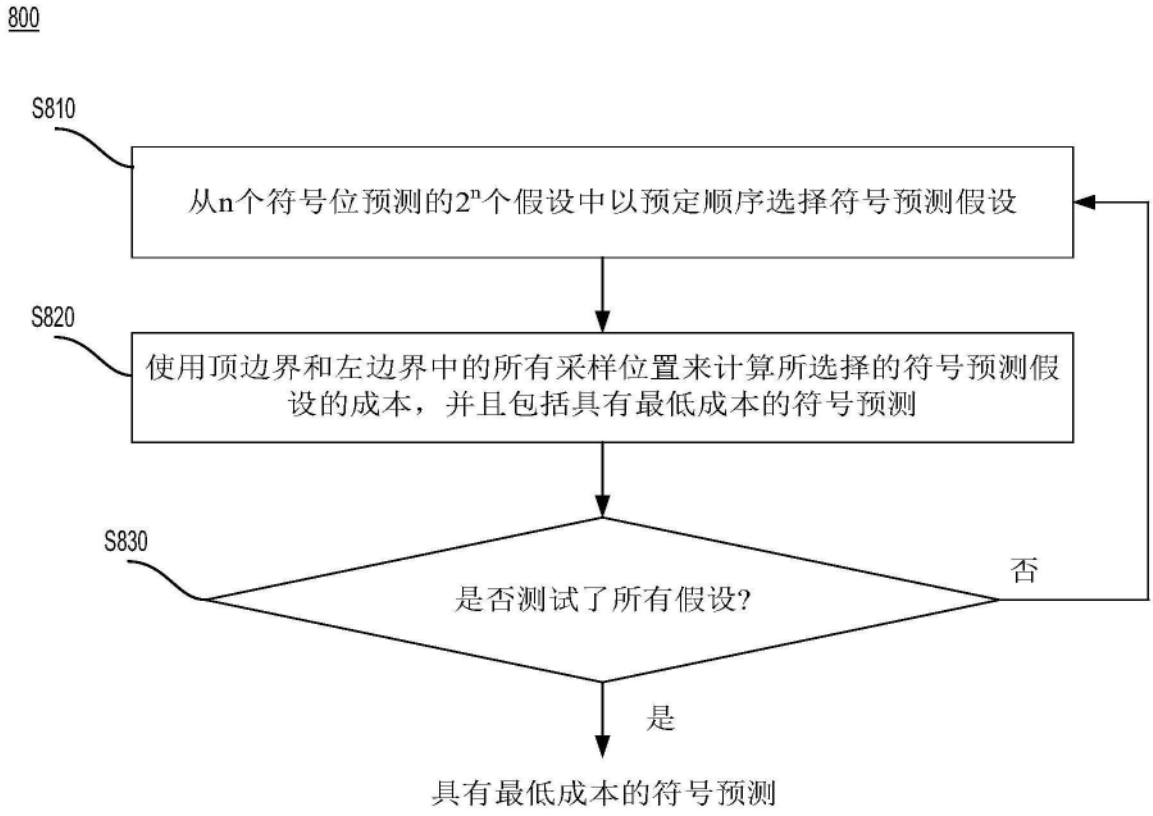

800
S810
从n个符号位预测的2^n个假设中以预定顺序选择符号预测假设
S820
使用顶边界和左边界中的所有采样位置来计算所选择的符号预测假设的成本，并且包括具有最低成本的符号预测
S830
是否测试了所有假设?
否
是
具有最低成本的符号预测

图8

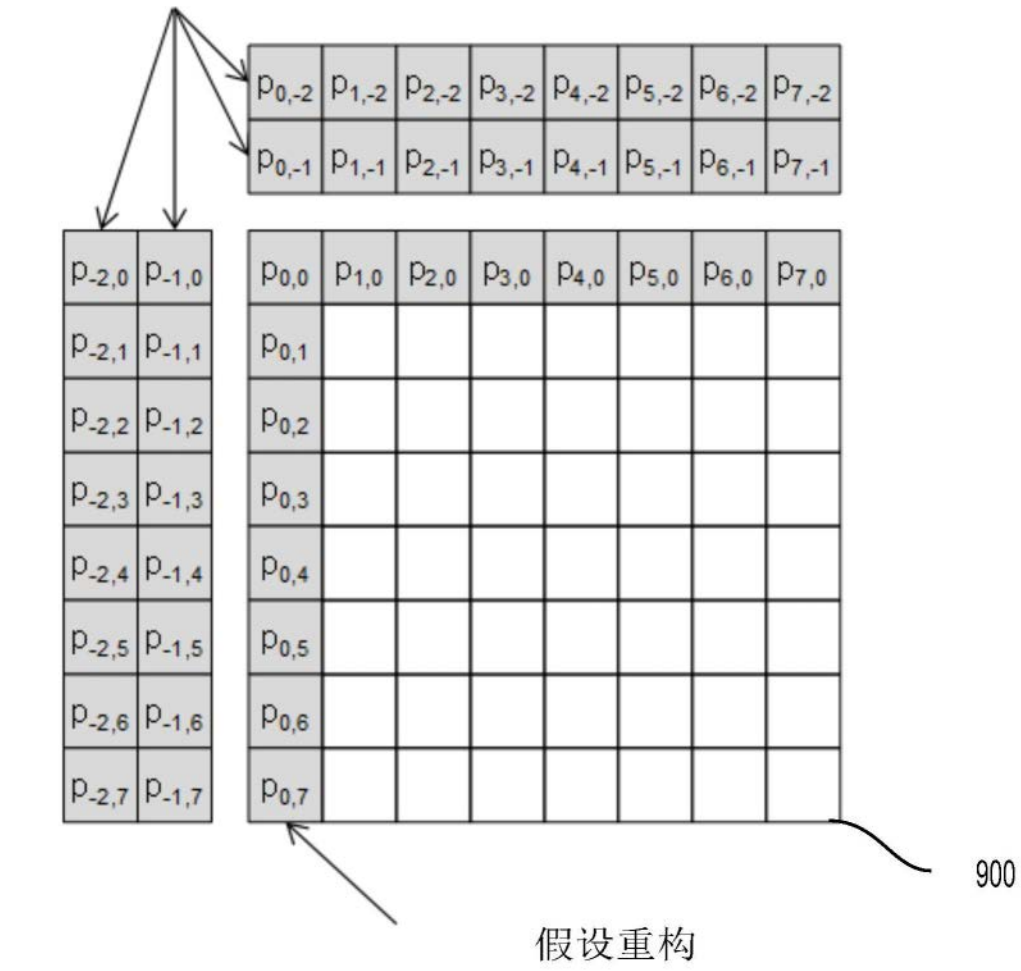

先前重构的相邻像素
$p_{0,-2}$ $p_{1,-2}$ $p_{2,-2}$ $p_{3,-2}$ $p_{4,-2}$ $p_{5,-2}$ $p_{6,-2}$ $p_{7,-2}$
$p_{0,-1}$ $p_{1,-1}$ $p_{2,-1}$ $p_{3,-1}$ $p_{4,-1}$ $p_{5,-1}$ $p_{6,-1}$ $p_{7,-1}$
$p_{-2,0}$ $p_{-1,0}$
$p_{-2,1}$ $p_{-1,1}$
$p_{-2,2}$ $p_{-1,2}$
$p_{-2,3}$ $p_{-1,3}$
$p_{-2,4}$ $p_{-1,4}$
$p_{-2,5}$ $p_{-1,5}$
$p_{-2,6}$ $p_{-1,6}$
$p_{-2,7}$ $p_{-1,7}$
$p_{0,0}$ $p_{1,0}$ $p_{2,0}$ $p_{3,0}$ $p_{4,0}$ $p_{5,0}$ $p_{6,0}$ $p_{7,0}$
$p_{0,1}$
$p_{0,2}$
$p_{0,3}$
$p_{0,4}$
$p_{0,5}$
$p_{0,6}$
$p_{0,7}$
900
假设重构

图9

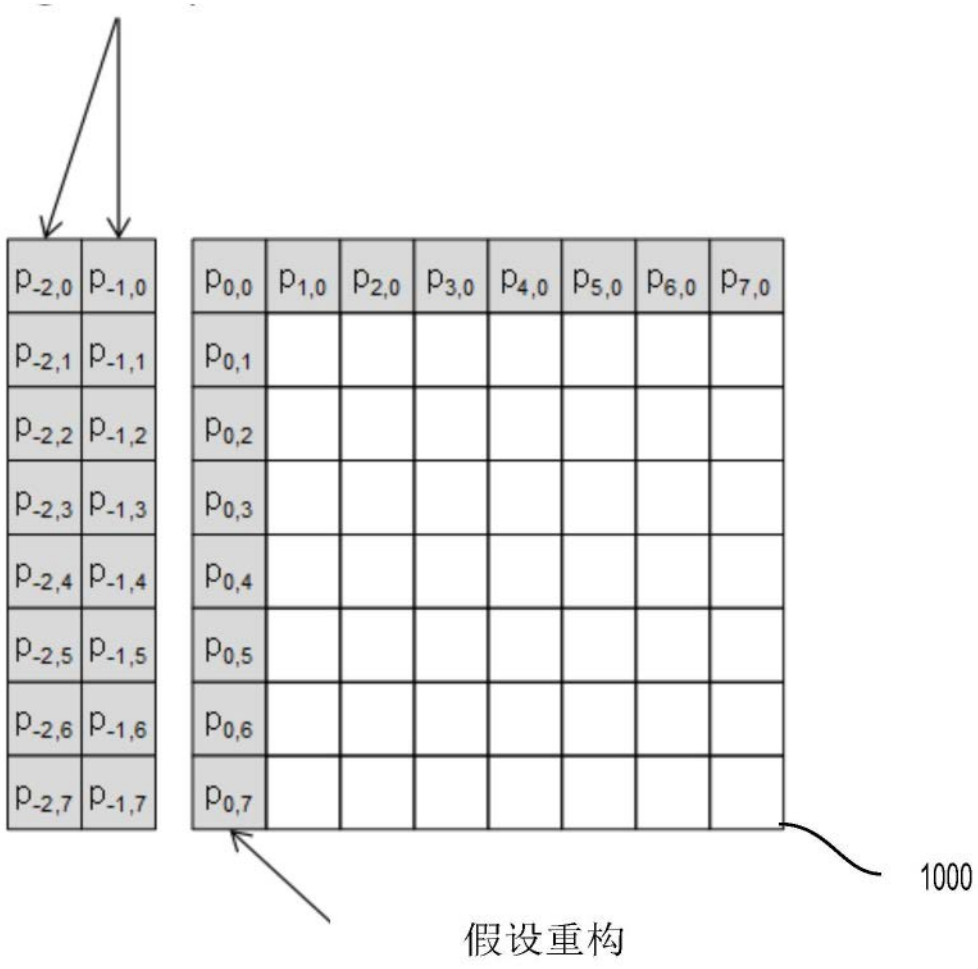

先前重构的相邻像素
$p_{-2,0}$ $p_{-1,0}$
$p_{-2,1}$ $p_{-1,1}$
$p_{-2,2}$ $p_{-1,2}$
$p_{-2,3}$ $p_{-1,3}$
$p_{-2,4}$ $p_{-1,4}$
$p_{-2,5}$ $p_{-1,5}$
$p_{-2,6}$ $p_{-1,6}$
$p_{-2,7}$ $p_{-1,7}$
$p_{0,0}$ $p_{1,0}$ $p_{2,0}$ $p_{3,0}$ $p_{4,0}$ $p_{5,0}$ $p_{6,0}$ $p_{7,0}$
$p_{0,1}$
$p_{0,2}$
$p_{0,3}$
$p_{0,4}$
$p_{0,5}$
$p_{0,6}$
$p_{0,7}$
1000
假设重构

图10

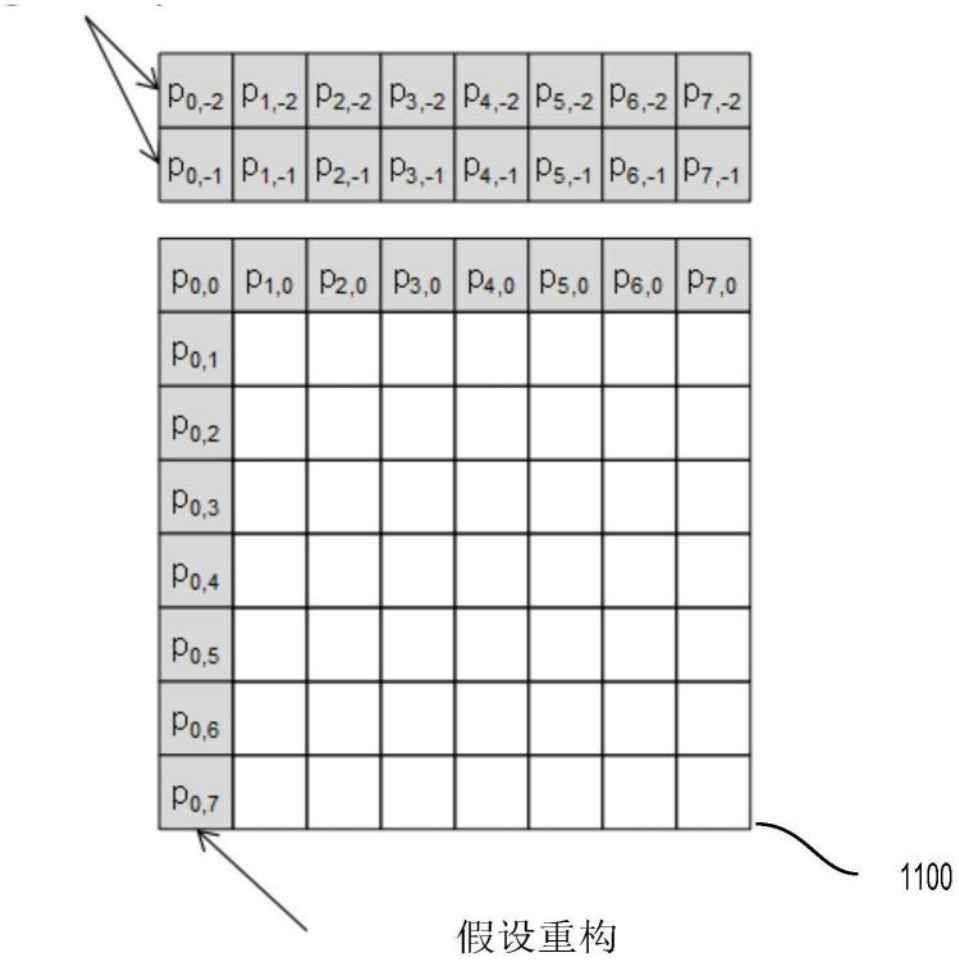

先前重构的相邻像素
$p_{0,-2}$ $p_{1,-2}$ $p_{2,-2}$ $p_{3,-2}$ $p_{4,-2}$ $p_{5,-2}$ $p_{6,-2}$ $p_{7,-2}$
$p_{0,-1}$ $p_{1,-1}$ $p_{2,-1}$ $p_{3,-1}$ $p_{4,-1}$ $p_{5,-1}$ $p_{6,-1}$ $p_{7,-1}$
$p_{0,0}$ $p_{1,0}$ $p_{2,0}$ $p_{3,0}$ $p_{4,0}$ $p_{5,0}$ $p_{6,0}$ $p_{7,0}$
$p_{0,1}$
$p_{0,2}$
$p_{0,3}$
$p_{0,4}$
$p_{0,5}$
$p_{0,6}$
$p_{0,7}$
1100
假设重构

图11

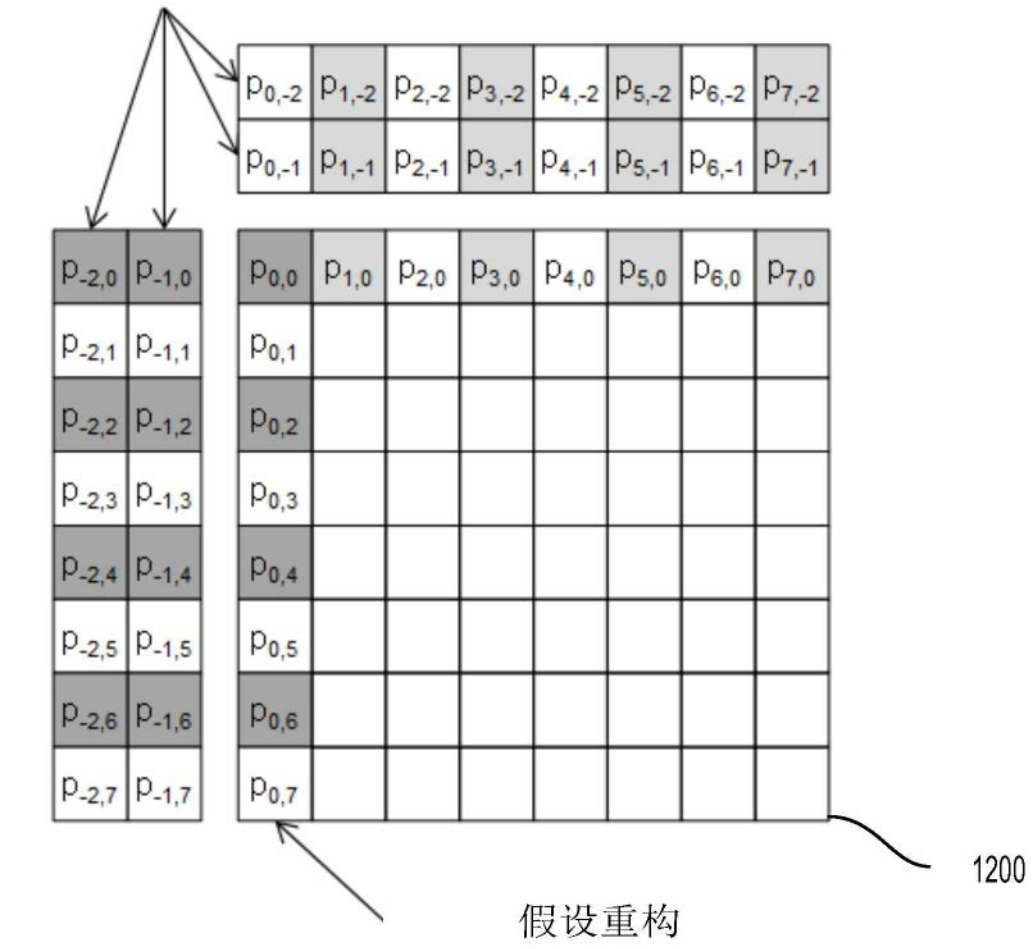

先前重构的相邻像素
$p_{0,-2}$ $p_{1,-2}$ $p_{2,-2}$ $p_{3,-2}$ $p_{4,-2}$ $p_{5,-2}$ $p_{6,-2}$ $p_{7,-2}$
$p_{0,-1}$ $p_{1,-1}$ $p_{2,-1}$ $p_{3,-1}$ $p_{4,-1}$ $p_{5,-1}$ $p_{6,-1}$ $p_{7,-1}$
$p_{-2,0}$ $p_{-1,0}$ $p_{0,0}$ $p_{1,0}$ $p_{2,0}$ $p_{3,0}$ $p_{4,0}$ $p_{5,0}$ $p_{6,0}$ $p_{7,0}$
$p_{-2,1}$ $p_{-1,1}$ $p_{0,1}$
$p_{-2,2}$ $p_{-1,2}$ $p_{0,2}$
$p_{-2,3}$ $p_{-1,3}$ $p_{0,3}$
$p_{-2,4}$ $p_{-1,4}$ $p_{0,4}$
$p_{-2,5}$ $p_{-1,5}$ $p_{0,5}$
$p_{-2,6}$ $p_{-1,6}$ $p_{0,6}$
$p_{-2,7}$ $p_{-1,7}$ $p_{0,7}$
1200
假设重构

图12

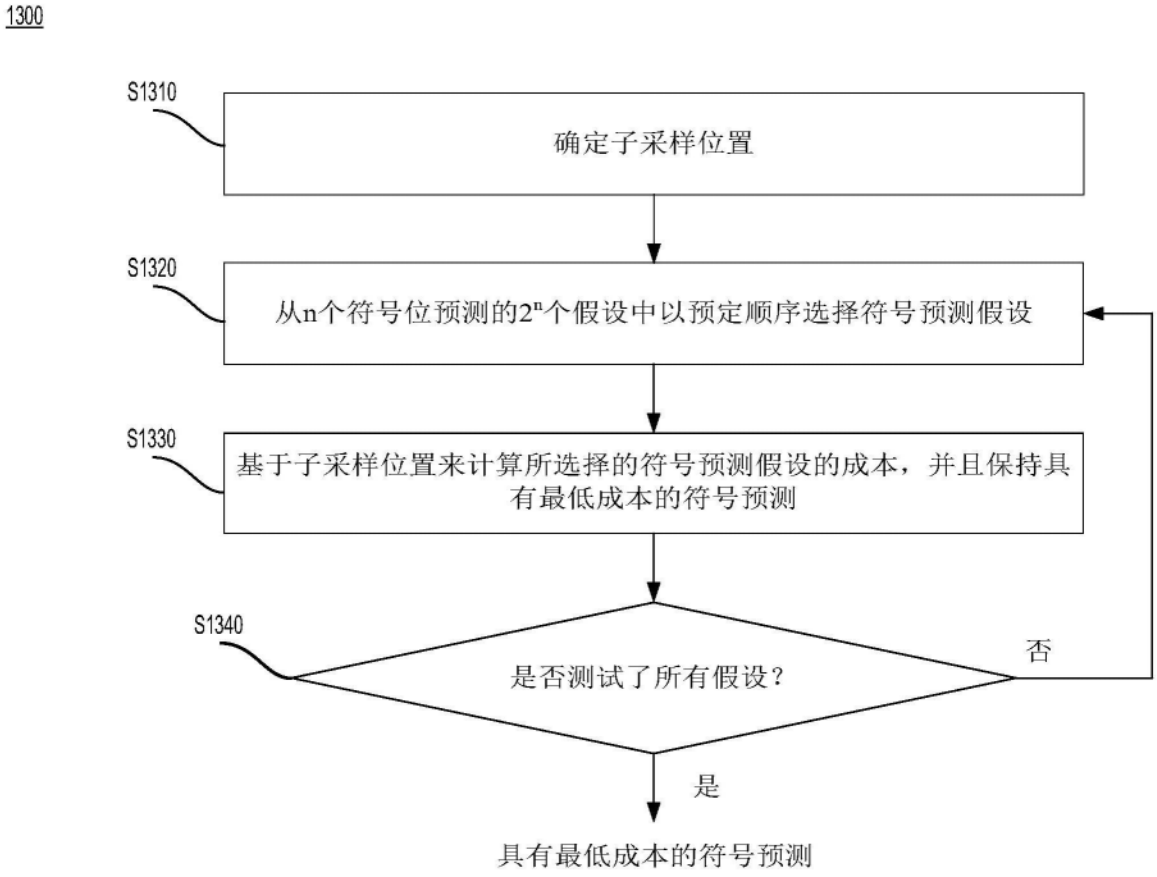

1300
S1310
确定子采样位置
S1320
从n个符号位预测的2ⁿ个假设中以预定顺序选择符号预测假设
S1330
基于子采样位置来计算所选择的符号预测假设的成本，并且保持具有最低成本的符号预测
S1340
是否测试了所有假设？
否
是
具有最低成本的符号预测

图13

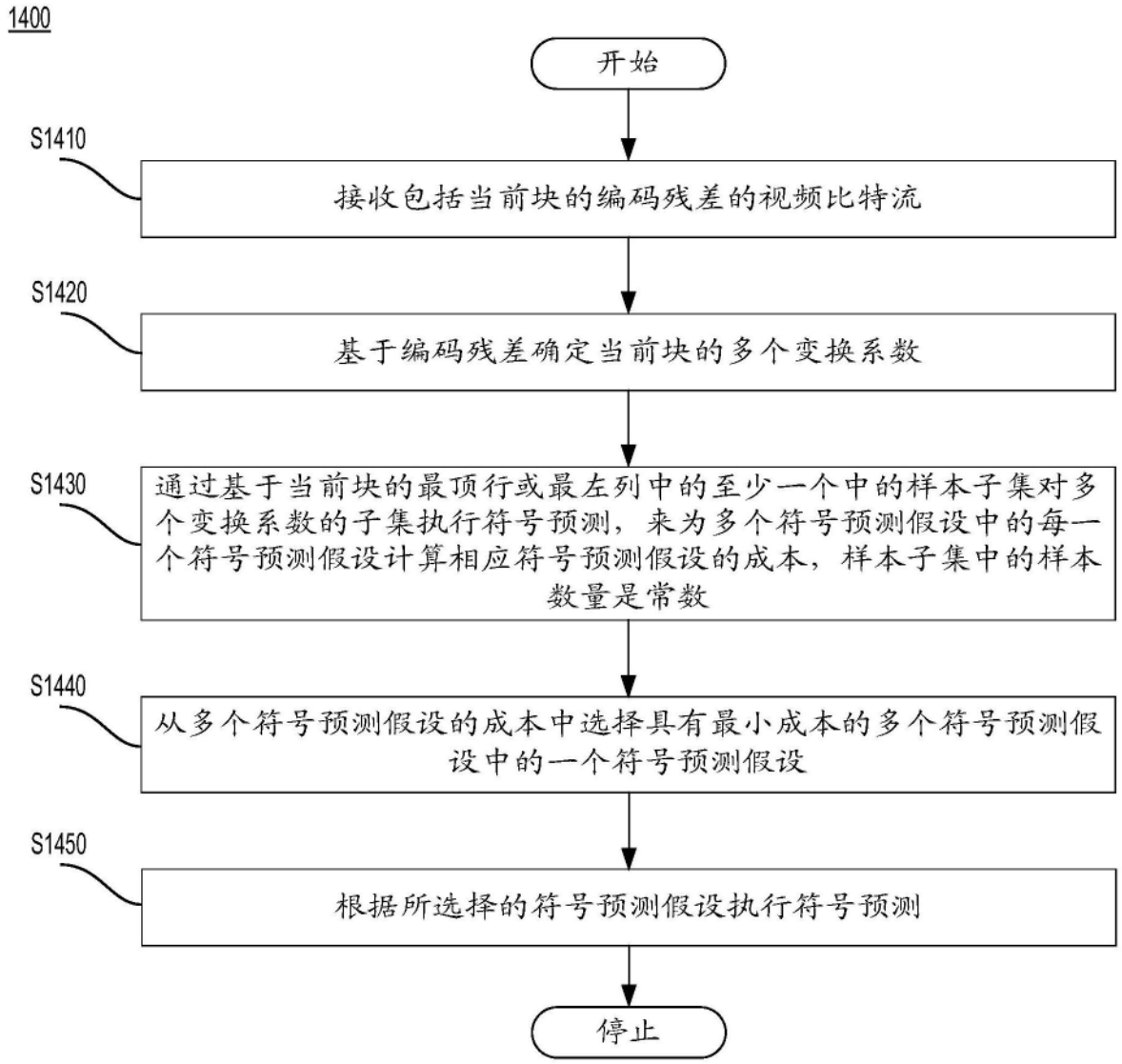

1400
开始
S1410
接收包括当前块的编码残差的视频比特流
S1420
基于编码残差确定当前块的多个变换系数
S1430
通过基于当前块的最顶行或最左列中的至少一个中的样本子集对多个变换系数的子集执行符号预测，来为多个符号预测假设中的每一个符号预测假设计算相应符号预测假设的成本，样本子集中的样本数量是常数
S1440
从多个符号预测假设的成本中选择具有最小成本的多个符号预测假设中的一个符号预测假设
S1450
根据所选择的符号预测假设执行符号预测
停止

图14

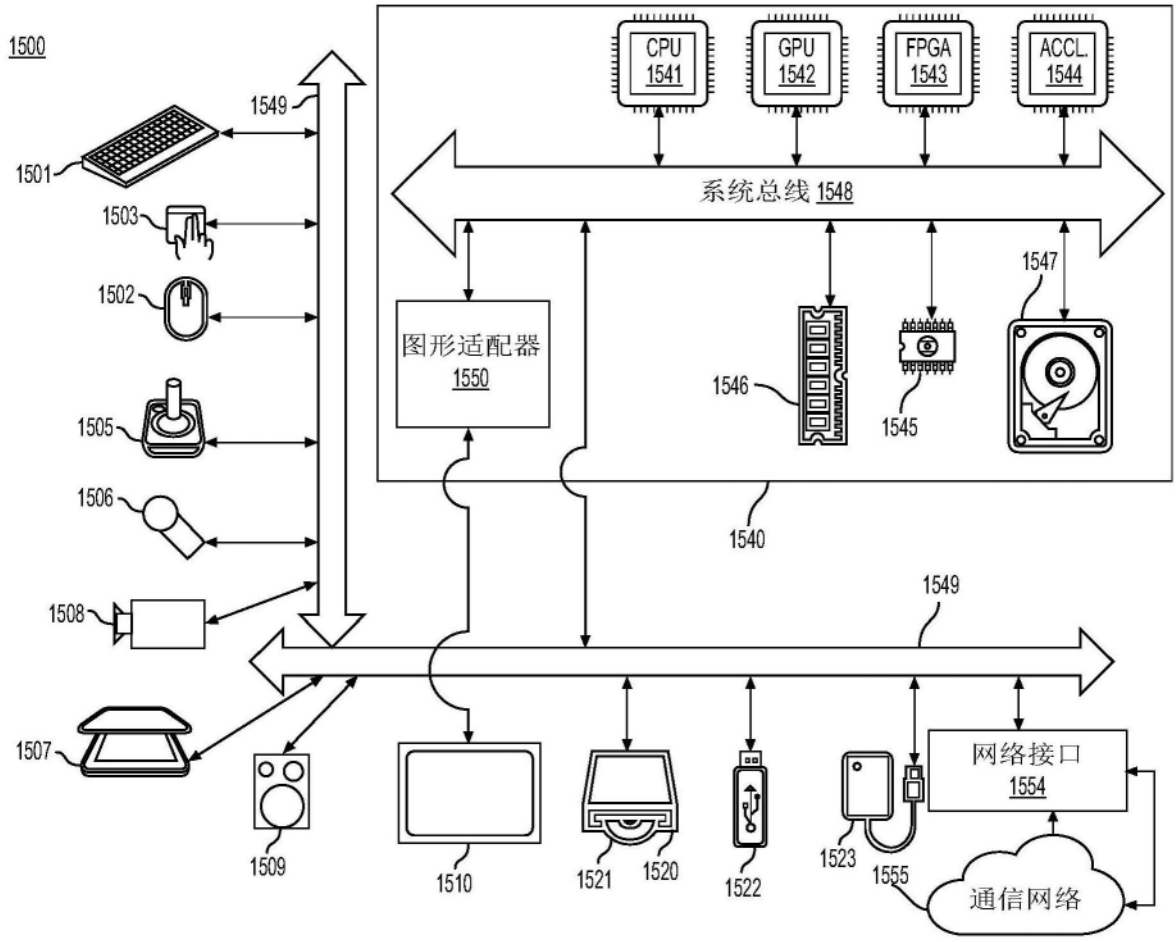

1500
1549
1501
1503
1502
1505
1506
1508
1507
1509
CPU 1541
GPU 1542
FPGA 1543
ACCL. 1544
系统总线 1548
图形适配器 1550
1546
1545
1547
1540
1549
1510
1521
1520
1522
1523
1555
网络接口 1554
通信网络

图15