



(12) 发明专利

(10) 授权公告号 CN 112292859 B

(45) 授权公告日 2023.09.22

(21) 申请号 201980020401.4

(74) 专利代理机构 深圳市深佳知识产权代理事

(22) 申请日 2019.08.28

务所(普通合伙) 44285

(65) 同一申请的已公布的文献号

专利代理人 王兆林

申请公布号 CN 112292859 A

(51) Int.CI.

H04N 19/169 (2006.01)

(43) 申请公布日 2021.01.29

H04N 19/196 (2006.01)

(30) 优先权数据

H04N 19/463 (2006.01)

62/730,885 2018.09.13 US

H04N 19/127 (2006.01)

16/232,677 2018.12.26 US

H04N 19/70 (2006.01)

(85) PCT国际申请进入国家阶段日

(56) 对比文件

2020.09.18

WO 2018033661 A1, 2018.02.22

(86) PCT国际申请的申请数据

CN 104604236 A, 2015.05.06

PCT/US2019/048538 2019.08.28

Stephan Wenger等.On VVC HLS
architecture and bitstream structure(
JVET-L0110).《Joint Video Experts Team
(JVET) of ITU-T SG 16 WP 3 and ISO/IEC
JTC 1/SC 29/WG 11 12th Meeting》.2018,

(87) PCT国际申请的公布数据

审查员 张海燕

W02020/055589 EN 2020.03.19

(73) 专利权人 腾讯美国有限责任公司

权利要求书2页 说明书15页 附图6页

地址 美国加利福尼亚州帕洛阿尔托公园大
道2747号

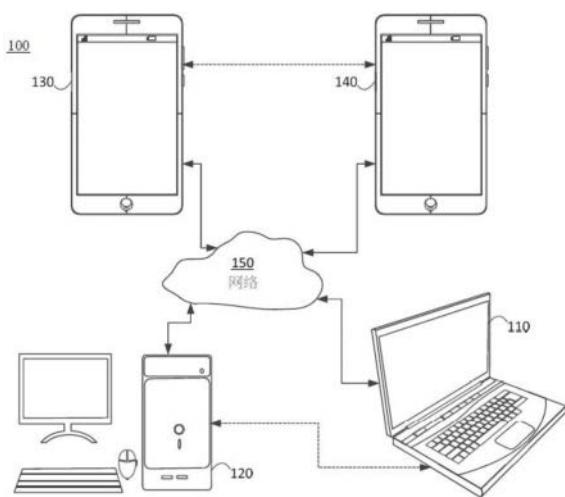
(72) 发明人 文格尔·史蒂芬 刘杉

(54) 发明名称

一种用于解码至少一个视频流的方法和装
置

(57) 摘要

一种用于解码至少一个视频流的方法和装
置,至少一个视频流中的每个视频流与各自的解
码器参数集相关联。该方法包括:解码器激活至
少一个视频流中的第一视频流的解码器参数集;
解码器处理在第一视频流之外的对解码器可用的
流结尾NAL单元;以及解码器响应于处理流结
尾NAL单元而停用第一视频流的解码器参数集。



1. 一种用于解码至少一个视频流的方法,其特征在于,所述至少一个视频流中的每个视频流与各自的解码器参数集相关联,所述方法包括:

解码器激活所述至少一个视频流中的第一视频流的解码器参数集,所述第一视频流的解码器参数集保持激活,直至所述解码器处理在所述第一视频流之外的对所述解码器可用的流结尾网络抽象NAL单元;

所述解码器处理在所述第一视频流之外的对所述解码器可用的流结尾网络抽象层NAL单元;以及

所述解码器响应于处理所述流结尾NAL单元,停用所述第一视频流的解码器参数集;

所述解码器激活所述至少一个视频流中的第二视频流的解码器参数集;所述至少一个视频流中的第二视频流的解码器参数集,至少有一个值不同于先前激活的解码器参数集。

2. 根据权利要求1所述的方法,其特征在于,

所述第二视频流的解码器参数集具有与所述第一视频流的解码器参数集不同的内容。

3. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

在所述解码器处理在所述第一视频流之外的对所述解码器可用的流结尾NAL单元之前,所述解码器接收所述第二视频流的解码器参数集。

4. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

在所述解码器处理在所述第一视频流之外的对所述解码器可用的流结尾NAL单元之后,所述解码器接收所述第二视频流的解码器参数集。

5. 根据权利要求1所述的方法,其特征在于,所述停用所述第一视频流的所述解码器参数集包括:

清空所述解码器的缓冲器,以及

停止输出经所述解码器解码的所述第一视频流的图片。

6. 根据权利要求1所述的方法,其特征在于,所述解码器激活所述第一视频流的解码器参数集包括:

所述解码器根据所述第一视频流的解码器参数集的参数,解码所述第一视频流。

7. 根据权利要求6所述的方法,其特征在于,所述解码器使用与所述第一视频流相关联的至少一个图片参数集PPS和至少一个图片头来解码所述第一视频流,其中,所述至少一个图片头和所述至少一个PPS比所述第一视频流的所述解码器参数集处于更低的语法层。

8. 根据权利要求7所述的方法,其特征在于,所述至少一个图片头和所述至少一个PPS处于同一语法层。

9. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

由所述解码器在所述第一视频流外接收所述流结尾NAL单元,使所述流结尾NAL单元在所述第一视频流之外对所述解码器可用。

10. 根据权利要求1所述的方法,其特征在于,所述第一视频流的解码器参数集对于第一视频流的多个第一编码视频序列保持激活。

11. 根据权利要求1所述的方法,其特征在于,所述第二视频流的解码器参数集对于第二视频流的多个第二编码视频序列保持激活。

12. 根据权利要求1所述的方法,其特征在于,所述至少一个视频流中的每个视频流包括多个第一编码视频序列,所述多个第一编码视频序列包括具有不同呈现时间的编码视频

序列,所述至少一个视频流中的每个视频流和与属于所述视频流的所有视频序列相关的解码器参数集相关联。

13.一种用于解码至少一个视频流的装置,其特征在于,所述至少一个视频流中的每个视频流与各自的解码器参数集相关联,所述装置包括解码器,所述解码器用于执行如权利要求1-12任意一项所述的方法。

14.一种存储指令的非暂态计算机可读介质,所述指令包括一个或多个指令,
在被设备的一个或多个处理器执行时,所述一个或多个指令用于执行权利要求1-12任意一项所述的方法。

15.一种用于解码至少一个视频流的装置,其特征在于,所述至少一个视频流中的每个视频流与各自的解码器参数集相关联,所述装置包括激活单元、处理单元和停用单元:

所述激活单元,用于激活所述至少一个视频流中的第一视频流的解码器参数集,所述第一视频流的解码器参数集保持激活,直至所述解码器处理在所述第一视频流之外的对所述解码器可用的流结尾网络抽象NAL单元;所述处理单元,用于处理在所述第一视频流之外的对所述解码器可用的流结尾网络抽象层NAL单元;以及

所述停用单元,用于响应于处理所述流结尾NAL单元,停用所述第一视频流的解码器参数集;

所述激活单元,还用于所述解码器激活所述至少一个视频流中的第二视频流的解码器参数集;所述至少一个视频流中的第二视频流的解码器参数集,至少有一个值不同于先前激活的解码器参数集。

16.一种计算机设备,其特征在于,所述设备包括处理器以及存储器;
所述存储器用于存储程序代码,并将所述程序代码传输给所述处理器;
所述处理器用于根据所述程序代码中的指令执行权利要求1-12中任意一项所述的用于解码至少一个视频流的方法。

一种用于解码至少一个视频流的方法和装置

[0001] 相关申请的交叉引用

[0002] 本申请要求于2018年9月13日提交、申请号为62/730,885的美国临时专利申请、以及于2018年12月26日提交、申请号为16/232,677的美国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

[0003] 本公开涉及视频处理技术，尤其涉及用于解码至少一个视频流的方法和装置。

背景技术

[0004] 当前已使用具有运动补偿的帧间预测来进行视频编码和解码。未压缩的数字视频可以包括一系列图片，每个图片在空间维度为例如 1920×1080 个亮度样本以及相关联的色度样本。该一系列图片可以具有固定的或可变的图片速率（也被非正式地称为帧速率），例如每秒60幅图片或60Hz。未压缩的视频对码率的要求较高。举例来说，每样本8比特的1080r60 4:2:0视频（帧速率60Hz，亮度样本分辨率 1920×1080 ）需要接近1.5Gbit/s的带宽。一小时的上述视频需要超过600GB的存储空间。

[0005] 视频编码和解码的一个目的可在于通过压缩来减少输入视频信号的冗余度。压缩可帮助降低上述带宽需求或存储空间需求，在一些情况下可降低两个或更多数量级。可以采用无损压缩、有损压缩及二者的组合。无损压缩是指可以从压缩的原始信号重建出原始信号的精确副本的技术。当使用有损压缩时，重建的信号可能与原始信号不同，但原始信号与重建的信号之间的失真足够小，使得重建的信号可用于目标应用。有损压缩在视频方面得到了广泛应用。容忍的失真量取决于应用，比如，某些消费者流媒体应用的用户可能比电视分配应用的用户容忍更高的失真。能够实现的压缩比可反映出：更高的可允许/可容忍的失真能产生更高的压缩比。

[0006] 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、以及熵编码，下面将介绍其中一些技术。

[0007] H.264之前的某些视频编解码器（例如MPEG-2Visual）使用了暂态头的层级结构，该暂态头包括序列头、图片组（GOP）头、图片头和切片头。包括在每个头中的语法元素涉及所有下层语法结构。例如，序列头的语法元素涉及包含在序列中的所有GOP、包含在这些GOP中的所有图片、以及包含在这些图片中的所有切片。GOP头的语法元素涉及包含在GOP中的所有图片、以及这些图片中的所有切片。这种层级结构可带来高效的编码，但具有次优的容错特性。例如，如果序列头的重要信息在传输中丢失，则序列的GOP、图片或切片均无法被解码。

[0008] 从2003年起，某些ITU和MPEG视频编解码器，即H.264和H.265，不在切片头之上使用暂态头，而是依赖于参数集。在每个语法层上，例如在序列层或图片层上，可通过解码器或通过外部装置从比特流中接收一个或多个参数集。使用这些（可能有许多）相同类型的参数集中的哪些参数集来解码给定的序列或图片，取决于在例如切片头（对于图片参数集

PPS) 或 PPS(对于序列参数集SPS) 中所编码的参照。该结构可以具有以下优点：即使比特流本身通过有损信道发送，也可以可靠地发送相关的参数集；或者，可以通过发送冗余副本来自增加参数集被接收的可能性——在参数集首次使用之前就发送可能会更好。该结构可能具有如下缺点：就相同数目和类型的语法元素所需要的比特数而言，发送参数集可能比发送MPEG-2类型的头占用更多资源。此外，此架构下每个切片头中可能以多个冗余副本的形式包含某些语法元素，这些语法元素虽在图片之间频繁变化，但在给定图片内保持恒定。虽然该做法能够使切片可独立解码(至少从解析依赖性和熵解码的角度来看)，但是也会占用更多的比特。

[0009] 在H.264的设计中，为了容错，切片的独立可解码性被当作主要的设计目标。然而，从2003年起，随着隐蔽丢失切片的效果越来越弱，对传送编码视频的网络架构的改进以及预测机制的进步显著地降低了切片独立可解码性的吸引力。

发明内容

[0010] 由于需求偏离了切片的独立可解码性，因此需要一种新的高级语法架构——在假定至少一些给定图片的损失可在解码器中被合理隐蔽的情况下，该架构保持良好的容错特性，并且在编码效率方面利用MPEG-2类型的头结构的优点。本发明的一些实施例提供该种高级语法架构，其维持了良好的容错特性及编码效率。

[0011] 根据本申请公开的一方面，提供了一种可用于解码至少一个视频流的方法，该至少一个视频流中的每个视频流与各自的解码器参数集相关联。该方法可以包括：解码器激活至少一个视频流中的第一视频流的解码器参数集。该方法还可以包括：解码器处理在第一视频流之外的对解码器可用的流结尾NAL单元。该方法还可以包括：解码器响应于处理流结尾网络抽象层(Network Abstraction Layer, NAL) 单元，停用第一视频流的解码器参数集。

[0012] 根据本申请公开的另一方面，提供了一种可用于解码至少一个视频流的装置，该至少一个视频流中的每个视频流与各自的解码器参数集相关联。该装置可以包括解码器，该解码器用于激活至少一个视频流中的第一视频流的解码器参数集。该解码器还可用于处理在第一视频流之外的对解码器可用的流结尾NAL单元。该解码器还可用于响应于处理流结尾NAL单元，停用第一视频流的解码器参数集。

[0013] 根据本申请公开的又一方面，可使用一种存储指令的非暂态计算机可读介质。该指令可以包括一个或多个指令，在被设备的一个或多个处理器执行时，该一个或多个指令使一个或多个处理器解码至少一个视频流，该至少一个视频流中的每个视频流与各自的解码器参数集相关联，其中，所述一个或多个处理器通过以下操作解码至少一个视频流：激活至少一个视频流中的第一视频流的解码器参数集；处理在第一视频流之外的对解码器可用的流结尾NAL单元；以及响应于处理流结尾NAL单元，停用第一视频流的解码器参数集。

[0014] 本公开内容的另一方面还提供了一种用于解码至少一个视频流的装置，所述至少一个视频流中的每个视频流与各自的解码器参数集相关联，所述装置包括激活单元、处理单元和停用单元：

[0015] 所述激活单元，用于激活所述至少一个视频流中的第一视频流的解码器参数集；

[0016] 所述处理单元，用于处理在所述第一视频流之外的对所述解码器可用的流结尾网

络抽象层NAL单元；以及

[0017] 所述停用单元，用于响应于处理所述流结尾NAL单元，停用所述第一视频流的解码器参数集。

[0018] 本公开内容的另一方面还提供了计算机设备，所述设备包括处理器以及存储器：

[0019] 所述存储器用于存储程序代码，并将所述程序代码传输给所述处理器；

[0020] 所述处理器用于根据所述程序代码中的指令执行上述用于解码至少一个视频流的方法。

[0021] 由此可见，在针对至少一个视频流的解码时，该至少一个视频流中的每个视频流与各自的解码器参数集相关联，解码器激活至少一个视频流中的第一视频流的解码器参数集；解码器处理在所述第一视频流之外的对所述解码器可用的流结尾NAL单元；解码器响应于处理所述流结尾NAL单元，停用所述第一视频流的解码器参数集。激活解码器参数集的停用意味着在没有语法冲突的情况下可以激活一个不同的解码器参数集，该不同的解码器参数集至少有一个值不同于先前激活的解码器参数集，这种新的高级语法架构保持良好的容错特性，并且在编码效率方面利用MPEG-2类型的头结构的优点。本发明的一些实施例提供该种高级语法架构，其维持了良好的容错特性及编码效率。

附图说明

[0022] 所公开的主题的其他特征、特性和各种优点将通过以下详细描述和附图而更加明显。在附图中：

[0023] 图1为一实施例提供的通信系统的简化框图的示意图；

[0024] 图2为一实施例提供的流媒体系统的简化框图的示意图；

[0025] 图3为一实施例提供的视频解码器和显示器的简化框图的示意图；

[0026] 图4为一实施例提供的视频编码器和视频源的简化框图的示意图；

[0027] 图5为一实施例提供的高级语法架构中的语法层级的示意图；

[0028] 图6为一实施例提供的图片头和图片参数集的示意图；

[0029] 图7为一实施例提供的带外接收流结尾时改变解码器参数集的流程示意图；

[0030] 图8为一实施例提供的计算机系统的示意图。

具体实施例

[0031] 图1示出了本申请一实施例中通信系统(100)的简化框图。系统(100)可以包括至少两个经由网络(150)互连的终端(110、120)。对于数据的单向传输，第一终端(110)可在本地位置编码视频数据，以经由网络(150)传输至另一终端(120)。第二终端(120)可从网络(150)接收另一终端的已编码视频数据，对已编码数据进行解码，并显示恢复的视频数据。单向数据传输可常见于诸如媒体服务应用中。

[0032] 图1示出了被提供用于支持已编码视频双向传输的第二对终端(130、140)，该双向传输可能发生在例如在视频会议期间。对于数据的双向传输，每个终端(130、140)可编码在本地位置捕获的视频数据，以经由网络(150)传输至另一终端。每个终端(130、140)还可接收由另一终端传输的已编码视频数据，可对已编码数据进行解码，并且可在本地显示设备处显示恢复的视频数据。

[0033] 在图1中,终端(110至140)可以是例如服务器、个人计算机、智能电话、和/或任何其他类型的终端。例如,终端(110-140)可以是膝上型计算机、平板电脑、媒体播放器和/或专用视频会议设备。网络(150)表示在终端(110至140)之间传送已编码视频数据的任意数目的网络,包括例如有线(连线的)和/或无线通信网络。通信网络(150)可在电路交换和/或分组交换信道中交换数据。该网络可包括电信网络、局域网、广域网和/或互联网。出于本申请论述的目的,除非在下文中有所解释,否则网络(150)的架构和拓扑对于本申请公开的操作来说可能是无关紧要的。

[0034] 作为本申请公开主题的一个实施例,图2示出了视频编码器和视频解码器在流式传输环境中的放置方式。本申请所公开主题可同等适用于其它支持视频的应用,包括例如视频会议、数字TV、在包括CD、DVD、存储棒等的数字介质上存储压缩视频等等。

[0035] 如图2所示,流系统(200)可包括捕获子系统(213),该捕获子系统包括视频源(201)和编码器(203)。该流系统(200)还可以包括至少一个流服务器(205)和/或至少一个流客户端(206)。

[0036] 视频源(201)可以创建例如未压缩的视频样本流(202)。视频源(201)可以为例如数字摄像机。采样流(202)可以由耦接至摄像机(201)的编码器(203)处理,与已编码的视频比特流相比,其被描绘为粗线以强调其高数据量。编码器(203)可以包括硬件、软件或其组合,以实现或实施以下更多细节中所描述的本申请公开的主题的各方面。编码器(203)还可生成经过编码的视频比特流(204)。视频比特流(204)可存储在流服务器(205)上以供将来使用,与未压缩视频样本流(202)相比,其被描绘为细线以强调较低数据量。一个或多个流客户端(206)可访问流服务器(205)以检索视频比特流(209),该视频比特流可以是已编码的视频比特流(204)的副本。

[0037] 流客户端(206)可以包括视频解码器(210)和显示器(212)。视频解码器(210)可解码例如视频比特流(209),该视频比特流为已编码的视频比特流(204)的输入拷贝,并且视频解码器(210)可创建输出视频样本流(211),该输出视频样本流可在显示器(212)或另一呈现设备(未描绘)上呈现。在一些流传输系统中,可以根据某些视频编码/压缩标准来编码视频比特流(204、209)。这种标准的示例包括但不限于ITU-T H.265建议书。被非正式地称为通用视频编码(Versatile Video Coding, VVC)的视频编码标准正处于开发之中。本申请实施例可以在VVC环境中使用。

[0038] 图3示出了本申请一实施例的附接至显示器(212)的视频解码器(210)的示例性功能框图。

[0039] 视频解码器(210)可包括信道(312)、接收器(310)、缓冲存储器(315)、熵解码器/解析器(320)、缩放器/逆变换单元(351)、帧内预测单元(352)、运动补偿预测单元(353)、聚合器(355)、环路滤波器单元(356)、参考图片存储器(357)、以及当前图片存储器(358)。在至少一个实施例中,视频解码器(210)可以包括集成电路、一系列集成电路、和/或其他电子电路。视频解码器(210)还可以部分或全部地实现为运行于具有关联存储器的一个或多个CPU上的软件。

[0040] 在本实施例和其他实施例中,接收器(310)可以逐个接收将由解码器(210)解码的一个或多个已编码视频序列,其中每个已编码视频序列的解码独立于其他已编码视频序列。可从信道(312)接收已编码视频序列,该信道可以是连接到存储已编码视频数据的存储

设备的硬件/软件链路。接收器(310)可在接受其他数据时一并接受已编码视频数据,该其他数据例如已编码音频数据和/或辅助数据流,并且该其他数据可被转发到各自的使用实体(未描绘)。接收器(310)可将已编码视频序列与其他数据分开。为了对抗网络抖动,可在接收器(310)与熵解码器/解析器(320)(下文称为“解析器”)之间耦接缓冲存储器(315)。当接收器(310)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,可不使用缓冲存储器(315)或者可使用小容量缓冲存储器(315)。为了应用于因特网等“尽力而为”包网络,可能需要缓冲存储器(315),该缓冲存储器可以相对较大,且可以具有适应性的大小。

[0041] 视频解码器(210)可以包括解析器(320),以从熵编码的视频序列中重建符号(321)。这些符号的类别包括例如用于管理解码器(210)运行的信息,也可能包括用于控制例如显示器(212)的呈现设备的信息,该呈现设备可耦合到如图2所示的解码器。用于(一个或多个)呈现设备的控制信息可以是例如辅助增强信息(SEI消息)或视频可用性信息(VUI)参数集片段(未描绘)的形式。解析器(320)可以对接收到的已编码视频序列进行解析/熵解码。已编码视频序列的编码可依据视频编码技术或标准,并且可遵循本领域技术人员公知的原理,包括:可变长度编码、霍夫曼编码(Huffman coding)、具有或不具有上下文敏感性的算术编码等。解析器(320)可基于与组相对应的至少一个参数,从已编码视频序列中提取用于视频解码器中至少一个像素子组的子组参数集。子组可以包括:图片组(GOP)、图片、图块、切片、宏块、编码单元(CU)、块、变换单元(TU)、预测单元(PU)等。解析器(320)还可以从已编码视频序列中提取例如变换系数、量化器参数值、运动矢量等信息。

[0042] 解析器(320)可以对从缓冲存储器(315)接收的视频序列执行熵解码/解析操作,从而创建符号(321)。

[0043] 取决于已编码视频图片或其部分(例如:帧间图片和帧内图片、帧间块和帧内块)的类型以及其他因素,符号(321)的重建可涉及多个不同单元。涉及哪些单元以及涉及这些单元的方式可由子组控制信息控制,该子组控制信息由解析器(320)从已编码视频序列中解析得到。简洁起见,这里未描绘解析器(320)与下述多个单元之间的子组控制信息流。

[0044] 除了已经提及的功能块以外,视频解码器(210)可以在概念上细分成下述的多个功能单元。在受商业约束运行的实际实施方式中,这些单元中的许多单元紧密交互,并且至少可以部分地相互集成。然而,为了描述所公开的主题,在概念上细分成下述功能单元是合适的。

[0045] 单元可以是缩放器/逆变换单元(351)。缩放器/逆变换单元(351)可从解析器(320)接收作为(一个或多个)符号(321)的量化变换系数以及控制信息,包括使用哪种变换、块大小、量化因子、量化缩放矩阵等。缩放器/逆变换单元(351)可输出包括样本值的块,该样本值可输入聚合器(355)中。

[0046] 在一些情况下,缩放器/逆变换(351)的输出样本可涉及帧内编码块,即该块不使用来自先前重建的图片的预测性信息、但可使用来自当前图片的先前重建部分的预测性信息。这些预测性信息可由帧内图片预测单元(352)提供。在一些情况下,帧内预测单元(352)使用从当前图片存储器(358)中当前(部分重建的)图片提取的周边已重建信息,来生成与正在重建的块在大小和形状上相同的块。在一些情况下,在每个样本的基础上,聚合器(355)将帧内预测单元(352)已经生成的预测信息添加至缩放器/逆变换单元(351)所提供

的输出样本信息。

[0047] 在其他情况下，缩放器/逆变换单元(351)的输出样本可涉及帧间编码块，且该块可能经过运动补偿。在该情况下，运动补偿预测单元(353)可访问参考图片存储器(357)以获取用于预测的样本。在根据涉及块的符号(321)对提取的样本进行运动补偿之后，聚合器(355)可将这些样本添加至缩放器/逆变换单元(351)的输出(在该情况下被称为残差样本或残差信号)，从而生成输出样本信息。运动补偿预测单元(353)从参考图片存储器(357)中的地址获取预测采样，该地址可由运动矢量控制。运动矢量可以符号(321)的形式对运动补偿预测单元(353)可用，该符号可具有例如x、Y、以及参考图片分量。运动补偿还可以包括：在使用子样本精确运动矢量时对从参考图片存储器(357)中提取的样本值进行内插、以及运动矢量预测机制等。

[0048] 在环路滤波器(356)中可对聚合器(355)的输出样本应用各种环路滤波技术。视频压缩技术可以包括环路内滤波器技术，其被包含在已编码视频比特流中的参数所控制，且该参数以来自解析器(320)的符号(321)的形式可用于环路滤波器(356)。然而，视频压缩技术还可以响应于在对已编码图片或已编码视频序列的前序(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的并经环路滤波的样本值。

[0049] 环路滤波器(356)的输出可以是样本流，该样本流可输出到例如显示器(212)的呈现设备，并存储在参考图片存储器(356)中以供将来帧间预测使用。

[0050] 某些已编码图片一旦被完全重建就可作为参考图片，以用于将来的预测。一旦已编码图片被完全重建并(例如通过解析器(320))被识别为参考图片，存储在当前图片存储器(358)中的当前参考图片可成为参考图片存储器(357)的一部分，并且在开始重建后续已编码图片之前可重新分配刷新的当前图片存储器。

[0051] 视频解码器(210)可根据预定视频压缩技术执行解码操作，该视频压缩技术可载于标准，例如ITU-T H.265标准建议书中。在已编码视频序列遵循视频压缩技术或标准的语法时，其可以符合由正在使用的视频压缩技术或标准所指定的语法，如在视频压缩技术文档或标准中所指定的语法，尤其是其中总则文档所指定的语法。此外，为了符合一些视频压缩技术或标准，已编码视频序列的复杂度可以处于由视频压缩技术或标准的层级所限定的界限内。在一些情况下，层级限制最大图片大小、最大帧率、最大重建样本速率(例如，以每秒兆样本为单位进行度量)、最大参考图片大小等。在一些情况下，层级所设置的界限可进一步通过假定参考解码器(HRD)规范以及在已编码视频序列中用信号表示的用于HRD缓冲器管理的元数据来限定。

[0052] 在一实施例中，接收器(310)可以接收附加(冗余)数据和已编码的视频。该附加数据可以被包含为已编码视频序列的一部分。视频解码器(210)可使用该附加数据以适当地解码数据和/或更准确地重建原始视频数据。附加数据可以为例如时间增强层、空间增强层、或者SNR增强层，以及冗余切片、冗余图片、前向纠错码等形式。

[0053] 图4示出了本申请一实施例提供的与视频源(201)相关联的视频编码器(203)的示例性功能框图。

[0054] 视频编码器(203)可包括例如作为源编码器(430)的编码器、编码引擎(432)、(本地)解码器(433)、参考图片存储器(434)、预测器(435)、传输器(440)、熵编码器(445)、控制器(450)和信道(460)。

[0055] 编码器(203)可以从视频源(201)(非编码器的一部分)接收视频样本,该视频源可捕获编码器(203)待编码的(一个或多个)视频图像。

[0056] 视频源(201)可以数字视频样本流的形式提供编码器(203)待编码的源视频序列,该数字视频样本流可具有任何合适的比特深度(例如:x比特、10比特、12比特、...)、任何颜色空间(例如,BT.601Y CrCb、RGB、...)、以及任何合适的采样结构(例如,Y CrCb 4:2:0、Y CrCb 4:4:4)。在媒体服务系统中,视频源(201)可以是存有提前准备好的视频的存储设备。在视频会议系统中,视频源(201)可以是捕获本地图像信息作为视频序列的摄像机。视频数据可以被提供为在顺序观看时形成运动的多个单独图片。图片自身可以被组织为像素的空间阵列,其中,取决于所使用的采样结构、颜色空间等,每个像素可包括一个或多个样本。本领域技术人员能够容易地理解像素与样本之间的关系。以下描述着眼于样本。

[0057] 在一实施例中,编码器(203)可以实时地或在由应用所要求的任何其他时间约束下,编码并压缩源视频序列的图片以得到已编码视频序列。控制器(450)可具有强制采用适当编码速度的功能。控制器(450)还可以控制下述的其他功能单元,并且可以在功能上耦接至这些单元。简洁起见,未示出该耦接。控制器(450)设置的参数可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 值等)、图片大小、图片组(GOP)布局、最大运动矢量搜索范围等。由于控制器(450)的其他功能可能涉及为特定系统设计而优化的视频编码器(203),本领域技术人员可以容易地确定这些功能。

[0058] 一些视频编码器以本领域技术人员容易理解为“编码循环”的方式运行。简而言之,编码循环可包括源编码器(430)的编码部分(负责基于待编码的输入图片和(一个或多个)参考图片创建符号)和嵌入编码器(203)中的(本地)解码器(433)。当某些视频压缩技术中符号与已编码视频比特流之间为无损压缩时,该解码器重构符号以创建(远程)解码器也会创建的样本数据。该重建的样本流可以输入参考图片存储器(434)。由于对符号流的解码产生与解码器位置(本地或远程)无关的位精确结果,因此参考图片存储器(434)中的内容在本地编码器与远程编码器之间也是按比特位精确对应的。换句话说,编码器的预测部分“看到”的参考图片样本与解码器将在解码期间使用预测时所“看到”的样本值完全相同。本领域技术人员知晓这种参考图片同步性基本原理(以及在例如因信道误差而无法维持同步性的情况下产生的漂移)。

[0059] “本地”解码器(433)的操作可与例如已在上文结合图4详细描述视频解码器(210)的“远程”解码器相同。然而,另外简要参考图3,当符号可用且熵编码器(445)和解析器(320)能够无损地将符号编码/解码为已编码视频序列时,包括信道(312)、接收器(310)、缓冲存储器(315)和解析器(320)在内的视频解码器(210)的熵解码部分,可能无法完全在本地解码器(433)中实施。

[0060] 此时可以观察到,除存在于解码器中的解析/熵解码之外的任何解码器技术,也必定以基本上相同的功能形式存在于对应的编码器中。出于此原因,本申请侧重于解码器操作。可简化编码器技术的描述,因为编码器技术与全面地描述的解码器技术互逆。仅在某些区域中需要更详细的描述,并且在下文提供。

[0061] 作为操作的一部分,源编码器(430)可执行运动补偿预测编码。参考来自视频序列中被指定为“参考图片”的一个或多个先前已编码图片,所述运动补偿预测编码对输入图片进行预测性编码。以此方式,编码引擎(432)对输入图片的像素块与参考图片的像素块之间

的差异进行编码,所述参考图片可被选作所述输入帧的预测参考。

[0062] 本地视频解码器(433)可基于源编码器(430)创建的符号,对可指定为参考帧的帧的已编码视频数据进行解码。编码引擎(432)的操作可为有损过程。当已编码视频数据可在视频解码器(图4中未示)处被解码时,重建的视频序列通常可以是带有一些误差的源视频序列的副本。本地视频解码器(433)复制解码过程,所述解码过程可由视频解码器对参考图片执行,且可使重建的参考帧存储在参考图片存储器(434)中。以此方式,视频编码器(203)可在本地存储重建的参考帧的副本,所述副本与将由远端视频解码器获得的重建参考帧具有共同内容(不存在传输误差)。

[0063] 预测器(435)可针对编码引擎(432)执行预测搜索。即,对于将要编码的新帧,预测器(435)可在参考图片存储器(434)中搜索可作为所述新帧的适当预测参考的样本数据(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。预测器(435)可基于样本块逐像素块操作,以找到合适的预测参考。在一些情况下,根据预测器(435)获得的搜索结果,可确定输入图片可具有从参考图片存储器(434)中存储的多个参考图片取得的预测参考。

[0064] 控制器(450)可管理源编码器(430)的编码操作,包括例如设置用于对视频数据进行编码的参数和子群参数。

[0065] 可在熵编码器(445)中对所有上述功能单元的输出进行熵编码。熵编码器(445)根据例如霍夫曼编码、可变长度编码、算术编码等技术对各种功能单元生成的符号进行无损压缩,从而将所述符号转换成已编码视频序列。

[0066] 传输器(440)可缓冲由熵编码器(445)创建的已编码视频序列,从而为通过信道(460)进行传输做准备,所述通信信道可以是通向将存储已编码的视频数据的存储装置的硬件/软件链路。传输器(440)可将来自源编码器(430)的已编码视频数据与要传输的其它数据合并,所述其它数据例如是已编码音频数据和/或辅助数据流(未示出来源)。

[0067] 控制器(450)可管理编码器(203)的操作。在编码期间,控制器(450)可以为每个已编码图片分配某一已编码图片类型,但这可能影响可应用于相应的图片的编码技术。例如,通常可将图片分配为帧内图片(I图片)、预测图片(P图片)、或者双向预测图片(B图片)。

[0068] 帧内图片(I图片)可以是不将序列中的任何其它图片用作预测源就可被编码和解码的图片。一些视频编解码器容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。所属领域的技术人员了解I图片的变体及其相应地应用和特征。

[0069] 预测性图片(P图片)可以是可使用帧内预测或帧间预测进行编码和解码的图片,所述帧内预测或帧间预测使用至多一个运动矢量和参考索引来预测每个块的样本值。

[0070] 双向预测性图片(B图片)可以是可以使用帧内预测或帧间预测进行编码和解码的图片,该帧内预测或帧间预测使用至多两个运动矢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单个块的重建。

[0071] 源图片通常可在空间上细分成多个样本块(例如,4×4、8×8、4×8或16×16个样本的块),且逐块进行编码。这些块可参考其它(已编码)块进行预测编码,根据应用于块的相应图片的编码分配来确定所述其它块。举例来说,I图片的块可进行非预测编码,或所述

块可参考同一图片的已经编码的块来进行预测编码(空间预测或帧内预测)。P图片的像素块可参考一个先前编码的参考图片通过空间预测或通过时域预测进行预测编码。B图片的块可参考一个或两个先前编码的参考图片通过空间预测或通过时域预测进行预测编码。

[0072] 视频编码器(203)可根据例如ITU-T H.265建议书的预定视频编码技术或标准执行编码操作。在操作中,视频编码器(203)可执行各种压缩操作,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,已编码视频数据可符合所用视频编码技术或标准指定的语法。

[0073] 在一实施例中,传输器(440)可在传输已编码的视频时传输附加数据传输附加数据和已编码的视频。源编码器(430)可将此类数据作为可以是已编码视频序列的一部分。附加数据可包括时间/空间/SNR增强层、冗余图片和切片等其它形式的冗余数据、SEI消息、VUI参数集片段等。

[0074] 以下描述本申请公开的一些实施例的方面,包括将在视频编解码器技术或标准中,例如在通用视频编码(VVC)中,实现的高级语法架构。

[0075] 由于H.264的NAL单元的概念已被证明有用,且至少一些系统规范(包括某些文件格式)依赖于该概念,因此高级语法架构可以包括这一概念。

[0076] 可选地,高级语法架构可以不包括(独立的、规则的)切片的概念。自2003年以来(H.264 1版的公开日期),由于帧内预测机制的数量和效率的不断增加,视频编码的进步使得在许多情况下基于切片的错误隐蔽实际上变得不可能。同时从编码效率的观点来看,这些预测机制使得在某些情况下使用切片变得非常占用资源。因此,近来极少有实现方式使用切片来实现其的预期目的(MTU大小匹配)。然而,基本上所有需要低延迟容错率的应用都依赖于基于图片的容错工具,例如帧内刷新、开放GOP、具有不均匀基层保护的可缩放性等。

[0077] 在移除切片的情况下,在熵等级上可独立解码(即,没有解析相关性)的高级语法架构的最小VCL语法单元可以例如图块或已编码图片。

[0078] 对某些应用场景来说,独立解码图块可能是有用的。例如,考虑立方体图(cube map)。从空间中任何给定视点来看,不透明立方体同时可见的表面不超过三个。因此,为了考虑基于给定视点的显示,仅需要解码构成立方体图的代码图片中假定的六个正方形图块中的三个正方形图块。为此,至少对于需要独立图块的应用而言,独立图块可以在高级语法架构中基本取代独立切片。换言之,H.263+附录K中所谓的矩形切片将取代扫描顺序切片。运动约束图块集也可以作为该高级语法架构的要求。

[0079] 帧内预测分解机制(in-picture prediction breaking mechanisms)的大致概念是在规范空间和实施空间中的拼缀。在一实施例中,高级语法架构可以包括单独的标记,每个标记用于一个预测机制,该标记管理针对给定图块的数据的预测导入,并且置于图块头或参数集中。因此,该实现方式可能是一种更好、更简洁、更灵活的解决方案。

[0080] 在使用高级语法架构的一实施例中,可基于所使用的形貌来使能图块排布。例如,支持直接并行化的一种非常基本的图块排布机制可以作为所有形貌的一部分。另外,可以仅为某些形貌指定更高级的技术。例如,使用立方体图的360°形貌可以允许使用为该应用定制的运动约束独立图块,即可以按某种方式布置的六个图块,比如按 3×2 的方式布置,或者按照十字交叉的方式布置。其他形貌可以适用于其他投影格式。例如,二十面体类型的投影可能需要更多的图块,或者需要理想地响应于投影形状的其他类似的分解预测机制。

[0081] 除上述由特殊应用驱动的需求之外,已编码图片成为分解预测的最小单元。当已编码图片是分解预测的最小单元时,所有帧内预测机制均不会被分解,仅帧间预测机制会被分解。例如,某些较老的视频编码标准的某些元数据的运动补偿和帧间预测可能被分解。为了高效地支持没有切片/图块的已编码图片,一实施例中的高级语法架构可以包括图片头以携带语法元素,该语法元素在H.264/H.265中被置于切片头中,但关联整个图片。该语法元素可作为指向图片参数集(PPS)的参考。如先前在切片头中所提供的那样,图片头仅关联与其有关的图片,不关联之后的图片。换言之,图片头的内容是暂态的,且在图片头之间不存在预测(否则,即便是基于图片的容错也不起作用)。

[0082] 忽略容错的方面,图片头可携带于图片的第一图块(或仅有的图块)中,或携带于其自身VCL的NAL单元中。前者更加有效,后者结构上更加简洁。

[0083] 在一实施例中,不论是就语法(个体NAL单元)而言还是就功能和持久性范围而言,高级语法架构可以包含如现有架构中所提供的图片参数集(PPS)和序列参数集(SPS)。

[0084] 在SPS之上,高级语法架构可以包括解码器参数集(DPS),以包含标志、子形貌等。可确保DPS的内容在视频流的存续期间内保持恒定,直至接收流结尾NAL单元。

[0085] 在使用高级语法架构的一实施例中,可能需要允许外部携带流结尾NAL单元。例如,当SIP的re-invite改变了流的基本参数(且已被解码系统确认)时,必须告知解码系统的解码器将出现不同的DPS。如果将该信息提供给解码器的唯一方式是将其放入比特流中、通过防止起始码仿真等等,则会带来缺陷。此外在某些超时场景中,在比特流中放置该信息可能实际上不起作用。

[0086] 在许多情况下,当在分组网络上传送已编码图片时,已编码图片会大于最大传输单元(MTU)尺寸。由于引入不必要的分解预测不利于编码效率来(毕竟,完全移除切片的目的正在于编码效率),优选地不依赖于图块。另外,依赖图块并非优选的原因在于,图块具有两个可能矛盾的功能——并行化和应用特定的图块排布。规范空间中视频编解码器内部是否需要分割机制从两方面都可以说得通。如果在视频编解码器中需要分割机制,则在一实施例中高级语法架构可以使用例如H.265的“从属切片(dependent slices)”,仅此而已。或者,可以在高级语法架构中的较高层级处提供分割(fragmentation)。需要注意的是,H.26x视频的许多RTP有效载荷格式除依赖于基于编码器的MTU大小匹配(用于网关不进行代码转换的网关场景)而使用的切片之外,还确实包含某种形式的分割。

[0087] 参照图5,考虑到以上描述,一实施例中高级语法架构的语法层级(501)可基本如下所述。

[0088] 语法层级可以包括在会话的存续期间内保持不变的解码器参数集(DPS)(502)。

[0089] 在一些实施例中,语法层级可以包括视频参数集(VPS)(503),以将可缩放层连结在一起,其中,该视频参数集在跨越层边界的IDR处中断。

[0090] 语法层级可以包含功能性与在H.265中基本类似的序列参数集(SPS)(504),范围为已编码视频序列。

[0091] 语法层级可以包括位于相同语义层和具有相似范围的图片参数集(PPS)(505)和图片头(PH)(506)。也就是说,图片参数集(505)和图片头(506)可以覆盖全部已编码图片,但也可以在已编码图片之间变化。在功能上,图片参数集(505)可以与在H.265中基本类似,其范围为已编码图片。图片头(506)可以携带可能在图片之间变化的图片恒定数据,并且还

可以携带指向图片参数集(505)的参照。

[0092] 在一些实施例中,对于需要图块的应用场景,语法层级可以包括图块头(507)。

[0093] 在一些实施例中,语法层级可以包括分割单元头(508),例如,可以为从属的切片头。

[0094] 语法层级可以包括含有编码单元(CU)数据的已编码图片的VCL数据(509)。

[0095] 上述各种语法元素和语法层的交互的各方面将在下文更为详细地描述。

[0096] 【图片头/图片参数集交互】

[0097] 参照图6,以下根据本申请实施例描述了图片头(Picture Header,PH)(601)和图片参数集(Picture Parameter Set,PPS)(602)的交互。其中,图片头(601)和图片参数集(602)两者在语法中处于相同的语法层,例如已编码图片。

[0098] 参照图6,PH(601)和PPS(602)两者均可包含某些命名的语法元素。如图6中所示,一实施例中可以包括PH(601)和PPS(602),二者均恰好包含四个语法元素。然而可以理解,PH(601)和PPS(602)也可例如具有任何尺寸、具有不同的尺寸、包含可选的元素等。这些语法元素之一为PH_pps_id(603),其可以为PH(601)中指向PPS(602)的参照。该语法元素的语义可相当于与较老视频编码标准的切片头中pps_id的语义,即,视情况而定激活PPS和任何下游较高层的参数集,例如SPS、VPS、DPS。在PPS(602)中,PPS_pps_id(604)可以为自参照,也可以在接收时为PPS的ID标识。图片参数集识别是语法元素的示例,其中在某些情况下,PH(601)和PPS(602)中对应语法元素的值对每一适用的比特流必须相同。

[0099] 某些语法元素可仅存在于PH(601)中,而不存在于PPS(602)中。至少在一些情况下,这些语法元素可以属于其PH(601)包含该语法元素的图片,且这些语法元素可能在图片之间变化。由于基本上每次解码新图片时都需要激活新的PPS(602),因此将这些语法元素放入参数集中,例如放入PPS(602)中,可能效率低下。这些语法元素的一个示例可以为当前正在被处理的图片的标识,例如,时间参照、图片次序序数(picture order count,POC)等。例如,PH(601)可以包括POC(605)。PPS(602)中的相应条目被标记为用于图片类型的pic_type(606),此为仅在PPS(602)中而不在PH(601)中存在的语法元素的示例。因此,对于所有激活了PPS(602)的图片,使用pic_type(606)的值。

[0100] 某些语法元素可仅存在于PPS(602)中,而不存在于PH(601)中。可以理解,这个类别中可能有大多数较大的语法元素,这些语法元素可能或很可能与多个已编码图片关联但不适用于整个已编码视频序列。当这种语法元素不太可能在图片之间变化、因此激活不同的PPS(602)将不构成负担时,这种语法元素也可以存在于PPS(602)中,但不存在于PH(601)中。举例来说,考虑复杂的且可能较大的数据集,例如缩放矩阵,该数据集可以允许若干(最多全部)变换系数单独地选择量化器参数。该数据不太可能在给定的图片类型(例如I图片、P图片和B图片)内的典型图片组(GOP)的进程中变化。将缩放列表信息置于PH中会带来以下缺点:由于PH本质上是暂态的,因此可能需要随着每个已编码图片重复传输相同的缩放列表。

[0101] 然而,可能存在第三类语法元素。这类语法元素可以具有类似的名称,例如pps_foo(608)和ph_foo(607),并且在PPS(602)和PH(601)中均可存在这类语法元素。这些语法元素之间的关系可在视频技术或标准中根据语法元素的性质来定义,并且可根据语法元素而有所不同。

[0102] 例如,在同一或另一实施例中,在一些情况下,PH(601)中的语法元素的值,例如ph_foo(607)的值,可以覆写PPS(602)中的命名类似且语义绑定的语法元素的值,例如pps_foo(608)的值。

[0103] 在同一或另一实施例中,在某些其他情况下,PH(601)中的另一语法元素的值,例如ph_bar(609)的值,使用PPS(602)中的命名类似(此处为“bar”)和语义绑定的语法元素,例如pps_bar(610),作为某种形式的预测值。例如,在一些情况下,可将基于PH的语法元素ph_bar(609)叠加到PPS(602)中命名类似且语义上绑定的语法元素pps_bar(610)上,可从所述PPS(602)中的命名类似且语义上绑定的语法元素pps_bar(610)中减去基于PH的语法元素ph_bar(609),等等。

[0104] 【解码器参数集及比特流终止】

[0105] 解码器参数集(DPS)(502)与MPEG-2的序列头的非常相似,但其为参数集。因此,与MPEG-2序列头不同,DPS(502)性质上不是暂态的。由于激活时间可分别不同于参数集或头的解码时间,某些激活规则可以应用于不同于头的参数集,例如不同于MPEG-2序列头的参数集。考虑到这一重要区别,SPS可类比于MPEG-2的GOP头,而DPS可类比于MPEG-2的序列头。

[0106] DPS(502)的范围可以为H.265中所谓的视频比特流。视频比特流可包括许多已解码视频序列(CVS)。H.264和H.265中存在某些元素,其范围超出给定的CVS,其中首先是HRD参数。在规范空间中,H.264和H.265通过将CVS层级以上的参数置于SPS中、并要求相关信息在每个已编码视频序列中激活的SPS之间保持恒定,以处理这些参数。本申请实施例的DPS可以将这些语法元素累积到结构中,该结构可针对许多CVS保持已知和恒定。

[0107] 先前未说明的一个方面在于如何用信号通知解码器,使其在给定的时间点必须准备好接收将会需要不同DPS的参数集。该参数集可例如DPS或SPS,其中需要恒定的参数发生变化。

[0108] 虽然H.264和H.265均包含流结尾(EOS)NAL单元,但至少部分地因为下述的结构缺点,NAL单元可能不被频繁使用。

[0109] 在H.264和H.265中,不同于参数集等一些其他NAL单元类型,EOS需要在编码视频比特流中传送,并且其位置有明确限定的约束。例如,在H.264或H.265中,EOS不能被放置在编码图片的VCL的NAL单元内。实际中,需要编码器的协同或(至少)另一知晓视频编码标准的高级语法约束的实体的协同,以在已编码视频比特流中的适当位置插入EOS NAL单元。在至少一些场景中,这样的协作是不实用的。例如,参照图1的通信系统,假设接收终端脱离网络覆盖,且该终端正处于接收属于已编码图片NAL单元的期间,则编码器由于不再与解码器连接而不能向解码器提供EOS NAL单元。由于在接收已编码图片的NAL单元时连接中断,接收机处EOS NAL单元中的拼接也无法实行,因而EOS不能被置于在已编码图片的NAL单元之间。在现实应用中,接收终端可以将其解码器重置到已知的最新状态,但是该操作可能消耗几秒钟。虽然这对于所呈现的场景来说可能可以接受,但可能存在其他场景,这些场景中可能需要解码器的反应更加快速和清晰。

[0110] 在本申请公开的同一或另一实施例中,EOS可以作为视频流的一部分被接收(如在H.264/H.265中)或被带外接收。

[0111] 参照图7,在同一或另一实施例中,当带外接收EOS(701)并处理EOS时,解码器可以停用视频流的激活的解码器参数集。激活解码器参数集(DPS)的停用意味着在没有语法冲

突的情况下可以激活一个不同的DPS，该不同的DPS至少有一个值不同于先前激活的DPS。

[0112] 例如，停用激活的DPS可以包括解码器立即清空其缓冲器(702)并且停止输出已重建的图片(703)。在停用先前激活的DPS之后，解码器可以使其自身准备接收新的视频流(704)，其中，新的视频流与先前DPS相比可具有不同的DPS内容。然后，解码器可以通过激活(可选地，解码以及激活)先前的DPS或新的DPS来开始解码新的视频流(705)，其中，新的DPS可以与之前的DPS不同。可以在任何时间，甚至在带外接收EOS之前接受和解码新的DPS。如对参数集所常见的，接收和解码参数集的时间与解码过程无关的，只要参数集在其被激活时存在即可。接下来，可以开始根据新的DPS解码新的CVS(706)。

[0113] 本申请实施例还提供了一种用于解码至少一个视频流的装置，所述至少一个视频流中的每个视频流与各自的解码器参数集相关联，所述装置包括激活单元、处理单元和停用单元：

[0114] 所述激活单元，用于激活所述至少一个视频流中的第一视频流的解码器参数集；

[0115] 所述处理单元，用于处理在所述第一视频流之外的对所述解码器可用的流结尾网络抽象层NAL单元；以及

[0116] 所述停用单元，用于响应于处理所述流结尾NAL单元，停用所述第一视频流的解码器参数集。

[0117] 上述技术可实现为计算机软件，该计算机软件使用计算机可读指令，并物理存储于一个或多个计算机可读介质中。例如，图8示出了适于实现本申请公开的主题的某些实施例的计算机系统(800)。

[0118] 可以使用任何合适的机器代码或计算机语言来编码计算机软件，可对机器代码或计算机语言进行汇编、编译、链接等操作以创建包括指令的代码，该指令可以由一个或多个计算机中央处理单元(CPU)、图形处理单元(GPU)等直接执行或者通过解译、微代码执行等来执行。

[0119] 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。

[0120] 图8中示出的用于计算机系统(800)的部件本质上是示例性的，其目的不在于对实现本申请公开的实施例中计算机软件的使用范围或功能进行任何限制。部件的配置也不应当被解释为依赖于或受制于计算机系统(800)的示例性实施例中示出的任一部件或部件组合。

[0121] 计算机系统(800)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由一个或多个人类用户的输入，该输入方式例如触觉输入(比如击键、滑动、数据手套移动)、音频输入(比如语音、拍打)、视觉输入(比如手势)、嗅觉输入(未示出)。人机接口设备还可以用于采集某种媒体，该媒体不一定与人类有意识的输入直接相关，比如，音频(诸如：语音、音乐、环境声音)、图像(比如扫描图像、从静态图像摄像装置获取的摄影图像)、视频(比如二维视频、包括立体视频的三维视频)。

[0122] 输入人机接口设备可以包括以下中的一个或多个(每项仅示出一个)：键盘(801)、鼠标(802)、触控板(803)、触摸屏(810)、数据手套(未示出)、操纵杆(805)、麦克风(806)、扫描仪(807)、摄像装置(808)。

[0123] 计算机系统(800)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/味道来刺激一个或多个人类用户的感觉。该人机接口输出设备可以包括触觉输出设备(比如使用触摸屏(810)、数据手套(未示出)或操纵杆(805)产生触觉反馈,但也可使用不作为输入设备的触觉反馈设备)、音频输出设备(比如扬声器(809)、头戴式耳机(未示出))、视觉输出设备(比如:屏幕(810),包括CRT屏幕、LCD屏幕、等离子屏幕、OLED屏幕,每种屏幕具有或不具有触摸屏输入能力,每中屏幕具有或不具有触觉反馈能力——其中的一些可能能够通过诸如立体图像输出手段输出二维视觉输出或三维以上输出;虚拟现实眼镜(未示出);全息显示器和烟罐(未示出))、和打印机(未示出)。

[0124] 计算机系统(800)还可以包括人类可访问存储设备及其相关联的介质,例如,包括具有CD/DVD等介质(821)的CD/DVD ROM/RW(820)的光学介质、拇指驱动器(822)、可移除硬盘驱动器或固态驱动器(823)、传统磁性介质,例如磁带和软盘(未示出)、基于专用ROM/ASIC/PLD的设备,例如安全加密狗(未示出),等。

[0125] 本领域技术人员还应当理解,结合本申请公开的主题所使用的术语“计算机可读介质”不包含传输介质、载波、或其他瞬时信号。

[0126] 计算机系统(800)还可以包括到一个或多个通信网络的接口。网络可以例如是无线、有线连接、光网络。网络还可以是局域网、广域网、城域网、车辆工业网络、实时网络、延时容忍网络等。网络的示例包括:诸如以太网的局域网,无线LAN,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有线连线或无线广域数字网络电视,包括CANBus的车辆工业网络等。某些网络通常需要附接至某些通用数据端口或外围总线(849)(例如,计算机系统(800)的USB端口)的外部网络接口适配器,其他通常通过下述方式附接至系统总线以集成到计算机系统(800)的核心中(例如,通过以太网接口集成到PC计算机系统,或通过蜂窝网络接口集成到智能电话计算机系统的)。计算机系统(800)可通过使用这些网络中的任何网络与其他实体进行通信。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,从CAN总线到某些CAN总线设备)、或双向的(例如使用局域或广域数字网络到其他计算机系统)。可以在上述的每个网络和网络接口上使用某些协议和协议栈。

[0127] 前述人机接口设备、人类可访问存储设备和网络接口可以附接至计算机系统(800)的核心(840)。

[0128] 核心(840)可以包括一个或多个中央处理单元(CPU)(841)、图形处理单元(GPU)(842)、现场可编程门阵列(FPGA)(843)形式的专用可编程处理单元、用于某些任务的硬件加速器(844)等。可以通过系统总线(848)将这些设备连同只读存储器(ROM)(845)、随机存取存储器(846)、诸如内部非用户可访问硬盘驱动器、SSD等内部大容量存储装置(847)连接在一起。在一些计算机系统中,可以以一个或多个物理插头的形式访问系统总线(848),以通过另外的CPU、GPU等实现扩展。外围设备可以直接地或通过外围总线(849)附接至核心的系统总线(848)。外围总线的架构包括PCI、USB等。

[0129] CPU(841)、GPU(842)、FPGA(843)和加速器(844)可以执行某些指令,这些指令可以组合构成前述的计算机代码。该计算机代码可以存储在ROM(845)或RAM(846)中。临时数据也可以存储在RAM(846)中,而永久数据可以存储在例如内部大容量存储装置(847)中。可以通过使用缓存存储器来实现存储设备中的任何存储设备的快速存储和检索,该缓存存储器

可以与一个或多个CPU(841)、GPU(842)、大容量存储装置(847)、ROM(845)、RAM(846)等紧密关联。

[0130] 计算机可读介质上可具有计算机代码,该代码用于执行各种由计算机实现的操作。介质和计算机代码可以是出于本申请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技术人员公知且可用的。

[0131] 作为示例而非限制,(一个或多个)处理器(包括CPU、GPU、FPGA、加速器等)执行一个或多个有形计算机可读介质中含有的软件,从而使具有架构(800)的计算机系统——特别是核心(840)——可以发挥功能。这样的计算机可读介质可以是与前文介绍的用户可访问的大容量存储装置相关联的介质,也可以是具有非暂态性核心(840)的某些存储装置,比如核心内大容量存储装置(847)或ROM(845)。可以在这样的设备中存储实现本申请公开的各种实施例的软件,并且由核新执行(840)所述软件。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。软件可以使核心(840)——特别是其中的处理器(包括CPU、GPU、FPGA等)——执行本文描述的特定处理或其中特定部分,包括定义存储在RAM(846)中的数据结构以及根据由软件定义的处理修改该数据结构。另外地或替选地,计算机系统可以逻辑硬连线的方式、或其他表现为电路(例如,加速器(844))的方式来发挥功能,该逻辑硬连线或者电路可以代替软件或与软件一起操作,以执行本文描述的特定处理或其中特定部分。在合适的情况下,所涉及的软件可以涵盖逻辑,反之,所涉及的逻辑也可以涵盖软件。在合适的情况下,所涉及的计算机可读介质可以涵盖存储所执行软件的电路(比如集成电路(IC))、含有所执行逻辑的电路、或以上两者。本申请公开涵盖任何合适的硬件与软件的组合。

[0132] 虽然本申请已描述了若干示例性实施例,但存在落入本申请公开的范围内的改变、置换和各种等效替代。因此能够理解,本领域技术人员能够设想出多种系统和方法,这些系统和方法虽然在本文中没有明确示出或描述,但体现本申请公开的原理并且因此落入其精神和范围内。

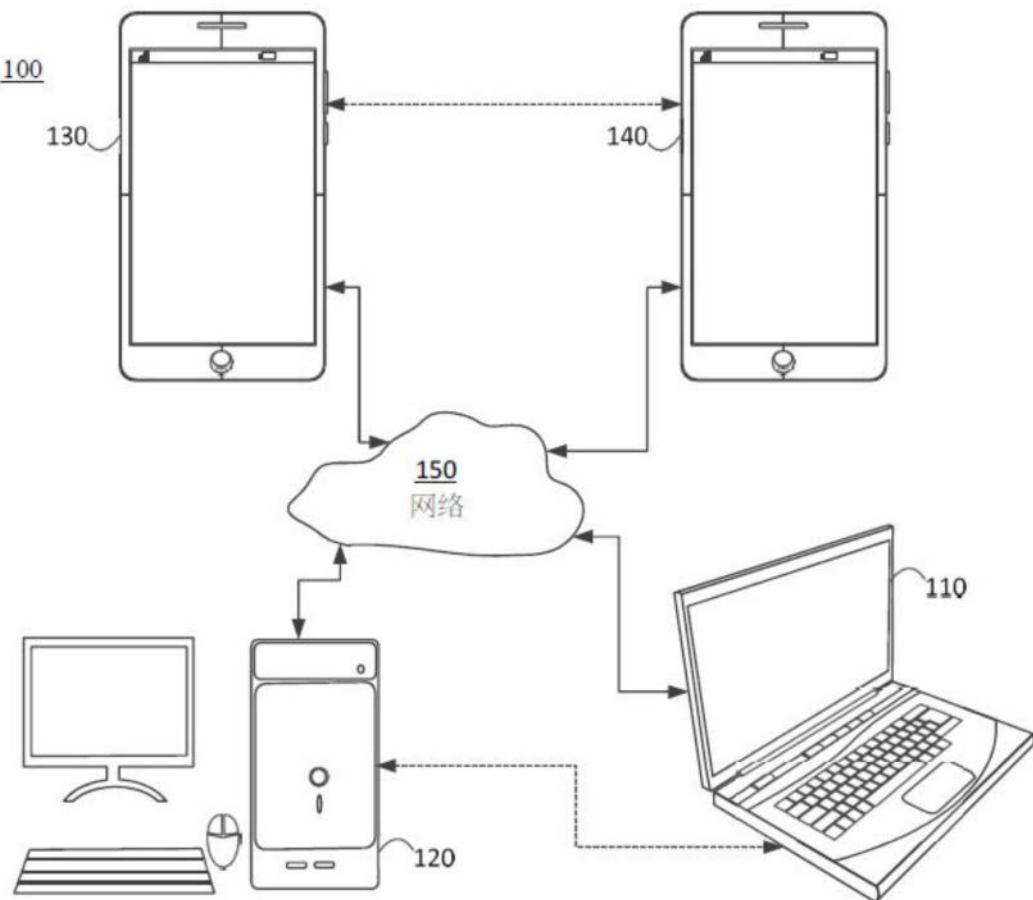


图1

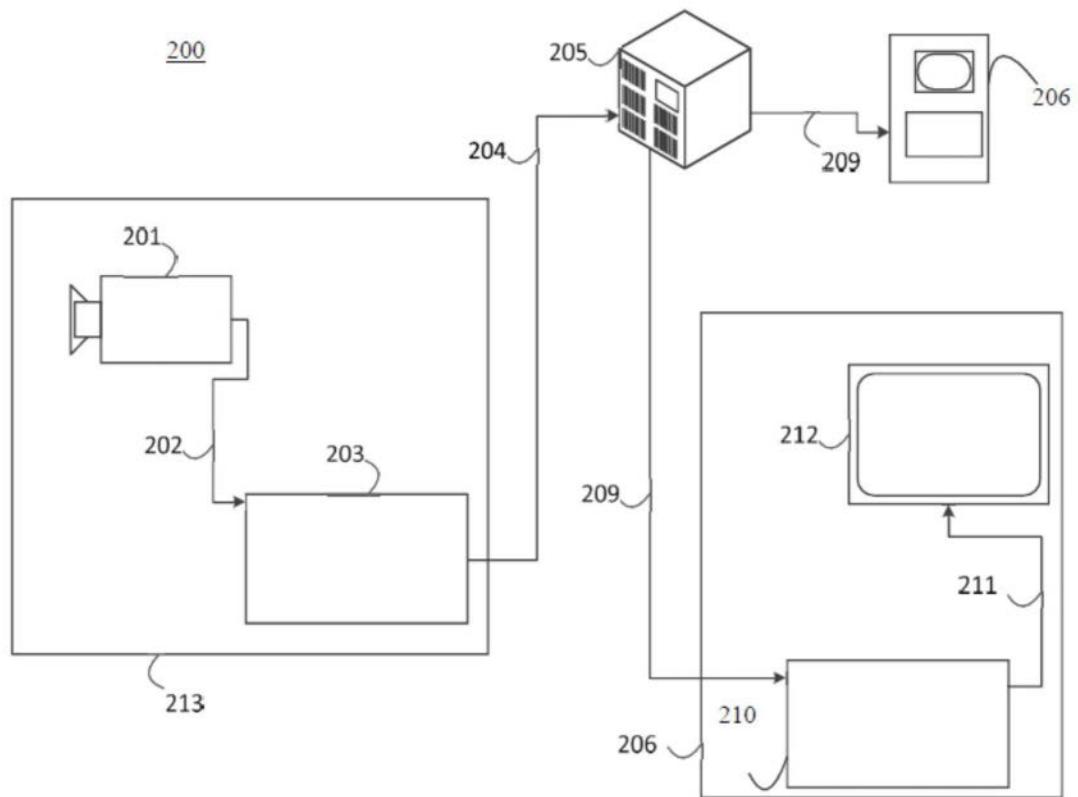


图2

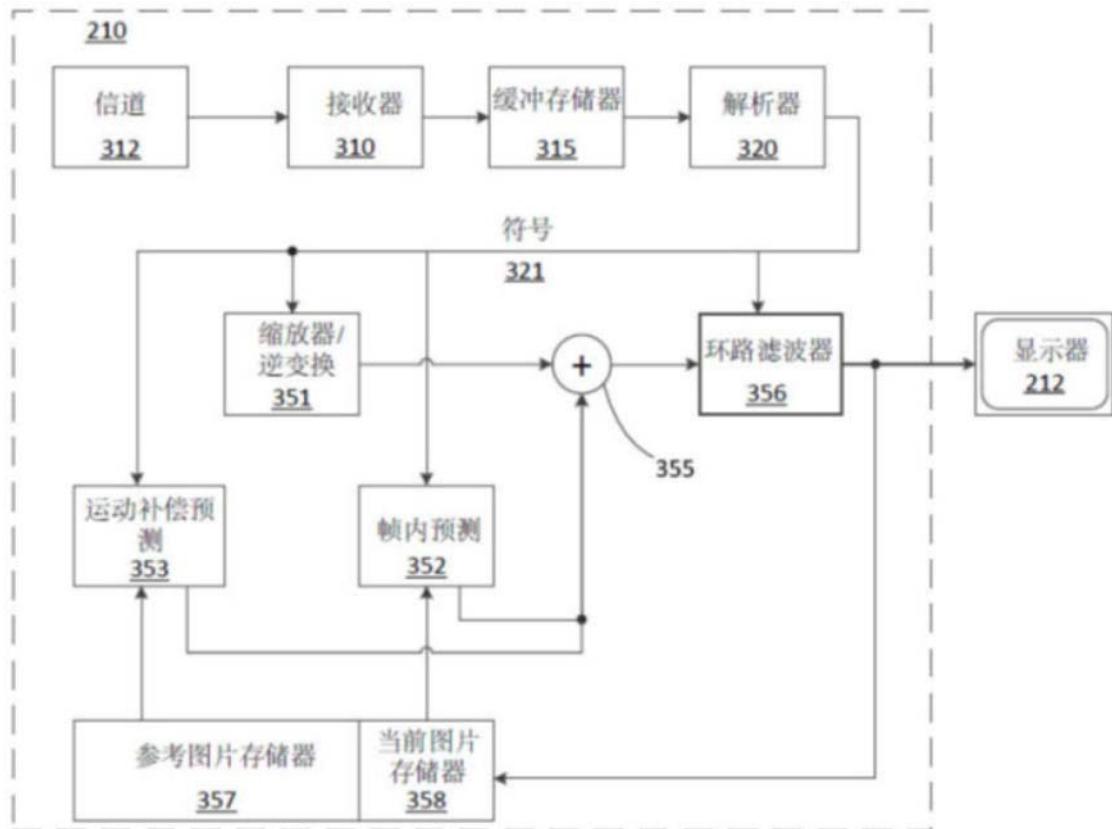


图3

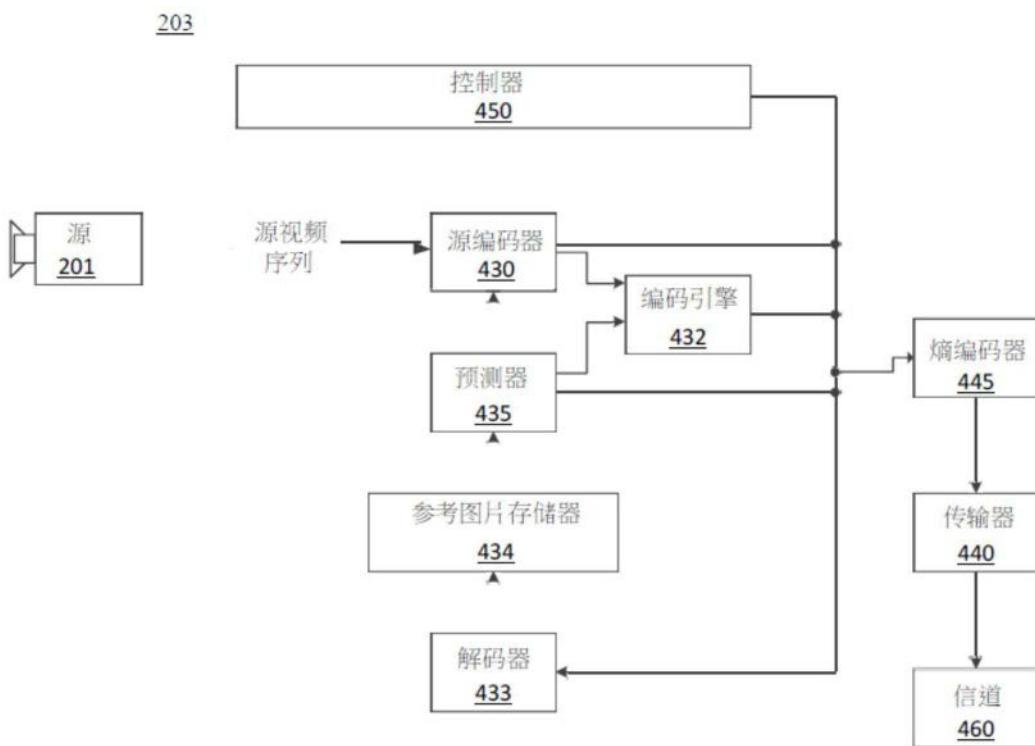


图4

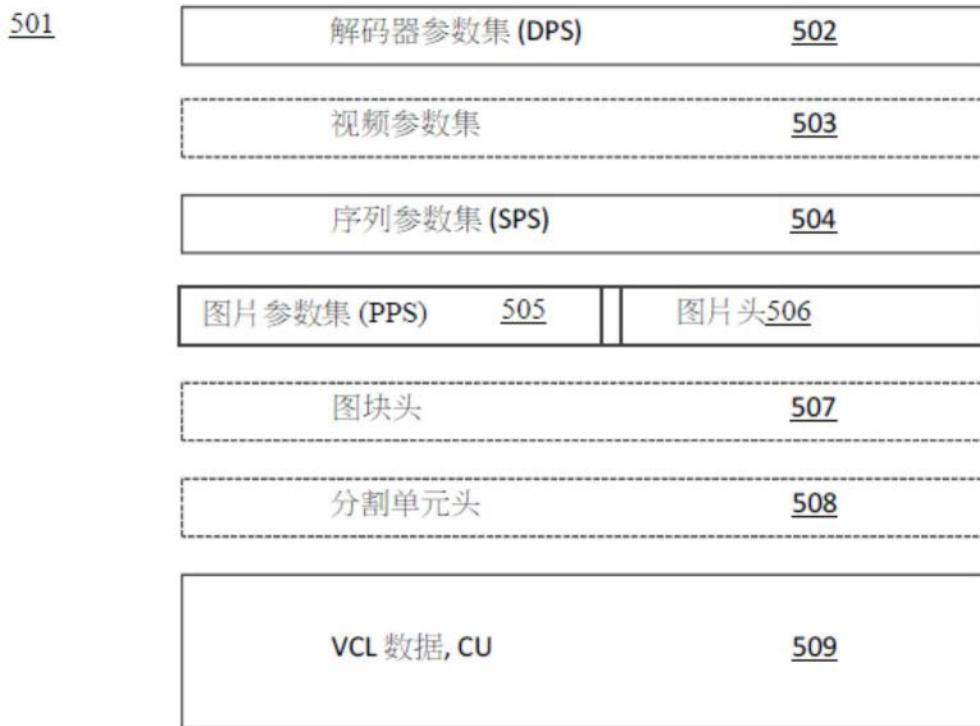


图5

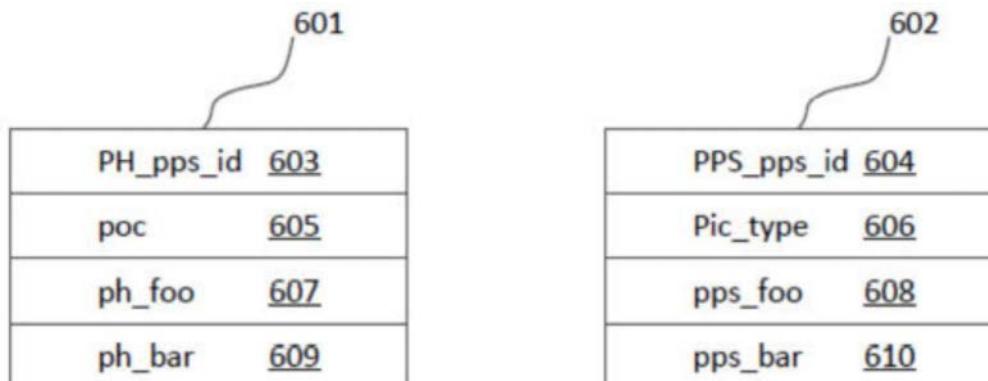


图6

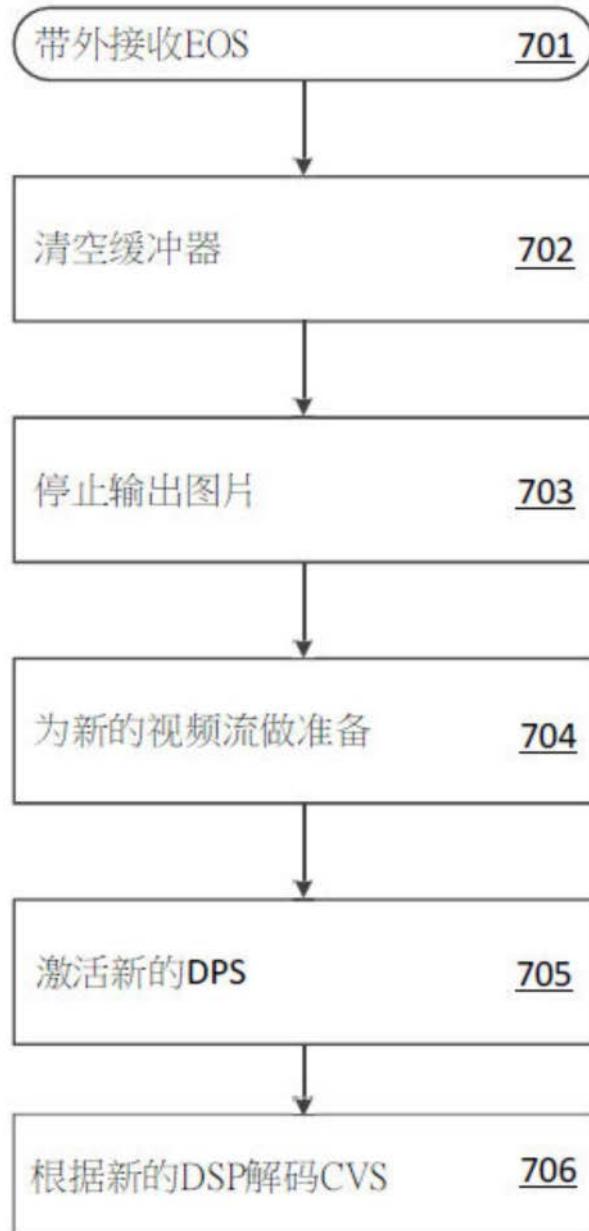


图7

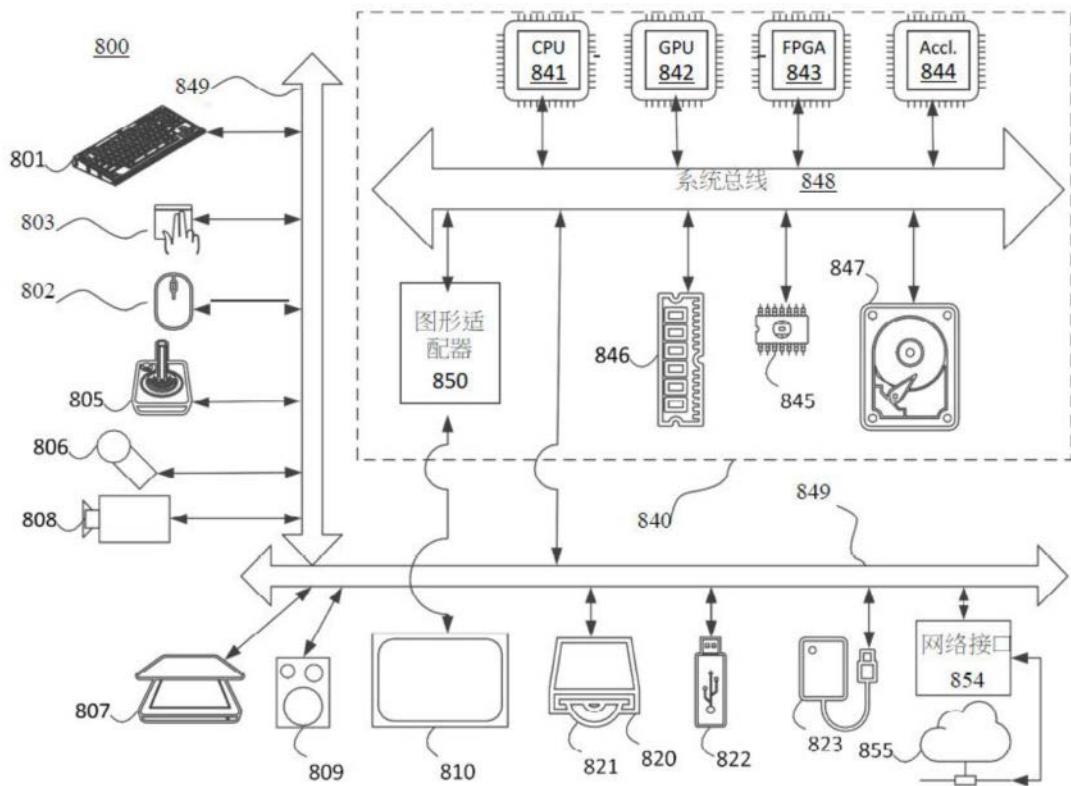


图8