



(12) 发明专利

(10) 授权公告号 CN 110896486 B

(45) 授权公告日 2023.05.19

(21) 申请号 201910830694.4

(22) 申请日 2019.09.04

(65) 同一申请的已公布的文献号
申请公布号 CN 110896486 A

(43) 申请公布日 2020.03.20

(30) 优先权数据

62/730,885 2018.09.13 US

16/232,675 2018.12.26 US

(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 史蒂芬·文格尔 刘杉

(74) 专利代理机构 北京德琦知识产权代理有限公司 11018

专利代理师 马媛媛 王琦

(51) Int.Cl.

H04N 19/70 (2014.01)

H04N 19/42 (2014.01)

H04N 19/44 (2014.01)

H04N 19/91 (2014.01)

H04N 19/184 (2014.01)

(56) 对比文件

CN 106664424 A, 2017.05.10

US 2016192027 A1, 2016.06.30

审查员 蒋碧波

权利要求书2页 说明书14页 附图7页

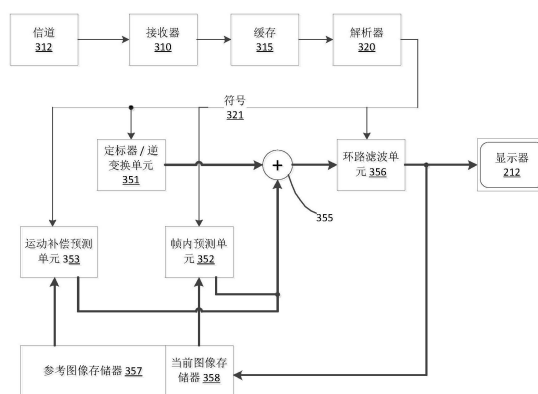
(54) 发明名称

利用高层语法架构进行编解码的方法和设
备

(57) 摘要

一种用于解码视频流的方法和设备,所述视频流包括至少两个编码视频序列,所述至少两个编码视频序列的每一个使用的序列参数集与其它编码视频序列使用的序列参数集至少有一个数值不相同,所述至少两个编码视频序列的每一个包括至少两个编码图像。该方法包括,在解码所述至少两个编码视频序列的任一个编码图像之前,解码器解码并启用与所述至少两个编码视频序列相关的单个解码器参数集。该方法还包括,所述解码器解码所述至少两个编码视频序列的至少一个编码图像。

解码器 210



1. 一种解码视频流的方法,其特征在于,所述视频流包括至少两个编码视频序列,所述至少两个编码视频序列中的每一个编码视频序列使用的序列参数集SPS (Sequence Parameter Set) 与其它编码视频序列使用的SPS至少有一个数值不相同,所述至少两个编码视频序列的每一个编码视频序列包括至少两个编码图像,该方法包括:

在解码所述至少两个编码视频序列的任一个编码图像之前,由解码器解码并启用与所述至少两个编码视频序列相关的单个解码器参数集DPS (Decoder Parameter Set),所述DPS适用于所述视频流包括的所有编码视频序列;

所述解码器解码所述至少两个编码视频序列的至少一个编码图像;及

响应于流结束网络抽象层NAL (Network Abstraction Layer) 单元接收,所述解码器停止使用所述DPS;

其中,所述解码器根据使用一语法架构的视频编解码器技术或标准解码所述至少一个编码图像,所述语法架构包括视频参数集VPS (Video Parameter Set),所述VPS用于将可缩放层结合在一起,所述VPS所处的语法等级低于所述DPS的语法等级。

2. 根据权利要求1所述的方法,其特征在于,进一步包括:

利用所述至少两个编码视频序列中的一个编码视频序列的序列参数集的参数,解码该编码视频序列的至少两个编码图像。

3. 根据权利要求1所述的方法,其特征在于,其中:

所述语法架构包括以下中的至少一个:图像头部PH (Picture Header) 和至少一个图像参数集PPS (Picture Parameter Set)。

4. 根据权利要求3所述的方法,其特征在于,其中:

所述语法架构使用包括PH和至少一个PPS的语法架构。

5. 根据权利要求4所述的方法,其特征在于,其中:

所述PH和所述至少一个PPS所处的语法等级低于所述DPS和所述SPS的语法等级。

6. 根据权利要求5所述的方法,其特征在于,其中:

所述图像头部和所述至少一个PPS位于同一语法等级。

7. 根据权利要求1所述的方法,其特征在于,其中:

所述语法架构不包括分片(slice)语法架构。

8. 根据权利要求1所述的方法,其特征在于,其中:

所述语法架构使用包括至少一个图块头部TH (Tile Header) 的语法架构。

9. 一种用于解码视频流的设备,其特征在于,所述视频流包括至少两个编码视频序列,所述至少两个编码视频序列的每一个编码视频序列使用的序列参数集SPS (Sequence Parameter Set) 与其它编码视频序列使用的SPS至少有一个数值不相同,所述至少两个编码视频序列的每一个编码视频序列包括至少两个编码图像,所述设备包括:

解码器,用于:

在解码所述至少两个编码视频序列的任一个编码图像之前,解码并启用与所述至少两个编码视频序列相关的单个解码器参数集DPS (Decoder Parameter Set),所述DPS适用于所述视频流包括的所有编码视频序列;

解码所述至少两个编码视频序列的至少一个编码图像;及

响应于流结束网络抽象层NAL (Network Abstraction Layer) 单元接收,停止使用所

述DPS;

其中,所述解码器根据使用一语法架构的视频编解码器技术或标准解码所述至少一个编码图像,所述语法架构包括视频参数集VPS(Video Parameter Set),所述VPS用于将可缩放层结合在一起,所述VPS所处的语法等级低于所述DPS的语法等级。

10.根据权利要求9所述的设备,其特征在于,其中:

所述解码器用于,通过利用该编码视频序列的相应序列参数集的参数来解码至少两个编码视频序列中的编码视频序列的至少两个编码图像。

11.根据权利要求9所述的设备,其特征在于,其中:

所述语法架构包括以下中的至少一个:图像头部PH(Picture Header)和至少一个图像参数集PPS(Picture Parameter Set)。

12.根据权利要求11所述的设备,其特征在于,其中:

所述视频编解码器技术或标准的所述语法架构使用包括PH和至少一个PPS的语法架构。

13.根据权利要求12所述的设备,其特征在于,其中:

所述图像头部和所述至少一个PPS所处的语法等级低于所述DPS和所述SPS的语法等级。

14.根据权利要求13所述的设备,其特征在于,其中:

所述图像头部和所述至少一个PPS位于同一语法等级。

15.根据权利要求9所述的设备,其特征在于,其中:

所述语法架构不包括分片(slice)语法架构。

16.根据权利要求9所述的设备,其特征在于,其中:

所述语法架构包括至少一个图块头部TH(Tile Header)。

17.一种非暂存计算机可读介质,存储有指令,所述指令包括一个或多个指令,由设备的一个或多个处理器执行时,使所述一个或多个处理器执行如权利要求1-8任一项所述的方法的操作。

利用高层语法架构进行编解码的方法和设备

[0001] 相关文件

[0002] 本申请要求2018年9月13日提交的美国临时申请号62/730,885,以及2018年12月26日提交的美国申请号16/232,675的优先权,其公开内容通过引用整体并入本文。

技术领域

[0003] 各实施例的方法和设备涉及视频编码和解码,特别地,涉及利用高层语法架构进行编解码的方法和设备。还公开了将高级语法参数的优先级和持续性编码在参数集中,例如序列和图像参数集,以及一些高级头部,例如图像头部。

背景技术

[0004] 利用具有运动补偿的图像间预测的视频编码和解码已得到应用。未压缩的数字视频可由图像序列组成,每个图像具有一定的空间维度,例如 $1920 \times 10 \times 0$ 的亮度采样和相关色度采样。图像序列可具有固定或可变的图像速率(俗称帧率),例如60张图像每秒或60Hz。未压缩的视频需要较高的比特率。例如,每个采样为x位的 $10 \times 0p604:2:0$ (60Hz帧率下的 $1920 \times 10 \times 0$ 亮度采样分辨率)的视频需要接近1.5G比特/秒的带宽。长度为一小时的这种视频需要600G字节以上的存储器。

[0005] 视频编码和解码的一个目的是,通过压缩来降低输入视频信号的冗余。在一些情况下,压缩可将带宽或存储器的需求减小至少两个数量级。可使用无损压缩、有损压缩,或其组合。无损压缩指可从压缩原始信号中重建原始信号的准确副本的技术。当利用有损压缩时,重建信号可与原始信号不一致,但原始信号和重建信号之间的失真足够小,以使重建信号能够实现所期望的用途。视频领域中广泛采用有损压缩。容许的失真量取决于应用。例如,消费型直播应用的用户比电视节目应用的用户能容忍更高的失真。可实现的压缩比可反映出:可允许/可容许失真的越高,可产生的压缩比越高。

[0006] 视频编码器和解码器可利用来自一些广泛类别的技术,例如包括运动补偿、变换、量化和熵编码,一其中一些将在下文中介绍。

[0007] 在H.264之前,一些视频编解码器,如可视化MPEG-2使用临时性头部组成的层次结构,包括序列头部、图像组(group of picture, GOP)头部、图像头部,以及分片(slice)头部。每个头部所包括的语法元素与所有下层的语法架构相关。例如,序列头部的语法元素与序列中的所有GOP、GOP中的所有图像、以及图像中的所有分片相关。GOP头部的语法元素与GOP中的所有图像、以及图像中的所有分片相关。这种层次结构可提高编码效率,但也会使抗差错性能不够理想。例如,如果序列头部的关键性信息在传输中丢失,序列中的GOP、图像或分片均无法被解码。

[0008] 2003年之后的一些ITU和MPEG视频编解码器,即H.264和H.265,在分片头部上不再使用临时头部,而是依赖参数集。在每个语法等级,如序列或图像等级中,解码器或外部装置可从比特流接收一个或多个参数集。这些(可能是多个)相同类型的参数集中哪一个将用于给定序列或图像的解码,取决于参照信息,该参照信息可以编码在例如分片头部(对于图

像参数集 (pictureparameterset, PPS) 来说) 或 PPS (对于序列参数集 (sequenceparameterset, SPS) 来说) 中。这种体系结构的优点是, 即使比特流本身通过有损信道发送, 也能够提高相关参数集的传输可靠性, 或者, 通过 (可能在相关参数集被首次使用之前) 发送相关参数集的冗余的副本, 提高了相关参数集被接收到的可能性。一个缺点可能是, 从相同数量和类型的语法元素所需要的比特数来看, 参数集比 MPEG-2 类型头部的发送成本更高。而且, 在这种体系结构下, 在不同图像间频繁变化而在给定图像中保持不变的一些语法元素可以以多个冗余副本的形式包括于每个分片头部中。尽管这样做可能使分片可被独立解码 (至少从语法依赖端的熵解码的角度), 但需要耗费更多比特。

[0009] 在 H.264 的设计期间, 出于错误复原的原因, 分片的独立可解码性是主要的设计目标。然而自 2003 年起, 由于传递编码视频的网络体系结构的改进, 以及预测机制的进步, 丢失分片的隐藏变得越来越低效, 使得分片的独立可解码性的吸引力显著下降分片。

发明内容

[0010] 由于对分片独立可解码性的需求已经转移, 需要新的高级语法架构, 使得在至少一些给定图像的丢失可在解码器中合理隐藏的情况下, 保持良好的抗差错性能, 并在编码效率上充分利用 MPEG-2 类型头部结构的优点。本申请公开的一些实施例提供了一种高级语法架构, 可以保持良好的抗差错性能和编码效率。

[0011] 根据本申请公开的一个方面, 一种方法可用于解码包括至少两个编码视频序列的视频流。该至少两个编码视频序列中的每一个序列使用的序列参数集与其它序列使用的序列参数集至少有一个数值互不相同。该至少两个视频序列中的每一个序列包括至少两个编码图像。该方法可包括, 在解码该至少两个视频序列的任一个编码图像之前, 由解码器解码和激活与该至少两个编码视频序列相关的单个解码器参数集 (DecoderParameterSet)。该方法还可以包括, 由该解码器解码该至少两个编码视频序列的至少一个编码图像。

[0012] 根据本申请公开的一个方面, 一种设备可用于解码包括至少两个编码视频序列的视频流。该至少两个编码视频序列中的每一个序列使用的序列参数集与其它序列使用的序列参数集至少有一个数值不相同。该至少两个视频序列中的每一个序列包括至少两个编码图像。该设备可包括解码器, 用于在解码该至少两个视频序列的任一个编码图像之前, 解码和激活与该至少两个编码视频序列相关的单个解码器参数集, 并解码该至少两个编码视频序列中的至少一个编码图像。

[0013] 根据本申请公开的一个方面, 可使用一种非暂存计算机可读介质存储指令。该指令可包括: 一个或多个指令, 当由设备的一个或多个处理器执行时, 可使一个或多个处理器在解码该至少两个视频序列的任一个编码图像之前, 解码和激活与该至少两个编码视频序列相关的单个解码器参数集, 并解码该至少两个编码视频序列中的至少一个编码图像。

附图说明

[0014] 如下详细描述和附图可使所公开的主题的进一步特征、本质和多种优点更加明显, 其中:

[0015] 图1是一个实施例的通信系统的简化框图的示意图。

[0016] 图2是一个实施例的流媒体系统的简化框图的示意图。

- [0017] 图3是一个实施例的视频解码器和显示器的简化框图的示意图。
- [0018] 图4是一个实施例的视频编码器和视频源的简化框图的示意图。
- [0019] 图5是一个实施例的高级语法架构中的语法层次结构的示意图。
- [0020] 图6是一个实施例的图像头部和图像参数集的示意图。
- [0021] 图7是一个实施例中,当流的末端从带外接收时,解码器参数集变化的流程示意图。
- [0022] 图8是一个实施例的计算机系统的示意图。

具体实施方式

[0023] 图1示出本公开实施例的通信系统(100)的简化框图。系统(100)可包括通过网络(150)互联的至少两个终端(110-120)。对于单向数据传输,第一终端(110)可在本地位置编码视频数据,用于经网络(150)传输至另一终端(120)。第二终端(120)可从网络(150)接收另一终端的编码视频数据,解码该编码数据并显示恢复出的视频数据。单向数据传输常用于媒体服务应用等。

[0024] 图1示出本申请实施例的通信系统(100)的简化框图。系统(100)可包括通过网络(150)互联的至少两个终端(110-120)。对于单向数据传输,第一终端(110)可在本地位置编码视频数据,用于经网络(150)传输至另一终端(120)。第二终端(120)可从网络(150)接收另一终端的编码视频数据,解码该编码数据并显示恢复出的视频数据。单向数据传输常用于媒体服务应用等。

[0025] 图1示出第二对终端(130,140),可以支持编码视频的双向传输,例如,在视频会议期间。对于双向数据传输,每个终端(130,140)可对在本地位置拍摄的视频数据进行编码,以便通过网络(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编码视频数据,可解码该编码数据,并可在本地显示器设备显示恢复出的视频数据。

[0026] 在图1中,终端(110-140)可以是诸如服务器、个人计算机以及智能电话和/或其它任何类型的终端。例如,终端(110-140)可以是笔记本电脑、平板电脑、媒体播放器、和/或专用的视频会议装置。网络(150)表示任何数量的网络,可以在终端(110-140)之间传输编码视频数据,可以包括,例如,有线和/或无线通信网络。通信网络(150)可在电路交换和/或包交换的信道中交换数据。代表性的网络包括电信网络、局域网、广域网、和/或互联网。为本文讨论的目的,除非有明确说明,网络(150)的体系结构和拓扑与本申请公开的操作无关。

[0027] 作为本申请公开主题的应用示例,图2示出一种在流媒体环境下视频编码器和解码器的部署方式。本申请公开的主题可与其它支持视频的应用共同使用,包括,例如视频会议、数字TV、将压缩视频存储到数字媒体,包括CD、DVD、记忆棒等,的应用等。

[0028] 如图2所示,流媒体系统(200)可包括拍摄子系统(213),其包括视频源(201)和编码器(203)。流媒体系统(200)还可以包括至少一个流媒体服务器(205)和/或至少一个流媒体客户端(206)。

[0029] 视频源(201)可以创建,例如,未压缩视频采样流(202)。视频源(201)可以是,例如数码照相机。采样流(202)(用粗线表示,强调比编码视频比特流的数据量大)可由耦合于照相机(201)的编码器(203)处理。编码器(203)可包括硬件、软件或其组合,以使能或实现如下文详细描述的本申请公开主题的各方面。编码器(203)还可生成编码视频比特流(204)。

编码视频比特流(204)(用细线表示,以强调比未压缩视频采样流(202)的数据量更小)可存储于流媒体服务器(205)以备后续使用。一个或多个流媒体客户端(206)可访问流媒体服务器(205),以获取视频比特流(209),视频比特流(209)可以是编码视频比特流(204)的副本。

[0030] 流媒体客户端(206)可包括视频解码器(210)和显示器(212)。视频解码器(210)可以,例如,解码视频比特流(209),其为接收到的编码视频比特流(204)的副本,并创建输出视频采样流(211),其可绘制在显示器(212)或另一绘制设备(未示出)上。在一些流媒体系统中,视频比特流(204,209)可根据一些视频编码/压缩标准进行编码。该标准的示例包括,但不限于,ITU-T建议H.265。目前正在开发的视频编码标准俗称多用途视频编码(VersatileVideoCoding,VVC)。本申请公开的各实施例可用于VVC环境。

[0031] 图3示出本申请实施例的连接至显示器(212)的视频解码器(210)的示例功能框图。

[0032] 视频解码器(210)可包括信道(312)、接收器(310)、缓存(315)、熵解码器/解析器(320)、定标器/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)、环路滤波单元(356)、参考图像存储器(357),以及当前图像存储器(358)。在至少一个实施例中,视频解码器(210)可包括集成电路、一系列集成电路、和/或其它电子电路。视频解码器(210)的部分或整体还可以由运行于具有关联存储器的一个或多个CPU上的软件实现。

[0033] 在本实施例以及其它实施例中,接收器(310)可接收需要视频解码器(210)解码的一个或多个编码视频序列,每次解码一个编码视频序列,其中每个编码视频序列的解码独立于其它编码视频序列。编码视频序列可接收自信道(312),其可以是连接到存储编码视频数据的存储设备的硬件/软件链路。接收器(310)可接收编码视频数据以及其它数据,例如,编码音频数据和/或辅助数据流,其可被前馈至其相应实体(未示出)。接收器(310)可分离编码视频序列和其它数据。为克服网络抖动,缓存(315)可耦合在接收器(310)和熵解码器/解析器(320)(下称“解析器”)之间。当接收器(310)从具有充足带宽和可控性的存储/前转设备或从等同步网络接收数据时,可不使用缓存(315)或使用小型缓存。对于尽力而为(besteffort)式的报文网络,如互联网,可能需要使用缓存(315),可以是较大容量的缓存,或是自适应调整容量的缓存。

[0034] 视频解码器(210)可以包括解析器(320),用于从熵编码视频序列中重建符号(321)。这些符号的分类包括,例如,用于管理视频解码器(210)的操作的信息,还可以包括控制绘制设备的信息。绘制设备可以是耦合于图2所示的解码器的显示器(212)。用于控制绘制设备的信息可以是,例如,补充增强信息(SupplementaryEnhancementInformation,SEI消息)或视频可用性信息(SupplementaryEnhancementInformation,VUI)参数集片段(未示出)的形式。解析器(320)可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依据视频编码技术或标准,并可遵循本领域技术人员公知的原理,包括可变量长度编码、哈夫曼编码、具有或不具有上下文敏感度的算术编码等。解析器(320)可以基于一组图像所对应的至少一个参数,从编码视频序列中提取对应于视频解码器中像素的至少一个子组的子组参数的集合。子组可包括图像组(GOP)、图像、图块(tile)、分片、宏块(macroblock)、编码单元(CU)、块、变换单元(TU)、预测单元(PU)等。解析器(320)还可以从编码视频序列中提取出如变换系数、量化器参数值、运动矢量等信息。

[0035] 解析器(320)可对从缓存(315)接收的视频序列执行熵解码/解析操作,以创建符号(321)。

[0036] 符号(321)的重建可根据编码视频图像或其中一些部分的类型(如:图像间和图像内,块间和块内)及其它因素涉及多个不同的单元。哪个单元被涉及、如何被涉及,可由解析器(320)由从编码视频序列中解析得到的子组控制信息控制。为清晰起见,未示出子组控制信息在解析器(320)和下述多个单元之间的流动。

[0037] 除已提到的功能块之外,视频解码器(210)还可从概念上细分为如下所述的多个功能单元。在商业约束下的实际实现中,多个单元相互密切交互,并且至少部分单元可以相互集成。然而,为描述本申请公开的主题,下文进行功能单元的概念细分是合适的。

[0038] 一个单元可为定标器/逆变换单元(351)。定标器/逆变换单元(351)可接收量化变换系数以及控制信息,包括使用哪种变换、块的大小、量化因子、量化缩放矩阵等,作为来自解析器(320)的符号(321)。定标器/逆变换单元(351)可输出包括采样值的块,这些采样值可以输入聚合器(355)。

[0039] 在一些情况下,定标器/逆变换(351)的输出采样可与内部编码块相关;内部编码块是指,未使用来自先前已重建的图像的预测信息,但可使用来自当前图像的先前已重建部分的预测信息的块。该预测信息可由内部图像预测单元(352)提供。在一些情况下,内部图像预测单元(352)生成与当前正在重建的块的尺寸和形状相同的块,生成该块利用了从当前图像存储器(358)获取的当前(已部分重建的)图像中周围的已重建的信息。在一些情况下,对每个采样,聚合器(355)可以将帧内预测单元(352)已生成的预测信息加入由定标器/逆变换单元(351)提供的输出采样信息。

[0040] 在其它情况下,定标器/逆变换单元(351)的输出采样可从属于帧间编码的、可能用作运动补偿的块。在该情况下,运动补偿预测单元(353)可访问参考图像存储器(357),以获取用于预测的采样。在根据与该块相关的符号(321)对获取的采样进行运动补偿之后,这些采样可由聚合器(355)加入到定标器/逆变换单元(351)的输出(在这种情况下,称为残差采样或残差信号),以生成输出采样信息。运动补偿预测单元(353)获取的预测采样在参考图像存储器(357)中的地址可由运动矢量控制。运动矢量可以符号(321)的形式由矢量运动补偿预测单元(353)使用,可包括,例如x,y和参考图像分量。当使用子采样的精确运动矢量时,矢量运动补偿还可包括对从参考图像存储器(357)获取的采样值进行插值、运动矢量预测机制等。

[0041] 聚合器(355)的输出采样可在环路滤波单元(356)中采用多种环路滤波技术进行处理。视频压缩技术可包括环内滤波技术,环内滤波技术由编码视频比特流中的参数控制,并可由环路滤波单元(356)使用,作为来自解析器(320)的符号(321),但也可响应于对编码图像或编码视频序列的先前(按照解码顺序)部分进行解码的期间获得的元信息,还可以响应于先前重建并经过环路滤波的采样值。

[0042] 环路滤波单元(356)的输出可为采样流,可输出至绘制设备,例如显示器(212),也可以存储在参考图像存储器(357)中,以备后续在图像间预测中使用。

[0043] 一旦被完全重建,一些编码图像可在后续预测中用作参考图像。一旦编码图像被完全重建并且被(例如解析器(320))标识为参考图像,当前图像存储器(358)存储的当前参考图像可成为参考图像存储器(357)的一部分,并在开始重建后续编码图像之前,可以重新

分配新的当前图像存储器。

[0044] 视频解码器(210)可根据标准中记载的预定义视频压缩技术,如ITU-TRec.H.265,来执行解码操作。编码视频序列可遵循如视频压缩技术文档或标准中,特别是其概要文档中,指定的视频压缩技术或标准的语法,从这个意义上来说,编码视频序列遵循所使用的视频压缩技术或标准指定的语法。并且,为遵循一些视频压缩技术或标准,编码视频序列的复杂度可在视频压缩技术或标准的等级所限定的界限之内。在一些情况下,各等级对最大图像尺寸、最大帧率、最大重建采样率(例如以每秒兆次采样测量)、最大参考图像尺寸等进行限制。在一些情况下,等级设置的限制可进一步通过假设参考解码器(HypotheticalReferenceDecoder,HRD)规范和编码视频序列中指示的HRD缓存管理的元数据来限定。

[0045] 在一个实施例中,接收器(310)可连同已编码视频一起接收附加(冗余)数据。附加数据可以是编码视频序列的一部分。附加数据可由视频解码器(210)用以对数据进行适当解码和/或较准确地重建原始视频数据。附加数据可以为时间、空间或信噪比(signalnoiseratio,SNR)增强层、冗余分片、冗余图像、前向纠错码等形式。

[0046] 图4示出本申请实施例的与视频源(201)关联的视频编码器(203)的功能框图。

[0047] 视频编码器(203)可包括一个编码器,该编码器可以是,例如源编码器(430)、编码引擎(432)、(本地)解码器(433)、参考图像存储器(434)、预测器(435)、传输器(440)、熵编码器(445)、控制器(450),以及信道(460)。

[0048] 编码器(203)可从视频源(201)(并非编码器的一部分)接收视频采样,视频源(201)可采集将由编码器(203)编码的视频图像。

[0049] 视频源(201)可以任何合适位的深度(例如:x位、10位、12位,…)、任何色彩空间(例如,BT.601YCrCb、RGB、…)和任何合适的采样结构(例如YCrCb4:2:0、YCrCb4:4:4)的数字视频采样流的形式来提供将由编码器(203)编码的源视频序列。在媒体服务系统中,视频源(201)可以是存储先前已准备好的视频的存储设备。在视频会议系统中,视频源(203)可以是采集本地图像信息作为视频序列的相机。可将视频数据提供为多个单独的图像,当按顺序观看时,这些图像呈现出运动效果。图像自身可被组织为空间像素阵列,其中每个像素可根据所用的采样结构、色彩空间等包括一个或多个采样。本领域技术人员可以很容易理解像素和采样之间的关系。下文侧重描述采样。

[0050] 根据实施例,编码器(203)可实时或在应用需要的任何其它时间约束下,对源视频序列的图像进行编码并压缩为编码视频序列(443)。施行适当的编码速度是控制器(450)的一个功能。控制器(450)还可控制如下文所述的其它功能单元,并可在功能上耦接到这些单元。为清楚起见,图中未标示耦接。由控制器(450)设置的参数可包括速率控制相关参数(图像跳过、量化器、率失真优化技术的 λ 值等)、图像尺寸、图像组(GOP)布局、最大运动矢量搜索范围等。本领域技术人员很容易理解控制器(450)具有的其它功能,这些功能属于针对特定系统设计而优化的视频编码器(203)。

[0051] 一些视频编码器在本领域技术人员所理解的“编码环路”中进行操作。简单来说,当一些视频压缩技术中的符号与编码视频比特流之间的压缩无损时,编码环路可包括源编码器(430)的编码部分(负责基于待编码的输入图像和参考图像创建符号),以及嵌入编码器(203)中的、用于重建符号以创建(远程)解码器也会创建的采样数据的(本地)解码器

(433)。重建的采样流可输入到参考图像存储器(434)。由于符号流的解码产生与解码器位置(本地或远程)无关的比特级精确结果,因此参考图像存储器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准对应的。换言之,编码器的预测部分“看到”的参考图像样本与解码期间使用预测时所“看到”的样本值完全相同。这种参考图像同步性的基本原理(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)是本领域技术人员已知的。

[0052] “本地”解码器(433)的操作可与“远程”解码器(210)基本相同,“远程”解码器(210)已在上文结合图3详细描述。然而,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/解码为编码视频序列时,解码器(210)的熵解码部分,包括信道(312)、接收器(310)、缓存(315)和解析器(320),不需要完整地实现于本地解码器(433)中。

[0053] 此时可以观察到,除解码器中存在的解析/熵解码之外,任何解码器技术也需要以基本相同的功能形式存在于对应的编码器中。因此,本申请公开的主题侧重于解码器操作。编码器技术的描述可被省略,因为其可以是已详尽描述的解码器技术的逆过程。仅在一些区域中需要更详细的描述,如下文所述。

[0054] 作为其操作的一部分,源编码器(430)可执行运动补偿预测编码,参考视频序列中一个或多个先前编码的帧(称为“参考帧”)对输入帧进行预测性编码。以此方式,编码引擎(432)对输入帧的像素块和参考帧的像素块之间的差异进行编码,参考帧可被选择作为输入帧的预测参考。

[0055] 本地视频解码器(433)可基于源编码器(430)创建的符号,对指定的参考帧的编码视频数据进行解码。有利地,编码引擎(432)的操作可以是有损过程。当编码视频数据在视频解码器(图4未示出)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)复制由视频解码器对参考帧执行的解码过程,且可使重建的参考帧存储在参考图像存储器(434)中。以此方式,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)。

[0056] 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)可在参考图像存储器(434)中搜索适合作为新图像的预测参考的采样数据(作为候选参考像素块)或一些元数据,如参考图像运动矢量、块的形状等。预测器(435)可对采样块逐像素块进行处理,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,输入图像可具有从存储于参考图像存储器(434)中存储的多个参考图像取得的预测参考。

[0057] 控制器(450)可管理视频编码器(430)的编码操作,包括,例如,设置用于编码视频数据的参数和子组参数。

[0058] 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。熵编码器根据例如哈夫曼编码、可变长度编码、算术编码等本领域技术人员已知的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编码视频序列。

[0059] 传输器(440)可缓冲由熵编码器(445)创建的编码视频序列,从而为通过通信信道(460)进行传输做准备,该通信信道(460)可以是通向存储编码视频数据的存储设备的硬件/软件链路。传输器(440)可将来自视频编码器(430)的编码视频数据与将要传输的其它

数据合并,其它数据可以是,例如,编码音频数据和/或辅助数据流(未示出来源)。

[0060] 控制器(450)可管理视频编码器(203)的操作。在编码期间,控制器(450)可以为每个编码图像分配编码图像类型,该编码图像类型能够影响可应用于相应图像的编码技术。例如,通常可将图像分配为内部图像(I图像)、预测性图像(P图像)或双向预测性图像(B图像)。

[0061] 内部图像(I图像)可以是不需要利用序列中其它任何帧作为预测源进行编码和解码的图像。一些视频编解码器允许不同类型的内部图像,包括,例如,独立解码器刷新(Independent Decoder Refresh, IDR)图像。本领域技术人员知晓I图像的变型及其相应应用和特征。

[0062] 预测性图像(P图像)是可使用帧内预测或帧间预测进行编码和解码的图像,使用最多一个运动矢量和参考索引预测每个块的采样值。

[0063] 双向预测性图像(B图像)是可使用帧内预测或帧间预测进行编码和解码的图像,利用最多两个运动矢量和参考索引预测每个块的采样值。类似地,多个预测性图像可使用多于两个参考图像和相关元数据以重建单个块。

[0064] 源图像通常可在空间上细分为多个采样块(例如,4×4、8×8、4×8或16×16个采样的块)并逐块进行编码。这些块可参考其它(已编码的)块进行预测编码,其它块可以根据应用于块的相应图像的编码任务来确定。例如,I图像的块可进行非预测编码,或可参考同一图像的已编码的块来进行预测编码(空间预测或帧内预测)。P图像的像素块可参考一个先前编码的参考图像进行空间预测或时域预测进行预测编码。B图像的块可参考一个或两个先前编码的参考图像,通过空间预测或时域预测进行预测编码。

[0065] 视频编码器(203)可根据如ITU-T Rec. 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,编码视频数据可符合所使用的视频编码技术或标准指定的语法。

[0066] 在实施例中,传输器(440)可传输附加数据和已编码的视频。视频编码器(430)可将此类数据作为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、冗余图像和分片等其它形式的冗余数据、补充增强信息(Enhancement Information, SEI)消息、可视可用性信息(Visual Usability Information, VUI)参数集片段等。

[0067] 下面描述本申请实施例的一些方面,包括在视频编解码器技术或标准例如多用途视频编码(VVC)中实现的高级语法架构。

[0068] 该高级语法架构可包括已被证明可用的H. 264的NAL单元的概念,因为该概念已被至少一些系统规范(包括特定文件格式)采用。

[0069] 可选地,该高级语法架构可不包括(独立的,规则的)分片的概念。自2003年(H. 264第1版的发行日期)开始,视频编码的进展在于,由于图像内预测机制不断增多、效率不断提高,在很多情况下,基于分片的误差隐藏实际上已经变得不可能。同时,由于这些预测机制的存在,在一些情况下,从编码效率的角度来看,使用分片的代价已变得相当昂贵。因此,最近很少有实现方案使用分片来达到所需要的目的(MTU尺寸匹配)。取而代之的是,基本上所有要求低延迟抗差错的应用均使用基于图像的抗差错工具,如内部刷新、开放GOP、对基础层进行非均匀保护的可缩放性等。

[0070] 去除分片后,高级语法架构中,熵级别可独立解码(即,无解析依赖性)的最小VCL语法单元可以是,例如,图块或编码图像。

[0071] 图块的独立解码对于某些应用场景是有帮助的。例如,立方体图的场景中。从空间中的任何给定视点来看,不透明立方体至多有三个表面是同时可见的。相应地,针对给定视点的显示方案,组成立方体图的代码图像的假设6个方形图块中,只有3个需要被解码。为支持该功能,在高级语法架构中,至少对于需要独立图块的应用来说,独立图块可直接替代独立分片。换言之,按扫描顺序组织的分片可替换为H.263+AnnexK所称的矩形分片。运动约束图块集合也是高级语法架构的需求之一。

[0072] 图像内预测中断机制的一般概念是规范空间和实现空间中的一个补充。实施例中,高级语法架构可包括独立的标记,每个标记对应一个预测机制,这些标记用于管理给定图块的数据的预测输入,设置于图块头部或参数集中。因此,这种实现方式可能是更优、更彻底、更有弹性的解决方案。

[0073] 在应用高级语法架构的实施例中,可根据使用的概要文件(profile)来启用图块化(tiling)。例如,支持直接并行化的非常基础的图块化机制可包括在所有概要文件中。并且,更高级的技术可被指定仅用于特定的概要文件。例如,应用立方体图的360概要文件可允许为该应用设计的运动约束独立图块;即,6个图块可以特定方式排列,如 3×2 排列或交叉式排列。其它概要文件可用于其它投影格式。例如,二十面体投影可能需要更多的图块,或类似的可理想地处理该投影形状的预测中断机制。

[0074] 除前述特定应用驱动的需求之外,编码图像成为中断预测的最小单元。当编码图像是中断预测的最小单元时,所有的图像内预测机制不会被中断,仅图像间预测机制会被中断。例如,一些元数据和一些较旧的视频编码标准的运动补偿和图像间预测可被中断。为有效支持不具有分片/图块的编码图像,各实施例的高级语法架构可包括图像头部,用于携带H.264/H.265中设置在分片头部但属于整个图像的语法元素。语法元素之一可以是对图像参数集(PPS)的引用信息。与先前分片头部中提供的信息相同的是,图像头部仅与其关联的图像有关,而与后续图像不相关。换言之,图像头部的内容是临时性的,图像头部之间不存在预测(否则,基于图像的抗差错功能都会失效)。

[0075] 忽略抗差错方面,图像头部可携带在图像的第一个(或唯一的)图块或其自身的VCLNAL单元中。前者效率更高,后者的结构更清晰。

[0076] 实施例中,高级语法架构可在语法(单个NAL单元)、功能性以及持续性范围上包括先前体系结构中提供的图像参数集(PictureParameterSet,PPS)和序列参数集(SequenceParameterSet,SPS)。

[0077] 在SPS之上,高级语法架构可包括解码器参数集(DecoderParameterSet,DPS),从而包括标记、子概要文件等。在视频流的存在期间内直到流结束NAL单元被接收之前,DPS的内容可保持不变。

[0078] 在利用高级语法架构的实施例中,实施例可能需要允许外部携带流结束NAL单元。例如,当SIP重新邀请改变了流的基本参数(已由解码系统确认)时,必须告知解码系统的解码器使用不同的DPS。如果只能通过比特流将该信息发送至解码器,该信息需要经过起始码防止竞争等处理,会产生一些不利效果。并且,实际应用时,一些超时情形下,通过比特流传输该信息可能是行不通的。

[0079] 在很多情况下,当通过报文网络传输编码图像时,编码图像的大小可能大于最大传输单元(MTU)的尺寸。引入不必要的预测中断不利于编码效率(毕竟,取消分片正是为了这一目的),所以最好不依赖图块。图块已经承担了并行化和应用专用图块化这两个可能相抵触的功能,基于这个原因,最好也不依赖图块。可以从是否需要在规范空间中的视频编解码器中使用分段机制两方面来讨论。如果需要在视频编解码器中使用分段机制,各实施例的高级语法架构可使用分段机制,例如,H.265的“非独立分片”。或者,也在高级语法架构的更高的层中提供分段。应当注意,多种H.26x视频的RTP有效载荷格式不仅依赖分片进行基于编码器的MTU尺寸匹配(用于网关的场景,其中网关不进行转码),还包括一些形式的分片机制。

[0080] 参考图5,考虑以上描述,各实施例的高级语法架构的语法层次结构(501)可基本如下:

[0081] 该语法层次结构可包括解码器参数集(DPS)(502),存在于会话期间。

[0082] 一些实施例中,该语法层次结构可包括视频参数集(VideoParameterSet,VPS)(503),用于将可缩放层结合在一起,其中视频参数集在各层的边界上的IDR处中断。

[0083] 该语法层次结构可包括序列参数集(SPS)(504),序列参数集与在H.265中的功能基本相似,其范围是编码视频序列。

[0084] 该语法层次结构可包括处于同一语义层级并具有相似范围的图像参数集(PictureParameterSet,PPS)(505)和图像头部(PictureHeader,PH)(506)。即,图像参数集(505)和图像头部(506)可覆盖全部编码图像,但在各编码图像中可不相同。图像参数集(505)可与在H.265中的功能基本相似,且其范围为一个编码图像。图像头部(506)可携带在各图像中可能不同的图像内恒定的数据,还可携带对图像参数集(505)的引用信息。

[0085] 一些实施例中,语法层次结构可包括图块头部(TileHeader)(507),用于需要图块的应用场景。

[0086] 一些实施例中,该语法层次结构可包括分片单元头部(Fragmentation Unit Header)(508),可以是,例如,非独立分片头部。

[0087] 该语法层次结构可包括编码图像的VCL数据,包括编码单元(CU)数据(509)。

[0088] 下面对上述各种语法元素和语法等级的交互进行详细描述。

[0089] [图像头部/图像参数集的交互]

[0090] 参考图6,下面结合本申请实施例描述图像头部(PH)(601)和图像参数集(PPS)(602)的交互,其中,图像头部(601)和图像参数集(602)在语法上处理同一语法等级的数据,即例如编码图像(509)。

[0091] 参考图6,PH(601)和PPS(602)可包括一些指定的语法元素。如图6所示,该实施例的PH(601)和PPS(602)均包括整整四个语法元素。然而,可以想到,PH(601)和PPS(602)可以,例如,具有任何尺寸,具有不同的尺寸,包括可选的元素等。这些语法元素之一,PH_pps_id(603),可以是在PH(601)中PPS(602)的引用信息。该语法元素的语义可与之前的视频编解码标准中的分片头部中pps_id的语义相似,即,用于激活PPS和任何下游更高级的参数集,如SPS、VPS、DPS,以及其它可能的情况。在PPS(602)中,PPS_pps_id(604)可为自引用信息,以及在接收时作为PPS的ID的标识。图像参数集标识是语法元素的一个例子。在一些情况下,对于每个遵守该语法架构的比特流,PH(601)和PPS(602)中相应的语法元素的值必须是

相同的。

[0092] 一些语法元素可仅出现在PH(601)中,而在PPS(602)中不出现。至少在一些情况下,这些语法元素可与其所在的PH(601)所属的图像有关,并可在不同图像中变化。因此,将这些语法元素加入参数集,如PPS(602),可能是低效的,因为基本上每次解码新的图像,都需要激活新的PPS(602)。这种语法元素的一个例子可以是当前处理的图像的标识,例如,时间参考信息、图像顺序计数,以及类似信息。例如,PH(601)可包括POC(605)。PPS(602)中的相应条目被标为pic_type(606),用于表示图像类型;这是仅出现在PPS(602)中而不出现在PH(601)中的语法元素的一个例子。相应地,激活PPS(602)的所有图像,均使用pic_type(606)的值。

[0093] 一些语法元素可仅出现在PPS(602)中,而不出现在PH(601)中。可以想到,可以或有可能与多个编码图像有关、但不用于整个编码视频序列的很多较大的语法元素可能属于这种类型。这种语法元素不太可能在不同的图像中变化时,可能出现在PPS(602)中,而不出现在PH(601)中,因此,激活另一个PPS(602)不会造成负担。例如,考虑复杂并且可能较大的数据集合,如缩放矩阵,该数据集合可允许一些(可以是所有)变换系数独立地选择量化器参数。该数据在给定图像类型,如I图像、P图像和B图像,的典型图像组(GOP)过程中不太可能改变。在PH中设置缩放列表信息的缺点在于,随每一个编码图像,都需要重新传输可能是完全相同的缩放列表,因为PH本质上是临时性的。

[0094] 然而,可能存在第三类语法元素。这些语法元素可具有相似的名称,例如pps_foo(608)和ph_foo(607),并且可能在PPS(602)和PH(601)中均有出现。根据语法元素的本质,这些语法元素之间的关系可定义在视频技术或标准中,并可随该类型中语法元素的不同而不同。

[0095] 例如,在相同或另一实施例中,一些情况下,PH(601)中的语法元素,如ph_foo(607),的值可覆盖PPS(602)中具有相似命名且语义相关的语法元素,如pps_foo(608),的值。

[0096] 在相同或另一实施例中,在其它一些情况下,PH(601)中另一语法元素(如ph_bar(609))的值使用PPS(602)中具有类似命名(此处为“bar”)并且语义相关的语法元素,如pps_bar(610),作为某种形式的预测信息。例如,在一些情况下,基于PH的语法元素(609)可加上或减去PPS(602)中具有类似命名并且语义相关的语法元素(610)。

[0097] [解码器参数集和比特流终止]

[0098] 解码器参数集(DPS)(502)与MPEG-2的序列头部具有众多相似性,但它是一个参数集。因此,不同于MPEG-2的序列头部,DPS(502)不是临时性的。应用于参数集的一些激活规则与如MPEG-2的序列头部等头部的激活规则的不同之处在于,即激活时间可分别与参数集或头部的解码时间不同。考虑到这一重要区别,SPS可与MPEG-2的GOP头部类似,并且DPS可与MPEG-2的序列头部类似。

[0099] DPS(502)可具有一定范围,该范围在H.265中称为视频比特流。视频比特流可包括多个编码视频序列(CVS)。H.264和H.265中有一些元素超出了给定CVS的范围,首先且最重要的是HRD参数。在规范空间中,对于CVS等级之上的参数,H.264和H.265将这些参数放入SPS并要求每个编码视频序列中激活的SPS之间的相关信息保持不变。本申请实施例中,DPS可将这些语法元素累积形成对于多个CVS是可知且保持不变的结构。

[0100] 先前未设想的一个方面是,如何从给定的点及时通知解码器准备接受要求不同DPS的参数集。该参数集可为,例如,其中需要保持不变的参数被改变的DPS或SPS。

[0101] 尽管H.264和H.265均包括流结束NAL单元(EndofStream,EOS),该NAL单元不可被频繁使用,至少部分原因是由于下文所述的架构上的不足。

[0102] 在H.264和H.265中,与一些其它NAL单元类型,如参数集,不同的是,EOS需要在编码视频比特流中传输,并且针对其传输位置也已定义了一些约束条件。例如,在H.264或H.265中,不能将EOS设置于编码图像的VCLNAL单元中。在实现中,需要编码器或(至少)知晓视频编码标准的高级语法约束的另一实体的协作,以在编码视频比特流的合适位置插入EOSNAL单元。在至少一些情形下,这种协作是不现实的。例如,参见图1的通信系统,假设接收终端脱离了网络覆盖范围,并且该终端正在接收编码图像的NAL单元过程中,则编码器由于与解码器连接中断,无法向解码器提供EOSNAL单元。因为在接收编码图像的NAL单元时连接中断,且无法将EOS设置于编码图像的NAL单元之间,因此接收器也无法对EOSNAL单元进行拼接。在实际应用中,接收终端可将其解码器复位至已知的近期状态,但该操作可耗时若干秒。尽管该耗时对上述场景是可接受的,但可能还存在其它一些场景,要求解码器具有更快和更合理定义的反应。

[0103] 在本申请公开的相同或另一实施例中,EOS可作为视频流的一部分而被接收(如H.264/H.265中)或作为带外信息被接收。

[0104] 参考图7,在相同或另一实施例中,当EOS在频带外被接收(701)和处理时,解码器可停止使用视频流的有效解码器参数集。停止使用有效解码器参数集(DPS)表示,在不产生语法冲突的情况下,启用与先前的有效DPS具有至少一个不同的值的另一个DPS。

[0105] 例如,停止使用有效的DPS可包括,解码器立即清除其缓存(702)并停止输出重建的图像(703)。在停止使用先前的有效DPS之后,解码器可准备接收新的视频流(704),该新的视频流的DPS内容可与先前DPS不同。然后解码器可通过(可选的解码以及)启用先前或新的DPS(705),从而开始解码新的视频流,其中新DPS可与旧DPS不同。新的DPS的接收和解码可发生在任何时间,甚至在带外接收EOS之前。通常,参数集接收和解码的时机与解码过程无关,只要参数集在被启用时在场即可。其后,可开始(706)根据新的DPS解码新的CVS。

[0106] 上述高级语法的技术可由利用计算机可读指令的计算机软件实现,该计算机软件可物理存储于一个或多个计算机可读介质中。例如,图8示出了适用于实现本申请公开的一些实施例的计算机系统(800)。

[0107] 计算机软件可利用任何合适的机器代码或计算机语言来编码,可采用汇编、编译、链接或类似机制生成指令代码。这些指令代码可由计算机中央处理单元(CPU)、图形处理单元(GPU)等直接执行或通过代码解释、微代码执行等操作来执行。

[0108] 这些指令可在多种类型的计算机或计算机组件中执行,包括,例如,个人计算机、平板电脑、服务器、智能电话、游戏设备、物联网设备等。

[0109] 图8所示的用于计算机系统(800)的组件本质上是示例性的,而非意图对实现本申请实施例的计算机软件的使用或功能范围做任何限制。各组件的配置方式也不应被解释为依赖或需要该非限制性实施例的计算机系统(800)示出的任一个组件或组件组合。

[0110] 计算机系统(800)可包括一些人机接口输入设备。该人机接口输入设备可响应一个或多个个人类用户的输入,例如触感输入(如:按键、滑动操作、数字手套移动)、音频输入

(如:语音、拍手声)、可视输入(如:姿态)、嗅觉输入(未示出)。人机接口设备还可用于采集一些媒体信息,这些媒体信息不一定与人类有意识的输入直接相关,如音频(如:讲话、音乐、环境声音)、图像(如:扫描图像、从静态图像照相机获取的照片图像)、视频(如二维视频、三维视频,包括立体视频)。

[0111] 人机接口输入设备可包括以下一个或多个(每种仅示出一个):键盘(801)、鼠标(802)、触控板(803)、触摸屏(810)、数字手套、操纵杆(805)、麦克风(806)、扫描仪(807)、照相机(808)。

[0112] 计算机系统(800)还可包括一些人机接口输出设备。人机接口输出设备可激发一个或多个人类用户的感受,通过例如,触感输出、声音、光,以及气味/味道。该人机接口输出设备可包括触感输出设备(例如触摸屏(810)、数字手套或操纵杆(805)的触感反馈,但也可存在不作为输入设备的触感反馈设备)。例如,这种设备可以是音频输出设备(如:扬声器(809)、耳机(未示出))、通过图形适配器850耦合到系统总线848的可视输出设备(如屏幕(810),包括CRT屏幕、LCD屏幕、等离子屏幕、OLED屏幕,每一个可以具有或不具有触摸屏输入能力,每一个可以具有或不具有触感反馈能力——其中一些能够输出二维可视输出,或通过如立体显示输出等手段输出三维以上的输出;虚拟现实眼镜(未示出)、全息照相显示器和烟柜(未示出)),以及打印机(未示出)。

[0113] 计算机系统(800)还可包括人类可访问存储设备及其相关介质,例如包括具有CD/DVD的CD/DVDROM/RW(820)的光介质或类似介质(821)、指状驱动器(822)、可移除硬盘驱动器或固态驱动器(823)、传统磁性介质如磁带和软盘(未示出)、基于专业ROM/ASIC/PLD的设备如安全加密狗(未示出)等。

[0114] 本领域技术人员还应理解,结合本申请公开的主题有关的术语“计算机可读介质”不包含传输介质、载波或其它暂时性信号。

[0115] 计算机系统(800)还可包括可连接一个或多个通信网络的接口。网络可以是,例如,无线网络、有线网络、光网络。网络还可是局域网、广域网、城域网、车连网和工业网络、实时网络、容迟网络等。网络的示例包括局域网(例如以太网、无线LAN、蜂窝网络,包括GSM、3G、4G、5G、LTE等)、TV有线或无线广域数字网络(包括有线TV、卫星TV和地面广播TV)、车连网和工业网络(包括CAN总线)等。一些网络通常需要外部网络接口适配器,连接至一些通用数据端口或外围设备总线(849)(如计算机系统(800)的USB端口);其它网络通常通过连接到如下所述的系统总线而集成在计算机系统(800)的核中而(例如,集成在PC计算机系统种的以太网接口或集成在智能电话计算机系统种的蜂窝网络接口)。利用任一种网络,计算机系统(800)可与其它实体通信。该通信可为单向通信,例如仅接收(例如广播TV)、单向仅发送(例如从CAN总线到一些CAN总线设备)。该通信也可以为双向通信,例如利用局域或广域数字网络与其它计算机系统的通信。上述网络855和网络接口854的每一个可以采用某些协议和协议栈。

[0116] 前述人机接口设备、人类可访问存储设备和网络接口可连接至计算机系统(800)的核(840)。

[0117] 核(840)可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,用于特定任务844的硬件加速器等。以上设备,以及只读存储器(ROM)(845)、随机存取存储器(846)、内部大容量存储如内

部非用户可访问的硬盘驱动、SSD等(847),可连接到系统总线(848)。在一些计算机系统中,系统总线(848)可以通过一个或多个物理插头的形式访问,从而能够利用额外的CPU、GPU等进行扩展。外围设备可直接连接到核的系统总线(848),也可连接到外围设备总线(849)。外围设备总线的体系结构包括PCI、USB等。

[0118] CPU(841)、GPU(842)、FPGA(843)和加速器(844)可执行一些指令,这些指令组合起来可构成前述计算机代码。该计算机代码可存储于ROM(845)或RAM(846)中。中间数据还可存储在RAM(846)中,永久数据可存储在,例如,内部大容量存储设备(847)中。可通过使用缓存设备实现对任何存储设备的快速存储和读取,该缓存设备可紧密关联到一个或多个CPU(841)、GPU(842)、大容量存储器(847)、ROM(845)、RAM(846)等。

[0119] 计算机可读媒体中可存储有计算机代码,用于执行各种计算机实现的操作。介质和计算机代码可为本申请公开的目的专门设计和建造,或者可为计算机软件领域的技术人员所公知类型的、可用的类型。

[0120] 作为示例而非限定性的、具有体系架构(800)的计算机系统,特别是核(840),的功能可通过处理器(包括CPU、GPU、FPGA、加速器等)执行由一个或多个有形的计算机可读介质体现的软件产生。该计算机可读介质可以是以上关联于用户可访问的大容量存储设备的介质,以及核(840)中非暂存性质的一些存储设备,如核内部的大容量存储设备(847)或ROM(845)。实现本申请公开各实施例的软件可存储于上述设备并由核(840)执行。根据需要,计算机可读介质可包括一个或多个存储设备或芯片。上述软件可使核(840),特别是其中的处理器(包括CPU、GPU、FPGA等)执行本文描述的过程,或上述过程的一些部分,包括定义存储于RAM(846)的数据结构,根据软件定义的过程修改该数据结构。作为附加方案或替代方案,该计算机系统的功能可由电路(例如:加速器(844))提供,该电路可以通过逻辑硬连线或其它方式实现。该电路运转时,可代替软件或与软件配合的方式,以执行本文描述的过程,或过程的一些部分。在适当的情况下,所说的软件可包括逻辑,反之亦然。在适当的情况下,所说的计算机可读介质可包括存储软件用于执行的(如集成电路(IC))、用于实现需要执行的逻辑的电路,或其组合。本申请公开包括任何合适的硬件和软件的组合。

[0121] 尽管本申请公开已描述一些非限定性实施例,但本申请公开的范围内仍然存在一些具有小改动的实施方式、变换顺序的实施方式、以及各种替代性的等同实施方式。因此应当理解,尽管本文未明确示出或描述,本领域技术人员能够提出多种能够实现本申请公开的原理的系统和方案,因此这些系统和方案仍然属于本申请公开的精神和范围内。

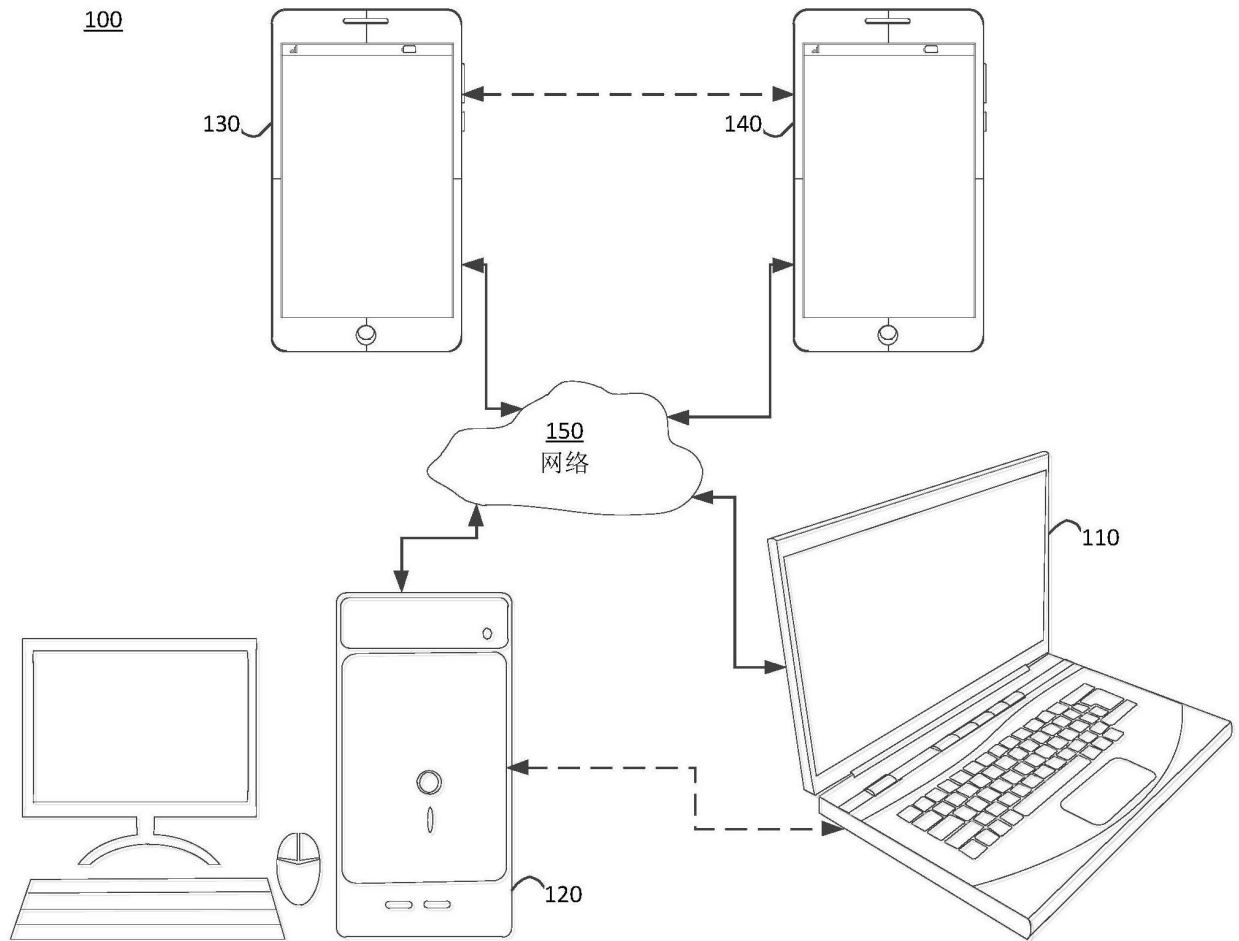


图1

流媒体系统 200

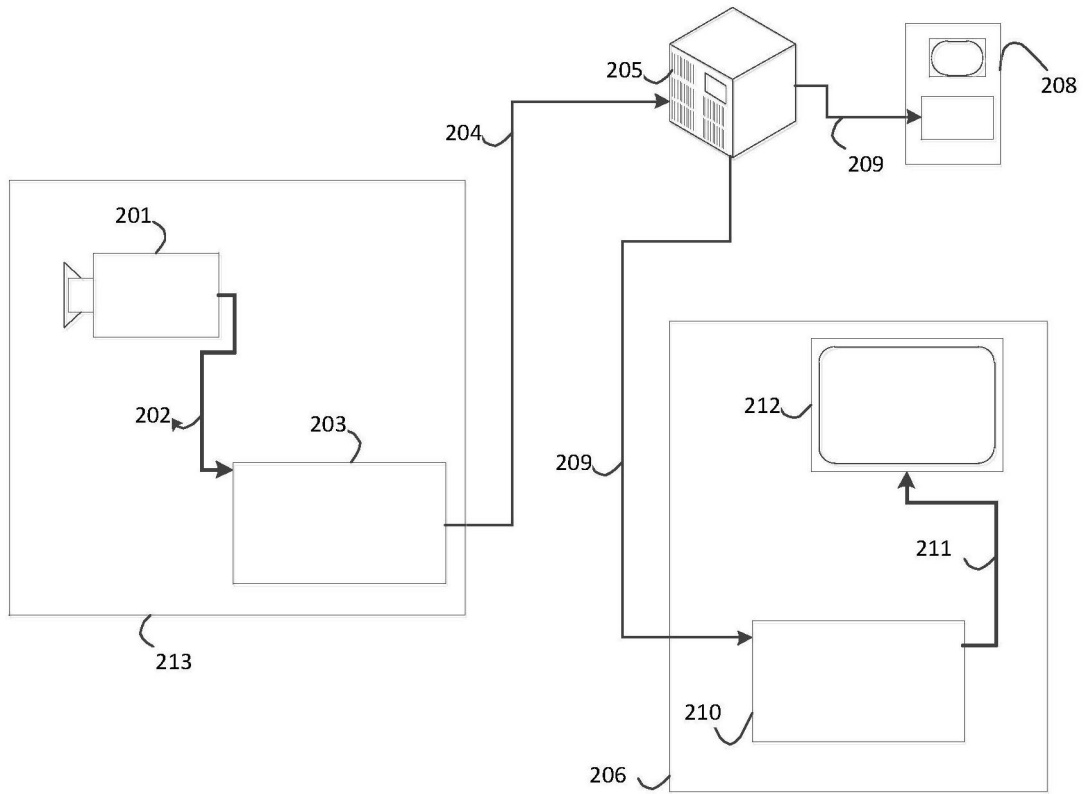


图2

解码器 210

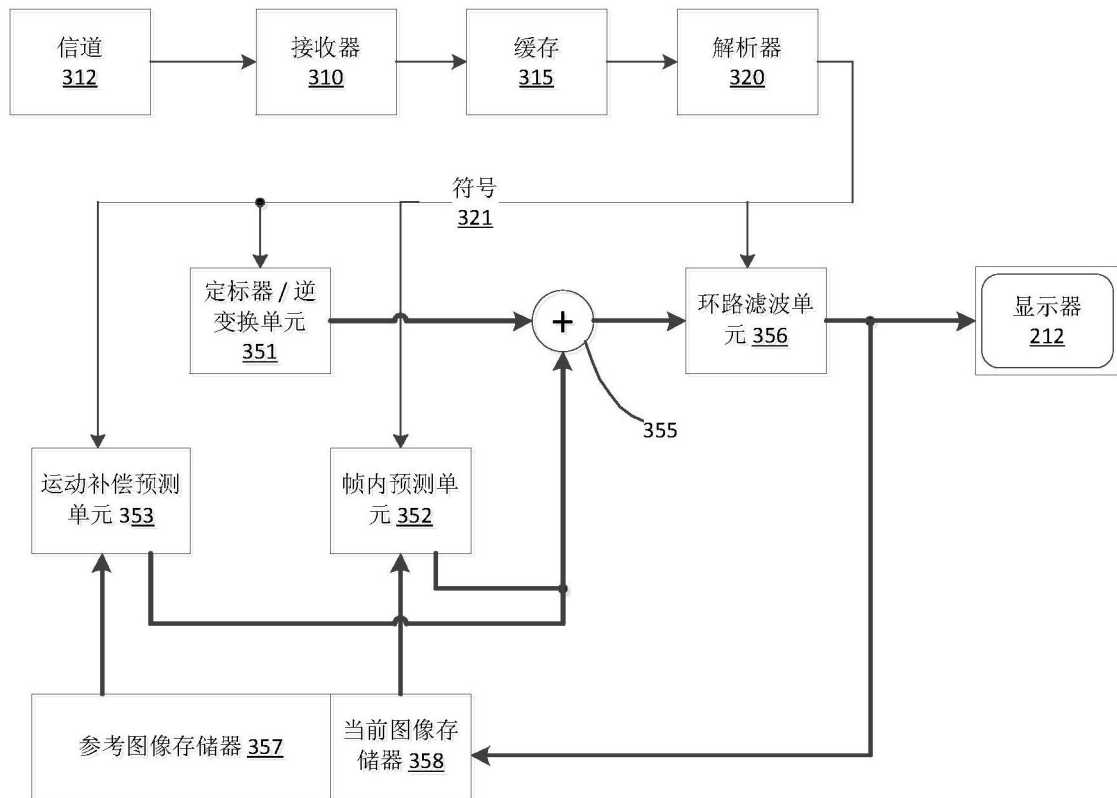


图3

编码器 203

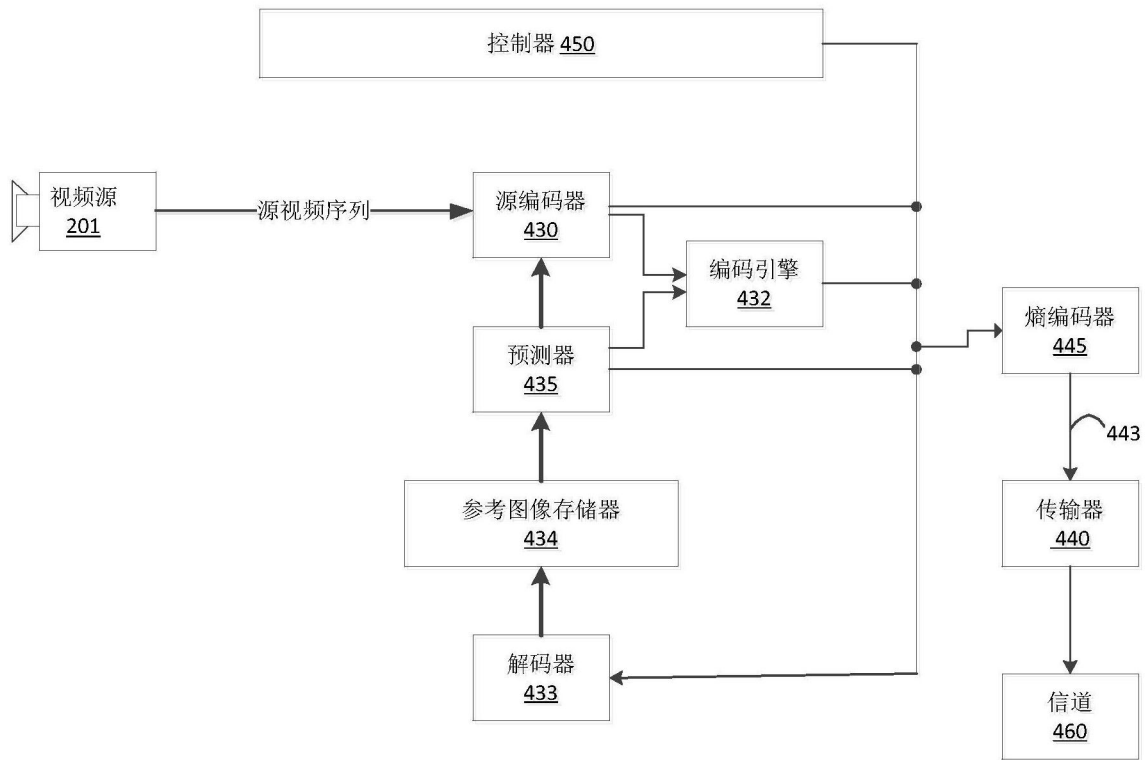


图4

501

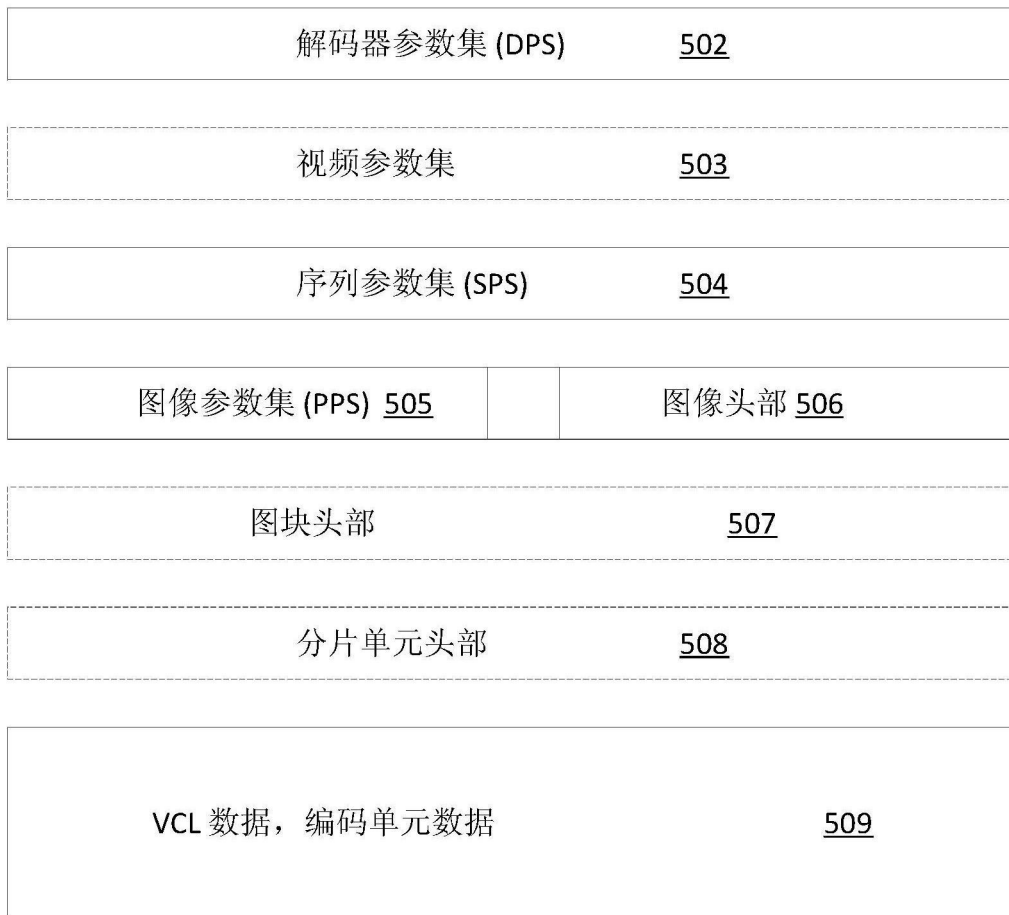


图5



图6

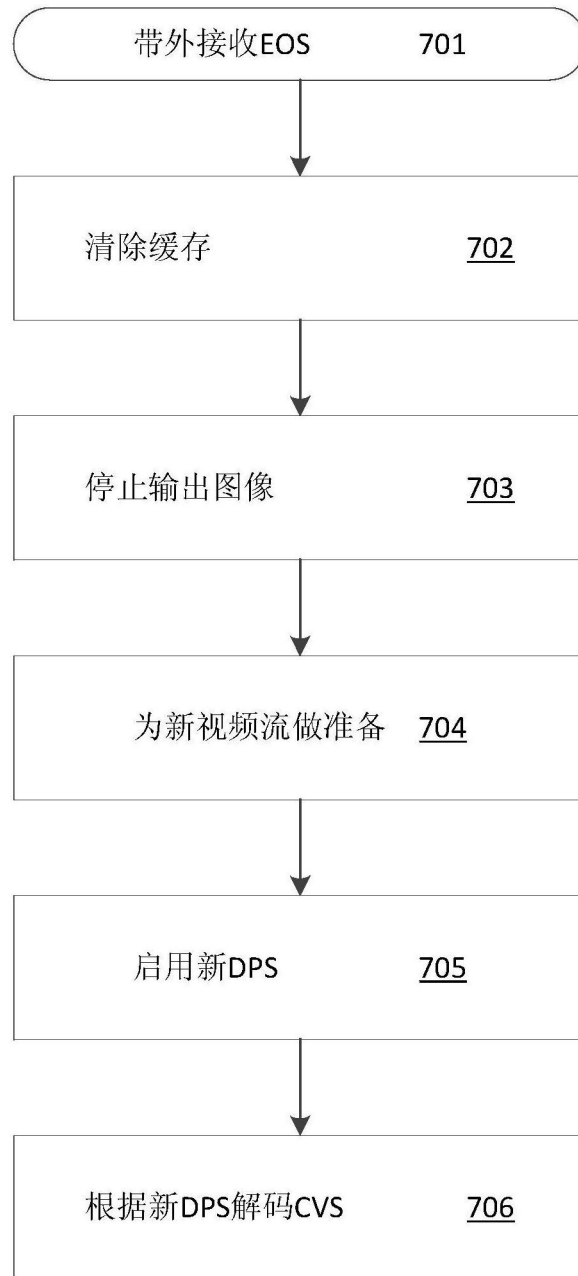


图7

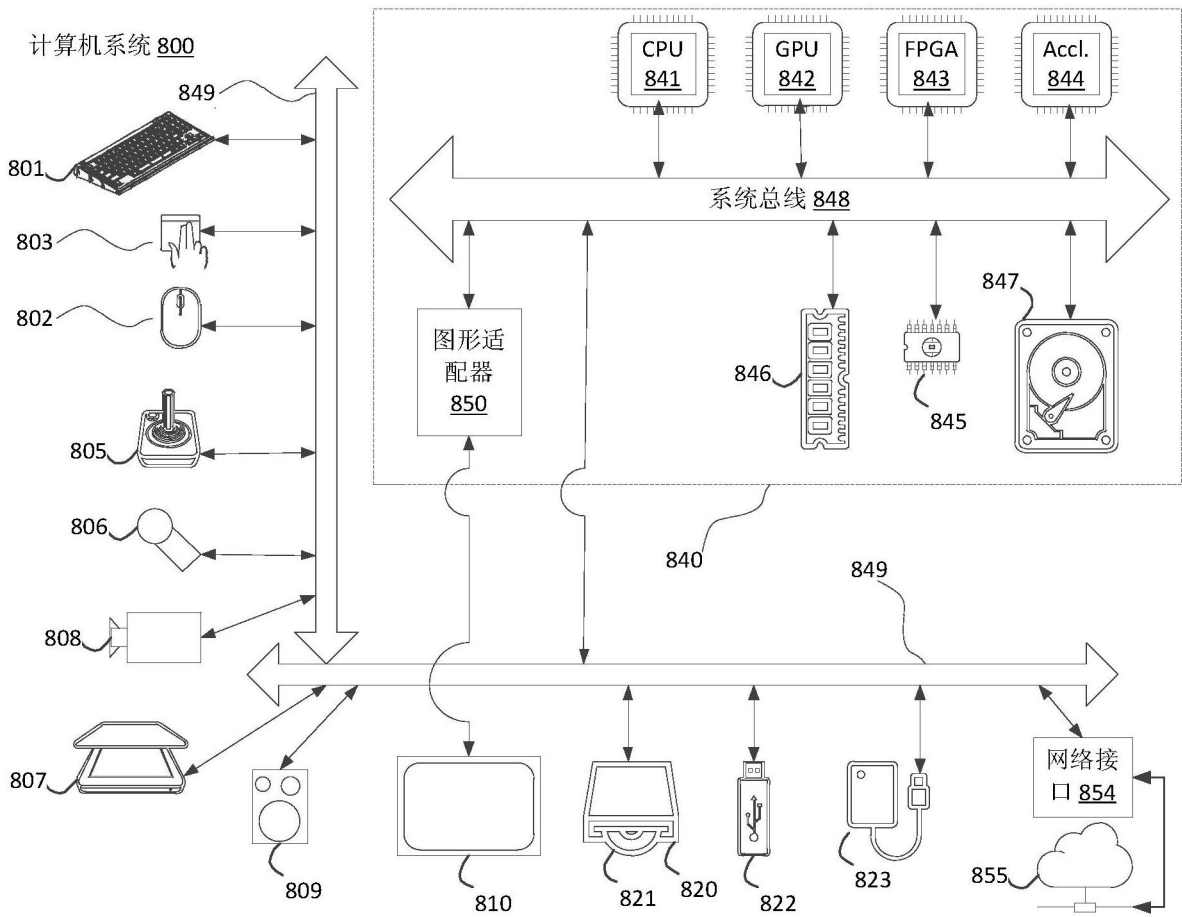


图8