



(12) 发明专利申请

(10) 申请公布号 CN 118714339 A

(43) 申请公布日 2024. 09. 27

(21) 申请号 202410557605.4

(22) 申请日 2019.12.19

(30) 优先权数据

62/783,152 2018.12.20 US

16/403,799 2019.05.06 US

(62) 分案原申请数据

201980064949.9 2019.12.19

(71) 申请人 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 崔秉斗 史蒂芬·文格尔 刘杉

(74) 专利代理机构 深圳市深佳知识产权代理事

务所(普通合伙) 44285

专利代理师 陈松浩

(51) Int.Cl.

H04N 19/44 (2014.01)

H04N 19/46 (2014.01)

H04N 19/70 (2014.01)

H04N 19/129 (2014.01)

H04N 19/17 (2014.01)

H04N 19/176 (2014.01)

权利要求书1页 说明书15页 附图9页

(54) 发明名称

一种用于视频解码的方法和相关装置

(57) 摘要

用于视频解码的方法和设备包括:对包括定长码字的高级语法结构中的携带图片片段的标识的二进制编码语法元素进行解码;以及重建图片片段。使得解码器或MANE可以通过访问定长码字的高级语法结构实现对图片片段的重建,解码器或MANE可以便捷的标识编码图片中的图块等,降低了处理复杂度。

| C | | 描述符 |
|--------------------|----|-----|
| nal_unit_header { | 全码 | 01 |
| forbidden_zero_bit | 全码 | 01 |
| nal_ref_idc | 全码 | 01 |
| nal_unit_type | 全码 | 01 |
| ... | | |
| } | | |

101

| 描述符 | |
|-----------------------|----|
| nal_unit_header { | 01 |
| forbidden_zero_bit | 01 |
| nal_unit_type | 01 |
| nal_ref_idc | 01 |
| nal_temporal_id_plus1 | 01 |
| } | |

102

1. 一种用于视频解码的方法,其特征在于,所述方法包括:

对包括定长码字的网络抽象层(NAL)单元报头中的二进制编码的编码单元(CU)地址进行解码,所述CU地址携带CU的标识;

基于对所述NAL单元报头中二进制编码的CU地址的解码,确定所述CU地址包含在待解码CU的列表中;

基于确定所述CU地址包含在待解码CU的列表中,重建所述图块;

其中,所述NAL单元报头仅包括nal_unit_type语法元素、nuh_cu_address语法元素和nuh_temporal_id_plus1语法元素;所述nal_unit_type语法元素表示NAL单元携带的数据的类型,所述nuh_cu_address语法元素表示CU地址,所述nuh_temporal_id_plus1语法元素表示NAL单元所属的编码图片的时间层。

2. 根据权利要求1所述的方法,其中,对于给定的帧,所述帧中的每个CU具有唯一的CU地址,并且所述唯一的CU地址是根据所述帧中的图片片段的扫描顺序来分配的。

3. 根据权利要求2所述的方法,其中,所述扫描顺序为从所述帧的左上角开始,按照从上至下、从右至左的扫描顺序,基于扫描顺序依次对扫描到的CU地址的值加一。

4. 根据权利要求1所述的方法,其中,所述CU地址在所述NAL单元报头中处于固定位置。

5. 根据权利要求1所述的方法,其中,所述CU地址的大小通过所述NAL单元报头中的高级语法结构语法元素的预定义值来确定。

6. 根据权利要求1所述的方法,其中,所述CU地址的大小通过参数集中的至少一个语法元素来确定,所述参数集对于所述NAL单元报头所属的图片是有效的。

7. 一种用于视频解码的装置,其特征在于,所述装置包括解码单元和重建单元:

所述解码单元用于对包括定长码字的网络抽象层(NAL)单元报头中的二进制编码的编码单元(CU)地址进行解码,所述CU地址携带CU的标识;

所述重建单元用于基于对所述NAL单元报头中二进制编码的CU地址的解码,确定所述CU地址包含在待解码CU的列表中;基于确定所述CU地址包含在待解码CU的列表中,重建所述图块;

其中,所述NAL单元报头仅包括nal_unit_type语法元素、nuh_cu_address语法元素和nuh_temporal_id_plus1语法元素;所述nal_unit_type语法元素表示NAL单元携带的数据的类型,所述nuh_cu_address语法元素表示CU地址,所述nuh_temporal_id_plus1语法元素表示NAL单元所属的编码图片的时间层。

8. 一种用于视频解码的设备,其特征在于,所述设备包括:

至少一个存储器,被配置成存储程序代码;以及

至少一个处理器,被配置成读取所述程序代码并按照所述程序代码的指示执行如权利要求1-6中任意一项所述的方法。

9. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质用于存储计算机程序,所述计算机程序用于被处理器执行以实现权利要求1-6中任意一项所述的方法。

一种用于视频解码的方法和相关装置

[0001] 相关申请的交叉引用

[0002] 本申请根据35U.S.C§119要求于2018年12月20日在美国专利商标局提交的美国临时申请第62/783,152号和于2019年5月6日在美国专利商标局提交的美国申请第16/403,799号的优先权,这两个申请在此通过引用以其整体并入本文中。

[0003] 本申请对申请号为201980064949.9,申请日为2019年12月19日,发明名称为“一种用于视频解码的方法和相关装置”的中国专利申请提出分案申请。

技术领域

[0004] 所公开的主题涉及视频编解码,并且更具体地涉及一种用于视频解码的方法和相关装置。

背景技术

[0005] 使用经运动补偿的帧间图片预测的视频编码和解码已知了数十年。未压缩的数字视频可以包括一系列图片,每个图片的空间尺寸为例如1920x1080亮度样本和相关联的色度样本。这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。未压缩的视频具有高的比特率要求。例如,每个样本8比特的1080p60 4:2:0视频(60Hz帧速率下的1920x1080亮度样本分辨率)需要接近1.5Gbit/s的带宽。一小时这样的类视频需要超过600GB的存储空间。

[0006] 视频编码和解码的一个目的可以通过压缩来减少输入视频信号中的冗余。压缩可以帮助降低前述带宽或存储空间要求,在一些情况下将前述带宽或存储空间要求降低两个数量级或更多。可以采用无损压缩和有损压缩二者以及它们的组合。无损压缩是指可以根据压缩后的原始信号重建原始信号的精确副本的技术。当使用有损压缩时,重建的信号可能与原始信号不同,但是原始信号与重建的信号之间的失真很小,足以使重建的信号对于预期的应用有用。在视频的情况下,广泛采用有损压缩。容忍的失真量取决于应用;例如,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。可达到的压缩比可以反映出:更高的可允许/可容忍的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来自几种广泛类别的技术,包括例如运动补偿、变换、量化和熵编码,下面将介绍其中的一些。

[0008] 将编码视频比特流划分为包以通过分组网络进行传输的概念已经使用了数十年。早期,大多数视频编码标准和技术已针对面向比特的传输和定义的比特流进行了优化。分组发生在例如以实时传输协议(Real-time Transport Protocol,RTP)有效载荷格式指定的系统层接口中。随着适于在因特网上大量使用视频的因特网连接的出现,视频编码标准通过视频编码层(video coding layer,VCL)和网络抽象层(network abstraction layer,NAL)的概念区分反映了该突出的使用情况。NAL单元在2003年引入在H.264中,以及由于此后仅作了一些修改就保留在某些视频编码标准和技术中。

[0009] 在许多情况下,NAL单元可以看作是解码器可以在其上作用的最小实体,而不必对

编码视频序列的所有先前NAL单元进行解码。就此而言,NAL单元通过诸如选择性转发单元(Selective Forwarding Unit,SFU)或多点控制单元(Multipoint Control Unit,MCU)的媒体感知网络元件(Media Aware Network Element,MANE)使某些错误恢复技术以及某些比特流操纵技术能够包括比特流修剪。

[0010] 在不同的接收器所需子集可能不同的场景下,MANE将为每个接收器量身定制不同的比特流。目前,这样定制比特流以及以参数集的形式保持状态等需要MANE处理复杂的变长编码切片报头,效率很低。

发明内容

[0011] 为此,本申请公开了用于视频解码的方法,实现了在视频编解码器中的包括定长码字的语法结构中包括图块标识的技术,以允许基于语法结构来改进图块的标识。

[0012] 一方面,本申请实施例提供了一种用于视频解码的方法,包括:对包括定长码字的高级语法结构中的二进制编码语法元素进行解码,所述语法元素携带图片片段的标识;以及重建图片片段。

[0013] 另一方面本申请实施例提供了一种用于视频解码的设备,其特征在于,所述设备包括:至少一个存储器,被配置成存储程序代码;以及至少一个处理器,被配置成读取所述程序代码并按照所述程序代码的指示执行以上方面所述的方法。

[0014] 另一方面,本申请实施例提供了一种计算机可读介质,所述计算机程序用于执行以上方面的方法。

[0015] 由此可见,对包括定长码字的高级语法结构中的二进制编码语法元素进行解码,所述语法元素携带图片片段的标识,使得解码器或MANE可以通过访问定长码字的高级语法结构实现对图片片段的重建,解码器或MANE可以便捷的标识编码图片中的图块等,降低了处理复杂度。

附图说明

[0016] 通过下面的详细描述和附图,所公开的主题的其他特征、性质和各种优点将更加明显,在附图中:

[0017] 图1是根据H.264和H.265的NAL单元报头的示意图示;

[0018] 图2是根据一个实施方式的通信系统的简化框图的示意图示;

[0019] 图3是根据另一个实施方式的通信系统的简化框图的示意图示;

[0020] 图4是根据一个实施方式的解码器的简化框图的示意图示;

[0021] 图5是根据一个实施方式的编码器的简化框图的示意图示;

[0022] 图6是根据一个实施方式的包括CU地址或图块ID语法元素的NAL单元报头的示意图示;

[0023] 图7是根据一个实施方式的图块布局的示意图示;

[0024] 图8是根据一个实施方式的NAL单元解码/转发的示意图示;

[0025] 图9是根据一个实施方式的计算机系统的示意图示。

具体实施方式

[0026] 图1描绘了根据H.264(101)和H.265(102)的NAL单元报头的语法图的相关部分,在两种情况下都没有对它们进行任何相应的扩展。在两种情况下,“forbidden_zero_bit”是用于在某些系统层环境中防止启动代码仿真的零比特。“nal_unit_type”语法元素是指NAL单元携带的数据的类型,其可以是例如某些切片类型、参数集类型、补充增强信息(Supplementary Enhancement Information, SEI)消息等中的一种。H.265NAL单元报头还包括“nuh_layer_id”和“nuh_temporal_id_plus1”,这两个指示NAL单元所属的编码图片的空间/SNR和时间层。

[0027] 可以观察到,NAL单元报头仅包括可容易解析的定长码字,这些码字相比于比特流中的其他数据诸如例如其他NAL单元报头、参数集等不具有任何解析依赖性。当NAL单元报头是NAL单元中的第一八位字节时,MANE可以容易地对它们进行提取,对它们进行解析,并对它们进行作用。相比之下,其他高级语法元素例如切片或图块报头不太容易被MANE访问,因为它们可能要求保存参数集上下文和/或处理变长或算术编码码点。然而,即使诸如图块群组报头的结构可以被设计成具有使报头容易被MANE访问的性质,可是现有的视频压缩技术和标准可能还没有这样做。

[0028] 还可以观察到,如图1中示出的NAL单元报头不包括可以将NAL单元与编码图片的片段(例如切片、图块或比特流的表示编码图片的空间区域的类似部分)相关联的信息。在相关技术中,这样的信息在某些情况下以宏块或CU地址的形式存在于切片报头中。在一些情况下,该地址是整数n,其指示在从图片的左上角开始计数时,片段、切片、图块按扫描顺序从第n宏块/CU开始。因此,n可以取决于图片大小和宏块/CU大小二者,并且对于小的图片大小可以较小(例如,适合于8位二进制代码),或者对于大的图片大小可以较大(例如32400,需要16位二进制代码);在两种情况下,假设宏块/CU大小为16x16个样本。

[0029] 从历史上看,诸如图块或切片的图片片段主要用于便于比特流划分以匹配最大传输单元大小的约束和并行化。在两种情况下,通常不需要在MANE、SFU或类似设备中标识图块或切片。解码器可以结合从参数集的解码获得的状态来从相对复杂的切片报头和/或类似信息中获得相关信息。

[0030] 然而,最近,图片片段尤其是图块(以及图块群组,它们是图块按扫描顺序、矩形顺序或任何其他合适的顺序的集合)已经用于诸如收集表示合成的360度投影中的某些视图的CU以及其他应用的目的。在那些应用中的一些应用中,在对于应用而言不需要时,MANE和SFU可以有利地从编码图片中删除某些图块或其他片段。例如,当使用立方体投影时,从外部视点渲染场景需要六个立方体表面中的最多三个。将表示其余的最少三个表面的片段和CU传输至端点会浪费资源。然而,在发送方可以向MANE发送完整表示(包括立方体投影的所有六个表面)并且MANE仅将所需子集转发至可能的多个接收器,并且不同接收器所需子集可能不同的场景下,MANE将为每个接收器量身定制包含可能不同的立方体表面的比特流。目前,这样定制比特流以及以参数集的形式保持状态等需要MANE处理复杂的变长编码切片报头,如对切片报头进行解码所需的那样。

[0031] 图2示出了根据本公开内容的实施方式的通信系统(200)的简化框图。系统(200)可以包括经由网络(250)互连的至少两个终端(210至220)。对于数据的单向传输,第一终端(210)可以对本地位置处的视频数据进行编码,以经由网络(250)传输至另一终端(220)。第

二终端 (220) 可以从网络 (250) 接收另一终端的编码视频数据, 对编码数据进行解码并显示恢复的视频数据。单向数据传输在媒体服务应用等中会是常见的。

[0032] 图2示出了被提供以支持编码视频的双向传输的第二对终端 (230, 240), 该双向传输可以例如在视频会议期间发生。对于数据的双向传输, 每个终端 (230, 240) 可以对在本地位置处捕获的视频数据进行编码, 以经由网络 (250) 传输至另一终端。每个终端 (230, 240) 还可以接收由另一终端发送的编码视频数据, 可以对编码数据进行解码并且可以在本地显示设备上显示恢复的视频数据。

[0033] 在图2中, 终端 (210至240) 可以被示为服务器、个人计算机和智能电话, 但是本公开内容的原理可以不限于此。本公开内容的实施方式适用于膝上型计算机、平板电脑、媒体播放器和/或专用视频会议设备。网络 (250) 表示在终端 (210至240) 之间传送编码视频数据的任意数量的网络, 包括例如有线和/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。代表性网络包括电信网络、局域网、广域网和/或因特网。出于本论述的目的, 除非下面在本文中说明, 否则网络 (250) 的架构和拓扑对于本公开内容的操作来说可能是不重要的。

[0034] 作为所公开主题的应用的示例, 图3示出了视频编码器和解码器在流媒体环境中的布局。所公开的主题可以等同地适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、记忆棒等的数字介质上存储压缩视频等等。

[0035] 流式传输系统可以包括捕获子系统 (313), 该捕获子系统可以包括视频源 (301) 例如数码相机, 视频源 (301) 创建例如未压缩的视频样本流 (302)。该样本流 (302) (被描绘为粗线以强调在与编码视频比特流相比时高数据量) 可以由耦接至相机 (301) 的编码器 (303) 来处理。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地描述的所公开主题的各方面。编码视频比特流 (304) (被描绘为细线以强调在与样本流相比时较低数据量) 可以被存储在流式传输服务器 (305) 上以供将来使用。一个或更多个流式传输客户端 (306, 308) 可以访问流式传输服务器 (305) 以检索编码视频比特流 (304) 的副本 (307, 309)。客户端 (306) 可以包括视频解码器 (310), 该视频解码器对编码视频比特流的传入副本 (307) 进行解码并且创建可以在显示器 (312) 或其他呈现设备 (未描绘) 上呈现的传出视频样本流 (311)。在一些流式传输系统中, 可以根据某些视频编码/压缩标准来对视频比特流 (304, 307, 309) 进行编码。这些标准的示例包括ITU-T H.265建议书。正在开发的视频编码标准被非正式地称为通用视频编码或VVC。所公开的主题可以在VVC的背景下使用。

[0036] 图4可以是根据本发明的实施方式的视频解码器 (310) 的功能框图。

[0037] 接收器 (410) 可以接收要由解码器 (310) 解码的一个或更多个编解码器视频序列; 在同一个实施方式或另一个实施方式中, 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 (412) 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/软件链路。接收器 (410) 可以接收编码视频数据以及其他数据例如可以被转发至它们各自的使用实体 (未描绘) 的编码音频数据和/或辅助数据流。接收器 (410) 可以将编码视频序列与其他数据分开。为了对抗网络抖动, 可以在接收器 (410) 与熵解码器/解析器 (420) (下文中称为“解析器”) 之间耦接缓冲存储器 (415)。当接收器 (410) 正在从具有足够带宽和可控性的存储/转发设备或从等时同步网络接收数据时, 可以不需要缓冲存储器 (415) 或者缓冲存储器 (415) 可以很小。为了在最优分组网络例如因特

网上使用,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。

[0038] 视频解码器(310)可以包括用于根据熵编码视频序列重建符号(421)的解析器(420)。这些符号的类别包括用于管理解码器(310)的操作的信息以及用于控制诸如显示器(312)的呈现设备的潜在信息,该呈现设备不是解码器的组成部分而是可以耦接至解码器,如图3中所示。用于呈现设备的控制信息可以是补充增强信息(SEI消息)或视频可用性信息(Video Usability Information,VUI)参数集片段(未描绘)的形式。解析器(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,并且可以遵循本领域技术人员公知的原理,包括变长编码、霍夫曼编码、具有或不具有上下文敏感性的算术编码等等。解析器(420)可以基于与群组相对应的至少一个参数,从编码视频序列中提取用于视频解码器中的像素子群组中的至少一个子群组的子群组参数集。子群组可以包括图片群组(Group of Pictures,GOP)、图片、图块、切片、宏块、编码单元(Coding Unit,CU)、块、变换单元(Transform Unit,TU)、预测单元(Prediction Unit,PU)等。熵解码器/解析器还可以从编码视频序列中提取诸如变换系数、量化器参数值、运动矢量等的信息。

[0039] 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,从而创建符号(421)。

[0040] 取决于编码视频图片或其部分的类型(例如:帧间图片和帧内图片、帧间块和帧内块)以及其他因素,符号(421)的重建可以涉及多个不同的单元。涉及哪些单元以及涉及哪些重建方式可以通过由解析器(420)从编码视频序列解析的子群组控制信息来控制。为了清楚起见,未描绘这种子群组控制信息在解析器(420)与下面的多个单元之间的流动。

[0041] 除了已经提到的功能块之外,解码器310可以在概念上细分为如下所述的多个功能单元。在商业约束下运行的实际实现方式中,这些单元中的许多单元彼此紧密交互并且可以至少部分地彼此集成。然而,出于描述所公开的主题的目的,概念上细分成下面的功能单元是适当的。

[0042] 第一单元是缩放器/逆变换单元(451)。缩放器/逆变换单元(451)从解析器(420)接收作为符号(421)的量化变换系数以及控制信息,包括要使用哪种变换方式、块大小、量化因子、量化缩放矩阵等。缩放器/逆变换单元(451)可以输出包括样本值的块,样本值可以输入到聚合器(455)中。

[0043] 在一些情况下,缩放器/逆变换单元(451)的输出样本可以属于帧内编码的块;即:不使用来自先前重建的图片的预测性信息,但是可以使用来自当前图片的先前重建部分的预测性信息的块。这样的预测性信息可以由帧内预测单元(452)提供(例如帧内图片预测单元)。在一些情况下,帧内图片预测单元使用从当前(部分重建)图片458获取的周围已经重建的信息来生成大小和形状与正在重建的块相同的块。在一些情况下,聚合器(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缩放器/逆变换单元(451)提供的输出样本信息。

[0044] 在其他情况下,缩放器/逆变换单元(451)的输出样本可以属于帧间编码且潜在运动补偿的块。在这种情况下,运动补偿预测单元(453)可以访问参考图片缓冲器(457)以获取用于预测的样本。在根据属于块的符号(421)对获取的样本进行运动补偿之后,这些样本

可以由聚合器 (455) 添加到缩放器/逆变换单元的输出 (在这种情况下, 被称为残差样本或残差信号), 以生成输出样本信息。运动补偿单元从参考图片存储器内的地址提取预测样本可以受运动矢量控制, 运动矢量以符号 (421) 的形式可供运动补偿单元使用, 符号可以具有例如 X、Y 和参考图片分量。运动补偿还可以包括当使用子样本精确运动矢量时从参考图片存储器中提取的样本值的插值、运动矢量预测机制等。

[0045] 聚合器 (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内滤波器技术, 所述环路内滤波器技术由编码视频比特流中包括的且作为来自解析器 (420) 的符号 (421) 可用于环路滤波器单元 (456) 的参数来控制, 但是也可以响应于在对编码图片或编码视频序列的先前 (按解码顺序) 部分进行解码期间获得的元信息, 以及响应于先前重建且经过环路滤波的样本值。

[0046] 环路滤波器单元 (456) 的输出可以是样本流, 样本流可以被输出至呈现设备 (312) 以及被存储在参考图片存储器 (457) 中以供在将来的帧间图片预测中使用。

[0047] 某些编码图片一旦完全重建就可以用作参考图片以供在将来预测中使用。一旦编码图片被完全重建并且编码图片已被标识为参考图片 (例如, 通过解析器 (420)), 作为参考图片的当前图片 (458) 就可以成为参考图片缓冲器 (457) 的一部分, 并且可以在开始重构随后的编码图片之前重新分配新的当前图片存储器。

[0048] 视频解码器 420 可以根据可以记录在诸如 ITU-T H.265 建议书的标准中的预定视频压缩技术执行解码操作。如在视频压缩技术文档或标准中指定的且明确地在其中的配置文件文档中指定的, 在编码视频序列遵循视频压缩技术或标准的语法的意义上, 编码视频序列可以符合由正使用的视频压缩技术或标准指定的语法。对于合规性, 还要求编码视频序列的复杂度在视频压缩技术或标准的层级限定的范围内。在一些情况下, 层级会限制最大图片大小、最大帧速率、最大重建样本率 (以例如每秒兆个样本为单位进行测量)、最大参考图片大小等。在一些情况下, 由层级设置的限制可以通过假设参考解码器 (Hypothetical Reference Decoder, HRD) 规范以及在编码视频序列中用信号表示的 HRD 缓冲器管理的元数据来进一步限制。

[0049] 在实施方式中, 接收器 (410) 可以连同编码视频一起接收附加 (冗余) 数据。附加数据可以被包括为编码视频序列的一部分。视频解码器 (420) 可以使用附加数据来适当地对数据进行解码以及/或者更准确地对原始视频数据进行重建。附加数据可以呈例如时间、空间或 SNR 增强层、冗余切片、冗余图片、前向纠错码等的形式。

[0050] 图 5 可以是根据本公开内容的实施方式的视频编码器 (303) 的功能框图。

[0051] 编码器 (303) 可以从视频源 (301) (其不是编码器的一部分) 接收视频样本, 该视频源可以捕获将由编码器 (303) 编码的视频图像。

[0052] 视频源 (301) 可以提供要由编码器 (303) 编码的呈数字视频样本流形式的源视频序列, 该数字视频样本流可以具有任何合适的位深度 (例如: 8 位、10 位、12 位、...)、任何色彩空间 (例如, BT.601 Y CrCb、RGB 等) 和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 (301) 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 (301) 可以是捕获本地图像信息作为视频序列的相机。可以将视频数据提供为多个单独图片, 在顺序观看时, 这些图片被赋予运动。图片本身可以被组织为空间像素阵列, 其中, 取决于所使用的采样结构、色彩空间等, 每个像素可以包括一个或更多个样

本。本领域技术人员可以容易地理解像素与样本之间的关系。下面的描述侧重于样本。

[0053] 根据实施方式,编码器(303)可以实时地或在应用所要求的任何其他时间约束下对源视频序列的图片进行编码并将其压缩为编码视频序列(543)。施行适当的编码速度是控制器(550)的一个功能。控制器控制如下所述的其他功能单元并且在功能上耦接至这些单元。为了清楚起见,未描绘耦接。由控制器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 值等)、图片大小、图片群组(GOP)布局、最大运动矢量搜索范围等。本领域技术人员可以容易地识别控制器(550)的其他功能,因为这些功能可能涉及针对特定系统设计而优化的视频编码器(303)。

[0054] 一些视频编码器以如下方式进行操作,本领域技术人员容易将该方式理解为“编码环路”。作为简化的描述,编码环路可以包括编码器(530)的编码部分(下文中称为“源编码器”)(负责基于要编码的输入图片和参考图片创建符号);以及嵌入在编码器(303)中的(本地)解码器(533),解码器(533)以类似于(远程)解码器创建样本数据的方式重建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,符号与编码视频比特流之间的任何压缩都是无损的)。该重建的样本流被输入至参考图片存储器(534)。由于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,因此参考图片缓冲器内容在本地编码器与远程编码器之间也是位精确的。换言之,编码器的预测部分“看到”的参考图片样本与解码器在解码期间使用预测时将“看到”的样本值完全相同。这种参考图片同步性的基本原理(以及如果例如由于信道错误而不能保持同步性,则导致漂移)是本领域技术人员公知的。

[0055] “本地”解码器(533)的操作可以与“远程”解码器(310)的操作相同,上面已经结合图4详细描述了“远程”解码器。然而,另外简要地参照图4,当符号可用并且由熵编码器(545)和解析器(420)能够无损地将符号编码/解码为编码视频序列时,解码器(310)的包括信道(412)、接收器(410)、缓冲器(415)和解析器(420)的熵解码部分可能不能完全在本地解码器(533)中实现。

[0056] 此时可以观察到,除了存在于解码器中的解析/熵解码之外的任何解码器技术也都必定以基本上相同的功能形式存在于相应的编码器中。由于这个原因,所公开的主题侧重于解码器操作。可以简化编码器技术的描述,因为编码器技术与全面描述和解码器技术互逆。仅在某些区域中需要更详细的描述并在下面提供该更详细的描述。

[0057] 作为其操作的一部分,源编码器(530)可以执行运动补偿预测性编码,运动补偿预测编码参考来自视频序列中的被指定为“参考帧”的一个或多个先前编码帧对输入帧进行预测性编码。以这种方式,编码引擎(532)对输入帧的像素块与参考帧的像素块之间的差异进行编码,该参考帧可以被选作输入帧的预测参考。

[0058] 本地视频解码器(533)可以基于由源编码器(530)创建的符号来对可以被指定为参考帧的帧的编码视频数据进行解码。编码引擎(532)的操作可以有利地是有损过程。当编码视频数据可以在视频解码器(图5中未示出)处被解码时,重建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)复制可以由视频解码器对参考帧执行的解码过程,并且可以使重建的参考帧被存储在参考图片存储器(534)中。以这种方式,编码器(303)可以在本地存储重建的参考帧的副本,这些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(不存在传输错误)。

[0059] 预测器 (535) 可以针对编码引擎 (532) 执行预测搜索。也就是说,对于要被编码的新帧,预测器 (535) 可以在参考图片存储器 (534) 中搜索可以作为新图片的适当预测参考的样本数据 (作为候选参考像素块) 或某些元数据,诸如参考图片运动矢量、块形状等。预测器 (535) 可以基于样本块逐像素块进行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参考图片存储器 (534) 中存储的多个参考图片中得出的预测参考。

[0060] 控制器 (550) 可以管理视频编码器 (530) 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组参数。

[0061] 所有前述功能单元的输出可以在熵编码器 (545) 中经受熵编码。熵编码器通过根据本领域技术人员已知的技术例如霍夫曼编码、变长编码、算术编码等对符号进行无损压缩来将由各种功能单元生成的符号转换为编码视频序列。

[0062] 传输器 (540) 可以缓冲由熵编码器 (545) 创建的编码视频序列,从而为通过通信信道 (560) 进行传输做准备,该通信信道可以是通向将存储编码视频数据的存储设备的硬件/软件链路。传输器 (540) 可以将来自视频编码器 (530) 的编码视频数据与要传输的其他数据例如编码音频数据和/或辅助数据流 (未示出的源) 合并。

[0063] 控制器 (550) 可以管理编码器 (303) 的操作。在编码期间,控制器 (550) 可以给每个编码图片分配某种编码图片类型,这会影响可以应用于相应图片的编码技术。例如,图片通常会被分配为以下帧类型之一:

[0064] 帧内图片 (I图片) 可以是在不使用序列中的任何其他帧作为预测源的情况下可以被编码和解码的图片。一些视频编解码器允许不同类型的帧内图片,包括例如独立解码器刷新图片。本领域技术人员了解I图片的这些变型以及它们各自的应用和特征。

[0065] 预测性图片 (P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一个运动矢量和参考索引来预测每个块的样本值。

[0066] 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超过两个的参考图片和相关联的元数据来重建单个块。

[0067] 通常,源图片可以在空间上细分为多个样本块 (例如,每个4x4、8x8、4x8或16x16样本的块) 并逐块被编码。这些块可以参考如由应用于块的相应图片的编码分配确定的其他 (已经编码的) 块进行预测性编码。例如,可以对I图片的块进行非预测性编码,或者可以参考相同图片的已经编码的块对它们进行预测性地编码 (空间预测或帧内预测)。可以对P图片的像素块进行非预测性编码,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图片的像素块进行编码。可以对B图片的块进行非预测性编码,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对B图片的块进行非预测性编码。

[0068] 视频编码器 (303) 可以根据预定视频编码技术或标准例如ITU-T H.265建议书执行编码操作。在其操作中,视频编码器 (303) 可以执行各种压缩操作,包括利用输入视频序列中的时间和空间冗余的预测性编码操作。因此,编码视频数据可以符合由使用的视频编码技术或标准指定的语法。

[0069] 在实施方式中,传输器 (540) 可以传输附加数据连同编码视频。视频编码器 (530) 可以包括这样的数据作为编码视频序列的一部分。附加数据可以包括时间/空间/SNR增强

层、其他形式的冗余数据例如冗余图片和切片、补充增强信息 (SEI) 消息、视觉可用性信息 (VUI) 参数集片段等。

[0070] 根据实施方式,将标识诸如图块、图块群组、切片、块群组 (GOB) 等的图像片段(下文中为图块)的信息放置在易于访问的高级语法结构或类似结构中,诸如放置在NAL单元报头 (NAL Unit Header, NUH) 中,这些结构包括定长码字并被设计成易于由MANE (下文中为NUH) 处理。

[0071] 标识图块的信息可以采用不同的形式。在设计该信息时,应牢记一些设计考虑因素。下面列出了这些设计考虑因素中的一些。

[0072] 在与例如传统视频编码技术或标准中的可能的切片数量相比时,给定图片中的图块数量可能会很小。例如,在H.264中,可能(对于某些图片大小)具有涵盖单个宏块的切片,从而允许存在与宏块一样多的切片。相比之下,当表示平铺的立方体映射时,六个图块可能就足够了,这与图片的分辨率无关。在许多实际情况下,可以安全地假定最大图块数量为64、128或256。

[0073] 图块布局可以是固定的,并且虽然视频编码技术本身可以允许图块布局从图片到图片的灵活性,但是系统标准或技术可以将该灵活性限制到在整个会话期间图块布局保持在相同的点。这可以允许例如在会话建立期间通过非视频比特流特定的方式使图块布局可供MANE使用。由此可以禁止MANE操作与视频编码中的参数集之间的不期望的上下文依赖性。

[0074] 至少在上述假设下,在与诸如H.264和H.265的相关技术相比时,可以显著简化用于标识由NAL单元携带的图块以允许MANE删除NAL单元的机制。例如,在H.264和H.265中,MANE必须标识正确的序列参数集,以了解切片报头中的切片/图块地址码字的长度。这样的长度信息在序列参数集中被编码为变长码字;因此,MANE至少遵循参数集的激活序列以标识当前活动的序列参数集,并且(由于参数集是解析独立的,因此可能不是以此顺序排列)对变长码字进行解码以标识切片报头中携带的二进制编码切片/图块地址的长度。然后,MANE对切片报头中的变长码字进行解码,以便获得起始宏块/CU地址。该信息与从参数集解码的图块布局相匹配,用以标识图块。

[0075] 在同一个实施方式或另一个实施方式中,图块的标识信息可以是图块的第一宏块/CU的地址。实际上,这样的机制会将起始地址从切片报头移动至NUH。虽然这样做可以是编解码器设计的最小更改方法,但是它具有显著增加NUH大小的缺点。然而,即使从编码效率的视点来看,NUH大小的增加也是可以容忍的,因为从切片/图块报头中也删除了相同量的位。

[0076] 如上面所指出的,对于小的图片大小和大的宏块/CU大小,宏块/CU地址可以相当小,或者对于小的CU大小和大的图片大小,宏块/CU地址可以相当大。出于该原因,H.265的SPS包含切片报头中携带的宏块/CU地址的长度的指示。在同一个实施方式或另一个实施方式中,可以为NAL单元报头保留该机制。然而,这样做会有两个缺点。首先,通过参数集值确定NAL单元报头中的语法元素的大小而建立的上下文依赖性会要求MANE跟随参数集激活,这会很繁琐。第二,至少到目前为止,NAL单元报头是八位字节对齐的,以简化MANE中的处理。在由参数集用信号表示的宏块/CU地址的大小与其余的NAL单元报头语法元素加起来不会达到能够被8整除的位数的这样的情况下,保持该八位字节对齐会要求填充浪费位。

[0077] 在同一个实施方式或另一个实施方式中,宏块/CU地址的大小(或NAL单元报头中的任何其他语法元素)可以由NAL单元报头中的其他字段确定。这种机制避免了参数集与NAL单元报头之间的上下文依赖性,并且在许多情况下会更可取。缺点可能是在NAL单元报头的其他字段中使用位或码点。下面提供了更多详细信息。

[0078] 然而,当不考虑传统意义上的切片,而仅考虑图块或图块群组或CU向比特流实体的类似CU分配机制时,可以使用更高级的选项。为了描述这些选项,应简要回顾术语切片和图块。切片可以是CU或宏块的集合,通常按扫描顺序排列,并且由以下两个因素标识:起始宏块/CU地址——通常被编码在切片报头中;以及切片的结尾——通常由新切片的开头标识(这又通过下一个切片报头的存在来指示)。虽然某些视频压缩技术和标准对切片的数量和布局施加了一些相对较小的限制,但在大多数情况下,切片布局可以基于编码图片的变化而变化,并且通常由诸如速率控制和MTU大小匹配的机制确定。

[0079] 另一方面,图块可以指代CU的典型的矩形形状,并且矩形的大小和形状(其中矩形图块和其他矩形图块组合起来构成图片)被编码在参数集中。换言之,图块布局在某种程度上是静态的,因为从一种图块布局到另一种图块布局的改变需要激活不同的参数集。此外,通过限制图块的数量可以有利于实现有效的硬件实现。结果可能是,在许多视频压缩技术和标准中,例如8位的相对短的定长二进制码字允许在实际使用中对所有图片大小的最大数量的图块进行寻址。因此,用于标识图块ID的定长码字可以用于标识NAL单元报头中的图块,从而避免了在标识图块的NAL单元报头码字与参数集之间的上下文依赖性和解析。类似地,图块群组ID的定长码字可以用于标识图块群组。当然,如果期望的话,支持NAL单元报头中的宏块/CU地址的变长码字的机制可以以类似的架构缺陷为代价同样地应用于图块ID码字。

[0080] 参照图6,介绍了关于NAL单元报头设计的一些实现选项。

[0081] NAL单元(601)可以是编码视频比特流的一部分。在一些情况下,NAL单元是八位字节对齐的,并且小于或等于数据网络的通用最大传输单元(Maximum Transfer Unit,MTU)大小。一种这样的常见MTU大小约为1500个八位字节,这是源自早期以太网技术的某些限制。NAL单元可以在其开始处包括NAL单元报头(602)。编码视频比特流内的NAL单元的成帧可以通过起始代码,通过与基本的面向包的传输网络的分组结构对齐等来进行。

[0082] 再次参照图6,还示出了示例性NAL单元报头(603)的语法图,类似于H.265中使用的语法图。所公开的主题同样可以与类似结构的NAL单元报头一起使用,例如H.264或VVC或包括定长码字的任何其他高级语法结构的NAL单元报头。在NAL单元报头(603)中,可以包括语法元素CU地址或图块ID(604)。该语法元素的长度可以是固定的,并且可以被选择为使得NAL单元报头继续是八位字节对齐的。语法元素(604)可以具有不仅可由视频编码器和解码器而且可由MANE容易处理的格式。通过示例而非作为限制的方式,CU地址或图块ID(604)可以由6位无符号整数表示,如描述符u(6)所表示的。在所示示例中,CU地址或图块ID占用H.265中用于layer_id的相同位。以NAL单元报头(605)以及CU地址或图块ID(606)示出了相似主题的不同表示。

[0083] 还示出了NAL单元(607),其保留了H.265NAL单元报头的字段。语法元素(608)被添加在例如NAL单元报头的结尾处。该语法元素的位置仅是示例性的;也可以将其插入NAL单元报头的其他语法元素的中间的某处。该语法元素可以具有固定或可变大小,并且当具有

可变大小,其大小可以通过上面提到的机制(例如,通过参数集语法元素,通过NAL单元类型等)中的任一种机制或任何其他合适的机制来确定。

[0084] 语法元素(608)可以携带任何形式的图片片段标识信息;例如,宏块/CU地址或图块标识,诸如图块编号或表示图块群组的编号。语法元素的编号范围可以被预先确定。在宏块/CU地址的情况下,编号范围可以从0到图片中宏块/CU的最大数量。对于图块ID,编号范围可以取决于图块的最大数量,其可以由本领域技术人员已知的机制来定义,诸如例如配置文件、层级、层、编码在参数集中的最大或实际图片大小等。如果语法元素存在于非图块/切片NAL单元(例如参数集NAL单元、SEI NAL单元或类似元素)中,则可以将语法元素的值限制为某个数字,例如0。可替代地,语法元素的存在可以在NAL单元类型上进行选择,因此,某些NAL单元类型中可能不存在语法元素。可替代地,在某些非图块/切片NAL单元类型的情况下,可以将上述语义之外的涵盖语义分配给语法元素。

[0085] 在同一个实施方式或另一个实施方式中,例如,可以如下标识图块ID。图7示出了空间域中的图片(701),其由通过粗体线指示的示例性图块布局(702)细分。呈现的图块布局可以由在例如H.265中可用的语法来表示,或者可以由表示图块布局(702)所需的更高级的语法来表示。图块布局中的每个图块可以具有图块ID,图块ID通过任何合适的编号机制但优选地通过对图块的扫描顺序编号来分配。在图7中,通过图块ID 1至ID 8示出了扫描顺序图块编号;例如,按扫描顺序的第二图块具有分配的图块ID 2(703)。

[0086] 编码器可以以类似于编写本领域技术人员已知的现有NAL单元报头语法的方式来编写NAL单元报头,该NAL单元报头包括涵盖如上所述填充的宏块/CU地址或图块ID的语法元素。

[0087] 解码器或MANE可以以本领域技术人员已知的方式,根据编码视频比特流对NAL单元报头进行解析——更准确地说,对构成NAL单元报头的语法元素进行解析,而不管携带宏块/CU地址或图块ID或其他形式的图块标识信息的语法元素的存在或不存在。然而,应当注意,在如上所述的一些情况下,语法元素可以在无需状态信息的情况下被编码,并且可以是可访问的熵编码格式,例如定长二进制代码。就此而言,超出语法元素tile_id本身的实际存在,根据所公开的主题对NAL单元报头进行解析可能不包括对解码器或MANE的额外繁重的操作。

[0088] 然而,根据所公开的主题,在与缺少所公开的主题所需的操作相比时,解码器或MANE可以毫不费力地标识编码图片中的图块。作为示例,假设已经通过外部非视频编码装置通知了解码器或MANE对于特定应用不需要重建特定图块。例如,考虑如图7中示出的场景,即村庄中的街道。假设街道已被观察相机捕获。考虑具有图块ID 2(703)的图块。该图块主要涵盖墙壁;假设观察系统的配置者认为该区域与观察无关。因此,虽然相机可能正在对所有图块进行编码,但对于该应用可能不需要具有ID 2的图块。因此,如果将相机创建的比特流通过一个或多个MANE路由到其最终目的地,并且一个MANE观察到带宽不足并必须从比特流中删除某些内容,则可能会有利地删除该图块,因为该图块可能对于该应用不需要。缺少所公开的主题,该技术至少需要将NAL单元(切片或图块)的有效载荷解析到所需的程度,以提取图块中第一宏块的宏块/CU地址,然后进行映射相对于图块布局的宏块/CU地址(如果正在使用图块)。取决于所使用的且如上所述的视频编码技术或标准,这可能需要在MANE中处理变长码字并保持参数集上下文;从实现方式和计算复杂度的视点来看,这二者

都是不期望的。替代地,在同一个实施方式或另一个实施方式中,MANE可以通过二进制编码的码字的NAL单元报头处理来获得标识哪个图块由NAL单元携带所需的所有信息。

[0089] 参照图8,解码器或MANE可以例如如下使用所公开的主题。

[0090] 解码器可以根据视频比特流对NAL单元报头进行解析,该NAL单元报头包括涵盖宏块/CU地址或图块ID的语法元素(801)。使用该信息,解码器或MANE可以标识图块ID(802)。可以直接对图块ID进行编码,或者解码器/MANE可以将关于图块布局的先验信息(例如通过解码参数集和随后的激活序列而建立的先验信息)与被编码在NAL单元报头中的宏块/CU地址相匹配。解码器可以将图块ID与需要分别通过解码器或MANE重建或转发的图块的列表相匹配(803)。如果存在匹配(804),则解码器可以重建或者MANE可以转发携带该图块的NAL单元。然而,如果不存在匹配(805),则解码器或MANE可能会静静地丢弃NAL单元。

[0091] 可以使用计算机可读指令将用于上述网络抽象单元报头中的图片参考的技术实现为计算机软件,并且将其物理地存储在一个或多个计算机可读介质中。例如,图9示出了适合于实现所公开的主题的某些实施方式的计算机系统900。

[0092] 可以使用任何合适的机器代码或计算机语言对计算机软件进行编码,机器代码或计算机语言可以经受汇编、编译、链接等机制以创建包括指令的代码,所述指令可以由计算机中央处理单元(Central Processing Unit,CPU)、图形处理单元(Graphics Processing Unit,GPU)等直接执行或通过解释、微代码执行等执行。

[0093] 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。

[0094] 图9中示出的用于计算机系统900的部件本质上是示例性的,并且不旨在对实现本公开内容的实施方式的计算机软件的使用范围或功能提出任何限制。部件的配置也不应被解释为具有与计算机系统900的示例性实施方式中示出的部件中的任何一个部件或部件的组合有关的任何依赖性要求。

[0095] 计算机系统900可以包括某些人机接口输入设备。这样的人机接口输入设备可以响应于一个或多个人类用户通过例如触觉输入(例如:击键、滑动、数据手套移动)、音频输入(例如:语音、拍手)、视觉输入(例如:手势)、嗅觉输入(未描绘)的输入。人机接口设备还可以用于捕获不一定与人的意识输入直接有关的某些媒体,例如音频(例如:语音、音乐、环境声音)、图像(例如:扫描图像、从静态图像相机获得的摄影图像)、视频(例如二维视频、包括立体视频的三维视频)。

[0096] 输入人机接口设备可以包括以下中的一个或多个(描绘的每种中的仅一个):键盘901、鼠标902、触控板903、触摸屏910、数据手套904(图中未示出)、操纵杆905、麦克风906、扫描仪907、相机908。

[0097] 计算机系统900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受。这样的人机接口输出设备可以包括触觉输出设备(例如,通过触摸屏910、数据手套904或操纵杆905的触觉反馈,但是也可以存在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909、耳机(未描绘))、视觉输出设备(例如包括CRT屏幕、LCD屏幕、等离子屏幕、OLED屏幕的屏幕910,每个屏幕具有或不具有触摸屏输入能力,每个屏幕具有或不具有触觉反馈能力,其中一些屏幕可能能够通过诸如立体图像输出、虚拟现实眼镜(未描绘)、全息显示器

和发烟器(未描绘)以及打印机(未描绘)的方式输出二维视觉输出或多于三维输出。

[0098] 计算机系统900还可以包括人类可访问的存储设备及其相关联的介质,例如包括带有CD/DVD等介质921的CD/DVD ROM/RW 920的光学介质、拇指驱动器922、可移动硬盘驱动器或固态驱动器923、诸如磁带和软盘(未描绘)的传统磁性介质、诸如安全加密狗(未描绘)的基于专用ROM/ASIC/PLD的设备等。

[0099] 本领域技术人员还应当理解,结合当前公开的主题使用的术语“计算机可读介质”不涵盖传输介质、载波或其他暂态信号。

[0100] 计算机系统900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的、有线的、光学的。网络还可以是本地的、广域的、城域的、车载和工业的、实时的、耐延迟的等。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电视有线或无线广域数字网络、包括CAN总线的车载和工业的等。某些网络通常需要附接至某些通用数据端口或外围总线(949)的外部网络接口适配器(诸如,例如计算机系统900的USB端口);其他网络通常通过附接至如下所述的系统总线而集成到计算机系统900的核中(例如,到PC计算机系统的以太网接口或到智能手机计算机系统的蜂窝网络接口)。使用这些网络中的任一个,计算机系统900可以与其他实体进行通信。这样的通信可以是单向的,仅接收的(例如,广播电视),单向仅发送的(例如,到某些CAN总线设备的CAN总线)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。可以在如上所述的这些网络和网络接口中的每一个上使用某些协议和协议栈。

[0101] 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至计算机系统900的核940。

[0102] 核940可以包括一个或多个中央处理单元(CPU)941、图形处理单元(GPU)942、现场可编程门区(Field Programmable Gate Area,FPGA)943形式的专用可编程处理单元、某些任务的硬件加速器944等。这些设备以及只读存储器(Read-only memory,ROM)945、随机存取存储器946、例如内部非用户可访问的硬盘驱动器、SSD等的内部大容量存储装置947可以通过系统总线948连接。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948,以通过附加CPU、GPU等实现扩展。外围设备可以直接附接至核的系统总线948,或者通过外围总线949附接至核的系统总线948。外围总线的架构包括PCI、USB等。

[0103] CPU 941、GPU 942、FPGA 943和加速器944可以执行某些指令,这些指令组合起来可以构成前述计算机代码。该计算机代码可以存储在ROM 945或RAM 946中。过渡数据也可以存储在RAM 946中,而永久性数据可以存储在例如内部大容量存储装置947中。可以通过使用高速缓冲存储器来实现到存储器设备中的任何存储器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 941、GPU 942、大容量存储装置947、ROM 945、RAM 946等紧密相关联。

[0104] 计算机可读介质可以在其上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出于本公开内容的目的而专门设计和构造的介质和计算机代码,或者介质和计算机代码可以是计算机软件领域的技术人员公知且可用的类型。

[0105] 作为示例但不作为限制,具有架构的计算机系统900,特别是核940可以由于处理器(包括CPU、GPU、FPGA、加速器等)执行在一种或更多种有形的计算机可读介质中实施的软

件而提供功能。这样的计算机可读介质可以是与以下各项相关联的介质：如上面所介绍的用户可访问的大容量存储、以及核940的具有非暂态性质的某些存储装置例如核内部大容量存储装置947或ROM 945。实现本公开内容的各种实施方式的软件可以存储在这样的设备中并由核940执行。根据特定需要，计算机可读介质可以包括一个或更多个存储器设备或芯片。软件可以使核940特别是其中的处理器（包括CPU、GPU、FPGA等）执行本文中所述的特定过程或特定过程的特定部分，包括根据软件定义的过程定义存储在RAM 946中的数据结构和修改这样的数据结构。另外地或者作为备选方案，计算机系统可以由于逻辑硬连线或以其他方式实施在电路（例如：加速器944）中而提供功能，该电路可以代替软件或与软件一起运行以执行本文中所述的特定过程或特定过程的特定部分。在适当的情况下，提及软件可以涵盖逻辑，反之提及逻辑也可以涵盖软件。在适当的情况下，提及计算机可读介质可以涵盖存储用于执行的软件的电路（例如集成电路（integrated circuit, IC））、实施用于执行的逻辑的电路或上述二者。本公开内容涵盖硬件和软件的任何合适的组合。

[0106] 基于上述实施例的说明，相应的，本申请实施例提供了一种用于视频解码的方法，所述方法包括：

[0107] 对包括定长码字的高级语法结构中的二进制编码语法元素进行解码，所述语法元素携带图片片段的标识；以及

[0108] 重建所述图片片段。

[0109] 在一种可能的实现方式中，所述图片片段是图块、图块群组或子图片中的一种。

[0110] 在一种可能的实现方式中，所述语法元素是所述图片片段中相应的宏块地址、第一宏块或编码单元的编码单元地址或第一编码单元的图块地址中的一个。

[0111] 在一种可能的实现方式中，所述语法元素是图块标识符。

[0112] 在一种可能的实现方式中，对于给定的图块布局，所述图块布局中的每个图块具有唯一的图块标识符，并且所述图块标识符是根据所述图块布局中的图片片段的扫描顺序来分配的。

[0113] 在一种可能的实现方式中，所述扫描顺序为从所述图块布局的图块的左上角开始，按照从上至下、从左至右的扫描顺序，基于扫描顺序依次对扫描到图块的图块标识符的值加一。

[0114] 在一种可能的实现方式中，所述语法元素在所述高级语法结构中处于固定位置。

[0115] 在一种可能的实现方式中，所述语法元素的大小通过所述高级语法结构中的高级语法结构语法元素的预定义值来确定。

[0116] 在一种可能的实现方式中，所述语法元素的大小通过参数集中的至少一个语法元素来确定，所述参数集对于所述高级语法结构所属的图片是有效的。

[0117] 在一种可能的实现方式中，所述重建所述图片片段包括：

[0118] 基于对所述语法元素进行解码，确定与所述语法元素相关联的图块标识符包括在待解码图块的列表中；以及

[0119] 基于确定所述图块标识符包括在待解码图块的列表中，重建所述图片片段。

[0120] 虽然本公开内容描述了几个示例性实施方式，但是存在落入本公开内容的范围内的变更、置换和各种替代等同物。因此，将认识到，本领域技术人员将能够设计出尽管未在本文中明确示出或描述但实施了本公开内容的原理并因此在本公开内容的精神和范围内

的许多系统和方法。

| | | |
|--|----------|------|
| <code>nal_unit(NumBytesInNALunit) {</code> | C | 描述符 |
| <code> forbidden_zero_bit</code> | 全部 | f(1) |
| <code> nal_ref_idc</code> | 全部 | u(2) |
| <code> nal_unit_type</code> | 全部 | u(5) |
| <code> ...</code> | | |
| <code>}</code> | | |

101

| | |
|--------------------------------------|------|
| <code>nal_unit_header() {</code> | 描述符 |
| <code> forbidden_zero_bit</code> | f(1) |
| <code> nal_unit_type</code> | u(6) |
| <code> nuh_layer_id</code> | u(6) |
| <code> nuh_temporal_id_plus1</code> | u(3) |
| <code>}</code> | |

102

图1

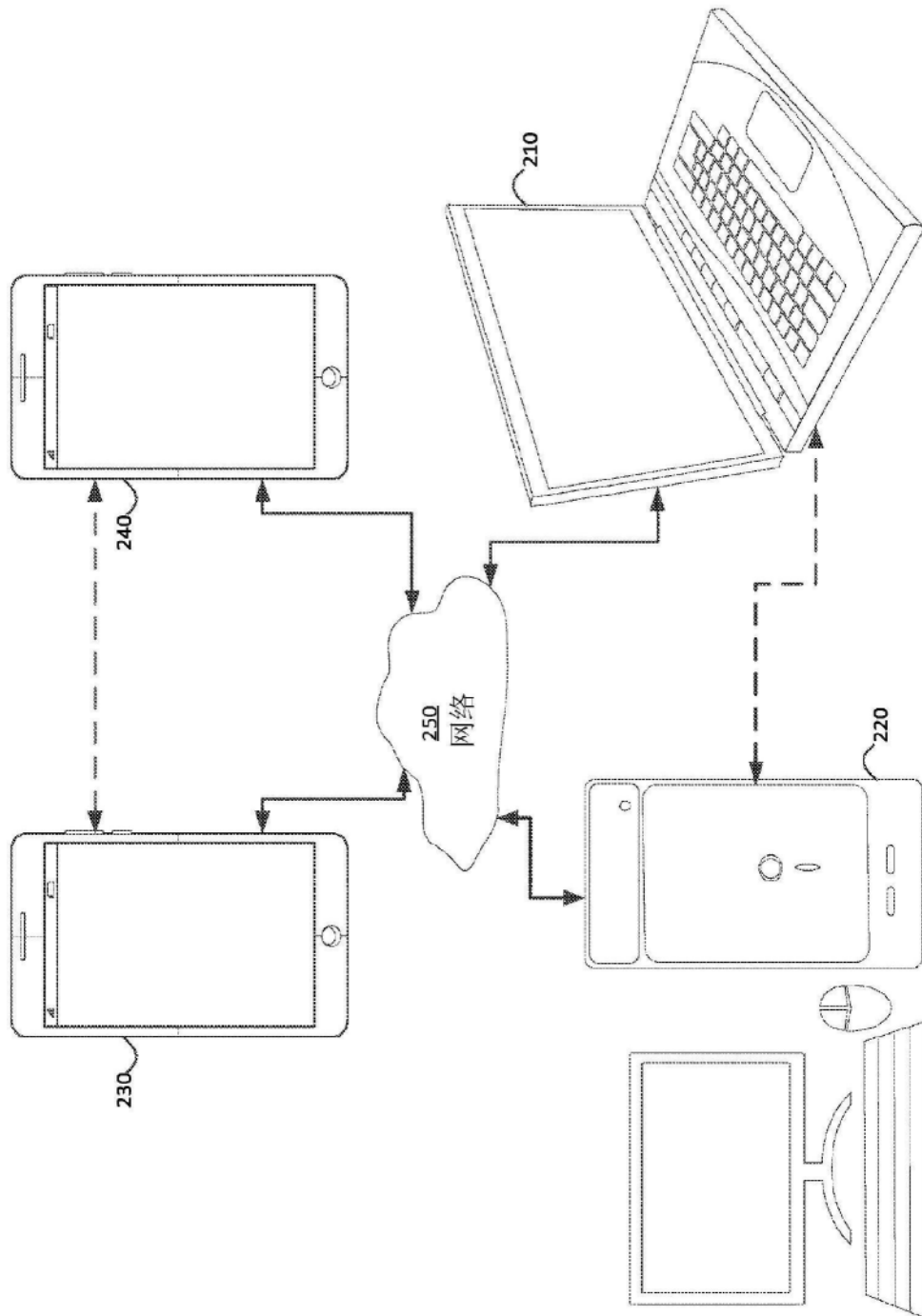


图2

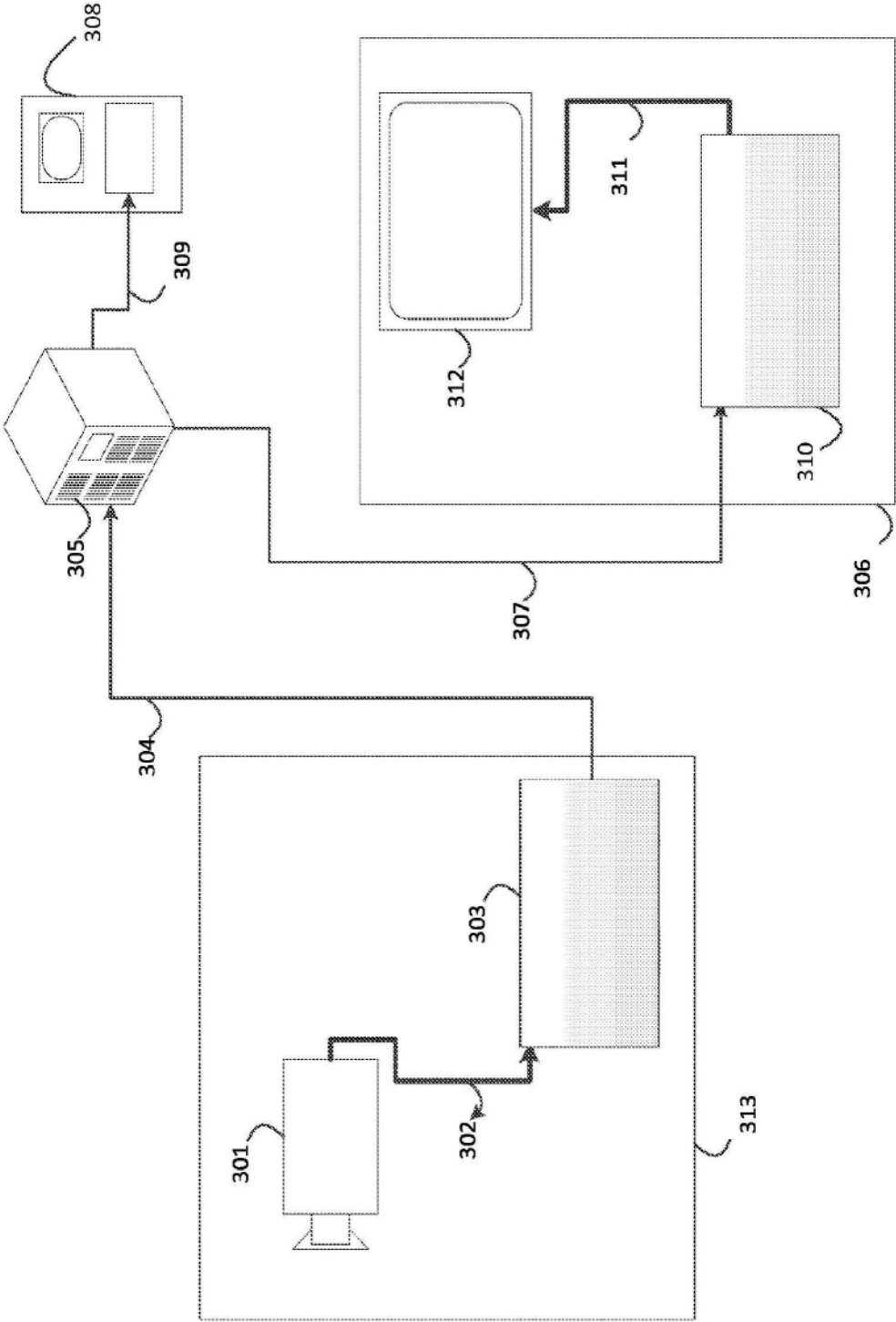


图3

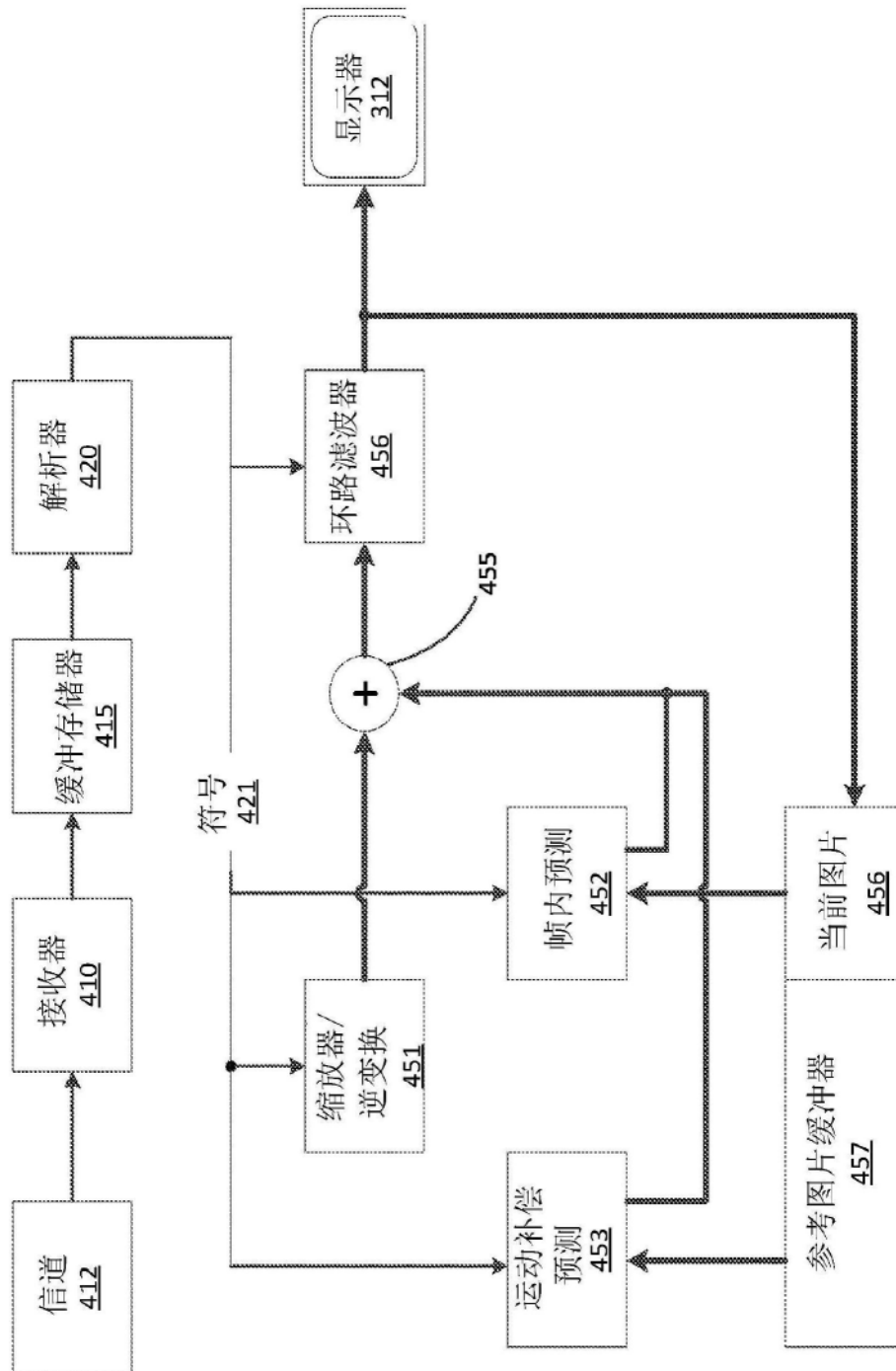


图4

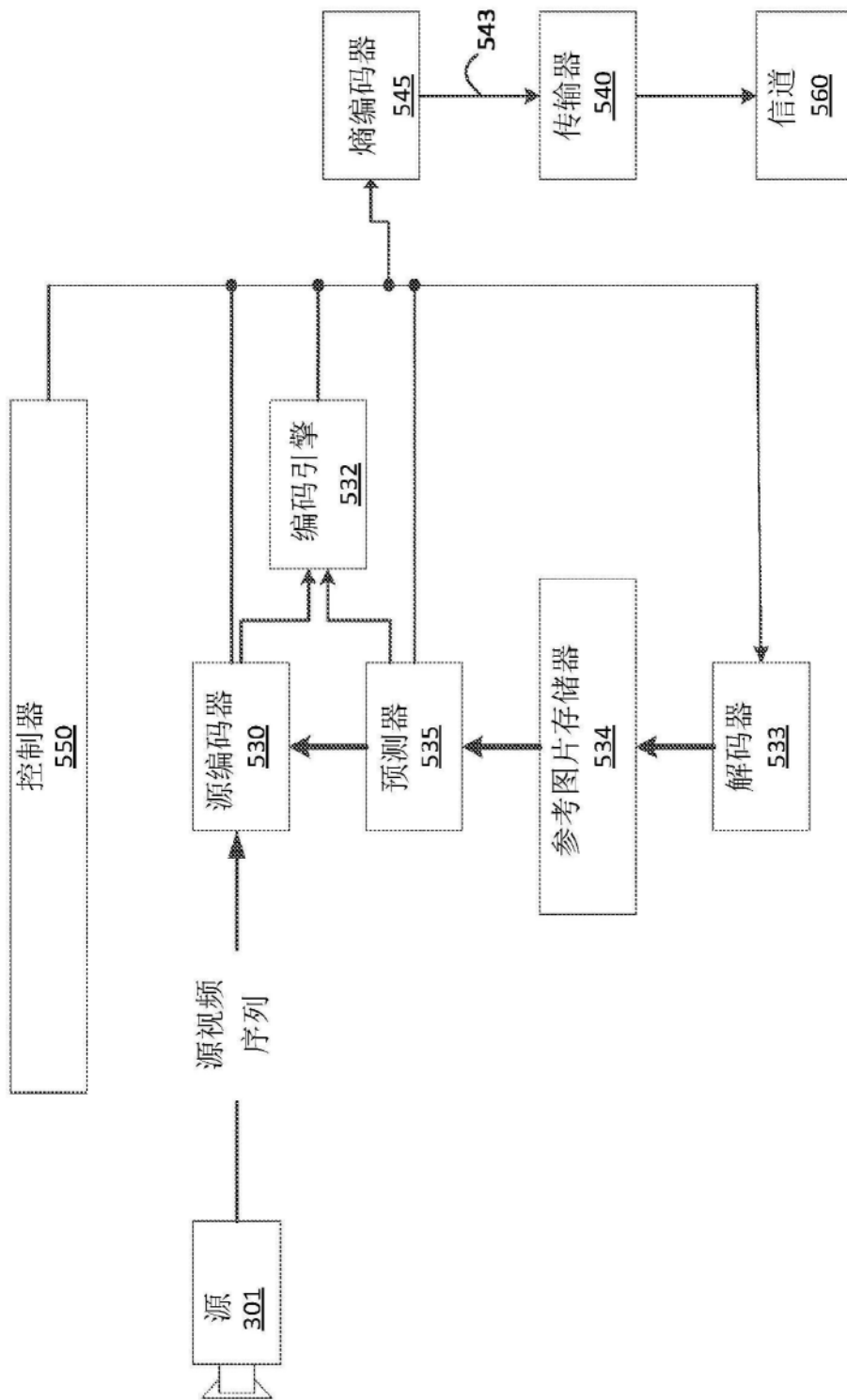


图5

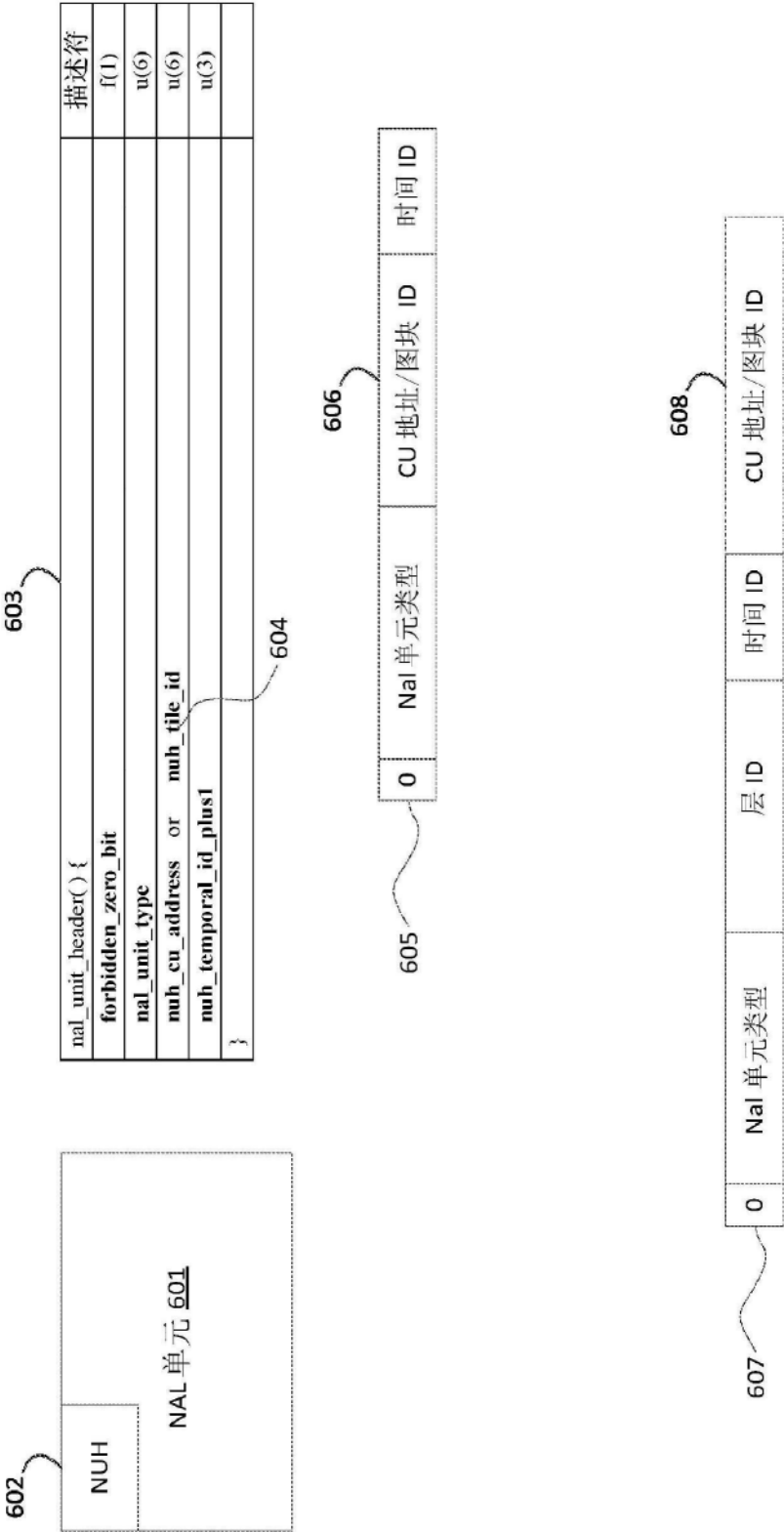


图6

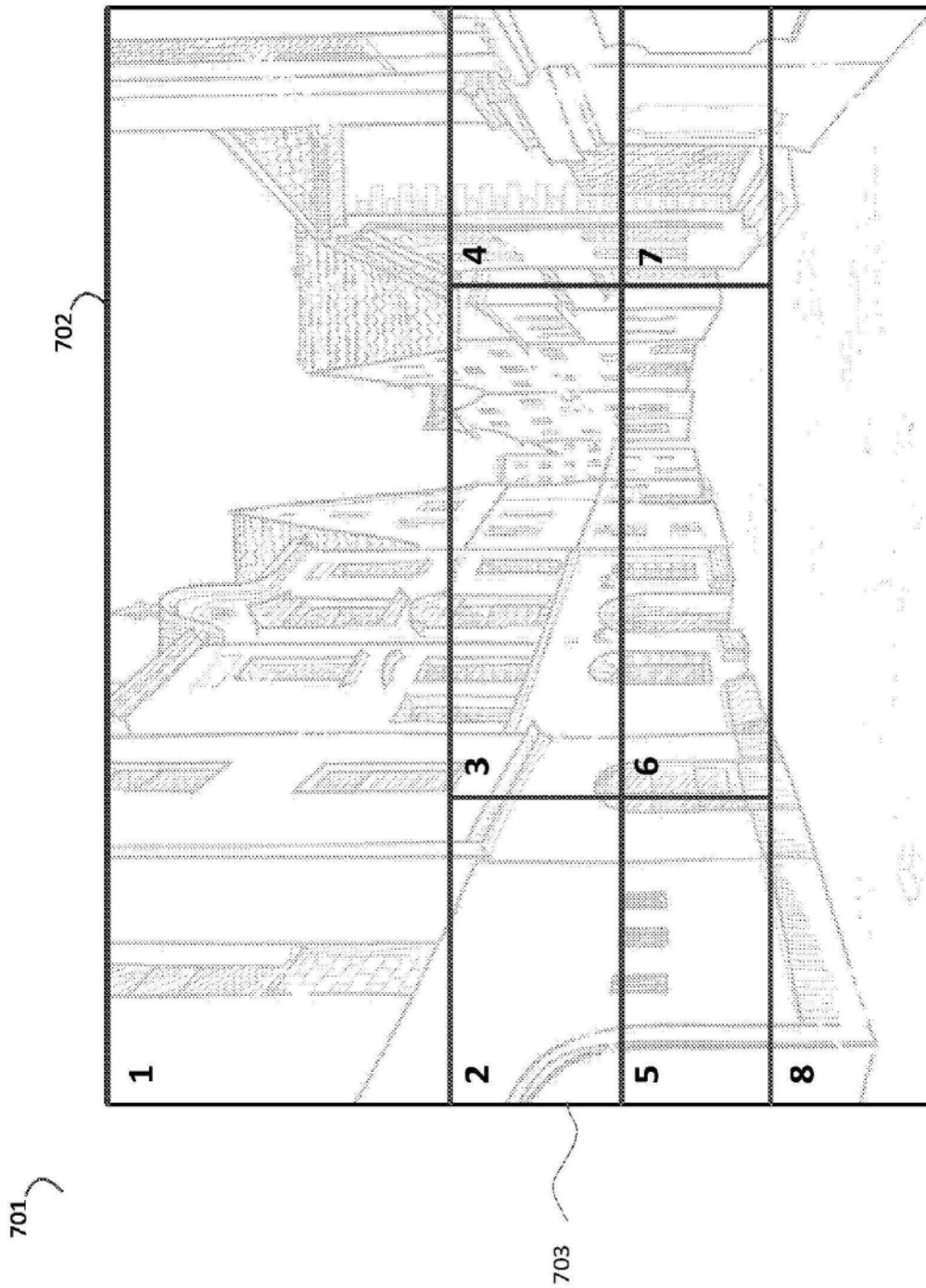


图7

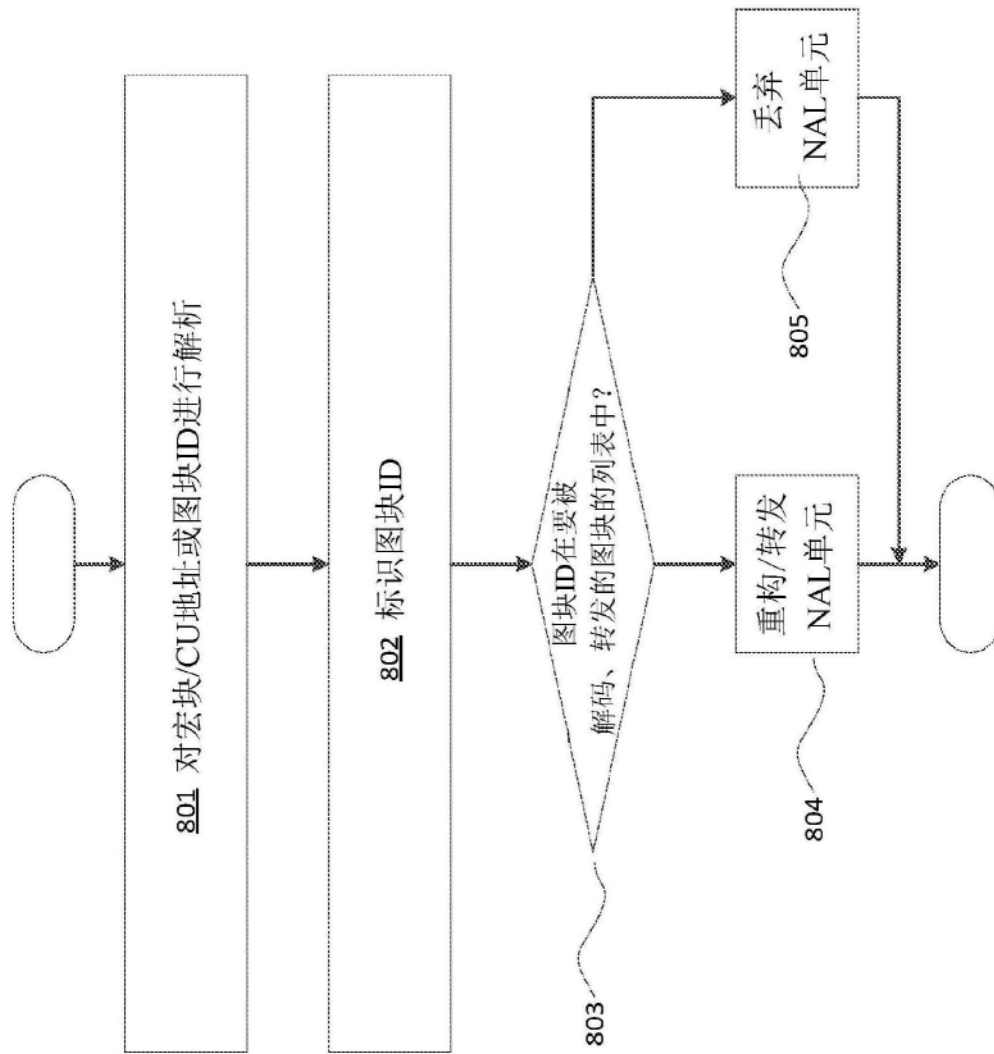


图8

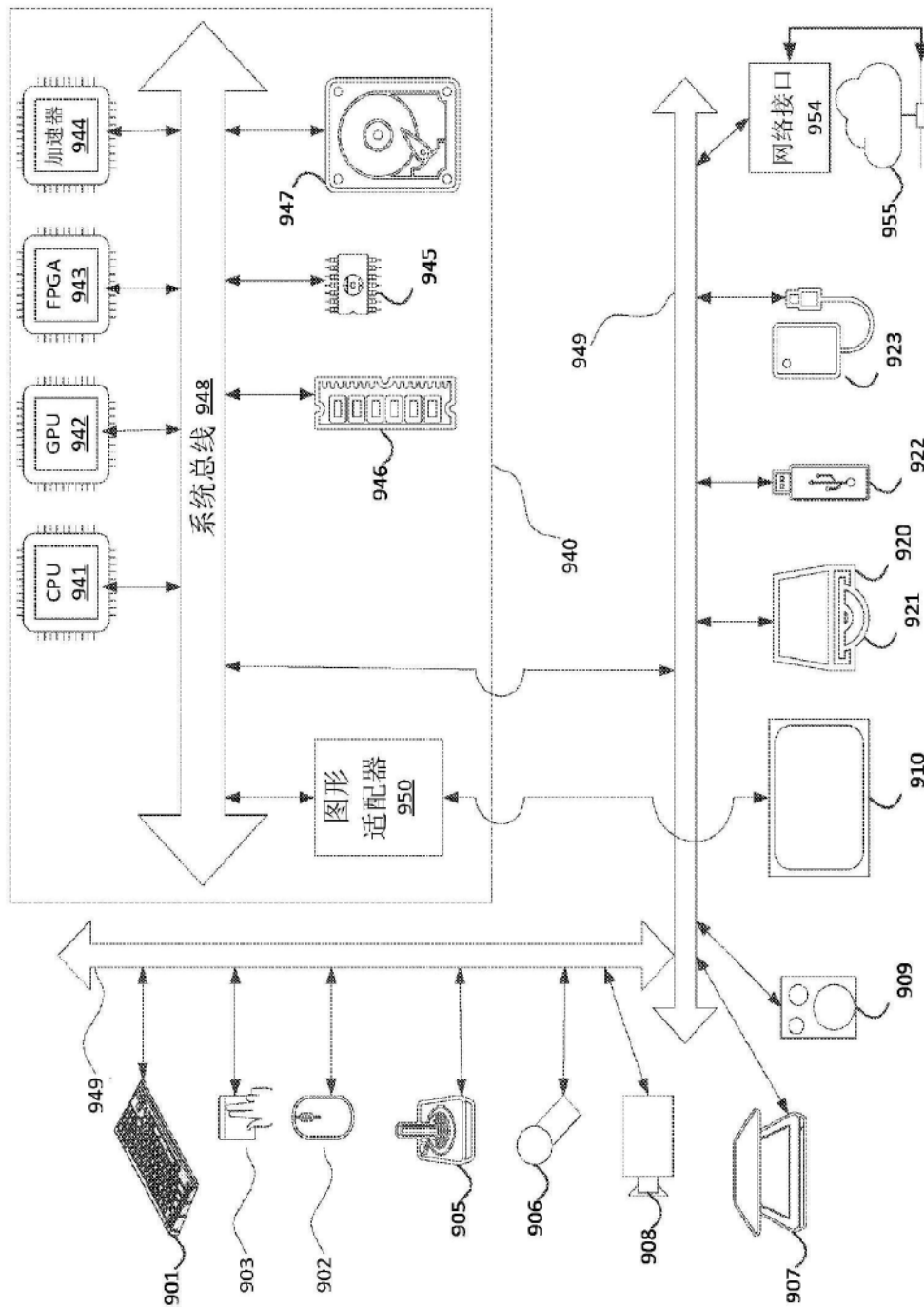


图9