



(12) 发明专利

(10) 授权公告号 CN 113796080 B

(45) 授权公告日 2024.06.14

(21) 申请号 202080031695.3

(22) 申请日 2020.09.18

(65) 同一申请的已公布的文献号  
申请公布号 CN 113796080 A

(43) 申请公布日 2021.12.14

(30) 优先权数据  
62/903,660 2019.09.20 US  
17/021,243 2020.09.15 US

(85) PCT国际申请进入国家阶段日  
2021.10.25

(86) PCT国际申请的申请数据  
PCT/US2020/051477 2020.09.18

(87) PCT国际申请的公布数据  
W02021/055738 EN 2021.03.25

(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 崔秉斗 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 程杰 王琦

(51) Int.Cl.  
H04N 19/33 (2006.01)  
H04N 19/70 (2006.01)  
H04N 19/597 (2006.01)  
H04N 19/172 (2006.01)  
H04N 19/177 (2006.01)

(56) 对比文件  
US 2015373361 A1, 2015.12.24  
US 2016227233 A1, 2016.08.04  
审查员 李若云

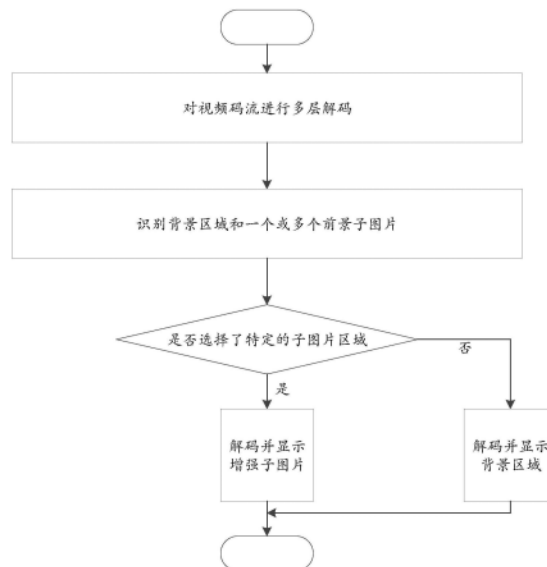
权利要求书3页 说明书28页 附图20页

(54) 发明名称

视频编解码的方法、装置、系统及计算机可读介质

(57) 摘要

本申请提供了一种视频解码的方法、装置、系统及计算机可读介质。其中,一种视频解码的方法,该方法可以包括:接收具有多个层的视频数据;识别一个或多个第一语法元素,所述一个或多个第一语法元素指定一个或多个输出层集,所述一个或多个输出层集对应于所述接收的视频数据的多个层中的一个或多个输出层;以及解码和显示与指定的所述一个或多个输出层集对应的所述一个或多个输出层。



1. 一种视频解码的方法,其特征在于,包括:

接收具有多个层的视频数据;

识别一个或多个第一语法元素,所述一个或多个第一语法元素在高级语法结构中被发信号通知,指定一个或多个输出层集,所述一个或多个输出层集对应于所述接收的视频数据的多个层中的一个或多个输出层;

识别在高级语法结构中发信号通知的一个或多个第三语法元素,所述一个或多个第三语法元素指示子图片标识符sub\_pic\_id\_layer[i][j][k],所述子图片标识符sub\_pic\_id\_layer[i][j][k]对应于与所述一个或多个输出层集相关联的每个输出层,其中,sub\_pic\_id\_layer[i][j][k]指示存在于第i个输出层集的第j层中的第k个子图片,用于对于指定输出层集的每个层,确定对哪一个子图片进行解码和输出,以及

解码与指定的所述一个或多个输出层集对应的所述一个或多个输出层;

其中,所述高级语法结构包括以下之一:视频参数集、依赖性参数集、序列参数集、图片参数集、自适应参数集和补充增强信息消息;所述序列参数集在被一个或多个图片参数集引用之前用于解码。

2. 根据权利要求1所述的方法,进一步包括:

指定所述一个或多个输出层集的数目;以及

通过设置输出层标志,来发信号通知输出层的数目。

3. 根据权利要求1所述的方法,进一步包括:识别一个或多个第二语法元素,所述一个或多个第二语法元素为所述一个或多个输出层集中的每一个输出层集,指定配置文件层级信息。

4. 根据权利要求3所述的方法,其中,在所述高级语法结构中发信号通知所述一个或多个第二语法元素。

5. 根据权利要求3所述的方法,进一步包括:

为所述一个或多个输出层集中的每一个输出层集,指定所述配置文件层级信息的数目;以及

对于所述一个或多个输出层集中的每一个输出层集,发信号通知与所述配置文件层级信息对应的一组语法元素,或者发信号通知一个索引,该索引用于指示所述配置文件层级信息中的一个配置文件层级信息条目。

6. 一种用于产生码流的视频编码的方法,其特征在于,包括:

编码具有多个层的视频数据;

设置一个或多个第一语法元素,所述一个或多个第一语法元素在高级语法结构中被发信号通知,指定一个或多个输出层集,所述一个或多个输出层集对应于所述视频数据的多个层中的一个或多个输出层;以及

设置一个或多个第三语法元素,所述一个或多个第三语法元素在高级语法结构中发信号通知,指示子图片标识符sub\_pic\_id\_layer[i][j][k],所述子图片标识符sub\_pic\_id\_layer[i][j][k]对应于与所述一个或多个输出层集相关联的每个输出层,其中,sub\_pic\_id\_layer[i][j][k]指示存在于第i个输出层集的第j层中的第k个子图片,用于对于指定输出层集的每个层,确定对哪一个子图片进行解码和输出;

编码与指定的所述一个或多个输出层集对应的所述一个或多个输出层;

其中,所述高级语法结构包括以下之一:视频参数集、依赖性参数集、序列参数集、图片参数集、自适应参数集和补充增强信息消息;所述序列参数集在被一个或多个图片参数集引用之前用于编码。

7.一种计算机系统,其特征在于,所述计算机系统包括:

一个或多个计算机可读非易失性存储介质,被配置为存储计算机程序代码;以及

一个或多个计算机处理器,被配置为存取所述计算机程序代码并按照所述计算机程序代码的指令进行操作,以执行如权利要求1-6中任一项所述的方法。

8.一种视频编码装置,其特征在于,包括:

一个或多个可读非易失性存储介质,被配置为存储程序代码;以及

一个或多个处理器,被配置为存取所述程序代码并按照所述程序代码的指令进行操作,以执行如权利要求6所述的方法。

9.一种视频解码的装置,其特征在于,包括:

接收模块,被配置为使一个或多个计算机处理器接收具有多个层的视频数据;

第一标识模块,被配置为使所述一个或多个计算机处理器识别一个或多个第一语法元素,所述一个或多个第一语法元素在高级语法结构中被发信号通知,指定一个或多个输出层集,所述一个或多个输出层集对应于来自接收的视频数据的所述多个层中的一个或多个输出层;识别在高级语法结构中发信号通知的一个或多个第三语法元素,所述一个或多个第三语法元素指示子图片标识符 $sub\_pic\_id\_layer[i][j][k]$ ,所述子图片标识符 $sub\_pic\_id\_layer[i][j][k]$ 对应于与所述一个或多个输出层集相关联的每个输出层,其中, $sub\_pic\_id\_layer[i][j][k]$ 指示存在于第 $i$ 个输出层集的第 $j$ 层中的第 $k$ 个子图片,用于对于指定输出层集的每个层,确定对哪一个子图片进行解码和输出;以及

解码模块,被配置为使所述一个或多个计算机处理器,解码指定的所述一个或多个输出层集对应的所述一个或多个输出层;

其中,所述高级语法结构包括以下之一:视频参数集、依赖性参数集、序列参数集、图片参数集、自适应参数集和补充增强信息消息;所述序列参数集在被一个或多个图片参数集引用之前用于解码。

10.根据权利要求9所述的装置,进一步包括:

第一指定模块,被配置为使所述一个或多个计算机处理器指定所述一个或多个输出层集的数目;以及

第一信令模块,被配置为使所述一个或多个计算机处理器通过设置输出层标志来发信号通知输出层的数目。

11.根据权利要求9或10所述的装置,进一步包括第二标识模块,其被配置为使所述一个或多个计算机处理器识别一个或多个第二语法元素,所述一个或多个第二语法元素为所述一个或多个输出层集中的每一个输出层集,指定配置文件层级信息。

12.根据权利要求11所述的装置,其特征在于,进一步包括:

第二指定模块,被配置为使所述一个或多个计算机处理器为所述一个或多个输出层集中的每一个输出层集,指定所述配置文件层级信息的数目;以及

第二信令模块,被配置为使所述一个或多个计算机处理器针对所述一个或多个输出层集中的每一个输出层集,发信号通知对应于所述配置文件层级信息的一组语法元素,或者

发信号通知一个索引,该索引用于指示所述配置文件层级信息中的一个配置文件层级信息条目。

13. 一种非易失性计算机可读介质,其上存储有计算机程序,所述计算机程序被配置为使一个或多个计算机处理器:执行如权利要求1-6中任一项所述的方法。

## 视频编解码的方法、装置、系统及计算机可读介质

[0001] 相关申请的交叉引用

[0002] 本申请要求于2019年9月20日提交的第62/903,660号美国临时专利申请和于2020年9月15日提交的第17/021,243号美国专利申请的优先权,这两个专利申请的全部内容并入本文。

### 技术领域

[0003] 本公开总体上涉及视频编码及解码领域,并且更具体地涉及已编码视频码流中的参数集参考和范围。

### 背景技术

[0004] 通过使用具有运动补偿的帧间图片预测技术,进行视频的编码和解码,已为人所知有几十年。未压缩的数字视频可包括一系列图片,每个图片具有例如1920×1080亮度样本及相关色度样本的空间维度。所述系列图片具有固定的或可变的图片速率(也非正式地称为帧率),例如每秒60个图片或60Hz。未压缩的视频具有非常大的比特率要求。例如,每个样本8比特的1080p60 4:2:0的视频(1920x1080亮度样本分辨率,60Hz帧率)要求接近1.5Gbit/s带宽。一小时这样的视频就需要超过600GB的存储空间。

[0005] 视频编码和解码的一个目的,是通过压缩减少输入视频信号的冗余信息。视频压缩可以帮助降低对上述带宽或存储空间的要求,在某些情况下可降低两个或更多数量级。无损和有损压缩,以及两者的组合均可采用。无损压缩是指从压缩的原始信号中重建原始信号精确副本的技术。当使用有损压缩时,重建信号可能与原始信号不完全相同,但是原始信号和重建信号之间的失真足够小,使得重建信号可用于预期应用。有损压缩广泛应用于视频。容许的失真量取决于应用。例如,与电视分发应用的用户相比,某些消费流媒体应用的用户可以容忍更高的失真。可实现的压缩比反映出:较高的允许/容许失真可产生较高的压缩比。

[0006] 视频编码器和解码器可以使用几大类技术,包括例如运动补偿、变换、量化和熵编解码。下文将介绍这几类中的一些技术。

[0007] 历史上,视频编码器和解码器倾向于在给定的图片大小上进行操作,在大多数情况下,图片大小对于已编码视频序列(Coded Video Sequence, CVS)、图片组(Group of Pictures, GOP)或类似的多图片时间帧,是定义好的并保持恒定。例如,在MPEG-2中,系统设计会根据场景活动等因素,改变水平分辨率(以及从而改变图片大小),但仅在I图片处是这样,因此这通常适用于GOP。例如从ITU-T推荐标准H.263的附录P可知,为了在CVS中使用不同的分辨率,可以重采样参考图片。然而,因为这里的图片大小并没有改变,仅对参考图片进行重采样,从而可能导致只使用了图片画布的一部分(在下采样的情况下),或者只采集了场景的一部分(在上采样的情况下)。此外,H.263附录Q允许向上或向下将单个宏块的重采样为两倍(在每个维度上)。同样,图片大小保持不变。宏块的大小在H.263中是固定的,并且因此不需要被发信号通知。

[0008] 在现代视频编解码中,在预测图片中改变图片大小变得更加主流。例如,VP9允许对整个图片进行参考图片重采样和分辨率改变。类似地,针对VVC提出的某些建议(包括,例如亨得里等人的“关于VVC的自适应分辨率改变(ARC)”,联合视频小组文件JVET-M0135-v1,2019年1月9日至19日,其整体并入本文中),允许将整个参考图片进行重采样为不同的更高或更低的分辨率。在该文件中,建议将不同的候选分辨率编码在序列参数集中,并由图片参数集中的每个图片的语法元素来引用。

## 发明内容

[0009] 本申请实施例涉及视频解码的方法、装置、系统和计算机可读介质。

[0010] 根据本申请一个方面,提供了一种视频解码的方法。该方法可以包括:接收具有多个层的视频数据;识别一个或多个第一语法元素,所述一个或多个第一语法元素指定一个或多个输出层集,所述一个或多个输出层集对应于所述接收的视频数据的多个层中的一个或多个输出层。解码与指定的所述一个或多个输出层集对应的所述一个或多个输出层。

[0011] 根据本申请另一个方面,提供了一种视频解码的计算机系统。计算机系统可以包括:一个或多个处理器、一个或多个计算机可读存储器、一个或多个计算机可读有形存储设备,以及存储在一个或多个存储设备中的至少一个上的程序指令,以供一个或多个处理器中的至少一个经由一个或多个存储器中的至少一个执行,由此计算机系统能够执行一种视频解码方法。

[0012] 根据本申请又一个方面,提供了一种视频解码的装置,包括:接收模块,被配置为使所述一个或多个计算机处理器接收具有多个层的视频数据;第一标识模块,被配置为使所述一个或多个计算机处理器识别一个或多个第一语法元素,所述一个或多个第一语法元素指定一个或多个输出层集,所述一个或多个输出层集对应于来自接收的视频数据的所述多个层中的一个或多个输出层;以及解码模块,被配置为使所述一个或多个计算机处理器,解码指定的所述一个或多个输出层集对应的所述一个或多个输出层。

[0013] 根据本申请又一个方面,提供了一种视频解码的非易失性计算机可读介质,计算机可读介质可以包括一个或多个计算机可读存储设备和存储在一个或多个有形存储设备中的至少一个上的程序指令,这些程序指令可由处理器执行,用于执行相应地可包括以下内容的方法:接收具有多个层的视频数据;识别一个或多个第一语法元素,所述一个或多个第一语法元素指定一个或多个输出层集,所述一个或多个输出层集对应于所述接收的视频数据的多个层中的一个或多个输出层;以及解码与指定的所述一个或多个输出层集对应的所述一个或多个输出层。

## 附图说明

[0014] 这些和其它目的、特征和优点将从以下结合附图阅读的说明性实施例的详细描述中变得明显。由于图示是为了便于本领域技术人员结合详细描述进行清楚的理解,因此附图的各种特征并非按比例绘制。在附图中:

[0015] 图1是根据实施例的通信系统的简化框图的示意图;

[0016] 图2是根据实施例的流媒体系统的简化框图的示意图;

[0017] 图3是根据实施例的解码器的简化框图的示意图;

- [0018] 图4是根据实施例的编码器的简化框图的示意图；
- [0019] 图5是根据实施例的用于发信号通知ARC参数的选项的示意图；
- [0020] 图6是根据实施例的语法表的示例；
- [0021] 图7是根据实施例的计算机系统的示意图；
- [0022] 图8是用于具有自适应分辨率改变的可缩放性的预测结构的示例；
- [0023] 图9是根据实施例的语法表的示例；
- [0024] 图10是对每个存取单元的poc周期和存取单元计数值进行解析和解码的简化框图的示意图；
- [0025] 图11是包括多层子图片的视频码流结构的示意图；
- [0026] 图12是具有增强分辨率的所选子图片的显示的示意图；
- [0027] 图13是包括多层子图片的视频码流的解码和显示过程的框图；
- [0028] 图14是具有子图片的增强层的360视频显示的示意图；
- [0029] 图15是子图片及其对应层和图片预测结构的布局信息的示例；
- [0030] 图16是具有局部区域的空间可缩放性形态的子图片及其对应层和图片预测结构的布局信息的示例；
- [0031] 图17是用于子图片布局信息的语法表的示例；
- [0032] 图18是用于子图片布局信息的SEI消息的语法表的示例；
- [0033] 图19是用于指示每个输出层集的输出层和配置文件/层级/级别信息的语法表的示例；
- [0034] 图20是用于指示每个输出层集的输出层模式的语法表的示例；以及
- [0035] 图21是用于指示每个输出层集的每个层的当前子图片的语法表的示例。

### 具体实施方式

[0036] 本文公开了所要求保护的结构和方法的详细实施例；然而，可以理解的是，所公开的实施例仅仅是图示了可以以各种形式体现的所要求保护的结构和方法。然而，这些结构和方法可以以许多不同的形式体现，并且不应当被解释为限于本文阐述的示例性实施例。相反，提供这些示例性实施例是为了使本公开将是彻底和完整的，并且将向本领域技术人员充分传达范围。在描述中，可以省略公知特征和技术的细节，以避免不必要地模糊所呈现的实施例。

[0037] 实施例总体上涉及数据处理领域，并且更具体地涉及媒体处理。以下描述的示例性实施例提供了一种系统、方法和计算机程序，以除其它外，允许发信号通知已编码视频数据的输出层集。因此，一些实施例具有通过改进的视频编码和解码来改进计算领域的的能力。

[0038] 如前所述，视频编码器和解码器倾向于在给定的图片大小上进行操作，在大多数情况下，该图片大小对于已解码视频序列(CVS)、图片群组(GOP)或类似的多图片时间帧，是定义好的并保持恒定。例如，在MPEG-2中，已知系统设计会根据诸如场景的活动等因素改变水平分辨率(并且由此改变图片大小)，但仅限在I图片处，因此通常用于GOP。例如从ITU-T推荐标准H.263的附录P可知，为了在CVS中使用不同的分辨率，可以重采样参考图片。然而，这里的图片大小并没有改变，仅对参考图片进行重采样，从而可能导致只使用了图片画布的一部分(在下采样的情况下)，或者只采集了场景的一部分(在上采样的情况下)。此外，

H.263附录Q允许向上或向下将单个宏块的重采样为两倍(在每个维度上)。同样,图片大小保持不变。宏块的大小在H.263中是固定的,并且因此不需要发信号通知。然而,在例如360度编解码或某些监控应用程序的背景下,多个语义上独立的源图片(例如立方体投影的360度场景的六个立方体表面,或者在多摄像机监控设置的情况下的各个摄像机输入)可能需要单独的自适应分辨率设置,以在给定时间点处理不同的每个场景的活动。换句话说,在给定的时间点,编码器可以选择对构成整个360度或监控场景的不同的语义独立图片,使用不同的重采样因子。当组合成单个图片时,对于已编码图片的部分,这反过来又要求执行参考图片重采样,以及自适应分辨率编解码信令是可用的。因此,为更好地对视频层进行标识、编码、解码和显示,使用可用的自适应分辨率编解码信令数据可能是有利的。

[0039] 图1是根据本申请公开的实施例的通信系统(100)的简化框图。通信系统(100)包括至少两个终端装置(110、120),所述至少两个终端装置可通过网络(150)彼此互连。对于单向数据传输,第一终端装置(110)可在本地对视频数据编码,以通过网络(150)传输到另一终端装置(120)。第二终端装置(120)可从网络(150)接收另一终端装置的已编码视频数据,对已编码视频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等应用中是较常见的。

[0040] 图1图示了用于支持已编码视频的双向传输的第二对终端装置(130、140),所述双向传输可例如在视频会议期间发生。对于双向数据传输,第三终端装置(130)和第四终端装置(140)中的每个终端装置可对在本地采集的视频数据进行编码,以通过网络(150)传输到第三终端装置(130)和第四终端装置(140)中的另一终端装置。第三终端装置(130)和第四终端装置(140)中的每个终端装置还可接收由第三终端装置(130)和第四终端装置(140)中的另一终端装置传输的已编码视频数据,对所述已编码视频数据进行解码,并在本地的显示装置上显示恢复的视频数据。

[0041] 在图1的实施例中,第一终端装置(110)、第二终端装置(120)、第三终端装置(130)和第四终端装置(140)可为服务器、个人计算机和智能电话,但本申请公开的原理可不限于此。本申请公开的实施例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(150)表示在第一终端装置(110)、第二终端装置(120)、第三终端装置(130)和第四终端装置(140)之间传送已编码视频数据的任何数目的网络,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。

[0042] 作为所公开主题的申请的实施例,图2示出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地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。

[0043] 流式传输系统可包括采集子系统(213),所述采集子系统可包括数码相机等视频源(201),所述视频源创建例如未压缩的视频样本流(202)。相较于已编码的视频流,视频样本流(202)被描绘为粗线以强调高数据量的视频样本流,视频样本流(202)可由耦接到数码相机(201)的视频编码器(203)处理。视频编码器(203)可包括硬件、软件或软硬件组合,以实现或实施下文更详细地描述的所公开主题的各方面。相较于视频样本流(202),已编码视频码流(204)被描绘为细线,以强调较低数据量的已编码视频码流(204),其可存储在流

式传输服务器(205)上以供将来使用。至少一个流式传输客户端(206、208)可访问流式传输服务器(205),以检索已编码视频码流(204)的副本(207)和副本(209)。客户端(206)可包括视频解码器(210),视频解码器(210)对已编码视频码流的传入副本(207)进行解码,且产生可在显示器(212)或其它呈现装置(未描绘)上呈现的输出视频样本流(211)。在一些流式传输系统中,可根据某些视频编码/压缩标准,对视频码流(204、207、209)进行编码。这些标准的实施例包括ITU-T推荐标准H.265。正在开发的视频编解码标准非正式地称为下一代视频编解码(Versatile Video Coding,VVC),本申请公开主题的实施例可用于VVC标准的上下文中。

[0044] 图3是根据本申请公开的实施例的视频解码器(210)的框图。

[0045] 接收器(310)可接收将由视频解码器(210)解码的至少一个编解码器视频序列。在同一实施例或另一实施例中,接收器(310)可以一次接收一个已编码视频序列,其中每个已编码视频序列的解码独立于其它已编码视频序列。可从信道(312)接收已编码视频序列,所述信道可以是通向存储已编码的视频数据的存储装置的硬件/软件链路。接收器(310)可接收已编码的视频数据以及其它数据,例如,可转发到它们各自的使用实体(未标示)的已编码音频数据和/或辅助数据流。接收器(310)可将已编码视频序列与其它数据分开。为了防止网络抖动,缓冲存储器(315)可耦接在接收器(310)与熵解码器/解析器(320)(此后称为“解析器”)之间。当接收器(310)从具有足够带宽和可控性的存储/转发装置或从等时同步网络接收数据时,也可能不需要配置缓冲存储器(315),或可以将所述缓冲存储器做得较小。为了在互联网等尽力而为分组网络上使用,也可能需要缓冲存储器(315),所述缓冲存储器可相对较大且可有利地具有自适应性大小。

[0046] 视频解码器(210)可包括解析器(320)以根据已熵编码视频序列重建符号(321)。这些符号的类别包括用于管理视频解码器(210)的操作的信息,以及用以控制显示器(212)等显示装置的潜在信息,所述显示屏不是视频解码器的组成部分,但可耦接到视频解码器,如图2中所示。用于显示装置的控制信息可以是辅助增强信息(Supplementary Enhancement Information,SEI消息)或视频可用性信息(Video Usability Information,VUI)的参数集片段(未标示)。解析器(320)可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根据视频编码技术或标准进行,且可遵循本领域技术人员熟知的原理,包括可变长度编码、霍夫曼编码(Huffman coding)、具有或不具有上下文灵敏度的算术编码等等。解析器(320)可基于对应于群组的至少一个参数,从已编码视频序列提取用于视频解码器中的像素的子群中的至少一个子群的子群参数集。子群可包括图片组(Group of Pictures,GOP)、图片、图块、条带、宏块、编码单元(Coding Unit,CU)、块、变换单元(Transform Unit,TU)、预测单元(Prediction Unit,PU)等等。熵解码器/解析器还可从已编码视频序列提取信息,例如变换系数、量化器参数值、运动矢量等等。

[0047] 解析器(320)可对从缓冲存储器(315)接收的视频序列执行熵解码/解析操作,从而创建符号(321)。

[0048] 取决于已编码视频图片或一部分已编码视频图片(例如:帧间图片和帧内图片、帧间块和帧内块)的类型以及其它因素,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方式可由解析器(320)从已编码视频序列解析的子群控制信息控制。为了简洁起见,未描述解析器(320)与下文的多个单元之间的此类子群控制信息流。

[0049] 除已经提及的功能块以外,视频解码器(210)可在概念上细分成如下文所描述的数个功能单元。在商业约束下运行的实际实施例中,这些单元中的许多单元彼此紧密交互并且可以彼此集成。然而,出于描述所公开主题的目的,概念上细分成下文的功能单元是适当的。

[0050] 第一单元是缩放器/逆变换单元(351)。缩放器/逆变换单元(351)从解析器(320)接收作为符号(321)的量化变换系数以及控制信息,包括使用哪种变换方式、块大小、量化因子、量化缩放矩阵等。缩放器/逆变换单元(351)可输出包括样本值的块,所述样本值可输入到聚合器(355)中。

[0051] 在一些情况下,缩放器/逆变换单元(351)的输出样本可属于帧内编码块;即:不使用来自先前已重建的图像的预测性信息,但可使用来自当前图像的先前已重建部分的预测性信息的块。此类预测性信息可由帧内图片预测单元(352)提供。在一些情况下,帧内图片预测单元(352)采用从当前(部分已重建)图片存储器(358)提取的周围已重建信息,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(355)基于每个样本,将帧内预测单元(352)生成的预测信息添加到由缩放器/逆变换单元(351)提供的输出样本信息中。

[0052] 在其它情况下,缩放器/逆变换单元(351)的输出样本可属于帧间编码和潜在运动补偿块。在此情况下,运动补偿预测单元(353)可访问参考图片缓冲存储器(357)以提取用于预测的样本。在根据符号(321)对提取的样本进行运动补偿之后,这些样本可由聚合器(355)添加到缩放器/逆变换单元(351)的输出(在这种情况下被称作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)从参考图片缓冲存储器(357)内的地址获取预测样本可受到运动矢量控制,且所述运动矢量以所述符号(321)的形式而供运动补偿预测单元(353)使用,所述符号(321)例如是包括X、Y和参考图片分量。运动补偿还可包括在使用子样本精确运动矢量时,从参考图片缓冲存储器(357)提取的样本值的内插、运动矢量预测机制等等。

[0053] 聚合器(355)的输出样本可在环路滤波器单元(356)中被各种环路滤波技术采用。视频压缩技术可包括环路内滤波器技术,所述环路内滤波器技术受控于包括在已编码视频码流中的参数,且所述参数作为来自解析器(320)的符号(321)可用于环路滤波器单元(356),视频压缩技术还可响应于在解码已编码图片或已编码视频序列的先前(按解码次序)部分期间获得的元信息,以及响应于先前已重建且已经过环路滤波的样本值。

[0054] 环路滤波器单元(356)的输出可以是样本流,所述样本流可输出到显示器(212)以及存储在当前图片存储器(358),以用于后续的帧间图片预测。

[0055] 一旦完全重建,某些已编码图片就可用作参考图片以用于将来预测。举例来说,一旦已编码图片被完全重建,且已编码图片(通过例如解析器(320))被识别为参考图片,则当前图片存储器(358)可变为参考图片缓冲存储器(357)的一部分,且可在开始重建后续已编码图片之前重新分配新的当前图片存储器(358)。

[0056] 视频解码器(210)可根据例如ITU-T推荐H.265标准中记录的预定视频压缩技术执行解码操作。已编码视频序列可符合所使用的视频压缩技术或标准指定的语法,从某种意义上说,已编码视频序列遵守视频压缩技术或标准的语法,如视频压缩技术文档或标准中所指定,特别是其中的配置文件。对于合规性,还要求已编码视频序列的复杂度处于如视频压缩技术或标准的层级所限定的范围内。在一些情况下,层级限制最大图片大小、最大帧

率、最大重建采样率(以例如每秒兆(mega)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,由层级设定的限制可通过假想参考解码器(Hypothetical Reference Decoder,HRD)规范和已在已编码视频序列中发信号通知的HRD缓冲器管理的元数据来进一步限定。

[0057] 在实施例中,接收器(310)可连同已编码视频一起接收附加(冗余)数据。所述附加数据可以是已编码视频序列的一部分。所述附加数据可由视频解码器(320)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。附加数据可呈例如时间、空间或信噪比(signal noise ratio,SNR)增强层、冗余条带、冗余图片、前向纠错码等形式。

[0058] 图4是根据本申请公开的实施例的视频编码器(203)的框图。

[0059] 视频编码器(203)可从视频源(201)(并非编码器的一部分)接收视频样本,所述视频源可采集将由视频编码器(203)编码的视频图像。

[0060] 视频源(201)可提供将由视频编码器(203)编码的呈数字视频样本流形式的源视频序列,所述数字视频样本流可具有任何合适位深度(例如:8位、10位、12位……)、任何色彩空间(例如BT.601Y CrCb、RGB……)和任何合适采样结构(例如Y CrCb 4:2:0、Y CrCb4:4:4)。在媒体服务系统中,视频源(201)可以是存储先前已准备的视频的存储装置。在视频会议系统中,视频源(201)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片,当按顺序观看时,这些图片被赋予运动。图片自身可构建为空间像素阵列,其中取决于所用的采样结构、色彩空间等,每个像素可包括至少一个样本。所属领域的技术人员可以很容易理解像素与样本之间的关系。下文侧重于描述样本。

[0061] 根据实施例,视频编码器(203)可实时或在由应用所要求的任何其它时间约束下,将源视频序列的图片编码且压缩成已编码视频序列(443)。施行适当的编码速度是控制器(450)的一个功能。控制器(450)控制如下文所描述的其它功能单元且在功能上耦接到这些单元。为了简洁起见,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(图片跳过、量化器、率失真优化技术的 $\lambda$ 值等)、图片大小、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(450)的其它功能,这些功能涉及针对某一系统设计优化的视频编码器(203)。

[0062] 一些视频解码器以本领域技术人员识别为“编码环路”的方式进行操作。作为简单的描述,编码环路可包括编码器(430)(在下文称为“源编码器”),负责基于待编码的输入图片和参考图片创建符号)的编码部分和嵌入于视频编码器(203)中的(本地)解码器(433)。解码器(433)重建符号,以创建(远程)解码器也要创建的样本数据(因为在本公开主题中所考虑的视频压缩技术中,符号与已编码视频码流之间的任何压缩是无损的)。将重建的样本流输入到参考图片存储器(434)。由于符号流的解码产生与解码器位置(本地或远程)无关的位精确结果,因此参考图片存储器(434)的缓冲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。本领域技术人员熟知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)。

[0063] “本地”的解码器(433)的操作可与已在上文结合图3详细描述“远程”解码器(210)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的解码器(433)中实施。

[0064] 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。

[0065] 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。参考来自视频序列中被指定为“参考帧”的至少一个先前已编码帧,所述运动补偿预测编码对输入帧进行预测性编码。以此方式,编码引擎(432)对输入帧的像素块与参考帧的像素块之间的差异进行编码,所述参考帧可被选作所述输入帧的预测参考。

[0066] 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帧执行,且可使已重建的参考帧存储在参考图片高速缓存(434)中。以此方式,视频编码器(203)可在本地存储已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已重建参考帧具有共同内容(不存在传输误差)。

[0067] 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。

[0068] 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。

[0069] 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、算术编码等本领域技术人员所熟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已编码视频序列。

[0070] 传输器(440)可缓冲由熵编码器(445)创建的已编码视频序列,从而为通过通信信道(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)。

[0071] 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以下任一种帧类型:

[0072] 帧内图片(I图片),其可以是不将序列中的任何其它帧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)图片。所属领域的技术人员了解I图片的变体及其相应的应用和特征。

[0073] 预测性图片(P图片),其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。

[0074] 双向预测性图片(B图片),其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,多个预测性图片可使用多于两个参考图片和相关联元数据以用于重建单个块。

[0075] 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。

[0076] 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。

[0077] 在实施例中,传输器(440)可在传输已编码的视频时传输附加数据。源编码器(430)可将此类数据作为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。在更详细地描述本申请所公开的主题的某些方面之前,需要先介绍将在本说明书的其余部分中提及的几个术语。

[0078] “子图片”,在下文中,在某些情况下指的是在语义上分组的、可以用变化的分辨率独立编码的样本、块、宏块、编码单元或类似实体的矩形排列。一个图片可以有一个或多个子图片。一个或多个已编码子图片可以形成一个已编码图片。一个或多个子图片可以组装成一个图片,并且可以从一个图片中提取一个或多个子图片。在某些环境中,可以在压缩域中将一个或多个已编码子图片组装为一个已编码图片,而无需将其转码到样本级,在同一情况下或在某些其它情况下,可以从压缩域中的已编码图片中提取一个或多个已编码子图片。

[0079] “自适应分辨率改变(Adaptive Resolution Change,ARC)”,在下文指的是,允许通过例如参考图片重采样,来改变已编码视频序列内的图片或子图片的分辨率的机制。“ARC参数”在下文指的是执行自适应分辨率改变所需的控制信息,其可以包括例如滤波器参数、缩放因子、输出图片和/或参考图片的分辨率、各种控制标志等等。

[0080] 以上的描述着重于对单个的、语义上独立的已编码视频图片的编码和解码。在描述具有独立的ARC参数的多个子图片的编码/解码的含义及其隐含的附加复杂性之前,将先描述用于发信号通知ARC参数的选项。

[0081] 参考图5,示出了用于发信号通知ARC参数的几个新选项。如每个选项所指出的,从编码效率、复杂度和体系结构的角度来看,它们具有某些优点和缺点。视频编码标准或技术可以选择这些选项中的一个或多个,或选择现有技术中已知的选项,用于发信号通知ARC参数。这些选项可以不是互斥的,并且可以根据应用需要、所涉及的标准技术或编码器的选择进行互换。

[0082] ARC参数的类别可以包括:

[0083] 在X维度和Y维度上分开或组合的上采样因子和/或下采样因子。

[0084] 加上时间维度的上采样因子和/或下采样因子,用于指示对给定数量的图片进行恒速放大和/或缩小。

[0085] 以上两者中的任一者可涉及对一个或多个可能较短的语法元素进行编码,所述一个或多个可能较短的语法元素可能指向一个包含所述因子(多个因子)的表格。

[0086] 分辨率:在X或Y维度上,以样本、块、宏块、CU或任何其他合适的粒度为单位的,输入图片、输出图片、参考图片、编码图片的组合或单独的分辨率。如果存在一个以上的分辨率(例如,一个分辨率用于输入图片,一个分辨率用于参考图片),则在某些情况下,可以从一组值中推断出另一组值。例如,可以通过使用标志选通分辨率。有关更详细的示例,请参阅下文。

[0087] “变形”(warping)坐标:类似于在H.263标准附录P中使用的坐标,可以具有如上所述的适当粒度。H.263标准附录P定义了一种对这种变形坐标进行编码的有效方式,但是可以想象,还可以设想其它可能更有效的方法。例如,附录P的变形坐标的可变长度可逆的“霍夫曼”式编码可以用适当长度的二进制编码来代替,其中,二进制码字的长度可以例如根据最大图片大小推导出(可能乘以某个因子并偏移某个值),以便允许在最大图片大小的边界之外进行“变形”。

[0088] 上采样滤波器参数或下采样滤波器参数:在最简单的情况下,可能只有单个的滤波器用于进行上采样和/或下采样。然而,在某些情况下,允许滤波器的设计有较大的灵活性,可能是有利的,这可能需要通过发信号通知滤波器参数来实现。此类参数可以通过一个可能的滤波器设计列表中的索引来选择,可以充分地指定滤波器(例如,通过使用适当的熵编解码技术的滤波器系数列表),可以通过上采样比率和/或下采样比率隐含地选择滤波器,所述上采样比率和/或下采样比率反过来是根据上面提到的机制中的任何机制发信号通知的,等等。

[0089] 在下文中,本说明书假设,通过码字指示,对上采样因子和/或下采样因子(在X维度和Y维度使用相同的因子)的有限集合进行编码。所述码字可以通过例如使用视频编码规范(例如H.264和H.265)中的某些语法元素所共有的指数哥伦布(Exp-Golomb)编码,进行有利地可变长度编码。

[0090] 可以根据应用的需要以及视频压缩技术或标准中可用的放大和缩小机制的能力,设计许多类似的码字的值与上采样因子和/或下采样因子之间的映射。该映射对应的表可以扩展到更多的值。码字的值还可以用除Exp-Golomb码之外的熵编解码机制来表示,例如使用二进制编码。当在视频处理引擎(最重要的是编码器和解码器)本身之外,例如MANE(Media Aware Network Element,媒体感知网元)对重采样因子感兴趣时,使用二进制编码可能具有某些优点。应当注意,对于(可能)最常见的、不需要改变分辨率的情况,可以选择较短的Exp-Golomb码;在上述表中,Exp-Golomb码例如只有单个比特。对于该最常见的情况,Exp-Golomb码可以比使用二进制码具有编码效率的优势。

[0091] 上述表中条目的数量及其语义可以是完全可配置的或部分可配置的。例如,上述表的基本形态可以在诸如序列参数集或解码器参数集等的“高层”参数集中传送。可替换地或另外,可以在视频编解码技术或标准中定义一个或多个此类表,并且可以通过例如解码

器或序列参数集来选择。

[0092] 下面描述如何将如上所述编码的上采样因子/下采样因子 (ARC信息) 包含在视频编解码技术或标准语法中。类似的考虑可以应用于控制上采样滤波器/下采样滤波器的一个或几个码字。关于滤波器或其它数据结构何时需要相对大量的数据, 参见下文的讨论。

[0093] H. 263标准附录P将四个变形坐标形式的ARC信息502包含在图片头501中, 更具体地, 是包含在H. 263标准增强P类型 (PLUSPTYPE) (503) 头扩展中。当a) 存在可用的图片头, 以及b) 预计ARC信息会频繁改变时, 这可能是一个明智的设计选择。然而, 使用H. 263型信令的开销可能相当高, 并且缩放因子可能不适用于图片边界, 因为图片头可能具有瞬时性质。

[0094] 上文提到的JVCET-M135-v1, 包括位于图片参数集 (504) 中的ARC参考信息 (505) (索引), 用于对包括目标分辨率的表 (506) 进行索引, 而表 (506) 位于序列参数集 (507) 内。根据作者做出的口头陈述, 可能的分辨率在序列参数集 (SPS) (507) 中的表 (506) 中的位置, 通过在能力交换期间使用SPS作为互操作性协商点, 可以证明可能的分辨率在序列参数集 (SPS) (507) 中的表 (506) 中的位置是合理的。在表 (506) 中的值所设置的限制内, 通过参考适当的图片参数集 (Picture Parameter Set, PPS) (504), 分辨率可以随图片而改变。

[0095] 仍参照图5, 可以存在以下附加选项, 以在视频码流中传送ARC信息。这些选项中的每一个选项都具有优于上述现有技术的某些优点。这些选项可以同时存在于同一视频编码技术或标准中。

[0096] 在一个实施例中, 诸如重采样 (缩放) 因子等ARC信息 (509) 可以出现在条带头、GOB (Group of Blocks, 块组) 头、图块头、或图块组头 (以下称为图块组头) (508) 中。ARC信息很小, 这样可能是足够的, 例如, ARC信息是几个比特的单个可变长度ue(v) 的码字或固定长度的码字, 例如, 如上所示。ARC信息可以应用于例如由图块组表示的子图片, 而不是用于整个图片, 将ARC信息直接包含在图块组头中具有额外的优点。另见下文。此外, 即使视频压缩技术或标准仅设想了整个图片的自适应分辨率改变 (与例如基于图块组的自适应分辨率改变相反), 从容错的角度来看, 将ARC信息放入图块组头中, 相较于将其放入H. 263型图片头中, 具有一定的优势。

[0097] 在同一实施例或另一实施例中, ARC信息 (512) 本身可以存在于适当的参数集 (511) 中, 例如, 图片参数集、头参数集、图块参数集、适配参数集 (Adaptation Parameter Set, APS) 等等 (本文中描述了适配参数集)。该参数集的范围可以有利地不大于一个图片, 例如为图块组。ARC信息 (512) 的使用可以通过激活相关参数集而隐含。例如, 当视频编码技术或标准只考虑基于图片的ARC时, 图片参数集或等同参数集可能适用。

[0098] 在同一实施例或另一实施例中, ARC参考信息 (513) 可以存在于图块组头 (514) 或类似数据结构中。该ARC参考信息 (513) 可以指范围超出了单个图片的参数集 (516) (例如, 序列参数集或解码器参数集) 中可用的ARC信息 (515) 的子集。

[0099] 如JVET-M0135-v1中使用的, 从图块组头、PPS、SPS间接暗示激活PPS的附加级别似乎是不必要的, 因为图片参数集可以与序列参数集一样, 可以用于 (在RFC3984等某些标准中已经使用) 能力协商或通告。然而, 如果ARC信息应当适用于例如也是由图块组表示的子图片, 则激活范围限制在图块组的参数集 (例如适配参数集或头参数集) 可能是更好的选择。而且, 如果ARC信息具有超过可忽略的大小——例如包含多个滤波器系数等滤波器控制信息——则从编码效率的角度, 选择使用参数集可能是比选择直接使用图块组头 (508) 更

好,因为那些设置可以由将来的图片或子图片通过参考相同的参数集复用。

[0100] 当使用序列参数集或范围跨越多个图片的另一更高层参数集时,可能需要考虑以下几点:

[0101] 用于存储ARC信息表(516)的参数集在一些情况下可以是序列参数集,但是在其它情况下有利地可以是解码器参数集。解码器参数集可以具有多个CVS(Coded Video Sequence,已编码视频序列)(即,已编码视频流)的激活范围,即,从会话开始直到会话结束期间的所有已编码视频比特。此类范围可能是更合适的,因为可能的ARC因子可能是可能以硬件实现的解码器特征,并且硬件特征倾向于不随任何CVS(其在至少一些娱乐系统中是长度为一秒或更短的图片组)而改变。也就是说,将所述表放入序列参数集中是明确地包括在本申请描述的放置选项中。

[0102] ARC参考信息(513)可以有利地直接放置在图片/条带/图块/GOB/图块组头(下文为图块组头)(514)中,而不是像JVCEI-M0135-v1中那样放置在图片参数集中。原因如下:当编码器想要改变图片参数集中的单个值时,例如ARC参考信息,则编码器可能必须创建新的PPS并参考该新的PPS。假设仅ARC参考信息改变,但是诸如PPS中的量化矩阵信息等其它信息保留。此类信息可能很大,并且会需要重新发送以使新的PPS完整。由于ARC参考信息可能是单个码字,诸如ARC参考信息(513)中的索引,并且这将是惟一改变的值,因此重新传输所有的例如量化矩阵信息将非常麻烦和浪费。因此,从编码效率的角度来看,到目前为止,可以更好地避免如在JVET-M0135-v1中所提出的通过PPS的间接访问。类似地,将ARC参考信息放入PPS具有另外的缺点,即由ARC参考信息(513)参考的ARC信息必须应用于整个图片而不是子图片,因为图片参数集激活的范围是图片。

[0103] 在同一个实施例或另一个实施例中,ARC参数的信令可以遵循图6中概述的详细示例。图6描绘了至少自1993年以来在视频编码标准中使用的表示的语法图。此类语法图的符号大致遵循C风格的编程。粗体线表示码流中存在的语法元素,无粗体线通常表示控制流或变量的设置。

[0104] 图块组头(601),作为可应用于图片的(可能是矩形的)一部分的头的示例性语法结构,可以有条件地包含可变长度的Exp-Golomb编码的语法元素`dec_pic_size_idx`(602)(以粗体示出)。该语法元素在图块组头中的存在可以通过使用自适应分辨率(603)来选通。在此,标志的值未以黑体示出,这意味着该标志在码流中出现的位置点是其出现在语法表中。可以以码流内部或外部的任何高级语法结构来发信号通知自适应分辨率是否用于该图片或其一部分。在所示的示例中,如下所述在序列参数集中发信号通知自适应分辨率。

[0105] 仍然参考图6,还示出了序列参数集(610)的摘录。所示的第一语法元素是`adaptive_pic_resolution_change_flag`(611)。当为真时,该标志可以指示使用了自适应分辨率,而自适应分辨率可能需要特定的控制信息。在该示例中,此类控制信息基于标志的值有条件地存在,该标志基于序列参数集(612)和图块组头(601)中的`if()`语句。

[0106] 当使用自适应分辨率时,在本例子中,输出分辨率是以样本(613)为单位进行编码。数字613指的是`output_pic_width_in_luma_samples`和`output_pic_height_in_luma_samples`,它们可以一起定义输出图片的分辨率。在视频编码技术或标准的其它地方,可以定义对任一值的某些限制。例如,级别定义可以限制总输出采样的数量,其可以是上述两个语法元素的值的乘积。此外,某些视频编码技术或标准,或外部技术或标准(例如,系统标

准)可限制数值范围(例如,一个维度或两个维度都必须可被2的幂数除尽)或纵横比(例如,宽度和高度必须具有例如4:3或16:9的关系)。可以引入此类限制以便于硬件实现或用于其它原因,并且在本领域中是公知的。

[0107] 在某些应用中,建议编码器指示解码器使用某个参考图片大小,而不是隐含地假定其大小为输出图片大小。在该示例中,语法元素reference\_pic\_size\_present\_flag (614)选通参考图片尺寸(615)的有条件存在(同样,该数字指的是宽度和高度两者)。

[0108] 最后,示出了一个可能的解码图片宽度和高度的表。此表可以例如通过表指示(num\_dec\_pic\_size\_in\_luma\_samples\_minus1) (616)来表示。“minus1”(减1)可以指该语法元素的值的解释。例如,如果编码值为零,则存在一个表条目。如果值为五,则存在六个表条目。对于表中的每一“行”,已解码的图片宽度和高度然后被包含在语法(617)中。

[0109] 所呈现的表条目(617)可以使用图块组头中的语法元素dec\_pic\_size\_idx (602)进行索引,从而允许每个图块组具有不同的解码大小(实际上是缩放因子)。

[0110] 上文所描述的发信号通知自适应分辨率参数的技术可实施为使用计算机可读指令的且以物理方式存储在至少一个计算机可读存储介质中的计算机软件。例如,图7示出适于实施本申请主题的某些实施例的计算机系统700。

[0111] 所述计算机软件可使用任何合适的机器代码或计算机语言来编码,所述机器代码或计算机语言可经受汇编、编译、链接或类似机制以创建包括指令的代码,所述指令可直接或通过解译、微码执行等而由计算机中央处理单元(central processing unit,CPU)、图形处理单元(Graphics Processing Unit,GPU)等执行。

[0112] 所述指令可在各种类型的计算机或计算机组件上执行,所述计算机或计算机组件包括例如个人计算机、平板电脑、服务器、智能电话、游戏装置、物联网装置等。

[0113] 图7中所示的用于计算机系统700的组件在本质上是示范性的,并非旨在暗示关于实施本申请的实施例的计算机软件的使用或功能的范围的任何限制。也不应将组件的配置解释为对计算机系统700的示范性实施例中所示的组件中的任一个组件或组件组合有任何依赖或需求。

[0114] 计算机系统700可包括某些人机接口输入装置。此类人机接口输入装置可响应于至少一个人类用户通过例如触觉输入(例如:按键、滑动、数据手套移动)、音频输入(例如:语音、拍击)、视觉输入(例如:手势)、嗅觉输入(未描绘)进行的输入。人机接口装置还可用于捕获未必与人的有意识输入直接相关的某些媒体,例如音频(例如:话语、音乐、环境声)、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)。

[0115] 输入人机接口装置可包括以下至少一个(每种仅描绘一个):键盘701、鼠标702、轨迹垫703、触摸屏710、数据手套(704)、操纵杆705、麦克风706、扫描仪707、相机708。

[0116] 计算机系统700还可包括某些人机接口输出装置。此类人机接口输出装置可通过例如触觉输出、声音、光和气味/味道刺激至少一个人类用户的感受。此类人机接口输出装置可包括触觉输出装置(例如触摸屏710、数据手套(704)或操纵杆705的触觉反馈,但还可存在不充当输入装置的触觉反馈装置)、音频输出装置(例如:扬声器709、头戴式耳机(未描绘))、视觉输出装置(例如屏幕710,包括CRT屏幕、LCD屏幕、等离子体屏幕、OLED屏幕,各自具有或不具有触摸屏输入能力,各自具有或不具有触觉反馈能力,其中的一些能够通过例

如立体平画输出的方式输出二维视觉输出或大于三维的输出;虚拟现实眼镜(未描绘)、全息显示器和烟雾箱(未描绘),以及打印机(未描绘)。

[0117] 计算机系统700还可包括人类可访问的存储装置和存储装置的相关联介质,例如光学介质,包括具有CD/DVD等介质721的CD/DVD ROM/RW720、拇指驱动器722、可移动硬盘驱动器或固态驱动器723、磁带和软盘(未描绘)等传统磁性媒体、基于ROM/ASIC/PLD的专用装置,例如安全保护装置(未描绘),等等。

[0118] 所属领域的技术人员还应理解,结合当前公开的主题使用的术语“计算机可读介质”并未涵盖传输介质、载波或其他瞬时信号。

[0119] 计算机系统700还可包括到至少一个通信网络的接口。所述网络可例如是无线的、有线的、光学的。所述网络还可以是本地的、广域的、城域的、车载和工业的、实时的、容忍延迟的等等。所述网络的示例包括例如以太网、无线LAN的局域网、包括GSM、3G、4G、5G、LTE等的蜂窝网络、包括有线TV、卫星TV和地面广播TV的TV有线或无线广域数字网络、包括CAN总线的车载网络和工业网络等。某些网络通常需要附接到某些通用数据端口或外围总线749(例如,计算机系统700的USB端口)的外部网络接口适配器;其他网络通常通过附接到如下文所描述的系统总线而集成到计算机系统700的核心中(例如通过以太网接口集成到PC计算机系统中,或通过蜂窝网络接口集成到智能电话计算机系统中)。通过使用这些网络中的任一网络,计算机系统700可与其他实体通信。此类通信可以是仅单向接收(例如广播TV)、仅单向发送(例如连到某些CAN总线装置的CAN总线)或是双向的,例如使用局域数字网络或广域数字网络连接到其他计算机系统。可在如上文所描述的那些网络和网络接口中的每一个上使用某些协议和协议栈。

[0120] 上述人机接口装置、人类可访问存储装置和网络接口可附接到计算机系统700的核心740。

[0121] 核心740可包括至少一个中央处理单元(Central Processing Unit,CPU)741、图形处理单元(Graphics Processing Unit,GPU)742、现场可编程门区域(Field Programmable Gate Areas,FPGA)形式的专用可编程处理单元743、用于某些任务的硬件加速器744等等。这些装置连同只读存储器(read-only memory,ROM)745、随机存取存储器746、例如内部非用户可访问的硬盘驱动器、SSD等内部大容量存储装置747可通过系统总线748连接。在一些计算机系统中,系统总线748可通过至少一个物理插头形式访问以实现通过额外CPU、GPU等来扩展。外围装置可直接或通过外围总线749附接到核心的系统总线748。用于外围总线的架构包括PCI、USB等等。

[0122] CPU741、GPU742、FPGA743和加速器744可执行某些指令,所述指令组合起来可构成上述计算机代码。计算机代码可存储在ROM745或RAM746中。过渡数据也可存储在RAM746中,而永久性数据可例如存储在内部大容量存储装置747中。可通过使用高速缓冲存储器来实现对任一存储器装置的快速存储和检索,所述高速缓冲存储器可与至少一个CPU741、GPU742、大容量存储装置747、ROM745、RAM746等紧密关联。

[0123] 计算机可读介质上可具有用于执行各种计算机实施的操作的计算机代码。所述介质和计算机代码可以是专为本申请的目的设计和构建的介质和计算机代码,或可属于计算机软件领域中的技术人员众所周知且可用的种类。

[0124] 举例来说但不作为限制,具有架构700且尤其是核心740的计算机系统可提供因处

理器(包括CPU、GPU、FPGA、加速器等)执行以至少一个有形计算机可读介质体现的软件而产生的功能。此类计算机可读介质可以是与上文所介绍的用户可访问大容量存储装置以及核心740的非易失性质的某些存储装置(例如核心内部大容量存储装置747或ROM745)相关联的介质。实施本申请的各种实施例的软件可存储在此类装置中且由核心740执行。根据特定需求,计算机可读介质可包括至少一个存储器装置或芯片。软件可使核心740且具体地说使其中的处理器(包括CPU、GPU、FPGA等等)执行本文中所描述的特定过程或特定过程的特定部分,包括限定存储在RAM746中的数据结构以及根据由软件限定的过程修改此类数据结构。另外或作为替代方案,计算机系统可提供由硬连线的或以其他方式体现于电路(例如:加速器744)中的逻辑所产生的功能,所述逻辑可代替或连同软件一起操作以执行本文描述的特定过程或特定过程的特定部分。适当时,对软件的引用可涵盖逻辑,且反之亦然。适当时,对计算机可读介质的引用可涵盖存储用于执行的软件的电路(例如集成电路(IC))、体现用于执行的逻辑的电路或这两种电路。本申请涵盖硬件与软件的任何合适的组合。

[0125] 某些视频编解码技术或标准(例如VP9)通过结合时间可缩放性,实施某些形式的参考图片重采样(以与所公开的主题完全不同的方式发信号通知),以支持空间可缩放性,从而实现空间可缩放性。具体地,可以使用ARC风格技术将某些参考图片上采样到更高的分辨率,以形成空间增强层的基础。可以使用高分辨率的正常预测机制来修正那些上采样的图片,以便添加细节。

[0126] 所公开的主题可用于这种环境。在某些情况下,在相同或另一实施例中,NAL单元头中的值(例如,时间ID字段)不仅可以用于指示时间层,还可以用于指示空间层。这样做对于某些系统设计具有某些优点;例如,基于NAL单元头的时间ID值,为时间层选定的转发而创建和优化的现有选定的转发单元(SFU),可以在不修改的情况下,用于可缩放环境。为了实现这一点,可能需要由NAL单元头中的时间ID字段来指示已编码图片大小和时间层之间的映射。

[0127] 在一些视频编解码技术中,存取单元(AU)可以是指,在给定时间实例中捕获并组成相应图片/条带/图块/NAL单元码流的,一个或多个已编码图片、一个或多个条带、一个或多个图块、一个或多个NAL单元等。该时间实例可以是合成时间。

[0128] 在HEVC和某些其它视频编解码技术中,图片顺序计数(POC)值可以用于指示在已解码图片缓冲器(DPB)中存储的多个参考图片中选择的参考图片。当存取单元(AU)包括一个或多个图片、条带或图块时,属于同一AU的每个图片、条带或图块可以携带相同的POC值,从中可以导出它们是从相同合成时间的内容创建的。换言之,在两个图片/条带/图块携带相同的给定POC值的场景中,相同的给定POC值可以指示这两个图片/条带/图块属于相同的AU并且具有相同的合成时间。相反地,具有不同的POC值的两个图片/图块/条带可以指示,那些图片/条带/图块属于不同的AU并且具有不同的合成时间。

[0129] 在所公开主题的实施例中,可以放宽前述严格关系,因为存取单元可以包括具有不同的POC值的图片、条带或图块。通过在AU内允许不同POC值,可以使用POC值来识别具有相同呈现时间的潜在独立可解码的图片/条带/图块。这又使得能够在不改变参考图片选择信令(例如,参考图片集信令或参考图片列表信令)的情况下支持多个可缩放层,如下面更详细描述。

[0130] 然而,仍然希望能够仅从POC值,相对于其它具有不同POC值的图片/条带/图块,来

识别图片/条带/图块所属的AU。这是可以实现的,如下所述。

[0131] 在相同或其它实施例中,在高级语法结构(诸如NAL单元头、条带头、图块组头、SEI消息、参数集或AU定界符)中,发信号通知存取单元计数(AUC)。AUC的值可以用于识别哪些NAL单元、图片、条带或图块属于给定的AU。AUC的值可以对应于不同的合成时间实例。AUC值可以等于POC值的倍数。通过将POC值除以整数值,可以计算AUC值。在某些情况下,除法运算可能对解码器实现造成一定负担。在这种情况下,在AUC值的编号空间中的小限制可以允许用移位运算代替除法运算。例如,AUC值可以等于POC值范围的最高有效位(MSB)值。

[0132] 在相同的实施例中,在高级语法结构(诸如NAL单元头、条带头、图块组头、SEI消息、参数集或AU定界符)中,发信号通知每个AU的POC周期的值(`poc_cycle_au`)。`poc_cycle_au`可以指示多少个不同且连续的POC值可以与同一AU相关联。例如,如果`poc_cycle_au`的值等于4,则POC值等于0至3(含)的图片、条带或图块与AUC值等于0的AU相关联,并且POC值等于4至7(含)的图片、条带或图块与AUC值等于1的AU相关联。因此,可以通过将POC值除以`poc_cycle_au`的值来推断AUC的值。

[0133] 在相同或另一实施例中,可以从例如位于视频参数集(VPS)中的,标识已编码视频序列中的空间或SNR层的数目的,信息中导出`poc_cycle_au`的值。下面简要描述这种可能的关系。虽然如上所述的推导可以在VPS中节省一些比特,并且因此可以提高编解码效率,但是在视频参数集之下分层地将`poc_cycle_au`显式编码在适当的高级语法结构中可能是有利的,以便能够针对码流的给定小部分(诸如图片)最小化`poc_cycle_au`。由于可以在低级语法结构中编码POC值(和/或间接参考POC的语法元素的值),因此这种优化可以节省比通过上述推导过程所节省的更多的比特。

[0134] 图8示出了具有自适应分辨率改变的`temporal_id`、`layer_id`、POC和AUC值的组合的视频序列结构的示例。在该示例中,AUC=0的第一AU中的图片、条带或图块可以具有`temporal_id=0`和`layer_id=0`或1,而AUC=1的第二AU中的图片、条带或图块可以分别具有`temporal_id=1`和`layer_id=0`或1。不管`temporal_id`和`layer_id`的值如何,每个图片的POC的值增加1。在该示例中,`poc_cycle_au`的值可以等于2。优选地,可以将`poc_cycle_au`的值设置为等于(空间可缩放性)层的数目。因此,在该示例中,POC的值增加2,而AUC的值增加1。

[0135] 在以上实施例中,可以通过使用HEVC中的现有参考图片集(RPS)信令或参考图片列表(RPL)信令,来支持图片间或层间预测结构和参考图片指示的全部或子集。在RPS或RPL中,通过发信号通知当前图片和所选参考图片之间的POC的值或POC的增量值,来指示所选参考图片。对于所公开的主题,RPS和RPL可以用于指示图片间或层间预测结构而不改变信令,但是具有以下限制。如果参考图片的`temporal_id`的值大于当前图片的`temporal_id`的值,则当前图片可以不使用参考图片进行运动补偿或其它预测。如果参考图片的`layer_id`的值大于当前图片的`layer_id`的值,则当前图片可以不使用参考图片进行运动补偿或其它预测。

[0136] 在相同和其它实施例中,可以在存取单元内的多个图片上禁用基于POC差的用于时间运动矢量预测的运动矢量缩放。因此,尽管每个图片在存取单元内可以具有不同的POC值,但是运动矢量未缩放并且用于存取单元内的时间运动矢量预测。这是因为在相同AU中具有不同POC的参考图片可以认为是具有相同时间实例的参考图片。因此,在该实施例中,

当参考图片属于与当前图片相关联的AU时,运动矢量缩放函数可以返回1。

[0137] 在相同和其它实施例中,当参考图片的空间分辨率不同于当前图片的空间分辨率时,可以在多个图片上可选地禁用基于POC差的用于时间运动矢量预测的运动矢量缩放。当允许运动矢量缩放时,基于当前图片和参考图片之间的POC差和空间分辨率的比值来缩放运动矢量。

[0138] 在相同或另一实施例中,对于时间运动矢量预测,尤其当poc\_cycle\_au具有非均匀值时(当vps\_contant\_poc\_cycle\_per\_au=0时),可以基于AUC差而不是POC差来缩放运动矢量。否则(当vps\_contant\_poc\_cycle\_per\_au=1时),基于AUC差的运动矢量缩放可能与基于POC差的运动矢量缩放相同。

[0139] 在相同或另一实施例中,当基于AUC差对运动矢量进行缩放时,不基于AUC差对与当前图片相同的AU(具有相同的AUC值)中的参考运动矢量进行缩放,并且不进行缩放或基于当前图片和参考图片之间的空间分辨率的比值进行缩放而用于运动矢量预测。

[0140] 在相同和其它实施例中,AUC值用于标识AU的边界并用于假想参考解码器(HRD)的操作,这需要具有AU粒度的输入定时和输出定时。在大多数情况下,可以输出AU中具有最高层的已解码图片进行显示。AUC值和layer\_id值可以用于标识输出图片。

[0141] 在实施例中,图片可以由一个或多个子图片组成。每个子图片可以覆盖图片的局部区域或整个区域。由子图片支持的区域可以与由另一个子图片支持的区域重叠或不重叠。由一个或多个子图片组成的区域可以覆盖或不覆盖图片的整个区域。如果图片由一个子图片组成,则由该子图片所支持的区域与由该图片所支持的区域相同。

[0142] 在相同的实施例中,可以通过与用于已编码图片的编解码方法类似的编解码方法,对子图片进行编码。子图片可以独立编码或者可以根据另一个子图片或已编码图片进行编码。子图片可能具有或不具有对另一个子图片或已编码图片的任何解析依赖性。

[0143] 在相同的实施例中,已编码子图片可以包含在一个或多个层中。层中的已编码子图片可以具有不同的空间分辨率。原始子图片可以在空间上重采样(上采样或下采样),用不同的空间分辨率参数进行编码,并且包含在对应于层的码流中。

[0144] 在相同或另一实施例中,具有(W,H)的子图片可以编码并包含在对应于层0的已编码码流中,其中W指示子图片的宽度,并且H指示子图片的高度。而来自具有原始空间分辨率、具有(W\*Sw,k,H\*Sh,k)的子图片中的,上采样(或下采样)的子图片可以编码并包含在对应于层k的已编码码流中,其中Sw,k、Sh,k指示水平的重采样率和垂直的重采样率。如果Sw,k、Sh,k的值大于1,则重采样等于上采样。然而,如果Sw,k、Sh,k的值小于1,则重采样等于下采样。

[0145] 在相同或另一实施例中,一层中的已编码子图片可以具有与相同子图片或不同子图片中的另一层中的已编码子图片不同的视觉质量。例如,层n中的子图片i用量化参数 $Q_{i,n}$ 进行编码,而层m中的子图片j用量化参数 $Q_{j,m}$ 进行编码。

[0146] 在相同或另一实施例中,一层中的已编码子图片可以是独立可解码的(independently decodable),而对同一局部区域的另一层中的已编码子图片没有任何解析依赖性或解码依赖性。可以在不参考同一局部区域的另一子图片层的情况下独立可解码的子图片层是独立子图片层。独立子图片层中的已编码子图片可以具有或不具有对同一子图片层中的先前已编码子图片的解码依赖性或解析依赖性,但是已编码子图片对来自另一

子图片层中的已编码图片可以不具有任何依赖性。

[0147] 在相同或另一实施例中,一层中的已编码子图片可以是依赖可解码的(dependently decodable),具有对来自相同局部区域的另一层中的已编码子图片的任何解析依赖性或解码依赖性。通过参考同一局部区域的另一子图片层进行依赖解码的子图片层是依赖子图片层。依赖子图片中的已编码子图片可以参考属于相同子图片的已编码子图片、相同子图片层中的先前已编码子图片,或者同时参考这两种参考子图片。

[0148] 在相同或另一实施例中,已编码子图片由一个或多个独立子图片层和一个或多个依赖子图片层组成。然而,对于已编码子图片,可以存在至少一个独立子图片层。独立子图片层可以具有等于0的层标识符(layer\_id)的值,其可以存在于NAL单元头或另一高级语法结构中。layer\_id等于0的子图片层是基本子图片层。

[0149] 在相同或另一实施例中,图片可以由一个或多个前景子图片和一个背景子图片组成。由背景子图片支持的区域可以等于图片的区域。由前景子图片支持的区域可以与由背景子图片支持的区域重叠。背景子图片可以是基本子图片层,而前景子图片可以是非基本(增强)子图片层。一个或多个非基本子图片层可以参考相同的基本层进行解码。layer\_id等于a的每个非基本子图片层可以参考layer\_id等于b的非基本子图片层,其中a大于b。

[0150] 在相同或另一实施例中,图片可以由具有或不具有背景子图片的一个或多个前景子图片组成。每个子图片可以具有其自己的基本子图片层和一个或多个非基本(增强)层。每个基本子图片层可以供一个或多个非基本子图片层进行参考。layer\_id等于a的每个非基本子图片层可以参考layer\_id等于b的非基本子图片层,其中a大于b。

[0151] 在相同或另一实施例中,图片可以由具有或不具有背景子图片的一个或多个前景子图片组成。(基本或非基本)子图片层中的每个已编码子图片可以供属于相同子图片的一个或多个非基本层子图片,和不属于相同子图片的一个或多个非基本层子图片进行参考。

[0152] 在相同或另一实施例中,图片可以由具有或不具有背景子图片的一个或多个前景子图片组成。层a中的子图片可以进一步分区成同一层中的多个子图片。层b中的一个或多个已编码子图片可以参考层a中的已分区的子图片。

[0153] 在相同或另一实施例中,已编码视频序列(CVS)可以是一组已编码图片。CVS可以由一个或多个已编码子图片序列(CSPS)组成,其中CSPS可以是覆盖图片的相同局部区域的一组已编码子图片。CSPS可以具有与已编码视频序列相同或不同的时间分辨率。

[0154] 在相同或另一实施例中,CSPS可以编码并包含在一个或多个层中。CSPS可以由一个或多个CSPS层组成。对与CSPS对应的一个或多个CSPS层进行解码,可以重建对应于相同局部区域的子图片序列。

[0155] 在相同或另一实施例中,对应于一个CSPS的CSPS层的数目可以与对应于另一CSPS的CSPS层的数目相同或不同。

[0156] 在相同或另一实施例中,一个CSPS层可以具有与另一CSPS层不同的时间分辨率(例如帧速率)。原始(未压缩)子图片序列可以在时间上重采样(上采样或下采样),用不同的时间分辨率参数进行编码,并且包含在对应于层的码流中。

[0157] 在相同或另一实施例中,具有帧速率F的子图片序列可以编码并包含在对应于层0的已编码码流中,而来自具有 $F \cdot St, k$ 的原始子图片序列的时间上采样(或下采样)的子图片序列,可以编码并包含在对应于层k的已编码码流中,其中 $St, k$ 指示层k的时间采样率。如果

$St,k$ 的值大于1,则时间重采样过程等于帧速率上变换。然而,如果 $St,k$ 的值小于1,则时间重采样过程等于帧速率下变换。

[0158] 在相同或另一实施例中,当具有CSPS层a的子图片供具有CSPS层b的子图片参考,用于运动补偿或任何层间预测时,如果CSPS层a的空间分辨率不同于CSPS层b的空间分辨率,则CSPS层a中的已解码像素进行重采样并用作参考。重采样过程可能需要上采样滤波或下采样滤波。

[0159] 图9示出了在条带头中的语法表的示例,该语法表用于发信号通知VPS(或SPS)中的语法元素 $vps\_poc\_cycle\_au$ ,该语法元素 $vps\_poc\_cycle\_au$ 指示用于已编码视频序列中的所有图片/条带的 $poc\_cycle\_au$ ,以及该语法表发信号通知语法元素 $slice\_poc\_cycle\_au$ ,语法元素 $slice\_poc\_cycle\_au$ 在条带头中指示当前条带的 $poc\_cycle\_au$ 。如果每个AU的POC值均匀地增加,则将VPS中的 $vps\_contant\_poc\_cycle\_per\_au$ 设置为等于1,并且在VPS中发信号通知 $vps\_poc\_cycle\_au$ 。在这种情况下,不显式地发信号通知表示 $slice\_poc\_cycle\_au$ ,并且通过将POC的值除以 $vps\_poc\_cycle\_au$ 来计算每个AU的AUC值。如果每个AU的POC值不均匀地增加,则将VPS中的 $vps\_contant\_poc\_cycle\_per\_au$ 设置为等于0。在这种情况下,不发信号通知 $vps\_access\_unit\_cnt$ ,而在每个条带或图片的条带头中发信号通知表示 $slice\_access\_unit\_cnt$ 。每个条带或图片可以具有不同的 $slice\_access\_unit\_cnt$ 值。通过将POC的值除以 $slice\_poc\_cycle\_au$ ,来计算每个AU的AUC值。图10示出了图示相关工作流程的框图。

[0160] 在相同或其它实施例中,即使图片、条带或图块的POC的值可能不同,对应于具有相同AUC值的AU的图片、条带或图块,也可以与相同的解码或输出时间实例相关联。因此,在同一AU中的图片、条带或图块之间没有任何相互解析依赖性/解码依赖性的情况下,与同一AU相关联的图片、条带或图块的全部或子集可以并行解码,并且可以在同一时间实例输出。

[0161] 在相同或其它实施例中,即使图片、条带或图块的POC值可能不同,对应于具有相同AUC值的AU的图片、条带或图块,也可以与相同的合成时间实例/显示时间实例相关联。当合成时间包含在容器格式中时,即使图片对应于不同的AU,如果图片具有相同的合成时间,也可以在相同时间实例处显示图片。

[0162] 在相同或其它实施例中,每个图片、条带或图块可以在相同的AU中具有相同的时间标识符( $temporal\_id$ )。对应于一个时间实例的图片、条带或图块的全部或子集可以与相同的时间子层相关联。在相同或其它实施例中,每个图片、条带或图块可以在相同的AU中具有相同或不同的空间层id( $layer\_id$ )。对应于一个时间实例的图片、条带或图块的全部或子集可以与相同或不同的空间层相关联。

[0163] 图11示出了包括 $layer\_id$ 等于0的背景视频CSPS层和多个前景CSPS层的示例视频码流。虽然已编码子图片可以由一个或多个CSPS层组成,但是不属于任何前景CSPS层的背景区域可以包括基本层。基本层可以包含背景区域和前景区域,而增强CSPS层包含前景区域。在相同区域处,增强CSPS层可以具有比基本层更好的视觉质量。增强CSPS层可以参考对应于相同区域的基本层的重建像素和运动矢量。

[0164] 在相同或另一实施例中,在视频文件中,对应于基本层的视频码流包含在轨道中,而对应于每个子图片的CSPS层包含在单独的轨道中。

[0165] 在相同或另一实施例中,对应于基本层的视频码流包含在轨道中,而具有相同

layer\_id的CSPS层包含在单独的轨道中。在该示例中,对应于层k的轨道仅包括对应于层k的CSPS层。

[0166] 在相同或另一实施例中,每个子图片的每个CSPS层存储在单独的轨道中。每个轨道可能具有也可能不具有对一个或多个其它轨道的解析依赖性 or 解码依赖性。

[0167] 在相同或另一实施例中,每个轨道可以包含对应于子图片的全部或子集的CSPS层中的层i至层j的码流,其中 $0 < i \leq j \leq k$ ,k是CSPS的最高层。

[0168] 在相同或另一实施例中,图片由一个或多个相关联的媒体数据组成,这些相关联的媒体数据包括深度图、 $\alpha$ 图、3D几何数据、占用图等。这种相关联的定时媒体数据可以划分成一个或多个数据子码流,每个数据子码流对应于一个子图片。

[0169] 在相同或另一实施例中,图12示出了基于多层子图片方法的视频会议的示例。在视频码流中,包含对应于背景图片的一个基本层视频码流和对应于前景子图片的一个或多个增强层视频码流。每个增强层视频码流对应于CSPS层。在显示器中,默认地显示对应于基本层的图片。它包含一个或多个用户的画中画(PIP)。当客户端控件选择特定用户时,对应于所选用户的增强CSPS层被解码并以增强的质量或空间分辨率显示。图13示出了操作示意图。

[0170] 在相同或另一实施例中,网络中间盒(诸如路由器)可以根据其带宽选择层的子集以发送给用户。图片/子图片组织可以用于带宽适配。例如,如果用户没有带宽,由于层和子图片的重要性或基于使用的设置,路由器剥离层或选择一些子图片,这可以动态地进行,以适应带宽。

[0171] 图14示出了360视频的用例。当球形360图片投影到平面图片上时,投影360图片可以分区成多个子图片作为基本层。对特定子图片的增强层进行编码并将其传输到客户端。解码器能够对包括所有子图片的基本层进行解码和对所选子图片的增强层进行解码。当当前视口与所选子图片相同时,所显示图片包括具有增强层的已解码子图片,所显示图片具有更高的质量。否则,可以低质量显示具有基本层的已解码图片。

[0172] 在相同或另一实施例中,用于显示的任何布局信息可以作为补充信息(诸如SEI消息或元数据)存在于文件中。一个或多个已解码子图片可以根据发信号通知表示的布局信息重新定位和显示。布局信息可以由流媒体服务器或广播器发信号通知表示,或者可以由网络实体或云服务器重新生成,或者可以由用户的定制设置来确定。

[0173] 在实施例中,当输入图片划分成一个或多个(矩形)子区域时,每个子区域可以被编码为独立层。对应于局部区域的每个独立层可以具有唯一的layer\_id值。对于每个独立层,可以发信号通知表示子图片大小和位置信息。例如,图片大小(宽度,高度)、左上角的偏移信息(x\_offset,y\_offset)。图15示出了划分的子图片的布局、其子图片大小和位置信息及子图片对应的图片预测结构的示例。在高级语法结构(诸如参数集、条带头或图块组头,或SEI消息)中发信号通知表示布局信息,该布局信息包括子图片大小和子图片位置。

[0174] 在相同的实施例中,对应于独立层的每个子图片可以在AU内具有其唯一的POC值。当通过使用RPS或RPL结构中的一个或多个语法元素,指示存储在DPB中的图片中的参考图片时,可以使用对应于层的每个子图片的一个或多个POC值。

[0175] 在相同或另一实施例中,为了指示(层间)预测结构,可以不使用layer\_id并且可以使用POC(增量)值。

[0176] 在相同的实施例中,具有对应于层(或局部区域)的POC值等于N的子图片,可以用作或可以不用作具有POC值等于N+K的子图片的参考图片,POC值等于N+K的子图片对应于相同层(或相同局部区域),参考图片用于运动补偿预测。在大多数情况下,数目K的值可以等于(独立)层的最大数目,数目K的值可以与子区域的数目相同。

[0177] 在相同或另一实施例中,图16示出了图15的扩展情况。当输入图片划分成多个(例如四个)子区域时,每个局部区域可以用一个或多个层进行编码。在该情况下,独立层的数目可以等于子区域的数目,并且一个或多个层可以对应于一个子区域。因此,每个子区域可以用一个或多个独立层以及零个或多个依赖层进行编码。

[0178] 在相同的实施例中,在图16中,输入图片可以划分成四个子区域。右上子区域可以编码为两个层,即层1和层4,而右下子区域可以被编码为两个层,即层3和层5。在这种情况下,层4可以参考层1进行运动补偿预测,而层5可以参考层3进行运动补偿。

[0179] 在相同或另一实施例中,可以(可选地)禁用跨层边界的环路滤波(诸如去块滤波、自适应环路滤波、整形器、双边滤波或任何基于深度学习的滤波)。

[0180] 在相同或另一实施例中,可以(可选地)禁用跨层边界的运动补偿预测或块内复制。

[0181] 在相同或另一实施例中,可以可选地进行用于运动补偿预测的边界填充或子图片边界处的环路滤波。在高级语法结构(诸如一个或多个参数集(VPS、SPS、PPS或APS)、条带头或图块组头,或SEI消息)中发信号通知指示是否进行边界填充的标志。

[0182] 在相同或另一实施例中,在VPS或SPS中,发信号通知一个或多个子区域(或者一个或多个子图片)的布局信息。图17示出了VPS和SPS中的语法元素的示例。在该示例中,在VPS中发信号通知`vps_sub_picture_dividing_flag`。该标志可以指示一个或多个输入图片是否被划分成多个子区域。当`vps_sub_picture_dividing_flag`的值等于0时,对应于当前VPS的一个或多个已编码视频序列中的一个或多个输入图片,可以不划分成多个子区域。在这种情况下,输入图片大小可以等于已编码图片大小(`pic_width_in_luma_samples`,`pic_height_in_luma_samples`),在SPS中被发信号通知输入图片大小。当`vps_sub_picture_dividing_flag`的值等于1时,一个或多个输入图片可以被划分成多个子区域。在这种情况下,在VPS中发信号通知语法元素`vps_full_pic_width_in_luma_samples`和`vps_full_pic_height_in_luma_samples`。`vps_full_pic_width_in_luma_samples`和`vps_full_pic_height_in_luma_samples`的值可以分别等于一个或多个输入图片的宽度和高度。

[0183] 在相同的实施例中,`vps_full_pic_width_in_luma_samples`和`vps_full_pic_height_in_luma_samples`的值可以不用于解码,但可以用于合成和显示。

[0184] 在相同的实施例中,当`vps_sub_picture_dividing_flag`的值等于1时,可以在对应于一个或多个特定层的SPS中发信号通知语法元素`pic_offset_x`和`pic_offset_y`。在这种情况下,在SPS中发信号通知的已编码图片大小(`pic_width_in_luma_samples`,`pic_height_in_luma_samples`),可以等于对应于特定层的子区域的宽度和高度。而且,可以在SPS中发信号通知子区域的左上角的位置(`pic_offset_x`,`pic_offset_y`)。

[0185] 在相同的实施例中,子区域的左上角的位置信息(`pic_offset_x`,`pic_offset_y`)可以不用于解码,但可以用于合成和显示。

[0186] 在相同或另一实施例中,在参数集或SEI消息中,发信号通知输入图片、层间的依

赖性信息的全部或子集子区域的布局信息(大小和位置)。图18示出了语法元素的示例,其用于指示子区域的布局信息、层之间的依赖性信息以及子区域与一个或多个层之间的关系的信息。在该示例中,语法元素num\_sub\_region指示当前已编码视频序列中(矩形)子区域的数目。语法元素num\_layers指示当前已编码视频序列中层的数目。num\_layers的值可以等于或大于num\_sub\_region的值。当任何子区域被编码为单个层时,num\_layers的值可以等于num\_sub\_region的值。当一个或多个子区域编码为多个层时,num\_layers的值可以大于num\_sub\_region的值。语法元素direct\_dependency\_flag[i][j]指示从第j层到第i层的依赖性。num\_layers\_for\_region[i]指示与第i个子区域相关联的层的数目。sub\_region\_layer\_id[i][j]指示与第i个子区域相关联的第j层的layer\_id。sub\_region\_offset\_x[i]和sub\_region\_offset\_y[i]分别指示第i个子区域的左上角的水平位置和垂直位置。sub\_region\_width[i]和sub\_region\_height[i]分别指示第i个子区域的宽度和高度。

[0187] 在一个实施例中,在高级语法结构(例如VPS、DPS、SPS、PPS、APS或SEI消息)中,发信号通知一个或多个语法元素。所述一个或多个语法元素指定输出层集,以指示要输出的具有或不具有配置文件层级信息的多个层中的一个。参考图19,在VPS中发信号通知语法元素num\_output\_layer\_sets,其指示参考VPS的已编码视频序列中的输出层集(OLS)的数目。对于每个输出层集,可以发信号通知与输出层的数目一样多的output\_layer\_flag。

[0188] 在相同的实施例中,等于1的output\_layer\_flag[i]指定输出第i层。等于0的vps\_output\_layer\_flag[i]指定不输出第i层。

[0189] 在相同或另一实施例中,可以在高级语法结构(例如VPS、DPS、SPS、PPS、APS或SEI消息)中发信号通知一个或多个语法元素,其指定用于每个输出层集的配置文件层级信息。仍然参考图19,可以在VPS中发信号通知语法元素num\_profile\_tile\_level,其指示参考VPS的已编码视频序列中的每个OLS的配置文件层级信息的数目。对于每个输出层集,以与输出层的数目相同的数目,发信号通知用于配置文件层级信息的语法元素集,或发信号通知指示配置文件层级信息中的条目中的特定配置文件层级信息的索引。

[0190] 在相同的实施例中,profile\_tier\_level\_idx[i][j],在VPS中的profile\_tier\_level()语法结构的列表中,指定应用于第i个OLS的第j层的profile\_tier\_level()语法结构的索引。

[0191] 在相同或另一实施例中,参考图20,当最大层的数目大于1(vps\_max\_layers\_minus1>0)时,可以发信号通知语法元素num\_profile\_tile\_level和/或num\_output\_layer\_sets。

[0192] 在相同或另一实施例中,参考图20,语法元素vps\_output\_layers\_mode[i]可以存在于VPS中,其指示用于第i个输出层集的输出层信令的模式。

[0193] 在相同的实施例中,等于0的vps\_output\_layers\_mode[i]指定:仅输出具有第i个输出层集的最高层。等于1的vps\_output\_layer\_mode[i]指定:输出具有第i个输出层集的所有层。等于2的vps\_output\_layer\_mode[i]指定:输出的层是vps\_output\_layer\_flag[i][j]等于1且具有第i个输出层集的层。可以保留更多的值。

[0194] 在相同的实施例中,根据用于第i个输出层集的vps\_output\_layers\_mode[i]的值来发信号通知也可以不发信号通知output\_layer\_flag[i][j]。

[0195] 在相同或另一实施例中,参考图20,存在用于第i个输出层集的标志vps\_ptl\_

signal\_flag[i]。取决于vps\_ptl\_signal\_flag[i]的值,可以发信号通知或不发信号通知用于第i个输出层集的配置文件的层级信息。

[0196] 在相同或另一实施例中,参考图21,可以在高级语法结构(例如VPS、DPS、SPS、PPS、APS或SEI消息)中发信号通知当前CVS中的子图片的数目max\_subpics\_minus1。

[0197] 在相同的实施例中,参考图21,当子图片的数目大于1(max\_subpics\_minus1>0)时,可以发信号通知用于第i个子图片的子图片标识符sub\_pic\_id[i]。

[0198] 在相同或另一实施例中,在VPS中,发信号通知一个或多个语法元素,其指示属于每个输出层集的每个层的子图片标识符。参考图21,sub\_pic\_id\_layer[i][j][k]指示存在于第i个输出层集的第j层中的第k个子图片。使用这些信息,解码器可以识别,对于特定输出层集的每个层,可以对哪一个子图片进行解码和输出。

[0199] 在实施例中,图片头(PH)是包含应用于已编码图片的所有条带的语法元素的语法结构。图片单元(PU)是一组NAL单元,该组NAL单元根据指定的分类规则彼此关联、在解码顺序上连续并且正好包含一个已编码图片的。PU可以包含图片头(PH)和组成已编码图片的一个或多个VCL NAL单元。

[0200] 在实施例中,SPS(RBSP)在其被参考之前可用于解码过程,包括在TemporalId等于0的至少一个AU中,或通过外部手段提供。

[0201] 在实施例中,SPS(RBSP)在其被参考之前可用于解码过程,包括在CVS(包含参考SPS的一个或多个PPS)中TemporalId等于0的至少一个AU中,或通过外部手段提供。

[0202] 在实施例中,SPS(RBSP)在其由一个或多个PPS参考之前可用于解码过程,包括在至少一个PU中,或通过外部手段提供。其中,所述至少一个PU的nuh\_layer\_id等于CVS(包含参考SPS的一个或多个PPS)中参考SPS NAL单元的PPS NAL单元的最低nuh\_layer\_id值。

[0203] 在实施例中,SPS(RBSP)在其由一个或多个PPS参考之前可用于解码过程,包括在至少一个PU中或通过外部手段提供。所述至少一个PU的TemporalId equal等于0,且nuh\_layer\_id等于参考SPS NAL单元的PPS NAL单元的最低nuh\_layer\_id值。

[0204] 在实施例中,SPS(RBSP)在其由一个或多个PPS参考之前可用于解码过程,包括在至少一个PU中或通过外部手段提供。所述至少一个PU的TemporalId等于0,且nuh\_layer\_id等于CVS(包含参考SPS的一个或多个PPS)中参考SPS NAL单元的PPS NAL单元的最低nuh\_layer\_id值。

[0205] 在相同或另一实施例中,pps\_seq\_parameter\_set\_id为所参考SPS指定sps\_seq\_parameter\_set\_id的值。pps\_seq\_parameter\_set\_id的值在由CVS中的已编码图片参考的所有PPS中可以相同。

[0206] 在相同或另一实施例中,CVS中具有特定sps\_seq\_parameter\_set\_id值的所有SPS NAL单元可以具有相同内容。

[0207] 在相同或另一实施例中,不管nuh\_layer\_id值如何,SPS NAL单元可以共享sps\_seq\_parameter\_set\_id的相同值空间。

[0208] 在相同或另一实施例中,SPS NAL单元的nuh\_layer\_id值可以等于参考SPS NAL单元的PPS NAL单元的最低nuh\_layer\_id值。

[0209] 在实施例中,当nuh\_layer\_id等于m的SPS被nuh\_layer\_id等于n的一个或多个PPS参考时,nuh\_layer\_id等于m的层可以与nuh\_layer\_id等于n的层相同,或可以与nuh\_

layer\_id等于m的层的(直接或间接)参考层相同。

[0210] 在实施例中,PPS(RBSP)在其被参考之前可用于解码过程,包括在TemporalId等于PPS NAL单元的TemporalId的至少一个AU中,或通过外部手段提供。

[0211] 在实施例中,PPS(RBSP)在其被参考之前可用于解码过程,包括在至少一个AU中或通过外部手段提供。所述至少一个AU的TemporalId等于CVS(其包含参考PPS的一个或多个PH(或已编码条带NAL单元))中的PPS NAL单元的TemporalId。

[0212] 在实施例中,PPS(RBSP)在其由一个或多个PH(或已编码条带NAL单元)参考之前可用于解码过程,包括在至少一个PU中或通过外部手段提供。所述至少一个PU的nuh\_layer\_id等于,CVS(其包含参考PPS的一个或多个PH(或已编码条带NAL单元))中参考PPS NAL单元的已编码条带NAL单元的最低nuh\_layer\_id值。

[0213] 在实施例中,PPS(RBSP)在其由一个或多个PH(或已编码条带NAL单元)参考之前可用于解码过程,包括在至少一个PU中或通过外部手段提供。所述至少一个PU的TemporalId等于PPS NAL单元的TemporalId,且nuh\_layer\_id等于CVS(包含参考PPS的一个或多个PH(或已编码条带NAL单元))中参考PPS NAL单元的已编码条带NAL单元的最低nuh\_layer\_id值。

[0214] 在相同或另一实施例中,PH中的ph\_pic\_parameter\_set\_id指定所使用的参考PPS的pps\_pic\_parameter\_set\_id的值。pps\_seq\_parameter\_set\_id的值在由CLVS中的已编码图片参考的所有PPS中可以相同。

[0215] 在相同或另一实施例中,PU内具有特定pps\_pic\_parameter\_set\_id值的所有PPS NAL单元应具有相同内容。

[0216] 在相同或另一实施例中,不管nuh\_layer\_id值如何,PPS NAL单元可以共享pps\_pic\_parameter\_set\_id的相同值空间。

[0217] 在相同或另一实施例中,PPS NAL单元的nuh\_layer\_id值可以等于参考NAL单元(其参考PPS NAL单元)的已编码条带NAL单元的最低nuh\_layer\_id值。

[0218] 在实施例中,当nuh\_layer\_id等于m的PPS被nuh\_layer\_id等于n的一个或多个已编码条带NAL单元参考时,nuh\_layer\_id等于m的层,可以与nuh\_layer\_id等于n的层相同,或可以与nuh\_layer\_id等于m的层的(直接或间接)参考层相同。

[0219] 在实施例中,PPS(RBSP)在其被参考之前可用于解码过程,包括在至少一个AU中,或通过外部手段提供。所述至少一个AU的TemporalId等于PPS NAL单元的TemporalId。

[0220] 在实施例中,PPS(RBSP)在其被参考之前可用于解码过程,包括在至少一个AU中或通过外部手段提供。所述至少一个AU的TemporalId等于CVS(其包含参考PPS的一个或多个PH(或已编码条带NAL单元))中的PPS NAL单元的TemporalId。

[0221] 在实施例中,PPS(RBSP)在其由一个或多个PH(或已编码条带NAL单元)参考之前可用于解码过程,包括在至少一个PU中或通过外部手段提供。所述至少一个PU的nuh\_layer\_id等于CVS(其包含参考PPS的一个或多个PH(或已编码条带NAL单元))中参考PPS NAL单元的已编码条带NAL单元的最低nuh\_layer\_id值。

[0222] 在实施例中,PPS(RBSP)在其由一个或多个PH(或已编码条带NAL单元)参考之前可用于解码过程,包括在至少一个PU中,或通过外部手段提供。所述至少一个PU的TemporalId等于PPS NAL单元的TemporalId,且nuh\_layer\_id等于CVS(包含参考PPS的一个或多个PH

(或已编码条带NAL单元))中参考PPS NAL单元的已编码条带NAL单元的最低nuh\_layer\_id值。

[0223] 在相同或另一实施例中,PH中的ph\_pic\_parameter\_set\_id为所参考PPS指定pps\_pic\_parameter\_set\_id的值。pps\_seq\_parameter\_set\_id的值,在由CLVS中的已编码图片参考的所有PPS中,可以相同。

[0224] 在相同或另一实施例中,PU内具有特定pps\_pic\_parameter\_set\_id值的所有PPS NAL单元,应具有相同内容。

[0225] 在相同或另一实施例中,不管nuh\_layer\_id值如何,PPS NAL单元可以共享pps\_pic\_parameter\_set\_id的相同值空间。

[0226] 在相同或另一实施例中,PPS NAL单元的nuh\_layer\_id值可以等于参考NAL单元(其参考PPS NAL单元)的已编码条带NAL单元的最低nuh\_layer\_id值。

[0227] 在实施例中,当nuh\_layer\_id等于m的PPS被nuh\_layer\_id等于n的一个或多个已编码条带NAL单元参考时,nuh\_layer\_id等于m的层可以与nuh\_layer\_id等于n的层相同,或可以与nuh\_layer\_id等于m的层的(直接或间接)参考层相同。

[0228] 在实施例中,当在DPS、VPS或SPS中发信号通知标志no\_temporal\_sublayer\_switching\_flag时,参考包含等于1的的标志的参数集的PPS的TemporalId值可以等于0,而参考包含等于1的的标志的参数集的PPS的TemporalId值可以等于或大于参数集的TemporalId值。

[0229] 在实施例中,每个PPS(RBSP)在其被参考之前可用于解码过程,包括在至少一个AU中或通过外部手段提供。其中,所述至少一个AU的TemporalId小于或等于参考所述至少一个AU的的已编码条带NAL单元(或PH NAL单元)的TemporalId,。当PPS NAL单元包括在,包含参考PPS的已编码条带NAL单元的AU之前的AU中时,启用时间上层切换的VCL NAL单元,或者nal\_unit\_type等于STSA\_NUT的VCL NAL单元,其指示VCL NAL单元中的图片可以是逐级时间子层存取(STSA)图片,在PPS NAL单元之后和在参考APS的已编码条带NAL单元之前可能不存在。

[0230] 在相同或另一实施例中,参考PPS的PPS NAL单元和已编码条带NAL单元(及其PH NAL单元)可以包括在同一AU中。

[0231] 在相同或另一实施例中,PPS NAL单元和STSANAL单元可以包括在同一AU中,该AU在参考PPS的已编码条带NAL单元(及其PH NAL单元)之前。

[0232] 在相同或另一实施例中,参考PPS的STSA NAL单元、PPS NAL单元和已编码条带NAL单元(及其PH NAL单元)可以存在于同一AU中。

[0233] 在相同的实施例中,包含PPS的VCL NAL单元的TemporalId值可以等于先前STSA NAL单元的TemporalId值。

[0234] 在相同的实施例中,PPS NAL单元的图片顺序计数(POC)值可以等于或大于STSA NAL单元的POC值。

[0235] 在相同的实施例中,参考PPS NAL单元的已编码条带或参考PPS NAL单元的PH NAL单元的,图片顺序计数(POC)值,可以等于或大于参考PPS NAL单元的POC值。

[0236] 在实施例中,sps\_max\_sublayer\_minus1的值在已编码视频序列中的所有层上应相同,因为AU中的所有VCL NAL单元应具有相同的TemporalId值。sps\_max\_sublayers\_

minus1的值在由CVS中的已编码图片参考的所有SPS中应相同。

[0237] 在实施例中,由层A中的一个或多个已编码图片参考的SPS的chroma\_format\_idc值,应等于由层B中的一个或多个已编码图片参考的SPS中的chroma\_format\_idc值,其中层A是层B的直接参考层。这是因为任何已编码图片应该具有与其参考图片相同的chroma\_format\_idc值。在CVS中,由层A中的一个或多个已编码图片参考的SPS的chroma\_format\_idc值,应等于由层A的直接参考层中的一个或多个已编码图片参考的SPS中的chroma\_format\_idc值。

[0238] 在实施例中,由层A中的一个或多个已编码图片参考的SPS的subpics\_present\_flag和sps\_subpic\_id\_present\_flag值,应等于由层B中的一个或多个已编码图片参考的SPS中的subpics\_present\_flag和sps\_subpic\_id\_present\_flag值,其中层A是层B的直接参考层。这是因为子图片布局需要跨层对准或跨层相关联。否则,可能无法正确提取具有多层的子图片。在CVS中,由层A中的一个或多个已编码图片参考的SPS的subpics\_present\_flag和sps\_subpic\_id\_present\_flag值,应等于由层A的直接参考层中的一个或多个已编码图片参考的SPS中的subpics\_present\_flag和sps\_subpic\_id\_present\_flag值。

[0239] 在实施例中,当层A中的STSA图片被在相同AU中的层A的直接参考层中的图片参考时,参考STSA的图片应为STSA图片。否则,时间子层向上切换不能跨层同步。当层A中的STSA NAL单元被在相同AU中的层A的直接参考层中的VCL NAL单元参考时,参考STSA NAL单元的VCL NAL单元的nal\_unit\_type值应等于STSA\_NUT。

[0240] 在实施例中,当层A中的RASL图片被在相同AU中的层A的直接参考层中的图片参考时,参考RASL的图片应为RASL图片。否则,无法正确解码图片。当层A中的RASL NAL单元被在相同AU中的层A的直接参考层中的VCL NAL单元参考时,参考RASL NAL单元的VCL NAL单元的nal\_unit\_type值应等于RASL\_NUT。

[0241] 根据本申请的实施例,提供一种视频解码的方法,包括:接收具有多个层的视频数据;识别一个或多个第一语法元素,所述一个或多个第一语法元素指定一个或多个输出层集,所述一个或多个输出层集对应于所述接收的视频数据的多个层中的一个或多个输出层;以及解码与指定的所述一个或多个输出层集对应的所述一个或多个输出层。

[0242] 在一些实施例中,在高级语法结构中发信号通知所述一个或多个第一语法元素。所述高级语法结构包括以下之一:视频参数集、依赖性参数集、序列参数集、图片参数集、自适应参数集和补充增强信息消息。

[0243] 在一些实施例中,上述方法进一步包括:指定所述一个或多个输出层集的数目;以及通过设置输出层标志,来发信号通知输出层的数目。

[0244] 在一些实施例中,上述方法进一步包括:识别一个或多个第二语法元素,所述一个或多个第二语法元素为所述一个或多个输出层集中的每一个输出层集,指定配置文件层级信息。

[0245] 在一些实施例中,在高级语法结构中发信号通知所述一个或多个第二语法元素。所述高级语法结构包括以下之一:视频参数集、依赖性参数集、序列参数集、图片参数集、自适应参数集和补充增强信息消息。

[0246] 在一些实施例中,上述方法进一步包括:

[0247] 为所述一个或多个输出层集中的每一个输出层集,指定所述配置文件层级信息的

数目;以及

[0248] 对于所述一个或多个输出层集中的每一个输出层集,发信号通知与配置文件层级信息对应的一组语法元素,或者发信号通知一个索引,该索引用于指示配置文件层级信息中的一个配置文件层级信息条目。

[0249] 在一些实施例中,上述方法进一步包括发信号通知一个或多个第三语法元素,所述一个或多个第三语法元素指示子图片标识符,所述子图片标识符对应于与所述一个或多个输出层集相关联的每个输出层。

[0250] 在一些实施例中,在高级语法结构中发信号通知所述子图片标识符,所述高级语法结构包括以下之一:视频参数集、依赖性参数集、序列参数集、图片参数集、自适应参数集和补充增强信息消息。

[0251] 根据本申请的实施例,提供一种视频解码的计算机系统,所述计算机系统包括:

[0252] 一个或多个计算机可读非易失性存储介质,被配置为存储计算机程序代码;以及

[0253] 一个或多个计算机处理器,被配置为存取所述计算机程序代码并按照所述计算机程序代码的指令进行操作,以执行视频解码的方法。

[0254] 根据本申请的实施例,提供一种视频解码的装置,包括:

[0255] 接收模块,被配置为使所述一个或多个计算机处理器接收具有多个层的视频数据;

[0256] 第一标识模块,被配置为使所述一个或多个计算机处理器识别一个或多个第一语法元素,所述一个或多个第一语法元素指定一个或多个输出层集,所述一个或多个输出层集对应于来自接收的视频数据的所述多个层中的一个或多个输出层;以及

[0257] 解码和显示模块,被配置为使所述一个或多个计算机处理器,解码指定的所述一个或多个输出层集对应的所述一个或多个输出层。

[0258] 在一些实施例中,上述装置进一步包括:

[0259] 第一指定模块,被配置为使所述一个或多个计算机处理器指定所述一个或多个输出层集的数目;以及

[0260] 第一信令模块,被配置为使所述一个或多个计算机处理器通过设置输出层标志来发信号通知输出层的数目。

[0261] 在一些实施例中,上述装置进一步包括第二标识模块,其被配置为使所述一个或多个计算机处理器识别一个或多个第二语法元素,所述一个或多个第二语法元素为所述一个或多个输出层集中的每一个输出层集,指定配置文件层级信息。

[0262] 在一些实施例中,上述装置进一步包括:

[0263] 第二指定模块,被配置为使所述一个或多个计算机处理器为所述一个或多个输出层集中的每一个输出层集,指定所述配置文件层级信息的数目;以及

[0264] 第二信令模块,被配置为使所述一个或多个计算机处理器针对所述一个或多个输出层集中的每一个输出层集,发信号通知对应于配置文件层级信息的一组语法元素,或者发信号通知一个索引,该索引用于指示所述配置文件层级信息中的一个配置文件层级信息条目。

[0265] 在一些实施例中,上述装置进一步包括第三信令模块,其被配置为使所述一个或多个计算机处理器发信号通知一个或多个第三语法元素,所述一个或多个第三语法元素指

示子图片标识符,所述子图片标识符对应于与所述一个或多个输出层集相关联的每个输出层。

[0266] 根据本申请的实施例,提供一种非易失性计算机可读介质,其上存储有视频解码的计算机程序,所述计算机程序被配置为使一个或多个计算机处理器:执行视频解码的方法。

[0267] 虽然本公开已经描述了若干示例性实施例,但是存在属于本公开的范围内的改变、变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,这些系统和方法虽然未在本文中明确示出或描述,但体现本公开的原理,并且因此在本公开的精神和范围内。

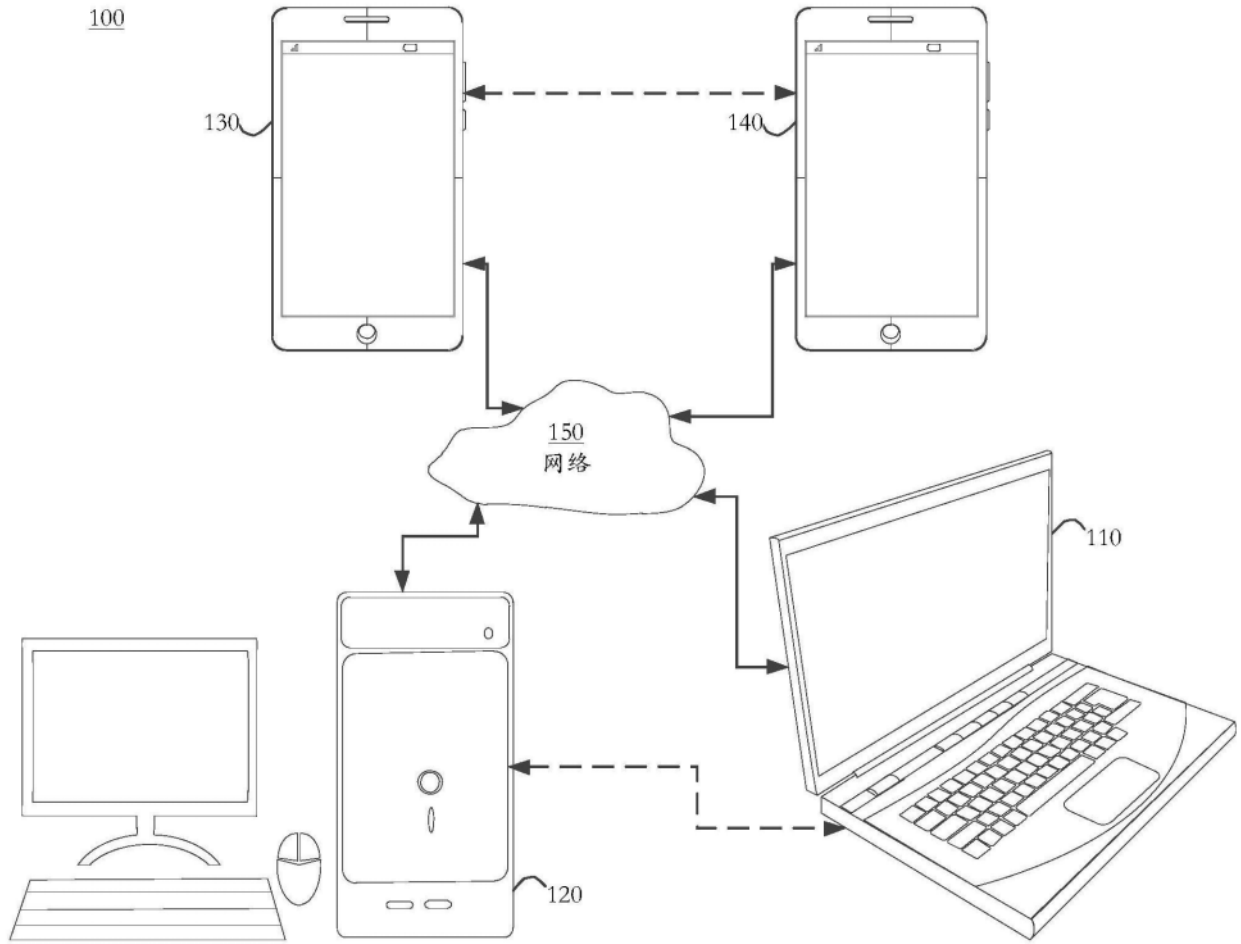


图1

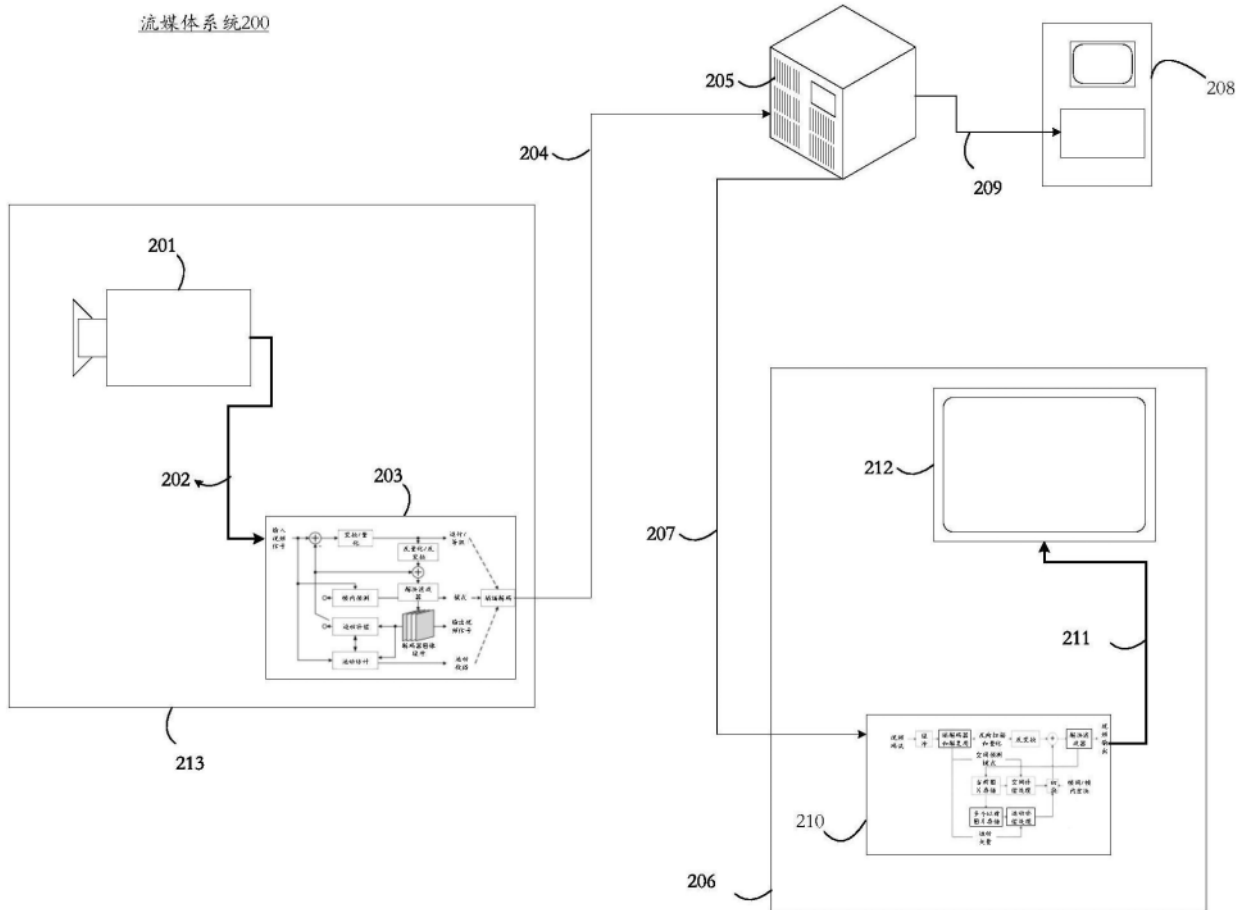


图2

解码器 210

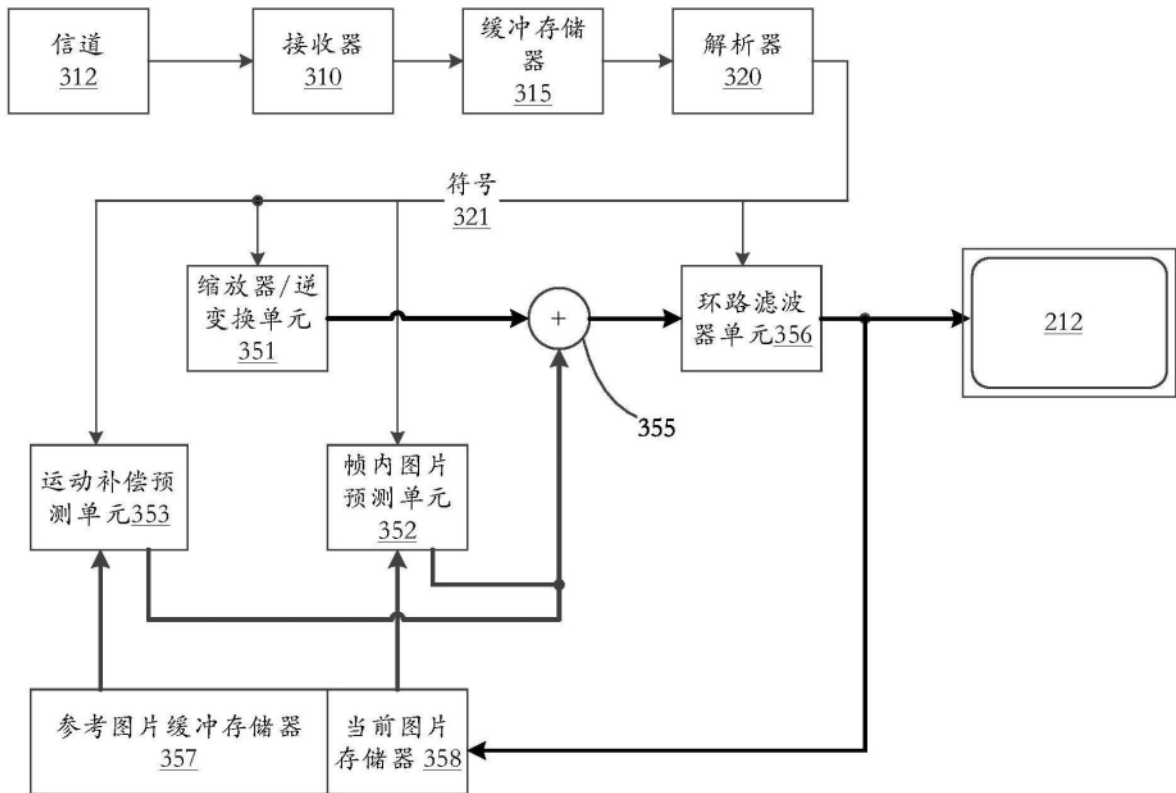


图3

视频编码器 203

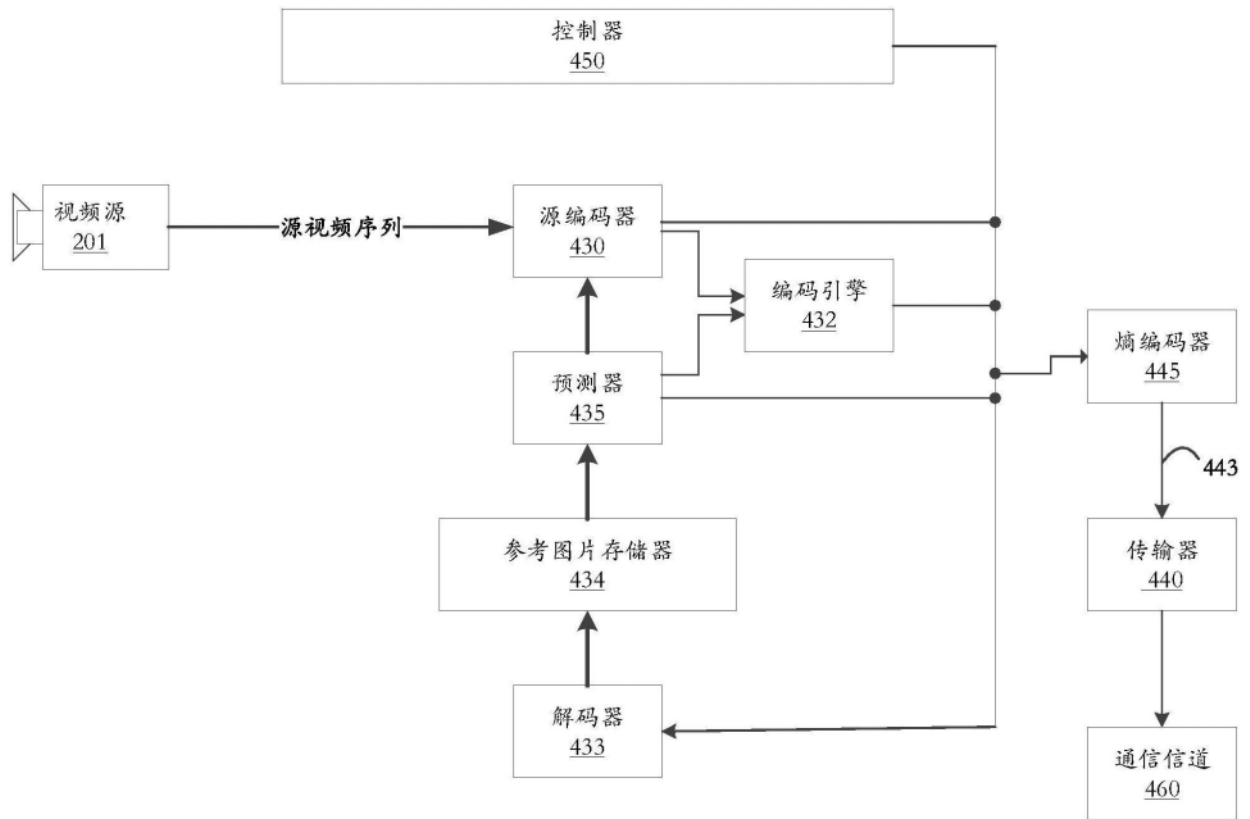


图4

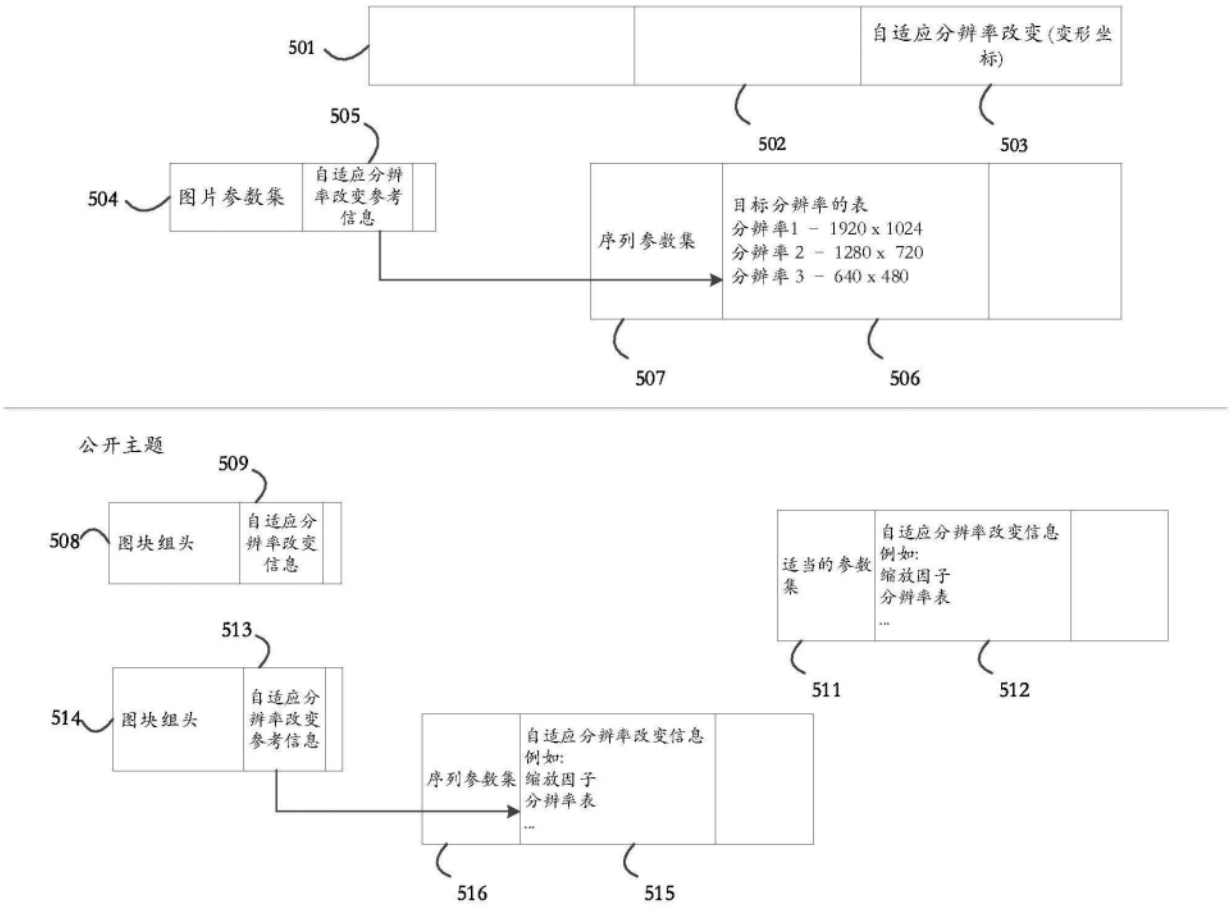


图5

601	tile_group_header() {	
	...	
603	if(adaptive_pic_resolution_change_flag) {	
602	dec_pic_size_idx	u(1)
	}	
	...	
	}	

		描述符
	seq_parameter_set_rbsp() {	
	...	
610	<b>adaptive_pic_resolution_change_flag</b>	u(1)
611	if(adaptive_pic_resolution_change_flag) {	
612	<b>output_pic_width_in_luma_samples</b>	ue(v)
613	<b>output_pic_height_in_luma_samples</b>	ue(v)
614	<b>reference_pic_size_present_flag</b>	u(1)
	if(reference_pic_size_present_flag)	
	{	
615	<b>reference_pic_width_in_luma_samples</b>	ue(v)
	<b>reference_pic_height_in_luma_samples</b>	ue(v)
616	}	
	<b>num_dec_pic_size_in_luma_samples_minus1</b>	ue(v)
617	for(i = 0; i <= num_dec_pic_size_in_luma_samples_minus1; i++) {	
	<b>dec_pic_width_in_luma_samples[ i ]</b>	ue(v)
	<b>dec_pic_height_in_luma_samples[ i ]</b>	ue(v)
	}	
	}	
	...	
	}	

图6

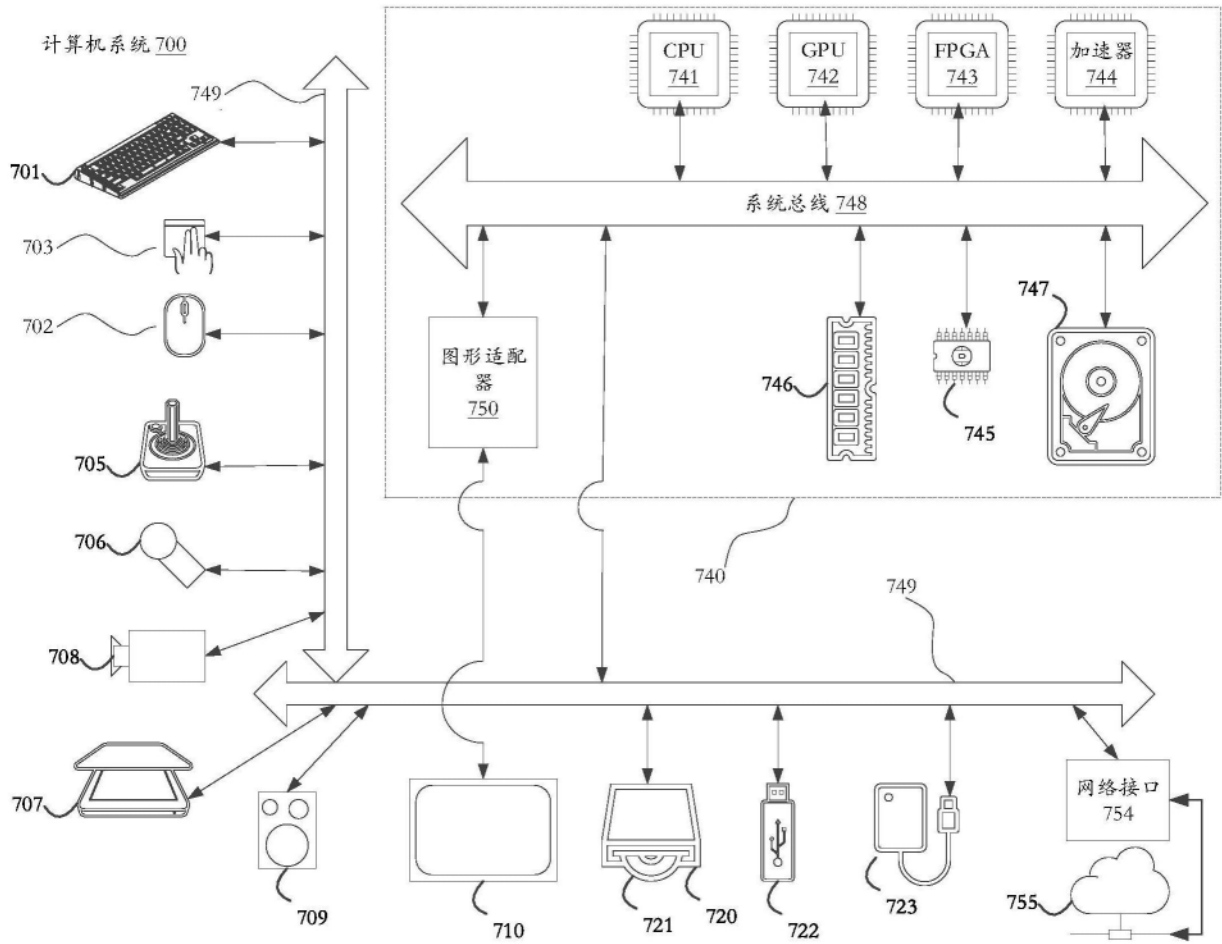


图7

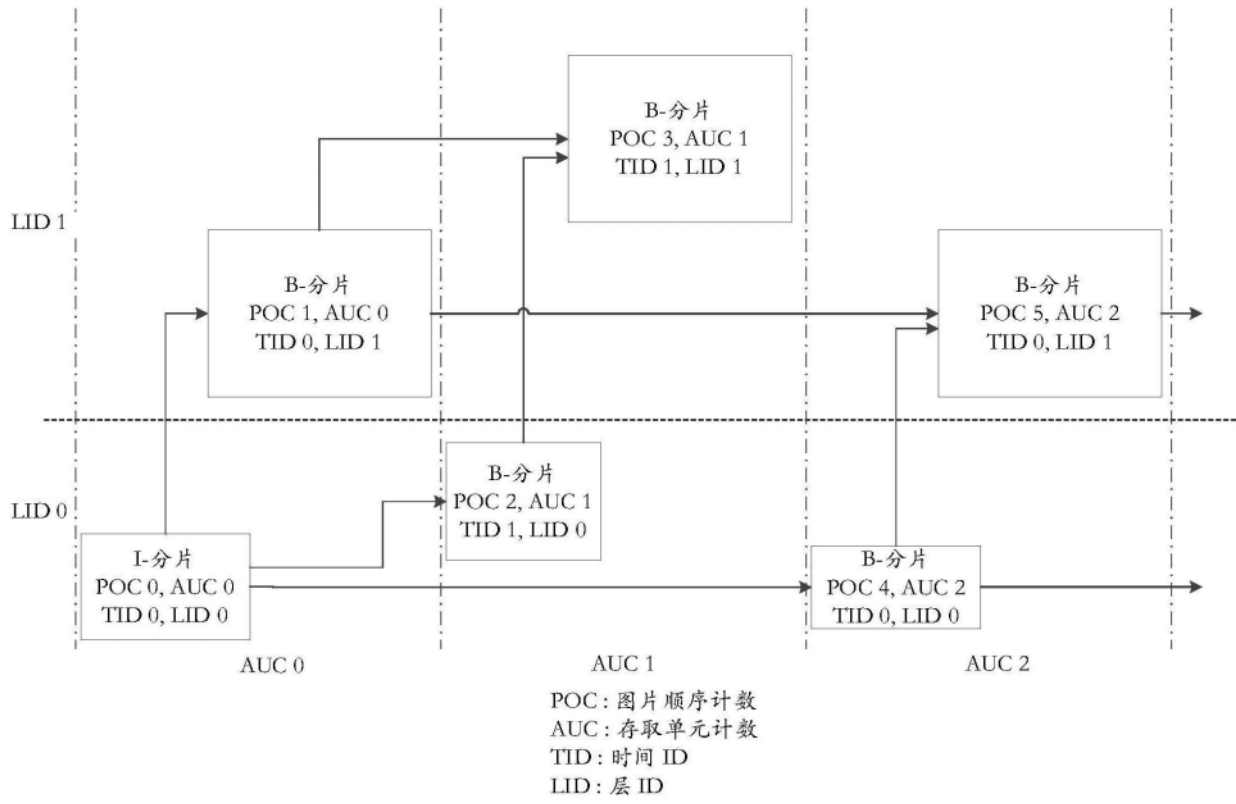


图8

<code>video_parameter_set_rbsp() {</code>	描述符
<code>  vps_video_parameter_set_id</code>	<code>u(4)</code>
<code>  vps_max_layers_minus1</code>	<code>u(8)</code>
<code>  for( i = 0; i &lt;= vps_max_layers_minus1; i++) {</code>	
<code>    vps_included_layer_id[ i ]</code>	<code>u(7)</code>
<code>    vps_reserved_zero_bit</code>	<code>u(1)</code>
<code>  }</code>	
<code>  vps_constraint_info_present_flag</code>	<code>u(1)</code>
<code>  vps_constant_poc_cycle_per_au</code>	<code>u(1)</code>
<code>  if(vps_constant_poc_cycle_per_au)</code>	
<code>    vps_poc_cycle_au</code>	<code>u(8)</code>
<code>  ... </code>	
<code>}</code>	
<code>  </code>	
<code>slice_header() {</code>	描述符
<code>  slice_pic_parameter_set_id</code>	<code>ue(v)</code>
<code>  if( rect_slice_flag    NumBricksInPic &gt; 1 )</code>	
<code>    slice_address</code>	<code>u(v)</code>
<code>  if( !rect_slice_flag &amp;&amp; !single_brick_per_slice_flag )</code>	
<code>    num_bricks_in_slice_minus1</code>	<code>ue(v)</code>
<code>  slice_type</code>	<code>ue(v)</code>
<code>  if( NalUnitType == GRA_NUT )</code>	
<code>    recovery_poc_cnt</code>	<code>se(v)</code>
<code>  slice_pic_order_cnt_lsb</code>	<code>u(v)</code>
<code>  ... </code>	
<code>  if(!vps_constant_poc_cycle_per_au)</code>	
<code>    slice_poc_cycle_au</code>	<code>u(8)</code>
<code>  </code>	
<code>  </code>	
<code>}</code>	

图9

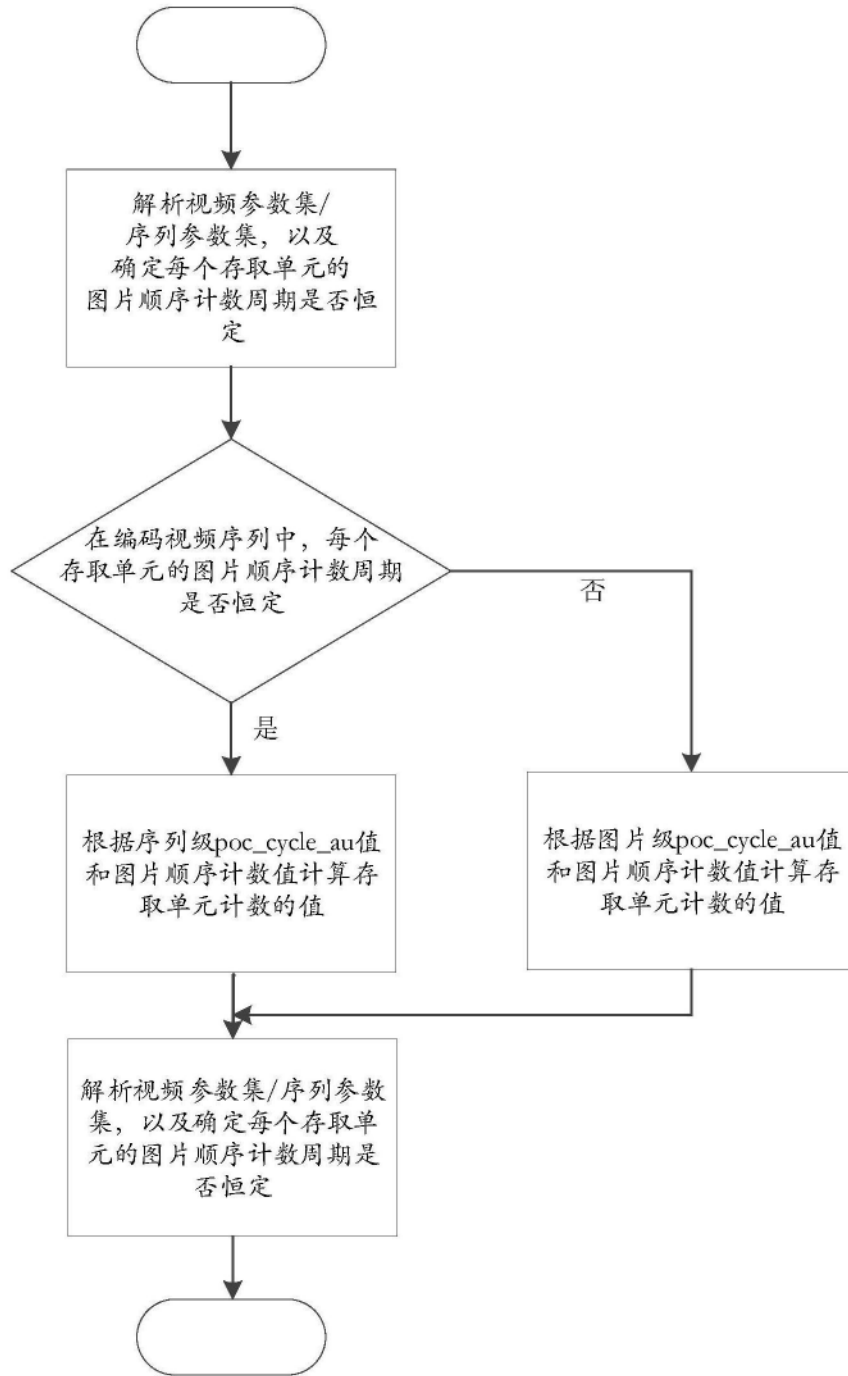


图10

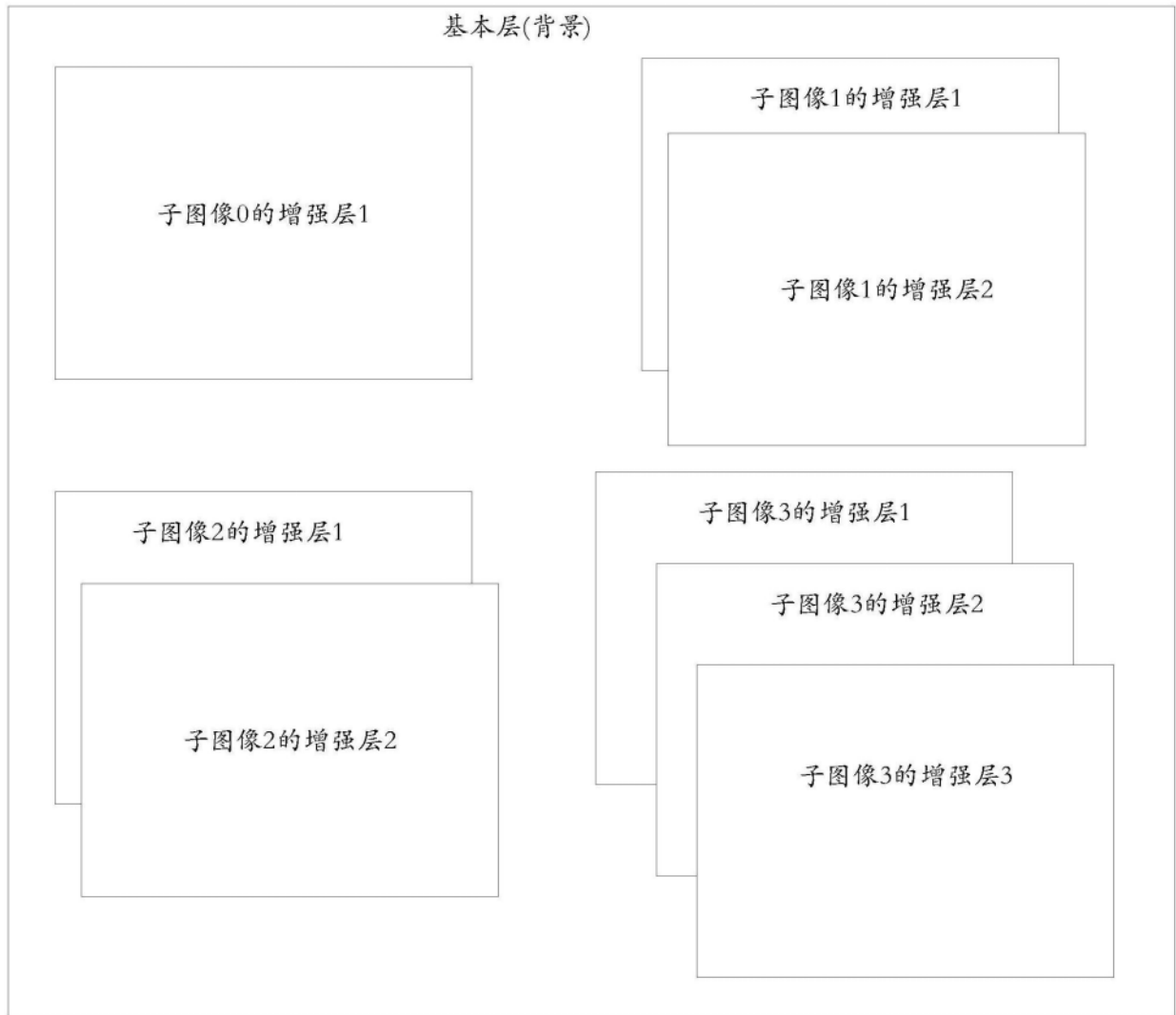


图11

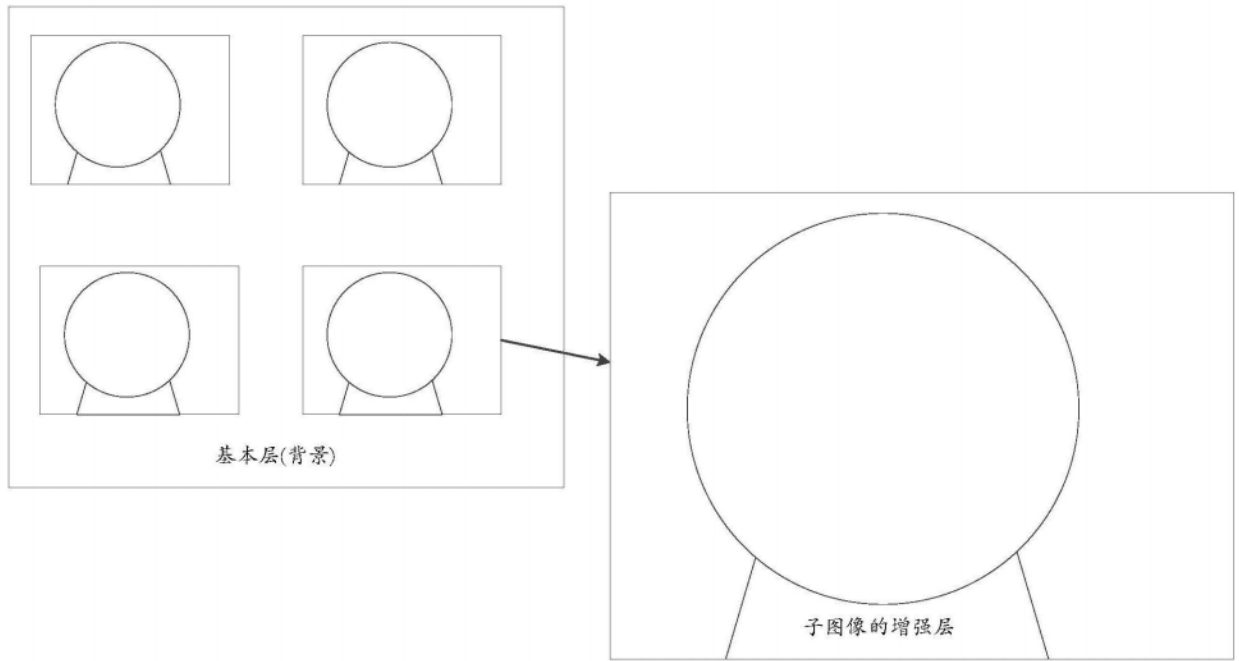


图12

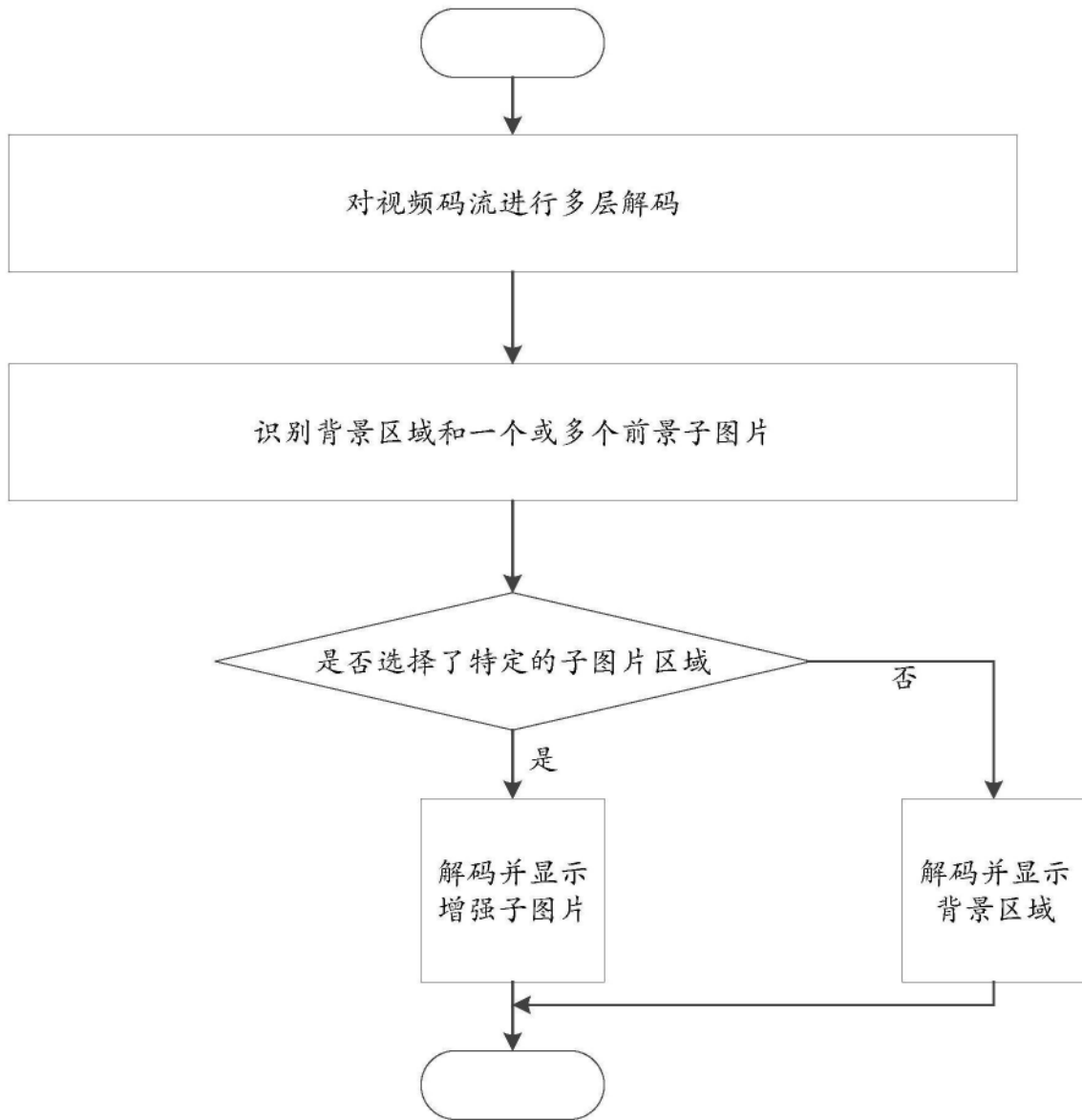


图13

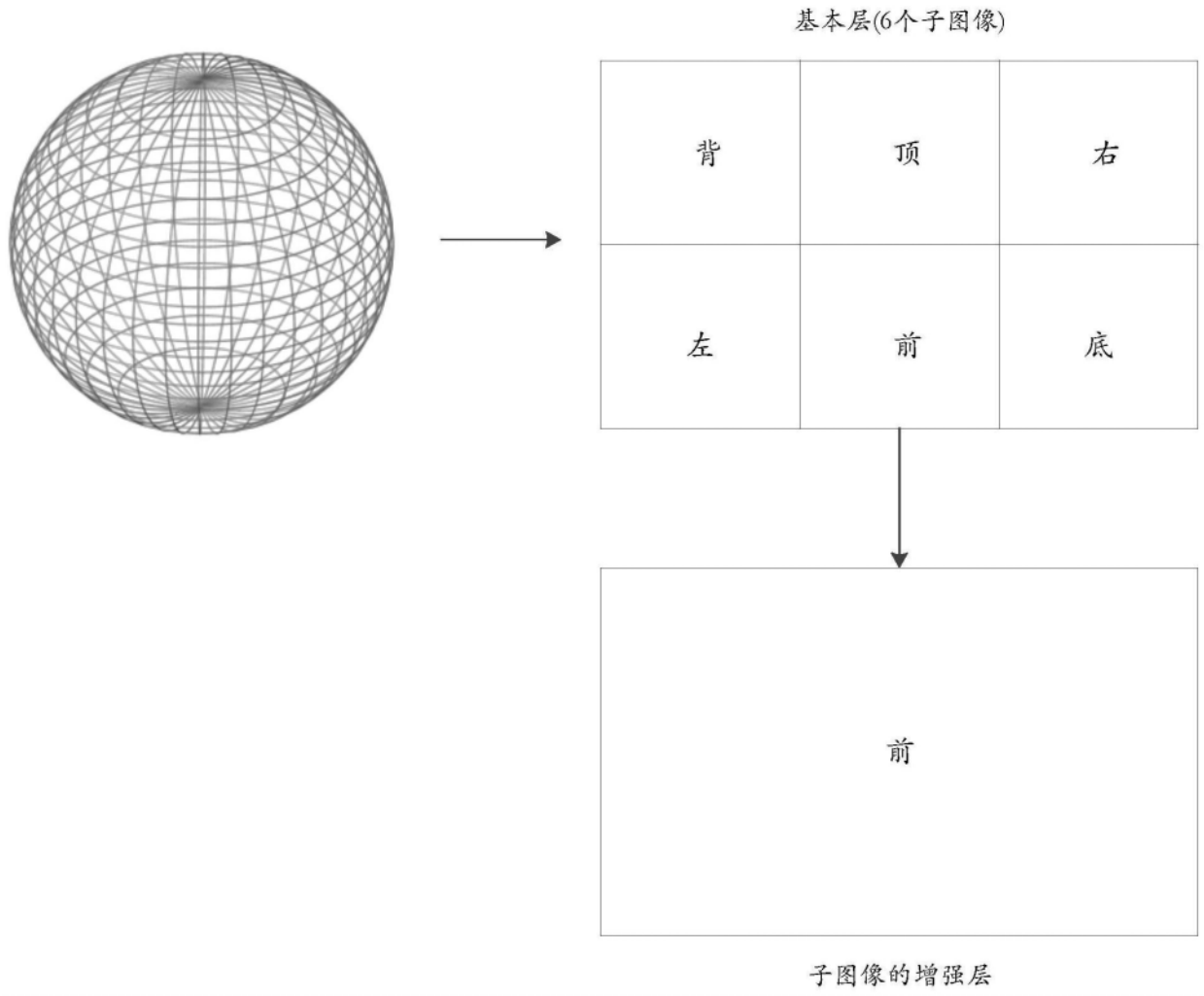


图14

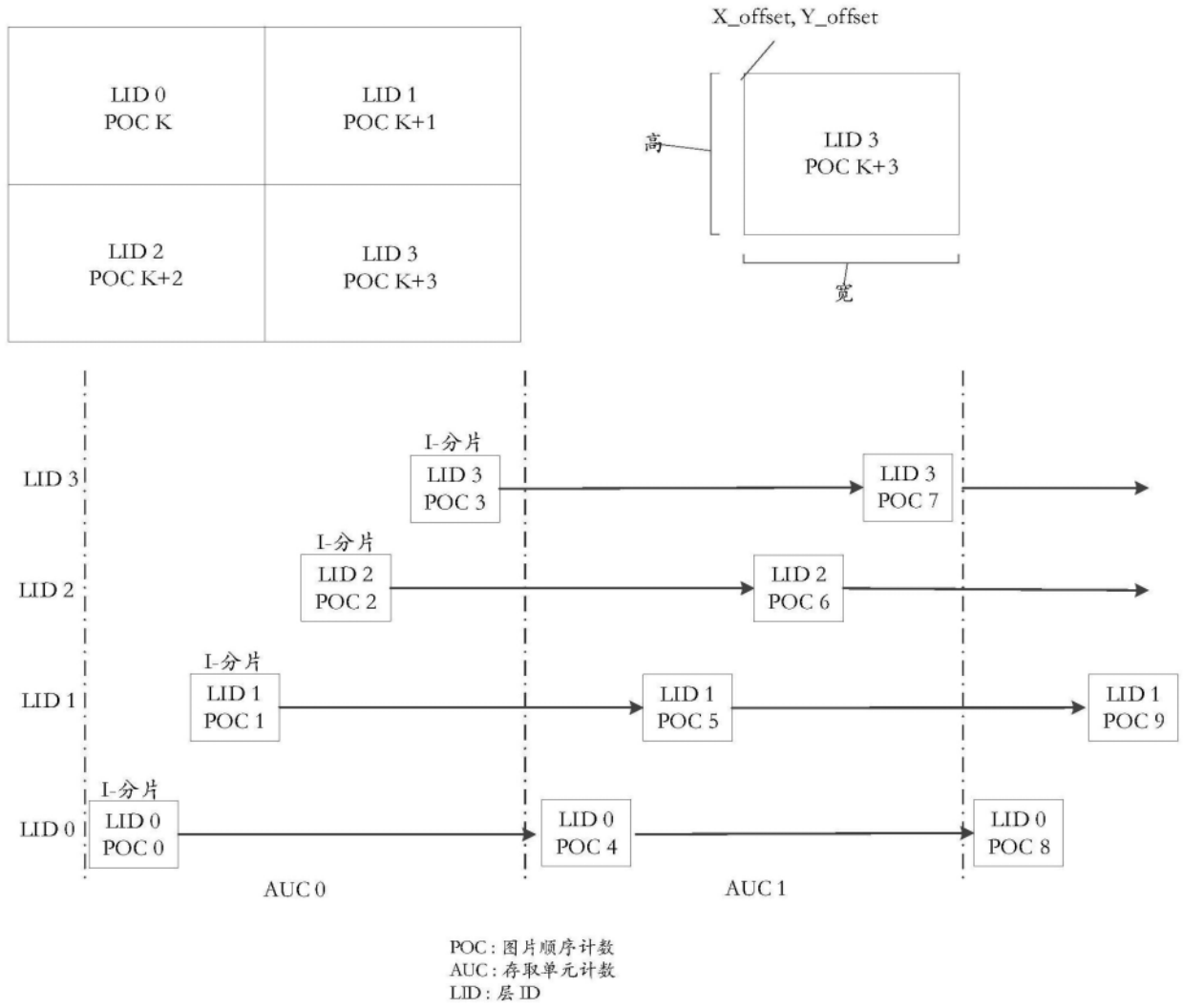


图15

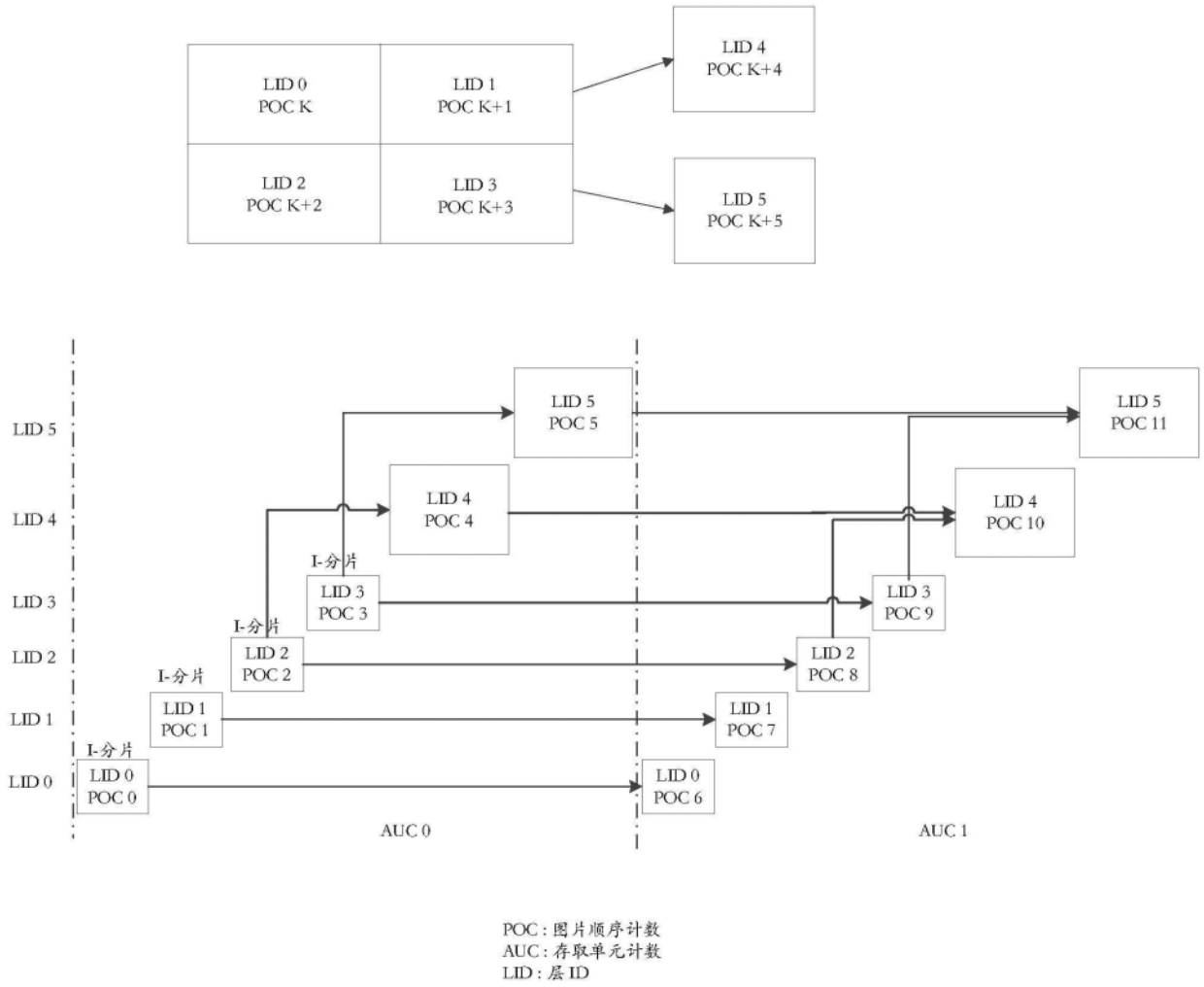


图16



<code>sub_region_partitioning_info( payloadSize ) {</code>	描述符
<code>num_sub_region</code>	<code>ue(v)</code>
<code>num_layers</code>	<code>ue(v)</code>
<code>for( j = 0; j &lt;= num_layers; j++ )</code>	
<code>layer_id [ i ]</code>	<code>ue(v)</code>
<code>for( j = 1; j &lt; num_layers; j++ )</code>	
<code>for( j = 0; j &lt; i; j++ )</code>	
<code>direct_dependency_flag[ i ][ j ]</code>	<code>u(1)</code>
<code>for( i = 0; i &lt; num_sub_region; i++ ) {</code>	
<code>num_layers_for_region[ i ]</code>	<code>ue(v)</code>
<code>for( j = 0; j &lt; num_layers_for_region[ i ]; j++ )</code>	
<code>sub_region_layer_id[ i ][ j ]</code>	<code>ue(v)</code>
<code>sub_region_offset_x[ i ]</code>	<code>ue(v)</code>
<code>sub_region_offset_y[ i ]</code>	<code>ue(v)</code>
<code>sub_region_width[ i ]</code>	<code>ue(v)</code>
<code>sub_region_height [ i ]</code>	<code>ue(v)</code>
<code>}</code>	
<code>...</code>	
<code>}</code>	

图18

<code>video parameter set rbsp() {</code>	描述符
<code>...</code>	
<code>vps max layers minus1</code>	<code>u(6)</code>
<code>num output layer sets</code>	<code>ue(v)</code>
<code>num profile tier level</code>	<code>ue(v)</code>
<code>for( i = 0; i &lt; num profile tier level; i++ )</code>	
<code>profile tier level( vps max sub layers minus1 )</code>	
<code>for( i = 0; i &lt; num output layer sets; i++ )</code>	
<code>for( j = 0; j &lt; NumLayersInIdList[ i ]; j++ ) {</code>	
<code>output layer flag[ i ][ j ]</code>	<code>u(1)</code>
<code>profile tier level idx[ i ][ j ]</code>	<code>u(v)</code>
<code>}</code>	
<code>...</code>	
<code>}</code>	

图19

	描述符
video parameter set rbsp() {	
...	
<b>vps max layers minus1</b>	u(6)
if(vps max layers minus1 > 0) {	
<b>num output layer sets</b>	ue(v)
<b>num profile tier level</b>	ue(v)
}	
for( i = 0; i < num profile tier level; i++ )	
profile tier level( vps max sub layers minus1 )	
for( i = 0; i < num output layer sets; i++ ) {	
<b>vps output layers mode[ i ]</b>	u(2)
<b>vps ptl signal flag[ i ]</b>	u(1)
for( j = 0; j < NumLayersInIdList[ I ]; j++ ) {	
if(vps output layers mode[ i ] == 2 )	
<b>output layer flag[ I ][ i ]</b>	u(1)
if( vps ptl signal flag[ i ] )	
<b>profile tier level idx[ I ][ i ]</b>	u(v)
}	
}	
}	
...	
}	

图20

	描述符
video parameter set rbsp() {	
...	
<b>vps max layers minus1</b>	u(6)
if(vps max layers minus1 > 0) {	
<b>num_output_layer_sets</b>	ue(v)
<b>num_profile_tier_level</b>	ue(v)
}	
<b>max_subpics_minus1</b>	u(8)
for( i = 0; i < max_subpics_minus1; i++ ) {	
<b>sub_pic_id[i]</b>	u(8)
}	
for( i = 0; i < num_profile_tier_level; i++ )	
profile_tier_level( vps_max_sub_layers_minus1 )	
for( i = 0; i < num_output_layer_sets; i++ ) {	
<b>vps_output_layers_mode[ i ]</b>	u(2)
<b>vps_ptl_signal_flag[ i ]</b>	u(1)
for( j = 0; j < NumLayersInIdList[ i ]; j++ ) {	
<b>num_output_subpic_layer[i][j]</b>	ue(v)
for( k = 0; k < num_output_subpic_layer[i][j]; k++ )	
<b>sub_pic_id_layer[i][j][k]</b>	u(8)
if(vps_output_layers_mode[i] == 2 )	
<b>output_layer_flag[ i ][ j ]</b>	u(1)
if( vps_ptl_signal_flag[ i ] )	
<b>profile_tier_level_idx[ i ][ j ]</b>	u(v)
}	
}	
}	
...	
}	

图21