



(12) 发明专利

(10) 授权公告号 CN 114514749 B

(45) 授权公告日 2023. 10. 20

(21) 申请号 202080070174.9

(22) 申请日 2020.10.06

(65) 同一申请的已公布的文献号
申请公布号 CN 114514749 A

(43) 申请公布日 2022.05.17

(30) 优先权数据
62/911,808 2019.10.07 US

(85) PCT国际申请进入国家阶段日
2022.04.07

(86) PCT国际申请的申请数据
PCT/US2020/054448 2020.10.06

(87) PCT国际申请的公布数据
W02021/022267 EN 2021.02.04

(73) 专利权人 华为技术有限公司
地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72) 发明人 王业奎

(51) Int.Cl.
H04N 19/52 (2006.01)
H04N 7/12 (2006.01)

(56) 对比文件
US 2019058895 A1, 2019.02.21
CN 105637878 A, 2016.06.01
US 2016255373 A1, 2016.09.01
US 2014185691 A1, 2014.07.03
董力菡等. 基于FPGA的HEVC网络适配层单元实现.《电气开关》.2019, (第02期),
Y-K WANG. JVET-P0115, AHG8:
Scalability-output layer sets: associated text specifications based upon VVC Draft 6.《JVET-P0115》.2019,

审查员 杨少魁

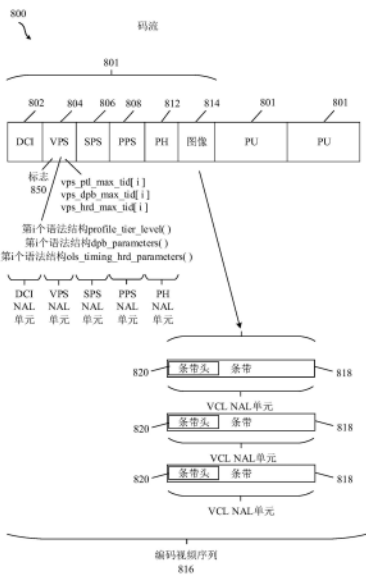
权利要求书3页 说明书50页 附图12页

(54) 发明名称

避免多层视频码流中的冗余指示

(57) 摘要

本发明提供了一种解码方法。所述方法包括：接收包括多层和视频参数集 (video parameterset, VPS) 的视频码流，其中，所述多层包括子层，所述VPS包括具有第一值的第一标志，具有所述第一值的所述第一标志表示级别信息、解码图像缓冲区 (decoded picture buffer, DPB) 参数和假设参考解码 (hypothetical reference decoder, HRD) 参数对应的最高子层表示的时间标识符 (identifier, ID) 不存在于所述VPS中并被推断为所述子层的最大数量，所述子层可以存在于由所述VPS指定的所述多层中的一层中；从所述VPS中获取所述最高子层表示的时间ID对应的所述级别信息、所述HRD参数和所述DPB参数；对所述多层中的一层中的图像进行解码，以得到经解码图像。本发明还提供了一种对应的编码方法。



1. 一种由视频解码器实现的解码方法,其特征在于,所述方法包括:

所述视频解码器接收包括多层和视频参数集(video parameter set,VPS)的视频码流,其中,所述多层包括子层,所述VPS包括具有第一值的第一标志,具有所述第一值的所述第一标志表示级别信息对应的最高子层表示的时间标识符(identifier,ID)、解码图像缓冲区(decoded picture buffer,DPB)参数对应的最高子层表示的时间标识符(identifier,ID)和假设解码器刷新(hypothetical decoder refresh,HRD)参数对应的最高子层表示的时间标识符(identifier,ID)不存在于所述VPS中并被推断为所述子层的最大数量,所述子层存在于由所述VPS指定的所述多层中的一层中,其中,所述级别信息存在于所述VPS中的第i个语法结构profile_tier_level()中;

所述视频解码器从所述VPS中获取所述最高子层表示的时间ID对应的所述级别信息、所述HRD参数和所述DPB参数;

所述视频解码器对所述多层中的一层中的图像进行解码,以得到经解码图像。

2. 根据权利要求1所述的方法,其特征在于,所述子层的最大数量的范围为1至7,所述第一值为1。

3. 根据权利要求1或2所述的方法,其特征在于,所述DPB参数存在于所述VPS中的第i个语法结构dpb_parameters()中。

4. 根据权利要求1至3中任一项所述的方法,其特征在于,所述HRD参数存在于所述VPS或序列参数集(sequence parameter set,SPS)中的第i个语法结构ols_timing_hrd_parameters()中。

5. 根据权利要求1至4中任一项所述的方法,其特征在于,所述图像使用所述级别信息和所述DPB参数中的一个或两个进行解码。

6. 根据权利要求1至5中任一项所述的方法,其特征在于,所述图像不使用所述HRD参数进行解码。

7. 根据权利要求1至6中任一项所述的方法,其特征在于,所述方法还包括:在电子设备的显示器上显示所述经解码图像。

8. 一种由视频编码器实现的编码方法,其特征在于,所述方法包括:

所述视频编码器生成包括子层的多层和包括第一标志的视频参数集(video parameter set,VPS);

所述视频编码器将所述第一标志设置为第一值以表示级别信息对应的最高子层表示的时间标识符(identifier,ID)、解码图像缓冲区(decoded picture buffer,DPB)参数对应的最高子层表示的时间标识符(identifier,ID)和假设解码器刷新(hypothetical decoder refresh,HRD)参数对应的最高子层表示的时间标识符(identifier,ID)不存在于所述VPS中并应该被推断为所述子层的最大数量,所述子层存在于由所述VPS指定的所述多层中的一层中,其中,所述级别信息存在于所述VPS中的第i个语法结构profile_tier_level()中;

所述视频编码器将所述多层和所述VPS编码到视频码流中;

所述视频编码器存储所述视频码流,以将所述视频码流发送给视频解码器。

9. 根据权利要求8所述的方法,其特征在于,所述子层的最大数量的范围为1至7。

10. 根据权利要求8或9所述的方法,其特征在于,所述DPB参数存在于所述VPS中的第i

个语法结构`dpb_parameters()`中,所述HRD参数存在于所述VPS或序列参数集(sequence parameter set,SPS)中的第*i*个语法结构`ols_timing_hrd_parameters()`中。

11.根据权利要求8至10中任一项所述的方法,其特征在于,所述方法还包括:使用所述HRD参数确定所述编码的视频码流是否一致。

12.一种解码设备,其特征在于,所述解码设备包括:

接收器,用于接收包括多层和视频参数集(video parameter set,VPS)的视频码流,其中,所述多层包括子层,所述VPS包括具有第一值的第一标志,具有所述第一值的所述第一标志表示级别信息对应的最高子层表示的时间标识符(identifier,ID)、解码图像缓冲区(decoded picture buffer,DPB)参数对应的最高子层表示的时间标识符(identifier,ID)和假设解码器刷新(hypothetical decoder refresh,HRD)参数对应的最高子层表示的时间标识符(identifier,ID)不存在于所述VPS中并被推断为所述子层的最大数量,所述子层存在于由所述VPS指定的所述多层中的一层中,其中,所述级别信息存在于所述VPS中的第*i*个语法结构`profile_tier_level()`中;

与存储器耦合的处理器,其中,所述处理器用于执行指令,以使得所述解码设备执行以下操作:

从所述VPS中获取所述最高子层表示的时间ID对应的所述级别信息、所述HRD参数和所述DPB参数;

对所述多层中的一层中的图像进行解码,以得到经解码图像。

13.根据权利要求12所述的解码设备,其特征在于,所述子层的最大数量的范围为1至7。

14.根据权利要求12或13所述的解码设备,其特征在于,所述DPB参数存在于所述VPS中的第*i*个语法结构`dpb_parameters()`中,所述HRD参数存在于所述VPS或序列参数集(sequence parameter set,SPS)中的第*i*个语法结构`ols_timing_hrd_parameters()`中。

15.根据权利要求12至14中任一项所述的解码设备,其特征在于,所述图像使用所述级别信息和所述DPB参数中的一个或两个进行解码。

16.根据权利要求12至15中任一项所述的解码设备,其特征在于,所述解码设备还包括用于显示所述经解码图像的显示器。

17.一种编码设备,其特征在于,所述编码设备包括:

包括指令的存储器;

与所述存储器耦合的处理器,其中,所述处理器用于实现所述指令,以使得所述编码设备执行以下操作:

生成包括子层的多层和包括第一标志的视频参数集(video parameter set,VPS);

将所述第一标志设置为第一值以表示级别信息对应的最高子层表示的时间标识符(identifier,ID)、解码图像缓冲区(decoded picture buffer,DPB)参数对应的最高子层表示的时间标识符(identifier,ID)和假设解码器刷新(hypothetical decoder refresh,HRD)参数对应的最高子层表示的时间标识符(identifier,ID)不存在于所述VPS中并应该被推断为所述子层的最大数量,所述子层存在于由所述VPS指定的所述多层中的一层中,其中,所述级别信息存在于所述VPS中的第*i*个语法结构`profile_tier_level()`中;

将所述多层和所述VPS编码到视频码流中;

与所述处理器耦合的发送器,其中,所述发送器用于将所述视频码流发送给视频解码器。

18.根据权利要求17所述的编码设备,其特征在于,所述子层的最大数量的范围为1至7。

19.根据权利要求17或18所述的编码设备,其特征在于,所述DPB参数存在于所述VPS中的第i个语法结构dpb_parameters()中,所述HRD参数存在于所述VPS或序列参数集(sequence parameter set,SPS)中的第i个语法结构ols_timing_hrd_parameters()中。

20.一种译码装置,其特征在于,所述译码装置包括:

接收器,用于接收图像进行编码或接收码流进行解码;

与所述接收器耦合的发送器,其中,所述发送器用于将所述码流发送给解码器或将经解码图像发送给显示器;

与所述接收器或所述发送器中的至少一个耦合的存储器,其中,所述存储器用于存储指令;

与所述存储器耦合的处理器,其中,所述处理器用于执行存储在所述存储器中的所述指令,以执行根据权利要求1至7中任一项所述的方法和根据权利要求8至11中任一项所述的方法。

21.根据权利要求20所述的译码装置,其特征在于,所述译码装置还包括用于显示经解码图像的显示器。

22.一种系统,其特征在于,所述系统包括:

编码器;

与所述编码器通信的解码器,其中,所述解码器包括根据权利要求12至16中任一项所述的解码设备,所述编码器包括根据权利要求17至19中任一项所述的编码设备。

23.一种译码模块,其特征在于,所述译码模块包括:

接收模块,用于接收图像进行编码或接收码流进行解码;

与所述接收模块耦合的发送模块,其中,所述发送模块用于将所述码流发送给解码模块或将经解码图像发送给显示模块;

与所述接收模块或所述发送模块中的至少一个耦合的存储模块,其中,所述存储模块用于存储指令;

与所述存储模块耦合的处理模块,其中,所述处理模块用于执行存储在所述存储模块中的所述指令,以执行根据权利要求1至7中任一项所述的方法和根据权利要求8至11中任一项所述的方法。

避免多层视频码流中的冗余指示

[0001] 相关申请交叉引用

[0002] 本专利申请要求王业奎于2019年10月7日提交的申请号为62/911,808、发明名称为“视频译码中的可适性(Scalability In Video Coding)”的美国临时专利申请的权益，其内容通过引用结合在本申请中。

技术领域

[0003] 本发明大体上描述了用于视频译码中多层视频码流的技术。更具体地，本发明描述了用于在视频译码中在多层码流中指示层对应的参数时减少冗余的技术。

背景技术

[0004] 即使视频相对较短，也需要大量的视频数据来描述，当数据要在带宽容量有限的通信网络中流式传输或以其它方式传输时，这样可能会造成困难。因此，视频数据通常要先压缩，然后通过现代电信网络进行传输。由于内存资源可能有限，当在存储设备上存储视频时，该视频的大小也可能是一个问题。视频压缩设备通常在信源侧使用软件和/或硬件对视频数据进行编码，然后传输或存储视频数据，从而减少表示数字视频图像所需的数据量。然后，对视频数据进行解码的视频解压缩设备在目的地侧接收压缩数据。在网络资源有限以及对更高视频质量的需求不断增长的情况下，需要改进压缩和解压缩技术，这些改进的技术能够在几乎不影响图像质量的情况下提高压缩比。

发明内容

[0005] 第一方面涉及一种由视频解码器实现的解码方法。所述方法包括：所述视频解码器接收包括多层和视频参数集(video parameter set,VPS)的视频码流，其中，所述多层包括子层，所述VPS包括具有第一值的第一标志，具有所述第一值的所述第一标志表示级别信息对应的最高子层表示的时间标识符(identifier,ID)、解码图像缓冲区(decoded picture buffer,DPB)参数对应的最高子层表示的时间标识符(identifier,ID)和假设参考解码(hypothetical reference decoder,HRD)参数对应的最高子层表示的时间标识符(identifier,ID)不存在于所述VPS中并被推断为所述子层的最大数量，所述子层可以存在于由所述VPS指定的所述多层中的一层中；所述视频解码器从所述VPS中获取所述最高子层表示的时间ID对应的所述级别信息、所述HRD参数和所述DPB参数；所述视频解码器对所述多层中的一层中的图像进行解码，以得到经解码图像。

[0006] 所述方法提供的技术确保不会在多层码流中为各层重复指示相同的档次、层次和级别(profile,tier,and level,PTL)参数、假设参考解码(hypothetical reference decoder,HRD)参数和解码图像缓冲区(decoded picture buffer,DPB)参数。仅指示相同的PTL参数、HRD参数和DPB参数一次减少了冗余并提高了多层视频码流的译码效率。因此，视频译码中的编码器/解码器(又称“编解码器”)相对于现有的编解码器有所改进。实际上，改进的视频译码过程在发送、接收和/或观看视频时提供了更好的用户体验。

[0007] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述子层的最大数量的范围为1至7,所述第一值为1,第二值为0。

[0008] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述级别信息包括在档次、层次和级别(profile,tier,and level,PTL)参数中。

[0009] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述级别信息存在于所述VPS中的第i个语法结构profile_tier_level()中。

[0010] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述DPB参数存在于所述VPS中的第i个语法结构dpb_parameters()中。

[0011] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述HRD参数存在于所述VPS或序列参数集(sequence parameter set,SPS)中的第i个语法结构ols_timing_hrd_parameters()中。

[0012] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述图像使用所述级别信息和所述DPB参数中的一个或两个进行解码。

[0013] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述图像不使用所述HRD参数进行解码。

[0014] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述方法还包括:在电子设备的显示器上显示所述经解码图像。

[0015] 第二方面涉及一种由视频编码器实现的编码方法。所述方法包括:所述视频编码器生成包括子层的多层和包括第一标志的视频参数集(video parameter set,VPS);所述视频编码器将所述第一标志设置为第一值以表示级别信息对应的最高子层表示的时间标识符(identifier,ID)、解码图像缓冲区(decoded picture buffer,DPB)参数对应的最高子层表示的时间标识符(identifier,ID)和假设参考解码(hypothetical reference decoder,HRD)参数对应的最高子层表示的时间标识符(identifier,ID)不存在于所述VPS中并应该被推断为所述子层的最大数量,所述子层可以存在于由所述VPS指定的所述多层中的一层中;所述视频编码器将所述多层和所述VPS编码到视频码流中;所述视频编码器存储所述视频码流,以将所述视频码流发送给视频解码器。

[0016] 所述方法提供的技术确保不会在多层码流中为各层重复指示相同的档次、层次和级别(profile,tier,and level,PTL)参数、假设参考解码(hypothetical reference decoder,HRD)参数和解码图像缓冲区(decoded picture buffer,DPB)参数。仅指示相同的PTL参数、HRD参数和DPB参数一次减少了冗余并提高了多层视频码流的译码效率。因此,视频译码中的编码器/解码器(又称“编解码器”)相对于现有的编解码器有所改进。实际上,改进的视频译码过程在发送、接收和/或观看视频时提供了更好的用户体验。

[0017] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述子层的最大数量的范围为1至7。

[0018] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述级别信息包括在档次、层次和级别(profile,tier,and level,PTL)参数中。

[0019] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述级别信息存在于所述VPS中的第i个语法结构profile_tier_level()中,所述DPB参数存在于所述VPS中的第i个语法结构dpb_parameters()中,所述HRD参数存在于所述VPS或序列参数集

(sequence parameter set,SPS)中的第i个语法结构ols_timing_hrd_parameters()中。

[0020] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述方法还包括:使用所述HRD参数确定所述编码的视频码流是否一致。

[0021] 第三方面涉及一种解码设备。所述解码设备包括:接收器,用于接收包括多层和视频参数集(video parameter set,VPS)的视频码流,其中,所述多层包括子层,所述VPS包括具有第一值的第一标志,具有所述第一值的所述第一标志表示级别信息对应的最高子层表示的时间标识符(identifier,ID)、解码图像缓冲区(decoded picture buffer,DPB)参数对应的最高子层表示的时间标识符(identifier,ID)和假设参考解码(hypothetical reference decoder,HRD)参数对应的最高子层表示的时间标识符(identifier,ID)不存在于所述VPS中并被推断为所述子层的最大数量,所述子层可以存在于由所述VPS指定的所述多层中的一层中;与所述存储器耦合的处理器,其中,所述处理器用于执行所述指令,以使得所述解码设备执行以下操作:从所述VPS中获取所述最高子层表示的时间ID对应的所述级别信息、所述HRD参数和所述DPB参数;对所述多层中的一层中的图像进行解码,以得到经解码图像。

[0022] 所述解码设备提供的技术确保不会在多层码流中为各层重复指示相同的档次、层次和级别(profile,tier,and level,PTL)参数、假设参考解码(hypothetical reference decoder,HRD)参数和解码图像缓冲区(decoded picture buffer,DPB)参数。仅指示相同的PTL参数、HRD参数和DPB参数一次减少了冗余并提高了多层视频码流的译码效率。因此,视频译码中的编码器/解码器(又称“编解码器”)相对于现有的编解码器有所改进。实际上,改进的视频译码过程在发送、接收和/或观看视频时提供了更好的用户体验。

[0023] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述子层的最大数量的范围为1至7。

[0024] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述级别信息包括在档次、层次和级别(profile,tier,and level,PTL)参数中。

[0025] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述级别信息存在于所述VPS中的第i个语法结构profile_tier_level()中,所述DPB参数存在于所述VPS中的第i个语法结构dpb_parameters()中,所述HRD参数存在于所述VPS或序列参数集(sequence parameter set,SPS)中的第i个语法结构ols_timing_hrd_parameters()中。

[0026] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述图像使用所述级别信息和所述DPB参数中的一个或两个进行解码。

[0027] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述解码设备还包括用于显示所述经解码图像的显示器。

[0028] 第四方面涉及一种编码设备。所述编码设备包括:包括指令的存储器;与所述存储器耦合的处理器,其中,所述处理器用于实现所述指令,以使得所述编码设备执行以下操作:生成包括子层的多层和包括第一标志的视频参数集(video parameter set,VPS);将所述第一标志设置为第一值以表示级别信息、解码图像缓冲区(decoded picture buffer,DPB)参数和假设参考解码(hypothetical reference decoder,HRD)对应的最高子层表示的时间标识符(TemporalID)不存在于所述VPS中并应该被推断为所述子层的最大数量,所述子层可以存在于由所述VPS指定的所述多层中的一层中;将所述多层和所述VPS编码到视

频码流中;与所述处理器耦合的发送器,其中,所述发送器用于将所述视频码流发送给视频解码器。

[0029] 所述编码设备提供的技术确保不会在多层码流中为各层重复指示相同的档次、层次和级别(profile,tier,and level,PTL)参数、假设参考解码(hypothetical reference decoder,HRD)参数和解码图像缓冲区(decoded picture buffer,DPB)参数。仅指示相同的PTL参数、HRD参数和DPB参数一次减少了冗余并提高了多层视频码流的译码效率。因此,视频译码中的编码器/解码器(又称“编解码器”)相对于现有的编解码器有所改进。实际上,改进的视频译码过程在发送、接收和/或观看视频时提供了更好的用户体验。

[0030] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述子层的最大数量的范围为1至7。

[0031] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述级别信息包括在档次、层次和级别(profile,tier,and level,PTL)参数中。

[0032] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述级别信息存在于所述VPS中的第i个语法结构profile_tier_level()中,所述DPB参数存在于所述VPS中的第i个语法结构dpb_parameters()中,所述HRD参数存在于所述VPS或序列参数集(sequence parameter set,SPS)中的第i个语法结构ols_timing_hrd_parameters()中。

[0033] 第五方面涉及一种译码装置。所述译码装置包括:接收器,用于接收图像进行编码或接收码流进行解码;与所述接收器耦合的发送器,其中,所述发送器用于将所述码流发送给解码器或将经解码图像发送给显示器;与所述接收器或所述发送器中的至少一个耦合的存储器,其中,所述存储器用于存储指令;与所述存储器耦合的处理器,其中,所述处理器用于执行存储在所述存储器中的所述指令,以执行本文公开的任一方法。

[0034] 所述译码装置提供的技术确保不会在多层码流中为各层重复指示相同的档次、层次和级别(profile,tier,and level,PTL)参数、假设参考解码(hypothetical reference decoder,HRD)参数和解码图像缓冲区(decoded picture buffer,DPB)参数。仅指示相同的PTL参数、HRD参数和DPB参数一次减少了冗余并提高了多层视频码流的译码效率。因此,视频译码中的编码器/解码器(又称“编解码器”)相对于现有的编解码器有所改进。实际上,改进的视频译码过程在发送、接收和/或观看视频时提供了更好的用户体验。

[0035] 可选地,根据任一上述方面,在所述方面的另一种实现方式中,所述译码装置还包括用于显示经解码图像的显示器。

[0036] 第六方面涉及一种系统。所述系统包括:编码器;与所述编码器通信的解码器,其中,所述编码器或所述解码器包括本文公开的解码设备、编码设备或译码装置。

[0037] 所述系统提供的技术确保不会在多层码流中为各层重复指示相同的档次、层次和级别(profile,tier,and level,PTL)参数、假设参考解码(hypothetical reference decoder,HRD)参数和解码图像缓冲区(decoded picture buffer,DPB)参数。仅指示相同的PTL参数、HRD参数和DPB参数一次减少了冗余并提高了多层视频码流的译码效率。因此,视频译码中的编码器/解码器(又称“编解码器”)相对于现有的编解码器有所改进。实际上,改进的视频译码过程在发送、接收和/或观看视频时提供了更好的用户体验。

[0038] 第七方面涉及一种译码模块。所述译码模块包括:接收模块,用于接收图像进行编码或接收码流进行解码;与所述接收模块耦合的发送模块,其中,所述发送模块用于将所述

码流发送给解码模块或将经解码图像发送给显示模块;与所述接收模块或所述发送模块中的至少一个耦合的存储模块,其中,所述存储模块用于存储指令;与所述存储模块耦合的处理模块,其中,所述处理模块用于执行存储在所述存储模块中的所述指令,以执行本文公开的任一方法。

[0039] 所述译码模块提供的技术确保不会在多层码流中为各层重复指示相同的档次、层次和级别(profile, tier, and level, PTL) 参数、假设参考解码(hypothetical reference decoder, HRD) 参数和解码图像缓冲区(decoded picture buffer, DPB) 参数。仅指示相同的 PTL 参数、HRD 参数和 DPB 参数一次减少了冗余并提高了多层视频码流的译码效率。因此,视频译码中的编码器/解码器(又称“编解码器”)相对于现有的编解码器有所改进。实际上,改进的视频译码过程在发送、接收和/或观看视频时提供了更好的用户体验。

[0040] 为了清楚起见,任一上述实施例可以与上述其它任何一个或多个实施例组合以创建在本发明范围内的新实施例。

[0041] 根据以下结合附图和权利要求书的详细描述,将会更清楚地理解这些和其它特征。

附图说明

[0042] 为了更全面地理解本发明,现在参考以下结合附图和具体实施方式进行的简要描述,其中,相似的附图标记表示相似的部件。

[0043] 图1为对视频信号进行译码的示例性方法的流程图。

[0044] 图2为用于视频译码的示例性编码和解码(编解码)系统的示意图。

[0045] 图3为示例性视频编码器的示意图。

[0046] 图4为示例性视频解码器的示意图。

[0047] 图5示出了用于空间可适性的多层译码的一个示例。

[0048] 图6示出了使用输出层集(output layer set, OLS)的多层译码的一个示例。

[0049] 图7示出了用于空间可适性的多层视频序列的一个示例。

[0050] 图8示出了视频码流的一个实施例。

[0051] 图9为对经编码视频码流进行解码的方法的一个实施例。

[0052] 图10为对经编码视频码流进行编码的方法的一个实施例。

[0053] 图11为视频译码设备的示意图。

[0054] 图12为译码模块的一个实施例的示意图。

具体实施方式

[0055] 首先应当理解,尽管下文提供一个或多个实施例的说明性实现方式,但所公开的系统 and/或方法可以使用任意数量的技术来实施,无论这些技术是当前已知的还是现有的。本发明决不应限于下文所说明的说明性实现方式、附图和技术,包括本文所说明并描述的示例性设计和实现方式,而是可以在所附权利要求书的范围以及其等效物的完整范围内修改。

[0056] 以下术语的定义如下所述,除非在本文相反的上下文中使用。具体地,以下定义旨在更加清晰地描述本发明。但是,术语在不同的上下文中可能会有不同的描述。因此,以下

定义应当视为补充信息,而不应当视为对此处为这些术语提供的描述的任何其它定义进行限制。

[0057] 码流是包括视频数据的一系列比特,这些视频数据会进行压缩以在编码器和解码器之间传输。编码器是一种用于通过编码过程将视频数据压缩到码流中的设备。解码器是一种用于通过解码过程从码流中重建视频数据以进行显示的设备。图像是创建帧或其场的亮度样本和/或色度样本组成的阵列。为了清楚论述,正被编码或解码的图像可以称为当前图像。参考图像是包括参考样本的图像,这些样本可以在根据帧间预测和/或层间预测通过参考对其它图像进行译码时使用。参考图像列表是用于帧间预测和/或层间预测的一系列参考图像。一些视频译码系统使用两个参考图像列表,这两个参考图像列表可以表示为参考图像列表1和参考图像列表0。参考图像列表结构是包括多个参考图像列表的可寻址语法结构。

[0058] 档次、层次和级别(profile,tier,and level,PTL)参数是一种提供档次、层次和/或级别信息的语法结构(例如第i个语法结构profile_tier_level())。PTL参数可以表示为语法结构profile_tier_level()。

[0059] 帧间预测是一种通过参考与当前图像不同的参考图像中的指示样本对当前图像中的样本进行译码的机制,其中,参考图像和当前图像位于同一层。参考图像列表结构条目是参考图像列表结构中表示与参考图像列表关联的参考图像的可寻址位置。

[0060] 条带头(slice header)是经编码条带的一部分,包括与在条带中表示的一个分块内的所有视频数据相关的数据元素。图像参数集(picture parameter set,PPS)是包括与整个图像相关的数据的参数集。更具体地,PPS是一种包括适用于0个或0个以上完整的经编码图像的语法元素的语法结构,0个或0个以上完整的经编码图像通过每个图像头中的语法元素确定。序列参数集(sequence parameter set,SPS)是包括与图像序列相关的数据的参数集。接入单元(access unit,AU)是与同一显示时间(例如同一个图像顺序编号)关联的一个或多个经编码图像的集合,这些图像从解码图像缓冲区(decoded picture buffer,DPB)输出(例如显示给用户)。接入单元分隔符(access unit delimiter,AUD)是用于指示AU起始或AU之间边界的指示符或数据结构。解码视频序列(decoded video sequence)是一系列已经由解码器重建以准备显示给用户的图像。

[0061] 网络抽象层(network abstraction layer,NAL)单元是一种包括原始字节序列载荷(Raw Byte Sequence Payload,RBSP)形式的数据、指示数据类型并根据需要穿插预防混淆字节的语法结构。视频编码层(video coding layer,VCL)NAL单元是编码后包括视频数据的NAL单元,例如图像中的经编码条带(coded slice)。非VCL NAL单元是包括非视频数据的NAL单元,例如支持对视频数据进行解码、一致性检查的性能或其它操作的语法和/或参数。层是一组VCL NAL单元和关联的非VCL NAL单元,其中,这些VCL NAL单元共用指定特征(例如,公共分辨率、帧率、图像大小等)。层中的VCL NAL单元可以共用NAL单元头层标识符(nuh_layer_id)的特定值。经编码图像(coded picture)是图像的经编码表示,包括接入单元(access unit,AU)内的具有特定NAL单元头层标识符(nuh_layer_id)值的VCL NAL单元并包括该图像中的所有编码树单元(coding tree unit,CTU)。经解码图像(decoded picture)是通过将解码过程应用于经编码图像而产生的图像。

[0062] 输出层集(output layer set,OLS)是一组层,这组层中的一层或多层被指定为一

个或多个输出层。输出层是指定用于输出(例如,输出给显示器)的层。第0个OLS只包括最低层(具有最小层标识符的层),因此只包括输出层。视频参数集(video parameter set,VPS)是一种包括与整个视频相关的参数的数据单元。层间预测是通过参考一个参考层中的参考图像对当前层中的当前图像进行译码的机制,其中,当前图像和参考图像包括在同一个AU中,参考层的nuh_layer_id小于当前层的nuh_layer_id。

[0063] 档次(profile)是一组定义的译码工具,用于创建符合性或一致性码流。每个档次指定与该档次一致的所有解码器应支持的算法特征和限制的子集。层次和级别是在最大码率、最大亮度采样率、最大亮度图像大小、最小压缩比、条带的最大允许数量和分块的最大允许数量方面定义码流的约束条件。低层比高层的约束多,低级别比高级别的约束多。档次、层次和级别(profile,tier,and level,PTL)参数是一种提供档次、层次和/或级别信息的语法结构。

[0064] 时间可适性码流(temporal scalable bitstream)是在多层中进行译码的码流,提供不同的时间分辨率/帧率(例如,每个层经过译码以支持不同的帧率)。子层是时间可适性码流的时间可适性层,时间可适性码流包括具有特定时间标识符值的VCL NAL单元和关联的非VCL NAL单元。例如,时间子层是包括与指定帧率关联的视频数据的层。子层表示是包括特定子层和较低子层的NAL单元的码流的子集。因此,可以组合一个或多个时间子层以实现子层表示,该子层表示可以被解码以产生具有指定帧率的视频序列。OLS索引是唯一标识对应OLS的索引。时间标识符(TemporalID或时间ID)是一个表示数据对应于视频序列中的时间位置的数据元素。AU中的所有VCL NAL单元应具有相同的TemporalId值。经编码图像、PU或AU的TemporalId值是该经编码图像、PU或AU的VCL NAL单元的TemporalId值。子层表示的TemporalId值是子层表示中的所有VCL NAL单元的最大TemporalId值。

[0065] 最高子层表示是离基本子层最远的子层表示。也就是说,其它子层都在最高子层表示之下。

[0066] 子码流提取过程是一个从码流中移除不属于由目标OLS索引和目标最高时间ID确定的目标集合的NAL单元的过程。子码流提取过程会产生输出子码流,该输出子码流包括码流中属于目标集合的NAL单元。

[0067] HRD是在编码器上运行的解码器模型,它检查编码过程产生的码流的可变性,以验证是否符合指定约束。码流一致性测试(bitstream conformance test)是一种确定经编码码流是否符合通用视频编码(Versatile Video Coding,VVC)等标准的测试。HRD参数是初始化和/或定义HRD的操作条件的语法元素。序列级HRD参数是适用于整个编码视频序列的HRD参数。最大HRD时间ID(hrd_max_tid[i])表示HRD参数包括在第i个OLS HRD参数集中的最高子层表示的时间ID。通用HRD参数(general_hrd_parameters)语法结构是一种包括序列级HRD参数的语法结构。操作点(operation point,OP)是OLS的时间子集,由OLS索引和最大时间ID标识。测试OP(targetOp)是在HRD执行一致性测试时选择的OP。目标OLS是在从码流中提取时选择的OLS。解码单元HRD参数存在标志(decoding_unit_hrd_params_present_flag)表示对应的HRD参数是在解码单元(decoding unit,DU)级还是在AU级操作。编码图像缓冲区(coded picture buffer,CPB)是HRD中的先进先出缓冲区,包括解码顺序中的在码流一致性检查过程中使用的经编码图像。解码图像缓冲区(decoded picture buffer,DPB)是用于保存经解码图像以供参考、输出重排序和/或输出延迟的缓冲区。

[0068] 解码图像缓冲区(decoded picture buffer,DPB)参数是一种提供DPB大小并可选地提供最大图像重排序次数和最大延迟信息的语法结构。最大图像重排序次数和最大延迟信息可以统一称为MRML。

[0069] 本文中使用的以下缩略语:编码树块(Coding Tree Block,CTB)、编码树单元(Coding Tree Unit,CTU)、编码单元(Coding Unit,CU)、编码视频序列(Coded Video Sequence,CVS)、联合视频专家组(Joint Video Experts Team,JVET)、运动约束分块集(Motion-Constrained Tile Set,MCTS)、最大传输单元(Maximum Transfer Unit,MTU)、网络抽象层(Network Abstraction Layer,NAL)、图像顺序编号(Picture Order Count,POC)、图像参数集(Picture Parameter Set,PPS)、原始字节序列载荷(Raw Byte Sequence Payload,RBSP)、序列参数集(Sequence Parameter Set,SPS)、通用视频编码(Versatile Video Coding,VVC)和工作草案(Working Draft,WD)。

[0070] 图1为对视频信号进行译码的示例性操作方法100的流程图。具体地,在编码器侧对视频信号进行编码。编码过程通过使用各种机制来压缩视频信号,以减小视频文件大小。文件较小使得压缩视频文件可以发送给用户,同时降低相关的带宽开销。然后,解码器对压缩视频文件进行解码,以重建原始视频信号向终端用户进行显示。解码过程通常是编码过程的逆过程,以使解码器重建的视频信号可以与编码器侧的视频信号保持一致。

[0071] 在步骤101中,将视频信号输入到编码器中。例如,视频信号可以是存储在存储器中的未压缩视频文件。又如,视频文件可以由摄像机等视频捕获设备捕获,并且进行编码以支持视频的直播流传输。视频文件可以同时包括音频分量和视频分量。视频分量包括一系列图像帧。这些图像帧按顺序观看时,给人以运动的视觉效果。这些帧包括以光表示的像素,在本文中称为亮度分量(或亮度样本),还包括以颜色表示的像素,称为色度分量(或颜色样本)。在一些示例中,这些帧还可以包括深度值,以支持三维观看。

[0072] 在步骤103中,将视频分割成块。分割包括将每帧中的像素细分成正方形块和/或矩形块进行压缩。例如,在高效视频编码(High Efficiency Video Coding,HEVC)(还称为H.265和MPEG-H第2部分)中,首先将帧分成编码树单元(coding tree unit,CTU),即预定义大小(例如64个像素×64个像素)的块。这些CTU包括亮度样本和色度样本。编码树可以用于将CTU分成块,然后重复细分这些块,直到获得支持进一步编码的配置。例如,帧的亮度分量可以细分到各个块包括相对均匀的亮度值。此外,帧的色度分量可以细分到各个块包括相对均匀的色值因此,分割机制因视频帧的内容而异。

[0073] 在步骤105中,使用各种压缩机制对在步骤103中分割得到的图像块进行压缩。例如,可以使用帧间预测和/或帧内预测。帧间预测的设计利用了公共场景中的对象往往出现在连续帧中这一事实。这样一来,描述参考帧中的对象的块不需要在相邻帧中重复描述。具体地,一个对象(例如一张桌子)可能在多个帧中保持在固定位置上。因此,该桌子被描述一次,而且相邻帧可以重新参考所述参考帧。模式匹配机制可以用于在多个帧上匹配对象。此外,由于对象移动或相机移动等原因,移动对象可以跨多个帧表示。在一个具体示例中,视频可以在多个帧上显示跨屏幕移动的汽车。运动矢量可以用来描述这种移动。运动矢量是一个二维矢量,提供对象在一个帧中的坐标到该对象在参考帧中的坐标的偏移。因此,帧间预测可以将当前帧中的图像块编码为运动矢量集,这些运动矢量表示当前帧中的图像块与参考帧中的对应块的偏移。

[0074] 帧内预测对公共帧中的块进行编码。帧内预测利用亮度分量和色度分量往往聚集在一个帧中这一事实。例如，一棵树某个部分的一片绿色往往与类似的几片绿色相邻。帧内预测使用多种方向性预测模式（例如，HEVC中有33种）、平面模式和直流（direct current, DC）模式。这些方向性模式表示当前块中的样本与对应方向上的相邻块中的样本相似/相同。平面模式表示一行/列（例如平面）中的一系列块可以根据该行的边缘上的相邻块进行插值。平面模式实际上通过使用变化值的相对恒定的斜率表示光/颜色跨行/列的平稳过渡。DC模式用于边界平滑，并表示块与所有相邻块中的样本的平均值相似/相同，这些相邻块与方向性预测模式的角度方向关联。相应地，帧内预测块可以将图像块表示为各种关系预测模式值而不是表示为实际值。此外，帧间预测块可以将图像块表示为运动矢量值而不是表示为实际值。在任一种情况下，预测块在一些情况下可能都无法准确表示图像块。任何差值都存储在残差块中。变换可以应用于这些残差块以进一步压缩文件。

[0075] 在步骤107中，可以使用各种滤波技术。在HEVC中，滤波器是根据环内滤波方案使用的。上文描述的基于块的预测可能会在解码器侧产生块状图像。此外，基于块的预测方案可以对块进行编码，然后重建经编码块，以便后续用作参考块。环内滤波方案迭代地将噪声抑制滤波器、去块效应滤波器、自适应环路滤波器和样本自适应偏移（sample adaptive offset, SAO）滤波器应用于块/帧。这些滤波器减少了块伪影，这样可以准确地重建经编码文件。此外，这些滤波器减少了重建参考块中的伪影，这样伪影不太可能在根据重建参考块编码的后续块中产生其它伪影。

[0076] 一旦视频信号完成分割、压缩和滤波，则在步骤109中，将所得数据编码到码流中。码流包括上文描述的数据以及支持在解码器侧进行适当的视频信号重建所需要的任何指示数据。例如，这些数据可以包括分割数据、预测数据、残差块和提供译码指令给解码器的各种标志。码流可以存储在存储器中，以便根据请求向解码器发送。码流还可以广播和/或组播到多个解码器。码流的创建是一个迭代过程。相应地，步骤101、步骤103、步骤105、步骤107和步骤109可以在多个帧和块中连续和/或同时执行。图1所示的顺序是为了清楚和便于论述的目的，并非旨在将视频译码过程限制于特定顺序。

[0077] 在步骤111中，解码器接收码流并开始解码过程。具体地，解码器使用熵解码方案将码流转换为对应的语法数据和视频数据。在步骤111中，解码器使用码流中的语法数据来确定帧的分割部分。分割应该与步骤103中的块分割的结果匹配。下面描述在步骤111中使用的熵编码/解码。编码器在压缩过程中做出许多选择，例如，根据一个或多个输入图像中的值的空间放置从若干个可能选择中选择块分割方案。指示确切的选择可能会使用大量的位元（bin）。本文所使用的“位元”是一个作为变量的二进制值（例如可以因内容而异的比特值）。熵编码使得编码器丢弃任何明显不适合特定情况的选项，从而留下一组可用选项。然后，为每个可用选项分配一个码字。码字的长度取决于可用选项的数量（例如，1个位元对应2个选项，2个位元对应3到4个选项，以此类推）。然后，编码器对所选选项的码字进行编码。这种方案减小了码字的大小，这是因为码字与预期的一样大，从而唯一地指示从可用选项的小子集中进行选择，而不是唯一地指示从所有可能选项的可能大集合中进行选择。然后，解码器通过以与编码器类似的方式确定这组可用选项对这一选择进行解码。通过确定这组可用选项，解码器可以读取码字并确定编码器做出的选择。

[0078] 在步骤113中，解码器执行块解码。具体地，解码器使用逆变换，以产生残差块。然

后,解码器使用残差块和对应的预测块,以根据分割来重建图像块。预测块可以包括编码器在步骤105中生成的帧内预测块和帧间预测块。接着,根据在步骤111中确定的分割数据将重建图像块放置在重建视频信号的帧中。还可以通过上述熵编码在码流中指示用于步骤113的语法。

[0079] 在步骤115中,以类似于编码器侧的步骤107的方式对重建视频信号的帧执行滤波。例如,噪声抑制滤波器、去块效应滤波器、自适应环路滤波器和SAO滤波器可以应用于帧,以去除块伪影。一旦对帧进行了滤波,则在步骤117中,可以将视频信号输出到显示器,以供终端用户观看。

[0080] 图2为用于视频译码的示例性编码和解码(编解码)系统200的示意图。具体地,编解码系统200提供功能来支持操作方法100的实现。编解码系统200广义地用于描述编码器和解码器中都使用的组件。编解码系统200接收视频信号并对视频信号进行分割,如参照操作方法100中的步骤101和步骤103所述,得到分割后的视频信号201。然后,编解码系统200在充当编码器时将分割后的视频信号201压缩到经编码码流中,如参照方法100中的步骤105、步骤107和步骤109所述。编解码系统200在充当解码器时从码流中生成输出视频信号,如参照操作方法100中的步骤111、步骤113、步骤115和步骤117所述。编解码系统200包括通用译码器控制组件211、变换缩放和量化组件213、帧内估计组件215、帧内预测组件217、运动补偿组件219、运动估计组件221、缩放和逆变换组件229、滤波器控制分析组件227、环内滤波器组件225、解码图像缓冲区组件223以及标头格式和上下文自适应二进制算术编码(context adaptive binary arithmetic coding,CABAC)组件231。这些组件如图所示相耦合。在图2中,黑线表示待编码/解码数据的移动,而虚线表示控制其它组件操作的控制数据的移动。编解码系统200中的组件都可以存在于编码器中。解码器可以包括编解码系统200中的组件的子集。例如,解码器可以包括帧内预测组件217、运动补偿组件219、缩放和逆变换组件229、环内滤波器组件225和解码图像缓冲区组件223。下面对这些组件进行描述。

[0081] 分割后的视频信号201是捕获到的已经通过编码树分割成像素块的视频序列。编码树使用各种划分模式将像素块细分成较小的像素块。然后,这些块可以进一步细分成较小的块。这些块可以称为编码树上的节点。较大的父节点划分成较小的子节点。节点进行细分的次数称为节点/编码树的深度。在一些情况下,划分得到的块可以包括在编码单元(coding unit,CU)中。例如,CU可以是CTU的子部分,包括亮度块、一个或多个红色差色度(Cr)块和一个或多个蓝色差色度(Cb)块以及CU对应的语法指令。划分模式可以包括二叉树(binary tree,BT)、三叉树(triple tree,TT)和四叉树(quad tree,QT),它们用于根据所使用的划分模式将节点分别分割成不同形状的2个、3个或4个子节点。将分割后的视频信号201转发给通用译码器控制组件211、变换缩放和量化组件213、帧内估计组件215、滤波器控制分析组件227和运动估计组件221进行压缩。

[0082] 通用译码器控制组件211用于根据应用约束条件做出与将视频序列中的图像译码到码流中相关的决策。例如,通用译码器控制组件211管理码率/码流大小相对于重建质量的优化。这些决策可以根据存储空间/带宽可用性和图像分辨率请求做出。通用译码器控制组件211还根据传输速度来管理缓冲区利用率,以缓解缓存欠载和超载问题。为了解决这些问题,通用译码器控制组件211管理由其它组件进行的分割、预测和滤波。例如,通用译码器控制组件211可以动态增加压缩复杂度以提高分辨率和带宽利用率,或者降低压缩复杂度

以降低分辨率和带宽利用率。因此,通用译码器控制组件211控制编解码系统200中的其它组件来平衡视频信号重建质量与码率。通用译码器控制组件211生成控制数据,这些控制数据用于控制其它组件的操作。还将控制数据转发给标头格式和CABAC组件231,以编码到码流中,从而指示解码器进行解码所使用的参数。

[0083] 还将分割后的视频信号201发送给运动估计组件221和运动补偿组件219进行帧间预测。分割后的视频信号201的帧或条带可以被分成多个视频块。运动估计组件221和运动补偿组件219相对于一个或多个参考帧中的一个或多个块对所接收到的视频块执行帧间预测译码,以提供时间预测。编解码系统200可以执行多个译码过程,以便为每个视频数据块选择合适的译码模式,等等。

[0084] 运动估计组件221和运动补偿组件219可以高度集成,但是出于概念目的,单独说明。由运动估计组件221执行的运动估计是生成运动矢量的过程,其中,这些运动矢量用于估计视频块的运动。例如,运动矢量可以表示经编码对象相对于预测块的位移。预测块是在像素差方面发现与待编码块高度匹配的块。预测块还可以称为参考块。这种像素差可以通过绝对差异和(sum of absolute difference, SAD)、平方差异和(sum of square difference, SSD)或其它差异度量来确定。HEVC使用几个经编码对象,包括CTU、编码树块(coding tree block, CTB)和CU。例如,CTU可以被分成CTB,CTB然后被分成CB,CB包含在CU中。CU可以被编码为包括预测数据的预测单元(prediction unit, PU)和/或包括CU的变换残差数据的变换单元(transform unit, TU)。运动估计组件221使用率失真分析作为率失真优化过程一部分来生成运动矢量、PU和TU。例如,运动估计组件221可以确定当前块/帧的多个参考块、多个运动矢量等,并且可以选择具有最佳率失真特性的参考块、运动矢量等。最佳率失真特性平衡了视频重建的质量(例如,压缩造成的数据丢失量)和译码效率(例如,最终编码的大小)。

[0085] 在一些示例中,编解码系统200可以计算存储在解码图像缓冲区组件223中的参考图像的子整数像素位置的值。例如,视频编解码系统200可以对参考图像的四分之一像素位置、八分之一像素位置或其它分数像素位置的值进行插值。因此,运动估计组件221可以相对于整像素位置和分数像素位置执行运动搜索,并输出具有分数像素精度的运动矢量。运动估计组件221通过将PU的位置与参考图像的预测块的位置进行比较,计算经帧间译码条带中的视频块的PU的运动矢量。运动估计组件221将计算得到的运动矢量作为运动数据输出到标头格式和CABAC组件231进行编码,并作为运动数据输出到运动补偿组件219。

[0086] 运动补偿组件219执行的运动补偿可以涉及根据运动估计组件221所确定的运动矢量获取或生成预测块。在一些示例中,运动估计组件221和运动补偿组件219可以在功能上集成。在接收到当前视频块的PU的运动矢量之后,运动补偿组件219可以定位运动矢量指向的预测块。然后,从被译码的当前视频块的像素值中减去预测块的像素值,得到像素差,如此形成残差视频块。一般而言,运动估计组件221相对于亮度分量执行运动估计,运动补偿组件219将根据亮度分量计算到的运动矢量用于色度分量和亮度分量。将预测块和残差块转发到变换缩放和量化组件213。

[0087] 还将分割后的视频信号201发送给帧内估计组件215和帧内预测组件217。与运动估计组件221和运动补偿组件219一样,帧内估计组件215和帧内预测组件217可以高度集成,但是出于概念目的,单独说明。帧内估计组件215和帧内预测组件217相对于当前帧中的

各块对当前块进行帧内预测,以替代如上所述的由运动估计组件221和运动补偿组件219在各帧之间执行的帧间预测。具体地,帧内估计组件215确定帧内预测模式以对当前块进行编码。在一些示例中,帧内估计组件215从多个测试的帧内预测模式中选择合适的帧内预测模式来对当前块进行编码。然后,将选定的帧内预测模式转发给标头格式和CABAC组件231进行编码。

[0088] 例如,帧内估计组件215对各种测试的帧内预测模式进行率失真分析来计算率失真值,并在测试的模式中选择具有最佳率失真特性的帧内预测模式。率失真分析通常确定经编码块与经编码以产生经编码块的原始未编码块之间的失真(或误差)量,以及确定用于产生经编码块的码率(例如比特数)。帧内估计组件215根据各种经编码块的失真和速率计算比率,以确定表现出块的最佳率失真值的帧内预测模式。另外,帧内估计组件215可以用于根据率失真优化(rate-distortion optimization,RDO),使用深度建模模式(depth modeling mode,DMM)对深度图像中的深度块进行译码。

[0089] 帧内预测组件217在编码器上实现时可以根据由帧内估计组件215确定的选定帧内预测模式从预测块中生成残差块,或者在解码器上实现时可以从码流中读取残差块。残差块包括预测块与原始块之间的差值,表示为矩阵。然后,将残差块转发给变换缩放和量化组件213。帧内估计组件215和帧内预测组件217可以对亮度分量和色度分量都进行操作。

[0090] 变换缩放和量化组件213用于进一步压缩残差块。变换缩放和量化组件213将离散余弦变换(discrete cosine transform,DCT)、离散正弦变换(discrete sine transform,DST)等变换或者概念上类似的变换应用于残差块,从而产生包括残差变换系数值的视频块。还可以使用小波变换、整数变换、子带变换或其它类型的变换。变换可以将残差信息从像素值域转换到变换域,例如频域。变换缩放和量化组件213还用于根据频率等对变换残差信息进行比例缩放。这种比例缩放涉及将比例缩放因子应用于残差信息,使得在不同的粒度下量化不同的频率信息,这可能会影响重建视频的最终视觉质量。变换缩放和量化组件213还用于量化变换系数以进一步降低码率。量化过程可以减小与部分或全部系数相关的位深度。量化程度可以通过调整量化参数来修改。在一些示例中,变换缩放和量化组件213随后可以对包括量化变换系数的矩阵执行扫描。将量化变换系数转发给标头格式和CABAC组件231,以将量化变换系数编码到码流中。

[0091] 缩放和逆变换组件229应用与变换缩放和量化组件213相反的操作以支持运动估计。缩放和逆变换组件229应用逆缩放、逆变换和/或反量化以重建像素域中的残差块。例如,残差块后续用作参考块。该参考块可以成为另一当前块的预测块。运动估计组件221和/或运动补偿组件219可以通过将残差块添加回对应的预测块来计算参考块,以用于后续块/帧的运动估计中。将滤波器应用于重建参考块,以减少在缩放、量化和变换过程中产生的伪影。当预测后续块时,这些伪影可能会使预测不准确(并产生额外的伪影)。

[0092] 滤波器控制分析组件227和环内滤波器组件225将滤波器应用于残差块和/或重建图像块。例如,可以将来自缩放和逆变换组件229的变换残差块与来自帧内预测组件217和/或运动补偿组件219的对应预测块组合以重建原始图像块。然后,可以将滤波器应用于重建图像块。在一些示例中,滤波器还可以应用于残差块。与图2中的其它组件一样,滤波器控制分析组件227和环内滤波器组件225高度集成,可以一起实现,但出于概念目的分开示出。将应用于重建参考块的滤波器应用于特定空间区域,这些滤波器包括多个参数以调整使用这

些滤波器的方式。滤波器控制分析组件227对重建参考块进行分析,以确定需要使用这些滤波器的位置并设置对应的参数。将这些数据作为滤波器控制数据转发给标头格式和CABAC组件231进行编码。环内滤波器组件225根据滤波器控制数据使用这些滤波器。这些滤波器可以包括去块效应滤波器、噪声抑制滤波器、SAO滤波器和自适应环路滤波器。这些滤波器可以根据示例应用于空域/像素域(例如重建像素块)或频域中。

[0093] 当作为编码器操作时,解码图像缓冲区组件223存储经滤波的重建图像块、残差块和/或预测块,以供后续用于运动估计中,如上所述。当作为解码器操作时,解码图像缓冲区组件223存储经滤波的重建块并将其转发给显示器,这些重建块是输出视频信号的一部分。解码图像缓冲区组件223可以是任何能够存储预测块、残差块和/或重建图像块的存储设备。

[0094] 标头格式和CABAC组件231从编解码系统200中的各种组件接收数据,并将这些数据编码到经编码码流中,以发送给解码器。具体地,标头格式和CABAC组件231生成各种标头以对控制数据(例如通用控制数据和滤波器控制数据)进行编码。此外,将预测数据(包括帧内预测数据和运动数据)以及以量化变换系数数据为形式的残差数据都编码到码流中。最终的码流包括解码器重建原始分割后的视频信号201所需要的所有信息。这些信息还可以包括帧内预测模式索引表(还称为码字映射表)、各种块的编码上下文的定义、最可能帧内预测模式的指示、分割信息的指示等。这些数据可以通过熵编码来编码。例如,这些信息可以通过上下文自适应可变长度编码(context adaptive variable length coding, CAVLC)、CABAC、基于语法的上下文自适应二进制算术编码(syntax-based context-adaptive binary arithmetic coding, SBAC)、概率区间分割熵(probability interval partitioning entropy, PIPE)编码或其它熵编码技术来编码。在熵编码之后,可以将经编码码流发送给另一设备(例如视频解码器)或存档以供后续发送或检索。

[0095] 图3为示例性视频编码器300的框图。视频编码器300可以用于实现编解码系统200的编码功能和/或执行操作方法100中的步骤101、步骤103、步骤105、步骤107和/或步骤109。编码器300对输入视频信号进行分割,得到分割后的视频信号301,其基本上类似于分割后的视频信号201。然后,编码器300中的组件压缩分割后的视频信号301并将其编码到码流中。

[0096] 具体地,将分割后的视频信号301转发给帧内预测组件317进行帧内预测。帧内预测组件317可以基本上类似于帧内估计组件215和帧内预测组件217。还将分割后的视频信号301转发给运动补偿组件321,以根据解码图像缓冲区组件323中的参考块进行帧间预测。运动补偿组件321可以基本上类似于运动估计组件221和运动补偿组件219。将来自帧内预测组件317和运动补偿组件321的预测块和残差块转发给变换和量化组件313进行残差块的变换和量化。变换和量化组件313可以基本上类似于变换缩放和量化组件213。将变换量化残差块和对应的预测块(与相关的控制数据一起)转发给熵编码组件331,以译码到码流中。熵编码组件331可以基本上类似于标头格式和CABAC组件231。

[0097] 还将变换量化残差块和/或对应的预测块从变换和量化组件313转发给逆变换和反量化组件329,以重建为参考块供运动补偿组件321使用。逆变换和反量化组件329可以基本上类似于缩放和逆变换组件229。还将环内滤波器组件325中的环内滤波器应用于残差块和/或重建参考块,具体取决于示例。环内滤波器组件325可以基本上类似于滤波器控制分

析组件227和环内滤波器组件225。环内滤波器组件325可以包括多个滤波器,如参照环内滤波器组件225所述。然后,将经滤波的块存储在解码图像缓冲区组件323中,以作为参考块供运动补偿组件321使用。解码图像缓冲区组件323可以基本上类似于解码图像缓冲区组件223。

[0098] 图4为示例性视频解码器400的框图。视频解码器400可以用于实现编解码系统200的解码功能和/或执行操作方法100中的步骤111、步骤113、步骤115和/或步骤117。解码器400从编码器300等接收码流,并根据所述码流生成重建输出视频信号,以向终端用户显示。

[0099] 码流由熵解码组件433接收。熵解码组件433用于执行熵解码方案,例如,CAVLC、CABAC、SBAC、PIPE编码或其它熵编码技术。例如,熵解码组件433可以使用标头信息来提供上下文以解析在码流中编码为码字的附加数据。经解码信息包括对视频信号进行解码所需的任何信息,例如,通用控制数据、滤波器控制数据、分割信息、运动数据、预测数据和残差块中的量化变换系数。将量化变换系数转发给逆变换和反量化组件429,以重建成残差块。逆变换和反量化组件429可以类似于逆变换和反量化组件329。

[0100] 将重建残差块和/或预测块转发给帧内预测组件417,以根据帧内预测操作重建为图像块。帧内预测组件417可以类似于帧内估计组件215和帧内预测组件217。具体地,帧内预测组件417使用预测模式来定位帧中的参考块,并将残差块应用于上述结果以重建帧内预测图像块。将重建帧内预测图像块和/或残差块以及对应的帧间预测数据通过环内滤波器组件425转发给解码图像缓冲区组件423。解码图像缓冲区组件423和环内滤波器组件425可以基本上分别类似于解码图像缓冲区组件223和环内滤波器组件225。环内滤波器组件425对重建图像块、残差块和/或预测块进行滤波。这些信息存储在解码图像缓冲区组件423中。将来自解码图像缓冲区组件423的重建图像块转发给运动补偿组件421进行帧间预测。运动补偿组件421可以基本上类似于运动估计组件221和/或运动补偿组件219。具体地,运动补偿组件421使用参考块的运动矢量来生成预测块,并将残差块应用于上述结果以重建图像块。还可以将所得到的重建块转发通过环内滤波器组件425转发给解码图像缓冲区组件423。解码图像缓冲区组件423继续存储其它重建图像块。这些重建图像块可以通过分割信息重建为帧。这些帧还可以放置在一个序列中。该序列作为重建输出视频信号输出到显示器。

[0101] 考虑到上述情况,视频压缩技术通过执行空间(帧内)预测和/或时间(帧间)预测来减少或去除视频序列中固有的冗余。对于基于块的视频译码,视频条带(即视频图像或视频图像的一部分)可以被分割成视频块,这些视频块还可以称为树块(treeblock)、编码树块(coding tree block,CTB)、编码树单元(coding tree unit,CTU)、编码单元(coding unit,CU)和/或编码节点(coding node)。图像中的经帧内编码(I)条带中的视频块是相对于同一图像中的相邻块中的参考样本使用空间预测进行编码的。图像中的经帧间编码(P或B)条带中的视频块可以相对于同一图像中的相邻块中的参考样本使用空间预测,也可以是相对于其它参考图像中的参考样本使用时间预测。图像可以称为帧,参考图像可以称为参考帧。POC是与每个图像关联的变量,该变量唯一地标识编码层视频序列(coded layer video sequence,CLVS)中的所有图像中的关联图像。POC表示何时从DPB输出关联图像,并表示关联图像在输出顺序中的位置相对于同一CLVS中的要从DPB输出的其它图像在输出顺序中的位置的关系。标志是一个变量或单比特语法元素,可以取两个可能值之一:0或1。

[0102] 空间预测或时间预测为待编码块产生预测块。残差数据表示待编码原始块与预测块之间的像素差。经帧间编码块是根据运动矢量和残差数据进行编码的,其中,运动矢量指向形成预测块的参考样本的块,残差数据表示经编码块和预测块之间的差值。经帧内编码块是根据帧内编码模式和残差数据进行编码的。为了进一步压缩,残差数据可以从像素域变换到变换域,从而产生可以接着进行量化的残差变换系数。量化变换系数最初以二维阵列排列,并且可以按顺序扫描,以产生变换系数的一维矢量。熵编码可以用于实现进一步压缩。

[0103] 图像和视频压缩发展迅速,由此产生了各种编码标准。这些视频编码标准包括ITU-TH.261、国际标准化组织/国际电工委员会(International Organization for Standardization/International Electrotechnical Commission,ISO/IEC)MPEG-1第2部分、ITU-TH.262或ISO/IEC MPEG-2第2部分、ITU-T H.263、ISO/IEC MPEG-4第2部分、高级视频编码(Advanced Video Coding,AVC)(还称为ITU-T H.264或ISO/IEC MPEG-4第10部分)以及高效视频编码(High Efficiency Video Coding,HEVC)(还称为ITU-T H.265或MPEG-H第2部分)。AVC包括可适性视频编码(Scalable Video Coding,SVC)、多视图视频编码(Multiview Video Coding,MVC)和多视图视频编码加深度(Multiview Video Coding plus Depth,MVC+D)、3D AVC(3D-AVC)等扩展版本。HEVC包括可适性HEVC(Scalable HEVC,SHVC)、多视图HEVC(Multiview HEVC,MV-HEVC)、3D HEVC(3D-HEVC)等扩展版本。

[0104] 还有一个新的视频编码标准,名为通用视频编码(Versatile Video Coding,VVC),正在由ITU-T和ISO/IEC的联合视频专家组(joint video experts team,JVET)开发。虽然VVC标准有几个工作草案,但本文特别参考了VVC的一个工作草案(WD),即B.Bross、J.Chen和S.Liu在2019年3月27日的第13次JVET会议上发表的JVET-N1001-v3“通用视频编码(草案5)(Versatile Video Coding(Draft 5))”(VVC草案5)。

[0105] 视频译码中的可适性通常使用多层译码技术来支持。多层码流包括基本层(base layer,BL)和一个或多个增强层(enhancement layer,EL)。可适性的示例包括空间可适性、质量/信噪比(signal to noise ratio,SNR)可适性、多视图可适性,等等。当使用的是多层译码技术时,图像或其一部分可以通过以下方式译码:(1)不使用参考图像,即,使用帧内预测,(2)参考同一层中的参考图像,即,使用帧间预测,或(3)参考其它一个或多个层中的参考图像,即,使用层间预测。当前图像的层间预测使用的参考图像称为层间参考图像(inter-layer reference picture,ILRP)。

[0106] 图5为基于层的预测500的一个示例的示意图。例如,在块压缩步骤105、块解码步骤113、运动估计组件221、运动补偿组件219、运动补偿组件321和/或运动补偿组件421处执行这种预测以确定运动矢量(motion vector,MV)。基于层的预测500与单向帧间预测和/或双向帧间预测共存,但也在不同层中的图像之间执行。

[0107] 基于层的预测500应用于不同层中的图像511、512、513和514与图像515、516、517和518之间。在所示的示例中,图像511、512、513和514是层N+1 532的一部分,而图像515、516、517和518是层N 531的一部分。层N 531和/或层N+1 532等是一组图像,这些图像都与类似的大小、质量、分辨率、信噪比、能力等类似的特征值关联。在所示的示例中,与层N 531相比,层N+1 532与更大的图像大小关联。因此,在本示例中,层N+1 532中的图像511、512、513和514比层N 531中的图像515、516、517和518大(例如,高度和宽度更大,因此样本更

多)。但是,这些图像可以根据其它特征分布在层N+1 532和层N 531之间。虽然只示出了两层:层N+1 532和层N 531,但一组图像可以根据关联特征划分成任意数量的层。层N+1 532和层N 531也可以通过层ID表示。层ID是与图像关联的数据项,表示图像是指示层的一部分。因此,图像511至518中的每个图像可以与对应的层ID关联,以表示层N+1 532或层N 531中的哪个层包括对应的图像。

[0108] 不同层531和532中的图像511至518被配置为交替显示。因此,不同层531和532中的图像511至518可以共用同一时间标识符(identifier, ID)并可以包括在同一AU中。本文所使用的“AU”是与同一显示时间关联的从DPB输出的一个或多个经编码图像组成的集合。例如,如果需要较小的图像,则解码器可以解码图像515并在当前显示时间显示;如果需要较大的图像,则解码器可以解码图像511并在当前显示时间显示。因此,高层N+1 532中的图像511至514与低层N 531中的对应图像515至518包括基本上相同的图像数据(虽然图像大小不同)。具体地,图像511与图像515包括基本上相同的图像数据,图像512与图像516包括基本上相同的图像数据,以此类推。

[0109] 图像511至518可以通过参考同一层N 531或N+1 532中的其它图像511至518进行译码。参考同一层中的一个图像对另一个图像进行译码即为帧间预测523,这与单向帧间预测和/或双向帧间预测一致。帧间预测523由实线箭头表示。例如,图像513可以使用将层N+1 532中的图像511、512和/或514中的一个或两个图像作为参考的帧间预测523进行译码,其中,单向帧间预测使用一个图像作为参考和/或双向帧间预测使用两个图像作为参考。此外,图像517可以使用将层N 531中的图像515、516和/或518中的一个或两个图像作为参考的帧间预测523进行译码,其中,单向帧间预测使用一个图像作为参考和/或双向帧间预测使用两个图像作为参考。当在执行帧间预测523时将一个图像作为同一层中的另一个图像的参考时,该图像可以称为参考图像。例如,图像512可以是用于根据帧间预测523对图像513进行译码的参考图像。帧间预测523还可以称为多层上下文中的层内预测。因此,帧间预测523是通过参考与当前图像不同的参考图像中的指示样本对当前图像中的样本进行译码的机制,其中,参考图像和当前图像在同一层中。

[0110] 图像511至518也可以通过参考不同层中的其它图像511至518进行译码。这个过程称为层间预测521,由虚线箭头表示。层间预测521是一种通过参考一个参考图像中的指示样本对当前图像中的样本进行译码的机制,其中,当前图像和参考图像在不同的层中,因此具有不同的层ID。例如,低层N 531中的图像可以作为对高层N+1 532中的对应图像进行译码的参考图像。在一个具体的示例中,图像511可以根据层间预测521通过参考图像515进行译码。在这种情况下,图像515用作层间参考图像。层间参考图像是层间预测521使用的参考图像。在大多数情况下,层间预测521被约束为使得当前图像(例如图像511)只能使用同一AU中包括的且在低层中的一个或多个层间参考图像,例如图像515。当多层(例如,两层以上)可用时,层间预测521可以根据级别比当前图像低的多个层间参考图像对当前图像进行编码/解码。

[0111] 视频编码器可以使用基于层的预测500来通过帧间预测523和层间预测521的许多不同组合和/或排列对图像511至518进行编码。例如,图像515可以根据帧内预测进行译码。然后,通过将图像515用作参考图像,图像516至518可以根据帧间预测523进行译码。此外,通过将图像515用作层间参考图像,图像511可以根据层间预测521进行译码。然后,通过将

图像511用作参考图像,图像512至514可以根据帧间预测523进行译码。因此,参考图像可以用作不同译码机制的单层参考图像和层间参考图像。通过根据低层N 531图像对高层N+1 532图像进行译码,高层N+1 532可以避免使用帧内预测,帧内预测的译码效率比帧间预测523和层间预测521的译码效率低得多。因此,译码效率低的帧内预测可以只限于最小/最低质量的图像,因此只限于对最少量的视频数据进行译码。用作参考图像和/或层间参考图像的图像可以在参考图像列表结构中包括的一个或多个参考图像列表的条目中指示。

[0112] 图5中的每个AU 506可以包括几个图像。例如,一个AU 506可以包括图像511和515。另一个AU 506可以包括图像512和516。实际上,每个AU 506是与同一显示时间(例如同一时间ID)关联的一个或多个经编码图像的集合,这些图像从解码图像缓冲区(decoded picture buffer,DPB)输出(例如显示给用户)。每个AUD 508是用于指示AU(例如AU 508)起始或AU之间边界的指示符或数据结构。

[0113] 以前的H.26x视频编码系列支持根据针对单层译码的一个或多个档次(profile)实现一个或多个单独档次中的可适性。可适性视频编码(scalable video coding,SVC)是AVC/H.264的可适性扩展版本,支持空间可适性、时间可适性和质量可适性。对于SVC,标志(flag)在EL图像中的每个宏块(macroblock,MB)中指示,以表示EL MB是否使用低层中的并置块进行预测。基于并置块的预测可以包括纹理、运动矢量和/或译码模式。SVC的实现方式不能在其设计中直接重用未经修改的H.264/AVC实现方式。SVC EL宏块语法和解码过程与H.264/AVC语法和解码过程不同。

[0114] 可适性HEVC(scalable HEVC,SHVC)是HEVC/H.265标准的扩展版本,支持空间可适性和质量可适性;多视图HEVC(multiview HEVC,MV-HEVC)是HEVC/H.265的扩展版本,支持多视图可适性;3D HEVC(3D-HEVC)是HEVC/H.264的扩展版本,支持比MV-HEVC更先进、更高效的三维(three dimensional,3D)视频译码。需要说明的是,时间可适性是单层HEVC编解码器的组成部分。HEVC的多层扩展版本的设计采用了如下理念:用于层间预测的经解码图像仅来自同一接入单元(access unit,AU),并作为长期参考图像(long-term reference picture,LTRP),而且分配有一个或多个参考图像列表中的参考索引和当前层中的其它时间参考图像。层间预测(inter-layer prediction,ILP)是通过将参考索引的值设置为参考一个或多个参考图像列表中的一个或多个层间参考图像在预测单元(prediction unit,PU)级实现的。

[0115] 需要说明的是,参考图像重采样和空间可适性特征都需要对参考图像或其一部分进行重采样。参考图像重采样(reference picture resampling,RPR)可以在图像级或编码块级实现。然而,当RPR称为译码特征时,它是单层译码的特征。即便如此,从编解码器设计的角度来看,可以或甚至优选使用相同的重采样滤波器,以实现单层译码的RPR特征和多层译码的空间可适性特征。

[0116] 图6为使用输出层集(output layer set,OLS)的基于层的预测600的一个示例的示意图。例如,在块压缩步骤105、块解码步骤113、运动估计组件221、运动补偿组件219、运动补偿组件321和/或运动补偿组件421处执行这种预测以确定MV。基于层的预测600与单向帧间预测和/或双向帧间预测共存,但也在不同层中的图像之间执行。图6中的基于层的预测类似于图5中的基于层的预测。因此,为了简洁起见,基于层的预测不再赘述。

[0117] 图6中的编码视频序列(coded video sequence,CVS) 690中的一些层包括在OLS

中。OLS是一组层,这组层中的一层或多层被指定为输出层。输出层是被输出的OLS中的一层。图6描述了3个不同的OLS,即OLS 1、OLS 2和OLS 3。如图所示,OLS 1包括层N 631和层N+1 632。OLS 2包括层N 631、层N+1 632、层N+2 633和层N+3 634。OLS 3包括层N 631、层N+1 632和层N+2 633。虽然示出了3个OLS,但在实际应用中可以使用不同数量的OLS。层N+4 635等其它层不包括在所示的OLS中(例如,OLS 1、2和3),但可以包括在其它OLS中。

[0118] 不同的OLS都可以包括任意数量的层。生成不同的OLS是为了适应具有不同译码能力的各种不同设备的译码能力。例如,可以生成只包括2层的OLS 1,以适应具有相对有限译码能力的手机。另一方面,可以生成包括4层的OLS 2,以适应大屏电视,大屏电视相比于手机能够解码更高层。可以生成包括3层的OLS 3,以适应个人计算机、笔记本电脑或平板电脑,这些相比于手机能够解码更高层,但相比于大屏电视不能解码最高层。

[0119] 图6中的这些层可以相互独立。也就是说,每层都可以不使用层间预测(inter-layer prediction, ILP)进行译码。在这种情况下,这些层称为同播层(simulcast layer)。图6中的这些层中的一层或多层还可以使用ILP进行译码。这些层是否为同播层或者其中一些层否使用ILP进行译码由视频参数集(video parameter set, VPS)中的标志指示,下文提供更全面的论述。当一些层使用ILP时,各层之间的层依赖关系也会在VPS中指示。

[0120] 在一个实施例中,当这些层是同播层时,只选择一层进行解码和输出。在一个实施例中,当一些层使用ILP时,所有层(例如,整个码流)都指定要解码,并且其中一些层被指定为输出层。一个或多个输出层可以是,例如,(1)仅最高层,(2)所有层或(3)最高层加上一组指示的低层。例如,当最高层加上一组指示的低层通过VPS中的标志被指定为输出时,输出OLS 2中的层N+3 634(即最高层)和层N 631和层N+1 632(即低层)。

[0121] 如图6所示,每层包括任意数量的图像。例如,层N 631包括图像615至618,层N+1 632包括图像611至614,层N+2 633包括图像641至644,层N+3 634包括图像651至654,层N+4 635包括图像661至664。

[0122] 图7为用于时间可适性的示例性多层视频序列700的示意图。根据方法100等,多层视频序列700可以由编码器(例如,编解码系统200和/或编码器300)编码,并由解码器(例如,编解码系统200和/或解码器400)解码。将多层视频序列700包括在内是为了描述编码视频序列中的各层的示例性应用。例如,多层视频序列700可以用作单独的实施例,或者可以与参照基于层的预测500或600描述的技术组合。

[0123] 多层视频序列700包括子层710、720和730。子层是时间可适性码流的时间可适性层,时间可适性码流包括具有特定时间标识符值的VCL NAL单元(例如图像)和关联的非VCL NAL单元(例如支持参数)。子层710可以称为基本层,子层720和730可以称为增强层。如图所示,子层710包括第一帧率(例如每秒30帧)下的图像711。子层710是基本层,因为子层710包括基本/最低帧率。子层720包括与子层710中的图像711存在时间偏移的图像721。因此,可以组合子层710和子层720,这样产生的帧率统一比单独的子层710的帧率高。例如,子层710和720可以具有每秒60帧的组合帧率。因此,子层720提高了子层710的帧率。此外,子层730包括也与子层720和710中的图像721和711存在时间偏移的图像731。因此,子层730可以与子层720和710组合,以进一步增强子层710。例如,子层710、720和730可以具有每秒90帧的组合帧率。

[0124] 子层表示740可以通过组合子层710、720和/或730来动态创建。子层表示740是包

括特定子层和较低子层中的NAL单元的码流的子集。在所示的示例中,子层表示740包括图像741,图像741是子层710、720和730中的组合图像711、721和731。相应地,通过选择包括一组期望的子层710、720和/或730的子层表示740,多层视频序列700可以在时间上缩放到期望的帧率。子层表示740可以通过使用包括子层710、720和/或730作为层的OLS来创建。在这种情况下,选择子层表示740作为输出层。因此,时间可适性是可以使用多层机制实现的几种机制之一。

[0125] 在一个实施例中,子层710、720和730分别与时间标识符(temporal identifier, ID)关联,时间标识符可以表示为TemporalID。例如,子层710可以具有TemporalID 755,子层720可以具有TemporalID 757,子层730可以具有TemporalID 759。TemporalID 755、757和759可以用于唯一地标识每个子层。

[0126] 在一个实施例中,语法元素vps_max_sublayers_minus1+1表示可以存在于由VPS指定的层(例如,层631、632等)中的时间子层(例如,子层710、720、730等)的最大数量。在一个实施例中,vps_max_sublayers_minus1的取值范围应为0至6(包括端值)。因此,多层视频序列700包括范围为1至7的子层的最大数量。

[0127] 在图7所示的示例中,子层730是离子层710(即基本层)最远的子层。也就是说,图7中的层都在子层730之下。因此,图7中的子层730可以称为最高子层表示。

[0128] HEVC、SHVC、MV-HEVC和VVC等视频编码标准规定并指示了档次、层次和级别。档次、层次和级别指定了对码流的限制条件和对解码这些码流需要的能力的限制。档次、层次和级别还可以用于指示各种解码器实现方式之间的互操作点。档次是一组定义的译码工具,用于创建符合性或一致性码流。每个档次指定与该档次一致的所有解码器应支持的算法特征和限制的子集。

[0129] 级别是码流的一组约束条件(例如,最大亮度采样率、分辨率对应的最大码流等)。也就是说,级别是一组指示播放指定档次的码流所需的解码器性能的约束条件。

[0130] 级别划分成两种层次:主层次和高层次。主层次比高层次低。层次是为了处理在最大码率方面不同的应用。主层次针对大多数应用设计,而高层次针对要求非常高的应用设计。对应于级别的一组约束条件和/或层次在本文可以称为级别信息。

[0131] 层次对应的每个级别指定了对特定标准(例如HEVC、VVC)中的语法元素可能使用的值的一组限制。所有档次都使用相同的一组层次和级别定义,但各种实现方式可以支持不同的层次和层次中每个支持档次对应的不同级别。对于任何给定的档次,层次的级别通常对应于特定的解码器处理负载和存储能力。

[0132] 使用HEVC标准,下表1提供了不同档次对应的级别和层次的一个代表性示例。

[0133] 表1

级别	主 4:4:4 12 档次对应的 最大码率 (1000 比特/秒)		高吞吐量 4:4:4 16 帧内档次对 应的最大码率 (1000 比特/秒)		最高帧率下的示例性图像 分辨率 (MaxDPBsize)
	主层次	高层次	主层次	高层次	
[0134] 5.2	180,000	720,000	5,760,000	23,040,000	4,096×2,160@120.0
6	180,000	720,000	5,760,000	23,040,000	8,192×4,320@30.0
6.1	180,000	720,000	11,520,000	46,080,000	8,192×4,320@60.0
6.2	720,000	2,400,000	23,040,000	76,800,000	8,192×4,320@120.0

[0135] 档次、层次和级别(profile, tier, and level, PTL) 参数是一种提供上表所示的档次、层次和/或级别的语法结构(例如第i个语法结构profile_tier_level())。解码图像缓冲区(decoded picture buffer, DPB) 参数是一种提供DPB大小并可选地提供最大图像重排序次数和最大延迟信息的语法结构(例如第i个语法结构dpb_parameters())。最大图像重排序次数和最大延迟信息可以统一称为MRML。HRD参数是一种初始化和/或定义HRD操作条件的语法结构(例如, 第一个语法结构ols_timing_hrd_parameters())。下文提供了关于PTL参数、DPB参数和HRD参数中的一个或多个的更多详细内容。

[0136] 在一个实施例中, 最大图像重排序次数表示为dpb_max_num_reorder_pics[i], 其表示在Htid等于i时OLS中的图像的最大允许数量, 这些图像可以在解码顺序中位于OLS中的任一图像之前且在输出顺序中位于该图像之后。dpb_max_num_reorder_pics[i]的取值范围应为0至dpb_max_dec_pic_buffering_minus1[i] (包括端值)。当i大于0时, dpb_max_num_reorder_pics[i]应大于或等于dpb_max_num_reorder_pics[i-1]。当dpb_max_num_reorder_pics[i] (其中, i的范围为0至MaxSubLayersMinus1-1, 包括端值) 不存在时, 由于subLayerInfoFlag等于0, 因此dpb_max_num_reorder_pics[i]被推断为dpb_max_num_reorder_pics[MaxSubLayersMinus1]。Htid是一个标识待解码的最高时间子层的变量。

[0137] dpb_max_dec_pic_buffering_minus1[i]+1表示在Htid等于i时DPB的以图像存储缓冲区为单位的最大所需大小。dpb_max_dec_pic_buffering_minus1[i]的取值范围应为0至MaxDpbSize-1(包括端值), 其中, MaxDpbSize如VVC标准中的第A.4.2小节所述。当i大于0时, dpb_max_dec_pic_buffering_minus1[i]应大于或等于dpb_max_dec_pic_buffering_minus1[i-1]。当dpb_max_dec_pic_buffering_minus1[i] (其中, i的范围为0至MaxSubLayersMinus1-1, 包括端值) 不存在时, 由于subLayerInfoFlag等于0, 因此dpb_max_dec_pic_buffering_minus1[i]被推断为dpb_max_dec_pic_buffering_minus1[MaxSubLayersMinus1]。

[0138] 在一个实施例中, 最大延迟信息表示为dpb_max_latency_increase_plus1[i]。dpb_max_latency_increase_plus1[i]不等于0用于计算MaxLatencyPictures[i]的值, MaxLatencyPictures[i]表示在Htid等于i时OLS中的图像的最大允许数量, 这些图像可以在解码顺序中位于OLS中的任一图像之前且在输出顺序中位于该图像之后。

[0139] MaxLatencyPictures[i]=dpb_max_num_reorder_pics[i]+dpb_max_latency_increase_plus1[i]-1。

[0140] 但是, 即使当一个或多个层具有相同的PTL参数、HRD参数和/或DPB参数时(例如,

当所有层具有相同数量的子层时),也在多层码流中为每层(例如,层631、632等)指示级别信息(和其它可能的PTL参数)、HRD参数和DPB参数。因此,在多层码流中重复指示相同参数,这样产生了冗余并导致译码过程效率低。

[0141] 本文公开的技术确保不会在多层码流中为各层重复指示相同的档次、层次和级别(profile,tier,and level,PTL)参数、假设参考解码(hypothetical reference decoder,HRD)参数和解码图像缓冲区(decoded picture buffer,DPB)参数。仅指示相同的PTL参数、HRD参数和DPB参数一次减少了冗余并提高了多层视频码流的译码效率。因此,视频译码中的编码器/解码器(又称“编解码器”)相对于现有的编解码器有所改进。实际上,改进的视频译码过程在发送、接收和/或观看视频时提供了更好的用户体验。

[0142] 图8示出了视频码流800的一个实施例。本文使用的视频码流800还可以称为经编码视频码流、码流或其变体。如图8所示,码流800包括至少一个图像单元(picture unit,PU)801。虽然图8示出了其中3个PU 801,但在实际应用中,码流800可以包括不同数量的PU801。每个PU 801是指一组NAL单元,这些NAL单元根据指定的分类规则相互关联、在解码顺序中是连续的且正好包括一个经编码图像(例如图像814)。

[0143] 在一个实施例中,每个PU 801包括以下各项中的一个或多个:解码能力信息(decoding capability information,DCI)802、视频参数集(video parameter set,VPS)804、序列参数集(sequence parameter set,SPS)806、图像参数集(picture parameter set,PPS)808、图像头(picture header,PH)812和图像814。DCI 802、VPS 804、SPS 806和PPS 808一般都可以称为参数集。在一个实施例中,码流800还可以包括图8未示出的其它参数集,例如自适应参数集(adaption parameter set,APS)。APS是一种包括应用于0个或0个以上条带的语法元素的语法结构,0个或0个以上条带通过条带头中的0个或0个以上语法元素确定。

[0144] DCI 802还可以称为解码参数集(decoding parameter set,DPS)或解码器参数集,是一种包括应用于整个码流的语法元素的语法结构。DCI 802包括的参数在视频码流(例如码流800)的生存期内保持不变,视频码流的生存期可以转换为一段视频序列码流的生存期。DCI802可以包括档次、级和子档次信息,以确定最大复杂度互操作点。即使会话中发生视频序列拼接,也会保证永远不会超过最大复杂度互操作点。DCI 802还可选地包括约束标志,其中,约束标志表示按照这些标志值指示,约束视频码流使用某些特征。如此,码流可以被标记为不使用某些工具,这样可以在解码器实现方式中实现资源分配,等等。与所有参数集类似,DCI 802在首次被参考时存在,并由视频序列中的第一个图像参考,表示DCI 802必须在码流800中的第一NAL单元之间发送。虽然码流800可以包括多个DCI 802,但这些DCI中的语法元素值在被参考时必须一致。

[0145] VPS 804包括用于增强层的参考图像集构建的解码依赖关系或信息。VPS 804提供可适性序列的整体信息,包括提供了哪些类型的操作点、操作点的档次、层次和级别以及码流800的其它一些高级属性,其中,这些高级属性可以用作会话协商和内容选择等的基础。

[0146] 在一个实施例中,VPS 804包括标志850。在一个实施例中,标志850表示为vps_default_ptl_dpb_hrd_max_tid_flag。标志850可以被设置为或具有第一值(例如1)或第二值(例如0)。当标志850不存在时,其值被推断为1。具有第一值的标志850表示语法元素vps_ptl_max_tid[i]、vps_dpb_max_tid[i]和vps_hrd_max_tid[i]不存在于VPS 804中并被推

断为默认值`vps_max_sublayers_minus1`。也就是说,档次、层次和级别(profile,tier,and level,PTL)参数、解码图像缓冲区(decoded picture buffer,DPB)参数和假设参考解码(hypothetical reference decoder,HRD)参数对应的最高子层表示(例如子层730)的时间标识符(tid)不存在于VPS 804中。

[0147] `vps_ptl_max_tid[i]`表示存在于VPS 804中的第*i*个语法结构`profile_tier_level()`中的级别信息对应的最高子层表示(例如子层730)的TemporalId和存在于OLS(OLS索引为`olsIdx`,使得`vps_ols_ptl_idx[olsIdx]`等于*i*)中的最高子层表示的TemporalId。

[0148] `vps_dpb_max_tid[i]`表示DPB参数可以存在于VPS 804中的第*i*个语法结构`dpb_parameters()`的最高子层表示(例如子层730)的TemporalId。`vps_dpb_max_tid[i]`的取值范围必须为0至`vps_max_sublayers_minus1`(包括端值)。当`vps_dpb_max_tid[i]`不存在时,`vps_dpb_max_tid[i]`的值被推断为`vps_max_sublayers_minus1`。

[0149] `vps_hrd_max_tid[i]`表示HRD参数包括在第*i*个语法结构`ols_timing_hrd_parameters()`的最高子层表示(例如子层730)的TemporalId。在一个实施例中,第*i*个语法结构`ols_timing_hrd_parameters()`存在于VPS 804或序列参数集(sequence parameter set,SPS) 806中。`vps_hrd_max_tid[i]`的取值范围应为0至`vps_max_sublayers_minus1`(包括端值)。当`vps_hrd_max_tid[i]`不存在时,`vps_hrd_max_tid[i]`的值被推断为`vps_max_sublayers_minus1`。

[0150] `vps_max_sublayers_minus1+1`表示可以存在于由VPS 804指定的层(例如,层631、632、633等)中的时间子层(例如,子层710、720、730)的最大数量。在一个实施例中,`vps_max_sublayers_minus1`的取值范围为0至6(包括端值)。因此,子层的最大数量的范围为1至7。也就是说,多层码流中存在1层或7层,取决于`vps_max_sublayers_minus1`的值。

[0151] 具有第二值的标志850表示语法元素`vps_ptl_max_tid[i]`、`vps_dpb_max_tid[i]`和`vps_hrd_max_tid[i]`存在于VPS 804中。也就是说,档次、层次和级别(profile,tier,and level,PTL)参数、解码图像缓冲区(decoded picture buffer,DPB)参数和假设参考解码(hypothetical reference decoder,HRD)参数对应的最高子层表示(例如子层730)的时间标识符(tid)存在于VPS 804中。

[0152] 在一个实施例中,标志850表示为`vps_all_layers_same_num_sublayers_flag`。标志850可以被设置为或具有第一值(例如1)或第二值(例如0)。具有第一值的标志850表示参考VPS804的每个编码视频序列(coded video sequence,CVS) 816中的所有层具有一样多的时间子层。具有第二值的标志850表示参考VPS 804的每个CVS 816中的层可以或可以不具有一样多的时间子层。当标志850不存在时,其值被推断为1。在一个实施例中,标志850的名称和语义从`vps_all_layers_same_num_sublayers_flag`变为`vps_default_ptl_dpb_hrd_max_tid_flg`,只是为了控制VPS语法,而不需要所有层具有一样多的子层。

[0153] SPS 806包括图像序列(sequence of picture,SOP)中的所有图像共有的数据。SPS 806是一种包括适用于0个或0个以上完整CLVS的语法元素的语法结构,0个或0个以上完整CLVS通过PPS中的语法元素的内容确定,而且PPS由每个图像头中的语法元素参考。相比之下,PPS 808包括整个图像共有的数据。PPS 808是一种包括应用于0个或0个以上完整经编码图像的语法元素的语法结构,0个或0个以上完整经编码图像通过每个图像头(例如

PH812)中的语法元素确定。

[0154] 在一个实施例中, VPS 804和/或SPS 806包括以下SPS语法和语义, 以实现本文公开的一个或多个实施例。

[0155] VPS语法包括以下内容。

[0156]	video_parameter_set_rbsp() {	描述符
	vps_video_parameter_set_id	u(4)
	vps_max_layers_minus1	u(6)
	vps_max_sub_layers_minus1	u(3)
	if(vps_max_layers_minus1 > 0 && vps_max_sub_layers_minus1 > 0)	
	vps_all_layers_same_num_sub_layers_flag	u(1)
	if(vps_max_layers_minus1 > 0)	
	vps_all_independent_layers_flag	u(1)
	...	
	vps_num_ptls	u(8)
	for(i = 0; i < vps_num_ptls; i++) {	
	if(i > 0)	
	pt_present_flag[i]	u(1)
	if(vps_max_sub_layers_minus1 > 0 && !vps_all_layers_same_num_sub_layers_flag)	
	ptl_max_temporal_id[i]	u(3)
	}	
	while(!byte_aligned())	
	vps_ptl_byte_alignment_zero_bit /*等于 0*/	u(1)
	for(i = 0; i < vps_num_ptls; i++)	
	profile_tier_level(pt_present_flag[i], ptl_max_temporal_id[i])	
	for(i = 0; i < TotalNumOls; i++)	
	if(NumLayersInOls[i] > 1 && vps_num_ptls > 1)	
	ols_ptl_idx[i]	u(8)
	if(!vps_all_independent_layers_flag)	
	vps_num_dpb_params	ue(v)
	if(vps_num_dpb_params > 0) {	

[0157]

same_dpb_size_output_or_nonoutput_flag	u(1)
if(vps_max_sub_layers_minus1 > 0)	
vps_sub_layer_dpb_params_present_flag	u(1)
}	
for(i = 0; i < vps_num_dpb_params; i++) {	
dpb_size_only_flag[i]	u(1)
if(vps_max_sub_layers_minus1 > 0 && !vps_all_layers_same_num_sub_layers_flag)	
dpb_max_temporal_id[i]	u(3)
dpb_parameters(dpb_size_only_flag[i], dpb_max_temporal_id[i], vps_sub_layer_dpb_params_present_flag)	
}	
for(i = 0; i < vps_max_layers_minus1 && vps_num_dpb_params > 1; i++) {	
if(!vps_independent_layer_flag[i])	
layer_output_dpb_params_idx[i]	ue(v)
if(LayerUsedAsRefLayerFlag[i] && !same_dpb_size_output_or_nonoutput_flag)	
layer_nonoutput_dpb_params_idx[i]	ue(v)
}	
general_hrd_params_present_flag	u(1)
if(general_hrd_params_present_flag) {	
num_units_in_tick	u(32)
time_scale	u(32)
general_hrd_parameters()	
}	
vps_extension_flag	u(1)
if(vps_extension_flag)	
while(more_rbsp_data())	
vps_extension_data_flag	u(1)
rbp_trailing_bits()	
}	

[0158] SPS原始字节序列载荷(raw byte sequence payload,RBSP)语法包括以下内容。

[0159]

seq_parameter_set_rbsp() {	描述符
sps_decoding_parameter_set_id	u(4)
sps_video_parameter_set_id	u(4)
sps_max_sub_layers_minus1	u(3)
sps_reserved_zero_4bits	u(4)
sps_ptl_dpb_present_flag	u(1)
if(sps_ptl_dpb_present_flag)	
profile_tier_level(1, sps_max_sub_layers_minus1)	
gdr_enabled_flag	u(1)
sps_seq_parameter_set_id	ue(v)
chroma_format_idc	ue(v)
...	
log2_max_pic_order_cnt_lsb_minus4	ue(v)
poc_msb_in_rap_pics_flag	u(1)
if(poc_msb_in_rap_pics_flag > 0)	
poc_msb_len_minus1	ue(v)
if(sps_max_sub_layers_minus1 > 0)	
sps_sub_layer_dpb_params_flag	u(1)
if(sps_ptl_dpb_present_flag)	
dpb_parameters(0, sps_max_sub_layers_minus1, sps_sub_layer_dpb_params_flag)	
long_term_ref_pics_flag	u(1)
...	
sps_scaling_list_enabled_flag	u(1)
vui_parameters_present_flag	u(1)
if(vui_parameters_present_flag)	
vui_parameters()	
sps_extension_flag	u(1)
if(sps_extension_flag)	
while(more_rbsp_data())	
sps_extension_data_flag	u(1)
rbsp_trailing_bits()	
}	

[0160]

[0161] DPB语法包括以下内容。

[0162]	dpb_parameters(dpbSizeOnlyFlag, maxSubLayersMinus1, subLayerInfoFlag) {	描述符
	for(i = (subLayerInfoFlag ? 0 : maxSubLayersMinus1);	
	i <= maxSubLayersMinus1; i++) {	
	max_dec_pic_buffering_minus1[i]	ue(v)
	if(!dpbSizeOnlyFlag) {	
	max_num_reorder_pics[i]	ue(v)
	max_latency_increase_plus1[i]	ue(v)
	}	
	}	
	}	

[0163] HRD参数语法包括通用HRD参数语法、OLS HRD参数语法和子层HRD参数语法。

[0164] 通用HRD参数语法包括以下内容。

[0165]	general_hrd_parameters() {	描述符
	general_nal_hrd_params_present_flag	u(1)
	general_vcl_hrd_params_present_flag	u(1)
	if(general_nal_hrd_params_present_flag general_vcl_hrd_params_present_flag) {	
	decoding_unit_hrd_params_present_flag	u(1)
	if(decoding_unit_hrd_params_present_flag) {	
	tick_divisor_minus2	u(8)
	decoding_unit_cpb_params_in_pic_timing_sei_flag	u(1)
	}	
	bit_rate_scale	u(4)
	cpb_size_scale	u(4)
	if(decoding_unit_hrd_params_present_flag)	
	cpb_size_du_scale	u(4)
	}	
	if(vps_max_sub_layers_minus1 > 0)	
	sub_layer_cpb_params_present_flag	u(1)

[0166]	if(TotalNumOlss > 1)	
	num_ols_hrd_params_minus1	ue(v)
	hrd_cpb_cnt_minus1	ue(v)
	for(i = 0; i <= num_ols_hrd_params_minus1; i++) {	
	if(vps_max_sub_layers_minus1 > 0 && !vps_all_layers_same_num_sub_layers_flag)	
	hrd_max_temporal_id[i]	u(3)
	ols_hrd_parameters(hrd_max_temporal_id[i])	
	}	
	if(num_ols_hrd_params_minus1 > 0)	
	for(i = 1; i < TotalNumOlss; i++)	
	ols_hrd_idx[i]	ue(v)
	}	

[0167] OLS HRD参数语法包括以下内容。

[0168]	ols_hrd_parameters(hrdMaxTid) {	描述符
	firstSubLayer = sub_layer_cpb_params_present_flag ? 0: hrdMaxTid	
	for(i = firstSubLayer; i <= hrdMaxTid; i++) {	
	fixed_pic_rate_general_flag[i]	u(1)
	if(!fixed_pic_rate_general_flag[i])	
	fixed_pic_rate_within_cvs_flag[i]	u(1)
	if(fixed_pic_rate_within_cvs_flag[i])	
	elemental_duration_in_tc_minus1[i]	ue(v)
	else if(hrd_cpb_cnt_minus1 == 0)	
	low_delay_hrd_flag[i]	u(1)
	if(general_nal_hrd_params_present_flag)	
	sub_layer_hrd_parameters(i)	
	if(general_vcl_hrd_params_present_flag)	
	sub_layer_hrd_parameters(i)	
	}	
	}	

[0169] 子层HRD参数语法包括以下内容。

[0170]	sub_layer_hrd_parameters(subLayerId) {	描述符
	for(j = 0; j <= hrd_cpb_cnt_minus1; j++) {	
	bit_rate_value_minus1[subLayerId][j]	ue(v)
	cpb_size_value_minus1[subLayerId][j]	ue(v)
	if(decoding_unit_hrd_params_present_flag) {	
	cpb_size_du_value_minus1[subLayerId][j]	ue(v)
	bit_rate_du_value_minus1[subLayerId][j]	ue(v)
	}	
	cbr_flag[subLayerId][j]	u(1)
	}	
	}	
	}	

[0171] VPS语义如下。

[0172] vps_max_layers_minus1+1表示参考VPS的每个CVS中的各层的最大允许数量。

[0173] vps_max_sub_layers_minus1+1表示参考VPS的每个CVS中可以存在的时间子层的最大数量。nvps_max_sub_layers_minus1的取值范围应为0至6(包括端值)。

[0174] vps_all_layers_same_num_sub_layers_flag等于1表示参考VPS的每个CVS中的所有层具有一样多的时间子层。vps_all_layers_same_num_sub_layers_flag等于0表示参考VPS的每个CVS中的各层可以或可以不具有一样多的时间子层。当vps_all_layers_same_num_sub_layers_flag不存在时,vps_all_layers_same_num_sub_layers_flag的值被推断为1。

[0175] vps_all_independent_layers_flag等于1表示CVS中的所有层都是独立译码的,而不使用层间预测。vps_all_independent_layers_flag等于0表示CVS中的一层或多层可以使用层间预测。当vps_all_independent_layers_flag不存在时,vps_all_independent_layers_flag的值被推断为1。当vps_all_independent_layers_flag等于1时,vps_independent_layer_flag[i]的值被推断为1。当vps_all_independent_layers_flag等于0时,vps_independent_layer_flag[0]的值被推断为1。

[0176] vps_direct_dependency_flag[i][j]等于0表示索引为j的层不是索引为i的层的直接参考层。vps_direct_dependency_flag[i][j]等于1表示索引为j的层是索引为i的层的直接参考层。当vps_direct_dependency_flag[i][j] (i和j的范围为0至vps_max_layers_minus1,包括端值)不存在时,vps_direct_dependency_flag[i][j]被推断为0。

[0177] 变量DirectDependentLayerIdx[i][j]表示第i层的第j个直接依赖层,变量LayerUsedAsRefLayerFlag[j]表示层索引为j的层是否用作任何其它层的参考层,这两个变量推导如下:

```

for( i = 0; i <= vps_max_layers_minus1; i++ )
[0178] LayerUsedAsRefLayerFlag[ j ] = 0
for( i = 1; i < vps_max_layers_minus1; i++ )

```



```

        if( !vps_independent_layer_flag[ i ] )
            for( j = i - 1, k = 0; j >= 0; j-- )
                if( vps_direct_dependency_flag[ i ][ j ] ) {
[0179]         DirectDependentLayerIdx[ i ][ k++ ] = j
                LayerUsedAsRefLayerFlag[ j ] = 1
                }
    
```

[0180] 变量variable GeneralLayerIdx[i]表示nuh_layer_id等于vps_layer_id[i]的层的层索引,该变量推导如下:

```
[0181] for( i = 0; i <= vps_max_layers_minus1; i++ )
```

```
[0182] GeneralLayerIdx[ vps_layer_id[ i ] ] = i
```

[0183] each_layer_is_an_ols_flag等于1表示每个输出层集只包括一层且码流中的每层本身是一个输出层集,其中,输出层集中包括的单层是唯一的输出层。each_layer_is_an_ols_flag等于0表示输出层集可以包括多层。如果vps_max_layers_minus1等于0,则each_layer_is_an_ols_flag的值被推断为1。否则,当vps_all_independent_layers_flag等于0时,each_layer_is_an_ols_flag的值被推断为0。

[0184] ols_mode_idc等于0表示:由VPS指定的OLS的总数量等于vps_max_layers_minus1+1,第i个OLS包括层索引从0到i(包括端值)的各层,只输出每个OLS中的最高层。ols_mode_idc等于1表示:由VPS指定的OLS的总数量等于vps_max_layers_minus1+1,第i个OLS包括层索引从0到i(包括端值)的各层,输出每个OLS中的所有层。ols_mode_idc等于2表示:由VPS指定的OLS的总数量被显式指示,输出每个OLS中的最高层和显式指示的一组较低层。ols_mode_idc的取值范围应为0至2(包括端值)。保留ols_mode_idc的值3,供ITU-T或ISO/IEC将来使用。当vps_all_independent_layers_flag等于1且each_layer_is_an_ols_flag等于0时,ols_mode_idc的值被推断为2。

[0185] 当ols_mode_idc等于2时,num_output_layer_sets_minus1+1表示由VPS指定的OLS的总数量。

[0186] 变量TotalNumOlss表示由VPS指定的OLS的总数量,该变量推导如下:

```
[0187] if( vps_max_layers_minus1 == 0 )
```

```
[0188] TotalNumOlss = 1
```

```
[0189] else if( each_layer_is_an_ols_flag || ols_mode_idc == 0 || ols_mode_idc == 1 )
```

```
[0190] TotalNumOlss = vps_max_layers_minus1 + 1
```

```
[0191] else if( ols_mode_idc == 2 )
```

```
[0192] TotalNumOlss = num_output_layer_sets_minus1 + 1
```

[0193] layer_included_flag[i][j]表示:当ols_mode_idc等于2时,第j层(即nuh_layer_id等于vps_layer_id[j]的层)包括在第i个OLS中。layer_included_flag[i][j]等于1表示第j层包括在第i个OLS中。layer_included_flag[i][j]等于0表示第j层不包括在第i个OLS中。

[0194] 变量NumLayersInOls[i]表示第i个OLS中的层的数量,变量LayerIdInOls[i][j]表示第i个OLS中的第j层的nuh_layer_id值,这两个变量推导如下:

```

        NumLayersInOls[ 0 ] = 1
[0195]   LayerIdInOls[ 0 ][ 0 ] = vps_layer_id[ 0 ]
        for( i = 1, i < TotalNumOls; i++ ) {
            if( each_layer_is_an_ols_flag ) {
                NumLayersInOls[ i ] = 1
                LayerIdInOls[ i ][ 0 ] = vps_layer_id[ i ]
            } else if( ols_mode_idc == 0 || ols_mode_idc == 1 ) {
                NumLayersInOls[ i ] = i + 1
                for( j = 0; j < NumLayersInOls[ i ]; j++ )
                    LayerIdInOls[ i ][ j ] = vps_layer_id[ j ]
[0196]   } else if( ols_mode_idc == 2 ) {
                for( k = 0, j = 0; k <= vps_max_layers_minus1; k++ )
                    if( layer_included_flag[ i ][ k ] )
                        LayerIdInOls[ i ][ j++ ] = vps_layer_id[ k ]
                NumLayersInOls[ i ] = j
            }
        }
    }
}

```

[0197] 变量OlsLayIdx[i][j]表示nuh_layer_id等于LayerIdInOls[i][j]的层的OLS层索引,该变量推导如下:

```

[0198]   for( i = 0, i < TotalNumOls; i++ )
[0199]   for  j = 0; j < NumLayersInOls[ i ]; j++ )
[0200]   OlsLayIdx[ i ][ LayerIdInOls[ i ][ j ] ] = j

```

[0201] 每个OLS中的最低层应为独立层。换句话说,当每个i的范围为0至TotalNumOls-1(包括端值)时,vps_independent_layer_flag[GeneralLayerIdx[LayerIdInOls[i][0]]]的值应等于1。

[0202] 每层应包括在由VPS指定的至少一个OLS中。换句话说,对于特定nuh_layer_id值nuhLayerId等于vps_layer_id[k](k的范围为0至vps_max_layers_minus1,包括端值)之一的每层,应存在至少一对i和j的值,其中,i的范围为0至TotalNumOls-1(包括端值),j的范围为0至NumLayersInOls[i]-1(包括端值),使得LayerIdInOls[i][j]的值等于nuhLayerId。

[0203] OLS中的任一层应是OLS中的输出层或OLS中的输出层的(直接或间接)参考层。

[0204] vps_output_layer_flag[i][j]表示:当ols_mode_idc等于2时,是否输出第i个OLS中的第j层。vps_output_layer_flag[i]等于1表示输出第i个OLS中的第j层。vps_output_layer_flag[i]等于0表示不输出第i个OLS中的第j层。当vps_all_independent_layers_flag等于1且each_layer_is_an_ols_flag等于0时,vps_output_layer_flag[i]的值被推断为1。

[0205] 变量OutputLayerFlag[i][j]的值等于1表示输出第i个OLS中的第j层,等于0表示不输出第i个OLS中的第j层,该变量推导如下:

```

        for( i = 0, i < TotalNumOls; i++ ) {
            OutputLayerFlag[ i ][ NumLayersInOls[ i ] - 1 ] = 1
[0206]         for( j = 0; j < NumLayersInOls[ i ] - 1; j++ )
                if( ols_mode_idc[ i ] == 0 )
                    OutputLayerFlag[ i ][ j ] = 0
                else if( ols_mode_idc[ i ] == 1 )
                    OutputLayerFlag[ i ][ j ] = 1
                else if( ols_mode_idc[ i ] == 2 )
[0207]                 OutputLayerFlag[ i ][ j ] = vps_output_layer_flag[ i ][ j ]
        }

```

[0208] 注：第0个OLS只包括最低层（例如nuh_layer_id等于vps_layer_id[0]的层），输出第0个OLS中包括的唯一层。

[0209] vps_num_ptls表示VPS中的语法结构profile_tier_level()的数量。

[0210] pt_present_flag[i]等于1表示VPS中的第i个语法结构profile_tier_level()中存在档次、层次和通用约束条件信息。pt_present_flag[i]等于0表示VPS中的第i个语法结构profile_tier_level()中不存在档次、层次和通用约束条件信息。pt_present_flag[0]的值被推断为0。当pt_present_flag[i]等于0时，VPS中的第i个语法结构profile_tier_level()中的档次、层次和通用约束条件信息被推断为与VPS中的第(i-1)个语法结构profile_tier_level()中的档次、层次和通用约束条件信息相同。

[0211] ptl_max_temporal_id[i]表示最高子层表示的TemporalId，VPS中的第i个语法结构profile_tier_level()中存在该最高子层表示的级别信息。ptl_max_temporal_id[i]的取值范围应为0至vps_max_sub_layers_minus1（包括端值）。当vps_max_sub_layers_minus1等于0时，ptl_max_temporal_id[i]的值被推断为0。当vps_max_sub_layers_minus1大于0且vps_all_layers_same_num_sub_layers_flag等于1时，ptl_max_temporal_id[i]的值被推断为vps_max_sub_layers_minus1。

[0212] vps_ptl_byte_alignment_zero_bit应等于0。

[0213] ols_ptl_idx[i]表示适用于第i个OLS的语法结构profile_tier_level()在VPS中的语法结构profile_tier_level()列表中的索引。当ols_ptl_idx[i]存在时，ols_ptl_idx[i]的取值范围应为0至vps_num_ptls-1（包括端值）。

[0214] 当NumLayersInOls[i]等于1时，适用于第i个OLS的语法结构profile_tier_level()存在于由第i个OLS中的层参考的SPS中。

[0215] vps_num_dpb_params表示VPS中的语法结构dpb_parameters()的数量。vps_num_dpb_params的取值范围应为0至16（包括端值）。当vps_num_dpb_params不存在时，vps_num_dpb_params的值被推断为0。

[0216] same_dpb_size_output_or_nonoutput_flag等于1表示VPS中不存在语法元素layer_nonoutput_dpb_params_idx[i]。same_dpb_size_output_or_nonoutput_flag等于0表示VPS中可以或可以不存在语法元素layer_nonoutput_dpb_params_idx[i]。

[0217] vps_sub_layer_dpb_params_present_flag用于控制VPS中的语法结构dpb_parameters()中是否存在语法元素max_dec_pic_buffering_minus1[]、max_num_

reorder_pics[]和max_latency_increase_plus1[]。当vps_sub_dpb_params_info_present_flag不存在时,vps_sub_dpb_params_info_present_flag被推断为0。

[0218] dpb_size_only_flag[i]等于1表示VPS中的第i个语法结构dpb_parameters()中不存在语法元素max_num_reorder_pics[]和max_latency_increase_plus1[]。dpb_size_only_flag[i]等于0表示VPS中的第i个语法结构dpb_parameters()中可以存在语法元素max_num_reorder_pics[]和max_latency_increase_plus1[]。

[0219] dpb_max_temporal_id[i]表示最高子层表示的TemporalId,VPS中的第i个语法结构dpb_parameters()中可以存在该最高子层表示的DPB参数。dpb_max_temporal_id[i]的取值范围应为0至vps_max_sub_layers_minus1(包括端值)。当vps_max_sub_layers_minus1等于0时,dpb_max_temporal_id[i]的值被推断为0。当vps_max_sub_layers_minus1大于0且vps_all_layers_same_num_sub_layers_flag等于1时,dpb_max_temporal_id[i]的值被推断为vps_max_sub_layers_minus1。

[0220] layer_output_dpb_params_idx[i]表示适用于第i层(当第i层是OLS中的输出层时)的语法结构dpb_parameters()在VPS中的语法结构dpb_parameters()列表中的索引。当layer_output_dpb_params_idx[i]存在时,layer_output_dpb_params_idx[i]的取值范围应为0至vps_num_dpb_params-1(包括端值)。

[0221] 如果vps_independent_layer_flag[i]等于1,则适用于第i层(当第i层是输出层时)的语法结构dpb_parameters()是存在于由该层参考的SPS中的语法结构dpb_parameters()。

[0222] 否则(vps_independent_layer_flag[i]等于0),以下内容适用:

[0223] 当vps_num_dpb_params等于1时,layer_output_dpb_params_idx[i]的值被推断为0。

[0224] 码流一致性的要求可以是:layer_output_dpb_params_idx[i]的值应使得dpb_size_only_flag[layer_output_dpb_params_idx[i]]等于0。

[0225] layer_nonoutput_dpb_params_idx[i]表示适用于第i层(当第i层是OLS中的非输出层时)的语法结构dpb_parameters()在VPS中的语法结构dpb_parameters()列表中的索引。当layer_nonoutput_dpb_params_idx[i]存在时,layer_nonoutput_dpb_params_idx[i]的取值范围应为0至vps_num_dpb_params-1(包括端值)。

[0226] 如果same_dpb_size_output_or_nonoutput_flag等于1,则以下内容适用:

[0227] 如果vps_independent_layer_flag[i]等于1,则适用于第i层(当第i层是非输出层时)的语法结构dpb_parameters()是存在于由该层参考的SPS中的语法结构dpb_parameters()。

[0228] 否则(vps_independent_layer_flag[i]等于0),layer_nonoutput_dpb_params_idx[i]的值被推断为layer_output_dpb_params_idx[i]。

[0229] 否则(same_dpb_size_output_or_nonoutput_flag等于0),当vps_num_dpb_params等于1时,layer_output_dpb_params_idx[i]的值被推断为0。

[0230] vps_extension_flag等于0表示VPS RBSP语法结构中不存在语法元素vps_extension_data_flag。vps_extension_flag等于1表示VPS RBSP语法结构中不存在语法元素vps_extension_data_flag。

[0231] vps_extension_data_flag可以取任意值。它的存在和值不影响解码器与本规范这一版本中指定的档次(profile)的一致性。与本规范这一版本一致的解码器应忽略所有语法元素vps_extension_data_flag。

[0232] SPS RBSP语义如下。

[0233] SPS RBSP在被参考之前应可用于解码过程,应包括在TemporalId等于0的至少一个接入单元中,或者应通过外部模块提供,而包括SPS RBSP的SPS NAL单元的nuh_layer_id应等于参考SPS NAL单元的PPS NAL单元的最小nuh_layer_id值。

[0234] CVS中具有特定sps_seq_parameter_set_id值的所有SPS NAL单元应具有相同的内容。

[0235] sps_decoding_parameter_set_id在大于0时表示由SPS参考的DPS的dps_decoding_parameter_set_id值。当sps_decoding_parameter_set_id等于0时,SPS不参考DPS且在对参考SPS的每个CLVS进行解码时不参考DPS。在由码流中的经编码图像参考的所有SPS中,sps_decoding_parameter_set_id的值应相同。

[0236] sps_video_parameter_set_id在大于0时表示SPS参考的VPS的vps_video_parameter_set_id值。当sps_video_parameter_set_id等于0时,SPS不参考VPS且在对参考SPS的每个CLVS进行解码时不参考VPS;GeneralLayerIdx[nuh_layer_id]的值被推断为0;vps_independent_layer_flag[GeneralLayerIdx[nuh_layer_id]]的值被推断为1。

[0237] 当vps_independent_layer_flag[GeneralLayerIdx[nuh_layer_id]]等于1时,由具有特定nuh_layer_id值(nuhLayerId)的CLVS参考的SPS的nuh_layer_id应等于nuhLayerId。

[0238] sps_max_sub_layers_minus1+1表示参考SPS的每个CLVS中可以存在的时间子层的最大数量。sps_max_sub_layers_minus1的取值范围应为0至vps_max_sub_layers_minus1(包括端值)。

[0239] 在与本规范这一版本一致的码流中,sps_reserved_zero_4bits应等于0。保留sps_reserved_zero_4bits的其它值,供ITU-T或ISO/IEC将来使用。

[0240] sps_ptl_dpb_present_flag等于1表示SPS中存在语法结构profile_tier_level()和语法结构dpb_parameters()。sps_ptl_dpb_present_flag等于0表示SPS中不存在语法结构profile_tier_level()和语法结构dpb_parameters()。sps_ptl_dpb_present_flag的值应等于vps_independent_layer_flag[nuh_layer_id]。

[0241] 如果vps_independent_layer_flag[GeneralLayerIdx[nuh_layer_id]]等于1,则变量MaxDecPicBuffMinus1被设置为SPS中的语法结构dpb_parameters()中的max_dec_pic_buffering_minus1[sps_max_sub_layers_minus1]。否则,MaxDecPicBuffMinus1被设置为VPS中的第layer_nonoutput_dpb_params_idx[GeneralLayerIdx[nuh_layer_id]]个语法结构dpb_parameters()中的max_dec_pic_buffering_minus1[sps_max_sub_layers_minus1]。

[0242] gdr_enabled_flag等于1表示参考SPS的CLVS中可以存在逐步解码刷新(gradual decoding refresh,GDR)图像。gdr_enabled_flag等于0表示参考SPS的CLVS中不存在GDR图像。

[0243] sps_sub_layer_dpb_params_flag用于控制SPS中的语法结构dpb_parameters()

中是否存在语法元素`max_dec_pic_buffering_minus1[i]`、`max_num_reorder_pics[i]`和`max_latency_increase_plus1[i]`。当`sps_sub_dpb_params_info_present_flag`不存在时，`sps_sub_dpb_params_info_present_flag`被推断为0。

[0244] `long_term_ref_pics_flag`等于0表示LTRP不用于CLVS中的任何经编码图像的帧间预测。`long_term_ref_pics_flag`等于1表示LTRP可以用于CLVS中的一个或多个经编码图像的帧间预测。

[0245] 通用档次、层次和级别语义如下。

[0246] 语法结构`profile_tier_level()`提供级别信息，可选地提供档次、层次、子档次和通用约束条件信息(表示为PT信息)。

[0247] 当语法结构`profile_tier_level()`包括在DPS中时，`0lsInScope`是包括参考DPS的整个码流中的所有层的OLS。当语法结构`profile_tier_level()`包括在VPS中时，`0lsInScope`是由VPS指定的一个或多个OLS。当语法结构`profile_tier_level()`包括在SPS中时，`0lsInScope`是只包括参考SPS的所有层中的最低层的OLS，其中，该最低层应为独立层。

[0248] `general_profile_idc`表示VVC标准附录A中指定的与`0lsInScope`一致的档次。码流包括的`general_profile_idc`值应为附录A中指定的值保留`general_profile_idc`的其它值，供ITU-T或ISO/IEC将来使用。

[0249] `general_tier_flag`表示用于解释附录A中指定的`general_level_idc`。

[0250] `num_sub_profiles`表示语法元素`general_sub_profile_idc[i]`的数量。

[0251] `general_sub_profile_idc[i]`表示根据ITU-T T.35建议书所述注册的第*i*个互操作元数据，其内容未在VVC标准中详述。

[0252] `general_level_idc`表示附录A中指定的与`0lsInScope`一致的级别。码流所包含的`general_level_idc`值应为附录A中指定的值。保留`general_level_idc`的其它值，供ITU-T或ISO/IEC将来使用。

[0253] 注1：`general_level_idc`的值越大，级别越高。在DPS中为`0lsInScope`指示的最大级别可能高于在SPS中为`0lsInScope`内包括的CVS指示的级别。

[0254] 注2：当`0lsInScope`与多个档次一致时，`general_profile_idc`应该表示提供优选解码结果或优选码流标识的档次，由编码器确定(以本规范未详述的方式)。

[0255] 注3：当语法结构`profile_tier_level()`包括在DPS中且`0lsInScope`的CVS与不同的档次一致时，`general_profile_idc`和`level_idc`应该表示能够对`0lsInScope`进行解码的解码器的档次和级别。

[0256] `sub_layer_level_present_flag[i]`等于1表示`TemporalId`等于*i*的子层表示的语法结构`profile_tier_level()`中存在级别信息。`sub_layer_level_present_flag[i]`等于0表示`TemporalId`等于*i*的子层表示的语法结构`profile_tier_level()`中不存在级别信息。

[0257] `ptl_alignment_zero_bits`应等于0。

[0258] 除了规定不存在值的推断之外，语法元素`sub_layer_level_idc[i]`的语义也与语法元素`general_level_idc`相同，但适用于`TemporalId`等于*i*的子层表示。

[0259] DPB语义如下。

[0260] 语法结构dpb_parameters(maxSubLayersMinus1,subLayerInfoFlag)为CVS中的每个CLVS提供DPB大小、最大图像重排序次数和最大延迟的信息。

[0261] 当VPS中包括语法结构dpb_parameters()时,语法结构dpb_parameters()适用于的OLS由VPS指定。当语法结构dpb_parameters()包括在SPS中时,该语法结构适用于只包括参考SPS的所有层中的最低层的OLS,其中,该最低层是独立层。

[0262] max_dec_pic_buffering_minus1[i]+1表示CVS中的每个CLVS在Htid等于i时的解码图像缓冲区的以图像存储缓冲区为单位的最大所需大小。max_dec_pic_buffering_minus1[i]的取值范围应为0至MaxDpbSize-1(包括端值),其中,MaxDpbSize在其它地方详述。当i大于0时,max_dec_pic_buffering_minus1[i]应大于或等于max_dec_pic_buffering_minus1[i-1]。当max_dec_pic_buffering_minus1[i](其中,i的范围为0至maxSubLayersMinus1-1,包括端值)不存在时,由于subLayerInfoFlag等于0,因此max_dec_pic_buffering_minus1[i]被推断为max_dec_pic_buffering_minus1[maxSubLayersMinus1]。

[0263] max_num_reorder_pics[i]表示CVS中的每个CLVS在Htid等于i时的其中图像的最大允许数量,这些图像可以在解码顺序中位于CLVS中的任一图像之前且在输出顺序中位于该图像之后。max_num_reorder_pics[i]的取值范围应为0至max_dec_pic_buffering_minus1[i](包括端值)。当i大于0时,max_num_reorder_pics[i]应大于或等于max_num_reorder_pics[i-1]。当max_num_reorder_pics[i](其中,i的范围为0至maxSubLayersMinus1-1,包括端值)不存在时,由于subLayerInfoFlag等于0,因此max_num_reorder_pics[i]被推断为max_num_reorder_pics[maxSubLayersMinus1]。

[0264] max_latency_increase_plus1[i]不等于0用于计算MaxLatencyPictures[i]的值,MaxLatencyPictures[i]表示CVS中的每个CLVS在Htid等于i时的其中图像的最大数量,这些图像可以在输出顺序中位于CLVS中的任一图像之前且在解码顺序中位于该图像之后。

[0265] 当max_latency_increase_plus1[i]不等于0时,MaxLatencyPictures[i]的值规定如下:

[0266] MaxLatencyPictures[i]=max_num_reorder_pics[i]+max_latency_increase_plus1[i]-1。

[0267] 当max_latency_increase_plus1[i]等于0时,不存在对应的限制。

[0268] max_latency_increase_plus1[i]的取值范围必须为0至232-2(包括端值)。当max_latency_increase_plus1[i](其中,i的范围为0至maxSubLayersMinus1-1,包括端值)不存在时,由于subLayerInfoFlag等于0,因此max_latency_increase_plus1[i]被推断为max_latency_increase_plus1[maxSubLayersMinus1]。

[0269] HRD参数语义包括通用HRD参数语义。

[0270] 通用HRD参数语义如下。

[0271] 语法结构general_hrd_parameters()提供HRD操作中使用的HRD参数。

[0272] num_ols_hrd_params_minus1+1表示语法结构general_hrd_parameters()中存在的语法结构ols_hrd_parameters()的数量。num_ols_hrd_params_minus1的取值范围应为0至63(包括端值)。当TotalNumOls大于1时,num_ols_hrd_params_minus1的值被推断为0。

[0273] hrd_cpb_cnt_minus1+1表示CVS的码流中的备选CPB规格的数量。hrd_cpb_cnt_

minus1的取值范围应为0至31(包括端值)。

[0274] hrd_max_temporal_id[i]表示最高子层表示的TemporalId,该最高子层表示的HRD参数包括在第i个语法结构layer_level_hrd_parameters()中。hrd_max_temporal_id[i]的取值范围应为0至vps_max_sub_layers_minus1(包括端值)。当vps_max_sub_layers_minus1等于0时,hrd_max_temporal_id[i]的值被推断为0。

[0275] ols_hrd_idx[i]表示适用于第i个OLS的语法结构ols_hrd_parameters()的索引。ols_hrd_idx[[i]]的取值范围应为0至num_ols_hrd_params_minus1(包括端值)。当ols_hrd_idx[[i]]不存在时,ols_hrd_idx[[i]]的值被推断为0。

[0276] 参考图像列表结构语义如下。

[0277] 语法结构ref_pic_list_struct(listIdx,rplsIdx)可以存在于SPS中,也可以存在于条带头中。根据语法结构是包括在条带头中还是包括在SPS中,以下内容适用:

[0278] -如果语法结构ref_pic_list_struct(listIdx,rplsIdx)存在于条带头中,则该语法结构表示当前图像(包括条带的图像)的参考图像列表listIdx。

[0279] -否则(语法结构ref_pic_list_struct(listIdx,rplsIdx)存在于SPS中),该语法结构表示参考图像列表listIdx的候选。本节其余内容规定的语义中的术语“当前图像”是指:(1)包括一个或多个条带的每个图像,其中,一个或多个条带包括的ref_pic_list_idx[listIdx]等于SPS中包括的语法结构ref_pic_list_struct(listIdx,rplsIdx)列表中的索引;(2)参考SPS的CVS中的每个图像。

[0280] num_ref_entries[listIdx][rplsIdx]表示语法结构ref_pic_list_struct(listIdx,rplsIdx)中的条目数量。num_ref_entries[listIdx][rplsIdx]的取值范围应为0至MaxDecPicBuffMinus1+14(包括端值)。

[0281] DCI 802、VPS 804、SPS 806和PPS 808包括在不同类型的网络抽象层(Network Abstraction Layer,NAL)单元中。NAL单元是一种包括要遵循的数据(例如经编码视频数据)类型的指示的语法结构。NAL单元分为视频编码层(video coding layer,VCL)NAL单元和非VCL NAL单元。VCL NAL单元包括表示视频图像中样本值的数据,而非VCL NAL单元包括任何关联的附加信息,例如,参数集(可以适用于多个VCL NAL单元的重要数据)和补充增强信息(定时信息和其它补充数据,其中,补充数据可以增强经解码视频信号的可用性,但不是解码视频图像中样本值所必需的)。

[0282] 在一个实施例中,DCI 802包括在指定为DCI NAL单元或DPS NAL单元的非VCL NAL单元中。也就是说,DCI NAL单元具有DCI NAL单元类型(NAL unit type,NUT),而DPS NAL单元具有DPS NUT。在一个实施例中,VPS 804包括在指定为VPS NAL单元的非VCL NAL单元中。因此,VPS NAL单元具有VPS NUT。在一个实施例中,SPS 806是指定为SPS NAL单元的非VCL NAL单元。因此,SPS NAL单元具有SPS NUT。在一个实施例中,PPS808包括在指定为PPS NAL单元的非VCL NAL单元中。因此,PPS NAL单元具有PPS NUT。

[0283] PH 812是一种包括应用于经编码图像(例如图像814)中的所有条带(例如条带818)的语法元素的语法结构。在一个实施例中,PH 812包括在指定为PH NAL单元的新类型非VCL NAL单元中。因此,PH NAL单元具有PH NUT(例如PH_NUT)。在一个实施例中,有且只有一个PH 812包括在每个PU 801中。也就是说,PU 801包括单个或唯一PH 812。在一个实施例中,码流800中的每个图像814正好包括一个PH NAL单元。

[0284] 图像814是黑白格式的亮度样本阵列或者是4:2:0、4:2:2和4:4:4颜色格式的亮度样本阵列和2个对应的色度样本阵列。在一个实施例中,有且只有一个图像814包括在每个PU 801中。因此,每个PU 801中只有一个PH 812且只有一个对应于该PH 812的图像814。也就是说,PU 801包括单个或唯一图像814。

[0285] 图像814可以是帧或场。然而,在一个CVS 816中,所有图像814都是帧,或者所有图像814都是场。CVS 816是视频码流800中的每个编码层视频序列(coded layer video sequence,CLVS)的编码视频序列。需要说明的是,当视频码流800包括单层时,CVS 816和CLVS相同。只有当视频码流800包括多层(例如,如图5和图6所示)时,CVS 816和CLVS才不同。

[0286] 每个图像814包括一个或多个条带818。条带818是图像(例如图像814)中的整数个完整的分块,或者是图像中的一个分块内的整数个连续完整的CTU行。每个条带818只包括在单个NAL单元(例如VCL NAL单元)中。分块(未示出)是图像(例如图像814)中的某一分块列和某一分块行内的CTU组成的矩形区域。CTU(未示出)是具有3个样本阵列的图像中的亮度样本组成的1个CTB和色度样本组成的2个对应CTB,或者是黑白图像中或使用3个单独颜色平面和语法结构进行译码的图像中的样本组成的1个CTB,其中,这些语法结构用于对上述样本进行译码。CTB(未示出)是一个 $N \times N$ 的样本块,其中, N 可以设为某个值,使得一个分量分成多个CTB,这就是一种分割方式。块(未示出)是一个 $M \times N$ (M 列 $\times N$ 行)的样本(例如像素)阵列,或者是一个 $M \times N$ 的变换系数阵列。

[0287] 在一个实施例中,每个条带818包括条带头820。条带头820是经编码条带818的一部分,包括与所有分块相关或者与在条带818中表示的一个分块内的CTU行相关的数据元素。也就是说,条带头820包括条带818的信息,例如条带类型、要使用的参考图像等。

[0288] 图像814及其条带818包括与正被编码或解码的图像或视频关联的数据。因此,图像814及其条带818可以简单地称为码流800中包括的载荷或数据。

[0289] 本领域技术人员将理解,码流800在实际应用中可以包括其它参数和信息。

[0290] 图9为由视频解码器(例如解码器400)实现的解码方法900的一个实施例。方法900可以在从视频编码器(例如视频编码器300)直接或间接接收到码流之后执行。通过确保不会在多层码流中为各层重复指示相同的档次、层次和级别(profile,tier,and level,PTL)参数、假设参考解码(hypothetical reference decoder,HRD)参数和解码图像缓冲区(decoded picture buffer,DPB)参数,方法900提高了解码过程。仅指示相同的PTL参数、HRD参数和DPB参数一次减少了冗余并提高了多层视频码流的译码效率。因此,视频译码中的编码器/解码器(又称“编解码器”)相对于现有的编解码器有所改进。实际上,改进的视频译码过程在发送、接收和/或观看视频时提供了更好的用户体验。

[0291] 在步骤902中,所述视频解码器接收视频码流。所述视频码流包括多层(例如,层531、532)和VPS(例如VPS 804),所述多层包括子层(例如,子层710、720、730)。所述VPS包括具有第一值(例如1)的第一标志(例如标志850)。具有所述第一值的所述第一标志表示级别信息、解码图像缓冲区(decoded picture buffer,DPB)参数和假设参考解码(hypothetical reference decoder,HRD)参数对应的最高子层表示(例如子层730)的时间标识符(TemporalID)不存在于所述VPS中并被推断为所述子层的最大数量,所述子层可以存在于由所述VPS指定的所述多层中的一层中。

[0292] 在一个实施例中,所述子层的最大数量的范围为1至7。在一个实施例中,所述级别信息包括在档次、层次和级别(profile,tier,and level,PTL)参数中。在一个实施例中,所述级别信息存在于所述VPS中的第i个语法结构profile_tier_level()中。在一个实施例中,所述DPB参数存在于所述VPS中的第i个语法结构dpb_parameters()中。在一个实施例中,所述HRD参数存在于所述VPS或序列参数集(sequence parameter set,SPS)中的第i个语法结构ols_timing_hrd_parameters()中。

[0293] 在步骤904中,从所述VPS中获取所述最高子层表示的时间ID对应的所述级别信息、所述HRD参数和所述DPB参数。

[0294] 在步骤906,中,所述视频解码器对所述多层中的一层中的图像进行解码,以得到经解码图像。在一个实施例中,所述图像不使用所述HRD参数进行解码。也就是说,所述HRD参数用于测试所述码流的一致性而不用于对其中包括的图像进行解码。

[0295] 在所述图像完成解码之后,所述图像可以用于生成或产生一个图像或视频序列,以在电子设备(例如,智能手机、平板电脑、笔记本电脑、个人计算机等)的显示器或屏幕上显示给用户。

[0296] 图10为由视频编码器(例如视频编码器300)实现的对视频码流进行编码的方法1000的一个实施例。方法1000可以在(例如视频中的)图像要编码到视频码流中并接着发送给视频解码器(例如视频解码器400)时执行。通过确保不会在多层码流中为各层重复指示相同的档次、层次和级别(profile,tier,and level,PTL)参数、假设参考解码(hypothetical reference decoder,HRD)参数和解码图像缓冲区(decoded picture buffer,DPB)参数,方法1000提高了编码过程。仅指示相同的PTL参数、HRD参数和DPB参数一次减少了冗余并提高了多层视频码流的译码效率。因此,视频译码中的编码器/解码器(又称“编解码器”)相对于现有的编解码器有所改进。实际上,改进的视频译码过程在发送、接收和/或观看视频时提供了更好的用户体验。

[0297] 在步骤1002中,所述视频编码器生成包括子层(例如,子层710、720、730)的多层(例如,层531、532)和包括第一标志(例如标志850)的VPS(VPS 804)。

[0298] 在步骤1004中,所述视频编码器将所述第一标志设置为第一值以表示级别信息、解码图像缓冲区(decoded picture buffer,DPB)参数和假设参考解码(hypothetical reference decoder,HRD)参数对应的最高子层表示(例如子层730)的时间标识符(Temporal ID)不存在于所述VPS中并应该被推断为所述子层的最大数量,所述子层可以存在于由所述VPS指定的所述多层中的一层。

[0299] 在一个实施例中,所述子层的最大数量的范围为1至7。在一个实施例中,所述级别信息包括在档次、层次和级别(profile,tier,and level,PTL)参数中。在一个实施例中,所述级别信息存在于所述VPS中的第i个语法结构profile_tier_level()中。在一个实施例中,所述DPB参数存在于所述VPS中的第i个语法结构dpb_parameters()中。在一个实施例中,所述HRD参数存在于所述VPS或序列参数集(sequence parameter set,SPS)中的第i个语法结构ols_timing_hrd_parameters()中。

[0300] 在一个实施例中,方法1000还包括:使用所述HRD参数确定所述编码的视频码流是否一致。也就是说,确定所述编码的码流是否与标准(例如VVC标准)一致。

[0301] 在步骤1006中,所述视频编码器将所述多层和所述VPS编码到视频码流中。在步骤

1008中,所述视频编码器存储所述视频码流,以将所述视频码流发送给视频解码器。所述视频码流可以存储在存储器中,直到所述视频码流被发送给所述视频解码器。在所述视频解码器接收到所述经编码视频码流之后,所述经编码视频码流可以解码(例如,如上所述)生成或产生一个图像或视频序列,以在电子设备(例如,智能手机、平板电脑、笔记本电脑、个人计算机等)的显示器或屏幕上显示给用户。

[0302] 下文提供了其它和/或补充细节。

[0303] 通用解码过程如下。

[0304] 该过程的输入是码流BitstreamToDecode。该过程的输出是经解码图像列表。

[0305] 解码过程的规定使得与指定档次和级别一致的所有解码器在为与该档次和级别一致的码流调用与该档次关联的解码过程时,产生数字上相同的裁剪后的经解码输出图像。与本文描述的过程产生相同的裁剪后的经解码输出图像(根据规定具有正确的输出顺序或输出时间)的任何解码过程都符合VVC标准的解码过程要求。

[0306] 对于码流中的每个IRAP AU,以下内容适用:

[0307] -如果AU是码流中在解码顺序中的第一个AU,每个图像都是即时解码刷新(instantaneous decoding refresh, IDR)图像,或者每个图像都是层中在解码顺序中位于序列结束NAL单元之后的第一个图像,则变量NoIncorrectPicOutputFlag被设置为1。

[0308] -否则,如果本规范中未详述的某一外部模块可用于将变量HandleCraAsCvsStartFlag设置为AU的一个值,则HandleCraAsCvsStartFlag被设置为通过外部模块提供的值,而且NoIncorrectPicOutputFlag被设置为HandleCraAsCvsStartFlag。

[0309] -否则,HandleCraAsCvsStartFlag和NoIncorrectPicOutputFlag都被设置为0。

[0310] 对于码流中的每个GDR AU,以下内容适用:

[0311] -如果AU是码流中在解码顺序中的第一个AU或者每个图像都是层中在解码顺序中位于序列结束NAL单元之后的第一个图像,则变量NoIncorrectPicOutputFlag被设置为1。

[0312] -否则,如果本规范中未详述的某一外部模块可用于将变量HandleGdrAsCvsStartFlag设置为AU的一个值,则HandleGdrAsCvsStartFlag被设置为通过该外部模块提供的值,而且NoIncorrectPicOutputFlag被设置为HandleGdrAsCvsStartFlag。

[0313] -否则,HandleGdrAsCvsStartFlag和NoIncorrectPicOutputFlag都被设置为0。

[0314] 注:需要上述针对IRAP图像和GDR图像的操作来标识码流中的CVS。

[0315] 为BitstreamToDecode中在解码顺序中的每个经编码图像重复调用第8.1.2节。

[0316] 用于参考图像列表构建的解码过程如下。

[0317] 该过程在开始非IDR图像中的每个条带的解码过程时调用。

[0318] 参考图像通过参考索引进行寻址。参考索引是参考图像列表中的索引。在对I条带进行解码时,条带数据的解码不使用参考图像列表。在对P条带进行解码时,条带数据的解码只使用参考图像列表0(即RefPicList[0])。在对B条带进行解码时,条带数据的解码既使用参考图像列表0,又使用参考图像列表1(即RefPicList[1])

[0319] 码流一致性的要求是以下约束条件适用:

[0320] -当每个i等于0或1时,num_ref_entries[i][RplIdx[i]]应不小于NumRefIdxActive[i]。

[0321] -由RefPicList[0]或RefPicList[1]中的每个激活条目参考的图像应存在于DPB中,该图像的TemporalId应小于或等于当前图像的TemporalId。

[0322] -由RefPicList[0]或RefPicList[1]中的每个条目参考的图像应不是当前图像,该图像的non_reference_picture_flag应等于0。

[0323] -图像中的条带的RefPicList[0]或RefPicList[1]中的短期参考图像(short term reference picture,STRP)条目和同一图像中的同一条带或不同条带的RefPicList[0]或RefPicList[1]中的LTRP条目应不参考同一图像。

[0324] -RefPicList[0]或RefPicList[1]中不存在LTRP条目,当前图像的PicOrderCntVal与由该条目参考的图像的PicOrderCntVal之间的差值大于或等于224。

[0325] -假设setOfRefPics是由RefPicList[0]中的与当前图像具有相同nuh_layer_id的所有条目和RefPicList[1]中的与当前图像具有相同nuh_layer_id的所有条目参考的唯一图像集。setOfRefPics中的图像数量应小于或等于MaxDecPicBuffMinus1,而且一个图像中的所有条带的setOfRefPics应是相同的。

[0326] -在当前图像是步进式时间子层接入(step-wise temporal sublayer access, STSA)图像时,RefPicList[0]或RefPicList[1]中应不存在与当前图像具有相同TemporalId的激活条目。

[0327] 在当前图像是解码顺序在与当前图像具有相同TemporalId的STSA图像之后的图像时,应不存在与当前图像具有相同TemporalId的图像,当前图像作为RefPicList[0]或RefPicList[1]中的解码顺序在STSA图像之前的激活条目。

[0328] -由当前图像中的条带的RefPicList[0]或RefPicList[1]中的每个ILRP条目参考的图像应与当前图像在同一接入单元中。

[0329] -由当前图像中的条带的RefPicList[0]或RefPicList[1]中的每个ILRP条目参考的图像应存在于DPB中,而且该图像的nuh_layer_id应小于当前图像的nuh_layer_id。

[0330] -条带的RefPicList[0]或RefPicList[1]中的每个ILRP条目都应是激活条目。

[0331] VVC标准的第C.1节中的HRD通用方面包括以下内容。

[0332] 本附录详述了假设参考解码(hypothetical reference decoder,HRD)及其用于检查码流和解码器一致性的用途。

[0333] 需要一组码流一致性测试来检查码流的一致性,该码流称为整个码流,表示为entireBitstream。这组码流一致性测试用于测试由VPS指定的每个OLS的每个OP的一致性。

[0334] 对于每个测试,以下按顺序执行的步骤按照列出的顺序适用,然后是本节中这些步骤之后描述的过程:

[0335] 1.测试操作点(表示为targetOp)是通过选择具有OLS索引op0lsIdx和最大TemporalId值opTid的目标OLS进行选择的。op0lsIdx的取值范围为0至TotalNum0lss-1(包括端值)。opTid的取值范围为0至vps_max_sub_layers_minus1(包括端值)。选择的每对op0lsIdx和opTid值应使得,通过调用第C.6节详述的以entireBitstream、op0lsIdx和opTid为输入的子码流提取过程而输出的子码流满足以下条件:

[0336] -BitstreamToDecode中存在nuh_layer_id等于LayerIdIn0ls[op0lsIdx]中的每个nuh_layer_id值的至少一个VCL NAL单元。

[0337] -BitstreamToDecode中存在TemporalId等于opTid的至少一个VCL NAL单元。

[0338] 2.如果targetOp中的层包括entireBitstream中的所有层且opTid等于或大于entireBitstream中的所有NAL单元之中的最大TemporalId值,则BitstreamToDecode被设置为与entireBitstream相同。否则,BitstreamToDecode被设置为调用以entireBitstream、op0lsIdx和opTid为输入的子码流提取过程(如第C.6节详述)的输出。

[0339] 3.Target0lsIdx和Htid的值分别被设置为targetOp的op0lsIdx和opTid。

[0340] 4.选择ScIdx的值。选择的ScIdx的范围应为0至hrd_cpb_cnt_minus1(包括端值)。

[0341] 5.BitstreamToDecode中与适用于Target0lsIdx的缓冲周期SEI消息(存在于TargetLayerBitstream中或通过本规范中未详述的外部机制获取)关联的接入单元被选择作为HRD初始化点,并称为目标OLS中的每层的接入单元0。

[0342] 6.适用于BitstreamToDecode的语法结构ols_hrd_parameters()和语法结构sub_layer_hrd_parameters()选择如下:

[0343] -选择VPS中(或通过本规范中未详述的外部模块提供)的第ols_hrd_idx[Target0lsIdx]个语法结构ols_hrd_parameters()。

[0344] -在选择语法结构ols_hrd_parameters()中,如果BitstreamToDecode是I类码流,则选择条件“if(general_vcl_hrd_params_present_flag)”紧跟着的语法结构sub_layer_hrd_parameters(Htid),而且变量NalHrdModeFlag被设置为0;否则(BitstreamToDecode是II类码流),选择条件“if(general_vcl_hrd_params_present_flag)”(在这种情况下,变量NalHrdModeFlag被设置为0)或条件“if(general_nal_hrd_params_present_flag)”(在这种情况下,变量NalHrdModeFlag被设置为1)紧跟着的语法结构sub_layer_hrd_parameters(Htid)。当BitstreamToDecode是II类码流且NalHrdModeFlag等于0时,除填充数据NAL单元外的所有非VCL NAL单元以及根据NAL单元流形成字节流(如附录B详述)的所有语法元素leading_zero_8bits、zero_byte、start_code_prefix_one_3bytes和trailing_zero_8bits(如果存在)从BitstreamToDecode中移除,而剩余的码流被分配给BitstreamToDecode。

[0345] 7.当decoding_unit_hrd_params_present_flag等于1时,调度CPB在接入单元级(在这种情况下,变量DecodingUnitHrdFlag被设置为0)或解码单元级(在这种情况下,变量DecodingUnitHrdFlag被设置为1)进行操作。否则,DecodingUnitHrdFlag被设置为0,而且调度CPB在接入单元级进行操作。

[0346] 8.对于从接入单元0开始的BitstreamToDecode中的每个接入单元,选择与该接入单元关联并适用于Target0lsIdx的缓冲周期SEI消息(存在于BitstreamToDecode中或通过本规范中未详述的外部模块获取),选择与该接入单元关联并适用于Target0lsIdx的图像时间SEI消息(存在于BitstreamToDecode中或通过本规范中未详述的外部模块获取),当DecodingUnitHrdFlag等于1且decoding_unit_cpb_params_in_pic_timing_sei_flag等于0时,选择与该接入单元中的解码单元关联并适用于Target0lsIdx的解码单元信息SEI消息(存在于BitstreamToDecode中或通过本规范中未详述的外部机制获取)。

[0347] 每个一致性测试包括上述每个步骤中的一个选项的组合。当一个步骤存在多个选项时,任何特定的一致性测试只选择一个选项。所有步骤的所有可能组合构成一整套一致性测试。对于每个测试操作点,待执行的码流一致性测试的次数等于 $n0 \times n1 \times n2 \times n3$,其中, $n0$ 、 $n1$ 、 $n2$ 和 $n3$ 的值指定如下:

- [0348] -n1等于hrd_cpb_cnt_minus1+1。
- [0349] -n1是BitstreamToDecode中与缓冲周期SEI消息关联的接入单元的数量。
- [0350] -n2推导如下：
- [0351] -如果BitstreamToDecode是I类码流，则n0等于1。
- [0352] -否则(BitstreamToDecode是II类码流)，n0等于2。
- [0353] -n3推导如下：
- [0354] -如果decoding_unit_hrd_params_present_flag等于0，则n3等于1。
- [0355] -否则，n3等于2。
- [0356] HRD包括码流提取器(可选存在)、编码图像缓冲区(coded picture buffer,CPB)、即时解码过程、概念上包括每层的子解码图像缓冲区(decoded picture buffer,DPB)的DPB以及输出裁剪。
- [0357] 对于每个码流一致性测试,CPB大小(比特数)为CpbSize[Htid][ScIdx](如第7.4.5.2节详述,其中,ScIdx和HRD参数如上在本节详述),每层的DPB参数max_dec_pic_buffering_minus1[Htid]、max_num_reorder_pics[Htid]和MaxLatencyPictures[Htid]存在于适用于该层的语法结构dpb_parameters()中或者根据该语法结构推导,取决于该层是否是独立层以及该层是否是目标OLS中的输出层。
- [0358] HRD操作如下。
- [0359] -HRD在解码单元0处初始化,CPB和DPB的每个子DPB都被设置为空(每个子DPB的子DPB满量被设置为0)。
- [0360] 注:在初始化之后,HRD不会通过后续缓冲周期SEI消息再次初始化。
- [0361] -与根据指定的到达时间表(arrival schedule)进入每个CPB的解码单元关联的数据通过假设流调度方法(hypothetical stream scheduler,HSS)发送。
- [0362] -与每个解码单元关联的数据在解码单元的CPB移除时间通过即时解码过程即时移除和解码。
- [0363] -每个经解码图像都放置在DPB中。
- [0364] -当一个经解码图像不再需要用于帧间预测参考,也不再需要用于输出时,从DPB中移除该经解码图像。
- [0365] 解码图像缓冲区的操作如下。
- [0366] 本节中的这些规范独立适用于选择的每个解码图像缓冲区(decoded picture buffer,DPB)参数集(如第C.1节详述)。
- [0367] 解码图像缓冲区在概念上包括子DPB,每个子DPB包括用于存储一层经解码图像的图像存储缓冲区。每个图像存储缓冲区可以包括标识为“用于参考”或保存用于后续输出的经解码图像。第C3.2节、第C3.3节和第C3.4节中详述的过程按照下文所述依次应用,并从OLS中的最低层开始,按照OLS中各层的nuh_layer_id值的递增顺序独立应用于每层。当这些过程应用于特定层时,只影响该特定层的子DPB。在这些过程的描述中,DPB是指特定层的子DPB,而该特定层称为当前层。
- [0368] 注:在输出时间DPB的操作中,同一接入单元中的PicOutputFlag等于1的经解码图像按照经解码图像的nuh_layer_id值的递增顺序连续输出。
- [0369] 假设图像n和当前图像是接入单元n中具有特定nuh_layer_id值的经编码图像或

经解码图像,其中,n为非负整数。

[0370] 在对当前图像进行解码之前从DPB中移除图像如下所述。

[0371] 在对当前图像进行解码之前(但在解析当前图像中的第一条带的条带头之后)从DPB中移除图像可以即时发生在接入单元n(包括当前图像)的第一解码单元的CPB移除时间并执行如下:

[0372] -调用第8.3.2节详述的用于参考图像列表构建的解码过程和第8.3.3节详述的用于参考图像标识的解码过程。

[0373] -在当前AU为不是AU 0的CVSS AU时,可以使用以下按顺序执行的步骤:

[0374] 1.为测试解码器推导变量NoOutputOfPriorPicsFlag,如下所示:

[0375] -如果为当前AU中的任何图像推导出的pic_width_max_in_luma_samples、pic_height_max_in_luma_samples、chroma_format_idc、separate_colour_plane_flag、bit_depth_luma_minus8、bit_depth_chroma_minus8或max_dec_pic_buffering_minus1[Htid]的值不同于为同一CLVS中的前一图像推导出的pic_width_in_luma_samples、pic_height_in_luma_samples、chroma_format_idc、separate_colour_plane_flag、bit_depth_luma_minus8、bit_depth_chroma_minus8或max_dec_pic_buffering_minus1[Htid]的值,则NoOutputOfPriorPicsFlag可以(但不是必须)被测试解码器设置为1,不考虑no_output_of_prior_pics_flag的值。

[0376] 注:虽然在这些条件下,可以优选将NoOutputOfPriorPicsFlag设置为no_output_of_prior_pics_flag,但是在这种情况下,允许测试解码器将NoOutputOfPriorPicsFlag设置为1。

[0377] -否则,NoOutputOfPriorPicsFlag被设置为no_output_of_prior_pics_flag。

[0378] 2.为测试解码器推导出的NoOutputOfPriorPicsFlag的值应用于HRD,使得当NoOutputOfPriorPicsFlag的值等于1时,DPB中的所有图像存储缓冲区都被清空,而不输出这些图像存储缓冲区包括的图像,并且DPB满度被设置为0。

[0379] -当以下两个条件对于DPB中的任何图像k都为真时,DPB中的所有这些图像k都从DPB中移除:

[0380] -图像k可以标识为“不用于参考”。

[0381] -图像k的PictureOutputFlag等于0或其DPB输出时间小于或等于当前图像n中的第一解码单元(表示为解码单元m)的CPB移除时间,即DpbOutputTime[k]小于或等于DuCpbRemovalTime[m]。

[0382] -对于从DPB中移除的每个图像,DPB满度减1。

[0383] 输出顺序DPB的操作如下所述。

[0384] 本节中的这些规范独立适用于选择的每个解码图像缓冲区(decoded picture buffer,DPB)参数集(如第C.1节详述)。

[0385] 解码图像缓冲区在概念上包括子DPB,每个子DPB包括用于存储一层经解码图像的图像存储缓冲区。每个图像存储缓冲区包括标识为“用于参考”或保存用于后续输出的经解码图像。

[0386] 在对当前图像进行解码之前调用从DPB中输出和移除图像的过程(如第C.5.2.2节详述),然后调用用于当前解码图像标识和存储的过程(如第C.3.4节详述),最后调用用于

额外缓冲 (bumping) 的过程 (如C.5.2.3节详述)。“缓冲”过程如第C.5.2.4节详述并如第C.5.2.2节和第C.5.2.3节详述调用。

[0387] 这些过程从OLS中的最低层开始,按照OLS中各层的nuh_layer_id值的递增顺序独立应用于每层。当这些过程应用于特定层时,只影响该特定层的子DPB。

[0388] 注:在输出顺序DPB的操作中,与在输出时间DPB的操作中一样,同一接入单元中的PicOutputFlag等于1的经解码图像按照经解码图像的nuh_layer_id值的递增顺序连续输出。

[0389] 假设图像n和当前图像是接入单元n中具有特定nuh_layer_id值的经编码图像或经解码图像,其中,n为非负整数。

[0390] 从DPB中输出和移除图像如下所述。

[0391] 在对当前图像进行解码之前(但在解析当前图像中的第一条带的条带头之后)从DPB中输出和移除图像即时发生在包括当前图像的接入单元中的第一解码单元从CPB中移除的时候并执行如下:

[0392] -调用第8.3.2节详述的用于参考图像列表构建的解码过程和第8.3.3节详述的用于参考图像标识的解码过程。

[0393] -如果当前AU为不是AU 0的CVSS AU时,可以使用以下按顺序执行的步骤:

[0394] 1.为测试解码器推导变量NoOutputOfPriorPicsFlag,如下所示:

[0395] -如果为当前AU中的任何图像推导出的pic_height_max_in_luma_samples、chroma_format_idc、separate_colour_plane_flag、bit_depth_luma_minus8、bit_depth_chroma_minus8或max_dec_pic_buffering_minus1[Htid]的值不同于为同一CLVS中的前一图像推导出的pic_width_in_luma_samples、pic_height_in_luma_samples、chroma_format_idc、separate_colour_plane_flag、bit_depth_luma_minus8、bit_depth_chroma_minus8或max_dec_pic_buffering_minus1[Htid]的值,则NoOutputOfPriorPicsFlag可以(但不是必须)被测试解码器设置为1,不考虑no_output_of_prior_pics_flag的值。

[0396] 注:虽然在这些条件下,可以优选将NoOutputOfPriorPicsFlag设置为no_output_of_prior_pics_flag,但是在这种情况下,允许测试解码器将NoOutputOfPriorPicsFlag设置为1。

[0397] -否则,NoOutputOfPriorPicsFlag被设置为no_output_of_prior_pics_flag。

[0398] 2.为测试解码器推导出的NoOutputOfPriorPicsFlag的值应用于HRD,如下所示:

[0399] -如果NoOutputOfPriorPicsFlag等于1,则DPB中的所有图像存储缓冲区都被清空,而不输出这些图像存储缓冲区包括的图像,并且DPB满度被设置为0。

[0400] -否则(NoOutputOfPriorPicsFlag等于0),包括标识为“不需要输出”且“不用于参考”的图像的所有图像存储缓冲区被清空(不输出),而DPB中的所有非空图像存储缓冲区通过反复调用第C.5.2.4节详述的“缓冲”过程来清空,并且DPB满度被设置为0。

[0401] -否则(当前图像不是CLVSS图像),包括标识为“不需要输出”且“不用于参考”的图像的所有图像存储缓冲区都被清空(不输出)。对于每个被清空的图像存储缓冲区,DPB满度减1。当以下一个或多个条件为真时,重复调用第C.5.2.4节详述的“缓冲”过程,同时对于每个被清空的其它图像存储缓冲区,进一步将DPB满度减1,直到以下条件都不为真:

[0402] DPB中标识为“需要输出”的图像的数量大于max_num_reorder_pics[Htid]。

[0403] `max_latency_increase_plus1[Htid]` 不等于0且DPB中存在至少一个图像标识为“需要输出”,该图像的关联变量`PicLatencyCount`大于或等于`MaxLatencyPictures[Htid]`。

[0404] DPB中的图像的数量大于或等于`max_dec_pic_buffering_minus1[Htid]+1`。

[0405] 额外缓冲如下所述。

[0406] 本节详述的过程即时发生在包括当前图像的接入单元n中的最后一个解码单元从CPB中移除的时候。

[0407] 在当前图像的`PictureOutputFlag`等于1时,对于DPB中标识为“需要输出”且输出顺序在当前图像之后的每个图像,关联变量`PicLatencyCount`被设置为`PicLatencyCount+1`。

[0408] 以下内容适用:

[0409] -如果当前经解码图像的`PictureOutputFlag`等于1,则当前经解码图像被标识为“需要输出”且其关联变量`PicLatencyCount`被设置为0。

[0410] -否则(当前经解码图像的`PictureOutputFlag`等于0),当前经解码图像被标识为“不需要输出”。

[0411] 当以下一个或多个条件为真时,重复调用第C.5.2.4节详述的“缓冲”过程,直到以下条件都不为真:

[0412] -DPB中标识为“需要输出”的图像的数量大于`max_num_reorder_pics[Htid]`。

[0413] -`max_latency_increase_plus1[Htid]` 不等于0且DPB中存在至少一个图像标识为“需要输出”,该图像的关联变量`PicLatencyCount`大于或等于`MaxLatencyPictures[Htid]`。

[0414] 缓冲过程如下所述。

[0415] “缓冲”过程包括以下按顺序执行的步骤:

[0416] 选择首先输出的图像作为DPB中标识为“需要输出”的所有图像中具有最小`PicOrderCntVal`值的图像。

[0417] 这些图像中的每个图像按照`nuh_layer_id`的递增顺序使用该图像的一致性裁剪窗口进行裁剪,输出裁剪后的图像,将该图像标识为“不需要输出”。

[0418] 包括标识为“不用于参考”且是裁剪后的输出图像之一的图像的每个图像存储缓冲区被清空,并且关联于DPB的满度减1。

[0419] 注:对于属于同一CVS并通过“缓冲过程”输出的任意两个图像`picA`和`picB`,当`picA`比`picB`先输出时,`picA`的`PicOrderCntVal`的值小于`picB`的`PicOrderCntVal`的值。

[0420] 子码流提取过程如下所述。

[0421] 该过程的输入包括码流`inBitstream`、目标OLS索引`targetOlsIdx`和目标最大TemporalId值`tIdTarget`。

[0422] 该过程的输出是子码流`outBitstream`。

[0423] 码流一致性要求是:对于任何输入码流,本节详述的以码流、等于由VPS指定的OLS列表中的索引的`targetOlsIdx`和等于0至6(包括端值)范围内任一值的`tIdTarget`为输入的过程输出的且满足以下条件的输出子码流应是一致性码流。

[0424] -输出子码流包括`nuh_layer_id`等于`LayerIdInOls[targetOlsIdx]`中的每个`nuh_layer_id`值的至少一个VCL NAL单元。

[0425] -输出子码流包括TemporalId等于`tIdTarget`的至少一个VCL NAL单元。

[0426] 注：一致性码流包括TemporalId等于0的一个或多个经编码条带NAL单元，但不必包括nuh_layer_id等于0的经编码条带NAL单元。

[0427] 输出子码流OutBitstream推导如下：

[0428] -码流outBitstream被设置为与码流inBitstream相同。

[0429] -从outBitstream中移除TemporalId大于tIdTarget的所有NAL单元。

[0430] -从outBitstream中移除nuh_layer_id不包括在列表LayerIdInOls[targetOlsIdx]中的所有NAL单元。

[0431] -从outBitstream中移除包括nesting_ols_flag等于1的可伸缩嵌套式SEI消息的所有SEI NAL单元，i的取值范围不是0至nesting_num_olss_minus1(包括端值)，使得NestingOlsIdx[i]等于targetOlsIdx。

[0432] -当targetOlsIdx大于0时，从outBitstream中移除包括payloadType等于0(缓冲周期)、1(图像时间)或130(解码单元信息)的不可伸缩嵌套的SEI消息的所有SEI NAL单元。

[0433] 下面提供了可伸缩嵌套式SEI消息语法。

[0434]	scalable_nesting(payloadSize) {	描述符
	nesting_ols_flag	u(1)
	if(nesting_ols_flag) {	
	nesting_num_olss_minus1	ue(v)
	for(i = 0; i <= nesting_num_olss_minus1; i++)	
	nesting_ols_idx_delta_minus1[i]	ue(v)
	} else {	
	nesting_all_layers_flag	u(1)
	if(!nesting_all_layers_flag) {	
	nesting_num_layers_minus1	ue(v)
	for(i = 1; i <= nesting_num_layers_minus1; i++)	
	nesting_layer_id[i]	u(6)
	}	
[0435]	}	
	nesting_num_seis_minus1	ue(v)
	while(!byte_aligned())	
	nesting_zero_bit /*等于 0*/	u(1)
	for(i = 0; i <= nesting_num_seis_minus1; i++)	
	sei_message()	
	}	
	}	

[0436] 下文提供了通用SEI载荷语义。

[0437] 以下内容适用于不可伸缩嵌套的SEI消息的适用层或OLS。

[0438] -对于不可伸缩嵌套的SEI消息，当payloadType等于0(缓冲周期)、1(图像时间)或130(解码单元信息)时，不可伸缩嵌套的SEI消息只适用于第0个OLS。

[0439] -对于不可伸缩嵌套的SEI消息,当payloadType等于VclAssociatedSeiList中的任一值时,不可伸缩嵌套的SEI消息只适用于VCL NAL单元的nuh_layer_id等于包括SEI消息的SEI NAL单元的nuh_layer_id的层。

[0440]

[0441] 码流一致性的要求是:以下限制条件适用于SEI NAL单元的nuh_layer_id值:

[0442] -当不可伸缩嵌套的SEI消息的payloadType等于0(缓冲周期)、1(图像时间)或130(解码单元信息)时,包括不可伸缩嵌套的SEI消息的SEI NAL单元的nuh_layer_id应等于vps_layer_id[0]。

[0443] -当不可伸缩嵌套的SEI消息的payloadType等于VclAssociatedSeiList中的任一值时,包括不可伸缩嵌套的SEI消息的SEI NAL单元的nuh_layer_id应等于与SEI NAL单元关联的VCL NAL单元的nuh_layer_id值。

[0444] -包括可伸缩嵌套式SEI消息的SEI NAL单元的nuh_layer_id应等于可伸缩嵌套的SEI消息适用于的所有层的最小nuh_layer_id值(当可伸缩嵌套式SEI消息的nesting_ols_flag等于0时)或等于可伸缩嵌套的SEI消息适用于的OLS中的所有层的最小nuh_layer_id值(当可伸缩嵌套式SEI消息的nesting_ols_flag等于1时)。

[0445] 下文提供了可伸缩嵌套式SEI消息语义。

[0446] 可伸缩嵌套式SEI消息提供一种将SEI消息与特定OLS或特定层关联的机制。

[0447] 可伸缩嵌套式SEI消息包括一个或多个SEI消息。可伸缩嵌套式SEI消息中包括的SEI消息还称为可伸缩嵌套的SEI消息。

[0448] 码流一致性的要求是,以下限制条件适用于将SEI消息包括在可伸缩嵌套式SEI消息中:

[0449] -payloadType等于132(经解码图像哈希(hash))或133(可伸缩嵌套式)的SEI消息可以不包括在可伸缩嵌套式SEI消息中。

[0450] -当可伸缩嵌套式SEI消息包括缓冲周期、图像时间或解码单元信息SEI消息时,可伸缩嵌套式SEI消息应不包括payloadType不等于0(缓冲周期)、1(图像时间)或130(解码单元信息)的任何其它SEI消息。

[0451] 码流一致性的要求是,以下限制条件应用于包括可伸缩嵌套式SEI消息的SEI NAL单元的nal_unit_type值:

[0452] -当可伸缩嵌套式SEI消息包括payloadType等于0(缓冲周期)、1(图像时间)、130(解码单元信息)、145(依赖RAP指示)或168(帧字段信息)的SEI消息时,包括可伸缩嵌套式SEI消息的SEI NAL单元的nal_unit_type应等于PREFIX_SEI_NUT。

[0453] nesting_ols_flag等于1表示可伸缩嵌套的SEI消息适用于特定OLS。nesting_ols_flag等于0表示可伸缩嵌套的SEI消息适用于特定层。

[0454] 码流一致性的要求是:以下限制条件适用于nesting_ols_flag值。

[0455] -当可伸缩嵌套式SEI消息包括payloadType等于0(缓冲周期)、1(图像时间)或130(解码单元信息)的SEI消息时,nesting_ols_flag值应等于1。

[0456] -当可伸缩嵌套式SEI消息包括payloadType等于VclAssociatedSeiList中的值的SEI消息时,nesting_ols_flag值应等于0。

[0457] nesting_num_olss_minus1+1表示可伸缩嵌套的SEI消息适用于的OLS的数量。

nesting_num_olss_minus1的取值范围应为0至TotalNumOlss-1(包括端值)。

[0458] nesting_ols_idx_delta_minus1[i]用于推导变量NestingOlsIdx[i],该变量表示可伸缩嵌套的SEI消息在nesting_ols_flag等于1时适用于的第i个OLS的OLS索引。nesting_ols_idx_delta_minus1[i]的取值范围应为0至TotalNumOlss-2(包括端值)。

[0459] 变量NestingOlsIdx[i]推导如下:

```

        if( i == 0 )
            NestingOlsIdx[ i ] = nesting_ols_idx_delta_minus1[ i ]
[0460]     else
            NestingOlsIdx[ i ] = NestingOlsIdx[ i - 1 ] + nesting_ols_idx_delta_minus1[ i ] + 1

```

[0461] nesting_all_layers_flag等于1表示可伸缩嵌套的SEI消息适用于nuh_layer_id大于或等于当前SEI NAL单元的nuh_layer_id的所有层。nesting_all_layers_flag等于0表示可伸缩嵌套的SEI消息可以或可以不适用于nuh_layer_id大于或等于当前SEI NAL单元的nuh_layer_id的所有层。

[0462] nesting_num_layers_minus1+1表示可伸缩嵌套的SEI消息适用于的层的数量。nesting_num_layers_minus1的取值范围应为0至vps_max_layers_minus1-GeneralLayerIdx[nuh_layer_id](包括端值),其中,nuh_layer_id是当前SEI NAL单元的nuh_layer_id。

[0463] nesting_layer_id[i]表示可伸缩嵌套的SEI消息在nesting_all_layers_flag等于0时适用于的第i层的nuh_layer_id值。nesting_layer_id[i]的值应大于nuh_layer_id,其中,nuh_layer_id是当前SEI NAL单元的nuh_layer_id。

[0464] 当nesting_ols_flag等于0时,变量NestingNumLayers和列表NestingLayerId[i](i的范围为0至NestingNumLayers-1,包括端值)可以推导如下,其中,变量NestingNumLayers表示可伸缩嵌套的SEI消息适用于的层的数量,列表NestingLayerId[i]表示可伸缩嵌套的SEI消息适用于的各层的nuh_layer_id值的列表,nuh_layer_id是当前SEI NAL单元的nuh_layer_id。

```

        if( nesting_all_layers_flag ) {
[0465]             NestingNumLayers =
                ps_max_layers_minus1 + 1 - GeneralLayerIdx[ nuh_layer_id ]
                for( i = 0; i < NestingNumLayers; i ++ )
                    NestingLayerId[ i ] = vps_layer_id[ GeneralLayerIdx[ nuh_layer_id ] + i ]
        } else {
[0466]             NestingNumLayers = nesting_num_layers_minus1 + 1
                for( i = 0; i < NestingNumLayers; i ++ )
                    NestingLayerId[ i ] = ( i == 0 ) ? nuh_layer_id : nesting_layer_id[ i ]
        }

```

[0467] nesting_num_seis_minus1+1表示可伸缩嵌套的SEI消息的数量。nesting_num_seis_minus1的取值范围应为0至63(包括端值)。

[0468] nesting_zero_bit应等于0。

[0469] 图11为本发明一个实施例提供的视频译码设备1100(例如,视频编码器300或视频

解码器400)的示意图。视频译码设备1100适用于实现本文描述的公开实施例。视频编码设备1100包括:用于接收数据的入端口1110和接收单元(Rx)1120;用于处理所述数据的处理器、逻辑单元或中央处理单元(central processing unit,CPU)1130;用于发送所述数据的发送单元(Tx)1140和出端口1150;用于存储所述数据的存储器1160。视频译码设备1100还可以包括与入端口1110、接收单元1120、发送单元1140和出端口1150耦合的光电(optical-to-electrical,OE)组件和电光(electrical-to-optical,E0)组件,用作光信号或电信号的出口或入口。

[0470] 处理器1130通过硬件和软件来实现。处理器1130可以实现为一个或多个CPU芯片、一个或多个核(例如多核处理器)、一个或多个现场可编程门阵列(field-programmable gate array,FPGA)、一个或多个专用集成电路(application specific integrated circuit,ASIC)和一个或多个数字信号处理器(digital signal processor,DSP)。处理器1130与入端口1110、接收器单元1120、发送单元1140、出端口1150和存储器1160通信。处理器1130包括译码模块1170。译码模块1170实现上文描述的公开实施例。例如,译码模块1170执行、处理、准备或提供各种编解码功能。因此,将译码模块1170包含在内为视频译码设备1100的功能提供了实质性的改进,并且影响了视频译码设备1100到不同状态的转换。可选地,以存储在存储器1160中并由处理器1130执行的指令来实现译码模块1170。

[0471] 视频译码设备1100还可以包括用于与用户进行数据通信的输入和/或输出(input/output,I/O)设备1180。I/O设备1180可以包括输出设备,例如,用于显示视频数据的显示器、用于输出音频数据的扬声器等。I/O设备1180还可以包括键盘、鼠标、轨迹球等输入设备和/或用于与这些输出设备交互的对应接口。

[0472] 存储器1160包括一个或多个磁盘、一个或多个磁带机和一个或多个固态硬盘,并且可以用作溢出数据存储设备,以在选择程序来执行时存储此类程序以及存储在执行程序过程中读取的指令和数据。存储器1160可以是易失性和/或非易失性的,并且可以是只读存储器(read-only memory,ROM)、随机存取存储器(random access memory,RAM)、三态内容寻址存储器(ternary content-addressable memory,TCAM)和/或静态随机存取存储器(static random-access memory,SRAM)。

[0473] 图12为译码模块1200的一个实施例的示意图。在一个实施例中,译码模块1200在视频译码设备1202(例如,视频编码器300或视频解码器400)中实现。视频译码设备1202包括接收模块1201。接收模块1201用于接收图像进行编码或接收码流进行解码。视频译码设备1202包括与接收模块1201耦合的发送模块1207。发送模块1207用于将码流发送给解码器或将经解码图像发送给显示模块(例如其中一个I/O设备1180)。

[0474] 视频译码设备1202包括存储模块1203。存储模块1203与接收模块1201或发送模块1207中的至少一个耦合。存储模块1203用于存储指令。视频译码设备1202还包括处理模块1205。处理模块1205与存储模块1203耦合。处理模块1205用于执行存储在存储模块1203中的指令,以执行本文公开的方法。

[0475] 还应当理解,本文中阐述的示例性方法的步骤不一定需要按照所描述的顺序执行,并且这些方法的步骤的顺序应当理解为仅仅是示例性的。同理,在与本发明各种实施例相一致的方法中,这些方法可以包括其它步骤,并且某些步骤可以省略或组合。

[0476] 虽然本发明提供了若干个实施例,但应当理解,在不脱离本发明的精神或范围的

情况下,所公开的系统和方法可以通过其它多种具体形式体现。本发明示例应被视为说明性而非限制性的,且本发明并不限于本文所给出的详细内容。例如,各种元件或组件可以组合或集成在另一系统中,或者某些特征可以省略或不实现。

[0477] 另外,在不脱离本发明的范围的情况下,各种实施例中描述和说明为离散或单独的技术、系统、子系统和方法可以与其它系统、模块、技术或方法组合或集成。展示或论述为彼此耦合或直接耦合或通信的其它项也可以采用电方式、机械方式或其它方式通过某一接口、设备或中间组件间接地耦合或通信。变化、替换、变更的其它示例可由本领域技术人员确定,并可以在不脱离本文公开的精神和范围的情况下举例。

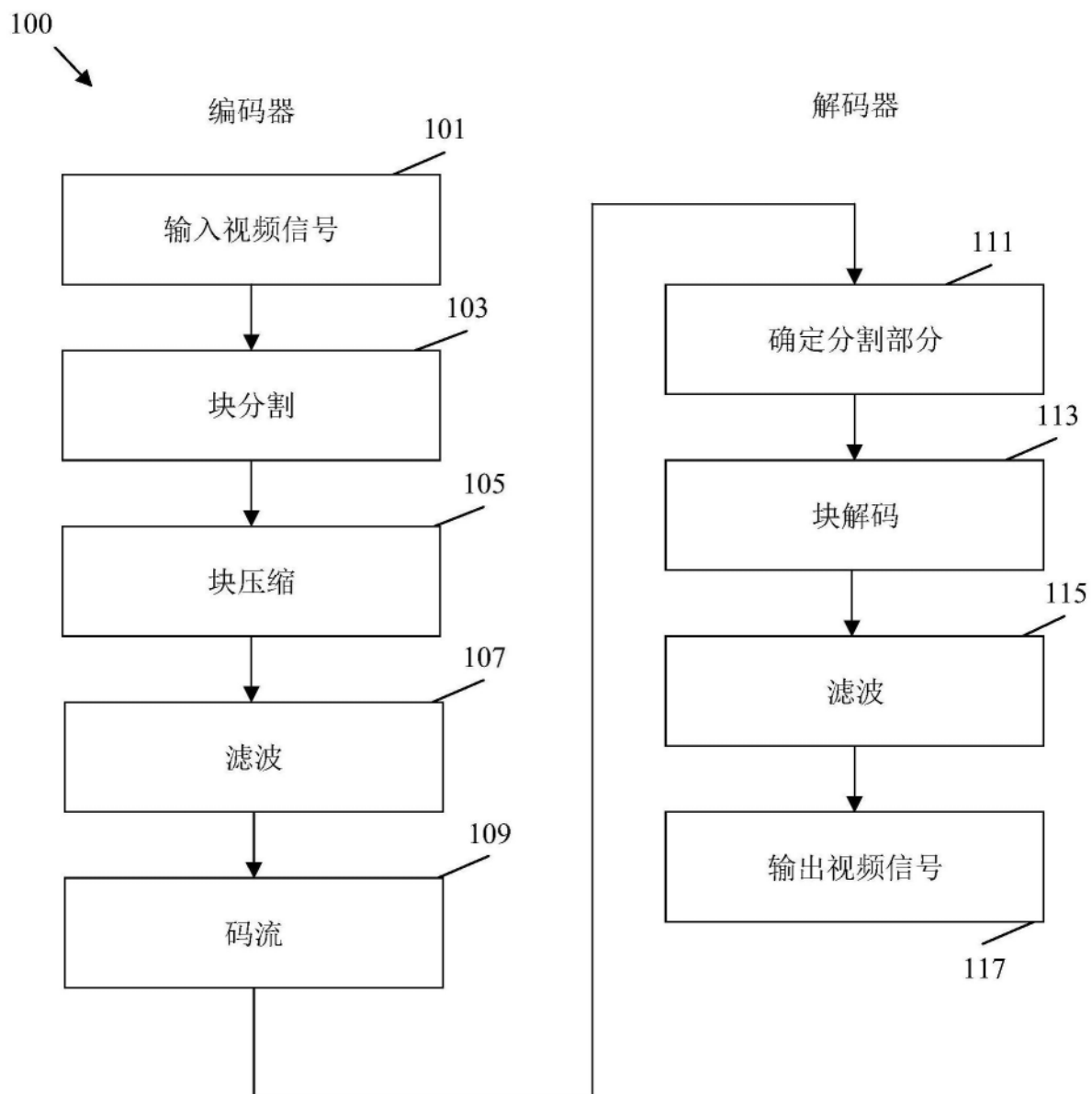


图1

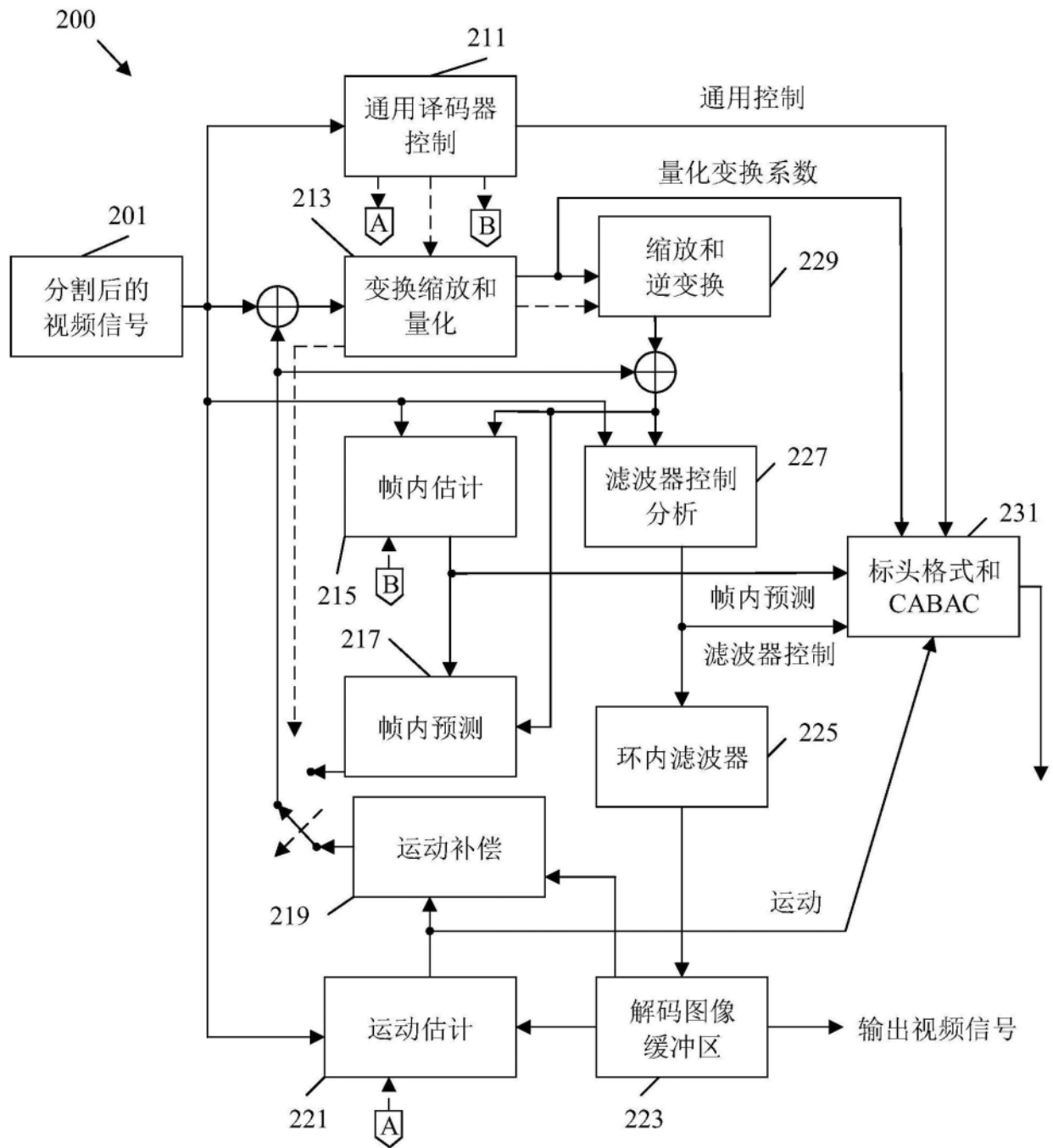


图2

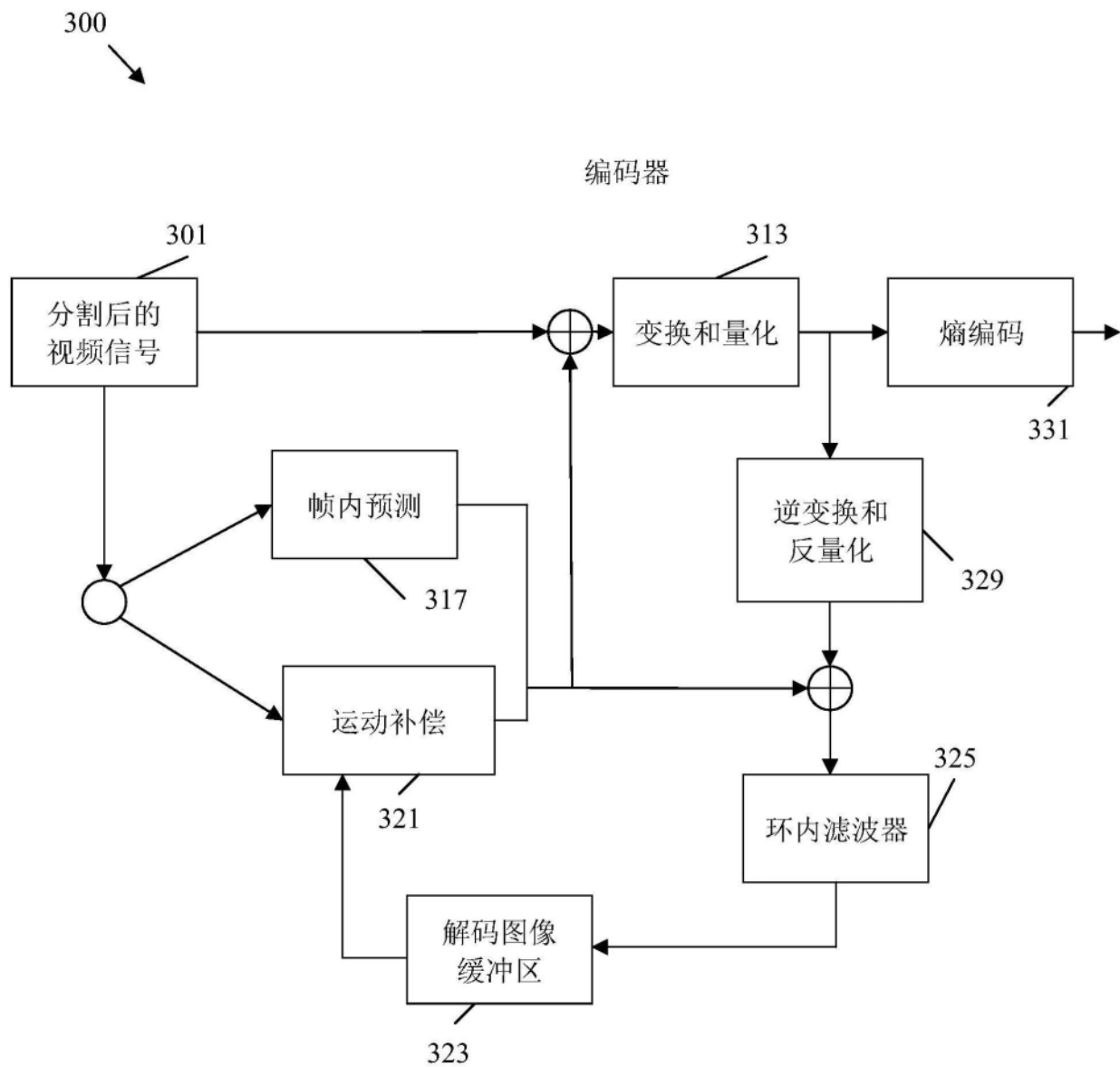


图3

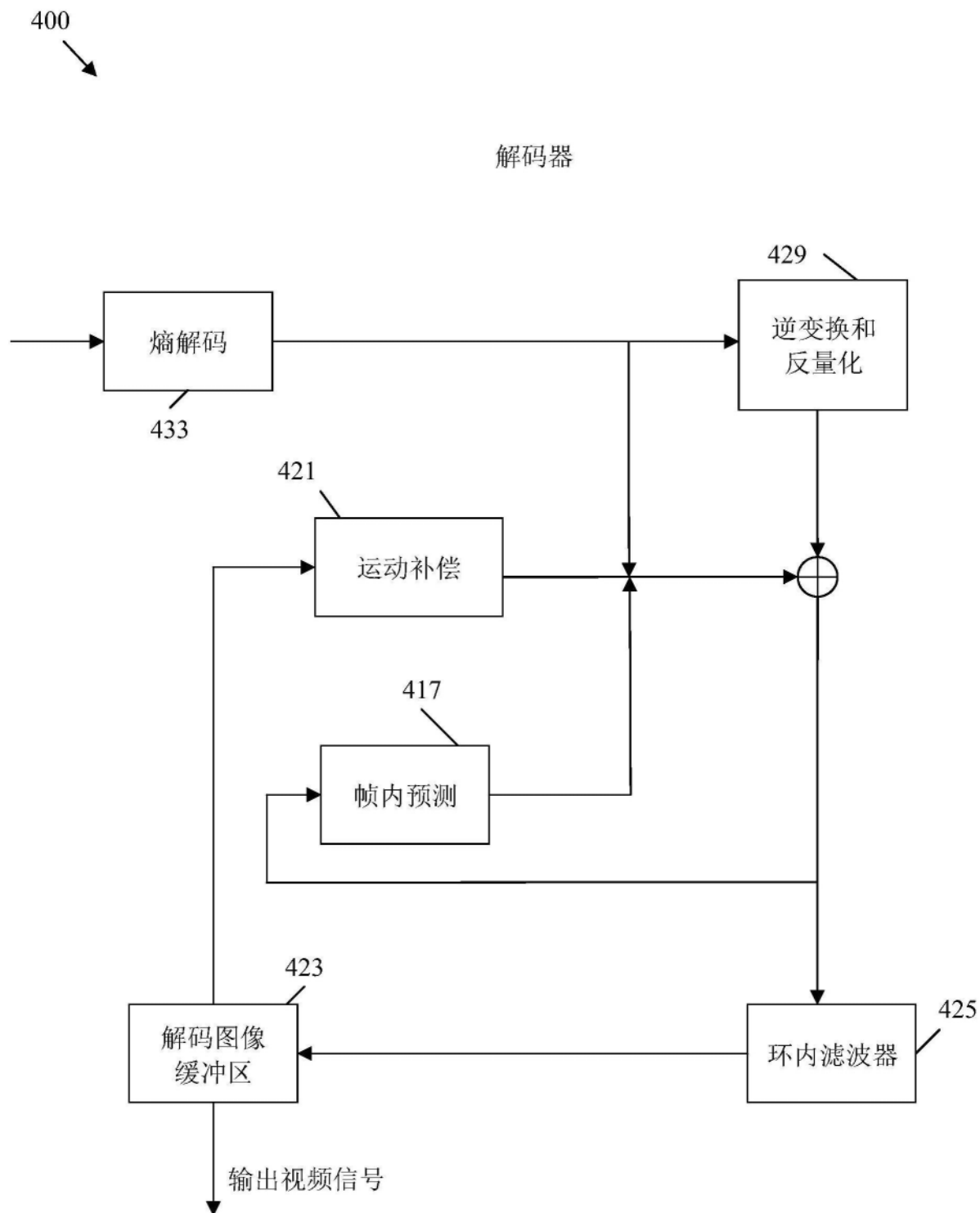


图4

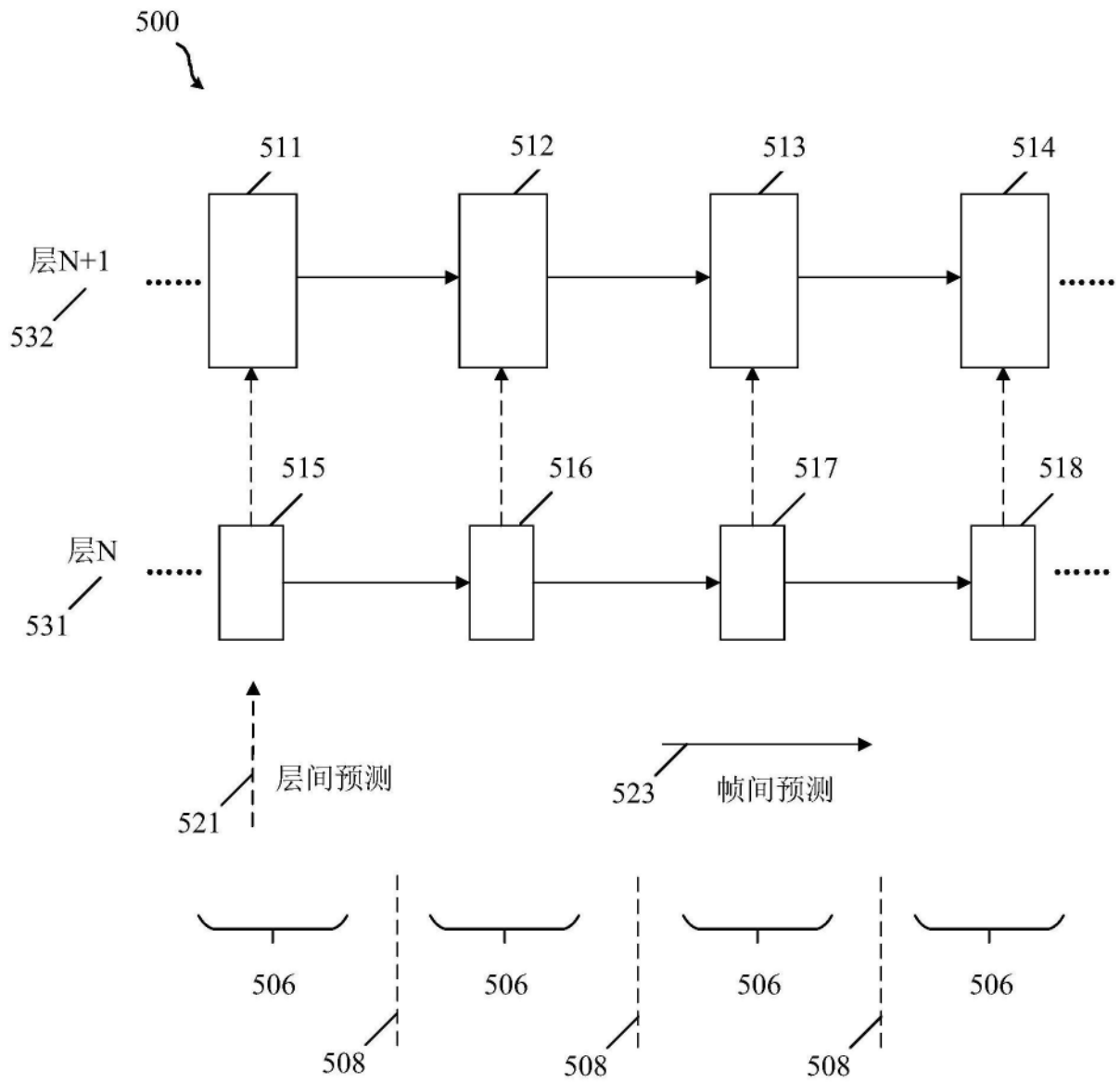


图5

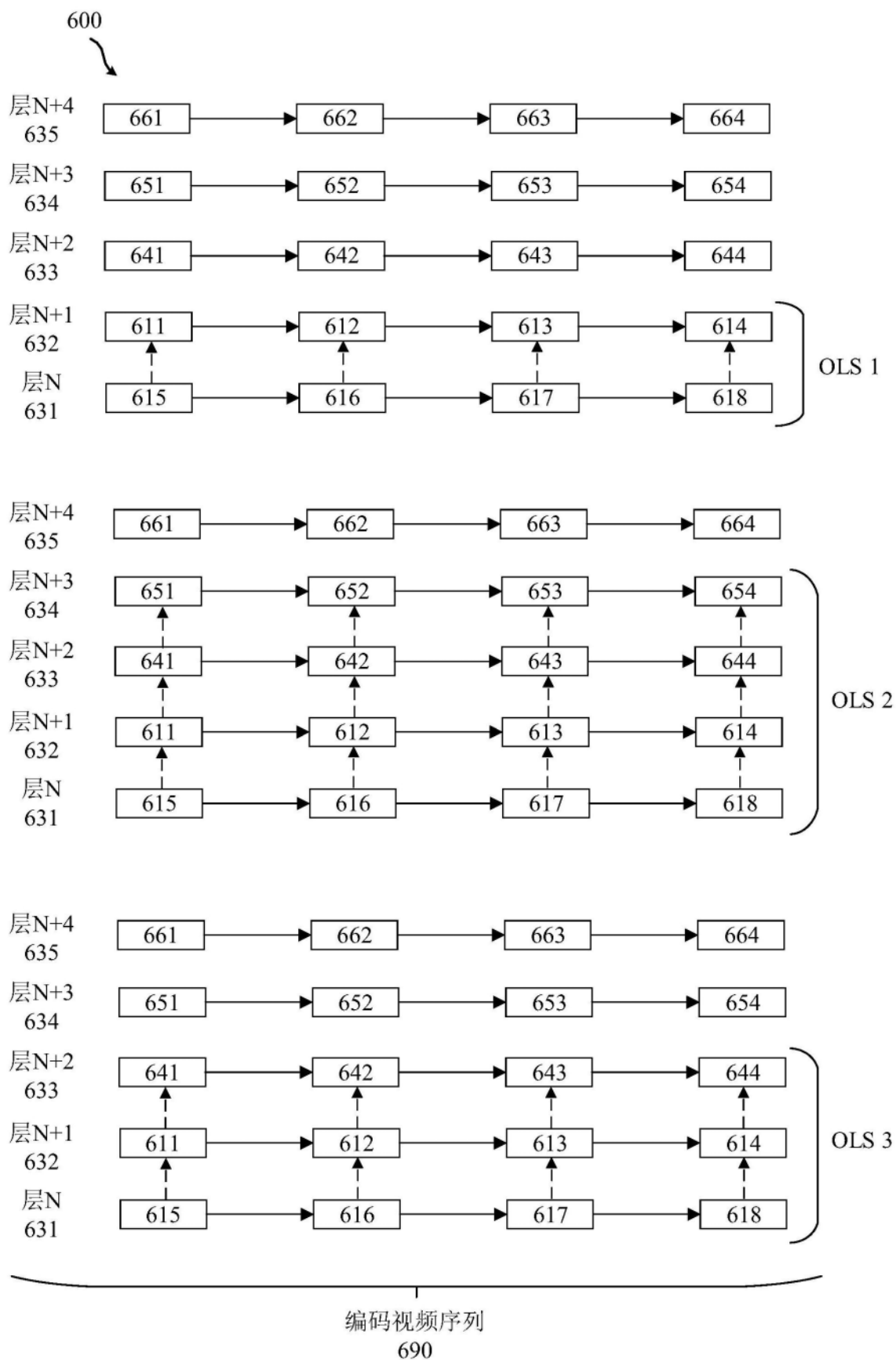


图6

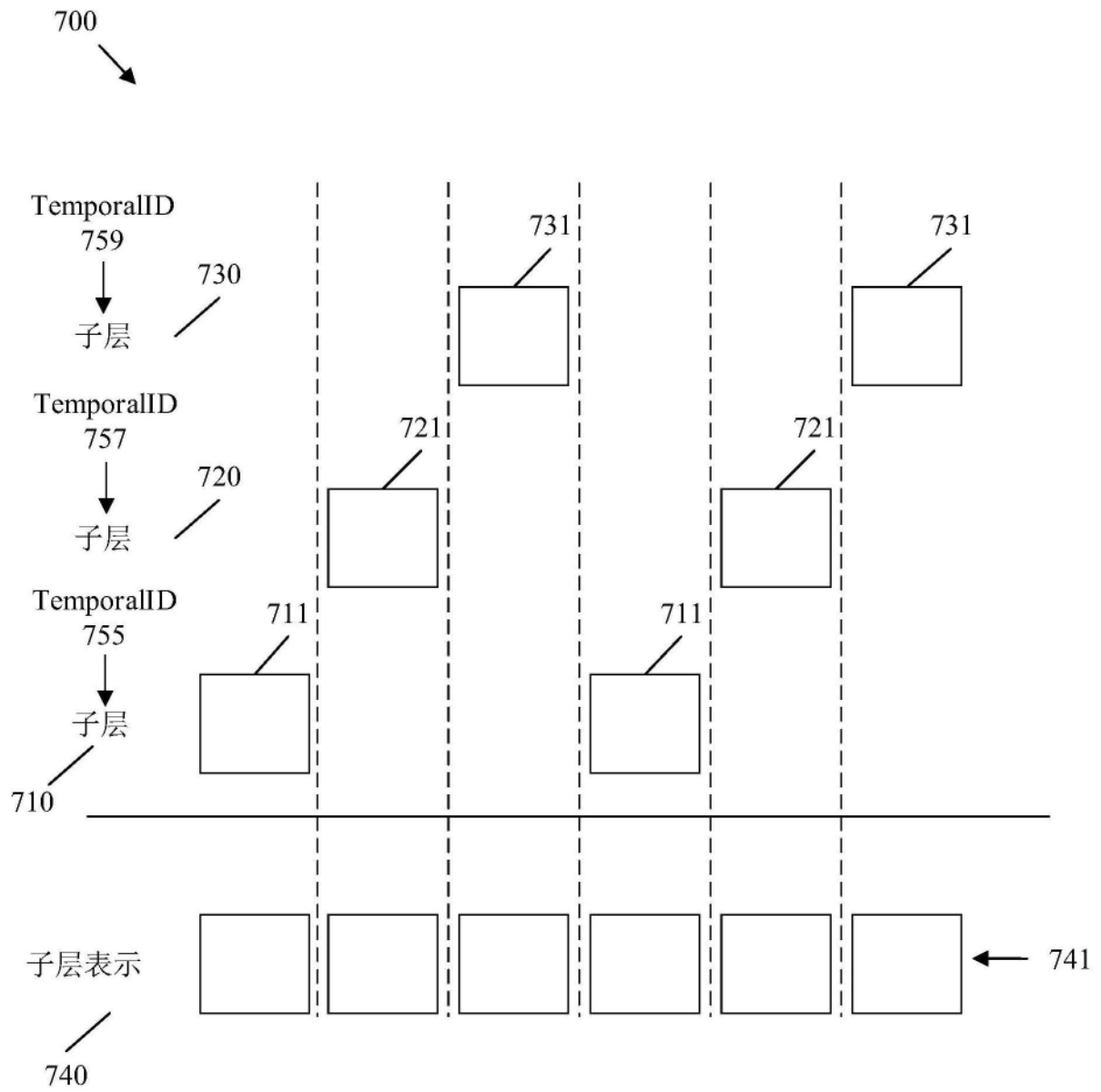


图7

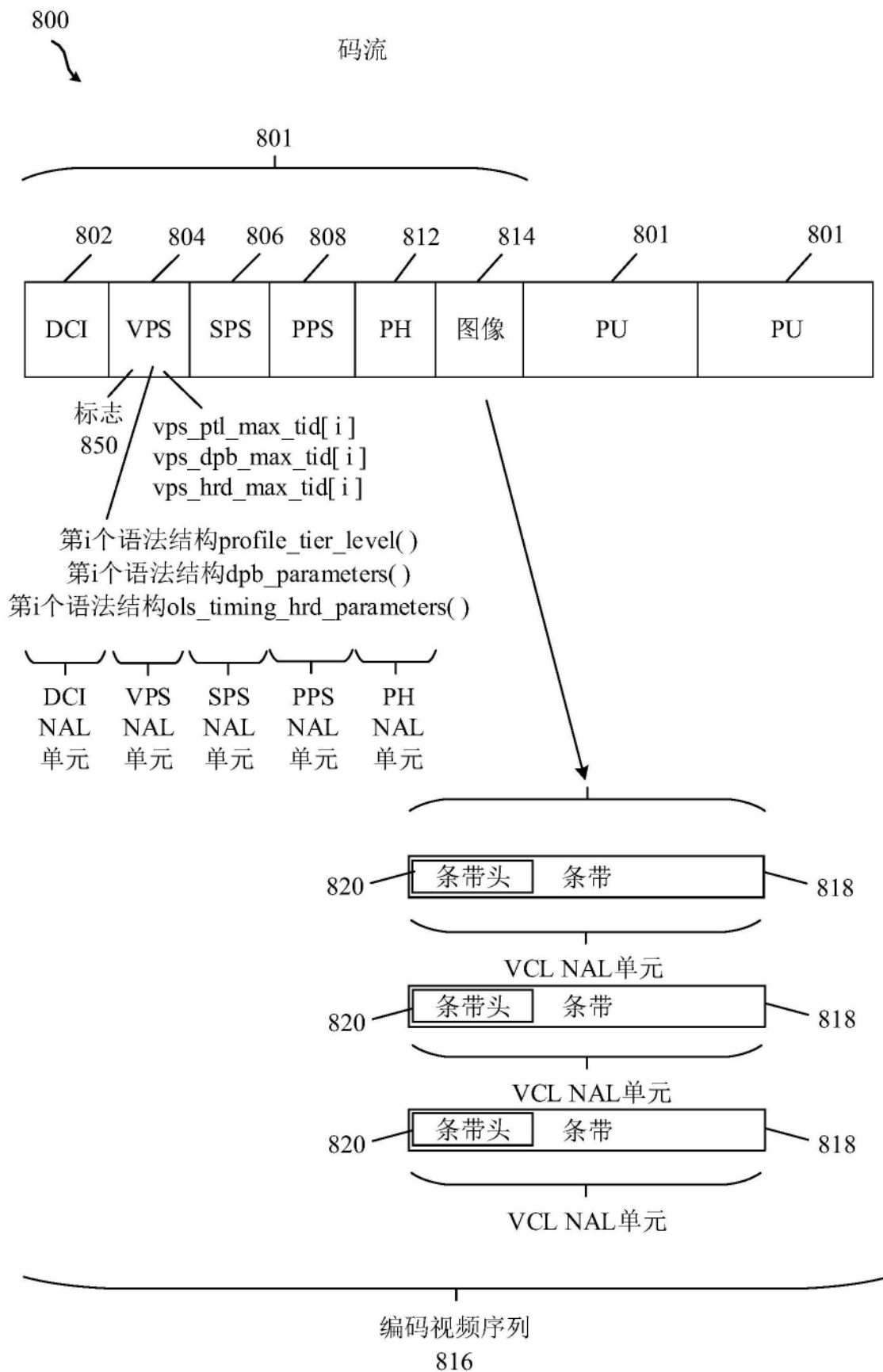


图8

900

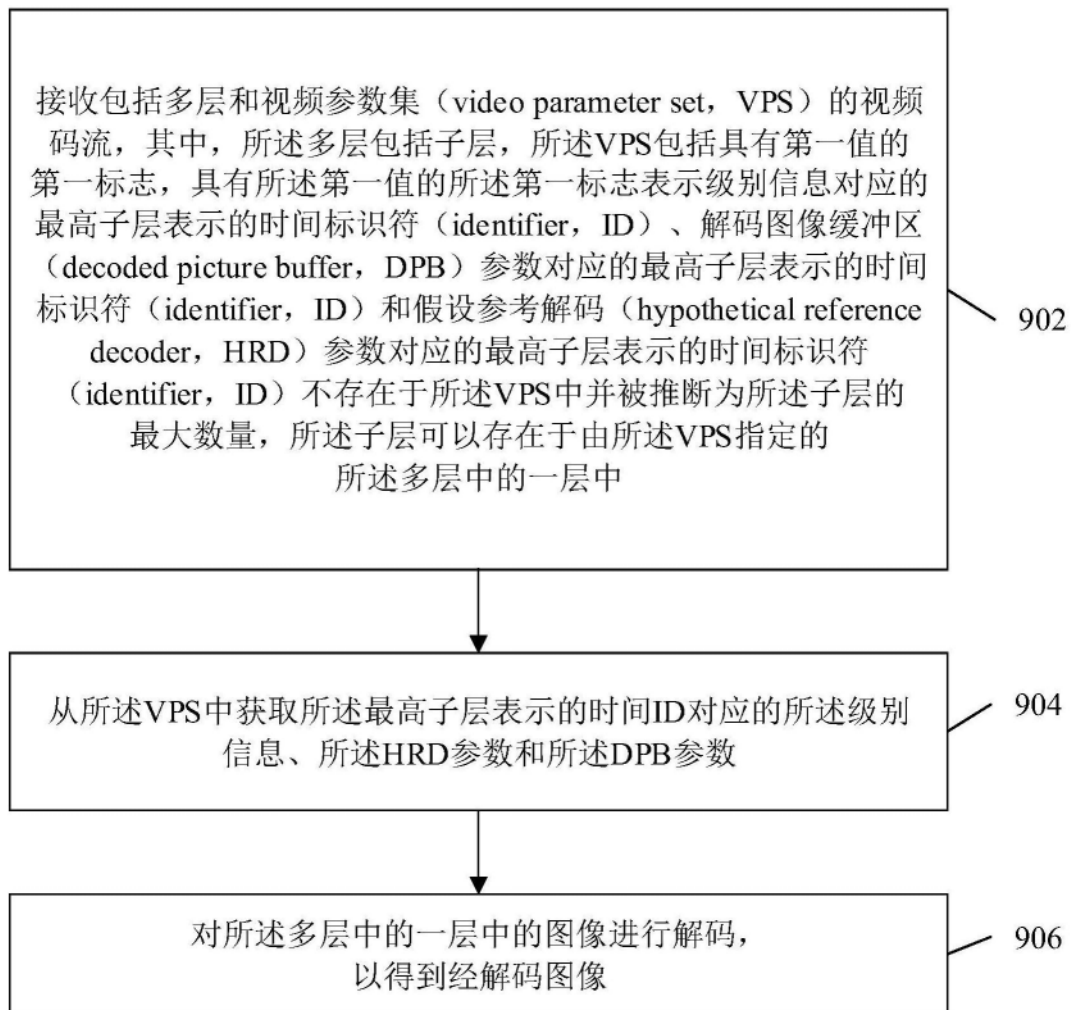


图9

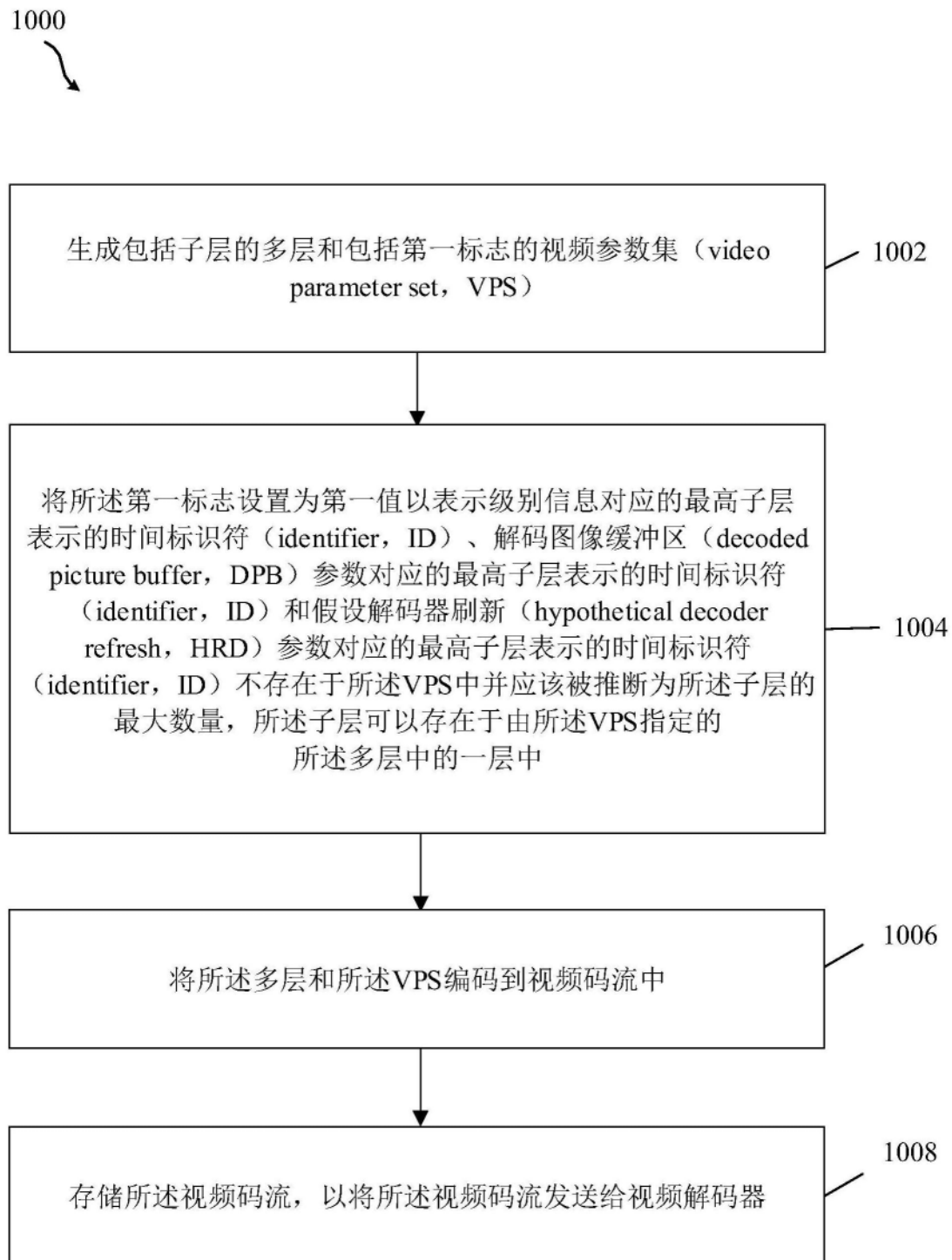


图10

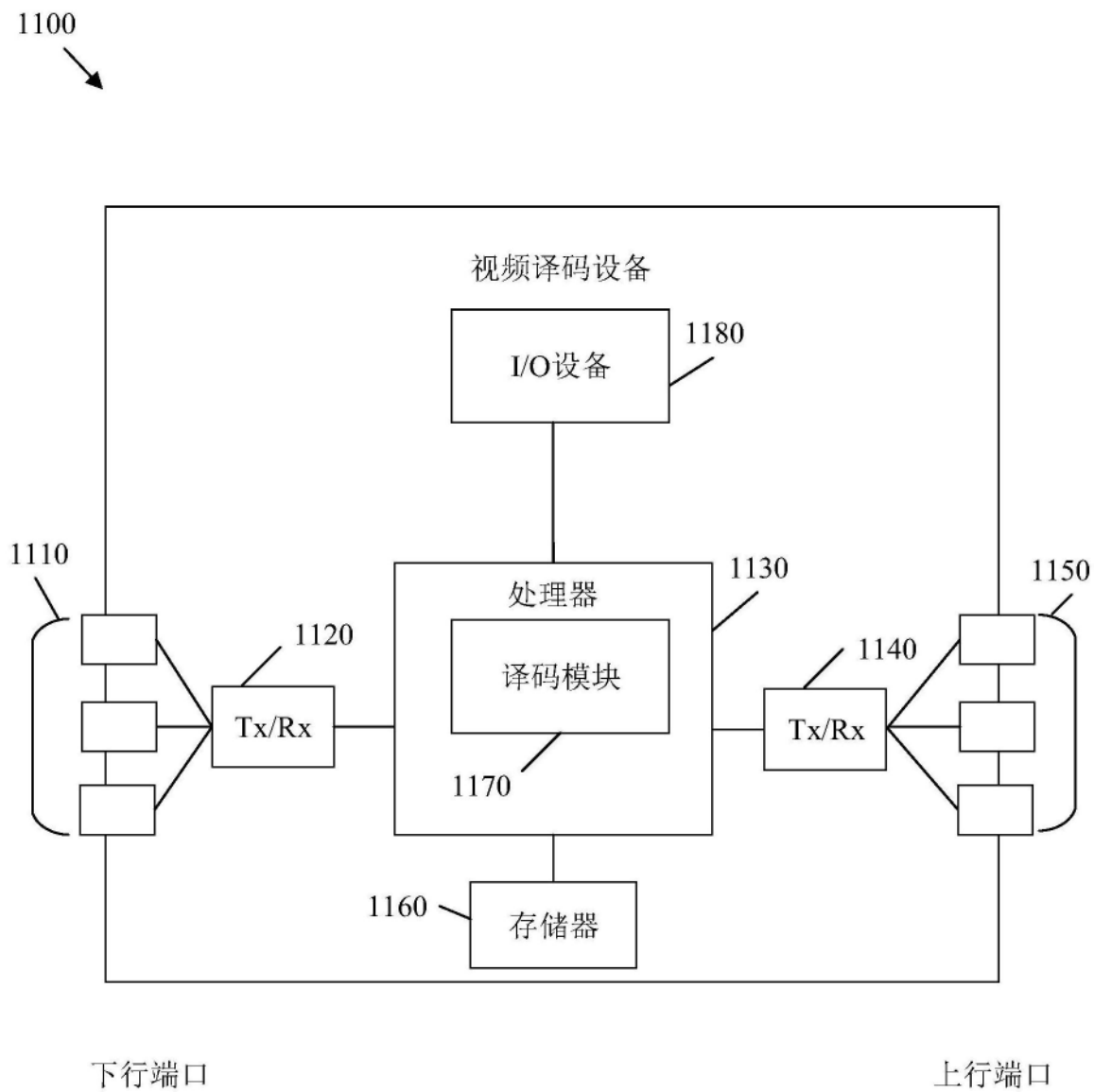


图11

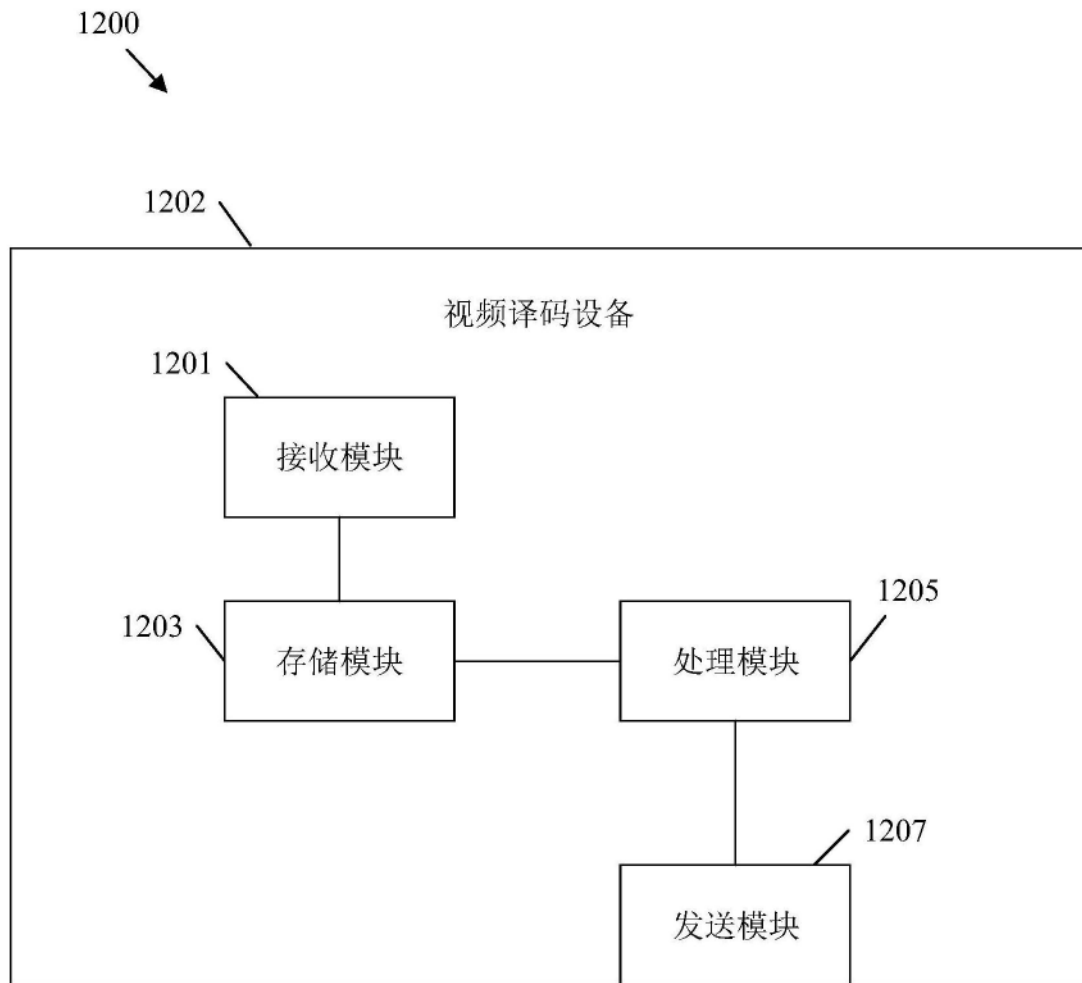


图12