

[19] Patents Registry
The Hong Kong Special Administrative Region
香港特別行政區
專利註冊處

[11] 1209552 B
CN 104685890 B

[12] **STANDARD PATENT (R) SPECIFICATION**
轉錄標準專利說明書

[21] Application no. 申請編號
15110263.2

[51] Int. Cl.
H04N 19/70 (2014.01) H04N 19/30 (2014.01)

[22] Date of filing 提交日期
19.10.2015

[43] METHODS, DEVICES AND APPRATUSES FOR PROCESSING AND ENCODING MULTI-LAYER VIDEO DATA
用於處理和編碼多層視頻數據的方法、裝置和設備

[30] Priority 優先權
02.10.2012 US 61/709,094
30.09.2013 US 14/042,416

[43] Date of publication of application 申請發表日期
01.04.2016

[45] Date of publication of grant of patent 批予專利的發表日期
13.03.2020

[86] International application no. 國際申請編號
PCT/US2013/062890

[87] International publication no. and date 國際申請發表編號及日期
WO2014/055536 10.04.2014

CN Application no. & date 中國專利申請編號及日期
CN 201380051398.5 01.10.2013

CN Publication no. & date 中國專利申請發表編號及日期
CN 104685890 03.06.2015

Date of grant in designated patent office 指定專利當局批予專利日期
12.03.2019

[73] Proprietor 專利所有人
QUALCOMM INCORPORATED
高通股份有限公司
5775 Morehouse Drive
San Diego, California 92121-1714
UNITED STATES OF AMERICA

[72] Inventor 發明人
WANG, Ye-Kui 王益魁

[74] Agent and / or address for service 代理人及/或送達地址
LI Wing Hong
20/F., Central Tower
28 Queen's Road Central
HONG KONG



(12)发明专利

(10)授权公告号 CN 104685890 B

(45)授权公告日 2019.03.12

(21)申请号 201380051398.5

(22)申请日 2013.10.01

(65)同一申请的已公布的文献号

申请公布号 CN 104685890 A

(43)申请公布日 2015.06.03

(30)优先权数据

61/709,094 2012.10.02 US

14/042,416 2013.09.30 US

(85)PCT国际申请进入国家阶段日

2015.03.31

(86)PCT国际申请的申请数据

PCT/US2013/062890 2013.10.01

(87)PCT国际申请的公布数据

W02014/055536 EN 2014.04.10

(73)专利权人 高通股份有限公司

地址 美国加利福尼亚州

(72)发明人 王益魁

(74)专利代理机构 北京律盟知识产权代理有限公司 11287

代理人 宋献涛

(51)Int.Cl.

H04N 19/70(2014.01)

H04N 19/30(2014.01)

(56)对比文件

US 2012183077 A1,2012.07.19,

CN 101491079 A,2009.07.22,

CN 101690220 A,2010.03.31,

CN 101523920 A,2009.09.02,

审查员 高静

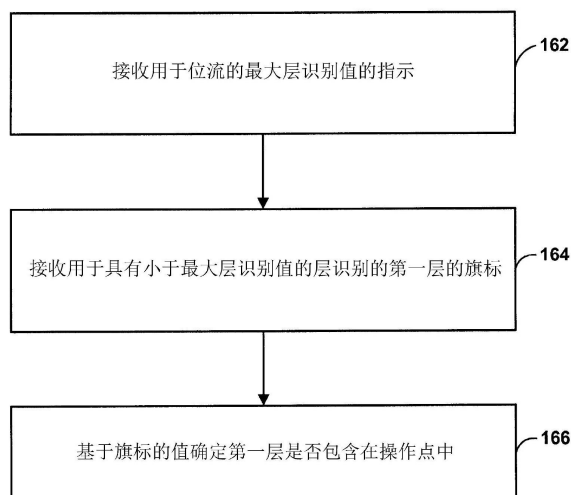
权利要求书5页 说明书24页 附图5页

(54)发明名称

用于处理和编码多层视频数据的方法、装置和设备

(57)摘要

本发明提供一种用于处理视频数据的装置,所述装置接收用于位流的最大层识别ID值的指示;接收用于具有小于所述最大层ID值的层ID值的第一层的旗标;及,基于所述旗标的值确定所述第一层是否包含在操作点中。



1. 一种处理多层视频数据的方法,所述方法包括:

在视频参数集中接收语法元素,所述语法元素对关联于第一层集合的任一NAL单元指定最大层识别ID值,其中所述多层视频数据包括两个或更多个层集合,其中所述两个或更多个层集合的每一者关联于所述多层视频数据的至少一个可独立解码的位流且包括两个或两个以上层;

对所述至少一个层集合的所述第一层集合,在所述视频参数集中接收用于具有小于所述最大层ID值的层ID值的第一层的含一位层ID旗标;

基于所述含一位层ID旗标的值确定所述第一层包含在所述第一层集合中;

接收多个NAL单元,所述多个NAL单元的各个NAL单元的每一者具有NAL单元标头,所述NAL单元标头包含层ID语法元素,所述层ID语法元素指定所述各个NAL单元所属的层;以及

基于对应于所述第一层的所述NAL单元的层ID语法元素而确定所述多个NAL单元的NAL单元被包含于所述第一层中。

2. 根据权利要求1所述的方法,其进一步包括:

在所述视频参数集中接收用于与关联于所述第一层的所述层ID不同的、介于1与所述最大层ID值之间的每一层ID值的含额外一位层ID旗标,其中用于每一含额外一位层ID旗标的值指示与各个层ID值相关联的层是否包含在所述第一层集合中。

3. 根据权利要求1所述的方法,其中用于所述含一位层ID旗标的第一值指示所述第一层包含在所述第一层集合中,且用于所述含一位层ID旗标的第二值指示所述第一层不包含在所述第一层集合中。

4. 根据权利要求1所述的方法,其中所述第一层集合包括操作点。

5. 根据权利要求1所述的方法,其中所述方法是由媒体感知网络元件MANE执行,且其中所述方法进一步包括:

响应于所述第一层包含在所述第一层集合中而从所述多个NAL单元中识别具有小于所述最大层ID值的所述层ID值的一个或多个NAL单元;

从所述多层视频数据中提取具有小于所述最大层ID值的所述层ID值的所述一个或多个NAL单元。

6. 根据权利要求1所述的方法,其中所述视频参数集包括适用于多个层的语法。

7. 根据权利要求1所述的方法,其中指定所述最大层ID值的所述语法元素以及所述含一位层ID旗标各为固定长度且非经熵编码的语法元素。

8. 一种用于处理多层视频数据的装置,所述装置包括:

存储器,所述存储器经配置以存储所述多层视频数据;及

一或多个处理器,所述一或多个处理器经配置以:

在视频参数集中接收语法元素,所述语法元素对关联于第一层集合的任一NAL单元指定最大层识别ID值,其中所述多层视频数据包括两个或更多个层集合,其中所述两个或更多个层集合的每一者关联于所述多层视频数据的至少一个可独立解码的位流且包括两个或两个以上层;

对所述至少一个层集合的所述第一层集合,在所述视频参数集中接收用于具有小于所述最大层ID值的层ID值的第一层的含一位层ID旗标;及

基于所述含一位层ID旗标的值确定所述第一层包含在所述第一层集合中;

接收多个NAL单元,所述多个NAL单元的各个NAL单元的每一者具有NAL单元标头,所述NAL单元标头包含层ID语法元素,所述层ID语法元素指定所述各个NAL单元所属的层;以及

基于对应于所述第一层的所述NAL单元的层ID语法元素而确定所述多个NAL单元的NAL单元被包含于所述第一层中。

9.根据权利要求8所述的装置,其中所述一或多个处理器经进一步配置以在所述视频参数集中接收用于介于1与所述最大层ID值之间的、与关联于所述第一层的所述层ID不同的每一层ID值的含额外一位层ID旗标,其中用于每一含额外一位层ID旗标的值指示与各个层ID值相关联的层是否包含在所述第一层集合中。

10.根据权利要求8所述的装置,其中用于所述含一位层ID旗标的第一值指示所述第一层包含在所述第一层集合中,且用于所述含一位层ID旗标的第二值指示所述第一层不包含在所述第一层集合中。

11.根据权利要求8所述的装置,其中所述第一层集合包括操作点。

12.根据权利要求8所述的装置,其中所述装置包括媒体感知网络元件MANE,且其中所述一或多个处理器经进一步配置以:

响应于所述第一层包含在所述第一层集合中而从所述多个NAL单元中识别具有小于所述最大层ID值的所述层ID值的一个或多个NAL单元;

从所述多层视频数据中提取具有小于所述最大层ID值的所述层ID值的所述一个或多个NAL单元。

13.根据权利要求8所述的装置,其中所述装置包括以下各者中的至少一者:

集成电路;

微处理器;或

无线通信装置。

14.根据权利要求8所述的装置,其中所述视频参数集包括适用于多个层的语法。

15.根据权利要求8所述的装置,其中指定所述最大层ID值的所述语法元素以及所述含一位层ID旗标各为固定长度且非经熵编码的语法元素。

16.一种用于处理多层视频数据的设备,所述设备包括:

用于在视频参数集中接收语法元素,所述语法元素对关联于第一层集合的任一NAL单元指定最大层识别ID值的装置,其中所述多层视频数据包括两个或更多个层集合,其中所述两个或更多个层集合的每一者关联于所述多层视频数据的至少一个可独立解码的位流且包括两个或两个以上层;

用于对所述至少一个层集合的所述第一层集合,在所述视频参数集中接收用于具有小于所述最大层ID值的层ID值的第一层的含一位层ID旗标的装置;

用于基于所述含一位层ID旗标的值确定所述第一层包含在所述第一层集合中的装置;

用于接收多个NAL单元的装置,所述多个NAL单元的各个NAL单元的每一者具有NAL单元标头,所述NAL单元标头包含层ID语法元素,所述层ID语法元素指定所述各个NAL单元所属的层;以及

用于基于对应于所述第一层的所述NAL单元的层ID语法元素而确定所述多个NAL单元的NAL单元被包含于所述第一层中的装置。

17.根据权利要求16所述的设备,其进一步包括:

用于在所述视频参数集中接收用于与关联于所述第一层的所述层ID不同的、介于1与所述最大层ID值之间的每一层ID值的含额外一位层ID旗标的装置,其中用于每一含额外一位层ID旗标的值指示与各个层ID值相关联的层是否包含在所述第一层集合中。

18.根据权利要求16所述的设备,其中用于所述含一位层ID旗标的第一值指示所述第一层包含在所述第一层集合中,且用于所述含一位层ID旗标的第二值指示所述第一层不包含在所述第一层集合中。

19.根据权利要求16所述的设备,其中所述第一层集合包括操作点。

20.根据权利要求16所述的设备,其中所述设备包括媒体感知网络元件MANE,且其中所述设备进一步包括用于从所述多个NAL单元中识别具有小于所述最大层ID值的所述层ID值的一个或多个NAL单元;以及用于从所述多层视频数据中提取具有小于所述最大层ID值的所述层ID值的所述一个或多个NAL单元的装置。

21.根据权利要求16所述的设备,其中所述视频参数集包括适用于多个层的语法。

22.根据权利要求16所述的设备,其中指定所述最大层ID值的所述语法元素以及所述含一位层ID旗标各为固定长度且非经熵编码的语法元素。

23.一种用于处理多层视频数据的非易失性计算机可读存储媒体,其存储指令,所述指令在由一或多个处理器执行时致使所述一或多个处理器进行以下操作:

在视频参数集中接收语法元素,所述语法元素对关联于第一层集合的任一NAL单元指定最大层识别ID值,其中所述多层视频数据包括两个或更多个层集合,其中所述两个或更多个层集合的每一者关联于所述多层视频数据的至少一个可独立解码的位流且包括两个或两个以上层;

对所述至少一个层集合的所述第一层集合,在所述视频参数集中接收用于具有小于所述最大层ID值的层ID值的第一层的含一位层ID旗标;

基于所述含一位层ID旗标的值确定所述第一层包含在所述第一层集合中;

接收多个NAL单元,所述多个NAL单元的各个NAL单元的每一者具有NAL单元标头,所述NAL单元标头包含层ID语法元素,所述层ID语法元素指定所述各个NAL单元所属的层;以及

基于对应于所述第一层的所述NAL单元的层ID语法元素而确定所述多个NAL单元的NAL单元被包含于所述第一层中。

24.根据权利要求23所述的非易失性计算机可读存储媒体,其存储其它指令,所述其它指令经配置以致使所述一或多个处理器进行以下操作:

在所述视频参数集中接收用于与关联于所述第一层的所述层ID不同的、介于1与所述最大层ID值之间的每一层ID值的含额外一位层ID旗标,其中用于每一含额外一位层ID旗标的值指示与各个层ID值相关联的层是否包含在所述第一层集合中。

25.根据权利要求23所述的非易失性计算机可读存储媒体,其中用于所述含一位层ID旗标的第一值指示所述第一层包含在所述第一层集合中,且用于所述含一位层ID旗标的第二值指示所述第一层不包含在所述第一层集合中。

26.根据权利要求23所述的非易失性计算机可读存储媒体,其中所述第一层集合包括操作点。

27.根据权利要求23所述的非易失性计算机可读存储媒体,其存储其它指令,所述其它指令致使所述一或多个处理器进行以下操作:

响应于所述第一层包含在所述第一层集合中而识别具有小于所述最大层ID值的所述层ID值的一个或多个NAL单元；

从所述多层视频数据中提取具有小于所述最大层ID值的所述层ID值的所述一个或多个NAL单元。

28. 根据权利要求23所述的非易失性计算机可读存储媒体，其中所述视频参数集包括适用于多个层的语法。

29. 根据权利要求23所述的非易失性计算机可读存储媒体，其中指定所述最大层ID值的所述语法元素以及所述含一位层ID旗标各为固定长度且非经熵编码的语法元素。

30. 一种编码多层视频数据的方法，所述方法包括：

针对所述多层视频数据确定层集合，其中所述层集合对应于至少一个所述多层视频数据的可独立解码的位流且包括两个或两个以上层；

产生语法元素以包含于视频参数集中，所述语法元素对关联于第一层集合的任一NAL单元指定最大层识别ID值，其中所述多层视频数据包括两个或更多个层集合，其中所述两个或更多个层集合的每一者关联于所述多层视频数据的至少一个可独立解码的位流且包括两个或两个以上层；

产生用于具有小于所述最大层ID值的层ID值的第一层的含一位层ID旗标以包含于所述视频参数集中，其中用于所述含一位层ID旗标的值发信表示所述第一层是否包含在所述第一层集合中；以及

输出包含所述多层视频数据的经编码位流。

31. 根据权利要求30所述的方法，其进一步包括：

产生用于介于1与所述最大层ID值之间的每一层ID值的含额外一位层ID旗标以包含于所述视频参数集中，其中用于每一含额外一位层ID旗标的值指示与各个层ID值相关联的层是否包含在所述层集合中。

32. 根据权利要求30所述的方法，其中用于所述含一位层ID旗标的第一值指示所述第一层包含在所述层集合中，且用于所述含一位层ID旗标的第二值指示所述第一层不包含在所述层集合中。

33. 根据权利要求30所述的方法，其中所述第一层集合包括操作点。

34. 根据权利要求30所述的方法，其中所述方法是由视频编码器执行。

35. 根据权利要求30所述的方法，其中所述视频参数集包括适用于多个层的语法。

36. 根据权利要求30所述的方法，其中指定所述最大层ID值的所述语法元素以及所述含一位层ID旗标各为固定长度且非经熵编码的语法元素。

37. 一种用于处理多层视频数据的视频译码装置，所述装置包括：

存储器，所述存储器经配置以存储所述多层视频数据；及

一或多个处理器，所述一或多个处理器经配置以：

针对所述多层视频数据确定层集合，其中所述层集合对应于至少一个包括所述多层视频数据的可独立解码的位流且包括两个或两个以上层；

产生语法元素以包含于视频参数集中，所述语法元素对关联于第一层集合的任一NAL单元指定最大层识别ID值，其中所述多层视频数据包括两个或更多个层集合，其中所述两个或更多个层集合的每一者关联于所述多层视频数据的至少一个可独立解码的位流且包

括两个或两个以上层;且,

产生用于具有小于所述最大层ID值的层ID值的第一层的含一位层ID旗标以包含于所述视频参数集中,其中用于所述含一位层ID旗标的值发信表示所述第一层是否包含在所述第一层集合中;以及

输出包含所述多层视频数据的经编码位流。

38.根据权利要求37所述的视频译码装置,其中所述一或多个处理器经进一步配置以产生用于介于1与所述最大层ID值之间的每一层ID值的含额外一位层ID旗标以包含于所述视频参数集中,其中用于每一含额外一位层ID旗标的值指示与各个层ID值相关联的层是否包含在所述层集合中。

39.根据权利要求37所述的视频译码装置,其中用于所述含一位层ID旗标的第一值指示所述第一层包含在所述层集合中,且用于所述含一位层ID旗标的第二值指示所述第一层不包含在所述层集合中。

40.根据权利要求37所述的视频译码装置,其中所述第一层集合包括操作点。

41.根据权利要求37所述的视频译码装置,其中所述装置包括视频编码器。

42.根据权利要求37所述的视频译码装置,其中所述视频参数集包括适用于多个层的语法。

43.根据权利要求37所述的视频译码装置,其中指定所述最大层ID值的所述语法元素以及所述含一位层ID旗标均为固定长度且非经熵编码的语法元素。

用于处理和编码多层视频数据的方法、装置和设备

[0001] 本申请案主张2012年10月2日申请的第61/709,094号美国临时专利申请案的权利,所述临时专利申请案的全部内容以引用的方式并入本文中。

技术领域

[0002] 本发明一般来说涉及处理视频数据且,更明确地说,涉及处理用于视频数据中的操作点。

背景技术

[0003] 数字视频能力可并入到广泛范围的装置中,所述装置包含数字电视、数字直播系统、无线广播系统、个人数字助理(PDA)、膝上型或桌上型计算机、平板计算机、电子书阅读器、数码相机、数字记录装置、数字媒体播放器、视频游戏装置、视频游戏控制台、蜂窝式或卫星无线电电话、所谓“智能型手机”、视频电话会议装置、视频流装置,及其类似者。数字视频装置实施视频译码技术,例如,由MPEG-2、MPEG-4、ITU-T H.263、ITU-T H.264/MPEG-4第10部分高级视频译码(AVC)定义的标准、目前正在开发中的高效率视频译码(HEVC)标准及此类标准的扩展中描述的视频译码技术。视频装置可通过实施此类视频译码技术来更有效率地发射、接收、编码、解码及/或存储数字视频信息。

[0004] 视频译码技术包含空间(图片内)预测及/或时间(图片间)预测以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,可将视频切片(例如,视频帧或视频帧的部分)分割成视频块,所述视频块还可被称作树块、译码单元(CU)及/或译码节点。图片的经帧内译码(I)切片中的视频块是使用相对于同一图片中的相邻块中的参考样本的空间预测而编码。图片的经帧间译码(P或B)切片中的视频块可使用相对于同一图片中的相邻块中的参考样本的空间预测或相对于其它参考图片中的参考样本的时间预测。可将图片称作帧,且可将参考图片称作参考帧。

[0005] 空间预测或时间预测导致译码块的预测性块。残余数据表示原始待译码块与预测性块之间的像素差。经帧间译码块是根据指向形成预测性块的参考样本的块的运动向量及指示经译码块与预测性块之间的差的残余数据而编码。经帧内译码块是根据帧内译码模式及残余数据而编码。为了进行进一步压缩,可将残余数据从像素域变换到变换域,从而产生残余变换系数,接着可将所述残余变换系数量化。可扫描最初布置成二维阵列的经量化的变换系数以便产生变换系数的一维向量,且可应用熵译码以实现甚至更多压缩。

发明内容

[0006] 一般来说,本发明描述用于在视频译码中发信用于操作点的层识别符的技术。

[0007] 在一个实例中,一种处理视频数据的方法包含:接收用于位流的最大层识别(ID)值的指示;接收用于具有小于所述最大层ID值的层ID值的第一层的旗标;及,基于所述旗标的值确定所述第一层是否包含在操作点中。

[0008] 在另一实例中,一种用于处理视频数据的装置包含一或多个处理器,所述一或多

个处理器经配置以：接收用于位流的最大层识别 (ID) 值的指示；接收用于具有小于所述最大层ID值的层ID值的第一层的旗标；及，基于所述旗标的值确定所述第一层是否包含在操作点中。

[0009] 在另一实例中，一种用于处理视频数据的设备包含：用于接收用于位流的最大层识别 (ID) 值的指示的装置；用于接收用于具有小于所述最大层ID值的层ID值的第一层的旗标的装置；及用于基于所述旗标的值确定所述第一层是否包含在操作点中的装置。

[0010] 在另一实例中，一种非暂时性计算机可读存储媒体存储指令，所述指令在由一或多个处理器执行时致使所述一或多个处理器进行以下操作：接收用于位流的最大层识别 (ID) 值的指示；接收用于具有小于所述最大层ID值的层ID值的第一层的旗标；及基于所述旗标的值确定所述第一层是否包含在操作点中。

[0011] 在另一实例中，一种编码视频数据的方法包含：产生用于位流的最大层识别 (ID) 值的指示；及产生用于具有小于所述最大层ID值的层ID值的第一层的旗标，其中用于所述旗标的值发信表示所述第一层是否包含在操作点中。

[0012] 在另一实例中，一种视频译码装置包含视频编码器，所述视频编码器经配置以：产生用于位流的最大层识别 (ID) 值的指示；及产生用于具有小于所述最大层ID值的层ID值的第一层的旗标，其中用于所述旗标的值发信表示所述第一层是否包含在操作点中。

[0013] 在另一实例中，一种用于视频译码的设备包含：用于产生用于位流的最大层识别 (ID) 值的指示的装置；及用于产生用于具有小于所述最大层ID值的层ID值的第一层的旗标的装置，其中用于所述旗标的值发信表示所述第一层是否包含在操作点中。

[0014] 在另一实例中，一种非暂时性计算机可读媒体存储指令，所述指令在由一或多个处理器执行时致使所述一或多个处理器进行以下操作：产生用于位流的最大层识别 (ID) 值的指示；及产生用于具有小于所述最大层ID值的层ID值的第一层的旗标，其中用于所述旗标的值发信表示所述第一层是否包含在操作点中。

[0015] 一或多个实例的细节陈述于随附图式及以下描述中。其它特征、目标及优点将从所述描述及所述图式以及从权利要求书变得显而易见。

附图说明

[0016] 图1为说明可利用本发明所描述的技术的实例视频编码及解码系统的框图。

[0017] 图2为说明可实施本发明所描述的技术的实例视频编码器的框图。

[0018] 图3为说明可实施本发明所描述的技术的实例视频解码器的框图。

[0019] 图4为说明形成网络的部分的实例装置集合的框图。

[0020] 图5为根据本发明所描述的技术的流程图，其说明用于编码视频数据的实例方法。

[0021] 图6为根据本发明所描述的技术的流程图，其说明用于处理视频数据的实例方法。

具体实施方式

[0022] 本发明涉及视频译码 (即，视频数据的编码及/或解码) 及视频处理，且更明确地说，涉及用于视频处理中的操作点语法。一般来说，本发明描述用于在视频译码中发信用于操作点的层识别符的技术。操作点是指可从在时间上可按比例调整及/或具有多个层或视图的原始位流提取的子位流。可基于识别位流的操作点的层识别符 (即，层ID) 及时间子层

识别符(即,时间ID)的值从所述位流提取所述子位流。通常,本发明使用术语“层识别符”及“层ID”来指空间层及/或视图的识别符,而术语“时间子层识别符”及“时间ID”是指时间子层的识别符。

[0023] 举例来说,可在位流内的参数集(例如,视频参数集(VPS))中发信操作点。对于所述操作点中的每一者,由(例如)视频编码器产生的操作点语法结构指定用以识别位流中属于给定操作点的子位流的网络抽象层(NAL)单元的层识别符集合。以此方式,网络物理(例如,媒体感知网络物理(MANE))可剖析NAL单元标头以从原始位流提取构成给定操作点的子位流的NAL单元。位流中的每一NAL单元可包含一层ID及一时间ID,且通过剖析层ID及时间ID,MANE可识别用于特定操作点的NAL单元。

[0024] 本发明的技术可通过改善用于操作点的层ID的发信而改善与操作点相关联的信息的发信效率。根据下文将更详细地解释的本发明的一种实例技术,可发信用于最大层ID的层识别值(即,层ID),且可以一系列旗标的形式发信额外层ID的存在。举例来说,假设位流包含具有各种时间及空间分辨率的六个子流,其中每一子流具有一层ID。可在位流中发信最大层ID值。出于此实例的目的,假设最大层ID值为9,此情形意味着潜在地存在可包含在操作点中的十个层,其中层ID为0到9。可使用九个旗标来发信用于操作点的剩余层ID值,其中第一旗标指示是否存在层ID值0,第二旗标指示是否存在层ID值1,等等,直到指示是否存在层ID值8的最后旗标。因此,为了发信层ID值2、5及9,可首先发信值9,后接续发信旗标序列001001000,其中用于第三位的1指示存在层ID值2,且用于第六位的1指示存在层ID值5。本发明还将论述用于发信层ID的其它技术。

[0025] 本发明一般将使用术语“视频译码”来指视频编码或视频解码。本发明还使用术语“视频处理”,所述视频处理一般意味着包含视频译码,但还意味着包含其它类型的视频处理,例如视频数据剖析、视频数据投送、视频位流剪接及其它此类过程。视频译码器一般可被视为指编码及/或解码视频数据的装置,而视频处理器或视频处理装置可被视为指译码视频数据的装置,但还指对视频数据执行其它过程的装置。

[0026] 图1为说明实例视频编码及解码系统10的框图,视频编码及解码系统10可利用本发明所描述的用于发信层ID的技术。如图1所展示,系统10包含源装置12,源装置12产生待由目的地装置14在稍后时间解码的经编码视频数据。可经由网络装置13将经编码视频数据从源装置12投送到目的地装置14,网络装置13可为网络装置的较大网络的部分。源装置12及目的地装置14可包括广泛范围的装置中的任一者,所述装置包含桌上型计算机、笔记型(即,膝上型)计算机、平板计算机、机顶盒、例如所谓“智能型”手机的电话手持机、所谓“智能型”平板计算机、电视、摄像机、显示装置、数字媒体播放器、视频游戏控制台、视频流装置或其类似者。在一些情况下,可装备源装置12及目的地装置14以用于实现无线通信。

[0027] 在图1的实例中,源装置12包含视频源18、视频编码器20及输出接口22。目的地装置14包含输入接口28、视频解码器30及显示装置32。在其它实例中,源装置12及目的地装置14可包含其它组件或布置。举例来说,源装置12可从外部视频源18(例如,外部摄像机)接收视频数据。同样地,目的地装置14可与外部显示装置介接,而不是包含集成式显示装置。

[0028] 图1的所说明系统10仅为一个实例。可通过任何数字视频译码及/或处理装置来执行本发明的技术。虽然一般通过视频编码装置或视频解码装置来执行所述技术,但还可通过视频编码器/解码器(通常称作“编码解码器(CODEC)”)来执行所述技术。此外,还可通过

视频预处理器来执行本发明的技术。源装置12及目的地装置14仅为其中源装置12产生用于发射到目的地装置14的经译码视频数据的此类译码装置的实例。在一些实例中,装置12、14可以实质上对称方式操作,使得装置12、14中的每一者包含视频编码及解码组件。因此,系统10可支持视频装置12、14之间的单向或双向视频发射,例如,用于视频流式传输、视频播放、视频广播或视频电话。

[0029] 在一个实例中,源装置12的视频编码器20可产生视频数据的位流。视频数据的VPS可定义对应于位流的子位流的多个操作点。视频编码器20可包含识别待包含在特定操作点中的层及时间子层的操作点语法。VPS中的操作点语法可包含用于操作点的最大层ID值的指示以及一或多个旗标。所述旗标指示具有小于最大层ID的层ID的层是否包含在操作点中。因此,在接收到具有最大层ID的VPS及旗标后,网络装置13即可识别用于操作点的NAL单元且将那些NAL单元投送到目的地装置14。在接收到NAL单元后,目的地装置14的视频解码器30即可解码经编码视频数据。视频解码器30可潜在地以与网络装置13相同的方式来剖析VPS中含有的操作点语法。举例来说,视频解码器30可剖析操作点语法以检查是否接收到所有预期层或确定待应用的假想参考解码器(HRD)参数集合。

[0030] 源装置12的视频源18可包含例如视频摄像机等视频俘获装置、含有先前所俘获视频的视频存档,及/或用以从视频内容提供者接收视频的视频馈送接口。作为另一替代例,视频源18可产生基于计算机图形的数据作为源视频,或直播视频、存档视频及计算机产生的视频的组合。在一些情况下,如果视频源18为视频摄像机,那么源装置12及目的地装置14可形成所谓的摄像机电话或视频电话。然而,如上文所提及,本发明所描述的技术一般来说可适用于视频译码,且可应用于无线及/或有线应用。

[0031] 在每一种情况下,可通过视频编码器20来编码所俘获、预先所俘获或计算机产生的视频。经编码视频数据可经由源装置12的输出接口22而直接发射到目的地装置14。经编码视频数据还可(或替代地)存储到存储装置上以供稍后由目的地装置14或其它装置存取,以用于解码及/或播放。

[0032] 链路16可包含:暂态媒体,例如无线广播或有线网络发射;或存储媒体(即,非暂时性存储媒体),例如硬盘、闪存盘、压缩光盘、数字视频光盘、蓝光光盘或其它计算机可读媒体。在一些实例中,网络服务器可从源装置12接收经编码视频数据且(例如)经由网络发射将所述经编码视频数据提供到目的地装置14。类似地,媒体生产设施(例如,光盘压印设施)的计算装置可从源装置12接收经编码视频数据且产生含有所述经编码视频数据的光盘。因此,在各种实例中,可将链路16理解为包含各种形式的一或多个计算机可读媒体。链路16可包括能够将经编码视频数据从源装置12移动到目的地装置14的任何类型的媒体或装置。在一个实例中,链路16可包括用以使得源装置12能够实时地直接将经编码视频数据发射到目的地装置14的通信媒体。经编码视频数据可根据通信标准(例如,无线通信协议)进行调制,且被发射到目的地装置14。通信媒体可包括任何无线或有线通信媒体,例如,射频(RF)频谱或一或多个物理传输线。通信媒体可形成基于包的的网络(例如,局域网、广域网或例如因特网等全域网络)的部分。通信媒体可包含路由器、交换器、基站,或可用以促进从源装置12到目的地装置14的通信的任何其它装备。

[0033] 目的地装置14的输入接口28从链路16接收信息,链路16可为计算机可读媒体。来自链路16的信息可包含由视频编码器20定义的语法信息(其也供视频解码器30使用),所述

语法信息包含描述块及其它经译码单元(例如,GOP)的特性及/或处理的语法元素。显示装置32可与目的地装置14集成或位于目的地装置14外部。显示装置32向用户显示经解码视频数据,且可包括例如以下各者等多种显示装置中的任一者:阴极射线管(CRT)、液晶显示器(LCD)、等离子显示器、有机发光二极管(OLED)显示器或另一类型的显示装置。

[0034] 或者,经编码数据可从输出接口22输出到存储装置34。类似地,可通过输入接口从存储装置34存取经编码数据。存储装置34可包含多种分散式或本地存取式数据存储媒体中的任一者,例如硬盘、蓝光光盘、DVD、CD-ROM、快闪存储器、易失性或非易失性存储器,或用于存储经编码视频数据的任何其它合适的数字存储媒体。在另一实例中,存储装置34可对应于可保持由源装置12产生的经编码视频的文件服务器或另一中间存储装置。目的地装置14可经由流式传输或下载从存储装置34存取所存储的视频数据。文件服务器可为能够存储经编码视频数据且将所述经编码视频数据发射到目的地装置14的任何类型的服务器。实例文件服务器包含web服务器(例如,用于网站)、FTP服务器、网络附接存储(NAS)装置或本地磁盘驱动器。目的地装置14可经由任何标准数据连接(包含因特网连接)存取经编码视频数据。此数据连接可包含适合于存取存储在文件服务器上的经编码视频数据的无线信道(例如,Wi-Fi连接)、有线连接(例如,DSL、电缆调制解调器,等等),或两者的组合。经编码视频数据从存储装置34的传输可为流式传输、下载传输或两者的组合。

[0035] 本发明的技术未必限于无线应用或设置。可将所述技术应用于支持多种多媒体应用(例如,(例如)经由因特网的空中电视广播、有线电视发射、卫星电视发射、流式视频发射)中的任一者的视频译码、用于存储在数据存储媒体上的数字视频的编码、存储在数据存储媒体上的数字视频的解码,或其它应用。在一些实例中,系统10可经配置以支持单向或双向视频发射以支持例如视频流式传输、视频播放、视频广播及/或视频电话等应用。

[0036] 视频编码器20及视频解码器30各自可实施为例如以下各者等多种合适的编码器或解码器电路系统中的任一者:一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程阵列(FPGA)、离散逻辑电路系统、软件、硬件、固件或其任何组合。当所述技术部分地在软件中实施时,装置可将用于软件的指令存储在合适的非暂时性计算机可读媒体中,且在硬件中使用一或多个处理器来执行所述指令以执行本发明的技术。视频编码器20及视频解码器30中的每一者可包含在一或多个编码器或解码器中,其中任一者可集成为相应装置中的组合式编码器/解码器(编码解码器)的部分。包含视频编码器20及/或视频解码器30的装置可包括集成电路、微处理器及/或无线通信装置(例如,蜂窝式电话)。

[0037] 视频编码器20及视频解码器30可根据视频译码标准(例如,目前正在开发中的HEVC标准)操作,且可符合HEVC测试模型(HM)。即将到来的HEVC标准的草案(被称作“HEVC工作草案8”或“WD8”)描述于布洛斯(Bross)等人的名为“高效率视频译码(HEVC)文本规范草案8(High efficiency video coding(HEVC)text specification draft 8)”的文档JCTVC-J1003_d7(ITU-T SG16 WP3及ISO/IEC JTC1/SC29/WG11的视频译码联合协作小组(JCT-VC),第10次会议:瑞典斯德哥尔摩,2012年7月)中。HEVC标准的工作草案8的全文借此以引用的方式并入本文中。HEVC标准的另一新近草案(被称作“HEVC工作草案10”或“WD10”)描述于布洛斯(Bross)等人的名为“高效率视频译码(HEVC)文本规范草案10(用于FDIS及Last Call)(High efficiency video coding(HEVC)text specification draft 10(for FDIS&Last Call))”的文档JCTVC-L1003v34(ITU-T SG16 WP3及ISO/IEC JTC1/SC29/WG11

的视频译码联合协作小组 (JCT-VC), 第12次会议: 瑞士日内瓦, 2013年1月14到23日) 中。WD10的全文借此以引用的方式并入本文中。

[0038] 替代地, 视频编码器20及视频解码器30可根据例如ITU-T H.264标准 (或者被称作MPEG-4第10部分高级视频译码 (AVC)) 等其它专属或工业标准或此类标准的扩展操作。然而, 本发明的技术并不限于任何特定译码标准。视频译码标准的其它实例包含MPEG-2及ITU-T H.263。在一些方面中, 视频编码器20及视频解码器30可各自与音频编码器及解码器集成, 且可包含适当的MUX-DEMUX单元或其它硬件及软件, 以处置共同数据流或单独数据流中的音频与视频两者的编码。在适用时, MUX-DEMUX单元可符合ITU H.223多路复用器协议或例如用户数据报协议 (UDP) 等其它协议。

[0039] ITU-T H.264/MPEG-4 (AVC) 标准是由ITU-T视频译码专家群 (VCEG) 连同ISO/IEC动画专家群 (MPEG) 一起制定作为被称为联合视频小组 (JVT) 的集体合作产品。在一些方面中, 可将本发明所描述的技术应用于一般符合H.264标准的装置。H.264标准描述于ITU-T研究群的日期为2005年3月的ITU-T推荐H.264 (用于一般视听服务的高级视频译码) 中, 所述标准在本文中可被称作H.264标准或H.264规范或H.264/AVC标准或规范。联合视频小组 (JVT) 继续致力于对H.264/MPEG-4AVC的扩展。

[0040] JCT-VC正致力于HEVC标准的开发。HEVC标准化努力是基于视频译码装置的演进模型, 其被称作HEVC测试模型 (HM)。HM假定视频译码装置相对于根据 (例如) ITU-T H.264/AVC的现有装置的若干额外能力。举例来说, 虽然H.264提供九种帧内预测编码模式, 但HM可提供多达三十三种帧内预测编码模式。

[0041] 一般来说, HM的工作模型描述视频帧或图片可划分成包含明度样本与色度样本两者的树块或最大译码单元 (LCU) 序列。位流内的语法数据可定义LCU的大小, LCU就像像素的数目来说为最大译码单元。切片包含按译码次序的数个连续树块。可将视频帧或图片分割成一或多个切片。每一树块可根据二叉树而分裂成若干译码单元 (CU)。一般来说, 二叉树数据结构包含每一CU一个节点, 其中根节点对应于树块。如果将CU分裂成四个子CU, 那么对应于所述CU的节点包含四个叶节点, 所述四个叶节点中的每一者对应于所述子CU中的一者。

[0042] 二叉树数据结构中的每一节点可提供针对对应CU的语法数据。举例来说, 二叉树中的节点可包含分裂旗标 (split flag), 所述分裂旗标指示是否将对应于节点的CU分裂成子CU。可递归地定义用于CU的语法元素, 且所述语法元素可取决于CU是否分裂成子CU。如果CU未被进一步分裂, 那么其被称作叶CU。在本发明中, 即使不存在原始叶CU的明确分裂, 叶CU的四个子CU也将被称作叶CU。举例来说, 如果 16×16 大小的CU未被进一步分裂, 那么四个 8×8 子CU也将被称作叶CU, 但所述 16×16 CU从未被分裂。

[0043] 除了CU不具有大小区别以外, CU具有类似于H.264标准的宏块的目的。举例来说, 树块可分裂成四个子节点 (还被称作子CU), 且每一子节点可又为父节点并分裂成另外四个子节点。被称作二叉树的叶节点的最终的未经分裂的子节点包括译码节点, 所述译码节点还被称作叶CU。与经译码位流相关联的语法数据可定义可分裂一树块的最大次数 (其被称作最大CU深度), 且还可定义译码节点的最小大小。因此, 位流还可定义最小译码单元 (SCU)。本发明使用术语“块”来指在HEVC的上下文中的CU、PU或TU中的任一者, 或在其它标准的上下文中的类似数据结构 (例如, 在H.264/AVC中的宏块及其子块)。

[0044] CU包含译码节点及与所述译码节点相关联的若干预测单元 (PU) 及变换单元 (TU)。

CU的大小对应于译码节点的大小,且形状必须为正方形。CU的大小的范围可从 8×8 像素直到具有最大 64×64 像素或大于 64×64 像素的树块的大小。每一CU可含有一或多个PU及一或多个TU。与CU相关联的语法数据可描述(例如)CU到一或多个PU的分割。分割模式可取决于CU被跳过或经直接模式编码、经帧内预测模式编码抑还是经帧间预测模式编码而不同。PU的形状可分割成非正方形。与CU相关联的语法数据还可描述(例如)CU根据四叉树到一或多个TU的分割。TU的形状可为正方形或非正方形(例如,矩形)。

[0045] HEVC标准允许根据TU进行的变换,所述变换可针对不同CU而不同。通常基于针对经分割的LCU定义的给定CU内的PU的大小设置TU大小,但并非总是如此情况。TU通常具有与PU相同的大小或小于PU。在一些实例中,可使用被称为“残余四叉树”(RQT)的四叉树结构将对应于CU的残余样本再分为更小的单元。RQT的叶节点可被称作变换单元(TU)。可变换与TU相关联的像素差值以产生可进行量化的变换系数。

[0046] 叶CU可包含一或多个预测单元(PU)。一般来说,PU表示对应于对应CU的全部或一部分的空间区域,且可包含用于检索PU的参考样本的数据。此外,PU包含与预测有关的数据。举例来说,当PU经帧内模式编码时,用于PU的数据可包含在残余四叉树(RQT)中,所述残余四叉树(RQT)可包含描述用于对应于PU的TU的帧内预测模式的数据。作为另一实例,当PU经帧间模式编码时,PU可包含定义所述PU的一或多个运动向量的数据。定义PU的运动向量的数据可描述(例如)运动向量的水平分量、运动向量的垂直分量、用于运动向量的分辨率(例如,四分之一像素精度或八分之一像素精度)、运动向量所指向的参考图片,及/或用于运动向量的参考图片列表(例如,列表0、列表1或列表C)。

[0047] 具有一或多个PU的叶CU还可包含一或多个变换单元(TU)。可使用RQT(还被称作TU四叉树结构)来指定变换单元,如上文所论述。举例来说,分裂旗标可指示叶CU是否分裂成四个变换单元。接着,每一变换单元可进一步分裂成其它子TU。当TU未被进一步分裂时,其可被称作叶TU。一般来说,对于帧内译码,属于叶CU的所有叶TU共享相同的帧内预测模式。即,一般应用相同的帧内预测模式来计算叶CU的所有TU的预测值。对于帧内译码,视频编码器可使用帧内预测模式将每一叶TU的残余值计算为CU的对应于TU的部分与原始块之间的差。TU未必限于PU的大小。因此,TU可能大于或小于PU。对于帧内译码,PU可与同一CU的对应叶TU并置。在一些实例中,叶TU的最大大小可对应于对应叶CU的大小。

[0048] 此外,叶CU的TU还可与被称作残余四叉树(RQT)的相应四叉树数据结构相关联。即,叶CU可包含指示如何将叶CU分割成TU的四叉树。TU四叉树的根节点一般对应于叶CU,而CU四叉树的根节点一般对应于树块(或LCU)。RQT的未被分裂的TU被称作叶TU。一般来说,除非另有提及,否则本发明分别使用术语“CU”及“TU”来指叶CU及叶TU。

[0049] 视频序列通常包含一系列视频帧或图片。图片群组(GOP)一般包括一系列的视频图片中的一或多者。GOP可在GOP的标头、图片中的一或多者的标头中或在别处包含描述包含在GOP中的图片数目的语法数据。图片的每一切片可包含描述所述相应切片的编码模式的切片语法数据。视频编码器通常对个别视频切片内的视频块操作,以便编码视频数据。视频块可对应于CU内的译码节点。视频块可具有固定或变化的大小,且可根据指定的译码标准而在大小方面不同。

[0050] 作为实例,HM支持以各种PU大小进行的预测。假设特定CU的大小为 $2N \times 2N$,那么HM支持以 $2N \times 2N$ 或 $N \times N$ 的PU大小进行的帧内预测,及以 $2N \times 2N$ 、 $2N \times N$ 、 $N \times 2N$ 或 $N \times N$ 的对称PU

大小进行的帧间预测。HM还支持以 $2N \times nU$ 、 $2N \times nD$ 、 $nL \times 2N$ 及 $nR \times 2N$ 的PU大小进行的帧间预测的不对称分割。在不对称分割中，CU的一个方向未被分割，而另一方向分割成25%及75%。CU的对应于25%分区的部分通过“n”后接续“上”、“下”、“左”或“右”的指示来指示。因此，例如，“ $2N \times nU$ ”是指在水平方向上以顶部 $2N \times 0.5N$ PU及底部 $2N \times 1.5N$ PU分割的 $2N \times 2N$ CU。

[0051] 在本发明中，“ $N \times N$ ”与“N乘N”可互换地使用以指视频块在垂直维度与水平维度方面的像素尺寸，例如， 16×16 像素或16乘16像素。一般来说， 16×16 块在垂直方向上将具有16个像素($y=16$)且在水平方向上将具有16个像素($x=16$)。同样地， $N \times N$ 块一般在垂直方向上将具有N个像素，且在水平方向上将具有N个像素，其中N表示非负整数值。可按行及列来布置块中的像素。此外，块未必需要在水平方向上与在垂直方向上具有相同数目个像素。举例来说，块可包括 $N \times M$ 个像素，其中M未必等于N。

[0052] 在使用CU的PU进行帧内预测性或帧间预测性译码之后，视频编码器20可计算用于CU的TU的残余数据。所述PU可包括描述一种在空间域(还称作像素域)中产生预测性像素数据的方法或模式的语法数据且所述TU可包括在将变换(例如，离散余弦变换(DCT)、整数变换、小波变换或概念上类似变换)应用于残余视频数据之后在变换域中的系数。残余数据可对应于未经编码图片的像素与对应于PU的预测值之间的像素差。视频编码器20可形成包含CU的残余数据的TU，且接着变换所述TU以产生用于CU的变换系数。

[0053] 在进行任何变换以产生变换系数之后，视频编码器20可执行变换系数的量化。量化一般是指如下过程：其中，将变换系数量化以可能地减少用以表示所述系数的数据的量，从而提供进一步压缩。所述量化过程可减少与所述系数中的一些或全部相关联的位深度。举例来说，可在量化期间将n位值降值舍位到m位值，其中n大于m。

[0054] 在量化之后，视频编码器可扫描变换系数，从而从包含经量化的变换系数的二维矩阵产生一维向量。所述扫描可经设计成将较高能量(及因此较低频率)系数放置在阵列前方且将较低能量(及因此较高频率)系数放置在阵列后方。在一些实例中，视频编码器20可利用预定义扫描次序来扫描经量化的变换系数，以产生可经熵编码的串行化向量。在其它实例中，视频编码器20可执行自适应性扫描。在扫描经量化的变换系数以形成一维向量之后，视频编码器20可(例如)根据上下文自适应性可变长度译码(CAVLC)、上下文自适应性二进制算术译码(CABAC)、基于语法的上下文自适应性二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码方法而熵编码所述一维向量。视频编码器20还可熵编码与经编码视频数据相关联的语法元素以供视频解码器30用于解码视频数据。

[0055] 为了执行CABAC，视频编码器20可将上下文模型内的上下文指配给待发射符号。所述上下文可能关于(例如)符号的相邻值是否为非零。为了执行CAVLC，视频编码器20可针对待发射符号选择可变长度码。可将VLC中的码字构造成使得相对较短码对应于更有可能的符号，而较长码对应于较不可能的符号。以此方式，相比于(例如)针对待发射的每一符号使用等长度码字来说，使用VLC可实现位节省。概率确定可基于指配给符号的上下文而进行。

[0056] 视频编码器20可进一步将语法数据(例如，基于块的语法数据、基于帧的语法数据，及基于GOP的语法数据)发送到视频解码器30(例如，在帧标头、块标头、切片标头或GOP标头中)。GOP语法数据可描述相应GOP中的帧的数目，且帧语法数据可指示用以编码对应帧的编码/预测模式。

[0057] HEVC经设计为在其希望服务(例如,兼容于)广泛范围的应用、位速率、分辨率、质量及服务的意义上为通用的。其中,由HEVC潜在地服务的应用尤其包含数字存储媒体、电视广播及实时通信。在建立HEVC的过程中,已考虑了来自典型应用的各种要求,已开发了必要的算法元素,且已将这些集成到单个语法中。因此,HEVC应促进在不同应用当中的视频数据互换。然而,考虑到实施HEVC的完整语法的实践性,还借助于“设置档(profile)”及“层级(level)”来约定语法的有限数目的子集。

[0058] 将“设置档”定义为由HEVC指定的完整位流语法的子集。在通过给定设置档的语法强加的界限内,仍有可能取决于位流中的语法元素所采取的值(例如,经解码图片的指定大小)而需要在编码器及解码器的性能方面非常大的变化。在许多应用中,当前实施能够处置特定设置档内的语法的所有假设性使用的解码器既不实际也不经济。

[0059] 为了处置此问题,在每一设置档内指定“阶层(tier)”及“层级”。阶层的层级是强加于位流中的语法元素的值的指定约束集合。这些约束可为对值的简单限制。替代地,其可采取对值的算术组合的约束的形式(例如,图片宽度乘以图片高度乘以每秒被解码的图片的数目)。针对较低阶层指定的层级比针对较高阶层指定的层级更受约束。针对所有设置档来定义相同层级集合,其中每一层级的定义的大多数方面跨越不同设置档而为共同的。在指定约束内,个别实施方案可支持每一受支持设置档的不同层级。在不同上下文中,层级为变换系数在按比例调整之前的值。在高效率视频译码(HEVC)工作草案8(WD8)的附件A中更详细地描述了设置档及层级。

[0060] 符合HEVC WD8的经译码视频内容使用共同语法。为了实现完整语法的子集,将发信表示稍后出现在位流中的语法元素的存在或不存在的旗标、参数及其它语法元素包含在位流中。

[0061] HEVC WD8将子层定义为时间可按比例调整位流的时间可按比例调整层,其由具有TemporalId变量的特定值的视频译码层(VCL)网络抽象层(NAL)单元及相关联的非VCLNAL单元组成。HEVC WD8进一步将子层表示定义为位流的子集,其由特定子层及较低子层的NAL单元组成。

[0062] HEVC 8的子条款10.1描述位流子集及用以产生子位流的提取过程。下文阐述子条款10.1。

[0063] 10.1子位流提取过程

[0064] 位流一致性要求为:包含在此子条款中所指定的过程的输出中的任何子位流应符合此推荐|国际标准,其中tIdTarget等于在0到6(包含0及6)的范围中的任何值,且其中targetDecLayerIdSet含有值0。

[0065] 注释—符合位流含有一或多个经译码切片NAL单元,其中nuh_reserved_zero_6bits等于0且TemporalId等于0。

[0066] 到此过程的输入为变量tIdTarget及列表targetDecLayerIdSet。

[0067] 此过程的输出为子位流。

[0068] 子位流是通过从位流去除TemporalId大于tIdTarget或nuh_reserved_zero_6bits不在targetDecLayerIdSet中的值当中的所有NAL单元而导出。

[0069] 一般来说,HEVC WD8描述基于识别位流的操作点的层识别符及时间子层识别符的值从位流提取子位流。

[0070] 一般通过nuh_reserved_zero_6bits值集合(表示为OpLayerIdSet)及TemporalId值(表示为OpTid)来识别操作点,且经导出作为子位流提取过程的输出的相关联的位流子集(如在HEVC WD8的子条款10.1中所指定,其中OpTid及OpLayerIdSet作为输入)是可独立解码的。一般将简单操作点模式视为如下操作点模式:其中对于每一操作点,OpLayerIdSet包含且仅包含nuh_reserved_zero_6bits的特定值及nuh_reserved_zero_6bits的所有其它值(小于nuh_reserved_zero_6bits的所述特定值)。

[0071] 以下表1展示VPS原始字节序列有效负载(RBSP)语法及语义的实例。

[0072] 表1

| | | |
|--------|---|-------|
| [0073] | video_parameter_set_rbsp() { | 描述符 |
| | video_parameter_set_id | u(4) |
| | vps_temporal_id_nesting_flag | u(1) |
| | vps_reserved_zero_2bits | u(2) |
| | vps_reserved_zero_6bits | u(6) |
| | vps_max_sub_layers_minus1 | u(3) |
| | profile_tier_level(1, vps_max_sub_layers_minus1) | |
| | vps_reserved_zero_12bits | u(12) |
| | for(i = 0; i <= vps_max_sub_layers_minus1; i++) { | |
| | vps_max_dec_pic_buffering[i] | ue(v) |
| | vps_max_num_reorder_pics[i] | ue(v) |
| | vps_max_latency_increase[i] | ue(v) |
| | } | |
| | vps_num_hrd_parameters | ue(v) |
| | for(i = 0; i < vps_num_hrd_parameters; i++) { | |
| | vps_simple_op_mode_flag[i] | u(1) |
| | if(i > 0) | |
| | operation_point(i) | |
| | hrd_parameters(i == 0, vps_max_sub_layers_minus1) | |
| | } | |
| | vps_extension_flag | u(1) |
| | if(vps_extension_flag) | |
| | while(more_rbsp_data()) | |
| | vps_extension_data_flag | u(1) |
| | } | |
| | rbbsp_trailing_bits() | |
| | } | |

[0074] 经设置为等于1的语法元素“vps_simple_op_mode_flag[i]”指定简单操作点模式用于第i个operation_point_layer_ids()语法结构。等于0的语法元素“vps_simple_op_mode_flag[i]”指定简单操作点模式并不用于第i个operation_point()语法结构。

[0075] 语法结构hrd_parameters(i,vps_max_sub_layers_minus1)及hrd_parameters(j,vps_max_sub_layers_minus1)的任何两个实例可能并不具有相同内容,其中i不等于j。

[0076] 以下表2展示设置档、阶层及层级语法及语义的实例。

[0077] 表2

| | | |
|--------|---|-------|
| [0078] | profile_tier_level(ProfilePresentFlag, MaxNumSubLayersMinus1) { | 描述符 |
| | if(ProfilePresentFlag) { | |
| | general_profile_space | u(2) |
| | general_tier_flag | u(1) |
| | general_profile_idc | u(5) |
| | for(i = 0; i < 32; i++) | |
| | general_profile_compatibility_flag[i] | u(1) |
| | general_reserved_zero_16bits [编辑(GJS): 相应地调整语义。] | u(16) |
| | } | |
| | general_level_idc | u(8) |
| | for(i = 0; i < MaxNumSubLayersMinus1; i++) { | |
| | if(ProfilePresentFlag) | |
| | sub_layer_profile_present_flag[i] | u(1) |
| | sub_layer_level_present_flag[i] | u(1) |
| | if(sub_layer_profile_present_flag[i]) | |
| | sub_layer_profile_idc[i] | u(5) |
| | if(sub_layer_level_present_flag[i]) | |
| | sub_layer_level_idc[i] | u(8) |
| | } | |
| | } | |

[0079] 当ProfilePresentFlag等于1时,经设置为等于1的语法元素“sub_layer_profile_present_flag[i]”指定对于具有等于i的TemporalId的子层的表示,设置档信息存在于profile_tier_level()语法结构中。等于0的sub_layer_profile_present_flag[i]指定对于具有等于i的TemporalId的子层的表示,设置档信息不存在于profile_tier_level()语法结构中。当不存在时,推断sub_layer_profile_present_flag[i]的值等于0。

[0080] 经设置为等于1的语法元素“sub_layer_level_present_flag[i]”指定对于具有等于i的TemporalId的子层的表示,层级信息存在于profile_tier_level()语法结构中。等于0的sub_layer_level_present_flag[i]指定对于具有等于i的TemporalId的子层的表示,层级信息不存在于profile_tier_level()语法结构中。

[0081] 语法元素“sub_layer_profile_idc[i]”及“sub_layer_level_idc[i]”分别具有与general_profile_idc及general_level_idc相同的语义,但仍适用于具有等于i的TemporalId的子层的表示。

[0082] 以下表3展示操作点语法及语义的实例。

[0083] 表3

| | | |
|--------|--|-------|
| [0084] | operation_point(opIdx) { | 描述符 |
| | op_first_present_layer_id[opIdx] | u(6) |
| | if(!vps_simple_op_mode_flag[opIdx]) { | |
| | op_num_layer_id_values_minus1[opIdx] | ue(v) |
| | for(i = 1; i <= op_num_layer_id_values_minus1[opIdx]; i++) | |
| | op_layer_id_delta_minus1[opIdx][i] | ue(v) |
| | } | |
| | } | |

[0085] operation_point(opIdx) 语法结构指定包含在应用有视频参数集中的第opIdx个

hrd_parameters() 语法结构的操作点的OpLayerIdSet中的nuh_reserved_zero_6bits值集合。

[0086] 当vps_simple_op_mode_flag[opIdx]经设置为等于0时,语法元素“op_first_present_layer_id[opIdx]”指定包含在应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet中的nuh_reserved_zero_6bits的第一(即,第0)值。当vps_simple_op_mode_flag[opIdx]等于1时,op_first_present_layer_id[opIdx]指定包含在应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet中的nuh_reserved_zero_6bits的最大值。

[0087] 当vps_simple_op_mode_flag[opIdx]等于0时,语法元素“op_num_layer_id_values_minus1[opIdx]+1指定包含在应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet中的nuh_reserved_zero_6bits值的数目。op_num_layer_id_values_minus1[opIdx]小于或等于63。

[0088] 当vps_simple_op_mode_flag[opIdx]等于0时,语法元素“op_layer_id_delta_minus1[opIdx][i]+1指定包含在应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet中的nuh_reserved_zero_6bits的第i值与nuh_reserved_zero_6bits的第(i-1)值之间的差。op_layer_id_delta_minus1[opIdx][i]的值是在0到63(包含0及63)的范围中。

[0089] 如下导出变量NumOpLayerIdsMinus1[opIdx]。

[0090] if(vps_simple_op_mode_flag[opIdx])

NumOpLayerIdsMinus1[opIdx] = op_first_present_layer_id[opIdx]

[0091] else

NumOpLayerIdsMinus1[opIdx] = op_num_layer_id_values_minus1[opIdx]

[0092] 推断NumOpLayerIdsMinus1[0]等于0。

[0093] 对于在0到NumOpLayerIdsMinus1[opIdx](包含0及NumOpLayerIdsMinus1[opIdx])的范围中的i,如下导出变量OpLayerId[opIdx][i]。

[0094]

OpLayerId[opIdx][0] = vps_simple_op_mode_flag[opIdx] ? 0 :
op_first_present_layer_id[opIdx]

for(i = 1; i <= NumOpLayerIdsMinus1[opIdx]; i++)

if(vps_simple_op_mode_flag[opIdx])

OpLayerId[opIdx][i] = i

else

OpLayerId[opIdx][i] =

OpLayerId[opIdx][i - 1] + op_layer_id_delta_minus1[opIdx][i] + 1

[0095] 推断OpLayerId[0][0]的值等于0。

[0096] 当i不等于j且i与j两者在0到NumOpLayerIdsMinus1[opIdx](包含0及NumOpLayerIdsMinus1[opIdx])的范围中时,无OpLayerId[opIdx][i]的值等于OpLayerId

[opIdx][j]。

[0097] 任何两个集合OpLayerId[opIdx1]及OpLayerId[opIdx2] (其中opIdx1不等于opIdx2) 不包含相同nuh_reserved_zero_6bits值集合。

[0098] 对于在0到NumOpLayerIdsMinus1[opIdx] (包含0及NumOpLayerIdsMinus1[opIdx]) 的范围中的i,应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet经设置为包含且仅包含等于OpLayerId[opIdx][i]的nuh_reserved_zero_6bits值。

[0099] 表4及下文描述替代操作点语法及语义。

[0100] 表4

| | | |
|--------|--|-------|
| [0101] | operation_point(opIdx) { | 描述符 |
| | if(!vps_simple_op_mode_flag[opIdx]) | |
| | op_num_layer_id_values_minus1[opIdx] | ue(v) |
| | for(i = 0; i <= op_num_layer_id_values_minus1[opIdx]; i++) | |
| | op_layer_id[opIdx][i] | u(6) |
| | } | |

[0102] operation_point(opIdx) 语法结构指定包含在应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet中的nuh_reserved_zero_6bits值集合。

[0103] 语法元素“op_num_layer_id_values_minus1[opIdx]”+1指定包含在应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet中的nuh_reserved_zero_6bits值的数目。op_num_layer_id_values_minus1[opIdx]小于或等于63。当不存在时,推断op_num_layer_id_values_minus1[opIdx]的值等于0。

[0104] 在符合本说明书的位流中,op_num_layer_id_values_minus1[opIdx]等于0。尽管在本说明书的此版本中要求op_num_layer_id_values_minus1[opIdx]的值等于0,但解码器允许其它值出现在op_num_layer_id_values_minus1[opIdx]语法中。

[0105] 语法元素“op_layer_id[opIdx][i]”指定包含在应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet中的nuh_reserved_zero_6bits的第i值。

[0106] 如下导出变量NumOpLayerIdsMinus1[opIdx]。

```
if( vps_simple_op_mode_flag[ opIdx ] )
```

```
    NumOpLayerIdsMinus1[ opIdx ] = op_layer_id[ opIdx ][ 0 ]
```

```
[0107] else
```

```
    NumOpLayerIdsMinus1[ opIdx ] = op_num_layer_id_values_minus1[ opIdx ]
```

[0108] 推断NumOpLayerIdsMinus1[0]等于0。

[0109] 对于在0到NumOpLayerIdsMinus1[opIdx] (包含0及NumOpLayerIdsMinus1[opIdx]) 的范围中的i,如下导出变量OpLayerId[opIdx][i]。

```
[0110] for( i=0; i<=NumOpLayerIdsMinus1[ opIdx ]; i++ )
```

```
[0111] OpLayerId[ opIdx ][ i ] = vps_simple_op_mode_flag[ opIdx ] ? i :
```

```
[0112] op_layer_id[ opIdx ][ i ]
```

[0113] 推断OpLayerId[0][0]的值等于0。

[0114] 当i不等于j且i与j两者在0到NumOpLayerIdsMinus1[opIdx] (包含0及NumOpLayerIdsMinus1[opIdx]) 的范围中时,无OpLayerId[opIdx][i]的值等于OpLayerId[opIdx][j]。

[0115] 任何两个集合OpLayerId[opIdx1]及OpLayerId[opIdx2]不包含相同nuh_reserved_zero_6bits值集合,其中opIdx1不等于opIdx2。

[0116] 对于在0到NumOpLayerIdsMinus1[opIdx] (包含0及NumOpLayerIdsMinus1[opIdx]) 的范围中的i,应用有视频参数集中的第opIdx个hrd_parameters()语法结构的操作点的OpLayerIdSet经设置为包含且仅包含等于OpLayerId[opIdx][i]的nuh_reserved_zero_6bits值。

[0117] JCTVC-K0204 (其以引用的方式并入本文中且可在http://phenix.int-evry.fr/jct/doc_end_user/documents/11_Shanghai/wg11/JCTVC-K0204-v1.zip得到) 提供如通过以下语法及语义描述的操作点的经修订发信:

[0118] 表5

| | | |
|--------|---|------------|
| [0119] | operation_point(opIdx) { | 描述符 |
| | for(i = 0; i <= max_num_layers_minus1; i++) | |
| | layer_present_in_op_flag[opIdx][i] | u(1) |
| | } | |

[0120] 经设置为等于1的语法元素“layer_present_in_op_flag[opIdx][i]”指定层i存在于操作点opIdx中,经设置为等于0的语法元素“layer_present_in_op_flag[opIdx][i]”指定层i不存在于操作点opIdx中。

[0121] 用于操作点的发信的现有方法可具有一些潜在缺点。举例来说,用于操作点的发信的现有方法可具有使用如HEVC WD8中所指定的ue(v)译码的经熵译码语法元素,或不支持大于max_num_layers_minus1的nuh_reserved_zero_6bits值(即,层ID)的发信,所述max_num_layers_minus1是在视频参数集(VPS)中发信。

[0122] 本发明提议可潜在地解决这些潜在缺点中的一些缺点的各种技术。在一种此类技术中,首先发信nuh_reserved_zero_6bits值的最大值(即,最大层ID值),后接续发信旗标的列表,其中所述旗标中的每一者指定具有小于最大层ID值的特定层ID值的层是否包含在应用有视频参数集中的第opIdx个hrd_parameters()语法结构的操作点的OpLayerIdSet中。在另一种技术中,发信M个旗标的列表,其中每一旗标指定具有特定的可能层ID值的层是否包含在应用有视频参数集中的第opIdx个hrd_parameters()语法结构的操作点的OpLayerIdSet中。M的值等于任何位流中的可能的不同层ID的总数。举例来说,M可等于 2^N ,其中N为用以表示nuh_reserved_zero_6bits(即,层ID)的位数。在又一技术中,发信nuh_reserved_zero_6bits值的最大值(即,最大层ID值)。如果并未使用简单操作点模式,那么发信旗标的列表,其中每一旗标指定具有小于最大层ID值的特定层ID值的层是否包含在应用有视频参数集中的第opIdx个hrd_parameters()语法结构的操作点的OpLayerIdSet中。

[0123] 现将描述以上技术的一些详细实例。下文所描述的实例一般与HEVC WD8一致,且因此,可假设下文未充分描述的方面与HEVC WD8中的方面相同。

[0124] 以下表6中展示用于第一实例的操作点语法及语义。

[0125] 表6

| | | |
|--------|---|------|
| [0126] | operation_point(opIdx) { | 描述符 |
| | op_max_layer_id [opIdx] | u(6) |
| | for(i = 0; i < op_max_layer_id [opIdx]; i++) | |
| | op_layer_id_included_flag [opIdx][i] | u(1) |
| | } | |

[0127] operation_point(opIdx) 语法结构指定 nuh_reserved_zero_6bits 值集合包含在应用有视频参数集中的第 opIdx 个 hrd_parameters() 语法结构的操作点的 OpLayerIdSet 中。

[0128] 语法元素“op_max_layer_id[opIdx]”指定 nuh_reserved_zero_6bits 的最大值包含在应用有视频参数集中的第 opIdx 个 hrd_parameters() 语法结构的操作点的 OpLayerIdSet 中。尽管表6展示语法元素“op_max_layer_id[opIdx]”是针对每一操作点而发信,但其还可在经编码位流中的别处发信,例如,在序列参数集或VPS中。

[0129] 经设置为等于0的语法元素“op_layer_id_included_flag[opIdx][i]”指定等于i的 nuh_reserved_zero_6bits 的值未包含在应用有视频参数集中的第 opIdx 个 hrd_parameters() 语法结构的操作点的 OpLayerIdSet 中。等于1的 op_layer_id_included_flag[opIdx][i] 指定等于i的 nuh_reserved_zero_6bits 的值包含在应用有视频参数集中的第 opIdx 个 hrd_parameters() 语法结构的操作点的 OpLayerIdSet 中。所有 op_layer_id_included_flag[opIdx][i] 的总和小于或等于 max_num_layers_minus1, 其中 i 在从0到 op_max_layer_id[opIdx]-1 (包含0及 op_max_layer_id[opIdx]-1) 的范围内。

[0130] 如下导出变量 NumOpLayerIdsMinus1[opIdx] 及变量 OpLayerId[opIdx][i], 其中 i 在0到 NumOpLayerIdsMinus1[opIdx] (包含0及 NumOpLayerIdsMinus1[opIdx]) 的范围中。

```
for( i = 0, j = 0; i < op_max_layer_id[ opIdx ]; i++ )
    if( op_layer_id_included_flag[ opIdx ][ i ] )
```

```
OpLayerId[ opIdx ][ j++ ] = i
OpLayerId[ opIdx ][ j ] = op_max_layer_id[ opIdx ]
NumOpLayerIdsMinus1[ opIdx ] = j
```

[0132] 推断 NumOpLayerIdsMinus1[0] 等于0。推断 OpLayerId[0][0] 的值等于0。

[0133] 任何两个集合 OpLayerId[opIdx1] 及 OpLayerId[opIdx2] 不包含相同 nuh_reserved_zero_6bits 值集合, 其中 opIdx1 不等于 opIdx2。

[0134] 应用有视频参数集中的第 opIdx 个 hrd_parameters() 语法结构的操作点的 OpLayerIdSet 经设置为包含且仅包含等于 OpLayerId[opIdx][i] 的 nuh_reserved_zero_6bits 值, 其中 i 在0到 NumOpLayerIdsMinus1[opIdx] (包含0及 NumOpLayerIdsMinus1[opIdx]) 的范围中。

[0135] 返回参看以上实例, 假设位流包含具有各种时间及空间分辨率的六个子流, 其中每一子流具有一层ID。对于通过 opIdx 识别的操作点, 可在位流中发信最大层ID值作为语法元素“op_max_layer_id[opIdx]”的值。出于此实例的目的, 假设最大层ID值为9, 使得 op_max_layer_id[opIdx] 等于9。可使用九个旗标来发信剩余层ID值, 其中第一旗标指示是否存在层ID值0, 第二旗标指示是否存在层ID值1, 等等。因此, 为了发信层ID值2、5及10, 可首

先发信值10,后接续发信旗标序列001001000,其中第三位的1指示存在层ID值2,且第六位的1指示存在层ID值5。在使用表6的语法的情况下,如下获得旗标序列001001000。对于i=0,用于op_layer_id_included_flag[opIdx][i]的旗标的值为0。对于i=1,用于op_layer_id_included_flag[opIdx][i]的旗标的值为0。对于i=3,用于op_layer_id_included_flag[opIdx][i]的旗标的值为0。对于i=4,用于op_layer_id_included_flag[opIdx][i]的旗标的值为0。对于i=5,用于op_layer_id_included_flag[opIdx][i]的旗标的值为1。对于i=6,用于op_layer_id_included_flag[opIdx][i]的旗标的值为0。对于i=7,用于op_layer_id_included_flag[opIdx][i]的旗标的值为0。对于i=8,用于op_layer_id_included_flag[opIdx][i]的旗标的值为0。对于i=9,i的值不小于op_max_layer_id[opIdx],所述op_max_layer_id[opIdx]也等于9。因此,视频解码器可确定接收到最后的旗标。

[0136] 以下表7中展示用于第二实例技术的操作点语法及语义。

[0137] 表7

| | | |
|--------|--|------|
| [0138] | operation_point(opIdx) { | 描述符 |
| | for(i = 0; i < 64; i++) | |
| | op_layer_id_included_flag[opIdx][i] | u(1) |
| | } | |

[0139] operation_point(opIdx) 语法结构指定nuh_reserved_zero_6bits值集合包含在应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet中。

[0140] 经设置为等于0的语法元素“op_layer_id_included_flag[opIdx][i]”指定等于i的nuh_reserved_zero_6bits的值未包含在应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet中。等于1的op_layer_id_included_flag[opIdx][i]指定等于i的nuh_reserved_zero_6bits的值包含在应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet中。所有op_layer_id_included_flag[opIdx][i]的总和小于或等于max_num_layers_minus1,其中i在从0到63(包含0及63)的范围内。

[0141] 如下导出变量NumOpLayerIdsMinus1[opIdx] 及变量OpLayerId[opIdx][i],其中i在0到NumOpLayerIdsMinus1[opIdx] (包含0及NumOpLayerIdsMinus1[opIdx]) 的范围中。

```
for( i = 0, j = 0; i < 64; i++ )
    if( op_layer_id_included_flag[ opIdx ][ i ] )
[0142]         OpLayerId[ opIdx ][ j++ ] = i
        NumOpLayerIdsMinus1[ opIdx ] = j
```

[0143] 推断NumOpLayerIdsMinus1[0] 等于0。推断OpLayerId[0][0]的值等于0。

[0144] 任何两个集合OpLayerId[opIdx1] 及OpLayerId[opIdx2] 不包含相同nuh_reserved_zero_6bits值集合,其中opIdx1不等于opIdx2。

[0145] 应用有视频参数集中的第opIdx个hrd_parameters() 语法结构的操作点的OpLayerIdSet经设置为包含且仅包含等于OpLayerId[opIdx][i]的nuh_reserved_zero_6bits值,其中i在0到NumOpLayerIdsMinus1[opIdx] (包含0及NumOpLayerIdsMinus1

[opIdx]) 的范围中。

[0146] 以下表8中展示用于第三实例的操作点语法及语义。在此实例中,还改变了VPS语法及语义,如表8所展示及如下文所描述。

[0147] 表8

| | | |
|--------|---|-------|
| [0148] | video_parameter_set_rbsp() { | 描述符 |
| | video_parameter_set_id | u(4) |
| | vps_temporal_id_nesting_flag | u(1) |
| | vps_reserved_zero_2bits | u(2) |
| | vps_reserved_zero_6bits | u(6) |
| | vps_max_sub_layers_minus1 | u(3) |
| | profile_tier_level(1, vps_max_sub_layers_minus1) | |
| | vps_reserved_zero_12bits | u(12) |
| | for(i = 0; i <= vps_max_sub_layers_minus1; i++) { | |
| | vps_max_dec_pic_buffering[i] | ue(v) |
| [0149] | vps_max_num_reorder_pics[i] | ue(v) |
| | vps_max_latency_increase[i] | ue(v) |
| | } | |
| | vps_num_hrd_parameters | ue(v) |
| | for(i = 0; i < vps_num_hrd_parameters; i++) { | |
| | vps_simple_op_mode_flag[i] | u(1) |
| | if(i > 0) | |
| | operation_point(i) | |
| | hrd_parameters(i == 0, vps_max_sub_layers_minus1) | |
| | } | |
| | vps_extension_flag | u(1) |
| | if(vps_extension_flag) | |
| | while(more_rbsp_data()) | |
| | vps_extension_data_flag | u(1) |
| | } | |
| | rbbsp_trailing_bits() | |
| | } | |

[0150] 经设置为等于1的语法元素“vps_simple_op_mode_flag[i]”指定简单操作点模式用于第i个operation_point()语法结构。等于0的vps_simple_op_mode_flag[i]指定简单操作点模式并不用于第i个operation_point()语法结构。

[0151] 语法结构hrd_parameters(i,vps_max_sub_layers_minus1)及hrd_parameters(j,vps_max_sub_layers_minus1)的任何两个实例并不具有相同内容,其中i不等于j。

[0152] 表9

| | | |
|--------|---|------|
| [0153] | operation_point(opIdx) { | 描述符 |
| | op_max_layer_id[opIdx] | u(6) |
| | if(!vps_simple_op_mode_flag[opIdx]) | |
| | for(i = 0; i < op_max_layer_id[opIdx]; i++) | |
| | op_layer_id_included_flag[opIdx][i] | u(1) |
| | } | |

[0154] 表9所展示的operation_point(opIdx)语法结构指定nuh_reserved_zero_6bits

值集合包含在应用有视频参数集中的第opIdx个hrd_parameters()语法结构的操作点的OpLayerIdSet中。

[0155] 语法元素“op_max_layer_id[opIdx]”指定nuh_reserved_zero_6bits的最大值包含在应用有视频参数集中的第opIdx个hrd_parameters()语法结构的操作点的OpLayerIdSet中。

[0156] 当vps_simple_op_mode_flag[opIdx]等于0时,经设置为等于0的语法元素“op_layer_id_incuded_flag[opIdx][i]”指定等于i的nuh_reserved_zero_6bits的值未包含在应用有视频参数集中的第opIdx个hrd_parameters()语法结构的操作点的OpLayerIdSet中。当vps_simple_op_mode_flag[opIdx]等于0时,等于1的op_layer_id_incuded_flag[opIdx][i]指定等于i的nuh_reserved_zero_6bits的值包含在应用有视频参数集中的第opIdx个hrd_parameters()语法结构的操作点的OpLayerIdSet中。所有op_layer_id_incuded_flag[opIdx][i]的总和小于或等于max_num_layers_minus1,其中i在从0到op_max_layer_id[opIdx]-1(包含0及op_max_layer_id[opIdx]-1)的范围内。

[0157] 如下导出变量NumOpLayerIdsMinus1[opIdx]及变量OpLayerId[opIdx][i],其中i在0到NumOpLayerIdsMinus1[opIdx](包含0及NumOpLayerIdsMinus1[opIdx])的范围内。

```

        if( vps_simple_op_mode_flag[ opIdx ] ) {
            for( i = 0; i <= op_max_layer_id[ opIdx ]; i++ )
                OpLayerId[ opIdx ][ i ] = i
            NumOpLayerIdsMinus1[ opIdx ] = op_max_layer_id[ opIdx ]
        } else {
[0158]     for( i = 0, j = 0; i < op_max_layer_id[ opIdx ]; i++ )
                if( op_layer_id_incuded_flag[ opIdx ][ i ] )
                    OpLayerId[ opIdx ][ j++ ] = i
                OpLayerId[ opIdx ][ j ] = max_layer_id[ opIdx ]
                NumOpLayerIdsMinus1[ opIdx ] = j
        }

```

[0159] 推断NumOpLayerIdsMinus1[0]等于0。推断OpLayerId[0][0]的值等于0。

[0160] 任何两个集合OpLayerId[opIdx1]及OpLayerId[opIdx2]不包含相同nuh_reserved_zero_6bits值集合,其中opIdx1不等于opIdx2。

[0161] 应用有视频参数集中的第opIdx个hrd_parameters()语法结构的操作点的OpLayerIdSet经设置为包含且仅包含等于OpLayerId[opIdx][i]的nuh_reserved_zero_6bits值,其中i在0到NumOpLayerIdsMinus1[opIdx](包含0及NumOpLayerIdsMinus1[opIdx])的范围内。

[0162] 图2为说明可实施本发明所描述的技术的视频编码器20的实例的框图。视频编码器20可执行视频切片内的视频块的帧内译码及帧间译码。帧内译码依赖于空间预测来减少或去除给定视频帧或图片内的视频中的空间冗余。帧间译码依赖于时间预测来减少或去除视频序列的邻近帧或图片内的视频中的时间冗余。帧内模式(I模式)可指若干基于空间的

译码模式中的任一者。帧间模式(例如,单向预测(P模式)或双向预测(B模式))可指若干基于时间的译码模式中的任一者。

[0163] 如图2所展示,视频编码器20接收待编码视频帧内的当前视频块。在图2的实例中,视频编码器20包含模式选择单元40、参考帧存储器64、求和器50、变换处理单元52、量化单元54及熵编码单元56。模式选择单元40又包含运动补偿单元44、运动估计单元42、帧内预测处理单元46及分割单元48。为了实现视频块重新构造,视频编码器20还包含反量化单元58、反变换处理单元60及求和器62。还可包含解块滤波器以对块边界进行滤波,从而从经重新构造的视频去除方块效应伪影。在需要时,解块滤波器通常将对求和器62的输出进行滤波。除解块滤波器之外,还可使用额外滤波器(回路内或回路后)。出于简洁起见而未展示此类滤波器,但在需要时,此类滤波器可对求和器50的输出进行滤波(作为回路内滤波器)。

[0164] 在编码过程期间,视频编码器20接收待译码的视频帧或切片。可将所述帧或切片划分成多个视频块。运动估计单元42及运动补偿单元44相对于一或多个参考帧中的一或多个块执行所接收视频块的帧间预测性译码,以提供时间预测。帧内预测处理单元46可替代地相对于与待译码的块相同的帧或切片中的一或多个相邻块执行所接收视频块的帧内预测性译码,以提供空间预测。视频编码器20可执行多个译码遍次(例如)以选择用于视频数据的每一块的适当译码模式。

[0165] 此外,分割单元48可基于先前译码遍次中的先前分割方案的评估将视频数据的块分割成若干子块。举例来说,分割单元48最初可将帧或切片分割成LCU,且基于速率-失真分析(例如,速率-失真最优化)将所述LCU中的每一者分割成子CU。模式选择单元40可进一步产生指示LCU到子CU的分割的四叉树数据结构。四叉树的叶节点CU可包含一或多个PU及一或多个TU。

[0166] 模式选择单元40可(例如)基于误差结果选择译码模式(帧内或帧间)中的一者,且将所得的经帧内或经帧间译码块提供到求和器50以产生残余块数据且提供到求和器62以重新构造供用作参考帧的经编码块。模式选择单元40还将例如运动向量、帧内模式指示符、分割信息及其它此类语法信息的语法元素提供到熵编码单元56。

[0167] 运动估计单元42及运动补偿单元44可高度地集成,但出于概念目的而分别加以说明。由运动估计单元42执行的运动估计为产生运动向量的过程,所述运动向量估计视频块的运动。举例来说,运动向量可指示当前视频帧或图片内的视频块的PU相对于参考帧(或其它经译码单元)内的预测性块(其相对于当前帧(或其它经译码单元)内正被译码的当前块)的位移。预测性块为被发现在像素差方面紧密地匹配待译码的块的块,所述像素差可通过绝对差总和(SAD)、平方差总和(SSD)或其它差度量予以确定。在一些实例中,视频编码器20可计算存储在参考帧存储器64中的参考图片的子整数像素位置的值。举例来说,视频编码器20可内插所述参考图片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。因此,运动估计单元42可执行相对于全像素位置及分数像素位置的运动搜索,且以分数像素精度输出运动向量。

[0168] 运动估计单元42通过比较经帧间译码切片中的视频块的PU的位置与参考图片的预测性块的位置来计算所述PU的运动向量。所述参考图片可选自第一参考图片列表(列表0)或第二参考图片列表(列表1),所述参考图片列表中的每一者识别存储在参考帧存储器64中的一或多个参考图片。运动估计单元42将所计算的运动向量发送到熵编码单元56及运

动补偿单元44。

[0169] 由运动补偿单元44执行的运动补偿可涉及基于由运动估计单元42确定的运动向量提取或产生所述预测性块。此外,在一些实例中,运动估计单元42及运动补偿单元44可在功能上集成。在接收到用于当前视频块的PU的运动向量后,运动补偿单元44即可将运动向量所指向的预测性块定位于参考图片列表中的一者中。求和器50通过从正被译码的当前视频块的像素值减去预测性块的像素值从而形成像素差值来形成残余视频块,如下文所论述。一般来说,运动估计单元42执行相对于明度分量的运动估计,且运动补偿单元44将基于所述明度分量计算的运动向量用于色度分量与明度分量两者。模式选择单元40还可产生与视频块及视频切片相关联的语法元素以供视频解码器30用于解码视频切片的视频块。

[0170] 如上文所描述,作为由运动估计单元42及运动补偿单元44执行的帧间预测的替代例,帧内预测处理单元46可对当前块进行帧内预测。明确地说,帧内预测处理单元46可确定待用以编码当前块的帧内预测模式。在一些实例中,帧内预测处理单元46可(例如)在单独编码遍次期间使用各种帧内预测模式编码当前块,且帧内预测处理单元46(或在一些实例中,模式选择单元40)可从所测试的模式中选择将使用的适当帧内预测模式。

[0171] 举例来说,帧内预测处理单元46可使用针对各种所测试的帧内预测模式的速率-失真分析来计算速率-失真值,且在所测试的模式当中选择具有最佳速率-失真特性的帧内预测模式。速率-失真分析一般确定经编码块与经编码以产生所述经编码块的原始的未经编码块之间的失真(或误差)的量,以及用以产生经编码块的位速率(即,位数)。帧内预测处理单元46可从失真及速率计算各种经编码块的比以确定哪一帧内预测模式展现块的最佳速率-失真值。

[0172] 在针对块选择帧内预测模式之后,帧内预测处理单元46可将指示针对所述块选择的帧内预测模式的信息提供到熵编码单元56。熵编码单元56可编码指示所述所选择的帧内预测模式的信息。视频编码器20可在所发射的位流配置数据中包含以下各者:编码各种块的上下文的定义,及待用于所述上下文中的每一者的最有可能的帧内预测模式、帧内预测模式索引表及经修改的帧内预测模式索引表的指示,所述位流配置数据可包含多个帧内预测模式索引表及多个经修改的帧内预测模式索引表(还被称作码字映射表)。

[0173] 视频编码器20通过从正被译码的原始视频块减去来自模式选择单元40的预测数据而形成残余视频块。求和器50表示执行此减法运算的一或多个组件。变换处理单元52将例如离散余弦变换(DCT)或概念上类似变换等变换应用于残余块,从而产生包括残余变换系数值的视频块。变换处理单元52可执行概念上类似于DCT的其它变换。还可使用小波变换、整数变换、子带变换或其它类型的变换。在任何情况下,变换处理单元52将变换应用于残余块,从而产生残余变换系数块。所述变换可将残余信息从像素值域转换到变换域(例如,频域)。变换处理单元52可将所得变换系数发送到量化单元54。

[0174] 量化单元54量化变换系数以进一步减小位速率。所述量化过程可减小与所述系数中的一些或全部相关联的位深度。可通过调整量化参数来修改量化程度。在一些实例中,量化单元54可接着执行包含经量化的变换系数的矩阵的扫描。替代地,熵编码单元56可执行所述扫描。

[0175] 在量化之后,熵编码单元56熵译码经量化的变换系数。举例来说,熵编码单元56可执行上下文自适应性可变长度译码(CAVLC)、上下文自适应性二进制算术译码(CABAC)、基

于语法的上下文自适应性二进制算术译码 (SBAC)、概率区间分割熵 (PIPE) 译码或另一熵译码技术。在基于上下文的熵译码的情况下,上下文可基于相邻块。在通过熵编码单元56进行的熵译码之后,可将经编码位流发射到另一装置(例如,视频解码器30)或经存档以供稍后发射或检索。

[0176] 反量化单元58及反变换处理单元60分别应用反量化及反变换以在像素域中重新构造残余块(例如)以供稍后用作参考块。运动补偿单元44可通过将所述残余块加到参考帧存储器64的帧中的一者的预测性块来计算参考块。运动补偿单元44还可将一或多个内插滤波器应用于所述经重新构造的残余块以计算供用于运动估计的子整数像素值。求和器62将所述经重新构造的残余块加到由运动补偿单元44产生的经运动补偿的预测块以产生供存储在参考帧存储器64中的经重新构造的视频块。所述经重新构造的视频块可供运动估计单元42及运动补偿单元44用作参考块以对后续视频帧中的块进行帧间译码。

[0177] 图3为说明可实施本发明所描述的技术的视频解码器30的实例的框图。在图3的实例中,视频解码器30包含熵解码单元70、运动补偿单元72、帧内预测处理单元74、反量化单元76、反变换处理单元78、参考帧存储器82及求和器80。在一些实例中,视频解码器30可执行一般与关于视频编码器20(如图2中所展示)所描述的编码遍次互逆的解码遍次。

[0178] 在解码过程期间,视频解码器30从视频编码器20接收表示经编码视频切片的视频块及相关联的语法元素的经编码视频位流。视频解码器30可从网络实体29接收经编码视频位流。网络实体29可(例如)为服务器、媒体感知网络元件(MANE)、视频编辑器/剪接器,或经配置以实施上文所描述的技术中的一或多者的其它此类装置。网络实体29可包含经配置以执行本发明的技术的外部装置。如上文所描述,在网络实体29将经编码视频位流发射到视频解码器30之前,可通过网络实体29来实施本发明中所描述的技术中的一些技术。在一些视频解码系统中,网络实体29及视频解码器30可为单独装置的部分,而在其它例子中,可通过包括视频解码器30的相同装置来执行关于网络实体29所描述的功能性。

[0179] 在一个实例中,网络实体29可存储或接收视频数据的原始位流,所述原始位流为可按比例调整的及/或包含多个层或视图。在所述原始位流中,参数集(例如,VPS)可包含上文所论述的操作点语法。所述操作点语法可供网络实体29使用以识别哪些层存在于操作点中。网络实体29可基于包含在VPS中的操作点语法及基于视频解码器30需要的内容或视频解码器30请求的内容而从所述原始位流选择多个操作点(即,子位流)中的一者。对于对应于所选择的操作点的子位流,网络实体29可将包括所述位流的VLC NAL单元及非VLC NAL单元转发到视频解码器30且不转发其它NAL单元。

[0180] 对于VPS中所识别的特定操作点,网络实体29可接收用于位流的最大层ID值的指示且接收用于具有小于最大层ID值的层ID值的层的一系列旗标。基于所述旗标的值,网络实体29可确定哪些层包含在所述操作点中。举例来说,如果最大层ID的值为M,那么层M包含在操作点中。对于层M-1,网络实体29接收旗标,其中所述旗标的第一值(例如,1或0)指示层M-1包含在操作点中,且所述旗标的第二值(例如,0或1)指示层M-1未包含在操作点中。对于层M-2,网络实体29接收第二旗标,其中所述第二旗标的第一值(例如,1或0)指示层M-2包含在操作点中,且所述第二旗标的第二值(例如,0或1)指示层M-2未包含在操作点中。网络实体29可同样地接收用于直到层0的所有剩余层的旗标。因此,如果最大层ID的值为M,那么网络实体29可接收用于所有层0到M-1的旗标。

[0181] 视频解码器30的熵解码单元70熵解码由网络实体29提供的位流以产生经量化的系数、运动向量或帧内预测模式指示符及其它语法元素(例如,上文所描述的操作点语法)。熵解码单元70将运动向量及其它语法元素转发到运动补偿单元72。视频解码器30可在经编码位流的不同部分处接收不同语法元素。举例来说,一些语法元素可在VPS层级、SPS层级或APS层级被接收,而其它语法元素是在视频切片层级及/或视频块层级接收。

[0182] 当视频切片经译码为经帧内译码(I)切片时,帧内预测处理单元74可基于所发信的帧内预测模式及来自当前帧或图片的先前经解码块的数据产生当前视频切片的视频块的预测数据。当视频帧经译码为经帧间译码(即,B、P或GPB)切片时,运动补偿单元72基于从熵解码单元70接收的运动向量及其它语法元素产生用于当前视频切片的视频块的预测性块。所述预测性块可从参考图片列表中的一者内的参考图片中的一者产生。视频解码器30可基于存储在参考帧存储器82中的参考图片使用默认构造技术构造参考帧列表(列表0及列表1)。

[0183] 运动补偿单元72通过剖析运动向量及其它语法元素来确定当前视频切片的视频块的预测信息,且使用所述预测信息产生正被解码的当前视频块的预测性块。举例来说,运动补偿单元72使用所接收语法元素中的一些语法元素来确定用以译码视频切片的视频块的预测模式(例如,帧内预测或帧间预测)、帧间预测切片类型(例如,B切片或P切片)、用于切片的参考图片列表中的一或多者的构造信息、用于切片的每一经帧间编码视频块的运动向量、用于切片的每一经帧间译码视频块的帧间预测状态及其它信息,以解码当前视频切片中的视频块。

[0184] 运动补偿单元72还可基于内插滤波器来执行内插。运动补偿单元72可使用如由视频编码器20在视频块的编码期间使用的内插滤波器,来计算参考块的子整数像素的经内插值。在此情况下,运动补偿单元72可从所接收的语法元素确定由视频编码器20使用的内插滤波器且使用所述内插滤波器来产生预测性块。

[0185] 反量化单元76将提供于位流中且通过熵解码单元70解码的经量化的变换系数反量化(即,解量化)。反量化过程可包含使用通过视频解码器30针对视频切片中的每一视频块所计算的量化参数 QP_V 来确定量化的程度且,同样地确定应应用的反量化的程度。反变换处理单元78将反变换(例如,反DCT、反整数变换或概念上类似的反变换过程)应用于变换系数以便在像素域中产生残余块。

[0186] 在运动补偿单元72基于运动向量及其它语法元素产生用于当前视频块的预测性块之后,视频解码器30通过将来自反变换处理单元78的残余块与通过运动补偿单元72产生的对应预测性块加总来形成经解码视频块。求和器90表示执行此求和运算的一或多个组件。在必要时,还可应用解块滤波器来对经解码块进行滤波以便去除方块效应伪影。其它回路滤波器(译码回路中或译码回路后)还可用以使像素转变平滑,或以其它方式改善视频质量。接着将给定帧或图片中的经解码视频块存储在参考帧存储器82中,所述参考帧存储器82存储用于随后的运动补偿的参考图片。参考帧存储器82还存储供稍后在显示装置(例如,图1的显示装置32)上呈现的经解码视频。

[0187] 图4为说明形成网络100的部分的实例装置集合的框图。在此实例中,网络100包含路由装置104A、104B(路由装置104)及转码装置106。路由装置104及转码装置106希望表示可形成网络100的部分的少量装置。其它网络装置(例如,交换器、集线器、网关、防火墙、网

桥及其它此类装置)也可包含在网络100内。此外,可沿服务器装置102与客户端装置108之间的网络路径来提供额外网络装置。在一些实例中,服务器装置102可对应于源装置12(图1),而客户端装置108可对应于目的地装置14(图1)。

[0188] 一般来说,路由装置104实施一或多个路由协议以通过网络100交换网络数据。在一些实例中,路由装置104可经配置以执行代理服务器或高速缓冲存储器操作。因此,在一些实例中,可将路由装置104称作代理服务器装置。一般来说,路由装置104执行路由协议以通过网络100来探索路径。通过执行此类路由协议,路由装置104B可探索自其自身经由路由装置104A到服务器装置102的网络路径。

[0189] 路由装置104及转码装置106为可实施本发明所描述的技术的装置的实例。举例来说,作为将视频数据从服务器装置102投送到客户端装置108的部分,路由装置104及/或转码装置106可接收包含操作点语法的VPS语法。所述操作点语法可(例如)包含位流的最大层ID值。路由装置104及转码装置106可另外在操作点语法中接收用于具有小于最大层ID值的层ID的层的一或多个旗标。基于所述最大层ID值及所述旗标,路由装置104及转码装置106可确定所述层包含在操作点中,且因此可识别包括操作点的子位流的NAL单元。

[0190] 图5展示根据本发明的技术的编码视频数据的实例方法。将参考视频编码器(例如,视频编码器20)来描述图5的技术。视频编码器20可在经编码视频数据的位流中产生用于位流的最大层ID值的指示(152)。视频编码器20还可产生用于具有小于最大层ID值的层ID值的第一层的旗标(154)。最大层ID的指示及旗标可(例如)为VPS中含有的操作点语法的部分。

[0191] 图6展示根据本发明的技术的处理视频数据的实例方法。将参考视频处理装置来描述图6的技术,所述视频处理装置可对应于视频解码器(例如,图1及图3的视频解码器30)或可对应于网络装置或网络实体(例如,图1的网络装置13、图3的网络实体29,或图4的路由装置104或转码装置106)。视频处理装置可在经编码视频数据的位流中接收用于位流的最大层ID值的指示(162)。视频处理装置还可接收用于具有小于最大层ID值的层ID值的第一层的旗标(164)。基于旗标的值,视频处理装置可基于旗标的值确定第一层是否包含在操作点中(166)。

[0192] 应认识到,取决于实例,本文所描述的技术中的任一者的某些动作或事件可以不同序列来执行、可被添加、合并或完全地省略(例如,并非所有所描述动作或事件对于实践所述技术来说为必要的)。此外,在某些实例中,可(例如)通过多线程处理、中断处理或多个处理器同时地而不是顺序地执行动作或事件。

[0193] 在一或多个实例中,可以硬件、软件、固件或其任何组合来实施所描述功能。如果以软件来实施,那么可将所述功能作为一或多个指令或代码存储在计算机可读媒体上或经由计算机可读媒体进行传输,且由基于硬件的处理单元来执行。计算机可读媒体可包含计算机可读存储媒体(其对应于例如数据存储媒体等有形媒体)或通信媒体,通信媒体包含(例如)根据通信协议促进计算机程序从一处传送到另一处的任何媒体。以此方式,计算机可读媒体一般可对应于:(1)非暂时性的有形计算机可读存储媒体;或(2)例如信号或载波等通信媒体。数据存储媒体可为可由一或多个计算机或一或多个处理器存取以检索指令、代码及/或数据结构以用于实施本发明中所描述的技术的任何可用媒体。计算机程序产品可包含计算机可读媒体。

[0194] 作为实例而不是限制,此类计算机可读存储媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置、磁盘存储装置或其它磁性存储装置、快闪存储器,或可用以存储呈指令或数据结构形式的所要程序代码且可由计算机存取的任何其它媒体。而且,可将任何连接恰当地称为计算机可读媒体。举例来说,如果使用同轴电缆、光缆、双绞线、数字订户线(DSL)或无线技术(例如,红外线、无线电及微波)从网站、服务器或其它远程源发射指令,那么将同轴电缆、光缆、双绞线、DSL或无线技术(例如,红外线、无线电及微波)包含在媒体的定义中。然而,应理解,计算机可读存储媒体及数据存储媒体不包含连接、载波、信号或其它暂时性媒体,而是有关于非暂时性的有形存储媒体。如本文所使用,磁盘及光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字影音光盘(DVD)、软性磁盘及蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘通过激光以光学方式再现数据。以上各者的组合也应包含在计算机可读媒体的范围内。

[0195] 可通过例如以下各者等一或多个处理器来执行指令:一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其它等效集成或离散逻辑电路系统。因此,如本文所使用,术语“处理器”可指上述结构或适于实施本文所描述的技术的任何其它结构中的任一者。另外,在一些方面中,可将本文所描述的功能性提供于经配置以用于编码及解码的专用硬件及/或软件模块内,或并入于组合式编码解码器中。而且,可将所述技术完全实施于一或多个电路或逻辑元件中。

[0196] 可将本发明的技术实施于广泛多种装置或设备中,所述装置或设备包含无线手持机、集成电路(IC)或IC集合(例如,芯片组)。本发明中描述各种组件、模块或单元以强调经配置以执行所揭示技术的装置的功能方面,但未必要求通过不同硬件单元来实现。确切地说,如上文所描述,可将各种单元组合于编码解码器硬件单元中,或通过互操作性硬件单元(包含如上文所描述的一或多个处理器)的集合且结合合适的软件及/或固件来提供所述单元。

[0197] 已描述各种实例。这些及其它实例在所附权利要求书的范围内。

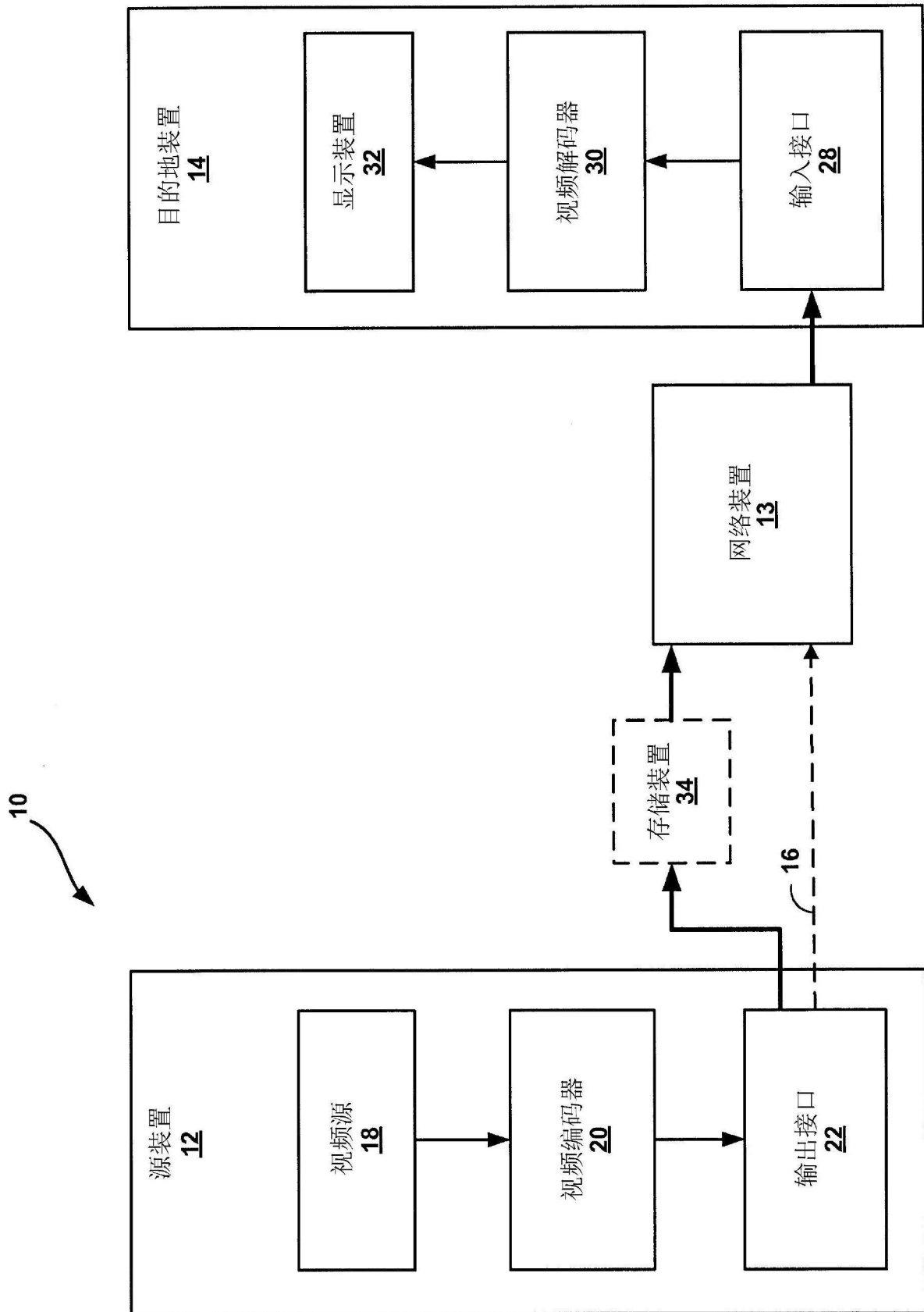


图1

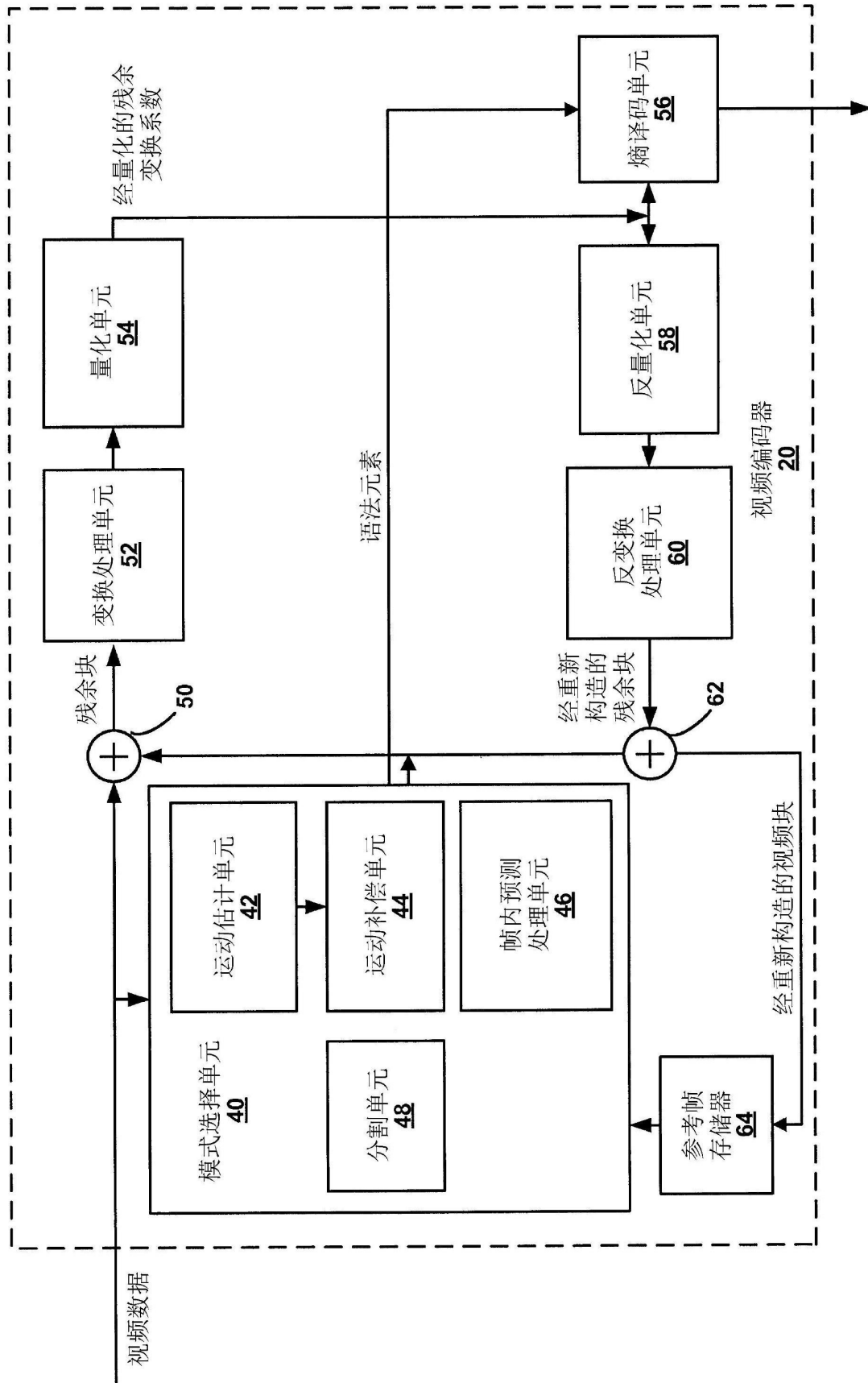


图2

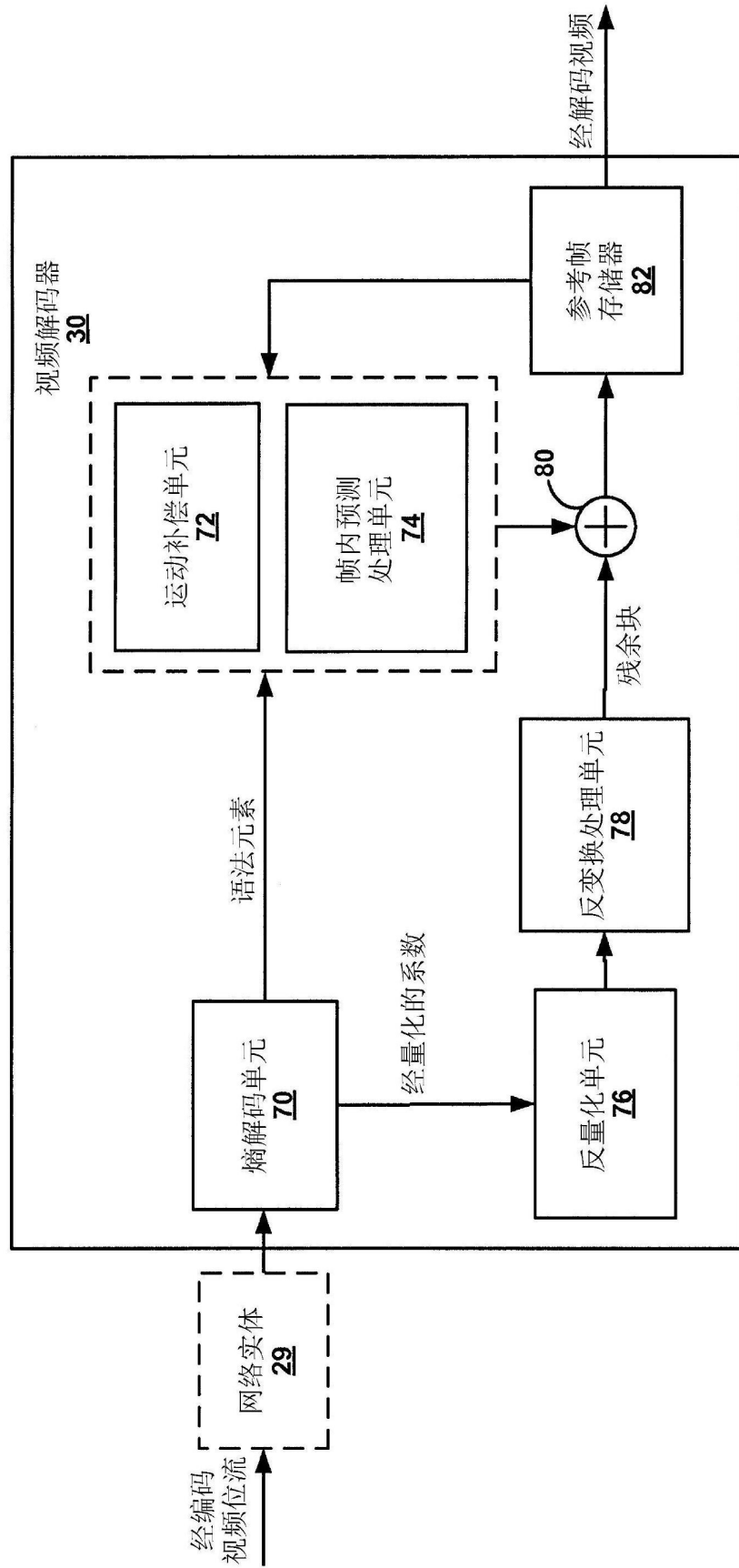


图3

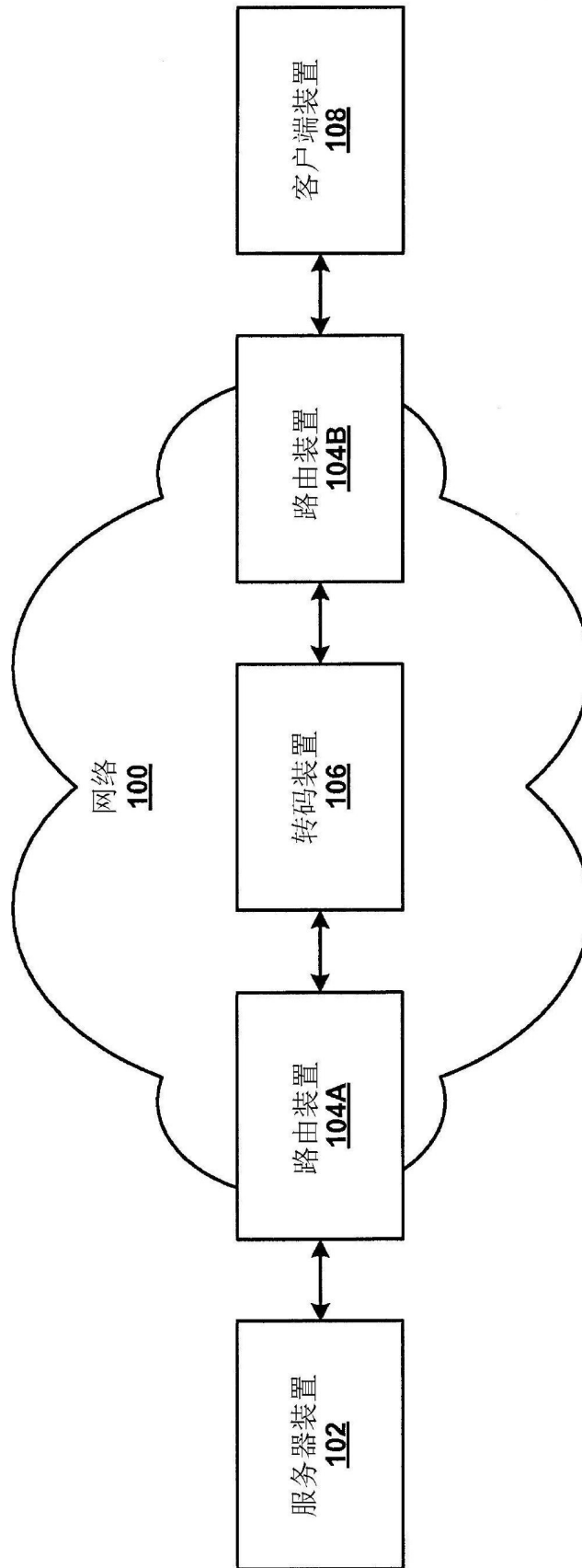


图4

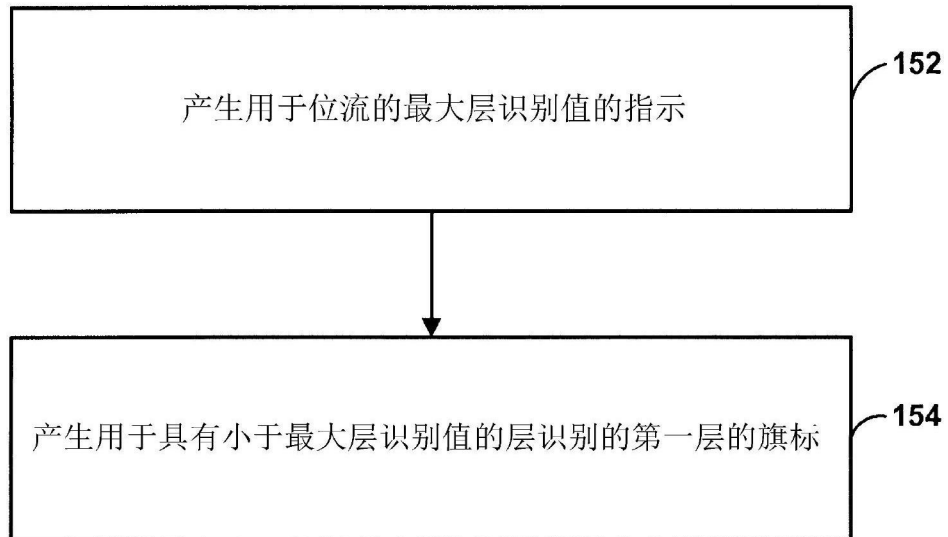


图5

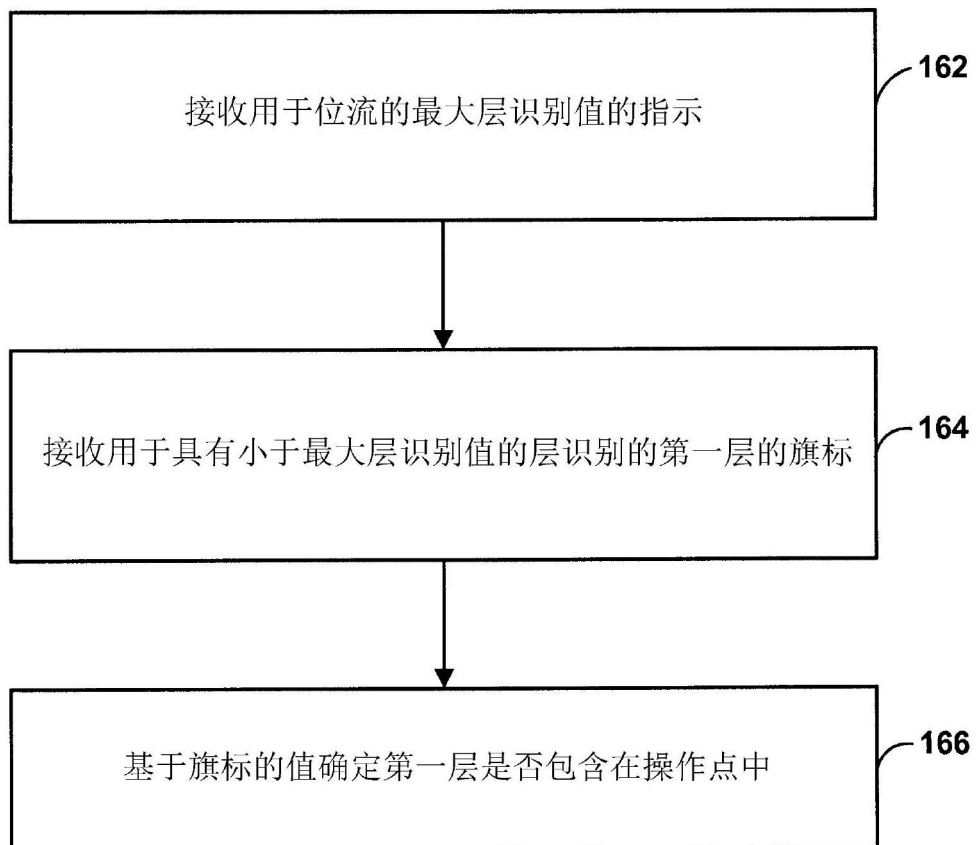


图6