



# (12)发明专利

(10)授权公告号 CN 105493506 B

(45)授权公告日 2019.04.02

(21)申请号 201480039783.2

(22)申请日 2014.07.14

(65)同一申请的已公布的文献号  
申请公布号 CN 105493506 A

(43)申请公布日 2016.04.13

(30)优先权数据

61/846,421 2013.07.15 US

14/329,416 2014.07.11 US

(85)PCT国际申请进入国家阶段日  
2016.01.12

(86)PCT国际申请的申请数据  
PCT/US2014/046544 2014.07.14

(87)PCT国际申请的公布数据  
W02015/009628 EN 2015.01.22

(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 克里希纳坎斯·拉帕卡

(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 19/187(2006.01)

H04N 19/597(2006.01)

H04N 19/70(2006.01)

H04N 19/30(2006.01)

H04N 19/46(2006.01)

H04N 19/184(2006.01)

(56)对比文件

CN 101523919 A,2009.09.02,全文.

W0 2012/170833 A1,2012.12.13,全文.

Ying Chen et al.AHG7: Comments on parameter sets for MV-HEVC.《Joint Collaborative Team on 3D Video Coding Extensions of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11》.2013,第1.1.2节.

Ying Chen et al.Unified NAL unit header design for HEVC and its extensions.《Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11》.2011,第2节.

审查员 袁敏

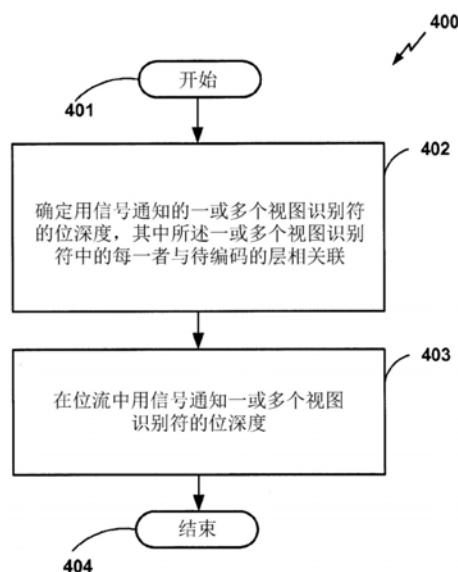
权利要求书3页 说明书30页 附图8页

(54)发明名称

在参数集中用信号通知视图ID位深度

(57)摘要

根据某些方面的一种用于编码视频信息的设备包含存储器及计算硬件。所述存储器经配置以存储视频信息。所述计算硬件经配置以确定用信号通知的一或多个视图识别符的位深度,其中所述一或多个视图识别符中的每一者与待编码的层相关联。所述计算硬件进一步经配置以在位流中用信号通知所述一或多个视图识别符的所述位深度。



1. 一种用于编码视频信息的设备,其包括:

存储器,其经配置以存储视频数据;及

计算硬件,其操作性地耦合到所述存储器且经配置以:

确定位流中将用信号通知的多个视图识别符ID值的每一者的位深度,其中所述多个视图ID值中的每一者与待编码的多个视图中的一者相关联;

在所述位流中用信号通知所述多个视图ID值的经确定的位深度的指示;

基于所述经确定的位深度的所述指示,在所述位流中用信号通知所述多个视图ID值以使得所述多个视图ID值能根据经索引的列表解析,所述经索引的列表具有对应于所述多个视图ID值的多个视图索引值,所述多个视图ID值的每一视图ID值指示了语法元素,所述语法元素的值使用由所述经确定的位深度指示的位数来指定所述多个视图中的对应一者的视图ID,及

基于所述位流中用信号通知的所述多个视图ID值来编码所述多个视图,

其中在视频参数集VPS中用信号通知所述经确定的位深度的所述指示及所述多个视图ID值。

2. 根据权利要求1所述的设备,其中所述计算硬件进一步经配置以基于待编码的视图的最大数目而确定所述位深度。

3. 根据权利要求1所述的设备,其中所述计算硬件进一步经配置以基于可缩放性掩码索引来确定一个或多个视图次序索引。

4. 根据权利要求1所述的设备,其中所述计算硬件经配置以用信号通知指示是否在所述位流中用信号通知所述多个视图ID值的旗标。

5. 根据权利要求1所述的设备,其中由所述用信号通知的位深度指示的所述位数在1与16之间。

6. 根据权利要求1所述的设备,其中所述设备包括以下各者中的一或多个:桌上型计算机、笔记本计算机、膝上型计算机、平板计算机、机顶盒、电话手持机、智能电话、智能板、电视、相机、显示装置、数字媒体播放器、视频游戏控制台,或视频流式传输装置。

7. 根据权利要求1所述的设备,其中所述经索引的列表是根据高效率视频译码HEVC标准指定的view\_id\_val[i],且所述经确定的位深度指定view\_id\_val[i]中的元素的以位为单位的长度。

8. 一种用于解码视频信息的设备,其包括:

存储器,其经配置以存储视频数据;及

计算硬件,其操作性地耦合到所述存储器且经配置以:

接收指示用于在位流中用信号通知多个视图识别符ID值的每一者的位的数目的位深度指示符,其中所述多个视图ID值中的每一者与待解码的多个视图中的一者相关联;

基于所述位深度指示符,接收所述位流中的所述多个视图ID值以作为经索引的列表,所述经索引的列表具有对应于所述多个视图ID值的多个视图索引值,

所述多个视图ID值的每一视图ID值指示了语法元素,所述语法元素的值使用由所述位深度指示符指示的位数来指定所述多个视图中的对应一者的视图ID,

及

基于所接收的所述位流中的所述多个视图ID值来解码所述多个视图,

其中在视频参数集VPS中接收所述位深度指示符及所述多个视图ID值。

9. 根据权利要求8所述的设备,其中由所述位深度指示符指示的所述位数目在1与16之间。

10. 根据权利要求8所述的设备,其中所述计算硬件进一步经配置以基于可缩放性掩码索引来确定一个或多个视图次序索引。

11. 根据权利要求8所述的设备,其中所述设备包括以下各者中的一或多个:桌上型计算机、笔记本计算机、膝上型计算机、平板计算机、机顶盒、电话手持机、智能电话、智能板、电视、相机、显示装置、数字媒体播放器、视频游戏控制台,或视频流式传输装置。

12. 根据权利要求8所述的设备,其中所述经索引的列表是根据高效率视频译码HEVC标准指定的view\_id\_val[i],且经确定的位深度指定view\_id\_val[i]中的元素的以位为单位的长度。

13. 一种用信号通知视图识别符位深度的方法,所述方法包括:

确定位流中将用信号通知的多个视图识别符ID值的每一者的位深度,其中所述多个视图ID值中的每一者与待编码的多个视图中的一者相关联;

在所述位流中用信号通知所述多个视图ID值的经确定的位深度的指示;

基于所述经确定的位深度的所述指示,在所述位流中用信号通知所述多个视图ID值以使得所述多个视图ID值能根据经索引的列表解析,所述经索引的列表具有对应于所述多个视图ID值的多个视图索引值,所述多个值的每一视图ID值指示了语法元素,所述语法元素的值使用由所述经确定的位深度指示的位数目来指定所述多个视图中的对应一者的视图ID,及

基于所述位流中用信号通知的所述多个视图ID值来编码所述多个视图,

其中在视频参数集VPS中用信号通知所述经确定的位深度的所述指示及所述多个视图ID值。

14. 根据权利要求13所述的方法,其进一步包括基于待编码的视图的最大数目而确定所述位深度。

15. 根据权利要求13所述的方法,其进一步包括基于可缩放性掩码索引来确定一个或多个视图次序索引。

16. 根据权利要求13所述的方法,其进一步包括用信号通知指示是否在所述位流中用信号通知所述多个视图ID值的旗标。

17. 根据权利要求13所述的方法,其中由所述用信号通知的位深度指示的所述位数目在1与16之间。

18. 一种解码视频信息的方法,其包括:

接收指示用于在位流中用信号通知多个视图识别符ID值的每一者的位的数目的位深度指示符,其中所述多个视图ID值中的每一者与待解码的多个视图中的一者相关联;

基于所述位深度指示符,接收所述位流中的所述多个视图ID值以作为经索引的列表,所述经索引的列表具有对应于所述多个视图ID值的多个视图索引值,所述多个视图ID值的每一视图ID值指示了语法元素,所述语法元素的值使用由所述位深度指示符指示的位数目来指定所述多个视图中的对应一者的视图ID,及

基于所接收的所述位流中的所述多个视图ID值来解码所述多个视图,

其中在视频参数集VPS中接收所述位深度指示符及所述多个视图ID值。

19. 根据权利要求18所述的方法,其中由所述位深度指示符指示的所述位数在1与16之间。

20. 根据权利要求18所述的方法,其进一步包括基于可缩放性掩码索引来确定一个或多个视图次序索引。

21. 一种包括指令的非暂时性计算机可读媒体,所述指令当在包括计算硬件的处理器上执行时致使所述处理器:

接收指示用于在位流中用信号通知多个视图识别符ID值的每一者的数目的位深度指示符,其中所述多个视图ID值中的每一者与待解码的多个视图中的一者相关联;

基于所述位深度指示符,接收所述位流中的所述多个视图ID值以作为经索引的列表,所述经索引的列表具有对应于所述多个视图ID值的多个视图索引值,所述多个视图ID值的每一视图ID值指示了语法元素,所述语法元素的值使用由所述位深度指示符指示的位数来指定所述多个视图中的对应一者的视图ID,及

基于所接收的所述位流中的所述多个视图ID值来解码所述多个视图,

其中在视频参数集VPS中接收所述位深度指示符及所述多个视图ID值。

22. 根据权利要求21所述的计算机可读媒体,其中由所述位深度指示符指示的所述位数在1与16之间。

23. 根据权利要求21所述的计算机可读媒体,其中所述指令进一步致使所述处理器基于可缩放性掩码索引来确定一个或多个视图次序索引。

24. 一种经配置以译码视频信息的设备,所述设备包括:

用于接收指示用于在位流中用信号通知多个视图识别符ID值的每一者的位的数目的位深度指示符的装置,其中所述多个视图ID值中的每一者与待解码的多个视图中的一者相关联;

用于基于所述位深度指示符,接收所述位流中的所述多个视图ID值以作为经索引的列表的装置,所述经索引的列表具有对应于所述多个视图ID值的多个视图索引值,所述多个视图ID值的每一视图ID值指示了语法元素,所述语法元素的值使用由所述位深度指示符指示的位数来指定所述多个视图中的对应一者的视图ID,及

用于基于所接收的所述位流中的所述多个视图ID值来解码所述多个视图的装置,

其中在视频参数集VPS中接收所述位深度指示符及所述多个视图ID值。

25. 根据权利要求24所述的设备,其中由所述位深度指示符指示的所述位数在1与16之间。

26. 根据权利要求24所述的设备,其进一步包括用于基于可缩放性掩码索引来确定一个或多个视图次序索引的装置。

## 在参数集中用信号通知视图ID位深度

### 技术领域

[0001] 本发明与视频译码和压缩的领域有关。明确地说,其与可缩放视频译码(SVC)有关,包含用于高级视频译码(AVC)的SVC以及用于高效率视频译码(HEVC)的SVC,其还被称作可缩放HEVC(SHVC)。其还涉及3D视频译码,例如HEVC的多视图扩展(被称作MV-HEVC和3D-HEVC)。各种实施例涉及用于在视频参数集(VPS)中用信号通知视图识别符(ID)位深度、用信号通知位速率信息及/或图片速率信息的系统和方法及相关过程。

### 背景技术

[0002] 数字视频能力可以并入到多种多样的装置中,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(所谓的“智能电话”)、视频电话会议装置、视频串流装置及其类似者。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、目前正在开发的高效率视频译码(HEVC)标准及此类标准的扩展中所描述的视频译码技术。视频装置可通过实施这些视频译码技术而更有效地发射、接收、编码、解码和/或存储数字视频信息。

[0003] 视频译码技术包含空间(图片内)预测和/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来说,视频切片(即,视频帧或视频帧的一部分)可以分割成视频块,视频块还可被称作树块、译码单元(CU)及/或译码节点。使用关于同一图片中的相邻块中的参考样本的空间预测编码图片的经帧内译码(I)切片中的视频块。图片的经帧间编码(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。图片可被称作帧,且参考图片可被称作参考帧。

[0004] 空间或时间预测导致待译码块的预测性块。残差数据表示待译码原始块与预测块之间的像素差。经帧间译码块是根据指向形成预测块的参考样本块的运动向量和指示经译码块与预测块之间的差的残差数据编码的。经帧内译码块是根据帧内译码模式和残差数据而编码。为了进一步压缩,可将残差数据从像素域变换到变换域,从而产生残差变换系数,接着可以对残差变换系数进行量化。可扫描一开始按二维阵列排列的经量化变换系数,以便产生变换系数的一维向量,且可应用熵译码以实现更多压缩。

### 发明内容

[0005] 一般来说,本发明描述与可缩放视频译码(SVC)相关的技术。下文描述的各种技术提供用于在VPS中用信号通知视图ID位深度且用信号通知位速率信息及/或图片速率信息的方法及装置。

[0006] 根据某些方面的用于编码视频信息的设备包含存储器及计算硬件。存储器经配置以存储视频信息。所述计算硬件经配置以确定用信号通知的一或多个视图识别符的位深

度,其中所述一或多个视图识别符中的每一者与待编码的层相关联。所述计算硬件进一步经配置以在位流中用信号通知所述一或多个视图识别符的位深度。

[0007] 根据某些方面的用于解码视频信息的设备包含存储器及计算硬件。存储器经配置以存储视频信息。所述计算硬件经配置以接收用于用信号通知指示一或多个视图识别符值的位的数目的位深度指示符,其中所述一或多个视图识别符值中的每一者与待解码的一或多个层中的一者相关联。所述计算硬件进一步经配置以接收所述一或多个视图识别符值中的每一者以作为具有所指示的位数目的值。

[0008] 根据某些方面的用于译码视频信息的设备包含存储器及计算硬件。存储器经配置以存储视频信息。所述计算硬件经配置以处理以下各者中的至少一者:指示一或多个层集合的至少一个子层是否具有用信号通知的位速率信息的第一信号,或指示所述一或多个层集合的至少一个子层是否具有用信号通知的图片速率信息的第二信号。

[0009] 在附图及以下描述中阐述一或多个实例的细节,其并不希望限制本文所述的发明性概念的完整范围。其它特征、目标和优点将从所述描述和图式以及权利要求书而显而易见。

## 附图说明

[0010] 在整个图式中,参考数字可再使用以指示参考元件之间的对应关系。提供图式以说明本文中描述的实例实施例,并且不希望限制本发明的范围。

[0011] 图1是说明可利用根据本发明中描述的方面的技术的实例视频编码及解码系统的框图。

[0012] 图2A为说明可实施根据本发明中描述的方面的技术的视频编码器的实例的框图。

[0013] 图2B是说明可实施根据本发明中描述的方面的技术的视频编码器的实例的框图。

[0014] 图3A为说明可实施本发明中描述的方面的技术的视频解码器的实例的框图。

[0015] 图3B是说明可实施根据本发明中描述的方面的技术的视频解码器的实例的框图。

[0016] 图4是说明根据本发明的方面的用于用信号通知视图ID位深度的方法的流程图。

[0017] 图5是说明根据本发明的方面的用于解码视图ID位深度的方法的流程图。

[0018] 图6是说明根据本发明的方面的用于在VPS中用信号通知位速率信息及/或图片速率信息的方法的流程图。

## 具体实施方式

[0019] 本发明中描述的技术一般涉及可缩放视频译码(SHVC,SVC)和多视图/3D视频译码(例如,多视图译码加上深度,MVC+D)。举例来说,所述技术可与高效率视频译码(HEVC)可缩放视频译码(SVC,有时称为SHVC)扩展相关,且与其一起或在其内使用。在SHVC、SVC扩展中,可存在多个视频信息层。视频信息的最低层级处的层可充当基础层(BL)或参考层(RL),且视频信息的最顶部(或最高层)处的层可充当增强型层(EL)。“增强型层”有时被称作“增强层”,且这些术语可互换地使用。“基础层”有时被称作“参考层”(RL),且这些术语也可互换地使用。基础层与顶层之间的所有层可充当额外EL和/或参考层。举例来说,给定层可为用于所述给定层下方(例如,在前面)的层(例如基础层或任何介入增强层)的EL。另外,给定层还可充当用于给定层上方(例如,在后面)的一或多个增强层的RL。基础层(例如,具有例如

设定为或等于“1”的层识别(ID)的最低层)与顶层(或最高层)之间的任何层可由比给定层高的层用作层间预测的参考,且可使用比给定层低的层作为用于层间预测的参考。举例来说,可使用比给定层低的层作为用于层间预测的参考来确定给定层。

[0020] 为简单起见,根据仅两个层来呈现实例:BL及EL;然而,应很好地理解,下文描述的构想和实施例也适用于具有多个层的情况。此外,为了易于解释,常常使用术语“帧”或“块”。然而,这些术语不打算具有限制性。举例来说,下文描述的技术可与多种视频单元中的任一者一起使用,所述视频单元包含但不限于像素、块(例如,CU、PU、TU、宏块等)、切片、帧、图片等。

#### [0021] 视频译码

[0022] 视频译码标准包含ITU-T H.261、ISO/IEC MPEG-1 Visual、ITU-T H.262或ISO/IEC MPEG-2 Visual、ITU-T H.263、ISO/IEC MPEG-4 Visual及ITU-T H.264(也被称作ISO/IEC MPEG-4 AVC),包含其可缩放视频译码(SVC)及多视图视频译码(MVC)及多视图译码加上深度(MVC+D)扩展。可从[http://phenix.int-evry.fr/jct/doc\\_end\\_user/documents/13\\_Incheon/wg11/JCTVC-M0432-v3.zip](http://phenix.int-evry.fr/jct/doc_end_user/documents/13_Incheon/wg11/JCTVC-M0432-v3.zip)获得最新的HEVC草案规范(且下文中被称作HEVC WD10)。HEVC的多视图扩展(即MV-HEVC)也正由JCT-3V开发。可从[http://phenix.it-sudparis.eu/jct2/doc\\_end\\_user/documents/4\\_Incheon/wg11/JCT3V-D1004-v4.zip](http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/4_Incheon/wg11/JCT3V-D1004-v4.zip)获得下文中的MV-HEVC WD4的最近工作草案(WD)。被称为SHVC的对HEVC的可缩放扩展也正由JCT-VC开发。可从[http://phenix.it-sudparis.eu/jct/doc\\_end\\_user/documents/13\\_Incheon/wg11/JCTVC-M1008-v3.zip](http://phenix.it-sudparis.eu/jct/doc_end_user/documents/13_Incheon/wg11/JCTVC-M1008-v3.zip)获得SHVC的最近工作草案(WD),且在下文被称作工作草案2。根据一个方面,JCT3V-D0196([http://phenix.it-sudparis.eu/jct2/doc\\_end\\_user/documents/4\\_Incheon/wg11/JCT3V-D0196-v1.zip](http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/4_Incheon/wg11/JCT3V-D0196-v1.zip))包含用于在视频参数集(VPS)中用信号通知视图ID的方法。根据一个方面,JCTVC-K0125([http://phenix.int-evry.fr/jct/doc\\_end\\_user/documents/11\\_Shanghai/wg11/JCTVC-K0125-v1.zip](http://phenix.int-evry.fr/jct/doc_end_user/documents/11_Shanghai/wg11/JCTVC-K0125-v1.zip))包含用于在VPS中用信号通知位速率及图片速率信息的方法。

[0023] 可缩放视频译码(SVC)可用于提供质量(也称为信噪比(SNR))可缩放性、空间可缩放性和/或时间可缩放性。举例来说,在一个实施例中,参考层(例如,基础层)包含足以在第一质量水平下显示视频的视频信息,且增强层包含相对于所述参考层的额外视频信息,以使得所述参考层和所述增强层一起包含足以在高于所述第一水平的第二质量水平(例如,更少的噪声、更大的分辨率、更好的帧速率等)下显示视频的视频信息。增强型层可具有与基础层不同的空间分辨率。例如,EL与BL之间的空间纵横比可为1.0、1.5、2.0或垂直及水平方向上的其它不同比率。换句话说,EL的空间纵横尺寸可等于BL的空间纵横尺寸的1.0、1.5或2.0倍。在一些实例中,EL的缩放因子可大于BL。举例来说,EL中的图片的大小可大于BL中的图片的大小。以此方式,可有可能(但不限于)EL的空间分辨率大于BL的空间分辨率。

[0024] 在涉及用于H.264的SVC扩展或用于H.265的SHVC扩展(如上文所论述)的SVC中,当前块的预测可使用为SVC提供的不同层来执行。此预测可称作层间预测。在SVC中可利用层间预测方法以便减少层间冗余。层间预测的一些实例可包含层间帧内预测、层间运动预测和层间残差预测。层间帧内预测使用基础层中的位于同一地点的块的重构来预测增强层中的当前块。层间运动预测使用基础层的运动信息(包含运动向量)来预测增强层中的运动。层间残差预测使用基础层的残差来预测增强层的残差。

**[0025] 概述**

**[0026]** 在MV-HEVC及SHVC的早期版本(例如,工作草案2)中,使用固定数目的位用信号通知层的视图ID。例如,SHVC的早期版本使用10个位用信号通知视图ID,而不管多少视图可用于一层。然而,使用固定位深度(例如,固定数目的位)用信号通知视图ID可导致低效,尤其当可用的视图的数目与可使用10个位用信号通知的视图的数目相比相对小(例如,1到4个视图)时。使用10个位的固定位深度可允许用于一层的至多1024 ( $2^{10}$ ) 视图的信令,但在大多数情况下,用于一层的总视图的数目远小于1024个视图。

**[0027]** 此外,在MV-HEVC及SHVC的早期版本中,针对每一层集合及每一层集合的每一子层而用信号通知位速率信息及图片速率信息。对于每一层集合,用信号通知包含位速率信息及图片速率信息的语法结构(例如,bit\_rate\_pic\_rate())。对于层集合的每一子层,用信号通知指示是否存在位速率信息的旗标,且用信号通知指示是否存在图片速率信息的旗标。即使可能不存在任何层集合或子层的任何位速率信息及图片速率信息也执行此过程。例如,用信号通知0以作为每一层集合及每一子层的旗标的值,以便指示全部层集合及子层不存在位速率信息及/或图片速率信息。此可例如导致低效及不必要的信令及/或处理,这是因为可存在许多层集合及所述层集合的许多子层。

**[0028]** 为了解决这些及其它挑战,本发明的技术可用信号通知视图ID的位深度,且使用经由所述位深度用信号通知的位的数目来用信号通知视图ID的值。例如,如果仅使用2个视图,那么视图ID的位深度可为1位,且可使用1位用信号通知视图ID的值。如果使用3个视图,那么视图ID的位深度可为2位,且可使用2位用信号通知视图ID的值。用于用信号通知视图ID的位深度可为变量(例如,在1位与16位之间)。以此方式,可在大多数情况下通过减少用于用信号通知视图ID的值的位数目而更有效地用信号通知视图ID;更可能的是,用信号通知的视图的数目小于可使用当前固定数目的位(例如,10位)用信号通知的最大数目。

**[0029]** 所述技术也可在VPS中用信号通知位速率信息及/或图片速率信息。例如,所述技术可指示任何层集合及/或层集合的任何子层是否具有位速率信息及/或图片速率信息。在一个实施例中,所述技术可在VPS中用信号通知指示任何层集合及/或子层是否存在位速率信息的全局旗标,且在VPS中用信号通知指示任何层集合及/或子层是否存在图片速率信息的全局旗标。通过在VPS中包含所述全局旗标,可仅当所述全局旗标指示至少一个层集合或层集合的子层存在位速率信息及/或图片速率信息时才用信号通知及/或存取位速率图片速率语法结构。如果全局旗标指示任何层集合不存在位速率信息及/或图片速率信息,那么不需要用信号通知及/或存取位速率图片速率语法结构,且不需要发送(例如,用信号通知)用于位速率图片速率语法结构内的个别层集合的个别子层的旗标。另外,全局旗标可允许以有效方式对位速率信息及图片速率信息的单独的处理。如果用于位速率信息的全局旗标指示至少一个层集合中不存在位速率信息,那么不需要用信号通知及/或处理用于任何子层的位速率信息的相应的旗标。类似地,如果用于图片速率信息的全局旗标指示至少一个层集合不存在图片速率信息,那么不需要用信号通知及/或处理用于任何子层的图片速率信息的相应的旗标。虽然先前还可能用信号通知位速率信息或图片速率信息中的仅一者,但需要用信号通知及处理多个个别旗标:用于每一层集合的每一子层的位速率信息的一个旗标,及用于每一层集合的每一子层的图片速率信息的一个旗标。

**[0030]** 下文参考附图更充分地描述新颖系统、设备及方法的各个方面。然而,本发明可以



许多不同形式来体现,且不应将其解释为限于贯穿本发明所呈现的任何特定结构或功能。而是,提供这些方面以使得本发明将为透彻且完整的,并且将向所属领域的技术人员充分传达本发明的范围。基于本文中的教导,所属领域的技术人员应了解,本发明的范围既定涵盖无论是独立于本发明的任何其它方面而实施还是与之组合而实施的本文中所揭示的新颖系统、设备及方法的任何方面。举例来说,可使用本文中所阐述的任何数目个方面来实施设备或实践方法。此外,本发明的范围既定涵盖使用作为本文中所陈述的本发明的各种方面的补充或替代的其它结构、功能性或结构与功能性来实践的此设备或方法。应理解,可通过权利要求的一或多个要素来体现本文中所揭示的任何方面。

[0031] 尽管本文描述了特定方面,但这些方面的许多变化及排列属于本发明的范围内。尽管提及了优选方面的一些益处及优点,但本发明的范围不欲限于特定益处、用途或目标。而是,本发明的方面既定广泛地适用于不同无线技术、系统配置、网络及发射协议,其中的一些是借助于实例而在图中以及在优选方面的以下描述中说明。具体实施方式和图式仅说明本发明,而不是限制由所附权利要求书和其等效物界定的本发明的范围。

#### [0032] 视频译码系统

[0033] 图1是说明可利用根据本发明中所描述的方面的技术的实例视频译码系统10的框图。如本文中所描述地使用,术语“视频译码器”一般指代视频编码器和视频解码器两者。在本发明中,术语“视频译码”或“译码”可一般指代视频编码和视频解码。

[0034] 如图1中所示,视频译码系统10包含源装置12和目的地装置14。源装置12产生经编码视频数据。目的地装置14可解码由源装置12产生的经编码视频数据。源装置12可经由通信信道16将视频数据提供到目的地装置14,通信信道16可包含计算机可读存储媒体或其它通信信道。源装置12及目的地装置14可包含广泛范围的装置,包含桌上型计算机、笔记本(例如,膝上型计算机)、计算机、平板计算机、机顶盒、电话手持机(例如所谓的“智能”电话、所谓的“智能”垫)、电视、相机、显示装置、数字媒体播放器、视频游戏控制台、车载计算机、视频流式传输装置或类似者。源装置12及目的地装置14可经装备以用于无线通信。

[0035] 目的地装置14可通路通信信道16接收待解码的经编码视频数据。通信信道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的媒体或装置类型。举例来说,通信信道16可包括通信媒体以使源装置12能够实时地将经编码视频数据直接传输到目的地装置14。可根据通信标准(例如,无线通信协议)调制经编码视频数据,并将其发射到目的地装置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网路(例如,局域网、广域网或全球网路,例如因特网)的部分。通信媒体可包含路由器、交换器、基站或任何其它可用于促进从源装置12到目的地装置14的通信的设备。

[0036] 在一些实施例中,经编码数据可从输出接口22输出到存储装置。在这些实施例中,信道16可对应于存储由源装置12产生的经编码视频数据的存储装置或计算机可读存储媒体。举例来说,目的地装置14可经由磁盘存取或卡存取来存取计算机可读存储媒体。类似地,可由输入接口28从计算机可读存储媒体存取经编码数据。计算机可读存储媒体可包含多种分布式或本地存取的数据储存媒体中的任一者,例如,硬盘驱动器、蓝光光盘、DVD、CD-ROM、闪存、易失性或非易失性存储器,或用于存储视频数据的其它数字存储媒体。计算机可读存储媒体可对应于可存储由源装置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计算机可读存储媒体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并且将经编码视频数据发射到目的地装置14的一类型的服务器。实例文件服务器包含网络服务器(例如,用于网站)、FTP服务器、网络连接存储(NAS)装置或本地磁盘驱动器。目的地装置14可通过标准数据连接(包含因特网连接)来存取经编码视频数据。此可包含无线通道(例如,Wi-Fi连接)、有线连接(例如,DSL、电缆调制解调器等),或适合于存取存储在文件服务器上的经编码视频数据的两者的组合。经编码视频数据从计算机可读存储媒体的传输可为流式传输、下载传输或两者的组合。

[0037] 本发明的技术可应用除无线应用或环境之外的应用或环境。所述技术可应用于视频译码以支持多种多媒体应用,例如空中协议电视广播、有线电视发射、卫星电视发射、因特网串流视频发射(例如动态自适应HTTP串流(DASH))、经编码到数据存储媒体上的数字视频,存储在数据存储媒体上的数字视频的解码,或其它应用。在一些实施例中,系统10可经配置以支持单向或双向视频传输,以支持例如视频流式传输、视频回放、视频广播和/或视频电话的应用。

[0038] 在图1中,源装置12包含视频源18、视频编码器20和输出接口22。目的地装置14包含输入接口28、视频解码器30和显示装置32。源装置12的视频编码器20可经配置以应用遵守多个标准或标准扩展的用于译码包含视频数据的位流的技术。在其它实施例中,源装置及目的地装置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(例如外部相机)接收视频数据。同样,目的地装置14可与外部显示装置介接,而非包含集成式显示装置。

[0039] 源装置12的视频源18可包含视频俘获装置,例如摄像机、含有先前所俘获视频的视频存档和/或用于从视频内容提供者接收视频的视频馈送接口。视频源18可产生基于计算机图形的数据(作为源视频),或实况视频、所存档视频和计算机产生的视频的组合。在一些实施例中,如果视频源18是摄像机,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。所俘获、预俘获或计算机产生的视频可由视频编码器20编码。经编码视频信息可由输出接口22输出到通信信道16,所述通信信道可包含计算机可读存储媒体,如上文所述。

[0040] 计算机可读存储媒体可包含瞬时媒体,例如无线广播或有线网络发射,或存储媒体(例如,非暂时性存储媒体),例如硬盘、快闪驱动器、压缩光盘、数字视频光盘、蓝光光盘或其它计算机可读媒体。网络服务器(未图示)可从源装置12接收经编码视频数据,且将经编码视频数据提供到目的地装置14(例如,经由网络传输)。媒体生产设施(例如,光盘冲压设施)的计算装置可从源装置12接收经编码视频数据,且生产含有所述经编码视频数据的光盘。因此,可将通信信道16理解为包含一或多个各种形式的计算机可读存储媒体。

[0041] 目的地装置14的输入接口28可从通信信道16接收信息。通信信道16的信息可包含由视频编码器20定义的语法信息,所述语法信息可供视频解码器30使用,所述语法信息包含描述块和其它经译码单元(例如,GOP)的特性和/或处理的语法元素。显示装置32向用户显示经解码视频数据,且可包含多种显示装置中的任一者,例如,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。

[0042] 视频编码器20和视频解码器30可以根据一种视频译码标准(例如目前正在开发的高效视频译码(HEVC)标准)来操作,并且可以符合HEVC测试模型(HM)。或者,视频编码器20和视频解码器30可根据其它专有或业界标准来操作,所述标准例如是ITU-T H.264标准,也

被称为MPEG-4第10部分,高级视频译码(AVC),或此类标准的扩展。但是,本发明的技术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和ITU-T H.263。虽然在图1中未展示,但在一些方面中,视频编码器20和视频解码器30可各自与音频编码器及解码器集成,且可包含适当的多路复用器-多路分用器单元或其它硬件和软件以处置共同数据流或单独数据流中的音频和视频两者的编码。如果适用,则多路复用器-多路分用器单元可符合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。

[0043] 图1仅为实例,且本发明的技术可适用于不一定包含编码装置与解码装置之间的任何数据通信的视频译码环境(例如,视频编码或视频解码)。在其它实例中,数据可从本地存储器检索、经由网络流式传输或类似者。编码装置可编码数据且将数据存储到存储器,及/或解码装置可从存储器检索数据且解码数据。在许多实例中,由并不彼此通信而是仅编码数据到存储器和/或从存储器检索数据且解码数据的装置执行编码和解码。

[0044] 视频编码器20和视频解码器30各自可实施为多种合适的编码器和解码器电路中的任一者,例如一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、离散逻辑、软件、硬件固件或其任何组合。当部分地用软件实施所述技术时,装置可将用于软件的指令存储在非暂时性计算机可读媒体中且使用一或多个处理器用硬件执行所述指令以执行本发明的技术。视频编码器20和视频解码器30中的每一者可包含在一或多个编码器或解码器中,所述编码器或解码器中的任一者可集成为相应装置中的组合编码器/解码器(CODEC)的部分。包含视频编码器20和/或视频解码器30的装置可包括集成电路、微处理器和/或无线通信装置(例如,蜂窝式电话)。

[0045] JCT-VC正在努力开发HEVC标准。HEVC标准化努力是基于被称作HEVC测试模型(HM)的视频译码装置的演进模型。HM根据例如ITU-T H.264/AVC假设视频译码装置相对于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。

[0046] 一般来说,HM的工作模型描述视频帧或图片可以分成包含明度及色度样本两者的一连串树块或最大译码单元(LCU)。位流内的语法数据可界定最大译码单位(LCU,其是在像素数目方面的最大译码单位)的大小。切片包含按译码次序的若干连续树块。视频帧或图片可以被分割成一或多个切片。每一树块可以根据二叉树分裂成译码单元(CU)。一般来说,二叉树数据结构包含每个CU一个节点,其中一个根节点对应于所述树块。如果一个CU分裂成四个子CU,那么对应于CU的节点包含四个叶节点,其中的每一者对应于所述子CU中的一者。

[0047] 二叉树数据结构的每一节点可提供用于对应CU的语法数据。举例来说,二叉树中的节点可包含分裂旗标,其指示对应于所述节点的所述CU是否分裂成子CU。用于CU的语法元素可递归地定义,且可取决于CU是否分裂成数个子CU。如果CU不进一步分裂,那么将其称为叶CU。在本发明中,叶CU的四个子CU也将被称作叶CU,即使不存在原始叶CU的明确分裂时也是如此。举例来说,如果16x16大小的CU不进一步分裂,那么这四个8x8子CU也将被称作叶CU,虽然16x16CU从未分裂。

[0048] CU具有与H.264标准的宏块类似的目的,除了CU不具有大小区别。举例来说,树块可分裂成四个子节点(也称为子CU),且每一子节点又可为父节点且可分裂成另外四个子节点。最后的未经分裂的子节点(被称作二叉树的叶节点)包括译码节点,也称为叶CU。与译码位流相关联的语法数据可定义树块可分裂的最大次数,被称作最大CU深度,且还可定义

译码节点的最小大小。因此,位流还可界定最小译码单元(SCU)。本发明使用术语“块”来指HEVC的上下文中的CU、PU或TU中的任一者,或者其它标准的上下文中的类似数据结构(例如,其在H.264/AVC中的宏块和子块)。

[0049] CU包含译码节点和与所述译码节点相关联的预测单元(PU)和变换单元(TU)。CU的大小对应于译码节点的大小并且形状必须是正方形。CU的大小可介于 $8 \times 8$ 个像素至多达具有最大 $64 \times 64$ 个像素或更大的树块大小的范围内。每一CU可含有一或多个PU和一或多个TU。举例来说,与CU相关联的语法数据可描述CU分割成一或多个PU。分割模式可在CU被跳过或经直接模式编码、经帧内预测模式编码或经帧间预测模式编码之间有所不同。PU可以分割成非正方形形状。举例来说,与CU相关联的语法数据还可描述CU根据四叉树到一或多个TU的分割。TU可以是正方形或非正方形(例如,矩形)形状。

[0050] HEVC标准允许根据TU变换,TU可针对不同CU而有所不同。TU的大小通常是基于针对经分割LCU定义的给定CU内的PU的大小而确定,但是情况可能并不总是如此。TU通常与PU大小相同或小于PU。在一些实例中,可以使用被称为“残差四叉树”(RQT)的四叉树结构将对应于CU的残差样本细分成较小单元。RQT的叶节点可被称为变换单元(TU)。可变换与TU相关联的像素差值以产生变换系数,所述变换系数可经量化。

[0051] 叶CU可包含一或多个预测单元(PU)。一般来说,PU表示对应于相对应的CU的全部或一部分的空间区域,并且可包含用于检索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的数据可以包含在残差四叉树(RQT)中,残差四叉树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PU的一或多个运动向量的数据。举例来说,定义PU的运动向量的数据可以描述运动向量的水平分量、运动向量的垂直分量、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)、运动向量指向的参考图片及/或运动向量的参考图片列表(例如,列表0、列表1或列表C)。

[0052] 具有一或多个PU的叶CU还可包含一或多个变换单元(TU)。变换单元可使用RQT(也称为TU四叉树结构)来指定,如上文所论述。举例来说,分裂旗标可以指示叶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更多个子TU。当TU未进一步分裂时,其可被称作叶TU。总体上,对于帧内译码,所有属于一叶CU的叶TU共享相同的帧内预测模式。也就是说,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的残差值计算为CU的对应于TU的部分与原始块之间的差。TU不一定限于PU的大小。因此,TU可大于或小于PU。对于帧内译码,PU可与相同CU的对应叶TU并置。在一些实例中,叶TU的最大大小可以对应于对应的叶CU的大小。

[0053] 此外,叶CU的TU还可与相应的四叉树数据结构(被称作残差四叉树(RQT))相关联。即,叶CU可包含指示叶CU如何分割成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)。未经分裂的RQT的TU被称作叶TU。一般来说,除非另外提及,否则本发明分别使用术语CU及TU来指叶CU及叶TU。

[0054] 视频序列通常包含一系列视频帧或图片。图片群组(GOP)一般包括一系列一或多个视频图片。GOP可包含描述GOP中所包含的图片数目的GOP的标头中、图片中的一或多者的标头中或别处的语法数据。图片的每一切片可包含描述用于相应的切片的编码模式的切片语法数据。视频编码器20通常对个别视频切片内的视频块进行操作以便对视频数据进行编

码。视频块可与CU内的译码节点相对应。视频块可具有固定或变化的大小,并且根据指定译码标准可在大小上有所不同。

[0055] 作为实例, HM支持各种PU大小的预测。假设特定CU的大小为 $2N \times 2N$ , 那么HM支持 $2N \times 2N$ 或 $N \times N$ 的PU大小的帧内预测, 及 $2N \times 2N$ 、 $2N \times N$ 、 $N \times 2N$ 或 $N \times N$ 的对称PU大小的帧间预测。HM还支持用于在 $2N \times nU$ 、 $2N \times nD$ 、 $nL \times 2N$ 及 $nR \times 2N$ 的PU大小下的帧间预测的不对称分割。在不对称分割中, 不分割CU的一个方向, 但是将另一方向分割成25%和75%。CU的对应于25%分区的部分通过“n”后面跟着“上”、“下”、“左”或“右”的指示来指示。因此, 例如, “ $2N \times nU$ ”是指经水平地分割的 $2N \times 2N$  CU, 其中顶部为 $2N \times 0.5N$  PU, 而底部为 $2N \times 1.5N$  PU。

[0056] 在本发明中, “ $N \times N$ ”与“N乘N”可互换地使用以依据垂直和水平尺寸来指代视频块的像素尺寸, 例如,  $16 \times 16$ 像素或16乘16像素。一般来说,  $16 \times 16$ 块将在垂直方向上具有16个像素( $y=16$ ), 且在水平方向上具有16个像素( $x=16$ )。同样,  $N \times N$ 块总体上在垂直方向上具有N个像素, 并且在水平方向上具有N个像素, 其中N表示非负整数值。块中的像素可布置成行和列。此外, 块未必需要在水平方向与垂直方向上具有相同数目的像素。举例来说, 块可包括 $N \times M$ 个像素, 其中M未必等于N。

[0057] 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可以计算用于CU的TU的残差数据。PU可包括描述在空间域(也称为像素域)中产生预测像素数据的方法或模式的语法数据, 且TU可包括在应用变换之后变换域中的系数, 所述变换例如离散正弦变换(DST)、离散余弦变换(DCT)、整数变换、小波变换或与残差视频数据概念上类似的变换。残差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可以形成包含用于CU的残差数据的TU, 并且接着变换TU以产生用于CU的变换系数。

[0058] 在用于产生变换系数的任何变换之后, 视频编码器20可执行变换系数的量化。量化为既定具有其最广泛普通意义的广义术语。在一个实施例中, 量化指变换系数经量化以可能减少用以表示系数的数据量从而提供进一步压缩的过程。量化过程可减少与系数中的一些系数或全部相关联的位深度。举例来说,  $n$ 位值可在量化期间被舍去成 $m$ 位值, 其中 $n$ 大于 $m$ 。

[0059] 在量化之后, 视频编码器可扫描变换系数, 从包括经量化变换系数的二维矩阵产生一维向量。扫描可经设计以将较高能量(并且因此较低频率)的系数放置在阵列的前面, 并且将较低能量(并且因此较高频率)的系数放置在阵列的背面。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被熵编码的串行化向量。在其它实例中, 视频编码器20可以执行自适应扫描。在扫描经量化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可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码方法熵编码一维向量。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在解码视频数据时使用。

[0060] 为了执行CABAC, 视频编码器20可向待发射的符号指派上下文模型内的上下文。上下文可涉及(例如)符号的相邻值是否为非零。为了执行CAVLC, 视频编码器20可选择用于待发射的符号的可变长度码。VLC中的码字可经构建以使得相对较短码对应于更有可能符号, 而较长码对应于不太可能符号。以此方式, 使用VLC可例如实现优于针对待发射的每一符号使用等长度码字的位节省。概率确定可基于指派到符号的上下文。

[0061] 视频编码器20可进一步例如在帧标头、块标头、切片标头或GOP标头中将例如基于块的语法数据、基于帧的语法数据及基于GOP的语法数据等语法数据发送到视频解码器30。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。

#### [0062] 视频编码器

[0063] 图2A为说明可实施根据本发明中描述的方面的技术的视频编码器的实例的框图。视频编码器20可经配置以处理视频位流的单层(例如针对HEVC)。此外,视频编码器20可经配置以执行本发明的技术中的任一者或全部,包括但不限于上文及下文相对于图4到6更详细地描述的在VPS中用信号通知视图ID位深度、用信号通知位速率信息及/或图片速率信息的方法,及相关过程。作为一个实例,层间预测单元66(当提供时)可经配置以执行本发明中描述的技术中的任一者或全部。然而,本发明的方面不限于此。在一些实例中,本发明中描述的技术可在视频编码器20的各种组件之间共享。在一些实例中,另外或替代地,处理器(未图示)可经配置以执行本发明中描述的技术中的任一者或全部。

[0064] 出于解释的目的,本发明描述在HEVC译码的上下文中的视频编码器20。然而,本发明的技术可以适用于其它译码标准或方法。图2A的编码器20说明编解码器的单层。然而,如将关于图2B进一步描述,可复制视频编码器20中的一些或全部以用于根据多层编解码器进行处理。

[0065] 视频编码器20可执行视频切片内的视频块的帧内、帧间和层间预测(有时被称作帧内、帧间或层间译码)。帧内译码依赖于空间预测来减少或移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移除视频序列的邻近帧或图片内的视频中的时间冗余。层间译码依赖于基于同一视频译码序列内的不同层内的视频的预测。帧内模式(I模式)可指代若干基于空间的译码模式中的任一者。例如单向预测(P模式)或双向预测(B模式)等帧间模式可指代若干基于时间的译码模式中的任一者。

[0066] 如图2A中所展示,视频编码器20接收待编码的视频帧内的当前视频块。在图2A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和熵编码单元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、帧内预测单元46、层间预测单元66和分割单元48。参考帧存储器64可包含经解码图片缓冲器。经解码图片缓冲器是具有其一般含义的广义术语,且在一些实施例中涉及视频编解码器管理的参考帧的数据结构。

[0067] 为了视频块重构,视频编码器20还包含反量化单元58、反变换单元60和求和器62。还可包含解块滤波器(图2A中未图示)以便对块边界进行滤波,以从经重构视频移除成块效应假影。在需要时,解块滤波器将通常滤波求和器62的输出。除了解块滤波器外,还可使用额外滤波器(回路中或回路后)。为简洁起见未展示这些滤波器,但在必要时,这些滤波器可以对求和器50的输出进行滤波(作为环路内滤波器)。

[0068] 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。所述帧或切片可划分成多个视频块。运动估计单元42和运动补偿单元44可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。帧内预测单元46可替代地相对于与待译码块相同的帧或切片中的一或多个相邻块执行对所接收的视频块的帧内预测性译码以提供空间预测。视频编码器20可以执行多个译码遍次,例如,以针对每一视频数据块

选择适当的译码模式。

[0069] 此外,分割单元48可以基于前述译码遍次中的前述分割方案的评估将视频数据块分割成子块。举例来说,分割单元48可初始地将帧或切片分割为LCU,且基于速率失真分析(例如,速率失真优化等)将LCU中的每一者分割为子CU。模式选择单元40可以进一步产生指示LCU分割成子CU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。

[0070] 模式选择单元40可例如基于误差结果选择译码模式(帧内、帧间或层间预测模式)中的一者,且将所产生的帧内、帧间或层间经译码块提供到求和器50来产生残差块数据并提供到求和器62来重构经编码块以用作参考帧。模式选择单元40还将语法元素(例如运动向量、帧内模式指示符、分割信息和其它此类语法信息)提供到熵编码单元56。

[0071] 运动估计单元42和运动补偿单元44可高度集成,但出于概念的目的分别加以说明。由运动估计单元42执行的运动估计是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置。预测性块是被发现在像素差方面与待译码块紧密匹配的块,像素差可通过绝对差总和(SAD)、平方差总和(SSD)或其它差度量来确定。在一些实例中,视频编码器20可计算存储于参考帧存储器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可相对于全像素位置和分数像素位置执行运动搜索并且输出具有分数像素精度的运动向量。

[0072] 运动估计单元42通过比较PU的位置与参考图片的预测性块的位置来计算用于经帧间译码切片中的视频块的PU的运动向量。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),其中的每一者识别存储在参考图片存储器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和运动补偿单元44。

[0073] 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确定的运动向量来获取或产生预测性块。在一些实例中,运动估计单元42与运动补偿单元44可在功能上整合。在接收到当前视频块的PU的运动向量后,运动补偿单元44可在参考图片列表中的一者中定位所述运动向量指向的预测块。求和器50通过从正经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差视频块,如下文所论述。在一些实施例中,运动估计单元42可相对于明度分量执行运动估计,且运动补偿单元44可针对色度分量和明度分量两者使用基于明度分量计算的运动向量。模式选择单元40可产生与视频块和视频切片相关联的语法元素供视频解码器30在解码视频切片的视频块时使用。

[0074] 作为如上文所描述由运动估计单元42和运动补偿单元44执行的帧间预测的替代方案,帧内预测单元46可帧内预测或计算当前块。确切地说,帧内预测单元46可以确定用来对当前块进行编码的帧内预测模式。在一些实例中,帧内预测单元46可以例如在分开的编码编次期间使用各种帧内预测模式对当前块进行编码,并且帧内预测单元46(或在一些实例中为模式选择单元40)可以从所述测试模式中选择适当帧内预测模式来使用。

[0075] 举例来说,帧内预测单元46可使用速率-失真分析计算针对各种测试的帧内预测模式的速率-失真值,且从所述测试的模式当中选择具有最好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,以及用于产生经编码块的位速率(也就是说,位数目)。帧内预测单元46可根据用于各种经编码块的失真和速率计算比率,以确定哪个帧内预测模式对于所述块展现最佳速率-失真值。

[0076] 在选择用于块的帧内预测模式后,帧内预测单元46可将指示用于块的选定帧内预测模式的信息提供到熵编码单元56。熵编码单元56可编码指示选定帧内预测模式的信息。视频编码器20在所发射的位流中可包含配置数据,其可包含多个帧内预测模式索引表和多个经修改的帧内预测模式索引表(也称为码字映射表),对用于各种块的上下文进行编码的定义,以及对最可能帧内预测模式、帧内预测模式索引表和经修改的帧内预测模式索引表的指示以用于所述上下文中的每一者。

[0077] 视频编码器20可包含层间预测单元66。层间预测单元66经配置以使用SVC中可用的一或多个不同层(例如,基础或参考层)来预测当前块(例如,EL中的当前块)。此预测可称作层间预测。层间预测单元66利用预测方法以减少层间冗余,进而提高译码效率且降低计算资源要求。层间预测的一些实例可包含层间帧内预测、层间运动预测及层间残差预测。层间帧内预测使用基础层中的位于同一地点的块的重构来预测增强层中的当前块。层间运动预测使用基础层的运动信息来预测增强层中的运动信息。层间残差预测使用基础层的残差来预测增强层的残差。当基础层与增强层具有不同的空间分辨率时,可由层间预测单元66执行使用时间缩放功能的空间运动向量缩放及/或层间位置映射,如下文更详细地描述。

[0078] 视频编码器20通过从正译码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差视频块。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)或概念上类似的变换等变换应用于残差块,从而产生包括残差变换系数的视频块。变换处理单元52可以执行概念上类似于DCT的其它变换。举例来说,还可使用离散正弦变换(DST)、小波变换、整数变换、子带变换或其它类型的变换。

[0079] 变换处理单元52可将变换应用于残差块,从而产生残差变换系数的块。所述变换可将残差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。量化单元54量化所述变换系数以进一步减小位速率。量化过程可减少与系数中的一些系数或全部相关联的位深度。可通过调节量化参数来修改量化程度。在一些实例中,量化单元54可以接着执行对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。

[0080] 在量化之后,熵编码单元56对经量化变换系数进行熵编码。举例来说,熵编码单元56可执行上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(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,上下文可以基于相邻块。在由熵编码单元56进行熵译码之后,可将经编码位流发射到另一装置(例如,视频解码器30),或者将所述经编码位流存档以用于稍后发射或检索。

[0081] 反量化单元58和反变换单元60分别应用反量化和反变换以在像素域中重构残差块(例如,用于稍后用作参考块)。运动补偿单元44可通过将残差块添加到参考帧存储器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经重构的残差块来计算用于在运动估计中使用的子整数像素值。求和器62将经重构的残差块添加到由运动补偿单元44产生的经运动补偿的预测块,以产生经重构的视频块以用于



存储于参考帧存储器64中。经重构的视频块可由运动估计单元42和运动补偿单元44用作参考块以对后续视频帧中的块进行帧间译码。

#### [0082] 多层视频编码器

[0083] 图2B是说明可实施根据本发明中描述的方面的技术的多层视频编码器21的实例的框图。视频编码器21可经配置以处理多层视频帧,例如,用于SHVC和多视图译码。此外,视频编码器21可经配置以执行本发明的技术中的任一者或全部。

[0084] 视频编码器21包含视频编码器20A和视频编码器20B,其中的每一者可经配置为图2A的视频编码器20且可执行上文关于视频编码器20所描述的功能。此外,如再使用参考数字所指示,视频编码器20A和20B可包含系统及子系统至少一些作为视频编码器20。虽然将视频编码器21说明为包含两个视频编码器20A和20B,但视频编码器21不受如此限制且可包含任何数目的视频编码器20层。在一些实施例中,视频编码器21可针对存取单元中的每一图片或帧包含视频编码器20。例如,包含五个图片的存取单元可由包含五个编码器层的视频编码器处理或编码。在一些实施例中,视频编码器21可包含比存取单元中的帧多的编码器层。在一些此类情况下,当处理一些存取单元时,一些视频编码器层可能不在作用中。

[0085] 除了视频编码器20A和20B外,视频编码器21还可包含再取样单元90。在一些情况下,再取样单元90可上取样所接收的视频帧的基础层以(例如)创建增强层。再取样单元90可对与帧的所接收的基础层相关联的特定信息上取样,但不对其余信息上取样。举例来说,再取样单元90可上取样基础层的空间大小或像素数目,但切片的数目或图片次序计数可保持恒定。在一些情况下,再取样单元90可不处理所接收的视频及/或可为任选的。例如,在一些情况下,模式选择单元40可执行上取样。在一些实施例中,再取样单元90经配置以对层进行上取样且重新组织、重新界定、修改或调整一或多个切片以符合一组切片边界规则及/或光栅扫描规则。虽然主要描述为对基础层或存取单元中的较低层进行上取样,但在一些情况下,再取样单元90可对层进行下取样。例如,如果在视频的流式传输期间减小带宽,那么可对帧进行下取样而不是上取样。再取样单元90还可进一步经配置以执行裁剪和/或填补操作。

[0086] 再取样单元90可经配置以从较低层编码器(例如,视频编码器20A)的经解码图片缓冲器114接收图片或帧(或与图片相关联的图片信息)且对所述图片(或所接收的图片信息)上取样。接着可将经上取样图片提供到较高层编码器(例如,视频编码器20B)的模式选择单元40,所述较高层编码器经配置以编码与较低层编码器相同的存取单元中的图片。在一些情况下,较高层编码器为从较低层编码器去除的一个层。在其它情况下,在图2B的层0编码器与层1编码器之间可存在一或多个较高层编码器。

[0087] 在一些情况下,可省略或绕过再取样单元90。在此些情况下,来自视频编码器20A的经解码图片缓冲器64的图片可直接提供到视频编码器20B的模式选择单元40,或至少不提供到再取样单元90。例如,如果提供到视频编码器20B的视频数据及来自视频编码器20A的经解码图片缓冲器64的参考图片具有相同的大小或分辨率,那么可将参考图片提供到视频编码器20B而没有任何再取样。

[0088] 在一些实施例中,视频编码器21使用下取样单元94下取样待提供到较低层编码器的视频数据,随后将所述视频数据提供到视频编码器20A。替代地,下取样单元94可为能够

对视频数据上取样或下取样的再取样单元90。在另外其它实施例中,可省略下取样单元94。

[0089] 如图2B中所说明,视频编码器21可进一步包含多路复用器98或mux。mux 98可从视频编码器21输出组合的位流。可通过从视频编码器20A和20B中的每一者取得位流且交替在给定时间输出哪一位流来创建组合的位流。虽然在一些情况下,可一次一个位地交替来自两个(或在两个以上视频编码器层的情况下,两个以上)位流的位,但在许多情况下,不同地组合所述位流。举例来说,可通过一次一个块地交替选定位流来产生输出位流。在另一实例中,可通过从视频编码器20A和20B中的每一者输出非1:1比率的块来产生输出位流。举例来说,可从视频编码器20B输出用于从视频编码器20A输出的每一块的两个块。在一些实施例中,可预编程来自mux 98的输出流。在其它实施例中,mux 98可基于从视频编码器21外部的系统(例如从源装置12上的处理器)接收的控制信号而组合来自视频编码器20A、20B的位流。可基于来自视频源18的视频的分辨率或位速率、基于信道16的带宽、基于与用户相关联的预订(例如,付费预订与免费预订)或基于用于确定来自视频编码器21的所要分辨率输出的任何其它因素而产生控制信号。

#### [0090] 视频解码器

[0091] 图3A是说明可实施本发明中描述的方面的技术的视频解码器的实例的框图。视频解码器30可经配置以处理视频位流的单层(例如,针对HEVC)。此外,视频解码器30可经配置以执行本发明的技术中的任一者或全部,包括但不限于在VPS中用信号通知视图ID位深度且用信号通知位速率信息及/或图片速率信息的方法,其在上文及下文相对于图4到6更详细地描述。作为一个实例,层间预测单元75可经配置以执行本发明中描述的技术中的任一者或全部。然而,本发明的方面不限于此。在一些实例中,本发明中描述的技术可在视频解码器30的各种组件之间共享。在一些实例中,另外或替代地,处理器(未图示)可经配置以执行本发明中描述的技术中的任一者或全部。

[0092] 出于解释的目的,本发明描述在HEVC译码的背景下的视频解码器30。然而,本发明的技术可以适用于其它译码标准或方法。图3A的解码器30说明编解码器的单层。然而,如将关于图3B进一步描述,可复制视频编码器30中的一些或全部以用于根据多层编解码器进行处理。

[0093] 在图3A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、帧内预测单元74、层间预测单元75、反量化单元76、反变换单元78、参考帧存储器82及求和器80。在一些实施例中,运动补偿单元72和/或帧内预测单元74可经配置以执行层间预测,在此情况下,可省略层间预测单元75。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(图2A)描述的编码遍次互逆的解码遍次)。运动补偿单元72可基于从熵解码单元70接收的运动向量产生预测数据,而帧内预测单元74可基于从熵解码单元70接收的帧内预测模式指示符产生预测数据。参考帧存储器82可包含经解码图片缓冲器。经解码图片缓冲器是具有其一般含义的广义术语,且在一些实施例中涉及视频编解码器管理的参考帧的数据结构。

[0094] 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和相关联的语法元素的经编码视频位流。视频解码器30的熵解码单元70对位流熵解码以产生经量化系数、运动向量或帧内预测模式指示符和其它语法元素。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可在视频切片层级及/或视频块层级接收语法元素。

[0095] 当视频切片经译码为经帧内译码(I)切片时,帧内预测单元74可以基于用信号通知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用于当前视频切片的视频块的预测数据。在视频帧经译码为经帧间译码(例如,B、P或GPB)切片时,运动补偿单元72基于从熵解码单元70接收的运动向量和其它语法元素产生用于当前视频切片的视频块的预测块。可从参考图片列表中的一者内的参考图片中的一者产生预测性块。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建技术构建参考帧列表——列表0和列表1。

[0096] 运动补偿单元72通过解析运动向量和其它语法元素确定用于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用于经解码的当前视频块的预测块。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确定用于对视频切片的视频块译码的预测模式(例如,帧内预测或帧间预测)、帧间预测切片类型(例如,B切片、P切片或GPB切片)、切片的参考图片列表中的一或多者的构建信息、切片的每一经帧间编码的视频块的运动向量、切片的每一经帧间译码的视频块的帧间预测状态,以及用以解码当前视频切片中的视频块的其它信息。

[0097] 运动补偿单元72还可以基于内插滤波器执行内插。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可根据接收的语法元素而确定由视频编码器20使用的内插滤波器并使用所述内插滤波器来产生预测性块。

[0098] 视频解码器30还可包含层间预测单元75。层间预测单元75经配置以使用SVC中可用的一或多个不同层(例如,基础或参考层)预测当前块(例如,EL中的当前块)。此预测可称作层间预测。层间预测单元75利用预测方法减少层间冗余,进而提高译码效率且降低计算资源要求。层间预测的一些实例可包含层间帧内预测、层间运动预测及层间残差预测。层间帧内预测使用基础层中的位于同一地点的块的重构来预测增强层中的当前块。层间运动预测使用基础层的运动信息来预测增强层中的运动。层间残差预测使用基础层的残差来预测增强层的残差。当基础及增强层具有不同的空间分辨率时,可通过层间预测单元75使用时间缩放功能执行空间运动向量缩放及/或层间位置映射,如下文更详细地描述。

[0099] 反量化单元76将提供于位流中且由熵解码单元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由视频解码器30针对视频切片中的每一视频块计算以确定应应用的量化程度及同样的反量化程度的量化参数QPY。

[0100] 反变换单元78对变换系数应用反变换(例如,反DCT、反DST、反整数变换或概念上类似的反变换过程),以便产生像素域中的残差块。

[0101] 在运动补偿单元72基于运动向量及其它语法元素产生当前视频块的预测性块后,视频解码器30通过对来自反变换单元78的残差块与由运动补偿单元72产生的对应预测性块求和而形成经解码的视频块。求和器90表示执行此求和运算的一或多个组件。视需要,还可应用解块滤波器以对经解码块进行滤波,以便移除成块假影。还可使用其它环路滤波器(在译码环路中或在译码环路之后)来使像素转变平滑或者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,所述参考图片存储器存储用于后续运动补偿的参考图片。参考帧存储器82还存储经解码视频以用于稍后呈现在显示装置(例如,图1的显示装置32)上。

### [0102] 多层解码器

[0103] 图3B是说明可实施根据本发明中描述的方面的技术的多层视频解码器31的实例的框图。视频解码器31可经配置以处理多层视频帧,例如,用于SHVC及多视图译码。另外,视频解码器31可经配置以执行本发明的技术中的任一者或全部。

[0104] 视频解码器31包含视频解码器30A和视频解码器30B,其中的每一者可经配置为图3A的视频解码器30且可执行上文关于视频解码器30所描述的功能。此外,如再使用参考数字所指示,视频解码器30A和30B可包含系统及子系统至少一些作为视频解码器30。虽然将视频解码器31说明为包含两个视频解码器30A和30B,但视频解码器31不受如此限制且可包含任何数目个视频解码器30层。在一些实施例中,视频解码器31可包含用于存取单元中的每一图片或帧的视频解码器30。举例来说,可由包含五个解码器层的视频解码器处理或解码包含五个图片的存取单元。在一些实施例中,视频解码器31可包含比存取单元中的帧更多的解码器层。在一些此类情况下,当处理一些存取单元时,一些视频解码器层可能不在作用中。

[0105] 除了视频解码器30A及30B之外,视频解码器31还可包含上取样单元92。在一些实施例中,上取样单元92可对所接收的视频帧的基础层进行上取样以产生将添加到用于帧或存取单元的参考图片列表的增强型层。此增强型层可存储在参考帧存储器82中(例如,其经解码图片缓冲器等中)。在一些实施例中,上取样单元92可包含相对于图2B的再取样单元90所描述的实施例中的某些或全部。在一些实施例中,上取样单元92经配置以对层进行上取样且重新组织、重新界定、修改或调整一或多个切片以符合一组切片边界规则和/或光栅扫描规则。在一些情况下,上取样单元92可为经配置以对所接收的视频帧的层上取样及/或下取样的再取样单元。

[0106] 上取样单元92可经配置以从较低层解码器(例如,视频解码器30A)的经解码图片缓冲器82接收图片或帧(或与图片相关联的图片信息)且对所述图片(或所接收的图片信息)进行上取样。随后可将此经上取样的图片提供到较高层解码器(例如,视频解码器30B)的层间预测单元75,所述较高层解码器经配置以解码与较低层解码器相同的存取单元中的图片。在一些情况下,较高层解码器为从较低层解码器移除的一个层。在其它情况下,在图3B的层0解码器与层1解码器之间可存在一或多个较高层解码器。

[0107] 在一些情况下,可省略或绕过上取样单元92。在此些情况下,可直接提供来自视频解码器30A的经解码图片缓冲器82的图片,或至少不提供到上取样单元92,不提供到视频解码器30B的层间预测单元75。例如,如果提供到视频解码器30B的视频数据及来自视频解码器30A的经解码图片缓冲器82的参考图片具有相同的大小或分辨率,那么可将参考图片提供到视频解码器30B而不需要上取样。另外,在一些实施例中,上取样单元92可为经配置以对从视频解码器30A的经解码图片缓冲器82接收的参考图片进行上取样或下取样的再取样单元90。

[0108] 如图3B中所说明,视频解码器31可进一步包含多路分用器99或demux。demux 99可将经编码视频位流分裂为多个位流,其中由demux 99输出的每一位流被提供到不同的视频解码器30A和30B。可通过接收位流来产生多个位流,且视频解码器30A和30B中的每一者在给定时间接收位流的一部分。虽然在一些情况下,可在视频解码器(例如,在图3B的实例中的视频解码器30A及30B)中的每一者之间一次一个位地交替来自在demux 99处接收的位流

的位,但在许多情况下,不同地划分所述位流。举例来说,可通过一次一个块地交替哪一视频解码器接收位流而划分所述位流。在另一实例中,可通过到视频解码器30A和30B中的每一者的块的非1:1比率来划分位流。举例来说,可针对提供到视频解码器30A的每一块将两个块提供到视频解码器30B。在一些实施例中,可预编程由demux 99对位流的划分。在其它实施例中,demux 99可基于从视频解码器31的外部的系统(例如从目的地装置14上的处理器)接收的控制信号而划分所述位流。可基于来自输入接口28的视频的分辨率或位速率、基于信道16的带宽、基于与用户相关联的预订(例如,付费预订与免费预订)或基于用于确定可由视频解码器31获得的分辨率的任何其它因素而产生控制信号。

[0109] 在VPS中用信号通知视图ID位深度且用信号通知位速率信息及/或图片速率信息

[0110] 在MV-HEVC及SHVC的早期版本(例如,工作草案2)中,使用固定数目的位用信号通知层的视图ID。例如,SHVC的早期版本使用10个位用信号通知视图ID,而不管多少视图可用于一层。然而,使用固定位深度用信号通知视图ID可导致低效,尤其当可用的视图的数目较小(例如,1到4个视图)时。使用10个位的固定位深度可允许用于一层的至多1024 ( $2^{10}$ ) 个视图的信令,但在大多数情况下,用于一层的总视图的数目远小于那个数目。

[0111] 此外,在MV-HEVC及SHVC的早期版本中,针对每一层集合及每一层集合的每一子层而用信号通知位速率信息及图片速率信息。对于每一层集合,用信号通知包含位速率及图片速率信息的语法结构(例如,bit\_rate\_pic\_rate())。对于层集合的每一子层,用信号通知指示是否存在位速率信息的相应的旗标,且用信号通知指示是否存在图片速率信息的相应的旗标。不管是否存在任何层集合或子层的任何位速率信息及/或图片速率信息而执行此过程。例如,用信号通知0以作为每一层集合及每一相应子层的旗标的值,以便指示全部层集合及子层不存在位速率信息及/或图片速率信息。此可例如导致低效及不必要的信令及/或处理,这是因为可存在许多层集合及所述层集合的许多子层。

[0112] 为了解决这些及其它挑战,本发明的技术可实现用信号通知视图ID的位深度,且使用所述用信号通知的位深度而用信号通知所述视图ID的值。例如,如果仅使用2个视图,那么视图ID的位深度可为1位,且可使用1位用信号通知视图ID的值。如果使用3个视图,那么视图ID的位深度可为2位,且可使用2位用信号通知视图ID的值。视图ID的位深度可为变量(例如,在1位与16位之间)。以此方式,可通过减少用于用信号通知视图ID的值的位数目而更有效地用信号通知视图ID。

[0113] 本发明的技术可实现在VPS中用信号通知位速率信息及/或图片速率信息。例如,所述技术可指示任何层集合及/或层集合的任何子层是否具有位速率信息及/或图片速率信息。在一个实施例中,所述技术可在VPS中用信号通知指示任何层集合及/或子层是否存在位速率信息的全局旗标,且在VPS中用信号通知指示任何层集合及/或子层是否存在图片速率信息的全局旗标。通过在VPS中包含所述全局旗标,可仅当所述全局旗标指示至少一个层集合或至少一个子层存在位速率信息及/或图片速率信息时才用信号通知及/或存取位速率图片速率语法结构。如果全局旗标指示任何层不存在位速率信息及/或图片速率信息,那么不需要用信号通知及/或存取位速率图片速率语法结构。另外,全局旗标可允许以有效方式对位速率信息及图片速率信息的单独的处理。如果用于位速率信息的全局旗标指示不存在位速率信息,那么不需要用信号通知及/或处理用于子层的位速率信息的旗标。类似地,如果用于图片速率信息的全局旗标指示子层不存在图片速率信息,那么不需要用信号

通知及/或处理用于子层的图片速率信息的旗标。虽然先前还可能用信号通知位速率信息或图片速率信息中的仅一者,但需要用信号通知及处理多个个别旗标:用于每一层集合的每一子层的位速率信息的一个旗标,及用于每一层集合的每一子层的图片速率信息的一个旗标。

[0114] 贯穿本发明所使用的各种术语是具有它们的一般含义的广义术语。另外,在一些实施方案中,某些术语涉及以下视频概念。图片可指视频图片,所述术语用于当前标准(例如,HEVC、SHVC、MV-HEVC等)中。视频参数集(VPS)可指全局应用于多个层且跨越存取单元序列的任何参数集合。补充增强信息(SEI)可指在相符位流中的图片的校正解码不一定需要但可用于提高用户体验(例如,帮助提高在存在传输错误等的情况下的视频质量)的任何信息。视图识别符(ID)可指视图(相机的表示)的识别符或触觉信号(例如,触觉传感器的表示)。会话协商可指能力交换、供应应答等。本发明的技术也可应用于用信号通知除位速率及图片速率之外的位流特性,例如随机存取周期、每一类型的经译码图片(经帧内译码、经单向预测图片、经双向预测等)的数目等。在一些实施例中,计算硬件可包含一或多个计算装置,包括计算机硬件。

#### [0115] 视图ID位深度的信令

[0116] 如上文所解释,可使用可变位深度用信号通知层的视图ID。在一个实施例中,位深度可为可变的,原因在于可例如基于将用信号通知的视图的数目而恰当地选择位深度。在某些实施例中,在VPS中用信号通知所述位深度。在一个实施例中,MV-HEVC及SHVC的早期版本中的vps\_extension()语法及语义可如斜体字所指示而改变。所述改变可从JCT3V-D0196中的方法改变。可通过鉴于view\_id\_len\_minus1用信号通知视图ID值view\_id\_val的长度而恰当地调整用于用信号通知视图ID的位数目。

[0117]

<b>vps_extension()</b> {	<b>描述符</b>
while( !byte_aligned( ) )	
<b>vps_extension_byte_alignment_reserved_one_bit</b>	u(1)
<b>avc_base_layer_flag</b>	u(1)
<b>splitting_flag</b>	u(1)
for( i = 0, NumScalabilityTypes = 0; i < 16; i++ ) {	
<b>scalability_mask_flag[i]</b>	u(1)
NumScalabilityTypes += scalability_mask_flag[i]	
}	
for( j = 0; j < ( NumScalabilityTypes - splitting_flag ); j++ )	
<b>dimension_id_len_minus1[j]</b>	u(3)
<b>vps_nuh_layer_id_present_flag</b>	u(1)
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
if( vps_nuh_layer_id_present_flag )	
<b>layer_id_in_nuh[i]</b>	u(6)
if( !splitting_flag )	
for( j = 0; j < NumScalabilityTypes; j++ )	
<b>dimension_id[i][j]</b>	u(v)
}	
<b>view_id_explicitly_signalled_flag</b>	u(1)
if( view_id_explicitly_signalled_flag ) {	
<b>view_id_len_minus1</b>	u(4)
<b>view_id_present_for_all_layers_flag</b>	u(1)
for( i = 0; i <= vps_max_layers_minus1; i++ )	
if( view_id_present_for_all_layers_flag    ( i % 2 == 0 ) )	
<b>view_id_val[i]</b>	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
for( j = 0; j < i; j++ )	
<b>direct_dependency_flag[i][j]</b>	u(1)
...	

[0118] 以上各种语法元素或变量可如下界定：

[0119] • scalability\_mask\_flag[i]等于1指示存在对应于表F-1中的第i可缩放性尺寸的dimension\_id语法元素。The scalability\_mask\_flag[i]等于0指示不存在对应于第i可缩放性尺寸的dimension\_id语法元素。

[0120] 表F-1-ScalabilityId到可缩放性尺寸的映射

[0121]

可缩放性掩码索引	可缩放性尺寸	ScalabilityId映射
----------	--------	-----------------

0	多视图	视图次序索引
1-15	保留	

[0122] • dimension\_id\_len\_minus1[j]加上1指定dimension\_id[i][j]语法元素的以位为单位的长度。

[0123] • 当splitting\_flag等于1时,以下适用:

[0124] -变量dimBitOffset[0]设定成等于0,且对于在1到NumScalabilityTypes-1(包括1和NumScalabilityTypes-1)的范围内的j,如下导出dimBitOffset[j]:

$$[0125] \quad \text{dimBitOffset}[j] = \sum_{\text{dimIdx}=0}^{j-1} (\text{dimension\_id\_len\_minus1}[\text{dimIdx}] + 1) \quad (\text{F-1})$$

[0126] -推断dimension\_id\_len\_minus1[NumScalabilityTypes-1]的值等于5? dimBitOffset[NumScalabilityTypes-1]。

[0127] -dimBitOffset[NumScalabilityTypes]的值设定成等于6。

[0128] -位流符合性要求是,dimBitOffset[NumScalabilityTypes-1]应小于6。

[0129] • vps\_nuh\_layer\_id\_present\_flag等于1指定对于从1到vps\_max\_layers\_minus1的i(包括1及vps\_max\_layers\_minus1),存在layer\_id\_in\_nuh[i]。vps\_nuh\_layer\_id\_present\_flag等于0指定对于从1到vps\_max\_layers\_minus1的i(包括1及vps\_max\_layers\_minus1),不存在layer\_id\_in\_nuh[i]。

[0130] • layer\_id\_in\_nuh[i]指定第i层的VCL NAL单元中的nuh\_layer\_id语法元素的值。对于在0到vps\_max\_layers\_minus1的范围内的i(包括0及vps\_max\_layers\_minus1),当不存在layer\_id\_in\_nuh[i]时,推断所述值等于i。

[0131] -当i大于0时,layer\_id\_in\_nuh[i]将大于layer\_id\_in\_nuh[i-1]。

[0132] -对于从0到vps\_max\_layers\_minus1的i(包括0及vps\_max\_layers\_minus1),变量LayerIdxInVps[layer\_id\_in\_nuh[i]]设定成等于i。

[0133] • dimension\_id[i][j]指定第j识别符呈现第i层的可缩放性尺寸类型。用于表示dimension\_id[i][j]的位数目是dimension\_id\_len\_minus1[j]+1位。

[0134] -如果splitting\_flag等于1,对于从0到vps\_max\_layers\_minus1的i(包括0及vps\_max\_layers\_minus1)及从0到NumScalabilityTypes-1的j(包括0及NumScalabilityTypes-1),推断dimension\_id[i][j]等于((layer\_id\_in\_nuh[i]&((1<<dimBitOffset[j+1])-1))>>dimBitOffset[j])。

[0135] -否则,对于从0到NumScalabilityTypes-1的j(包括0及NumScalabilityTypes-1),推断dimension\_id[0][j]等于0。

[0136] -如下导出指定第i层的第smIdx可缩放性尺寸类型的识别符的变量ScalabilityId[i][smIdx]及指定第i层是否为视图可缩放性扩展层的变量ViewScalExtLayerFlag:

```
[0137] for(i=0;i<=vps_max_layers_minus1;i++){
[0138]     lId=layer_id_in_nuh[i]
[0139]     for(smIdx=0,j=0;smIdx<16;smIdx++){
[0140]         if(scalability_mask_flag[smIdx])
[0141]             ScalabilityId[i][smIdx]=dimension_id[i][j++]
```



[0142]       ViewScalExtLayerFlag[lId] = (ScalabilityId[i][0]>0)

[0143]    }

[0144]    • view\_id\_explicitly\_signalled\_flag等于1指定针对由VPS指定的一些或全部层显式地指派视图识别符。view\_id\_explicitly\_signalled\_flag等于0指定不针对由VPS指定的层显式地指派视图识别符。

[0145]    • view\_id\_len\_minus1加上1指定view\_id\_val[i]语法元素的以位为单位的长度。

[0146]    • view\_id\_present\_for\_all\_layers\_flag等于1指定显式地用信号通知由VPS指定的每一层的视图识别符。view\_id\_present\_for\_all\_layers\_flag等于0指定针对由VPS指定的一些层显式地用信号通知视图识别符且针对由VPS指定的其它层导出视图识别符。

[0147]    • view\_id\_val[i]指定由VPS指定的第i层的视图识别符。view\_id\_val[i]语法元素由view\_id\_len\_minus1+1个位表示。

[0148]    -当view\_id\_explicitly\_signalled\_flag等于1时,view\_id\_present\_for\_all\_layers\_flag等于0,且i%2等于1,推断view\_id\_val[i]的值等于view\_id\_val[i-1]。

[0149]    -当view\_id\_explicitly\_signalled\_flag等于0时,推断view\_id\_val[i]的值等于ScalabilityId[i][0]。

[0150]    -对于具有等于nuhLayerId的nuh\_layer\_id的每一层,变量ViewId设定成等于view\_id\_val[LayerIdxInVps[nuhLayerId]]。层中的每一图片被视为与层的ViewId相关联。

[0151]    • 对于SHVC,可需要view\_id\_info\_present\_flag的值等于0。在包含深度的MV-HEVC的潜在扩展中,一个视图的纹理及深度分量将为两个相邻层且具有相同的view\_id。当纹理及深度始终配对时,将view\_id\_info\_present\_flag设定为等于1且将view\_id\_present\_for\_all\_layers\_flag设定为等于0将为合意的。

[0152]    在以上实施例中,用信号通知view\_id\_explicitly\_signalled\_flag以指示显式地用信号通知视图识别符。可在VPS中用信号通知view\_id\_explicitly\_signalled\_flag。如果view\_id\_explicitly\_signalled\_flag等于1,那么用信号通知view\_id\_len\_minus1。view\_id\_len\_minus1可指示用于用信号通知一或多个视图识别符的位深度减去1。在一个实例中,所述位深度可在1与16之间。使用由view\_id\_len\_minus1+1指示的位深度用信号通知view\_id\_val[i]。view\_id\_val[i]的长度可取决于需要用信号通知的视图的数目而为可变的。

[0153]    类似地,在解码器侧,例如在VPS中接收view\_id\_explicitly\_signalled\_flag。view\_id\_explicitly\_signalled\_flag可指示显式地用信号通知视图识别符。如果view\_id\_explicitly\_signalled\_flag等于1,那么处理view\_id\_len\_minus1。view\_id\_len\_minus1可指示用于用信号通知一或多个视图识别符的位深度减去1。可接收view\_id\_val[i]以作为具有view\_id\_len\_minus1+1的长度的值。

[0154]    以此方式,可变位深度可用于通过用信号通知视图ID的位深度而用信号通知层的视图ID。随后,可使用由位深度指示的位数来解码视图ID。在一些实施例中,所述位深度也可被称作视图ID的长度。

[0155]    在VPS中用信号通知位速率信息及/或图片速率信息

[0156] 在MV-HEVC及SHVC的早期版本中,在VPS中用信号通知可用于会话协商及内容选择的与简档、层及层级相关的信息。然而,不在VPS中用信号通知对于相同的目的也重要的其它信息,例如位速率及图片速率。SVC及MVC分别在可缩放性信息补充增强信息(SEI)消息及视图可缩放性信息SEI消息中支持位速率及图片速率信息的信令。在HEVC多层扩展中,在可缩放性信息SEI消息及视图可缩放性信息SEI消息(AVC扩展的)中传达的对于会话协商重要的信息中的一些或全部现在可包含在VPS中。

[0157] 因此,在本发明的一个方面中,在VPS中用信号通知位速率信息及/或图片速率信息。此信息的语义与例如国际标准化组织(ISO)基础媒体文件格式及其扩展文件格式等系统规范中的它们的对应部分对准。

[0158] 与JCTVC-K0125中的方法相比,本发明的技术可通过包含旗标bit\_rate\_present\_vps\_flag及pic\_rate\_present\_vps\_flag及它们在调节语法结构及语法元素中的用途而实现对位速率信息及图片速率信息中的仅一者的更有效的信令。

[0159] 在一个实施例中,可如斜体字所指示而改变MV-HEVC及SHVC的早期版本中的vps\_extension()语法及语义:

[0160]

vps_extension() {	描述符
...	
<i>bit_rate_present_vps_flag</i>	<i>u(1)</i>
<i>pic_rate_present_vps_flag</i>	<i>u(1)</i>
<i>if( bit_rate_present_vps_flag    pic_rate_present_vps_flag )</i>	
<i>for( i = 0; i &lt;= vps_number_layer_sets_minus1; i++ )</i>	
<i>bit_rate_pic_rate( )</i>	
<b>max_one_active_ref_layer_flag</b>	<b>u(1)</b>
...	
}	

[0161] • bit\_rate\_present\_vps\_flag等于1或pic\_rate\_present\_vps\_flag等于1指定在VPS中存在每一层集合的bit\_rate\_pic\_rate()语法结构。bit\_rate\_present\_vps\_flag等于0且pic\_rate\_present\_vps\_flag等于0指定没有bit\_rate\_pic\_rate()语法结构存在于VPS中。

[0162]

bit_rate_pic_rate() {	描述符
for( i = 0; i <= vps_max_sub_layers_minus1; i++ ) {	
if( bit_rate_present_vps_flag )	
bit_rate_present_flag[i]	u(1)
if( pic_rate_present_vps_flag )	
pic_rate_present_flag[i]	u(1)
if( bit_rate_present_flag[i] ) {	
avg_bit_rate[i]	u(16)
max_bit_rate[i]	u(16)
}	
if( pic_rate_present_flag[i] ) {	
constant_pic_rate_idc[i]	u(2)
avg_pic_rate[i]	u(16)
}	
}	
}	

[0163] • bit\_rate\_present\_flag[i]等于1指定存在层集合的第i子集的位速率信息。bit\_rate\_present\_flag[i]等于0指定不存在层集合的第i子集的位速率信息。层集合的第i子集是当使用作为输入的层集合i及与所述层集合相关联的层识别符列表调用时子位流提取过程的输出。当不存在时,推断bit\_rate\_present\_flag[i]的值等于0。

[0164] • pic\_rate\_present\_flag[i]等于1指定存在层集合的第i子集的图片速率信息。pic\_rate\_present\_flag[i]等于0指定不存在层集合的第i子集的图片速率信息。当不存在时,推断pic\_rate\_present\_flag[i]的值等于0。

[0165] • avg\_bit\_rate[i]指示以位/秒为单位的层集合的第i子集的平均位速率。通过BitRateBPS(avg\_bit\_rate[i])给出值,其中函数BitRateBPS()由下式指定

$$[0166] \quad \text{BitRateBPS}(x) = (x \& (2^{14}-1)) * 10^{(2+(x \gg 14))} \quad (\text{F-2})$$

[0167] -根据在SHVC WD2及MV-HEVC WD4的附件F.13中指定的存取单元移除时间而导出平均位速率。在下文中,bTotal是层集合的第i集合的全部NAL单元中的位数目,t<sub>1</sub>是VPS应用的第一存取单元的移除时间(以秒为单位),且t<sub>2</sub>是VPS应用的最后存取单元(在解码次序上)的移除时间(以秒为单位)。

[0168] -其中x指定avg\_bit\_rate[i]的值,以下适用:

[0169] ○如果t<sub>1</sub>不等于t<sub>2</sub>,以下条件将为真:

$$[0170] \quad (x \& (2^{14}-1)) == \text{Round}(b\text{Total} \div ((t_2-t_1) * 10^{(2+(x \gg 14))})) \quad (\text{F-3})$$

[0171] ○否则(t<sub>1</sub>等于t<sub>2</sub>),以下条件将为真:

$$[0172] \quad (x \& (2^{14}-1)) == 0 \quad (\text{F-4})$$

[0173] • max\_bit\_rate\_layer[i]指示附件F.13中指定的存取单元移除时间的任何一秒时间窗中的层集合的第i子集的位速率的上限。通过BitRateBPS(max\_bit\_rate\_layer[i])给出以位/秒为单位的位速率的上限。根据附件F.13中指定的存取单元移除时间而导出位

速率值。在下文中,  $t_1$  是任一时间点 (以秒为单位),  $t_2$  设定成等于  $t_1 + 1 \div 100$ , 且  $bTotal$  是存取单元的全部NAL单元中的位数目, 其中移除时间大于或等于  $t_1$  且小于  $t_2$ 。其中  $x$  指定  $max\_bit\_rate\_layer[i]$  的值,  $t_1$  的所有值遵守以下条件:

$$[0174] \quad (x \& (2^{14} - 1)) \geq bTotal \div ((t_2 - t_1) * 10^{(2 + (x \gg 14))}) \quad (F-5)$$

[0175] •  $constant\_pic\_rate\_idc[i]$  指示层集合的第  $i$  子集的图片速率是否为常量。在下文中, 时间片段  $tSeg$  是层集合的第  $i$  子集的在解码次序上的两个或更多个连续存取单元的任何集合,  $fTotal(tSeg)$  是时间片段  $tSeg$  中的存取单元的数目,  $t_1(tSeg)$  是时间片段  $tSeg$  的第一存取单元 (在解码次序上) 的移除时间 (以秒为单位),  $t_2(tSeg)$  是时间片段  $tSeg$  的最后存取单元 (在解码次序上) 的移除时间 (以秒为单位), 且  $avgFR(tSeg)$  是时间片段  $tSeg$  中的平均图片速率, 其通过下式给出:

$$[0176] \quad avgFR(tSeg) = Round(fTotal(tSeg) * 256 \div (t_2(tSeg) - t_1(tSeg))) \quad (F-6)$$

[0177] - 如果层集合的第  $i$  子集仅含有一个或两个存取单元, 或  $avgFR(tSeg)$  的值在所有时间片段上是常量, 那么图片速率是常量; 否则, 图片速率不是常量。

[0178] -  $constant\_pic\_rate\_idc[i]$  等于 0 指示层集合的第  $i$  子集的图片速率不是常量。 $constant\_pic\_rate\_idc[i]$  等于 1 指示层集合的第  $i$  子集的图片速率是常量。 $constant\_pic\_rate\_idc[i]$  等于 2 指示层集合的第  $i$  子集的图片速率可为或可不是常量。 $constant\_pic\_rate\_idc[i]$  的值将在 0 到 2 的范围内 (包括 0 及 2)。

[0179] •  $avg\_pic\_rate[i]$  指示层集合的第  $i$  子集的以每 256 秒的图片为单位的平均图片速率。其中  $fTotal$  是层集合的第  $i$  子集中的存取单元的数目,  $t_1$  是 VPS 应用的第一存取单元的移除时间 (以秒为单位), 且  $t_2$  是 VPS 应用的最后存取单元 (在解码次序上) 的移除时间 (以秒为单位), 以下适用:

[0180] - 如果  $t_1$  不等于  $t_2$ , 以下条件将为真:

$$[0181] \quad avg\_pic\_rate[i] = Round(fTotal * 256 \div (t_2 - t_1)) \quad (F-7)$$

[0182] - 否则 ( $t_1$  等于  $t_2$ ), 以下条件将为真:

$$[0183] \quad avg\_pic\_rate[i] = 0 \quad (F-8)$$

[0184] 在以上实施例中, 在 VPS 中用信号通知全局旗标  $bit\_rate\_present\_vps\_flag$  及  $pic\_rate\_present\_vps\_flag$ 。 $bit\_rate\_present\_vps\_flag$  指示一或多个层集合的至少一个子层是否具有用信号通知的位速率信息, 且  $pic\_rate\_present\_vps\_flag$  指示一或多个层集合的至少一个子层是否具有用信号通知的图片速率信息。如果  $bit\_rate\_present\_vps\_flag$  及  $pic\_rate\_present\_vps\_flag$  两者等于 0, 那么没有位速率图片速率语法结构在 VPS 中用信号通知。如果  $bit\_rate\_present\_vps\_flag$  或  $pic\_rate\_present\_vps\_flag$  等于 1, 那么用信号通知每一层集合的位速率图片速率语法结构。层集合可指一或多个层的群组。在每一层集合的位速率图片速率语法结构内, 如果  $bit\_rate\_present\_vps\_flag$  等于 1, 那么可用信号通知每一子层的  $bit\_rate\_present\_flag$ , 且如果  $pic\_rate\_present\_vps\_flag$  等于 1, 那么可用信号通知每一子层的  $pic\_rate\_present\_flag$ 。

[0185] 类似地, 在解码器侧, 在 VPS 中接收全局旗标  $bit\_rate\_present\_vps\_flag$  及  $pic\_rate\_present\_vps\_flag$ 。如果  $bit\_rate\_present\_vps\_flag$  及  $pic\_rate\_present\_vps\_flag$  两者等于 0, 那么不存取及/或处理位速率图片速率语法结构。如果  $bit\_rate\_present\_vps\_flag$  或  $pic\_rate\_present\_vps\_flag$  等于 1, 那么存取及/或处理每一层集合的位速率

图片速率语法结构。在每一层集合的位速率图片速率语法结构内,如果bit\_rate\_present\_vps\_flag等于1,那么可存取及/或处理每一子层的bit\_rate\_present\_flag,且如果pic\_rate\_present\_vps\_flag等于1,那么可存取及/或处理每一子层的pic\_rate\_present\_flag。

[0186] 以此方式,所述技术可通过在VPS中分别包含指示VPS中指示的所有层是否存在位速率信息及/或图片速率信息的全局旗标而减少用于编码及/或解码位速率信息及/或图片速率信息的资源。可仅当位速率图片速率语法结构包含位速率信息及/或图片速率信息时才存取位速率图片速率语法结构。另外,可单独地用信号通知及/或处理位速率信息及图片速率信息。例如,当仅需要位速率信息时,可仅用信号通知位速率信息而不必用信号通知图片速率信息,且反之亦然。

[0187] 下文参考图4到6描述与所述技术相关的某些细节。相对于图4所描述的所有特征及/或实施例可单独实施或与图4到6中所描述的其它特征及/或实施例以任何组合实施。

[0188] 用信号通知视图ID位深度的方法

[0189] 图4是说明根据本发明的方面的用于用信号通知或编码视图ID位深度的方法的流程图。取决于实施例,过程400可由编码器(例如,如图2A、2B等中所展示的编码器)、解码器(例如,如图3A、3B等中所展示的解码器),或任何其它组件执行。关于图2B中的编码器21描述过程400的框,但过程400可由如上文所提及的例如解码器等其它组件执行。取决于实施例,编码器21的层1视频编码器20B及/或编码器21的层0编码器20A可执行过程400。关于图4描述的所有实施例可单独地实施,或彼此组合实施。上文及下文例如关于图5和6解释与过程400相关的某些细节。

[0190] 过程400开始于框401处。编码器21可包含用于存储视频信息的存储器(例如,参考帧存储器64)。

[0191] 在框402处,编码器21确定用信号通知的一或多个视图识别符的位深度。一或多个视图识别符中的每一者可待编码的层相关联。可基于可例如在相同位流中编码的视图的最大数目而确定一或多个视图识别符的位深度。用于用信号通知视图识别符的位深度可例如在可取决于要编码的视图的数目(例如,最大数目)而恰当地选择的位深度的意义上是可变的。层可指与视频信息相关联的层,例如可缩放视频译码(例如,SHVC)中的层或3-D视频译码(例如,MV-HEVC)中的层。SHVC位流通常表示由一个相机俘获的视频信号,且位流可包含多个层,每一层对应于具有不同质量或不同空间分辨率的视频信号的表示。MV-HEVC位流通常表示由多个相机俘获的视频信号,且位流可包含多个层,每一层对应于由相异的相机俘获的视频信号的表示。MV-HEVC中的层也可被称作视图。

[0192] 在框403处,编码器21在位流中用信号通知一或多个视图识别符的位深度。在一些实施例中,在视频参数集(VPS)中用信号通知一或多个视图识别符的位深度。在一个实施例中,由用信号通知的位深度指示的位数目在1与16之间。用信号通知的位深度可由例如如结合图5所解释的解码器接收及解码。

[0193] 在某些实施例中,编码器21用信号通知是否在位流中显式地用信号通知视图识别符。在一个实施例中,编码器21通过用信号通知视图识别符显式地用信号通知的旗标而用信号通知是否在位流中显式地用信号通知视图识别符。在一些实施例中,编码器21使用由用信号通知的位深度指示的位数目而用信号通知一或多个视图识别符。

[0194] 过程400结束于框404处。可在过程400中添加及/或忽略框,其取决于实施例,且过程400的框可以不同次序执行,其取决于实施例。

[0195] 本发明中相对于再取样描述的任何特征和/或实施例可单独地或以其任何组合来实施。举例来说,结合图5到6描述的任何特征和/或实施例可以与结合图4描述的任何特征和/或实施例的任何组合来实施,且反之亦然。

[0196] 图5是说明根据本发明的方面的用于解码视图ID位深度的方法的流程图。取决于实施例,过程500可由编码器(例如,如图2A、2B等中所展示的编码器)、解码器(例如,如图3A、3B等中所展示的解码器),或任何其它组件执行。相对于图3B中的解码器31描述过程500的框,但过程500可由如上文所提及的例如编码器等其它组件执行。取决于实施例,解码器31的层1视频解码器30B和/或解码器31的层0解码器30A可执行过程500。相对于图5描述的所有实施例可单独地实施,或彼此组合实施。上文及下文例如相对于图4到6解释与过程500相关的某些细节。

[0197] 过程500开始于框501处。解码器31可包含用于存储视频信息的存储器(例如,参考帧存储器82)。

[0198] 在框502处,解码器31接收指示用于用信号通知一或多个视图识别符值的位的数目的位深度指示符。所述一或多个视图识别符值中的每一者可与待解码的一或多个层中的一者相关联。所述位深度指示符可为由编码器21编码或用信号通知的位深度,如上文结合图4所解释。在一个实施例中,由位深度指示符指示的位数目在1与16之间。所述位深度指示符可指示可编码的视图的最大数目。

[0199] 在框503处,解码器31接收一或多个视图识别符值中的每一者以作为具有所指示的位数值。可在VPS中接收位深度指示符及一或多个视图识别符值。

[0200] 过程500结束于框504处。可在过程500中添加及/或忽略框,其取决于实施例,且过程500的框可以不同次序执行,其取决于实施例。

[0201] 本发明中相对于再取样描述的任何特征和/或实施例可单独地或以其任何组合来实施。举例来说,结合图4及6描述的任何特征和/或实施例可以与结合图5描述的任何特征和/或实施例的任何组合来实施,且反之亦然。

[0202] 在VPS中用信号通知位速率信息及/或图片速率信息的方法

[0203] 图6是说明根据本发明的方面的用于在VPS中用信号通知位速率信息及/或图片速率信息的方法的流程图。取决于实施例,过程600可由编码器(例如,如图2A、2B等中所展示的编码器)、解码器(例如,如图3A、3B等中所展示的解码器),或任何其它组件执行。相对于图3B中的解码器31描述过程600的框,但过程600可由如上文所提及的例如编码器等其它组件执行。取决于实施例,解码器31的层1视频解码器30B和/或解码器31的层0解码器30A可执行过程600。相对于图6描述的所有实施例可单独地实施,或彼此组合实施。上文及下文例如相对于图4到5解释与过程600相关的某些细节。

[0204] 过程600开始于框601处。解码器31可包含用于存储视频信息的存储器(例如,参考帧存储器82)。

[0205] 在框602处,解码器31处理指示一或多个层集合的至少一个子层是否具有用信号通知的位速率信息的信号。层集合可指一或多个层的集合且可不同于层,原因在于层集合可包含一个以上层。例如,所述信号可为指示一或多个层集合的至少一个子层是否具有用

信号通知的位速率信息的全局旗标。所述全局旗标可包含在VPS中。在一个实施例中,处理所述信号是编码所述信号。在另一实施例中,处理所述信号是解码所述信号。在某些实施例中,计算装置可实施编码器及解码器两者的功能性。

[0206] 在框603处,解码器31处理指示一或多个层集合的至少一个子层是否具有用信号通知的图片速率信息的信号。例如,所述信号可为指示一或多个层集合的至少一个子层是否具有用信号通知的图片速率信息的全局旗标。所述全局旗标可包含在VPS中。在一个实施例中,处理所述信号是编码所述信号。在另一实施例中,处理所述信号是解码所述信号。在某些实施例中,计算装置可实施编码器及解码器两者的功能性。

[0207] 在某些实施例中,解码器31 (1) 当第一信号指示一或多个层集合的至少一个子层具有用信号通知的位速率信息时或 (2) 当第二信号指示一或多个层集合的至少一个子层具有用信号通知的图片速率信息时,处理位速率图片速率语法结构。在一些实施例中,解码器31通过当第一信号指示一或多个层集合的至少一个子层具有用信号通知的位速率信息时处理指示一或多个层集合的层的子层是否具有位速率信息的旗标,且当第二信号指示一或多个层集合的至少一个子层具有用信号通知的图片速率信息时处理指示一或多个层集合的层的子层是否具有图片速率信息的旗标,而处理位速率图片速率语法结构。在一个实施例中,处理所述位速率图片语法结构是编码位速率图片语法结构。在另一实施例中,处理位速率图片语法结构是解码位速率图片语法结构。在某些实施例中,计算装置可实施编码器及解码器两者的功能性。

[0208] 在一些实施例中,一个信号可用于指示一或多个层集合的至少一个子层是否具有用信号通知的位速率信息及一或多个层集合的至少一个子层是否具有用信号通知的图片速率信息。例如,相同的全局旗标可指示一或多个层集合及/或子层是否具有位速率信息及图片速率信息。在一个实施例中,用于每种类型的信息的全局旗标可合并为一个全局旗标。此全局旗标可包含在VPS中。

[0209] 根据某些方面,解码器31处理以下各者中的至少一者:指示一或多个层集合的至少一个子层是否具有用信号通知的位速率信息的第一信号,或指示一或多个层集合的至少一个子层是否具有用信号通知的图片速率信息的第二信号。例如,作为执行框602及603两者的替代,解码器31可例如在一个框中处理所述第一信号及所述第二信号中的至少一者。在一些实施例中,仅指示一或多个层集合的至少一个子层是否具有用信号通知的位速率信息的信号可包含在VPS中且由解码器31处理。在其它实施例中,仅指示一或多个层集合的至少一个子层是否具有用信号通知的图片速率信息的信号可包含在VPS中且由解码器31处理。

[0210] 过程600结束于框604处。可在过程600中添加及/或忽略框,其取决于实施例,且过程600的框可以不同次序执行,其取决于实施例。

[0211] 本发明中相对于再取样描述的任何特征和/或实施例可单独地或以其任何组合来实施。举例来说,结合图4到5描述的任何特征和/或实施例可以与结合图6描述的任何特征和/或实施例的任何组合来实施,且反之亦然。

[0212] 术语

[0213] 虽然以上揭示内容已描述特定实施例,但许多变化是可能的。举例来说,如上文所提及,以上技术可应用于3D视频编码。在3D视频的一些实施例中,参考层一基础层) 包含足

以显示视频的第一视图的视频信息,且增强层包含相对于所述参考层的额外视频信息,使得所述参考层及所述增强层一起包含足以显示视频的第二视图的视频信息。这两个视图可用于产生立体图像。如上文所论述,根据本发明的方面,来自参考层的运动信息可用于在编码或解码增强层中的视频单元时识别额外的隐式假设。此可提供对3D视频位流的更大译码效率。

[0214] 应认识到,取决于实例,本文中所描述的技术中的任一者的某些动作或事件可用不同次序执行、可添加、合并或全部省略(例如,实践所述技术并不需要所有的所描述动作或事件)。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而非循序地执行动作或事件。

[0215] 可使用多种不同技术和技艺中的任一者来表示本文中所揭示的信息和信号。例如,可通过电压、电流、电磁波、磁场或磁粒子、光场或光粒子或其任何组合来表示贯穿以上描述可能参考的数据、指令、命令、信息、信号、位、符号和码片。

[0216] 结合本文揭示的实施例所描述的各种说明性逻辑块、模块、电路及算法步骤可实施为电子硬件、计算机软件,或两者的组合。为清晰地说明硬件与软件的此可互换性,以上已大体就其功能性来描述了各种说明性组件、块、模块、电路和步骤。此类功能性是实施为硬件还是软件取决于具体应用及施加于整个系统的设计约束。熟练的技术人员可针对每一特定应用以不同方式来实施所描述的功能性,但这样的实施方案决策不应被解释为会引起脱离本发明的范围。

[0217] 本文中所描述的技术可以硬件、软件、固件或其任一组合来实施。所述技术可实施于多种装置中的任一者中,例如通用计算机、无线通信装置手持机或集成电路装置,其具有包含在无线通信装置手持机及其它装置中的应用的多种用途。被描述为模块或组件的任何特征可一起实施于集成逻辑装置中或分开来实施为离散但可互操作的逻辑装置。如果以软件实施,那么所述技术可至少部分地由包括程序代码的计算机可读数据存储媒体来实现,所述程序代码包含在执行时执行上文所描述的方法中的一或多者的指令。计算机可读数据存储媒体可形成计算机程序产品的一部分,所述计算机程序产品可包含封装材料。计算机可读媒体可包括存储器或数据存储媒体,例如随机存取存储器(RAM)(例如,同步动态随机存取存储器(SDRAM))、只读存储器(ROM)、非易失性随机存取存储器(NVRAM)、电可擦除可编程只读存储器(EEPROM)、快闪存储器、磁性或光学数据存储媒体及类似物。另外或作为替代,所述技术可至少部分地由计算机可读通信媒体来实现,所述计算机可读通信媒体以指令或数据结构的形式载运或传达程序代码且可由计算机存取、读取和/或执行(例如,传播的信号或波)。

[0218] 程序代码可由处理器执行,所述处理器可包含一或多个处理器,例如,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。此处理器可经配置以执行本发明中描述的技术中的任一者。通用处理器可为微处理器;但在替代方案中,处理器可为任何常规处理器、控制器、微控制器或状态机。处理器还可实施为计算装置的组合,例如,DSP与微处理器的组合、多个微处理器、一或多个微处理器与DSP核心的联合、或任何其它此类配置。因此,如本文中所使用的术语“处理器”可指前述结构中的任一者、前述结构的任何组合,或适用于实施本文中所描述的技术的任何其它结构或设备。另外,在一些方面中,可将本文中所描述的功能性提供于



经配置以用于编码和解码的专用软件模块或硬件模块内或并入组合的视频编码器-解码器(CODEC)中。

[0219] 本文论述的译码技术可为实例视频编码和解码系统中的实施例。系统包含提供经编码视频数据以在稍后时间由目的地装置解码的源装置。明确地说,源装置经由计算机可读媒体将视频数据提供到目的地装置。源装置和目的地装置可包括广泛范围的装置中的任一者,包含台式计算机、笔记本(即,膝上型)计算机、平板计算机、机顶盒、例如所谓的“智能”电话等电话手持机、所谓的“智能”板、电视机、相机、显示装置、数字媒体播放器、视频游戏控制台、视频串流装置,或类似物。在一些情况下,源装置和目的地装置可经装备以用于无线通信。

[0220] 目的地装置可经由计算机可读媒体接收待解码的经编码视频数据。计算机可读媒体可包括能够将经编码视频数据从源装置移动到目的地装置的任何类型的媒体或装置。在一个实例中,计算机可读媒体可包括使得源装置12能够实时将经编码视频数据直接发射到目的地装置的通信媒体。经编码视频数据可根据通信标准(例如无线通信协议)调制,且发射到目的地装置。通信媒体可以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发射线。通信媒体可形成基于包的网路(例如,局域网、广域网或全球网路,例如因特网)的部分。通信媒体可包含可用于促进从源装置到目的地装置的通信的路由器、交换机、基站或任何其它设备。

[0221] 在一些实例中,经编码数据可从输出接口输出到存储装置。类似地,经编码数据可由输入接口从存储装置存取。存储装置可包含多种分布式或本地存取的数据存储媒体中的任一者,例如硬盘驱动器、蓝光光盘、DVD、CD-ROM、快闪存储器、易失性或非易失性存储器或任何其它用于存储经编码视频数据的合适的数字存储媒体。在另一实例中,存储装置可以对应于文件服务器或可存储由源装置产生的经编码视频的另一中间存储装置。目的地装置可经由流式传输或下载从存储装置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经编码视频数据发射到目的地装置的任何类型的服务器。实例文件服务器包含网络服务器(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。目的地装置可通过任何标准数据连接(包含因特网连接)来存取经编码视频数据。此可包含无线通道(例如,Wi-Fi连接)、有线连接(例如,DSL、电缆调制解调器等),或适合于存取存储在文件服务器上的经编码视频数据的两者的组合。经编码视频数据从存储装置的发射可能是流式传输、下载传输或其组合。

[0222] 本发明的技术未必限于无线应用或环境。所述技术可以应用于视频译码以支持多种多媒体应用中的任一者,例如空中协议电视广播、有线电视发射、卫星电视发射、因特网流式视频发射(例如,经由HTTP的动态自适应流式传输(DASH))、经编码到数据存储媒体上的数字视频、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,系统可经配置以支持单向或双向视频传输,以支持例如视频流式传输、视频回放、视频广播和/或视频电话等应用。

[0223] 在一个实例中,源装置包含视频源、视频编码器和输出接口。目的地装置可包含输入接口、视频解码器和显示装置。源装置的视频编码器可经配置以应用本文所揭示的技术。在其它实例中,源装置和目的地装置可包括其它组件或布置。例如,源装置可从外部视频源(例如外部相机)接收视频数据。同样,目的地装置可与外部显示装置介接,而非包含集成显

示装置。

[0224] 以上实例系统仅是一个实例。用于并行处理视频数据的技术可由任何数字视频编码和/或解码装置来执行。尽管本发明的技术一般通过视频编码装置来执行,但是所述技术还可通过视频编码器/解码器(通常被称作“编解码器”)来执行。此外,本发明的技术还可通过视频预处理器执行。源装置和目的地装置仅为源装置在其中产生经译码视频数据以供发射到目的地装置的此类译码装置的实例。在一些实例中,源装置和目的地装置可以实质上对称的方式操作以使得所述装置中的每一者包含视频编码和解码组件。因此,实例系统可支持视频装置之间的单向或双向视频发射,例如用于视频串流、视频回放、视频广播或视频电话。

[0225] 视频源可包含视频俘获装置,例如,摄像机、含有先前所俘获视频的视频档案和/或用于从视频内容提供者接收视频的视频馈入接口。作为再一替代方案,视频源可产生基于计算机图形的数据作为源视频,或实况视频、存档视频与计算机产生的视频的组合。在一些情况下,如果视频源为摄像机,那么源装置和目的地装置可形成所谓的相机电话或视频电话。然而,如上文所提及,本发明中所描述的技术可大体上适用于视频译码,且可应用于无线和/或有线应用。在每一情况下,可由视频编码器编码所俘获、预先俘获或计算机产生的视频。经编码视频信息可随后由输出接口输出到计算机可读媒体上。

[0226] 如所提到,计算机可读媒体可包含瞬时媒体,例如,无线广播或有线网络传输,或存储媒体(即,非暂时性存储媒体),例如,硬盘、快闪驱动器、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(未图示)可从源装置接收经编码视频数据,且(例如)经由网络传输将经编码视频数据提供到目的地装置。类似地,媒体生产设施(例如,光盘冲压设施)的计算装置可从源装置接收经编码视频数据且生产含有经编码视频数据的光盘。因此,在各种实例中,计算机可读媒体可理解为包含各种形式的一或多个计算机可读媒体。

[0227] 目的地装置的输入接口从计算机可读媒体接收信息。计算机可读媒体的信息可包含由视频编码器定义的也供视频解码器使用的语法信息,所述语法信息包含描述块及其它经译码单元(例如,图片群组(GOP))的特性及/或处理的语法元素。显示装置向用户显示经解码视频数据,且可包括多种显示装置中的任一者,例如,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。已描述本发明的各种实施例。这些和其它实施例在所附权利要求书的范围内。

[0228] 已描述本发明的各种实施例。这些和其它实施例在所附权利要求书的范围内。

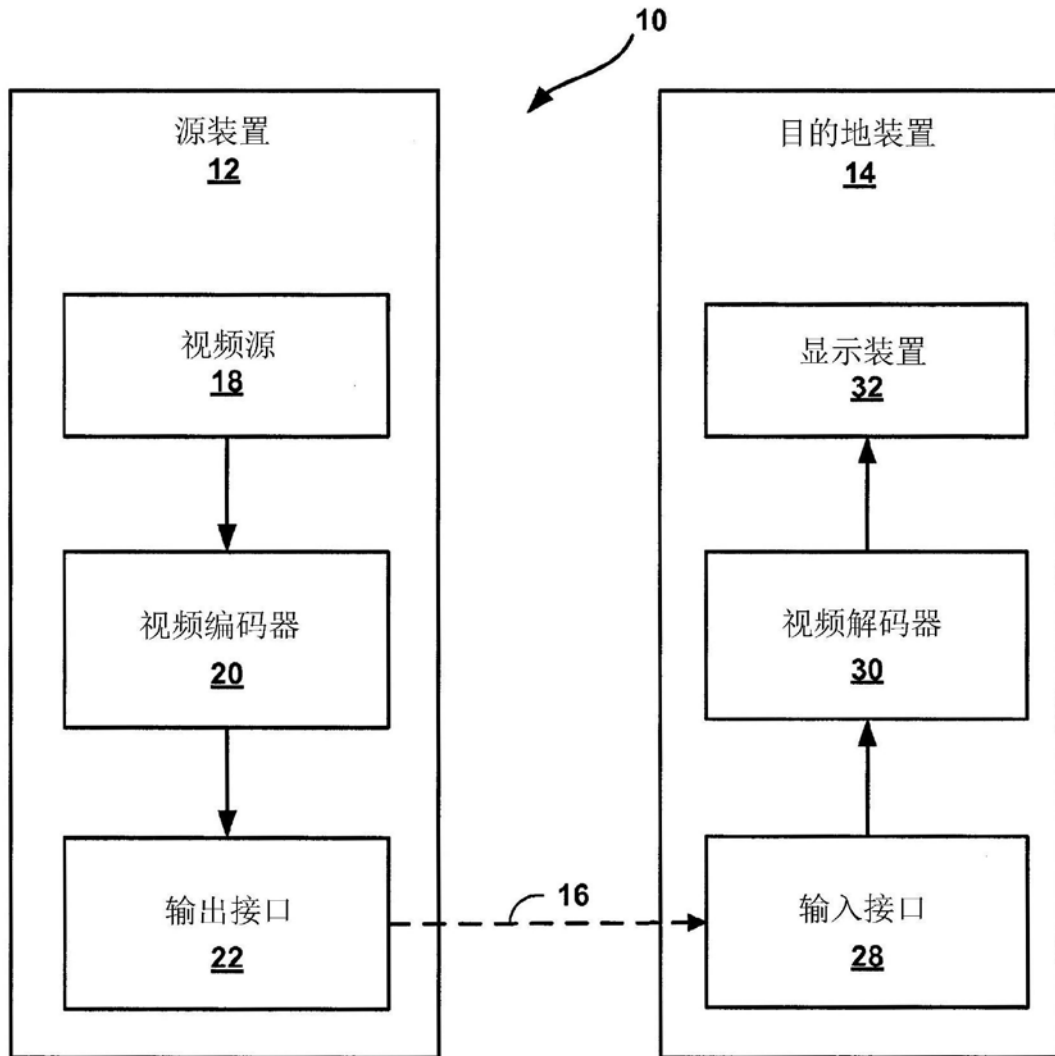


图1

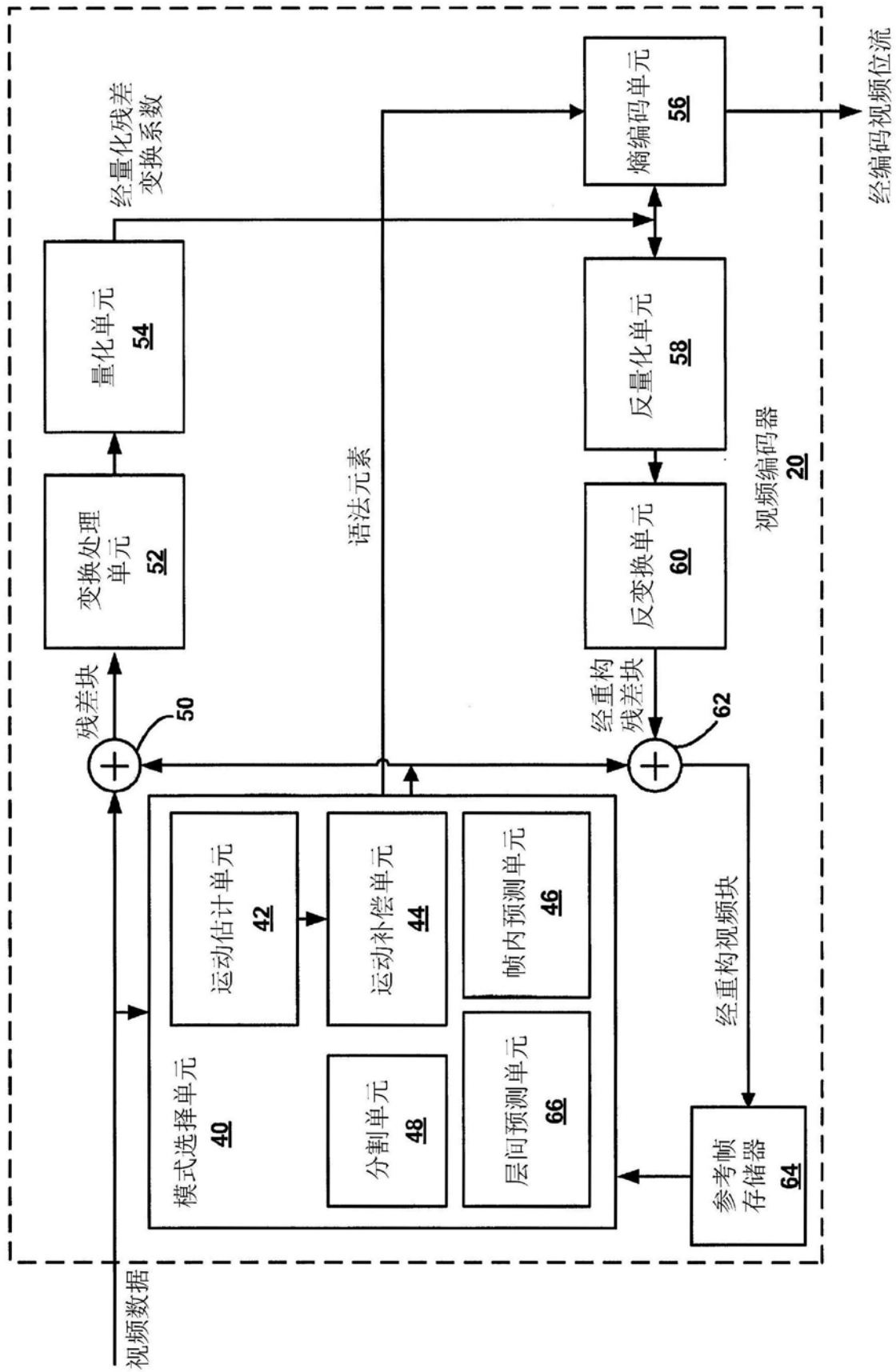


图2A

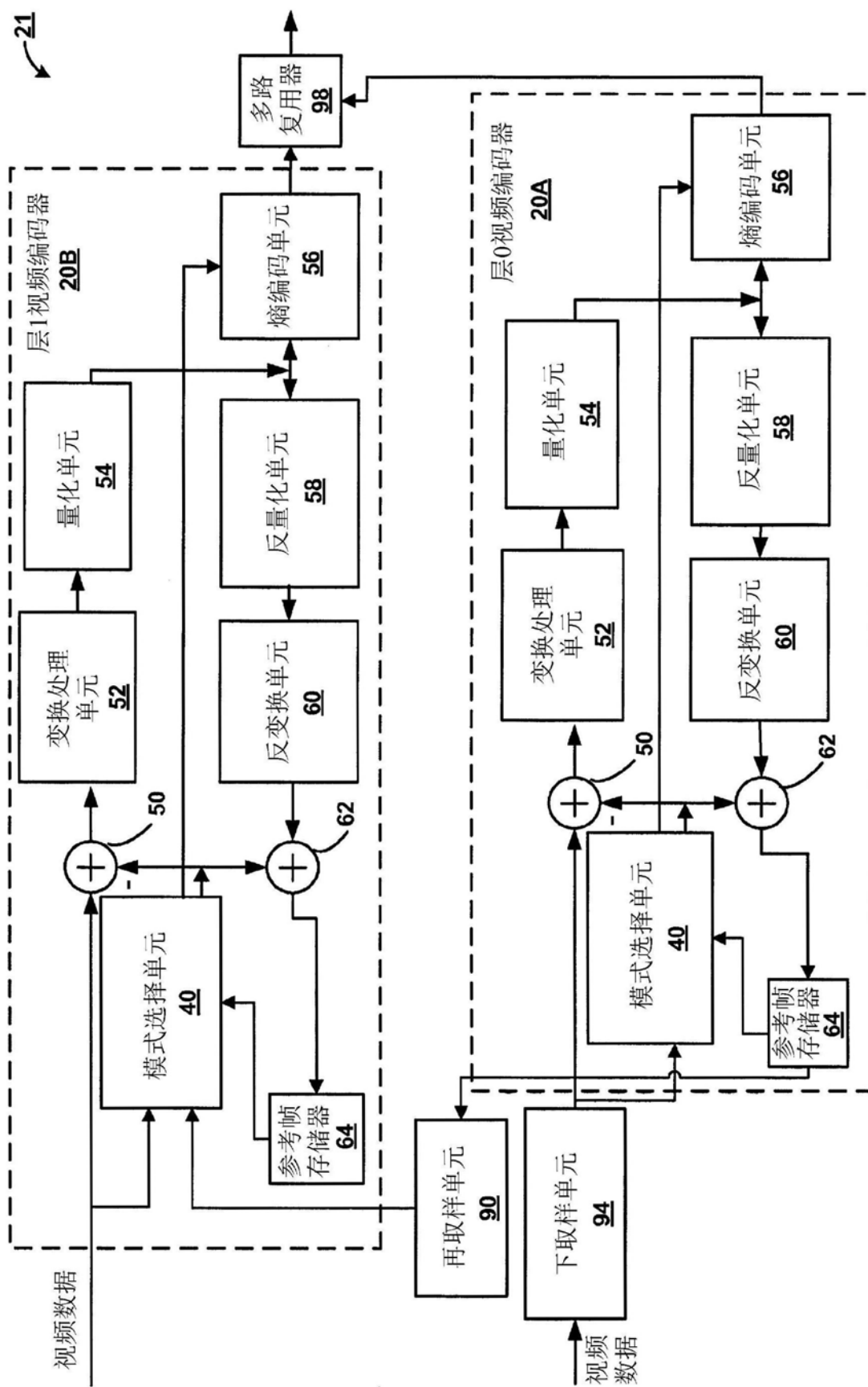


图2B

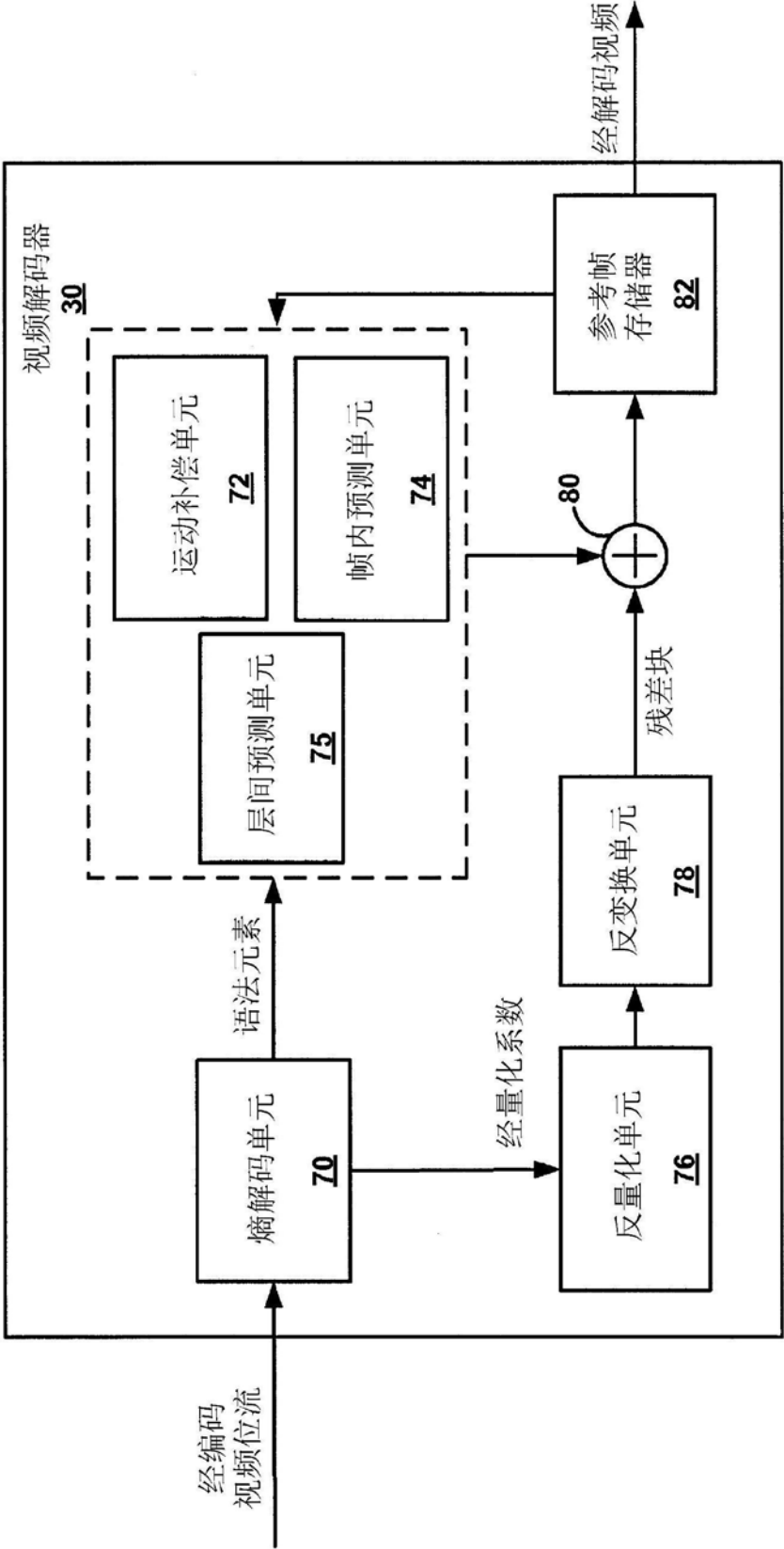


图3A

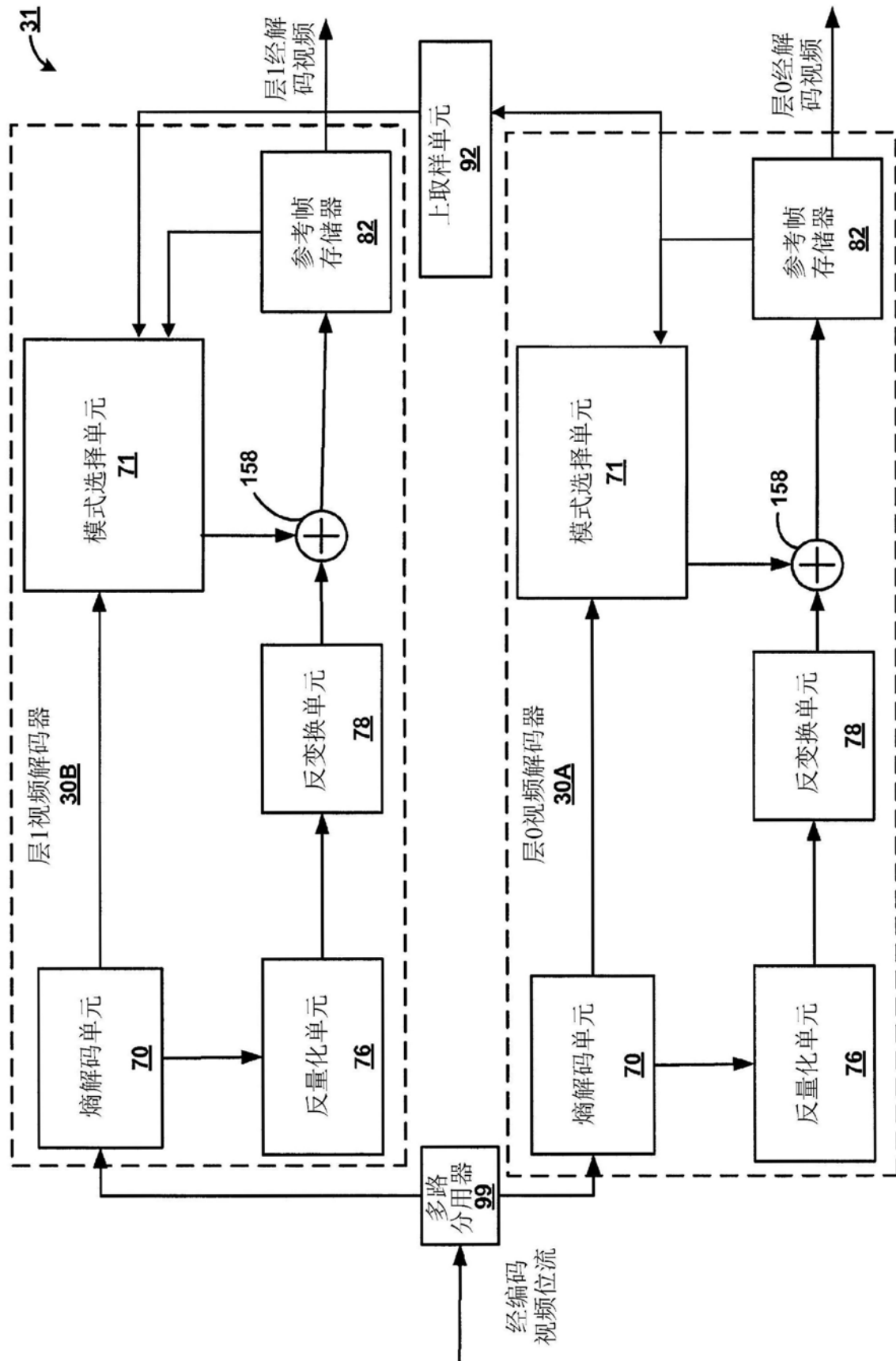


图3B

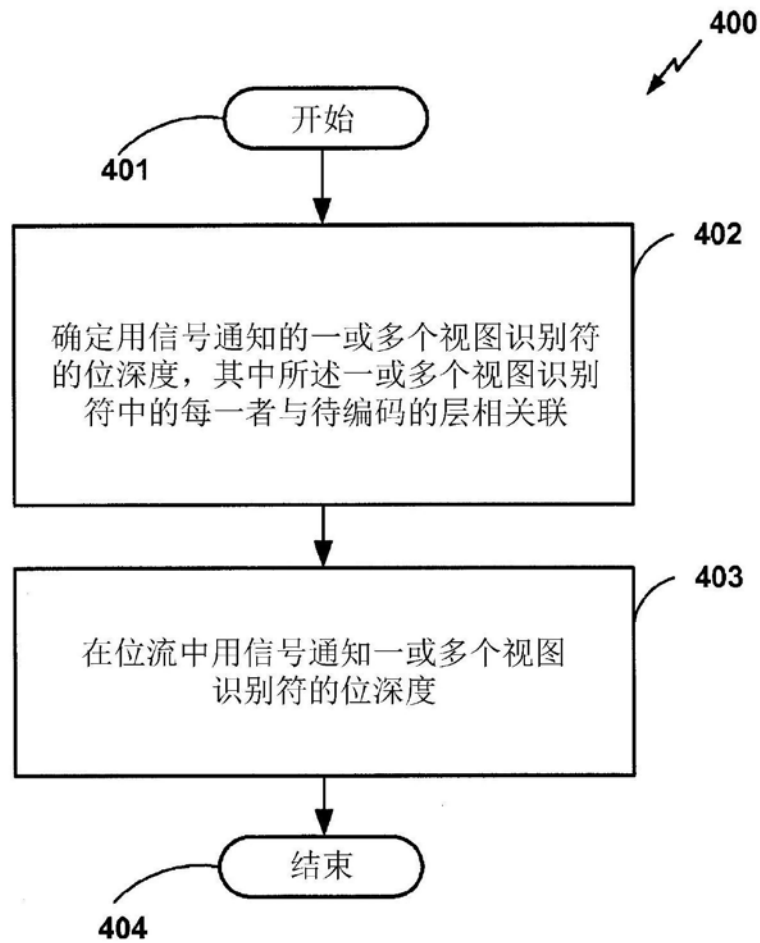


图4



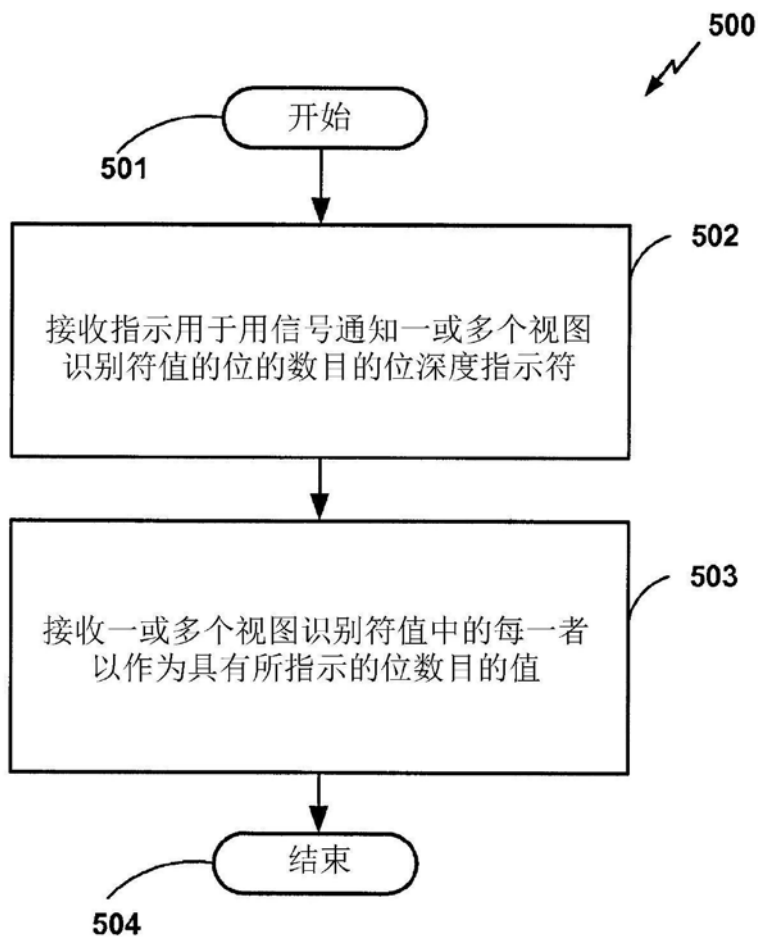


图5

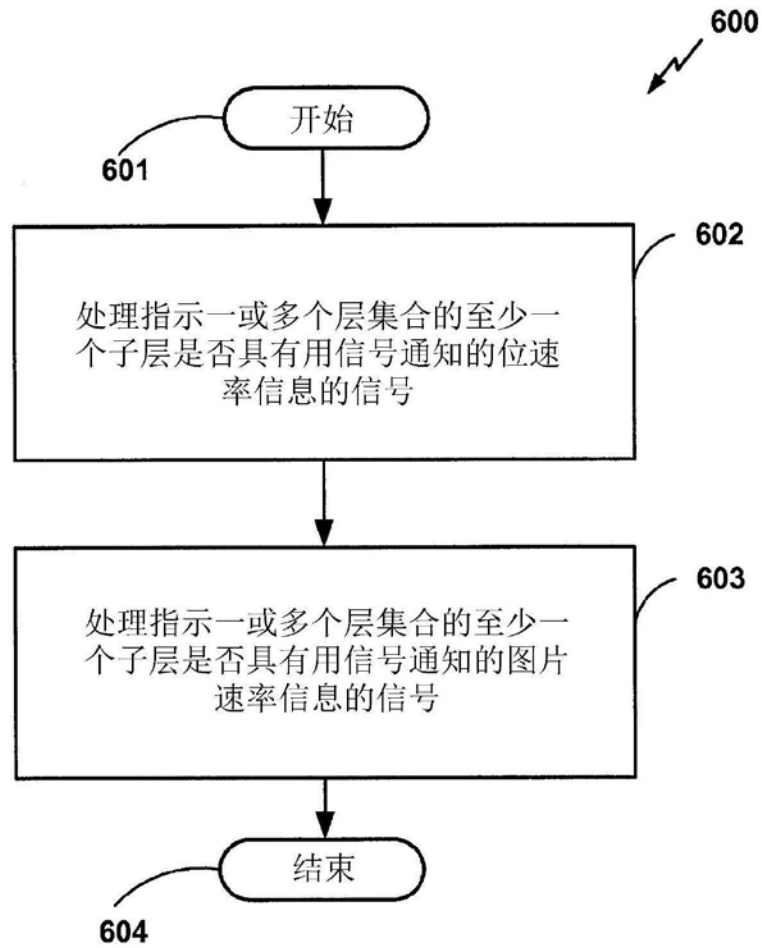


图6