



(12)发明专利

(10)授权公告号 CN 106105209 B

(45)授权公告日 2019.04.23

(21)申请号 201580012479.3

(22)申请日 2015.03.16

(65)同一申请的已公布的文献号
申请公布号 CN 106105209 A

(43)申请公布日 2016.11.09

(30)优先权数据

61/954,495 2014.03.17 US

14/656,573 2015.03.12 US

(85)PCT国际申请进入国家阶段日
2016.09.07

(86)PCT国际申请的申请数据
PCT/US2015/020662 2015.03.16

(87)PCT国际申请的公布数据
W02015/142694 EN 2015.09.24

(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
陈建乐 伏努·亨利
阿达许·克里许纳·瑞玛苏布雷蒙
尼安

(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 19/30(2014.01)

H04N 19/156(2014.01)

H04N 19/44(2014.01)

审查员 刘晶

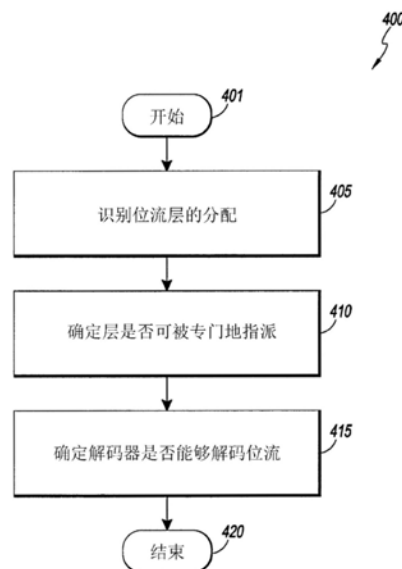
权利要求书2页 说明书34页 附图12页

(54)发明名称

多层视频编解码器的级别定义

(57)摘要

本发明揭示用于定义用于解码含有视频信息的多层位流的解码器能力的方法,其中所述解码器基于多个单层解码器核心而被实施。在一个方面中,所述方法可包含识别所述位流的层至少一次分配到至少一层组中。所述方法可进一步包含检测每一层组是否能够专门地指派给所述解码器核心中的一者以用于解码所述位流。所述方法还可包含至少部分地基于检测每一层组是否能够专门地指派给所述解码器核心中的一者而确定所述解码器是否能够解码所述位流。



1. 一种用于确定解码器是否能够解码含有视频信息的多层位流的方法,所述解码器基于多个单层解码器核心而被实施,所述方法包括:

识别所述位流的层至少一次分配到多层组中;

检测是否有足够的解码器核心以使得所述多层组的每一者可被专门地指派给所述解码器核心中的一者以用于解码所述位流;

至少部分地基于所述解码器的级别定义来确定所述解码器的最大图片大小;以及

至少部分地基于 (i) 所述多层组的每一者是否可被专门地指派给所述解码器核心中的一者以及 (ii) 经确定的所述最大图片大小而确定所述解码器是否能够解码所述位流。

2. 根据权利要求1所述的方法,其进一步包括指定与所述级别定义相关联的尺度,其中所述确定所述解码器是否能够解码所述位流进一步至少部分地基于所述尺度。

3. 根据权利要求2所述的方法,其进一步包括计算所述位流中的层的分辨率的总和,其中所述确定所述解码器是否能够解码所述位流进一步至少部分地基于所述所计算的分辨率总和。

4. 根据权利要求3所述的方法,其中所述确定所述解码器是否能够解码所述位流进一步至少部分地基于所述分辨率总和是否小于或等于通过将所述解码器的所述尺度乘以所述最大图片大小而产生的乘积。

5. 根据权利要求3所述的方法,其进一步包括:

确定所述位流的最高层的分辨率;

确定所述位流的标识符列表中的层的数目;以及

经由将所述标识符列表中的层的所述数目乘以所述最高层的所述分辨率来计算所述分辨率总和。

6. 根据权利要求2所述的方法,其进一步包括经由确定所述位流中的层的数目来确定所述位流的级别定义。

7. 根据权利要求6所述的方法,其进一步包括至少部分地基于所述位流的所述级别定义而确定所述位流的最大图片大小。

8. 根据权利要求7所述的方法,其中所述确定所述解码器是否能够解码所述位流进一步至少部分地基于所述位流的所述最大图片大小。

9. 根据权利要求8所述的方法,其进一步包括:

确定通过将所述位流中的层的所述数目乘以所述位流的所述最大图片大小而产生的第一乘积是否小于或等于通过将所述解码器的所述尺度乘以所述最大图片大小而产生的第二乘积;以及

至少部分地基于所述第一乘积是否小于或等于所述第二乘积而确定所述解码器是否能够解码所述位流。

10. 根据权利要求2所述的方法,其进一步包括对每一层、经解码图片缓冲器DPB或子DPB中的至少一者指定大小相关的限制。

11. 一种经配置以确定解码器是否能够解码含有视频信息的多层位流的视频译码装置,所述解码器基于多个单层解码器核心而被实施,所述视频译码装置包括:

用于识别所述位流的层至少一次分配到多层组中的装置;

用于检测是否有足够的解码器核心以使得所述多层组的每一者可被专门地指派给所

述解码器核心中的一者以用于解码所述位流的装置；

用于至少部分地基于所述解码器的级别定义来确定所述解码器的最大图片大小的装置；

用于至少部分地基于(i)所述多层组的每一者是否可被专门地指派给所述解码器核心中的一者以及(ii)经确定的所述最大图片大小而确定所述解码器是否能够解码所述位流的装置。

12. 根据权利要求11所述的视频译码装置,其进一步包括用于指定与所述级别定义相关联的尺度的装置,其中所述用于确定所述解码器是否能够解码所述位流的装置考虑所述尺度。

13. 根据权利要求12所述的视频译码装置,其进一步包括用于计算所述位流中的层的分辨率的总和的装置,其中所述用于确定所述解码器是否能够解码所述位流的装置考虑所述所计算的分辨率总和。

14. 根据权利要求13所述的视频译码装置,其中所述用于确定所述解码器是否能够解码所述位流的装置包括用于确定所述分辨率总和是否小于或等于通过将所述解码器的所述尺度乘以所述最大图片大小而产生的乘积的装置。

15. 根据权利要求14所述的视频译码装置,其进一步包括:

用于确定所述位流的最高层的分辨率的装置;

用于确定所述位流的标识符列表中的层的数目的装置;以及

用于经由将所述标识符列表中的层的所述数目乘以所述最高层的所述分辨率来计算所述分辨率总和的装置。

16. 一种非暂时性计算机可读存储媒体,其具有存储于其上的指令,所述指令在经执行时致使装置的处理器的处理器执行权利要求1-10的任一项所述的方法。

多层视频编解码器的级别定义

技术领域

[0001] 本发明涉及视频译码的领域,且更确切来说涉及在含有视频信息的多层位流的上下文中的解码器能力。

背景技术

[0002] 数字视频能力可并入到广泛范围的装置中,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)、膝上型或台式计算机、数码相机、数字记录装置、数字媒体播放器、视频游戏装置、视频游戏机、蜂窝式或卫星无线电电话、视频电话会议装置等等。数字视频装置实施视频压缩技术,例如在由运动图片专家组-2(MPEG-2)、MPEG-4、国际电信联盟电信标准化部门(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、高效率视频译码(HEVC)标准和这些标准的扩展中描述的视频压缩技术。视频装置可通过实施此类视频译码技术而更有效地传输、接收、编码、解码和/或存储数字视频信息。

[0003] 例如可缩放HEVC(SHVC)和多视图HEVC(MV-HEVC)等视频译码标准提供用于定义解码器能力的级别定义。在下文中,基于在进行本发明时SHVC的现有的级别定义和其它上下文而描述问题和解决方案,但解决方案也适用于MV-HEVC以及其它多层编解码器。

发明内容

[0004] 本发明的系统、方法和装置各自具有若干创新方面,其中没有单个方面单独负责本文中所揭示的合乎需要的属性。

[0005] 在一个方面中,一种用于确定基于多个单层解码器核心实施的解码器是否能够解码含有视频信息的多层位流的方法包括:识别所述位流的层至少一次分配到至少一层组中;检测每一层组是否能够专门地指派给所述解码器核心中的一者以用于解码所述视频信息;以及至少部分地基于检测每一层组是否能够专门地指派给所述解码器核心中的一者而确定所述解码器是否能够解码所述位流。

[0006] 在另一方面中,一种用于确定基于多个单层解码器核心实施的解码器是否能够解码含有视频信息的多层位流的设备包括至少一个处理器,所述至少一个处理器经配置以识别所述位流的层至少一次分配到至少一层组中、检测每一层组是否能够专门地指派给所述解码器核心中的一者以用于解码所述位流且至少部分地基于检测每一层组是否能够专门地指派给所述解码器核心中的一者而确定所述解码器是否能够解码所述位流。

[0007] 在另一方面中,一种非暂时性计算机可读存储媒体上面存储有指令,所述指令在经执行时致使装置的处理器识别含有视频信息的多层位流的层至少一次分配到至少一层组中以用于确定基于多个单层解码器核心实施的解码器是否能够解码所述位流,检测每一层组是否能够专门地指派给所述解码器核心中的一者以用于解码所述位流,且至少部分地基于检测每一层组是否能够专门地指派给所述解码器核心中的一者而确定所述解码器是否能够解码所述位流。

[0008] 在又一方面中,一种经配置以确定基于多个单层解码器核心实施的解码器是否能

够解码含有视频信息的多层位流的视频译码装置包括：用于识别所述位流的层至少一次分配到至少一层组中的装置；用于检测每一层组是否能够专门地指派给所述解码器核心中的一者以用于解码所述位流的装置；以及用于至少部分地基于检测每一层组是否能够专门地指派给所述解码器核心中的一者而确定所述解码器是否能够解码所述位流的装置。

附图说明

[0009] 图1A为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编码和解码系统的框图。

[0010] 图1B为说明可执行根据本发明中所描述的方面的技术的另一实例视频编码和解码系统的框图。

[0011] 图2A为说明可实施根据本发明中所描述的方面的技术的视频编码器的实例的框图。

[0012] 图2B为说明可实施根据本发明中所描述的方面的技术的视频编码器的另一实例的框图。

[0013] 图3A为说明可实施根据本发明中所描述的方面的技术的视频解码器的实例的框图。

[0014] 图3B为说明可实施根据本发明中所描述的方面的技术的视频解码器的另一实例的框图。

[0015] 图4A为根据本发明中所描述的方面的用于定义解码器能力的过程的示范性实施例的流程图。

[0016] 图4B为根据本发明中所描述的方面的用于进一步定义解码器能力的过程的示范性实施例的流程图。

[0017] 图5A说明根据本发明中所描述的方面的示范性多层位流。

[0018] 图5B说明根据本发明中所描述的方面的用于解码图5A的多层位流的多个示范性选项。

[0019] 图6A说明根据本发明中所描述的方面的另一示范性多层位流。

[0020] 图6B说明根据本发明中所描述的方面的用于解码图6A的多层位流的多个示范性选项。

具体实施方式

[0021] 大体上，本发明涉及对多层视频编解码器的上下文中的位流和解码器二者指定级别的改进，使得级别限制可随层的数目缩放从而允许以高效方式实现高图片质量以及其它益处。

[0022] 视频译码标准包含ITU-T H.261、ISO/IEC MPEG-1Visual、ITU-T H.262或ISO/IEC MPEG-2Visual、ITU-T H.263、ISO/IEC MPEG-4Visual和ITU-T H.264(也称为ISO/IEC MPEG-4AVC)，包含其可缩放视频译码(SVC)和多视图视频译码(MVC)扩展。

[0023] 另外，ITU-T视频译码专家组(VCEG)和ISO/IEC MPEG的视频译码联合合作小组(JCT-VC)已开发一种视频译码标准，即高效率视频译码(HEVC)。对HEVC草案10的完全引用为Bross等人的文件JCTVC-L1003，“高效率视频译码(HEVC)文本规范草案10(High

Efficiency Video Coding (HEVC) Text Specification Draft 10)” (ITU-T SG16WP3与ISO/IEC JTC1/SC29/WG11的视频译码联合合作小组 (JCT-VC), 第12次会议: 瑞士日内瓦, 2013年1月14日到2013年1月23日)。对HEVC的多视图扩展 (即, MV-HEVC) 和对HEVC的可缩放扩展 (名为SHVC) 也正分别由JCT-3V (ITU-T/ISO/IEC 3D视频译码扩展开发联合合作小组) 和JCT-VC开发。MV-HEVC的最近工作草案 (WD) 在下文中将被称作MV-HEVC WD7。SHVC的最近WD在下文中将被称作SHVC WD5。

[0024] 级别定义的现有方法有时并不提供足以定义解码器能力以用于高效解码多层位流的信息。举例来说, 为了解码多于4个各自具有720p分辨率的信噪比 (SNR) 可缩放层 (具有等效分辨率的层), 将需要5级或更高的解码器。因此, 亮度译码树块 (CTB) 大小将等于 32×32 或 64×64 (即, 不可使用较小的译码大小, 例如 16×16)。然而, 对于一些层 (例如, 具有720p或更低的分辨率的那些层), 此限制可能导致次最佳的译码效率。

[0025] 在一些情况下可通过重复使用多个现有单层解码器来制造解码器。在实例中, 按照现有级别定义, 由4个单层HEVC 3.1级解码器组成的SHVC解码器必须符合4级或更高以解码4个720p的SNR层。根据此定义, 解码器必须能够解码任何4级位流。然而, 除非改变解码器硬件, 否则此解码器不能够解码具有2个1080p分辨率的SNR层的SHVC 4级位流。

[0026] 现有HEVC级别定义的另一问题为, 以使得能够解码1080p的单层HEVC位流和720p的双层SHVC位流二者的方式实施的解码器被标注为3.1级。然而, 3.1级标签并不表达解码1080p的单层位流的能力。

[0027] 在另一实例中, 对于使用4个单层HEVC 3.1解码器实施以能够解码4个720p的SNR层的解码器, 按照现有级别定义, 解码器必须符合4级或更高。因此, 解码器需要能够解码具有多于3个平铺块行和多于3个平铺块列的位流, 每一平铺块具有256个亮度样本的宽度和144个亮度样本的高度。然而, 解码器的3.1级限制不能够解码一些此类位流。

[0028] 在现有的SHVC设计下, HEVC文本的子条款A.4.1中的所有项目被指定为应用于每一层。然而, 一些项目并非直接地可适用于每一层。举例来说, 对于关于经解码图片缓冲器 (DPB) 大小的第d项, 序列参数集 (SPS) 语法元素不可适用于增强层。另外, SHVC WD5中的DPB为共享子DPB设计, 因此第d项不可直接地应用于每一层。作为另一实例, 对于关于经译码图片缓冲器 (CPB) 大小的第h和i项, 针对位流特定的CPB操作, 所述参数不可应用于每一层。

[0029] 需要对CPB大小的位流特定限制 (根据HEVC文本的子条款A.4.1中的第h和i项)。然而, HEVC文本的子条款A.4.1中的第h和i项不可直接地应用于位流级别, 因为如果直接地应用, 那么对单层位流的相同CPB大小限制也将为对多层位流的限制。此限制不可随层的数目缩放且在存在许多层时仅允许低图片质量。

[0030] HEVC文本的子条款A.4.2中的第b、c、d、g、h、i和j项的限制仅被指定为层特定的。然而, 应指定这些项的位流特定限制, 无论其层特定的对应限制是否被指定。

[0031] 虽然本文中在HEVC和/或H.264标准的上下文中描述某些实施例, 但所属领域的技术人员可了解, 本文中所揭示的系统和方法可适用于任何合适的视频译码标准或非标准视频编解码器设计。举例来说, 本文中所揭示的实施例可适用于以下标准中的一或多个: 国际电信联盟 (ITU) 电信标准化部门 (ITU-T) H.261、国际标准化组织/国际电工委员会 (ISO/IEC) MPEG-1 Visual、ITU-T H.262或ISO/IEC MPEG-2 Visual、ITU-T H.263、ISO/IEC MPEG-4 Visual和ITU-T H.264 (也称为ISO/IEC MPEG-4 AVC), 包含可缩放和多视图扩展。

[0032] 在许多方面,HEVC一般遵循先前视频译码标准的框架。HEVC中的预测单元不同于某些先前视频译码标准中的预测单元(例如,宏块)。实际上,在HEVC中不存在如在某些先前视频译码标准中所理解的宏块的概念。宏块由基于四叉树方案的分层结构替换,分层结构可提供高灵活性以及其它可能益处。举例来说,在HEVC方案内,定义三个类型的块:译码单元(CU)、预测单元(PU)和变换单元(TU)。CU可指区分裂的基本单元。可将CU视为类似于宏块的概念,但HEVC不限制CU的最大大小,且可允许递归分裂成四个大小相等的CU以改进内容适应性。可将PU视为帧间/帧内预测的基本单元,且单个PU可含有多个任意形状分区以有效地对不规则图像图案译码。可将TU视为变换的基本单元。可独立于PU定义TU;然而,TU的大小可限于TU属于的CU的大小。此种将块结构分离成三个不同概念可允许根据单元的相应作用来优化每一单元,这可导致改进的译码效率。

[0033] 仅出于说明的目的,用仅包含视频数据的两个层(例如,例如基础层等较低层,和例如增强层等较高层)的实例来描述本文中所揭示的某些实施例。视频数据的“层”一般可指具有至少一个共同特性(例如,视图、帧速率、分辨率等等)的图片序列。举例来说,层可包含与多视图视频数据的特定视图(例如,透视图)相关联的视频数据。作为另一实例,层可包含与可缩放视频数据的特定层相关联的视频数据。因此,本发明可互换地提及视频数据的层和视图。也就是说,视频数据的视图可被称作视频数据的层,且视频数据的层可被称作视频数据的视图。另外,多层编解码器(还被称作多层视频译码器或多层编码器-解码器)可共同指多视图编解码器或可缩放编解码器(例如,经配置以使用MV-HEVC、3D-HEVC、SHVC或另一多层译码技术编码和/或解码视频数据的编解码器)。视频编码和视频解码可一般被称作视频译码。应理解,这些实例可适用于包含多个基础层和/或增强层的配置。另外,为了易于解释,参考某些实施例,以下揭示内容包含术语“帧”或“块”。然而,这些术语不打算限制性的。举例来说,下文所描述的技术可与任何合适的视频单元(例如,块(例如,CU、PU、TU、宏块等)、切片、帧等)一起使用。

[0034] 视频译码标准

[0035] 例如视频图像、TV图像、静态图像或由录像机或计算机产生的图像等数字图像可由布置成水平和垂直线的像素或样本构成。单个图像中的像素的数目通常有数万个。每一像素通常含有亮度和色度信息。在无压缩的情况下,将从图像编码器传达到图像解码器的信息的绝对量使实时图像传输变得不可能。为了减少待传输的信息量,已开发出例如JPEG、MPEG和H.263标准等数个不同压缩方法。视频译码标准包含本文中先前叙述的那些。

[0036] 视频译码系统

[0037] 下文参考附图更充分地描述新颖系统、设备和方法的各种方面。然而,本发明可以许多不同形式来体现,且不应被解释为限于贯穿本发明所呈现的任何特定结构或功能。实际上,提供这些方面以使得本发明将透彻且完整,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本文中所揭示的无论是独立于本发明的任何其它方面还是与之组合而实施的新颖系统、设备和方法的任何方面。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。另外,本发明的范围既定涵盖使用除本文中所阐述的本发明的各种方面之外的或不同于本文中所阐述的本发明的各种方面的其它结构、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。

[0038] 尽管本文中描述了特定方面,但这些方面的许多变化和排列属于本发明的范围内。尽管提到了优选方面的一些益处和优点,但本发明的范围不希望限于特定益处、用途或目标。实际上,本发明的方面希望广泛地适用于不同无线技术、系统配置、网络和传输协议,其中的一些是借助于实例而在图中和在优选方面的以下描述中加以说明。详细描述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定义的本发明的范围。

[0039] 附图说明若干实例。由附图中的参考数字指示的元件对应于在以下描述中由相同参考数字指示的元件。在本发明中,名称以序数词(例如,“第一”、“第二”、“第三”等等)开始的元件未必暗示所述元件具有特定次序。实际上,这些序数词仅用于指代相同或类似类型的不同元件。

[0040] 图1A为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。如本文中所使用,术语“视频译码器”一般指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”可一般地指视频编码和视频解码。除了视频编码器和视频解码器之外,本申请案中所描述的方面还可扩展到其它相关装置,例如,转码器(例如,可解码位流且重新编码另一位流的装置)和中间框(例如,可修改、变换和/或以其它方式操纵位流的装置)。

[0041] 如图1A中所展示,视频译码系统10包含源装置12,其产生在稍后时间由目的地装置14解码的经编码视频数据。在图1A的实例中,源装置12和目的地装置14构成单独装置。然而,应注意,源装置12和目的地装置14可在同一装置上或为其部分,如图1B的实例中所展示。

[0042] 再次参考图1A,源装置12和目的地装置14可分别包括广泛范围的装置中的任一者,包含台式计算机、笔记本(例如,膝上型)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”板、电视、相机、显示装置、数字媒体播放器、视频游戏机、视频流式传输装置等等。在各种实施例中,源装置12和目的地装置14可经装备以用于无线通信。

[0043] 目的地装置14可经由链路16接收待解码的经编码视频数据。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。在图1A的实例中,链路16可包括使得源装置12能够实时将经编码视频数据传输到目的地装置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)来调制,且被传输到目的地装置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。通信媒体可能形成基于数据包的网络(例如,局域网、广域网或全球网络,例如因特网)的部分。通信媒体可包含路由器、交换机、基站或可用于促进从源装置12到目的地装置14的通信的任何其它设备。

[0044] 替代地,经编码数据可从输出接口22输出到存储装置31(任选地存在)。类似地,可由(例如)目的地装置14的输入接口28从存储装置31存取经编码数据。存储装置31可包含多种分布式或本地存取数据存储媒体中的任一者,例如硬盘驱动器、闪存、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31可对应于文件服务器或可保持由源装置12产生的经编码视频的另一中间存储装置。目的地装置14可经由流式传输或下载从存储装置31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将所述经编码视频数据传输到目的地装置14的任何类型的

服务器。实例文件服务器包含网页服务器(例如,用于网站)、文件传送协议(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。此可包含无线信道(例如,无线局域网(WLAN)连接)、有线连接(例如,数字用户线(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31的传输可为流式传输、下载传输或两者的组合。

[0045] 本发明的技术不限于无线应用或设置。所述技术可应用于支持多种多媒体应用中的任一者的视频译码,例如空中电视广播、有线电视传输、卫星电视传输、例如经由因特网的流视频传输(例如,超文本传送协议(HTTP)动态自适应流式传输等)、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,视频译码系统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/或视频电话等应用。

[0046] 在图1A的实例中,源装置12包含视频源18、视频编码器20和输出接口22。在一些情况下,输出接口22可包含调制器/解调器(调制解调器)和/或发射器。在源装置12中,视频源18可包含例如视频捕获装置(例如,摄像机)、含有先前所捕获视频的视频存档、用以从视频内容提供者接收视频的视频馈入接口和/或用于产生计算机图形数据以作为源视频的计算机图形系统等源,或此类源的组合。作为一个实例,如果视频源18是摄像机,那么源装置12和目的地装置14可形成所谓的“相机电话”或“视频电话”,如图1B的实例中所说明。然而,本发明中所描述的技术一般可适用于视频译码,且可应用于无线和/或有线应用。

[0047] 可由视频编码器20对所捕获、预捕获或计算机产生的视频进行编码。可经由源装置12的输出接口22将经编码视频数据传输到目的地装置14。经编码视频数据还可(或替代地)存储到存储装置31上以供稍后由目的地装置14或其它装置存取以用于解码和/或回放。图1A和1B中所说明的视频编码器20可包括图2A中所说明的视频编码器20、图2B中所说明的视频编码器23或本文中所描述的任何其它视频编码器。

[0048] 在图1A的实例中,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。在一些情况下,输入接口28可包含接收器和/或调制解调器。目的地装置14的输入接口28可经由链路16和/或从存储装置31接收经编码视频数据。经由链路16传达或在存储装置31上提供的经编码视频数据可包含由视频编码器20产生的多种语法元素以供由视频解码器(例如,视频解码器30)用于解码视频数据。此类语法元素可与在通信媒体上传输、存储在存储媒体上或存储文件服务器的经编码视频数据包含在一起。图1A和1B中所说明的视频解码器30可包括图3A中所说明的视频解码器30、图3B中所说明的视频解码器33或本文中所描述的任何其它视频解码器。

[0049] 显示装置32可与目的地装置14集成或在目的地装置14外部。在一些实例中,目的地装置14可包含集成显示装置,且还可经配置以与外部显示装置介接。在其它实例中,目的地装置14可为显示装置。一般来说,显示装置32将经解码视频数据显示给用户,且可包括多种显示装置中的任一者,例如液晶显示器(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。

[0050] 在相关方面中,图1B展示实例视频译码系统10',其中源装置12和目的地装置14在装置11上或为其部分。装置11可为电话手持机,例如“智能”电话等等。装置11可包含与源装

置12和目的地装置14可操作通信的处理器/控制器装置13(任选地存在)。图1B的视频译码系统10'可进一步包含视频编码器20与输出接口22之间的视频处理单元21。在一些实施方案中,视频处理单元21为单独单元,如图1B中所说明;然而,在其它实施方案中,视频处理单元21可实施为视频编码器20和/或处理器/控制器装置13的一部分。视频译码系统10'还可包含跟踪器29(任选地存在),其可跟踪视频序列中的所关注对象。待跟踪的所关注对象可通过结合本发明的一或多个方面描述的技术来分段。在相关方面中,可由显示装置32单独地或与跟踪器29联合来执行跟踪。图1B的视频译码系统10'和其组件在其它方面类似于图1A的视频译码系统10和其组件。

[0051] 视频编码器20和视频解码器30可根据例如HEVC等视频压缩标准来操作,且可符合HEVC测试模型(HM)。替代地,视频编码器20和视频解码器30可根据其它专有或行业标准或此类标准的扩展来操作,所述标准例如ITU-T H.264标准,或者被称作MPEG-4第10部分AVC。然而,本发明的技术不限于任何特定译码标准。视频压缩标准的其它实例包含MPEG-2和ITU-T H.263。

[0052] 尽管图1A和1B的实例中未展示,但视频编码器20和视频解码器30可各自与音频编码器和解码器集成,且可包含适当的MUX-DEMUX单元或其它硬件和软件以处理共同数据流或单独数据流中的音频和视频两者的编码。在一些实例中,如果适用的话,那么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。

[0053] 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。当部分地在软件中实施所述技术时,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并使用一或多个处理器在硬件中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式编码器/解码器(CODEC)的部分。

[0054] 视频译码过程

[0055] 如上文简要提及,视频编码器20对视频数据进行编码。视频数据可包括一或多个图片。图片中的每一者为形成视频的部分的静态图像。在一些情况下,图片可被称作视频“帧”。当视频编码器20对视频数据进行编码时,视频编码器20可产生位流。位流可包含形成视频数据的经译码表示的一连串位。位流可包含经译码图片和相关联的数据。经译码图片为图片的经译码表示。

[0056] 为了产生位流,视频编码器20可对视频数据中的每一图片执行编码操作。当视频编码器20对图片执行编码操作时,视频编码器20可产生一系列经译码图片和相关联数据。相关联的数据可包含视频参数集(VPS)、序列参数集(SPS)、图片参数集(PPS)、自适应参数集(APS)和其它语法结构。SPS可含有适用于图片的零或更多序列的参数。PPS可含有适用于零或更多图片的参数。APS可含有适用于零或更多图片的参数。APS中的参数可为比PPS中的参数更可能改变的参数。

[0057] 为了产生经译码图片,视频编码器20可将图片分割成大小相等的视频块。视频块可为样本的二维阵列。视频块中的每一者与树块相关联。在一些情况下,树块可被称作最大译码单元(LCU)。HEVC的树块可广泛类似于例如H.264/AVC等先前标准的宏块。然而,树块不

必限于特定大小,且可包含一或多个译码单元(CU)。视频编码器20可使用四叉树分割将树块的视频块分割成与CU相关联的视频块(因此名称为“树块”)。

[0058] 在一些实例中,视频编码器20可将图片分割成多个切片。切片中的每一者可包含整数数目个CU。在一些情况下,切片包括整数数目个树块。在其它情况下,切片的边界可在树块内。

[0059] 作为对图片执行编码操作的部分,视频编码器20可对图片的每一切片执行编码操作。当视频编码器20对切片执行编码操作时,视频编码器20可产生与切片相关联的经编码数据。与切片相关联的经编码数据可被称作“经译码切片”。

[0060] 为了产生经译码切片,视频编码器20可对切片中的每一树块执行编码操作。当视频编码器20对树块执行编码操作时,视频编码器20可产生经译码树块。经译码树块可包括表示树块的经编码版本的数据。

[0061] 当视频编码器20产生经译码切片时,视频编码器20可根据光栅扫描次序对切片中的树块执行编码操作(例如,编码)。举例来说,视频编码器20可按以下次序对切片的树块进行编码:跨越切片中的树块的最顶行从左到右进行,接着跨越树块的下一较低行从左到右进行,以此类推,直到视频编码器20已对切片中的树块的每一者进行编码为止。

[0062] 作为根据光栅扫描次序对树块进行编码的结果,可能已对给定树块的上方和左方的树块进行编码,但尚未对给定树块的下方和右方的树块进行编码。因此,当对给定树块进行编码时,视频编码器20可能能够存取通过对在给定树块的上方和左方的树块进行编码而产生的信息。然而,当对给定树块进行编码时,视频编码器20可能不能够存取通过对在给定树块的下方和右方的树块进行编码而产生的信息。

[0063] 为了产生经译码树块,视频编码器20可对树块的视频块递归地执行四叉树分割以将所述视频块划分成逐渐变小的视频块。较小视频块中的每一者可与不同CU相关联。举例来说,视频编码器20可将树块的视频块分割成四个大小相等的子块、将所述子块中的一或多者分割成四个大小相等的子子块(sub-sub-block),以此类推。经分割CU可为视频块被分割成与其它CU相关联的视频块的CU。未经分割的CU可为视频块未被分割成与其它CU相关联的视频块的CU。

[0064] 位流中的一或多个语法元素可指示视频编码器20可分割树块的视频块的最大次数。CU的视频块可为正方形形状。CU的视频块的大小(例如,CU的大小)范围可从 8×8 像素直到具有 64×64 像素或更大的最大值的树块的视频块的大小(例如,树块的大小)。

[0065] 视频编码器20可根据z扫描次序对树块的每一CU执行编码操作(例如,编码)。换句话说,视频编码器20可按所述次序对左上CU、右上CU、左下CU和接着右下CU进行编码。当视频编码器20对经分割CU执行编码操作时,视频编码器20可根据z扫描次序对与经分割CU的视频块的子块相关联的CU进行编码。换句话说,视频编码器20可按所述次序对与左上子块相关联的CU、与右上子块相关联的CU、与左下子块相关联的CU和接着与右下子块相关联的CU进行编码。

[0066] 作为根据z扫描次序对树块的CU进行编码的结果,可已对给定CU的上方、左上方、右上方、左方和左下方的CU进行编码。尚未对给定CU的右下方的CU进行编码。因此,当对给定CU进行编码时,视频编码器20可能能够存取通过对与给定CU相邻的一些CU进行编码而产生的信息。然而,当对给定CU进行编码时,视频编码器20可能不能够存取通过对与给定CU相

邻的其它CU进行编码而产生的信息。

[0067] 当视频编码器20对未经分割的CU进行编码时,视频编码器20可产生用于所述CU的一或多个预测单元(PU)。CU的PU中的每一者可与CU的视频块内的不同视频块相关联。视频编码器20可产生CU的每一PU的经预测视频块。PU的经预测视频块可为样本块。视频编码器20可使用帧内预测或帧间预测来产生PU的经预测视频块。

[0068] 当视频编码器20使用帧内预测来产生PU的经预测视频块时,视频编码器20可基于与PU相关联的图像的经解码样本而产生PU的经预测视频块。如果视频编码器20使用帧内预测来产生CU的PU的经预测视频块,那么CU为经帧内预测CU。当视频编码器20使用帧间预测来产生PU的经预测视频块时,视频编码器20可基于不同于与PU相关联的图片的一或多个图片的经解码样本而产生PU的经预测视频块。如果视频编码器20使用帧间预测来产生CU的PU的经预测视频块,则所述CU为经帧间预测CU。

[0069] 此外,当视频编码器20使用帧间预测来产生PU的经预测视频块时,视频编码器20可产生PU的运动信息。PU的运动信息可指示PU的一或多个参考块。PU的每一参考块可为参考图片内的视频块。参考图片可为不同于与PU相关联的图片的图片。在一些情况下,PU的参考块也可被称作PU的“参考样本”。视频编码器20可基于PU的参考块而产生PU的经预测视频块。

[0070] 在视频编码器20产生CU的一或多个PU的经预测视频块之后,视频编码器20可基于CU的PU的经预测视频块而产生CU的残余数据。CU的残余数据可指示CU的PU的经预测视频块与CU的原始视频块中的样本之间的差。

[0071] 此外,作为对未经分割的CU执行编码操作的部分,视频编码器20可对CU的残余数据执行递归四叉树分割以将CU的残余数据分割成与CU的变换单元(TU)相关联的一或多个残余数据块(例如,残余视频块)。CU的每一TU可与不同残余视频块相关联。

[0072] 视频编码器20可将一或多个变换应用于与TU相关联的残余视频块以产生与TU相关联的变换系数块(例如,变换系数块)。在概念上,变换系数块可为变换系数的二维(2D)矩阵。

[0073] 在产生变换系数块之后,视频编码器20可对所述变换系数块执行量化过程。量化一般指对变换系数进行量化以可能地减少用以表示变换系数的数据的量从而提供进一步压缩的过程。量化过程可减小与变换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位变换系数向下舍入到m位变换系数,其中n大于m。

[0074] 视频编码器20可使每一CU与量化参数(QP)值相关联。与CU相关联的QP值可确定视频编码器20如何对与CU相关联的变换系数块进行量化。视频编码器20可通过调整与CU相关联的QP值来调整应用于与CU相关联的变换系数块的量化的程度。

[0075] 在视频编码器20对变换系数块进行量化之后,视频编码器20可产生表示经量化变换系数块中的变换系数的语法元素集。视频编码器20可将例如上下文自适应二进制算术译码(CABAC)操作等熵编码操作应用于这些语法元素中的一些。还可使用例如上下文自适应可变长度译码(CAVLC)、概率区间分割熵(PIPE)译码或其它二进制算术译码等其它熵译码技术。

[0076] 由视频编码器20产生的位流可包含一系列网络抽象层(NAL)单元。NAL单元中的每一者可为含有NAL单元中的数据类型 的指示和含有数据的字节的语法结构。举例来说,NAL

单元可含有表示视频参数集、序列参数集、图片参数集、经译码切片、补充增强信息 (SEI)、存取单元分隔符、填充数据或另一类型的数据的数据。NAL单元中的数据可包含各种语法结构。

[0077] 视频解码器30可接收由视频编码器20产生的位流。位流可包含由视频编码器20编码的视频数据的经译码表示。当视频解码器30接收到位流时,视频解码器30可对位流执行剖析操作。当视频解码器30执行剖析操作时,视频解码器30可从位流提取语法元素。视频解码器30可基于从位流提取的语法元素而重建视频数据的图片。基于语法元素而重建视频数据的过程一般可与由视频编码器20执行以产生语法元素的过程互逆。

[0078] 在视频解码器30提取与CU相关联的语法元素之后,视频解码器30可基于所述语法元素而产生CU的PU的经预测视频块。另外,视频解码器30可对与CU的TU相关联的变换系数块进行反量化。视频解码器30可对变换系数块执行反变换以重建与CU的TU相关联的残余视频块。在产生经预测视频块且重建残余视频块之后,视频解码器30可基于经预测视频块和残余视频块而重建CU的视频块。以此方式,视频解码器30可基于位流中的语法元素而重建CU的视频块。

[0079] 视频编码器

[0080] 图2A为说明可实施根据本发明中所描述的方面的技术的视频编码器20的实例的框图。视频编码器20可经配置以处理视频帧的单个层(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。在一些实例中,本发明中所描述的技术可在视频编码器20的各种组件间共享。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所描述的技术中的任一者或全部。

[0081] 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,本发明的技术可适用于其它译码标准或方法。图2A中所描绘的实例是针对单层编解码器。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处理多层编解码器。

[0082] 视频编码器20可对视频切片内的视频块执行帧内和帧间译码。帧内译码依赖于空间预测来减小或移除给定视频帧或图片内的视频的空间冗余。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。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。例如单向预测(P模式)或双向预测(B模式)等帧间模式可指若干基于时间的译码模式中的任一者。

[0083] 在图2A的实例中,视频编码器20包含多个功能组件。视频编码器20的功能组件包含预测处理单元100、残余产生单元102、变换处理单元104、量化单元106、反量化单元108、反变换单元110、重建单元112、滤波器单元113、经解码图片缓冲器114和熵编码单元116。预测处理单元100包含帧间预测单元121、运动估计单元122、运动补偿单元124、帧内预测单元126和层间预测单元128。在其它实例中,视频编码器20可包含更多、更少或不同功能组件。此外,运动估计单元122和运动补偿单元124可高度集成,但出于解释的目的而在图2A的实例中分开表示。

[0084] 视频编码器20可接收视频数据。视频编码器20可从各种源接收视频数据。举例来说,视频编码器20可从视频源18(例如,图1A或1B中所展示)或另一源接收视频数据。视频数据可表示一系列图片。为了对视频数据进行编码,视频编码器20可对图片中的每一者执行

编码操作。作为对图片执行编码操作的部分,视频编码器20可对图片的每一切片执行编码操作。作为对切片执行编码操作的部分,视频编码器20可对切片中的树块执行编码操作。

[0085] 作为对树块执行编码操作的部分,预测处理单元100可对树块的视频块执行四叉树分割以将所述视频块划分成逐渐变小的视频块。较小视频块中的每一者可与不同CU相关联。举例来说,预测处理单元100可将树块的视频块分割成四个大小相等的子块,将所述子块中的一或多者分割成四个大小相等的子子块,以此类推。

[0086] 与CU相关联的视频块的大小范围可从 8×8 样本直到具有 64×64 样本或更大的最大值的树块的大小。在本发明中,“ $N \times N$ ”和“ N 乘 N ”可互换使用来指在垂直和水平尺寸方面的视频块的样本尺寸,例如, 16×16 样本或16乘16样本。一般来说, 16×16 视频块在垂直方向上具有十六个样本($y=16$),且在水平方向上具有十六个样本($x=16$)。同样, $N \times N$ 块一般在垂直方向上具有 N 个样本,且在水平方向上具有 N 个样本,其中 N 表示非负整数值。

[0087] 此外,作为对树块执行编码操作的部分,预测处理单元100可产生树块的分层四叉树数据结构。举例来说,树块可对应于四叉树数据结构的根节点。如果预测处理单元100将树块的视频块分割成四个子块,那么根节点在四叉树数据结构中具有四个子级节点。所述子级节点中的每一者对应于与子块中的一者相关联的CU。如果预测处理单元100将子块中的一者分割成四个子子块,那么对应于与子块相关联的CU的节点可具有四个子级节点,其中的每一者对应于与子子块中的一者相关联的CU。

[0088] 四叉树数据结构的每一节点可含有对应树块或CU的语法数据(例如,语法元素)。举例来说,四叉树中的节点可包含分裂旗标,其指示对应于节点的CU的视频块是否被分割(例如,分裂)成四个子块。CU的语法元素可被递归地定义,且可取决于CU的视频块是否分裂成子块。视频块未被分割的CU可对应于四叉树数据结构中的叶节点。经译码树块可包含基于对应树块的四叉树数据结构的数据。

[0089] 视频编码器20可对树块的每一未经分割的CU执行编码操作。当视频编码器20对未经分割的CU执行编码操作时,视频编码器20产生表示未经分割的CU的经编码表示的数据。

[0090] 作为对CU执行编码操作的部分,预测处理单元100可在CU的一或多个PU间分割CU的视频块。视频编码器20和视频解码器30可支持各种PU大小。假定特定CU的大小为 $2N \times 2N$,那么视频编码器20和视频解码器30可支持 $2N \times 2N$ 或 $N \times N$ 的PU大小,以及 $2N \times 2N$ 、 $2N \times N$ 、 $N \times 2N$ 、 $N \times N$ 、 $2N \times nU$ 、 $nL \times 2N$ 、 $nR \times 2N$ 或类似的对称PU大小的帧间预测。视频编码器20和视频解码器30还可支持 $2N \times nU$ 、 $2N \times nD$ 、 $nL \times 2N$ 和 $nR \times 2N$ 的PU大小的不对称分割。在一些实例中,预测处理单元100可执行几何分割以沿并不按直角与CU的视频块的边会合的边界在CU的PU间分割CU的视频块。

[0091] 帧间预测单元121可对CU的每一PU执行帧间预测。帧间预测可提供时间压缩。为了对PU执行帧间预测,运动估计单元122可产生PU的运动信息。运动补偿单元124可基于运动信息和除与CU相关联的图片(例如,参考图片)以外的图片的经解码样本而产生PU的经预测视频块。在本发明中,由运动补偿单元124产生的经预测视频块可被称作经帧间预测视频块。

[0092] 切片可为I切片、P切片,或B切片。运动估计单元122和运动补偿单元124可取决于PU是在I切片、P切片还是B切片中而对CU的PU执行不同操作。在I切片中,所有PU都经帧内预测。因此,如果PU在I切片中,那么运动估计单元122和运动补偿单元124不对PU执行帧间预

测。

[0093] 如果PU在P切片中,那么含有PU的图片与被称作“列表0”的参考图片列表相关联。列表0中的参考图片中的每一者含有可用于对其它图片进行帧间预测的样本。当运动估计单元122关于P切片中的PU执行运动估计操作时,运动估计单元122可搜索列表0中的参考图片以找出PU的参考块。PU的参考块可为最紧密对应于PU的视频块中的样本的一组样本,例如,样本块。运动估计单元122可使用多种度量来确定参考图片中的一组样本对应于PU的视频块中的样本的紧密程度。举例来说,运动估计单元122可通过绝对差总和(SAD)、平方差总和(SSD)或其它差异度量来确定参考图片中的一组样本对应于PU的视频块中的样本的紧密程度。

[0094] 在识别出P切片中的PU的参考块之后,运动估计单元122可产生指示列表0中的含有参考块的参考图片的参考索引和指示PU与参考块之间的空间位移的运动向量。在各种实例中,运动估计单元122可以变化的精确度产生运动向量。举例来说,运动估计单元122可以四分之一样本精确度、八分之一样本精确度或其它分数样本精确度产生运动向量。在分数样本精确度的情况下,参考块值可从参考图片中的整数位置样本值内插。运动估计单元122可输出参考索引和运动向量作为PU的运动信息。运动补偿单元124可基于由PU的运动信息识别的参考块而产生PU的经预测视频块。

[0095] 如果PU在B切片中,那么含有PU的图片可与被称作“列表0”和“列表1”的两个参考图片列表相关联。在一些实例中,含有B切片的图片可与为列表0与列表1的组组合的列表组合相关联。

[0096] 此外,如果PU在B切片中,那么运动估计单元122可对PU执行单向预测或双向预测。当运动估计单元122对PU执行单向预测时,运动估计单元122可搜索列表0或列表1的参考图片以找出PU的参考块。运动估计单元122可随后产生指示列表0或列表1中的含有参考块的参考图片的参考索引和指示PU与参考块之间的空间位移的运动向量。运动估计单元122可输出参考索引、预测方向指示符和运动向量作为PU的运动信息。预测方向指示符可指示参考索引指示列表0还是列表1中的参考图片。运动补偿单元124可基于由PU的运动信息指示的参考块而产生PU的经预测视频块。

[0097] 当运动估计单元122对PU执行双向预测时,运动估计单元122可搜索列表0中的参考图片以找到PU的参考块,且还可搜索列表1中的参考图片以找到PU的另一参考块。运动估计单元122可随后产生指示列表0和列表1中的含有参考块的参考图片的参考索引和指示参考块与PU之间的空间位移的运动向量。运动估计单元122可输出PU的参考索引和运动向量作为PU的运动信息。运动补偿单元124可基于由PU的运动信息指示的参考块而产生PU的经预测视频块。

[0098] 在一些情况下,运动估计单元122不将PU的运动信息的完整集合输出到熵编码单元116。实际上,运动估计单元122可参考另一PU的运动信息而用信号表示PU的运动信息。举例来说,运动估计单元122可确定PU的运动信息足够类似于相邻PU的运动信息。在此实例中,运动估计单元122可在与PU相关联的语法结构中指示一值,所述值向视频解码器30指示PU具有与相邻PU相同的运动信息。在另一实例中,运动估计单元122可在与PU相关联的语法结构中识别相邻PU和运动向量差(MVD)。运动向量差指示PU的运动向量与所指示的相邻PU的运动向量之间的差。视频解码器30可使用所指示的相邻PU的运动向量和运动向量差来确

定PU的运动向量。通过在用信号表示第二PU的运动信息时参考第一PU的运动信息,视频编码器20可能能够使用较少位用信号表示第二PU的运动信息。

[0099] 作为对CU执行编码操作的部分,帧内预测单元126可对CU的PU执行帧内预测。帧内预测可提供空间压缩。当帧内预测单元126对PU执行帧内预测时,帧内预测单元126可基于同一图片中的其它PU的经解码样本而产生PU的预测数据。PU的预测数据可包含经预测视频块和各种语法元素。帧内预测单元126可对I切片、P切片和B切片中的PU执行帧内预测。

[0100] 为了对PU执行帧内预测,帧内预测单元126可使用多个帧内预测模式来产生PU的预测数据的多个集合。当帧内预测单元126使用帧内预测模式来产生PU的预测数据的集合时,帧内预测单元126可在与帧内预测模式相关联的方向和/或梯度上跨越PU的视频块从相邻PU的视频块扩展样本。假定对于PU、CU和树块采用从左到右、从上到下的编码次序,相邻PU可在PU的上方、右上方、左上方或左方。帧内预测单元126可取决于PU的大小而使用各种数目个帧内预测模式,例如33个定向帧内预测模式。

[0101] 预测处理单元100可从由运动补偿单元124针对PU产生的预测数据或由帧内预测单元126针对PU产生的预测数据中选择PU的预测数据。在一些实例中,预测处理单元100基于预测数据集合的速率/失真度量而选择PU的预测数据。

[0102] 如果预测处理单元100选择由帧内预测单元126产生的预测数据,那么预测处理单元100可用信号表示用以产生PU的预测数据的帧内预测模式,例如,所选帧内预测模式。预测处理单元100可以各种方式用信号表示所选帧内预测模式。举例来说,所选帧内预测模式有可能与相邻PU的帧内预测模式相同。换句话说,相邻PU的帧内预测模式可为当前PU的最可能模式。因此,预测处理单元100可产生用以指示所选帧内预测模式与相邻PU的帧内预测模式相同的语法元素。

[0103] 如上文所论述,视频编码器20可包含层间预测单元128。层间预测单元128经配置以使用SVC中可用的一或多个不同层(例如,基础层或参考层)来预测当前块(例如,EL中的当前块)。此预测可被称作层间预测。层间预测单元128利用预测方法来减少层间冗余,进而改进译码效率且减少计算资源要求。层间预测的一些实例包含层间帧内预测、层间运动预测和层间残余预测。层间帧内预测使用基础层中的协同定位块的重建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。层间残余预测使用基础层的残余来预测增强层的残余。

[0104] 在预测处理单元100选择CU的PU的预测数据之后,残余产生单元102可通过从CU的视频块减去(例如,由减号指示)CU的PU的经预测视频块而产生CU的残余数据。CU的残余数据可包含对应于CU的视频块中的样本的不同样本分量的2D残余视频块。举例来说,残余数据可包含对应于CU的PU的经预测视频块中的样本的亮度分量与CU的原始视频块中的样本的亮度分量之间的差的残余视频块。另外,CU的残余数据可包含对应于CU的PU的经预测视频块中的样本的色度分量与CU的原始视频块中的样本的色度分量之间的差的残余视频块。

[0105] 预测处理单元100可执行四叉树分割以将CU的残余视频块分割成子块。每一未经划分的残余视频块可与CU的不同TU相关联。与CU的TU相关联的残余视频块的大小和位置可或可不基于与CU的PU相关联的视频块的大小和位置。被称为“残余四叉树”(RQT)的四叉树结构可包含与残余视频块中的每一者相关联的节点。CU的TU可对应于RQT的叶节点。

[0106] 变换处理单元104可通过将一或多个变换应用于与CU的每一TU相关联的残余视频

块而产生用于所述TU的一或多个变换系数块。变换系数块中的每一者可为变换系数的2D矩阵。变换处理单元104可将各种变换应用于与TU相关联的残余视频块。举例来说,变换处理单元104可将离散余弦变换(DCT)、定向变换或概念上类似的变换应用于与TU相关联的残余视频块。

[0107] 在变换处理单元104产生与TU相关联的变换系数块之后,量化单元106可对所述变换系数块中的变换系数进行量化。量化单元106可基于与CU相关联的QP值而对与CU的TU相关联的变换系数块进行量化。

[0108] 视频编码器20可以各种方式使QP值与CU相关联。举例来说,视频编码器20可对与CU相关联的树块执行速率-失真分析。在速率-失真分析中,视频编码器20可通过对树块执行多次编码操作而产生树块的多个经译码表示。当视频编码器20产生树块的不同经编码表示时,视频编码器20可使不同QP值与CU相关联。当给定QP值与具有最低位速率和失真度量的树块的经译码表示中的CU相关联时,视频编码器20可用信号表示给定QP值与CU相关联。

[0109] 反量化单元108和反变换单元110可分别将反量化和反变换应用于变换系数块,以从变换系数块重建残余视频块。重建单元112可将经重建残余视频块添加到来自由预测处理单元100产生的一或多个经预测视频块的对应样本,以产生与TU相关联的经重建视频块。通过以此方式重建CU的每一TU的视频块,视频编码器20可重建CU的视频块。

[0110] 在重建单元112重建CU的视频块之后,滤波器单元113可执行解块操作以减少与CU相关联的视频块中的块假象。在执行一或多个解块操作之后,滤波器单元113可将CU的经重建视频块存储在经解码图片缓冲器114中。运动估计单元122和运动补偿单元124可使用含有经重建视频块的参考图片对后续图片的PU执行帧间预测。另外,帧内预测单元126可使用经解码图片缓冲器114中的经重建视频块对在CU相同的图片中的其它PU执行帧内预测。

[0111] 熵编码单元116可从视频编码器20的其它功能组件接收数据。举例来说,熵编码单元116可从量化单元106接收变换系数块,且可从预测处理单元100接收语法元素。当熵编码单元116接收到数据时,熵编码单元116可执行一或多个熵编码操作以产生经熵编码数据。举例来说,视频编码器20可对数据执行CAVLC操作、CABAC操作、可变到可变(V2V)长度译码操作、基于语法的上下文自适应二进制算术译码(SBAC)操作、概率区间分割熵(PIPE)译码操作或另一类型的熵编码操作。熵编码单元116可输出包含经熵编码数据的位流。

[0112] 作为对数据执行熵编码操作的一部分,熵编码单元116可选择上下文模型。如果熵编码单元116正执行CABAC操作,那么上下文模型可指示特定二进制数具有特定值的概率的估计。在CABAC的上下文中,术语“二进制数”用以指语法元素的二进制化版本的位。

[0113] 多层视频编码器

[0114] 图2B为说明可实施根据本发明中所描述的方面的技术的多层视频编码器23(也被简称为视频编码器23)的实例的框图。视频编码器23可经配置以处理多层视频帧(例如,针对SHVC和MV-HEVC)。此外,视频编码器23可经配置以执行本发明的技术中的任一者或全部。

[0115] 视频编码器23包含视频编码器20A和视频编码器20B,其中的每一者可被配置为视频编码器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可包含与视频编码器20相同的系统和子系统至少一些。尽管将视频编码器23说明为包含两个视频编码器20A和20B,但视频编码器23不受此限制且可包含任何数目个视频编码器20层。在一些实施例中,视频编码器23可包含针对存取单元中的

每一图片或帧的视频编码器20。举例来说,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23可包含比存取单元中的帧更多的编码器层。在一些此类情况下,当处理一些存取单元时,视频编码器层中的一些可能不起作用。

[0116] 除了视频编码器20A及20B之外,视频编码器23还可包含再取样本单元90。在一些情况下,再取样本单元90可对所接收视频帧的基础层进行升取样以(例如)创建增强层。再取样本单元90可对与帧的所接收基础层相关联的特定信息进行升取样,但不对其余信息进行升取样。举例来说,再取样本单元90可对基础层的空间大小或像素数目进行升取样,但切片的数目或图片次序计数可保持恒定。在一些情况下,再取样本单元90可不处理所接收视频且/或可为任选的。举例来说,在一些情况下,预测处理单元100可执行升取样。在一些实施例中,再取样本单元90经配置以对层进行升取样,且重新组织、重新定义、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。尽管主要描述为对基础层或存取单元中的较低层进行升取样,但在一些情况下,再取样本单元90可对层进行降取样。举例来说,如果在视频的流式传输期间减小带宽,那么可对帧进行降取样而不是升取样。

[0117] 再取样本单元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)并对所述图片(或所接收的图片信息)进行升取样。可随后将此经升取样图片提供到较高层编码器(例如,视频编码器20B)的预测处理单元100,所述较高层编码器经配置以与较低层编码器相同的存取单元中的图片进行编码。在一些情况下,较高层编码器与较低层编码器相隔一个层。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。

[0118] 在一些情况下,可省略或绕过再取样本单元90。在这些情况下,可将来自视频编码器20A的经解码图片缓冲器114的图片直接(或至少不提供到再取样本单元90)提供到视频编码器20B的预测处理单元100。举例来说,如果提供到视频编码器20B的视频数据和来自视频编码器20A的经解码图片缓冲器114的参考图片具有相同大小或分辨率,那么可将参考图片提供到视频编码器20B而无任何再取样。

[0119] 在一些实施例中,视频编码器23使用降取样本单元94对待提供到较低层编码器的视频数据进行降取样,之后将所述视频数据提供到视频编码器20A。替代地,降取样本单元94可为能够对视频数据升取样或降取样的再取样本单元90。在另外其它实施例中,可省略降取样本单元94。

[0120] 如图2B中所说明,视频编码器23可进一步包含多路复用器(或mux) 98。mux 98可从视频编码器23输出经组合的位流。可通过从视频编码器20A和20B中的每一者取得位流且使在给定时间输出哪一位流交替来产生经组合的位流。虽然在一些情况下,可一次一个位地交替来自两个(或在多于两个视频编码器层的情况下更多)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所选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,可针对从视频编码器20A输出的每一块从视频编码器20B输出两个块。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23外部的系统(例如,从包含源装置12的源装置上的处理器)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、基于链路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23的所要分辨率输出的任何其它因素而产生控制信号。

[0121] 视频解码器

[0122] 图3A为说明可实施根据本发明中所描述的方面的技术的视频解码器30的实例的框图。视频解码器30可经配置以处理视频帧的单个层(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。在一些实例中,本发明中所描述的技术可在视频解码器30的各种组件间共享。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所描述的技术中的任一者或全部。

[0123] 出于解释的目的,本发明描述在HEVC译码的上下文中的视频解码器30。然而,本发明的技术可适用于其它译码标准或方法。图3A中所描绘的实例是针对单层编解码器。然而,如将关于图3B进一步描述,可复制视频解码器30中的一些或全部以用于处理多层编解码器。

[0124] 在图3A的实例中,视频解码器30包含多个功能组件。视频解码器30的功能组件包含熵解码单元150、预测处理单元152、反量化单元154、反变换单元156、重建单元158、滤波器单元159和经解码图片缓冲器160。预测处理单元152包含运动补偿单元162、帧内预测单元164和层间预测单元166。在一些实例中,视频解码器30可执行与关于图2A的视频编码器20描述的编码遍次大体上互逆的解码遍次。在其它实例中,视频解码器30可包含较多、较少或不同的功能组件。

[0125] 视频解码器30可接收包括经编码视频数据的位流。位流可包含多个语法元素。当视频解码器30接收到位流时,熵解码单元150可对位流执行剖析操作。作为对位流执行剖析操作的结果,熵解码单元150可从位流提取语法元素。作为执行剖析操作的部分,熵解码单元150可对位流中的经熵编码语法元素进行熵解码。预测处理单元152、反量化单元154、反变换单元156、重建单元158和滤波器单元159可执行重建操作,重建操作基于从位流提取的语法元素而产生经解码视频数据。

[0126] 如上文所论述,位流可包括一系列NAL单元。位流的NAL单元可包含视频参数集NAL单元、序列参数集NAL单元、图片参数集NAL单元、SEI NAL单元等等。作为对位流执行剖析操作的一部分,熵解码单元150可执行剖析操作,所述剖析操作从序列参数集NAL单元提取序列参数集且对所述序列参数集进行熵解码、从图片参数集NAL单元提取图片参数集且对所述图片参数集进行熵解码、从SEI NAL单元提取SEI数据且对所述SEI数据进行熵解码,等等。

[0127] 另外,位流的NAL单元可包含经译码切片NAL单元。作为对位流执行剖析操作的部分,熵解码单元150可执行剖析操作,所述剖析操作从经译码切片NAL单元提取经译码切片且对经译码切片NAL单元进行熵解码。经译码切片中的每一者可包含切片标头和切片数据。切片标头可含有关于切片的语法元素。切片标头中的语法元素可包含识别与含有所述切片的图片相关联的图片参数集的语法元素。熵解码单元150可对经译码切片标头中的语法元素执行熵解码操作(例如,CABAC解码操作),以恢复切片标头。

[0128] 作为从经译码切片NAL单元提取切片数据的部分,熵解码单元150可执行剖析操作,所述剖析操作从切片数据中的经译码CU提取语法元素。所提取的语法元素可包含与变换系数块相关联的语法元素。熵解码单元150可随后对语法元素中的一些执行CABAC解码操

作。

[0129] 在熵解码单元150对未经分割的CU执行剖析操作之后,视频解码器30可对未经分割的CU执行重建操作。为了对未经分割的CU执行重建操作,视频解码器30可对CU的每一TU执行重建操作。通过对CU的每一TU执行重建操作,视频解码器30可重建与CU相关联的残余视频块。

[0130] 作为对TU执行重建操作的一部分,反量化单元154可对与TU相关联的变换系数块进行反量化(例如,解量化)。反量化单元154可以类似于针对HEVC所提议或由H.264解码标准定义的反量化过程的方式对变换系数块进行反量化。反量化单元154可使用由视频编码器20针对变换系数块的CU计算的量化参数QP来确定量化程度,且同样地,确定反量化单元154应用的反量化的程度。

[0131] 在反量化单元154对变换系数块进行反量化之后,反变换单元156可产生与变换系数块相关联的TU的残余视频块。反变换单元156可将反变换应用于变换系数块,以便产生TU的残余视频块。举例来说,反变换单元156可将反DCT、反整数变换、反卡忽南-拉维(Karhunen-Loeve)变换(KLT)、反旋转变换、反定向变换或另一反变换应用于变换系数块。在一些实例中,反变换单元156可基于来自视频编码器20的信令而确定待应用于变换系数块的反变换。在这些实例中,反变换单元156可基于在与变换系数块相关联的树块的四叉树的根节点处的用信号表示的变换而确定反变换。在其它实例中,反变换单元156可从例如块大小、译码模式等一或多个译码特性推断反变换。在一些实例中,反变换单元156可应用级联的反变换。

[0132] 在一些实例中,运动补偿单元162可通过基于内插滤波器执行内插而改良PU的经预测视频块。待用于具有子样本精确度的运动补偿的内插滤波器的标识符可包含在语法元素中。运动补偿单元162可使用由视频编码器20在产生PU的经预测视频块期间使用的相同内插滤波器来计算参考块的次整数样本的内插值。运动补偿单元162可根据接收到的语法信息确定由视频编码器20使用的内插滤波器且使用所述内插滤波器来产生经预测视频块。

[0133] 如果是使用帧内预测对PU进行编码,那么帧内预测单元164可执行帧内预测以产生PU的经预测视频块。举例来说,帧内预测单元164可基于位流中的语法元素而确定PU的帧内预测模式。位流可包含帧内预测单元164可用以确定PU的帧内预测模式的语法元素。

[0134] 在一些情况下,语法元素可指示帧内预测单元164将使用另一PU的帧内预测模式来确定当前PU的帧内预测模式。举例来说,当前PU的帧内预测模式有可能与相邻PU的帧内预测模式相同。换句话说,相邻PU的帧内预测模式可为当前PU的最可能模式。因此,在此实例中,位流可包含小语法元素,所述小语法元素指示PU的帧内预测模式与相邻PU的帧内预测模式相同。帧内预测单元164可接着使用帧内预测模式基于空间相邻PU的视频块而产生PU的预测数据(例如,经预测样本)。

[0135] 如上文所论述,视频解码器30还可包含层间预测单元166。层间预测单元166经配置以使用在SHVC中可用的一或多个不同层(例如,基础层或参考层)来预测当前块(例如,增强层中的当前块)。此预测可被称作层间预测。层间预测单元166利用预测方法来减少层间冗余,进而改进译码效率且减少计算资源要求。层间预测的一些实例包含层间帧内预测、层间运动预测和层间残余预测。层间帧内预测使用基础层中的协同定位块的重建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。层间残余预

测使用基础层的残余来预测增强层的残余。下文更详细地论述层间预测方案中的每一者。

[0136] 重建单元158可使用与CU的TU相关联的残余视频块和CU的PU的经预测视频块(例如,帧内预测数据或帧间预测数据,如果适用)来重建CU的视频块。因此,视频解码器30可基于位流中的语法元素而产生经预测视频块和残余视频块,且可基于经预测视频块和残余视频块而产生视频块。

[0137] 在重建单元158重建CU的视频块之后,滤波器单元159可执行解块操作以减少与CU相关联的块假象。在滤波器单元159执行解块操作以减少与CU相关联的块假象之后,视频解码器30可将CU的视频块存储在经解码图片缓冲器160中。经解码图片缓冲器160可提供参考图片以用于后续运动补偿、帧内预测和在例如图1A或1B的显示装置32等显示装置上呈现。举例来说,视频解码器30可基于经解码图片缓冲器160中的视频块而对其它CU的PU执行帧内预测或帧间预测操作。

[0138] 多层解码器

[0139] 图3B为说明可实施根据本发明中所描述的方面的技术的多层视频解码器33(也被简称为视频解码器33)的实例的框图。视频解码器33可经配置以处理多层视频帧(例如,针对SHVC和多视图译码)。此外,视频解码器33可经配置以执行本发明的技术中的任一者或全部。

[0140] 视频解码器33包含视频解码器30A和视频解码器30B,其中的每一者可被配置为视频解码器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可包含与视频解码器30相同的系统和子系统至少一些。尽管将视频解码器33说明为包含两个视频解码器30A和30B,但视频解码器33不受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33可包含针对存取单元中的每一图片或帧的视频解码器30。举例来说,包含五个图片的存取单元可由包含五个解码器层的视频解码器处理或解码。在一些实施例中,视频解码器33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起作用。

[0141] 除了视频解码器30A和30B之外,视频解码器33还可包含升取样单元92。在一些实施例中,升取样单元92可对所接收视频帧的基础层进行升取样以创建待添加到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经解码图片缓冲器160中。在一些实施例中,升取样单元92可包含关于图2A的再取样单元90描述的实施例中的某些或全部。在一些实施例中,升取样单元92经配置以对层进行升取样且重新组织、重新定义、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,升取样单元92可为经配置以对所接收视频帧的层进行升取样和/或降取样的再取样单元。

[0142] 升取样单元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160接收图片或帧(或与图片相关联的图片信息)并对所述图片(或所接收的图片信息)进行升取样。可随后将此经升取样图片提供到较高层解码器(例如,视频解码器30B)的预测处理单元152,所述较高层解码器经配置以对与较低层解码器相同的存取单元中的图片进行解码。在一些情况下,较高层解码器与较低层解码器相隔一个层。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。

[0143] 在一些情况下,可省略或绕过升取样单元92。在这些情况下,可将来自视频解码器

30A的经解码图片缓冲器160的图片直接(或至少不提供到升取样单元92)提供到视频解码器30B的预测处理单元152。举例来说,如果提供到视频解码器30B的视频数据和来自视频解码器30A的经解码图片缓冲器160的参考图片具有相同大小或分辨率,那么可将参考图片提供到视频解码器30B而不进行升取样。此外,在一些实施例中,升取样单元92可为经配置以对从视频解码器30A的经解码图片缓冲器160接收的参考图片进行升取样或降取样的再取样单元90。

[0144] 如图3B中所说明,视频解码器33可进一步包含多路分用器(或demux)99。demux 99可将经编码视频位流分裂成多个位流,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,图3B的实例中的视频解码器30A和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来划分所述位流。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,可针对被提供到视频解码器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33外部的系统(例如,从包含目的地装置14的目的地装置上的处理器)接收的控制信号而划分位流。可基于来自输入接口28的视频的分辨率或位速率、基于链路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33获得的分辨率的任何其它因素而产生控制信号。

[0145] 帧内随机存取点(IRAP)图片

[0146] 一些视频译码方案可提供贯穿位流的各种随机存取点,使得可从那些随机存取点中的任一者开始解码位流而无需解码位流中的在那些随机存取点之前的任何图片。在这些视频译码方案中,可正确地解码除随机存取跳过前导(RASL)图片以外的按解码次序在随机存取点之后的所有图片而不使用在随机存取点之前的任何图片。举例来说,即使位流的一部分在传输期间或在解码期间丢失,解码器仍可从下一个随机存取点开始恢复解码位流。对随机存取的支持可促进(例如)动态流式传输服务、搜寻操作、信道切换等。

[0147] 在一些译码方案中,这些随机存取点可由被称作帧内随机存取点(IRAP)图片的图片提供。举例来说,与存取单元(“auA”)中所含有的增强层(“layerA”)中的增强层IRAP图片相关联的随机存取点可提供层特定的随机存取,使得对于layerA的具有与存取单元(“auB”)中所含有的图片(其在层B中且按解码次序在auA之前)相关联的随机存取点(或auA中所含有的随机存取点)的每一参考层(“layerB”) (例如,参考层为用于预测layerA的层),可正确地解码layerA中按解码次序在auA之后的图片(包含位于auA中的那些图片)而无需解码layerA中在auA之前的任何图片。

[0148] IRAP图片可使用帧内预测来译码(例如,在不参考其它图片的情况下译码)和/或可使用层间预测来译码,且可包含(例如)即时解码器刷新(IDR)图片、清洁随机存取(CRA)图片和断链存取(BLA)图片。当位流中存在IDR图片时,按解码次序在IDR图片之前的所有图片不由在IDR图片之后的图片用于预测。当在位流中存在CRA图片时,在CRA图片之后的图片可或可不将按解码次序在CRA图片之前的图片用于预测。按解码次序在CRA图片之后但使用按解码次序在CRA图片之前的图片的那些图片可被称作RASL图片。按解码次序可在IRAP图

片之后且按输出次序在IRAP图片之前的另一类型的图片为随机存取可解码前导 (RADL) 图片,其可不含有对按解码次序在IRAP图片之前的任何图片的参考。如果在CRA图片之前的图片不可用,那么RASL图片可由解码器丢弃。BLA图片对解码器指示在BLA图片之前的图片可能不可用于解码器(例如,因为两个位流被拼接在一起且BLA图片为第二位流的按解码次序的第一图片)。含有作为IRAP图片的基础层图片(例如,层ID值为0)的存取单元(例如,由跨越多个层的与相同输出时间相关联的所有经译码图片组成的图片群组)可被称作IRAP存取单元。

[0149] 位流和编解码器的级别

[0150] 根据本发明的一或多个方面,可缩放视频经编码位流(也被简称为位流)具有参考层或基础层(BL)和一或多个增强层(EL)。在一些情况下,位流可为多视图视频经编码位流,其中所述层中的每一者可构成不同视图。多视图位流的一个实例为包含左眼视图层和右眼视图层的三维(3D)视频位流。

[0151] 位流可包含多个存取单元(AU)。每一存取单元可包含来自基础层的图片和来自每一增强层的图片。

[0152] 位流可由级别定义。级别定义在一些情况下可随层(包含基础层和增强层EL)的数目和/或层的分辨率的总和而变化。位流可具有标识符列表,标识符列表可识别位流的一或多个层。通过定义位流的级别,视频译码系统可预测解码器是否能够解码位流。在本发明的一些实施方案中,多层位流的级别还可由其它因素定义,且级别可与尺度相关联。

[0153] “配置文件”为整个位流语法的子集。在由给定配置文件的语法强加的界限内,仍可能需要取决于位流中的语法元素所取的值(例如,经解码图片的指定大小)而极大地改变编码器和解码器的性能。在许多应用中,实施能够处理特定配置文件内的语法的所有假设用途(包含语法元素的所有可能的值)的解码器既不实用也不经济。

[0154] 为了处理此问题,在每一配置文件内指定“层级”和“级别”。层级的级别为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,其可呈对值的算术组合的约束形式(例如,图片宽度乘以图片高度乘以每秒经解码的图片数目)。针对更低层级指定的级别比针对更高层级指定的级别受到更多约束。

[0155] 配置文件

[0156] 本文中描述两个实例配置文件:可缩放主(Scalable Main)和可缩放主10。可缩放主10被如此命名是因为每一层中的图片的每一分量限制为8位到10位,而对可缩放主的相同限制仅为8位。在本发明的一些实施方案中,如下文所描述而定义可缩放主和可缩放主10配置文件。

[0157]

H.11.1.2 可缩放主和可缩放主 10 配置文件

含有符合可缩放主或可缩放主 10 配置文件的输出层集合的位流应遵从对输出层集合的经导出子位流的以下约束，其中 `layerSetIdx` 为分别符合可缩放主或可缩放主 10 配置文件的输出层集合的层集合，所述层集合通过调用如子条款 F.10 中所指定的子位流提取过程而导出，其中 `tIdTarget` 等于 7，且其中 `TargetDecLayerIdList` 含有具有索引 `layerSetIdx` 的层集合的 `nuh_layer_id` 值。

含有符合可缩放主或可缩放主 10 配置文件的输出层集合的位流还应遵从对通过调用如子条款 F.10 中所指定的子位流提取过程而导出的基础层位流的以下约束，其中 `tIdTarget` 等于 7 且其中 `TargetDecLayerIdList` 仅含有等于 0 的一个 `nuh_layer_id` 值作为输入。

从符合可缩放主配置文件的位流导出的基础层位流应遵从以下约束：

- 基础层位流应遵从子条款 A.3.2 中所指定的主配置文件的所有约束。
- 基础层位流的 SPS 应具有等于 1 的 `general_profile_idc` 或等于 1 的 `general_profile_compatibility_flag[1]`。

从符合可缩放主 10 配置文件的位流导出的基础层位流应遵从以下约束：

- 基础层位流应遵从子条款 A.3.3 中所指定的主 10 配置文件的所有约束。
- 基础层位流的 SPS 应具有等于 1 或 2 的 `general_profile_idc`，或等于 1 的 `general_profile_compatibility_flag[1]` 或 `general_profile_compatibility_flag[2]`。

符合可缩放主或可缩放主 10 配置文件的输出层集合的经导出子位流应遵从以下约束：

- 子位流的所有有效 SPS 应仅具有等于 1 的 `chroma_format_idc`。
- 从子位流的任何有效 SPS 导出的 `CtbLog2SizeY` 应在 4 到 6 的范围(包含端点)内。
- 对于 `nuh_layer_id` 包含于具有索引 `layerSetIdx` 的层集合中的每一经解码图片，变量 `NumResampling`、`NumSampleResampling` 和 `NumMotionResampling` 应小于或等于 1。
- 对于不等于 2 的任何 `smIdx` 值且对于在 `layer_id_included_flag[layerSetIdx][j]` 等于 1 的情况下的任何 `j` 值，`ScalabilityId[j][smIdx]` 应等于 0。
- 对于层 `id iNuhLid` 等于包含于具有索引 `layerSetIdx` 的层集合中的 `nuh_layer_id` 中的任一者的层，`NumRefLayers[iNuhLid]` 的值(其指定直接和间接相关层的总数目且如 F.7.4.3.1 中被导出)应小于或等于 7。
- 所有有效 SPS 应仅在 `i` 等于 0 和在 2 到 6 的范围(包含端点)内的情况下具有等于 0 的 `sps_extension_type_flag[i]`。
- 所有有效 PPS 应仅在 `i` 在 0 到 6 的范围(包含端点)内的情况下具有等于 0 的 `pps_extension_type_flag[i]`。

符合可缩放主配置文件的输出层集合的经导出子位流应遵从以下约束：

- 子位流的所有有效 SPS 应仅具有等于 0 的 `bit_depth_luma_minus8`。
- 子位流的所有有效 SPS 应仅具有等于 0 的 `bit_depth_chroma_minus8`。

符合可缩放主 10 配置文件的输出层集合的经导出子位流应遵从以下约束：

- 子位流的所有有效 SPS 应具有在 0 到 2 的范围(包含端点)内的 `bit_depth_luma_minus8`。
- 子位流的所有有效 SPS 应具有在 0 到 2 的范围(包含端点)内的 `bit_depth_chroma_minus8`。

在此子条款的其余部分和子条款 0 中，对于输出层集合的子位流，`general_profile_idc`、`general_profile_compatibility_flag[i]`、`general_tier_flag` 和 `general_level_idc` 分别指输出层集合的 `profile_tier_level()` 语法结构中的 `general_profile_idc`、`general_profile_compatibility_flag[i]`、`general_tier_flag` 和 `general_level_idc`。

由等于 3 的 `general_profile_idc` 或等于 1 的 `general_profile_compatibility_flag[3]` 来指示输出层集合的子位流符合可缩放主配置文件。

[0158]

注释 1——当 `general_profile_compatibility_flag[3]` 等于 1 时, 且在 `i` 等于 1、2 的情况下, `general_profile_compatibility_flag[i]` 也应等于 1。
由等于 4 的 `general_profile_idc` 或等于 1 的 `general_profile_compatibility_flag[4]` 来指示输出层集合的子位流符合可缩放主 10 配置文件。

注释 2——当 `general_profile_compatibility_flag[4]` 等于 1 时, `general_profile_compatibility_flag[2]` 也应等于 1。

[0159] 表1——SHVC WD5和MV-HEVC WD 7中的配置文件

[0160] 层级和级别

[0161] 如上文所论述,层级的级别为强加于位流中的语法元素的值的指定约束集合。可针对位流和解码器二者指定级别。解码器的级别定义提供解码器的能力的上限。一旦对解码器定义级别,可预见的是,解码器能够解码级别等于或低于所述解码器级别的任何位流(前提是解码器可解码特定配置文件和层级的位流)。举例来说,预期4.0级解码器能够解码4级、3级、2级和1级的位流。然而,如在本文中描述,现有级别定义在一些情况下可能不准确地定义解码器能力。

[0162] 为了更准确地定义解码器能力,可将信息添加到约束。举例来说,在本发明的一些实施方案中,可如下文所论述而修改约束以引入一或多个变量,包含 `shvcScaleFactor` 和 `maxAuSliceSegs`。在一些实施例中, `shvcScaleFactor` 可允许针对单层位流定义的约束修改成适合于多层位流。举例来说,针对单层位流定义例如 `MaxCPB` 和 `MaxLumaSr` 等参数。然而,通过添加 `shvcScaleFactor`,可使用 `MaxCPB` 和 `MaxLumaSr` 而不定义这些参数值的新集合。在一些实施方案中, `shvcScaleFactor` 被设置成等于 $\text{Ceil}(\text{auSizeInSampleY} / \text{MaxLumaPs})$, 其中 `MaxLumaPs` 在表 A-1 中被指定用于输出层集合的子位流符合的级别。也就是说, `shvcScaleFactor` 提供子位流中的所有层的分辨率的总和(即, `auSizeInSamplesY`) 与所述级别的最大亮度图片大小的近似商。应注意,为了 SNR 可缩放性, `shvcScaleFactor` 的值等于子位流中的层的数目(即,子位流中的所有层具有相同空间分辨率)。当跨越层的空间可缩放性或图片速率可缩放性存在时,可使用比上文所定义更小的 `shvcScaleFactor` 的值。

[0163] 替代地, `shvcScaleFactor` 可被设置成等于 $\text{Ceil}((\text{NumLayersInIdList}[\text{LayerSetIdxForOutputLayerSet}[\text{optLsIdx}]] \times (\text{最高层的PicSizeInSamplesY})) / \text{MaxLumaPs})$ 。或者,换句话说, $\text{AuSizeInSamplesY} = \text{NumLayersInIdList}[\text{LayerSetIdxForOutputLayerSet}[\text{optLsIdx}]] \times (\text{最高层的PicSizeInSamplesY})$ 。以下所提议修改可适用于 `shvcScaleFactor` 的以上定义中的任一者。

[0164] 另外,变量 `maxAuSliceSegs` 可被设置成等于 `shvcScaleFactor` \times `MaxSliceSegmentsPerPicture`, 其中 `MaxSliceSegmentsPerPicture` 在 A-1 中被指定。

[0165] 实例实施例1

[0166] 在本发明的一些实施方案中,可如下文所描述而修改一般层级和级别限制以包含 `shvcScaleFactor`。在 SHVC 和 MV-HEVC 的较早版本(例如,SHVC WD 5 和 MV-HEVC WD 7)的上下文中提供本发明中的实例实施例。由斜体字和加下划线指示对 SHVC 和 MV-HEVC 的较早版本的添加部分,且以删除线指示对较早版本的删除部分。以粗体指示注释。

[0167]

H.11.2.1 一般层级和级别限制

应注意，此子条款中的以下所有文本被如下产生。首先从 P1003 (HEVC 缺陷报告 3)复制和粘贴子条款 A.4.1 的文本，随后接受更改，且随后用更改标记做出更改。

出于比较层级能力的目的，general_tier_flag 等于 0 的层级被视为比 general_tier_flag 等于 1 的层级更低的层级。

出于比较级别能力的目的，在特定级别的 general_level_idc 的值小于其它级别的 general_level_idc 的值时，特定层级的特定级别被视为比同一层级的某其它级别更低的级别。

指定以下部分以表达此子条款、子条款 0 和子条款 0 中的约束：

- 使存取单元 n 为按解码次序第 n 个存取单元，其中第一存取单元为存取单元 0 (即，第 0 个存取单元)。
- 使变量 fR 设置成等于 $1 \div 300$ 。
- 使变量 optLsIdx 为位流中所含有的输出层集合的输出层集合索引，且通过调用如子条款 F.10 中所指定的子位流提取过程而导出输出层集合的子位流，其中位流、tIdTarget 等于 7、且含有输出层集合的 nuh_layer_id 值的 TargetDecLayerIdList 作为输入。
- 使变量 numLayersInOptLayerSet 和 numLayersInSubBitstream 被如下导出：
 - numLayersInOptLayerSet = NumLayersInIdList[LayerSetIdxForOutputLayerSet[optLsIdx]]
 - 如果(由外部装置提供输出层集合中的基础层)
 - numLayersInSubBitstream = numLayersInOptLayerSet - 1
 - 否则
 - numLayersInSubBitstream = numLayersInOptLayerSet
- 使变量 auSizeInSamplesY 被如下导出：
 - auSizeInSamplesY = 0
 - 如果(由外部装置提供输出层集合中的基础层)
 - sIdx = 1
 - 否则
 - sIdx = 0
 - 对于(i = sIdx; i < numLayersInOptLayerSet; i++)
 - auSizeInSamplesY += 输出层集合中的第 i 层的 PicSizeInSamplesY
- 使变量 shvcScaleFactor 设置成等于 $\text{Ceil}(\text{auSizeInSamplesY} / \text{MaxLumaPs})$ ，其中 MaxLumaPs 在表 A-1 中被指定用于输出层集合的子位流符合的级别。**应注意，为了 SNR 可缩放性，shvcScaleFactor 的值等于子位流中的层的数目。替代地，当跨越层的空间可缩放性或图片速率可缩放性存在时，可使用较小的 shvcScaleFactor 的值。**替代地，使变量 shvcScaleFactor 设置成等于 $\text{Ceil}((\text{NumLayersInIdList}[\text{LayerSetIdxForOutputLayerSet}[\text{optLsIdx}]] \times \text{最高层的 PicSizeInSamplesY}) / \text{MaxLumaPs})$ 。或者换句话说， $\text{AuSizeInSamplesY} = \text{NumLayersInIdList}[\text{LayerSetIdxForOutputLayerSet}[\text{optLsIdx}]] \times \text{最高层的}$

[0168]

PicSizeInSamplesY, 其中 MaxLumaPs 在表 A-1 中被指定用于输出层集合的子位流符合的级别。

以下所提议修改可适用于 shvcScaleFactor 的以上定义中的任一者

- 使变量 maxAuSliceSegs 设置成等于 shvcScaleFactor×MaxSliceSegmentsPerPicture, 其中 MaxSliceSegmentsPerPicture 在表 A-1 中被指定。

应注意, 对于一般和配置文件特定的约束二者, 除显而易见为层特定的那些(例如, 亮度图片大小)以外, 其当前全被指定于存取单元级别上。替代地, 这些约束还被指定在图片级别下。

符合指定层级和级别下的配置文件的输出层集合的子位流应遵从针对如附录 C 中所指定的每一位流符合性测试的以下约束:

- a) 对于输出层集合中的每一层, PicSizeInSamplesY 应小于或等于 MaxLumaPs, 其中 MaxLumaPs 在表 A-1 中被指定。
- b) 对于输出层集合中的每一层, pic_width_in_luma_samples 的值应小于或等于 $\text{Sqrt}(\text{MaxLumaPs} \times 8)$ 。
- c) 对于输出层集合中的每一层, pic_height_in_luma_samples 的值应小于或等于 $\text{Sqrt}(\text{MaxLumaPs} \times 8)$ 。
- d) 对于输出层集合中的 nuh_layer_id 等于 currLayerId 的每一层, max_vps_layer_dec_pic_buff_minus1[optLsIdx][LayerIdxInVps[currLayerId]][HighestTid] 的值应小于或等于如由方程式 A-2 导出的 MaxDpbSize。
对于输出层集合的子 DPB 中的子 DPB 索引等于 subDpbIdx 的每一子 DPB, max_vps_dec_pic_buffering_minus1[optLsIdx][subDpbIdx][HighestTid] 的值应小于或等于 numLayersInSubDpb×MaxDpbSize, 其中 numLayersInSubDpb 为输出层集合中被指派子 DPB 的层的数目, 且 MaxDpbSize 由方程式 A-2 导出, 其中 PicSizeInSamplesY 为被指派子 DPB 的任何层的 PicSizeInSamplesY 的值。
- e) 对于 5 级和更高的级别, 输出层集合的子位流中的每一层的 CtbSizeY 的值应等于 32 或 64。**应注意, 通过在此处使用措辞“子位流中”, 限制随后也适用于由外部装置提供基础层的情况, 因为在所述情况下排除考虑基础层。在变量 CtbSizeY 不被指定用于由外部装置提供的基础层时, 此为需要的。**
- f) 输出层集合的子位流中的每一图片的 NumPicTotalCurr 的值应小于或等于 8。
- g) 当解码输出层集合的子位流中的每一经译码图片时, num_tile_columns_minus1 的值应小于 MaxTileCols, 且 num_tile_rows_minus1 应小于 MaxTileRows, 其中 MaxTileCols 和 MaxTileRows 在表 A-1 中被指定。
- h) 对于 VCL HRD 参数, 在 i 的至少一个值在 0 到 cpb_cnt_minus1[HighestTid] 的范围(包含端点)内的情况下, CpbSize[i] 应小于或等于 shvcScaleFactor×1000×MaxCPB, 其中 CpbSize[i] 基于如子条款 C.1 中所指定而选择的参数在子条款 E.3.3 中被指定, 且 MaxCPB 以 1000 位为单位在表 A-1 中被指定。
- i) 对于 NAL HRD 参数, 在 i 的至少一个值在 0 到 cpb_cnt_minus1[HighestTid] 的范围(包含端点)内的情况下, CpbSize[i] 应小于或等于 shvcScaleFactor×1100×MaxCPB, 其中 CpbSize[i] 基于如子条款 C.1 中所指定而选择的参数在子条款 E.3.3 中被指定, 且 MaxCPB 以 1100 位为单位在表 A-1 中被指定。

表 A-1 指定对每一层级的每一级别的限制。

位流符合的层级和级别由语法元素 general_tier_flag 和 general_level_idc 如下指示:

- 根据表 A-1 中所指定的层级约束, 等于 0 的 general_tier_flag 指示符合主层级, 且等于 1 的 general_tier_flag 指示符合高层级。针对低于 4 级的级别(对应于表 A-1 中标记有“-”的项), general_tier_flag 应等于 0。
- general_level_idc 应被设置成等于表 A-1 中所指定的级别数目的 30 倍的值。

[0169] 表2——SHVC WD 5和MV-HEVC WD 7中的一般层级和级别限制

[0170] 配置文件特定层级和级别限制

[0171] 实例实施例2

[0172] 在本发明的一些实施方案中,可如下文所描述而修改对可缩放主和可缩放主10配置文件的配置文件特定层级和级别限制。在SHVC和MV-HEVC的较早版本(例如,SHVC WD 5和MV-HEVC WD 7)的上下文中提供本发明中的实例实施例。由斜体字和加下划线指示对SHVC和MV-HEVC的较早版本的添加部分,且以删除线指示对较早版本的删除部分。以粗体指示注释。

[0173]

H.11.2.2 对可缩放主和可缩放主 10 配置文件的配置文件特定层级和级别限制

符合指定层级和级别下的可缩放主或可缩放主 10 配置文件的输出层集合的子位流应遵从针对如附录 C 中所指定的每一位流符合性测试的以下约束:

- a) 如子条款 C.2.3 中所指定, 存取单元 n (其中 n 大于 0) 自 CPB 的标称移除时间应满足以下约束: $AuNominalRemovalTime[n] - AuCpbRemovalTime[n-1]$ 大于或等于 $Max(auSizeInSamplesY \div (shvcScaleFactor \times MaxLumaSr), fR)$, 其中 $auSizeInSamplesY$ 为存取单元 $n-1$ 的 $auSizeInSamplesY$ 的值, 且 $MaxLumaSr$ 为表 A-2 中所指定的适用于存取单元 $n-1$ 的值。
- b) 如子条款 C.3.3 中所指定, 来自 DPB 的图片的连续输出时间之间的差应满足以下约束: $DpbOutputInterval[n]$ 大于或等于 $Max(auSizeInSamplesY \div (shvcScaleFactor \times MaxLumaSr), fR)$, 其中 $auSizeInSamplesY$ 为存取单元 n 的 $auSizeInSamplesY$ 的值, 且 $MaxLumaSr$ 为在表 A-2 中指定用于存取单元 n 的值, 前提是存取单元 n 具有被输出的图片且并非这些存取单元中的最后一个。
- c) 存取单元 0 的移除时间应满足以下约束: 存取单元 0 中的经译码切片段的数目小于或等于 $Min(Max(numLayersInSubBitstream, maxAuSliceSegs \times MaxLumaSr / MaxLumaPs \times (AuCpbRemovalTime[0] - AuNominalRemovalTime[0]) + MaxSliceSegmentsPerPicture \times auSizeInSamplesY / MaxLumaPs), maxAuSliceSegs)$, 其中 $MaxSliceSegmentsPerPicture$ 、 $MaxLumaPs$ 和 $MaxLumaSr$ 为表 A-1 和表 A-2 中所指定的值。
- d) 存取单元 n 和 $n-1$ (其中 n 大于 0) 的连续 CPB 移除时间之间的差应满足以下约束: 存取单元 n 中的经译码切片段的数目小于或等于 $Min((Max(numLayersInSubBitstream, maxAuSliceSegs \times MaxLumaSr / MaxLumaPs \times (AuCpbRemovalTime[n] - AuCpbRemovalTime[n-1])), maxAuSliceSegs)$, 其中 $numLayersInSubBitstream$ 和 $maxAuSliceSegs$ 为存取单元 n 的值, 且 $MaxLumaPs$ 和 $MaxLumaSr$ 为分别在表 A-1 和表 A-2 中指定的适用于存取单元 n 的值。
- e) 对于 VCL HRD 参数, 在 i 的至少一个值在 0 到 $cpb_cnt_minus1[HighestTid]$ 的范围(包含端点)内的情况下, 位速率 $[i]$ 应小于或等于 $shvcScaleFactor \times CpbBrVclFactor \times MaxBR$, 其中位速率 $[i]$ 基于如子条款 C.1 中所指定而选择的参数在子条款 E.3.3 中被指定, 且 $MaxBR$ 以 $CpbBrVclFactor$ 位/s 为单位在表 A-2 中被指定, 其中 $CpbBrVclFactor$ 在表 A-3 中被指定。
- f) 对于 NAL HRD 参数, 在 i 的至少一个值在 0 到 $cpb_cnt_minus1[HighestTid]$ 的范围(包含端点)内的情况下, 位速率 $[i]$ 应小于或等于 $shvcScaleFactor \times CpbBrNalFactor \times MaxBR$, 其中位速率 $[i]$ 基于如子条款 C.1 中所指定

[0174]

- 而选择的参数在子条款 E.3.3 中被指定, 且 MaxBR 以 CpbBrVclFactor 位/s 为单位在表 A-2 中被指定, 其中 CpbBrNallFactor 在表 A-3 中被指定。
- g) 存取单元 0 的 NumBytesInNalUnit 变量的总和应小于或等于 $\text{FormatCapabilityFactor} \times (\text{Max}(\text{auSizeInSamplesY}, \text{fR} \times \text{shvcScaleFactor} \times \text{MaxLumaSr}) + \text{shvcScaleFactor} \times \text{MaxLumaSr} \times (\text{AuCpbRemovalTime}[0] - \text{AuNominalRemovalTime}[0])) \div \text{MinCr}$, 其中 MaxLumaSr 和 MinCr 为在表 A-2 中指定的值, 且 FormatCapabilityFactor 为在表 A-3 中指定的值。
- h) 存取单元 n (其中 n 大于 0) 的 NumBytesInNalUnit 变量的总和应小于或等于 $\text{FormatCapabilityFactor} \times \text{shvcScaleFactor} \times \text{MaxLumaSr} \times (\text{AuCpbRemovalTime}[n] - \text{AuCpbRemovalTime}[n-1]) \div \text{MinCr}$, 其中 MaxLumaSr 和 MinCr 为在表 A-2 中指定的值, 且 FormatCapabilityFactor 为在表 A-3 中指定的值。
- i) 存取单元 0 的移除时间应满足以下约束: 存取单元 0 中的经译码图片中的平铺块的数目小于或等于 $\text{Min}(\text{Max}(\text{numLayersInSubBitstream}, \text{shvcScaleFactor} \times \text{MaxTileCols} \times \text{MaxTileRows} \times 120 \times (\text{AuCpbRemovalTime}[0] - \text{AuNominalRemovalTime}[0]) + \text{shvcScaleFactor} \times \text{MaxTileCols} \times \text{MaxTileRows}), \text{shvcScaleFactor} \times \text{MaxTileCols} \times \text{MaxTileRows})$, 其中 MaxTileCols 和 MaxTileRows 为在表 A-1 中指定的值。

存取单元 n 和 n-1 (其中 n 大于 0) 的连续 CPB 移除时间之间的差应满足以下约束: 存取单元 n 中的经译码图片中的平铺块的数目小于或等于 $\text{Min}(\text{Max}(\text{numLayersInSubBitstream}, \text{shvcScaleFactor} \times \text{MaxTileCols} \times \text{MaxTileRows} \times 120 \times (\text{AuCpbRemovalTime}[n] - \text{AuCpbRemovalTime}[n-1])) + \text{shvcScaleFactor} \times \text{MaxTileCols} \times \text{MaxTileRows})$, 其中 numLayersInSubBitstream 和 shvcScaleFactor 为存取单元 n 的值, 且 MaxTileCols 和 MaxTileRows 为在表 A-1 中指定的适用于存取单元 n 的值。

[0175] 表3—SHVC WD 5和MV-HEVC WD 7中的配置文件特定限制

[0176] 解码器能力

[0177] 在本发明的一些实施方案中, 指定多个种类的相符解码器。一个种类的解码器 (例如, 被称作I类解码器) 可解码符合更高级别但层数目少于与解码器级别相关联的尺度的一些位流, 而另一种类的解码器 (例如, 被称作II类解码器) 可仅解码符合相同或更低级别的位流 (假设配置文件和层级相同)。可通过重复使用现有单层HEVC解码器的硬件核心使用一个现有单层HEVC解码器或在根本不使用现有单层HEVC解码器的情况下 (即, 从零开始设计) 实施I类解码器。可通过原样重复使用硬件核心而不改变地使用多于一个现有单层HEVC解码器来实施II类解码器; 不可以此方式实施I类解码器。

[0178] 本发明的实施例可引入变量 (例如, 尺度N) 以限制位流或子位流中的一些或所有层的空间分辨率的总和。尺度可为在用于以下约束中时提供对解码器能力的大体上准确描述的任何数目。举例来说, 尺度可等于在2到63的范围内的任何值, 应注意, SHVC/MV-HEVC支持位流中的多达63层。因此, 解码器能力可由包含配置文件、层级、级别和尺度的多种描述词中的一或多者定义。

[0179] 实例实施例3

[0180] 在下文所描述的经修改解码器能力中, 可针对每一种类不同地定义能力。在SHVC和MV-HEVC的较早版本 (例如, SHVC WD 5和MV-HEVC WD 7) 的上下文中提供本发明中的实例实施例。由斜体字和加下划线指示对SHVC和MV-HEVC的较早版本的添加部分, 且以删除线指示对较早版本的删除部分。以粗体指示注释。

[0181]

H.11.3 解码器能力

具有特定尺度 N 的在特定层级的特定级别下的符合可缩放主配置文件的 I 类解码器应能够解码所有以下条件都适用的输出层集合的所有子位流：

- $general_profile_idc$ 等于 1 或 3，或在 i 等于 1 或 3 的情况下， $general_profile_compatibility_flag[i]$ 等于 1。
- $general_level_idc$ 表示低于或等于指定级别的级别。
- $general_tier_flag$ 表示低于或等于指定层级的层级。
- $auSizeInSamplesY$ 小于或等于 $N \times MaxLumaPs$ ，其中 $MaxLumaPs$ 在表 A-1 中被指定用于解码器的指定级别。

替代地，限制被指定为 “ $numlayersInSubBitstream \times MaxBsLumaPs$ 小于或等于 $N \times MaxDeLumaPs$ ，其中 $MaxBsLumaPs$ 为 $MaxLumaPs$ 在表 A-1 中被指定用于子位流的级别，且 $MaxDeLumaPs$ 为在表 A-1 中指定用于解码器的级别的 $MaxLumaPs$ 。”

在会话协商等确定解码器是否能够解码位流(对应于输出层集合)的容易程度上可存在差异。此处使用 $numlayersInSubBitstream$ 仅需要位流的配置文件、层级、级别和 $numlayersInSubBitstream$ ，而使用 $auSizeInSamplesY$ 需要 $auSizeInSamplesY$ 来代替位流的 $numlayersInSubBitstream$ 。无论以何种方式， $numlayersInSubBitstream$ 或 $auSizeInSamplesY$ 的值应在会话协商等中直接可用。按文件格式，值应在样本描述中用信号表示。按实时输送协议(RTP)有效负载格式，应针对此定义媒体类型参数。如果直接地用信号表示 $auSizeInSamplesY$ (例如，在 VPS 扩展或 VPS VUI 中)，那么使用 $auSizeInSamplesY$ 可能更好。

具有特定尺度 N 的在特定层级的特定级别下的符合可缩放主配置文件的 II 类解码器应能够解码所有以下条件都适用的输出层集合的所有子位流：

- $general_profile_idc$ 等于 1 或 3，或在 i 等于 1 或 3 的情况下， $general_profile_compatibility_flag[i]$ 等于 1。
- $general_level_idc$ 表示低于或等于指定级别的级别。
- $general_tier_flag$ 表示低于或等于指定层级的层级。
- $auSizeInSamplesY$ 小于或等于 $N \times MaxLumaPs$ ，其中 $MaxLumaPs$ 在表 A-1 中被指定用于解码器的指定级别。
- 在由外部装置提供输出层集合中的基础层的情况下使变量 $sIdx$ 设置成等于 1 且否则设置成等于 0，在 i 在 $sIdx$ 到 $numLayersInSubBitstream-1$ 的范围(包含端点)内的情况下变量 $picSize[i]$ 为输出层集合中的第 i 层的 $PicSizeInSamplesY$ 的值。存在具有元素 $nn[j]$ (j 在 0 到 k 的范围(包含端点)内)的正整数阵列 nn 和 p 在 0 到 $nn[j]-1$ 的范围(包含端点)内的 $LayerIdx[j][p]$ 的另一阵列，所述阵列对应于由 $nn[j]$ 表示的层，使得所有以下条件为真：
 - j 在 0 到 k 的范围内的情况下， $nn[j]$ 的总和等于 $numLayersInSubBitstream$ 。
 - i 在 $sIdx$ 到 $sIdx+nn[0]-1$ 的范围(包含端点)内的情况下， $picSize[i]$ 的总和小于或等于 $MaxLumaPs$ ，其中 $MaxLumaPs$ 在表 A-1 中被指定用于解码器的指定级别。
 - 当 k 大于 0 时，针对在 1 到 k 的范围(包含端点)内的 m 的每一值，在 i 在 $sIdx+nn[m-1]$ 到 $sIdx+nn[m]-1$ 的范围(包含端点)内的情况下， $picSize[i]$ 的总和小于或等于 $MaxLumaPs$ ，其中 $MaxLumaPs$ 在表 A-1 中被指定用于解码器的指定级别。

[0182]

- 在 i 在 $\text{LayerIdx}[0][0]$ 到 $\text{LayerIdx}[k][\text{nn}[k]]$ 的范围(包含端点)内的情况下, $\text{picSize}[i]$ 的总和小于或等于 MaxLumaPs , 其中 MaxLumaPs 在表 A-1 中被指定用于解码器的指定级别。替代地, 移除此条件, 且可移除以上以粗体标记的文本。

具有特定尺度 N 的在特定层级的特定级别下的符合可缩放主 10 配置文件的 I 类解码器应能够解码所有以下条件都适用的输出层集合的所有子位流:

- $\text{general_profile_idc}$ 等于 1、2、3 或 4, 或在 1 到 4 的范围(包含端点)内的任何 i 值的情况下, $\text{general_profile_compatibility_flag}[i]$ 等于 1。
- ~~general_level_idc 表示低于或等于指定级别的级别。~~
- general_tier_flag 表示低于或等于指定层级的层级。
- auSizeInSamplesY 小于或等于 $N \times \text{MaxLumaPs}$, 其中 MaxLumaPs 在表 A-1 中被指定用于解码器的指定级别。

具有特定尺度 N 的在特定层级的特定级别下的符合可缩放主 10 配置文件的 II 类解码器应能够解码所有以下条件都适用的输出层集合的所有子位流:

- $\text{general_profile_idc}$ 等于 1、2、3 或 4, 或在 1 到 4 的范围(包含端点)内的任何 i 值的情况下, $\text{general_profile_compatibility_flag}[i]$ 等于 1。
- general_level_idc 表示低于或等于指定级别的级别。
- general_tier_flag 表示低于或等于指定层级的层级。
- auSizeInSamplesY 小于或等于 $N \times \text{MaxLumaPs}$, 其中 MaxLumaPs 在表 A-1 中被指定用于解码器的指定级别。
- 在由外部装置提供输出层集合中的基础层的情况下使变量 sIdx 设置成等于 1 且否则设置成等于 0, 在 i 在 sIdx 到 $\text{numLayersInSubBitstream}-1$ 的范围(包含端点)内的情况下变量 $\text{picSize}[i]$ 为输出层集合中的第 i 层的 PicSizeInSamplesY 的值。存在具有元素 $\text{nn}[j]$ (j 在 0 到 k 的范围(包含端点)内)的正整数阵列 nn 和 p 在 0 到 $\text{nn}[j]-1$ 的范围(包含端点)内的 $\text{LayerIdx}[j][p]$ 的另一阵列, 所述阵列对应于由 $\text{nn}[j]$ 表示的层的索引, 使得所有以下条件为真:
 - j 在 0 到 k 的范围内的情况下, $\text{nn}[j]$ 的总和等于 $\text{numLayersInSubBitstream}$ 。
 - i 在 sIdx 到 $\text{sIdx}+\text{nn}[0]-1$ 的范围(包含端点)内的情况下, $\text{picSize}[i]$ 的总和小于或等于 MaxLumaPs , 其中 MaxLumaPs 在表 A-1 中被指定用于解码器的指定级别。
 - 当 k 大于 0 时, 针对在 1 到 k 的范围(包含端点)内的 m 的每一值, 在 i 在 $\text{sIdx}+\text{nn}[m-1]$ 到 $\text{sIdx}+\text{nn}[m]-1$ 的范围(包含端点)内的情况下, $\text{picSize}[i]$ 的总和小于或等于 MaxLumaPs , 其中 MaxLumaPs 在表 A-1 中被指定用于解码器的指定级别。

在 i 在 $\text{LayerIdx}[0][0]$ 到 $\text{LayerIdx}[k][\text{nn}[k]]$ 的范围(包含端点)内的情况下, $\text{picSize}[i]$ 的总和小于或等于 MaxLumaPs , 其中 MaxLumaPs 在表 A-1 中被指定用于解码器的指定级别。替代地, 移除此条件, 且可移除以上以粗体标记的文本。

[0183] 表4——SHVC WD 5和MV-HEVC WD 7中的解码器能力

[0184] 以上定义提供使用尺度 N 来定义解码器能力的示范性方法。然而, 存在其它选项。举例来说, 限制“ auSizeInSamplesY 小于或等于 $N \times \text{MaxLumaPs}$, 其中 MaxLumaPs 在表 A-1 中被指定用于解码器的指定级别”可用“ $\text{numLayersInSubBitstream} \times \text{MaxBsLumaPs}$ 小于或等于 $N \times \text{MaxDeLumaPs}$, 其中 MaxBsLumaPs 在表 A-1 中被指定用于子位流的级别, 且 MaxDeLumaPs 为在表 A-1 中被指定用于解码器的 MaxLumaPs ”替换。决定使用这些选项中的哪一个可取决于用户知晓变量 auSizeInSamplesY 和 $\text{numLayersInSubBitstream}$ 中的哪一个。无论以何种方式, $\text{numLayersInSubBitstream}$ 或 auSizeInSamplesY 的值应直接可用, 无论以会话协商还是其它方式。按文件格式, 值应在样本描述中用信号表示。按实时输送协议 (RTP) 有效负载格

式,应针对此定义媒体类型参数。如果直接地用信号表示auSizeInSamplesY(例如,在VPS扩展或VPS VUI中),那么使用auSizeInSamplesY可为有效选项。

[0185] 如先前所论述,可通过重复使用多个HEVC单层解码器而不改变硬件解码器核心来实施II类解码器。对II类解码器的以上解码器能力描述的关键点如下:存在所有层分配到至少一层组,且每一层组可专门地指派给解码器核心中的一者以用于解码。此包含每一层组的累积图片大小、位速率、CPB大小和DPB大小不得超出对应解码器核心的级别的最大图片大小、位速率、CPB大小和DPB大小的要求。分配所有层的一个特定实例为将每一层分配于一组中,使得每一层可专门地指派给一个解码器核心以用于解码。

[0186] 图4A说明根据本发明中所描述的方面的用于定义解码器能力的实例方法或过程400的流程图。在一个实施例中,过程400开始于框401处。

[0187] 在框405处,过程400可涉及识别位流层至少一次分配到至少一层组中。举例来说,具有五个层的多层位流可分配到五个层组中。

[0188] 在框410处,过程400可涉及检测每一层组是否能够专门地指派给解码器核心中的一者以用于解码位流。在实施例中,可使用多于一个现有单层解码器来实施一些解码器(即,II类解码器)。在框410处可确定此解码器是否能够解码多层位流的每一层组。举例来说,具有五个层组的多层位流的每一层组可能不能够专门地指派给仅包括四个单层解码器核心的解码器的解码器核心。

[0189] 在框415处,过程400可涉及至少部分地基于检测每一层组是否能够专门地指派给解码器核心中的一者而确定解码器是否能够解码位流。举例来说,如果多层位流的层组无法各自专门地指派给解码器的个别核心,那么可确定解码器不能够解码位流。然而,在一些实施例中,其它因素可促进确定解码器是否能够解码位流。在这些实施例中的一些中,可个别地或组合使用这些因素以定义解码器能力。

[0190] 过程400结束于框420处。取决于实施方案,可在过程400中添加和/或省略框,且取决于特定实施方案,可按不同次序执行过程400的框。

[0191] 在实施例中,为了进一步定义解码器能力,可指定解码器的级别定义。级别定义可传达关于解码器解码位流的能力的信息。在一些情况下,级别可提供解码器可解码的最大位流级别(例如,层的数目)。根据所定义级别,可预期解码器能够解码级别等于或小于解码器级别的任何位流。系统还可确定解码器可解码的最大亮度图片大小。在一些情况下,解码器的级别定义可指示解码器的最大亮度图片大小。举例来说,表A-1提供针对每一解码器级别的最大亮度图片大小。系统还可计算位流中的层的分辨率的总和。在一些情况下,可使用一个存取单元中的图片的分辨率来计算位流的分辨率的总和。用户还可指定尺度。可随后通过使尺度和级别定义与解码器相关联来定义解码器能力。

[0192] 图4B说明根据本发明中所描述的方面的用于进一步定义解码器能力的实例过程400'的流程图。应注意,过程400'可为上文参考图4A描述的过程400的部分,或可为分离于过程400的独立过程。在一个实施例中,过程400'开始于框450处。

[0193] 在框455处,过程400'可涉及指定解码器的级别定义。级别定义可传达关于解码器解码位流的能力的信息。在一些情况下,级别可提供解码器可解码的最大位流级别(例如,层的数目)。根据所定义级别,可预期解码器能够解码级别等于或小于解码器级别的任何位流。

[0194] 在框460处,过程400'可涉及确定解码器可解码的最大亮度图片大小。在一些情况下,解码器的级别定义可指示解码器的最大亮度图片大小。举例来说,表A-1提供针对每一解码器级别的最大亮度图片大小。

[0195] 在框465处,过程400'可涉及指定与位流的级别相关联的尺度。为了确定恰当尺度,用户或系统可验证解码器可解码如由本文中所描述的解码器能力使用特定尺度N指定的尽可能多的位流。

[0196] 在决定框470处,过程400'可涉及决定是否使用位流中的层的分辨率的总和来确定解码器能力。在一些情况下,使用位流中的层的分辨率的总和来确定解码器能力可为有利的。举例来说,如果直接地用信号表示auSizeInSamplesY(例如,在VPS扩展或VPS VUI中),那么使用auSizeInSamplesY可为确定解码器能力的有用方法。然而,系统可使用多种测试中的一或多者来确定是否使用分辨率的总和,包含位流中的层的数目还是auSizeInSamplesY更容易获得。对于SNR可缩放性(即,其中所有层具有相同分辨率),使用位流中的层的数目或auSizeInSamplesY为等效的。在此情况下,系统可使用位流中的层的数目,因为此选项更简单。

[0197] 在框475处,如果过程400'涉及决定使用分辨率的总和(如在决定框420处所确定),那么过程400'还可涉及计算分辨率的总和。

[0198] 然而,如果过程400'涉及决定不使用分辨率的总和(如在决定框470处所确定),那么过程400'可涉及在框480处确定位流中的层的数目。

[0199] 在框485处,过程400'可涉及使用位流中的层的数目和/或位流的任何其它识别信息来定义位流的级别定义。

[0200] 在框490处,正如解码器的情况,过程400'可涉及确定位流的最大亮度图片大小。举例来说,系统可使用表A-1以使用位流的级别定义来确定位流的最大亮度图片大小。

[0201] 在框495处,过程400'可涉及确定解码器是否能够解码位流。取决于所确定的因素,过程400'可涉及使用分辨率的总和或位流的层的数目来确定解码器能力。本文中揭示关于如何根据这些因素确定能力的其它信息。

[0202] 过程400'结束于框496处。取决于实施方案,可在过程400'中添加和/或省略框,且取决于特定实施方案,可按不同次序执行过程400'的框。

[0203] 添加尺度可提供可允许更准确地定义解码器能力的广泛有用信息。举例来说,通过使用现有级别定义,如果多于4个720p分辨率的SNR层(或等效数目个亮度像素,例如,更多具有经组合的空间和SNR层的层)待解码,那么将需要5级或更高的解码器。因此,亮度CTB大小将等于 32×32 或 64×64 ,对于720p或更低的分辨率,此可能导致次最佳的译码效率,因为无法使用较小CTB大小,例如 16×16 。然而,通过使尺度与4级解码器相关联,将清楚确定某些4级解码器适合于解码4个720p分辨率的SNR层。此将允许使用小于 32×32 的CTB大小,此可提供经改进的译码效率。

[0204] 在另一实例中,按照现有级别定义,由4个单层HEVC 3.1级解码器组成的II类解码器(即,通过重复使用多个现有单层解码器制造的那些解码器)必须符合4级或更高以解码4个720p的SNR层。根据此定义,解码器必须能够解码任何4级位流。然而,除非改变解码器硬件,否则此解码器不能够解码具有2个1080p分辨率的SNR层的SHVC4级位流。此问题通过所改变的级别定义来解决,确切地说是通过添加与级别相关联的尺度以及所改变的解码器能

力要求来解决。作为这些变化的结果,不需要能够解码具有4个720p分辨率的SNR层的位流的相符解码器来解码具有2个1080p分辨率的SNR层的位流。

[0205] 现有HEVC级别定义引起的另一示范性问题为,以使得能够解码1080p的单层HEVC位流和720p的双层SHVC位流二者的方式实施的解码器被标注为3.1级。然而,3.1级标签并不表达解码1080p的单层位流的能力。此问题通过使尺度与解码器相关联以进一步定义如上文所描述的解码器能力来解决。具体来说,3.1级解码器将与尺度2相关联,此传达解码具有1080p的单个层的位流和具有720p的两个层的位流二者的能力。

[0206] 在另一实例中,对于使用4个单层HEVC 3.1解码器实施以能够解码4个720p的SNR层的解码器,按照现有级别定义,解码器必须符合4级或更高。因此,解码器需要能够解码具有多于3个平铺块行和多于3个平铺块列的位流,每一平铺块具有256个亮度样本的宽度和144个亮度样本的高度。然而,解码器的3.1级限制将不能够解码一些此类位流。此问题通过如上文所描述的所改变的级别定义和所改变的解码器能力要求来解决,使得所描述的解码器符合3.1级且具有相关联尺度4,而非符合4级和尺度1。因此,3.1级的对平铺块行和列的数目的现有限制(而非4级的那些限制)将适用。

[0207] 因为通过重复使用现有解码器硬件核心来实施II类解码器,所以必须由解码器能力的定义中的约束禁止由两个解码器硬件核心联合地解码任何单个图片。此说明于图5A到5B和6A到6B中。

[0208] 图5A说明根据本发明中所描述的方面的示范性多层位流,其中位流为5层SHVC位流,其中层0(在510处)为四边高清(QHD) (960×540),且层1(在511处)、2(在512处)、3(在513处)和4(在514处)为高清(HD) (1080p)。

[0209] 图5B说明根据本发明中所描述的方面的用于解码图5A的多层位流的多个示范性选项。在本实施例中,选项1(在521处)表示使用五个4.0级解码器核心,其中每一解码器核心解码位流的一个层。选项2(在522处)表示使用四个4.0级解码器核心来解码所有五个层。通过基于auSizeInSamplesY限制($\text{auSizeInSamplesY} < 4 \times \text{MaxLumaPs}$)使用I类解码器,此选项为可能的。解码器核心解码一个完整层和一个部分层。然而,此不可能通过使用II类解码器(通过重复使用现有解码器硬件核心而不变化所实施的那些解码器)实现。选项3(在523处)表示使用四个有5.0级能力的解码器核心。在此选项中,解码器核心解码整数数目个完整层。此选项可能使用任一种类的解码器实现。

[0210] 图6A说明根据本发明的方面的另一示范性多层位流。在本实施例中,位流为5层SHVC位流,其中层0到4(在610到614处)为扩展图形阵列(XGA) (960×540)。位流表示SNR可缩放性(即,所有层具有相同分辨率)。

[0211] 图6B说明根据本发明的方面的用于解码图6A的多层位流的多个示范性选项。在本实施例中,选项1(在621处)表示使用五个3.1级解码器核心,其中每一解码器核心解码位流的一个层。选项2(在622处)表示使用四个3.1级解码器核心来解码所有五个层。出于上文参考图5B论述的原因,此选项可用于I类解码器但不可用于II类解码器。

[0212] 本发明的一些实施方案可包含一或多个经解码图片缓冲器(DPB)。每一DPB可进一步包含多个子DPB。DPB和子DPB可经配置以存储和传送数据。在一些实施方案中,可定义与DPB和/或子DPB相关的多个大小相关的限制。此类限制可基于多种因素,例如分辨率的数目、亮度图片大小、色度图片大小、经译码图片缓冲器大小、每一图片的切片段、平铺块行的

数目、列的数目等。此外,大小相关的限制一般可被指定用于位流的个别层以及子DPB。举例来说,大小相关的限制可被指定用于位流的每一层,且另一大小相关的限制可被指定用于每一子DPB。

[0213] 另外,对CPB大小的位流特定限制可以其可随层的数目缩放的方式被指定。以此方式,在存在许多层时可达成经改进的高图片质量。对应于HEVC文本的子条款A.4.2中的第b、c、d、g、h、i和j项的位流特定限制是以其可随层的数目缩放的方式被指定。

[0214] 本发明的实施例还可提供用信号表示输出层集合的位流或子位流中的层的数目的值的系统和方法。此用信号表示可由根据ISO媒体文件格式的文件的样本描述中的参数、RTP有效负载格式的媒体类型参数、会话描述协议(SDP)参数、HTTP动态自适应流式传输(DASH)内容的媒体呈现描述(MPD)中的MPD参数等来实现。

[0215] 此外,一些实施例可提供用信号表示位流中的输出层集合的位流或子位流中的所有层的亮度样本的数目总和的值的系统和方法。此用信号表示可例如由根据ISO媒体文件格式的文件的样本描述中的参数、RTP有效负载格式的媒体类型参数、SDP参数或DASH内容的MPD中的MPD参数来实现。

[0216] 结合本文中所示的实施例而描述的各种说明性逻辑框、模块、电路和算法步骤可被实施为电子硬件、计算机软件或两者的组合。为了清楚地说明硬件与软件的此可互换性,上文已大体上关于其功能性而描述了各种说明性组件、框、模块、电路和步骤。此功能性是实施为硬件还是软件取决于特定应用和施加于整个系统的设计约束。所属领域的技术人员可针对每一特定应用以不同方式来实现所描述的功能性,但此类实施方案决策不应被解释为会导致脱离本发明的范围。

[0217] 本文中所描述的技术可在硬件、软件、固件或其任一组合中实施。这些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具有多种用途(包含在无线通信装置手持机和其它装置中的应用)的集成电路装置。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实施为离散但可互操作的逻辑装置。如果在软件中实施,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被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)、闪存、磁性或光学数据存储媒体等等。另外或替代地,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携带或传达指令或数据结构的形式的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DSP、通用微处理器、ASIC、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所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、结合DSP核心的一或多个微处理器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,或适合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式视频编码器-解码器 (CODEC) 中。

[0219] 本文中所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的源装置,所述编码视频数据将在稍后时间由目的地装置解码。具体来说,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、电话手持机(例如,所谓的“智能”电话)、所谓的“智能”板、电视、相机、显示装置、数字媒体播放器、视频游戏机、视频流式传输装置等等。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源装置能够实时将经编码视频数据直接传输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)来调制,且被传输到目的地装置。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。通信媒体可能形成基于数据包的网络(例如,局域网、广域网或全球网络,例如因特网)的部分。通信媒体可包含路由器、交换机、基站或可用于促进从源装置到目的地装置的通信的任何其它设备。

[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,可由输入接口从存储装置存取经编码数据。存储装置可包含多种分布式或本地存取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、闪存、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,存储装置可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。目的地装置可经由串流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并将所述经编码视频数据传输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页服务器(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)或两者的组合。经编码视频数据从存储装置的传输可为流式传输、下载传输或其组合。

[0222] 本发明的技术未必限于无线应用或设置。所述技术可应用于支持多种多媒体应用中的任一者的视频译码,例如空中电视广播、有线电视传输、卫星电视传输、因特网流视频传输(例如,HTTP动态自适应流式传输(DASH))、被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,从而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中所揭示的技术。在其它实例中,源装置和目的地装置可包含其它组件或布置。举例来说,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] 以上实例系统仅为一个实例。用于并行地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由视频编码装置执行,但所述技术还可由视频编码器/解码器(通常被称作“CODEC”)执行。此外,本发明的技术还可由视频预处理器执行。源装置和目的地装置仅为源装置产生经译码视频数据以供传输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大体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传输,例如用于视频流式传输、视频回放、视频广播或视频电话。

[0225] 视频源可包含视频捕获装置,例如摄像机、含有先前所捕获视频的视频存档和/或用以从视频内容提供者接收视频的视频馈入接口。作为另一替代方案,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,如果视频源为视频相机,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一般可适用于视频译码,且可应用于无线和/或有线应用。在每一情况下,可由视频编码器对所捕获、预捕获或计算机产生的视频进行编码。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] 如所提及,计算机可读媒体可包含短暂性媒体,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、闪存驱动器、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,网页服务器(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语法信息,所述语法信息也由视频解码器使用,所述语法信息包含描述块和其它经译码单元(例如,图片群组(GOP))的特性和/或处理的语法元素。显示装置将经解码视频数据显示给用户,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。

10 ↗

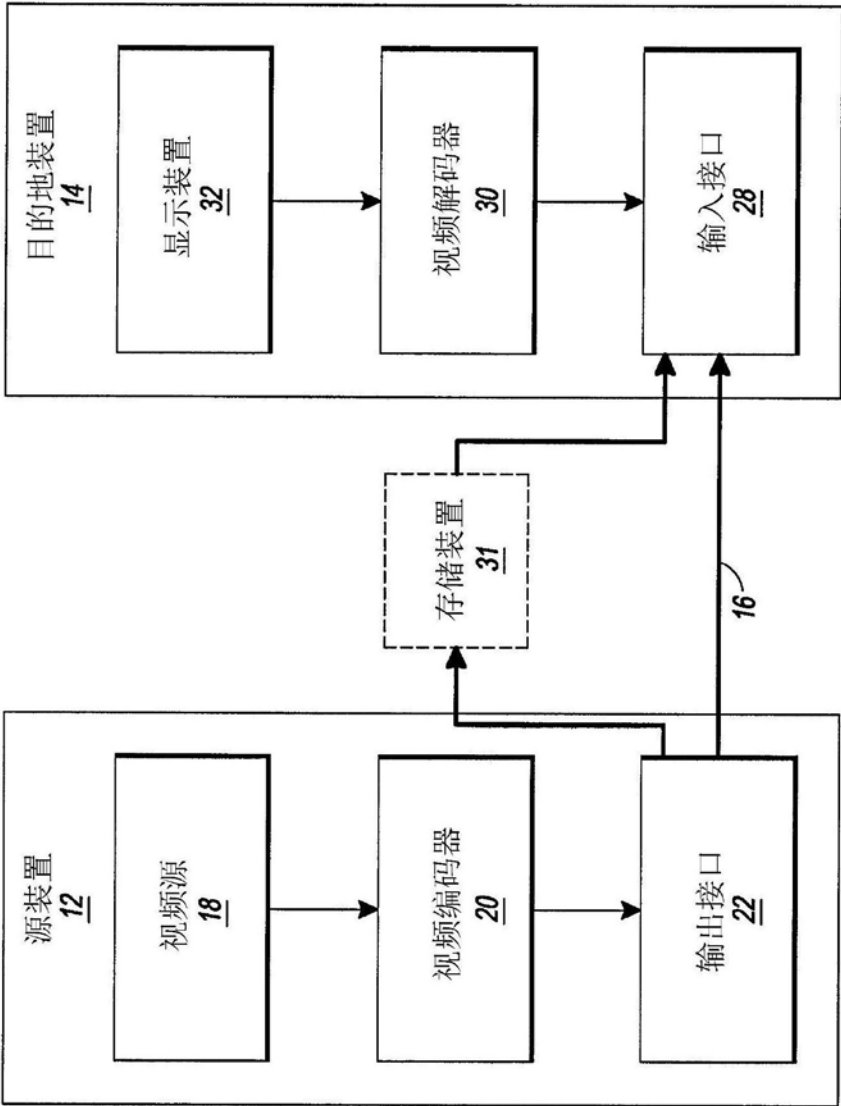


图1A

10'

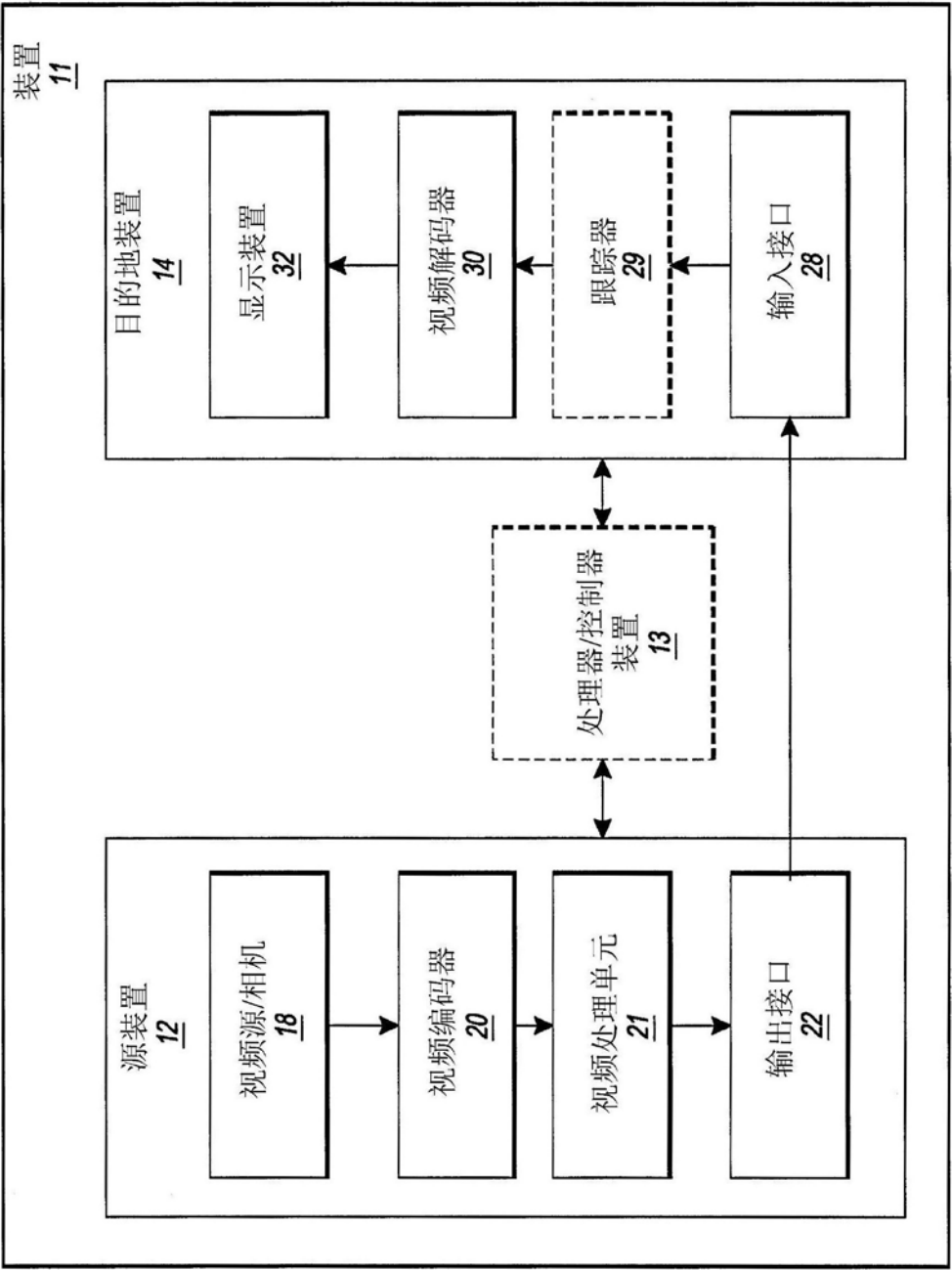


图1B

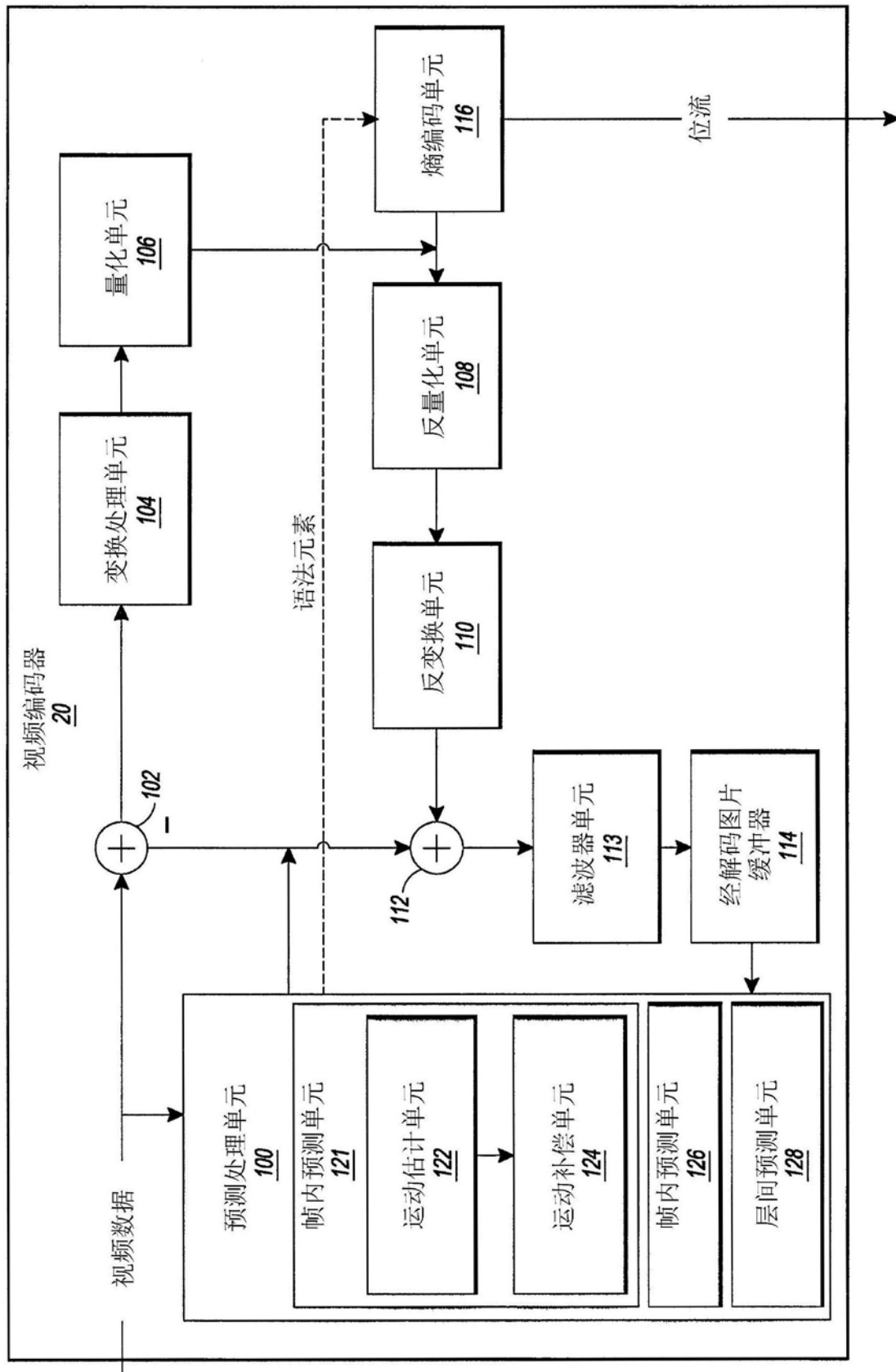


图2A

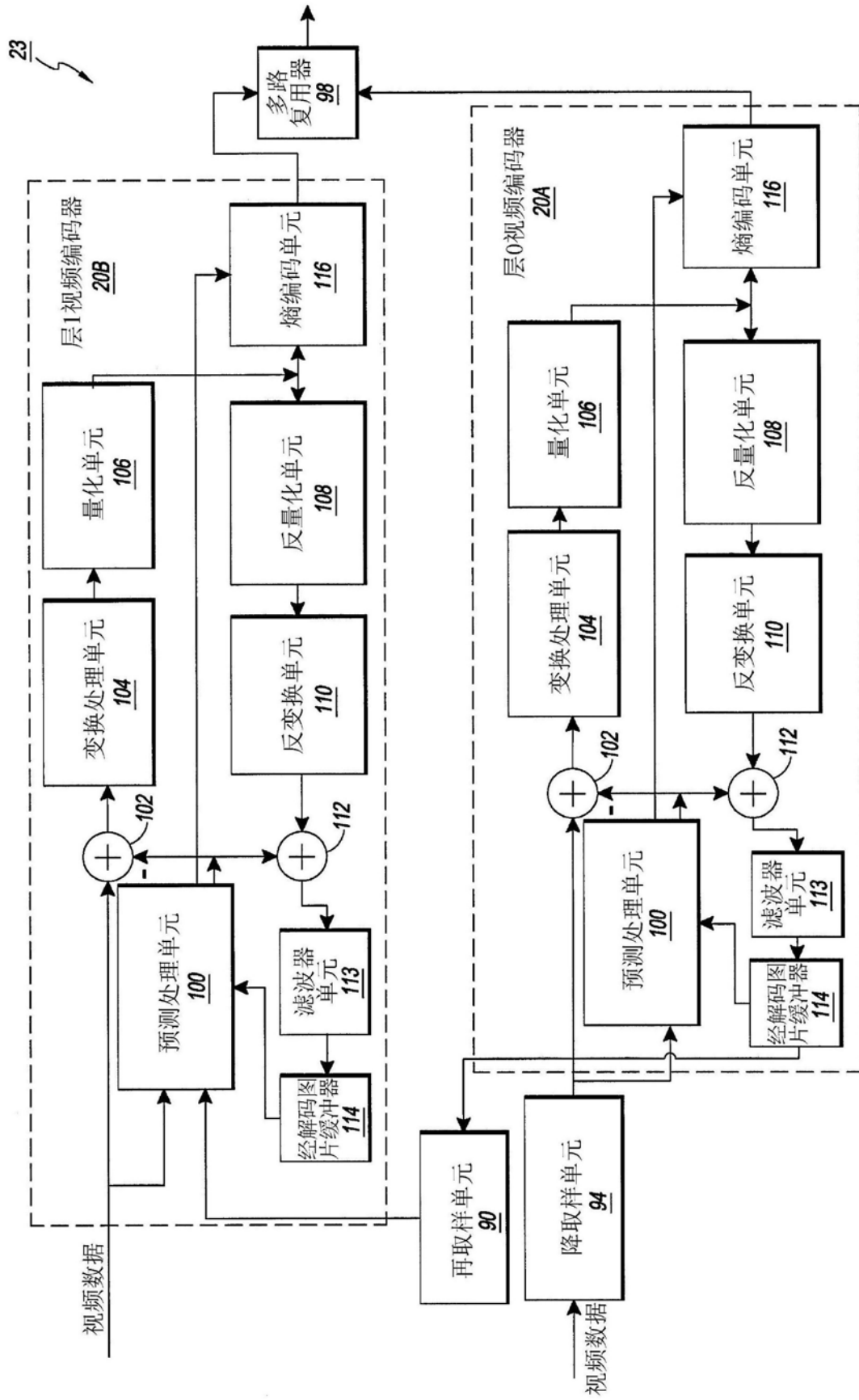


图2B

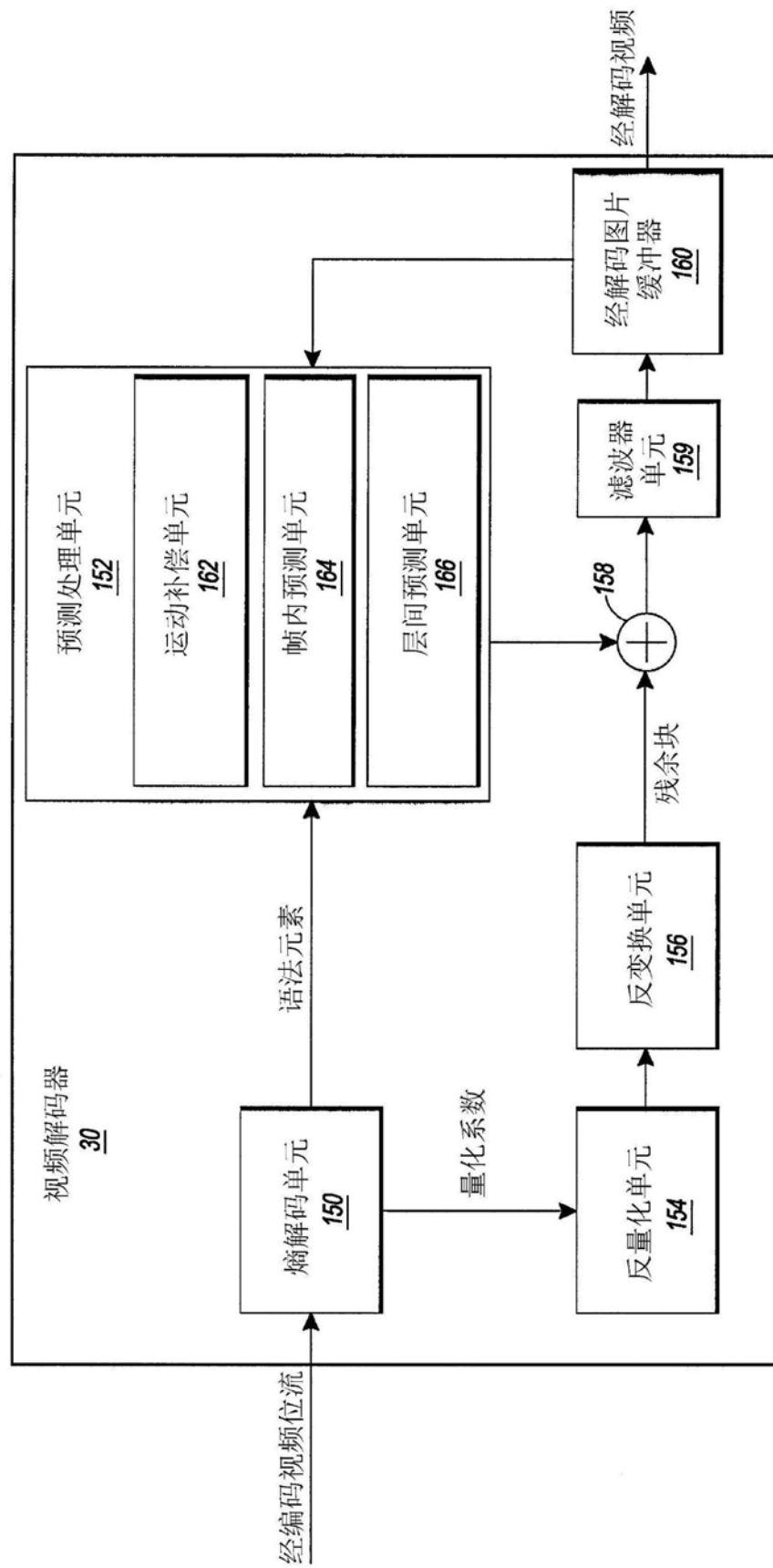


图3A

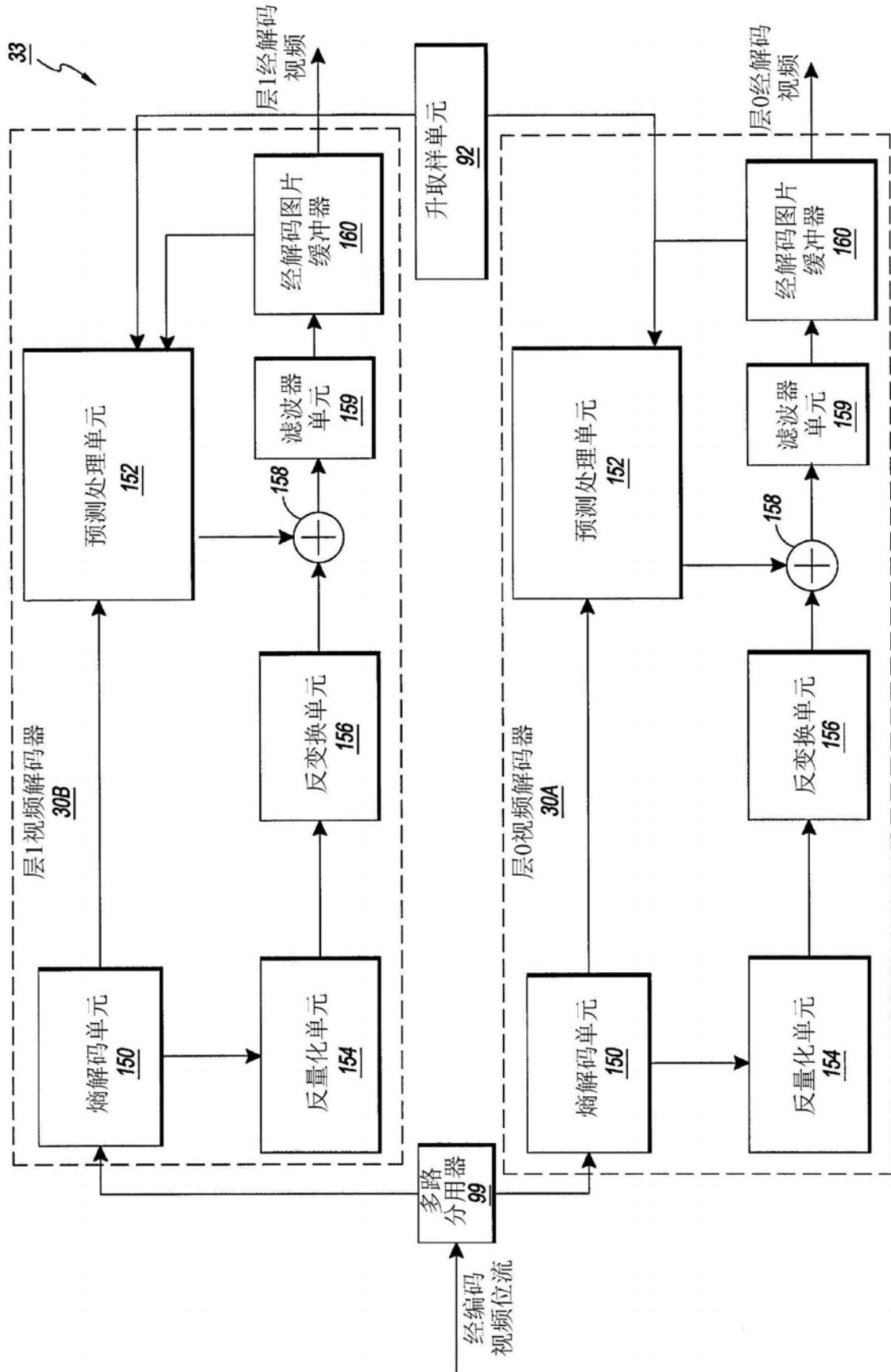


图3B

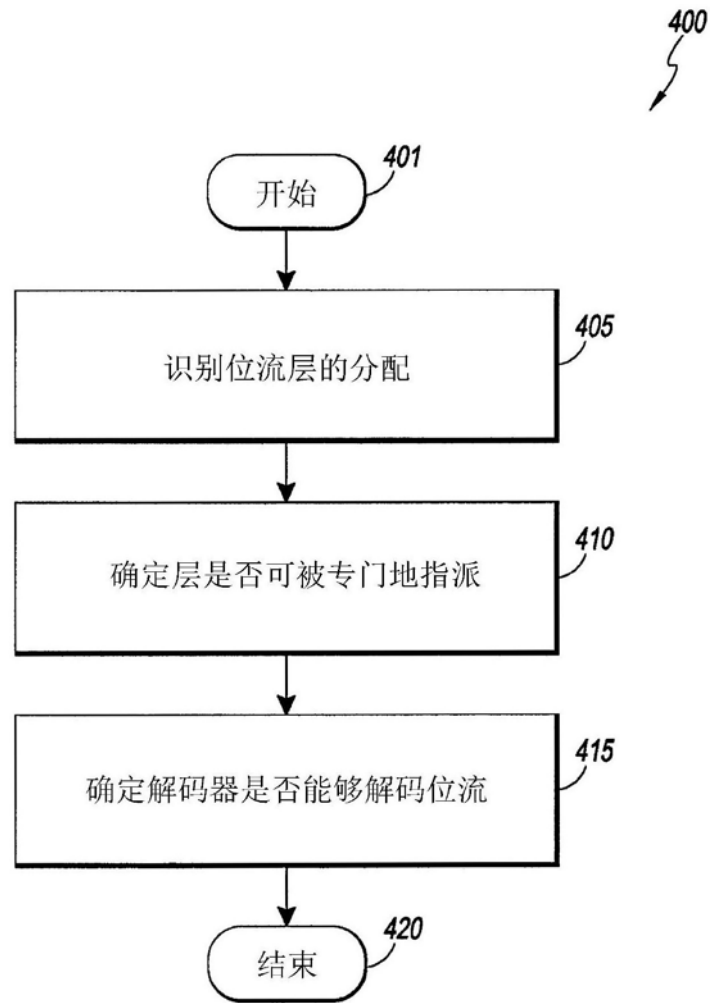


图4A

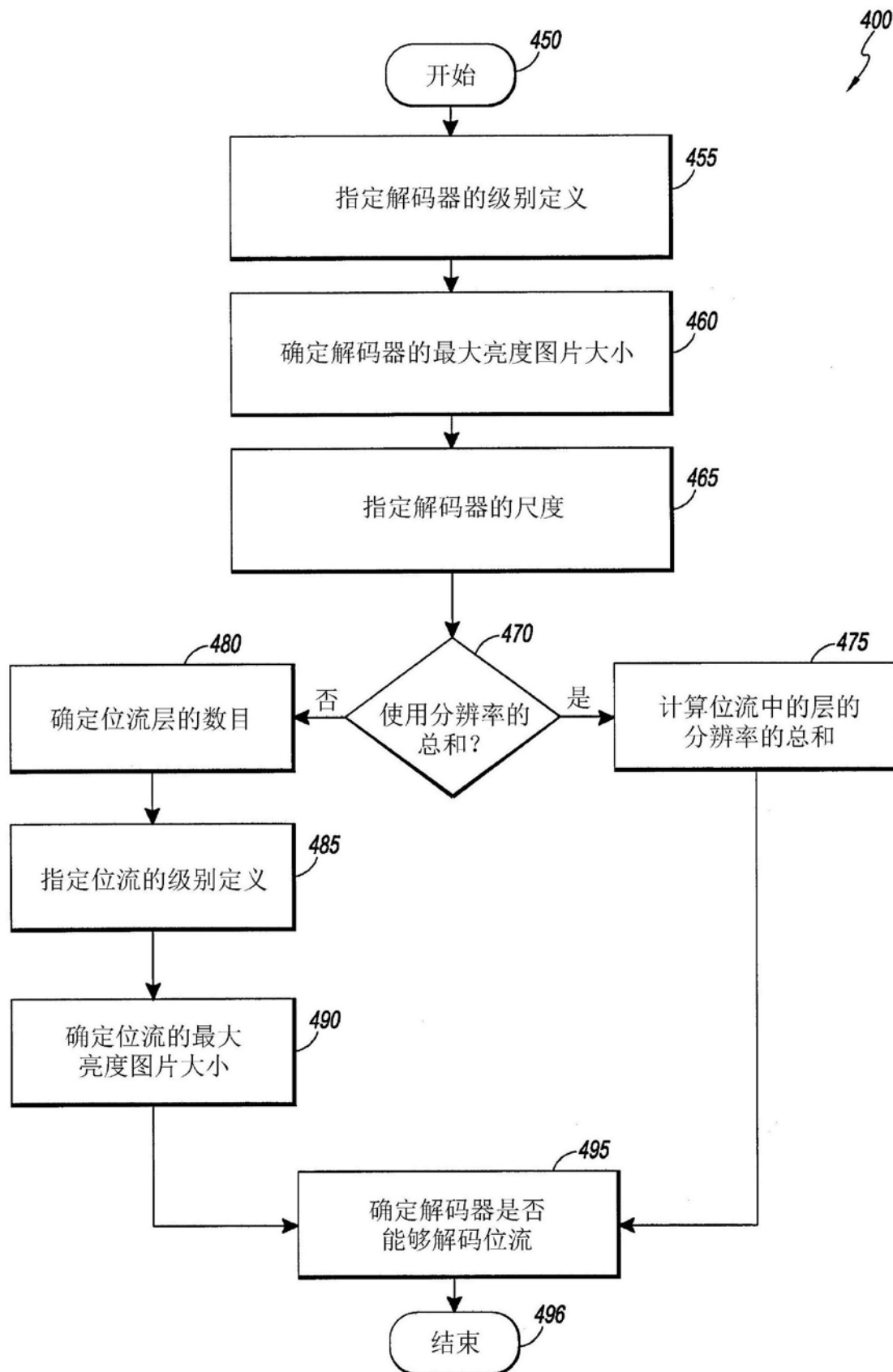


图4B

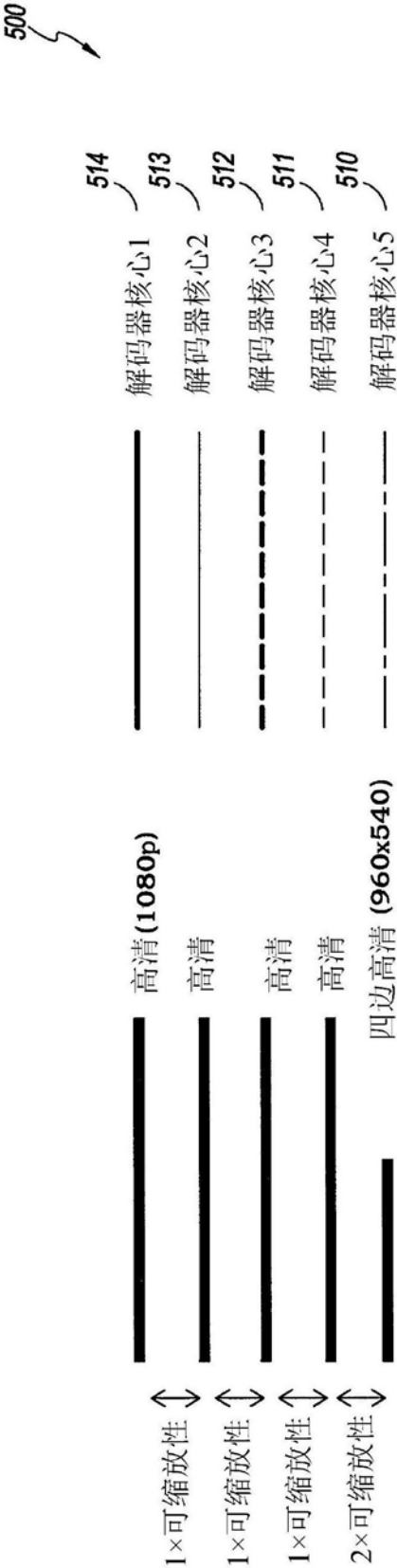


图5A

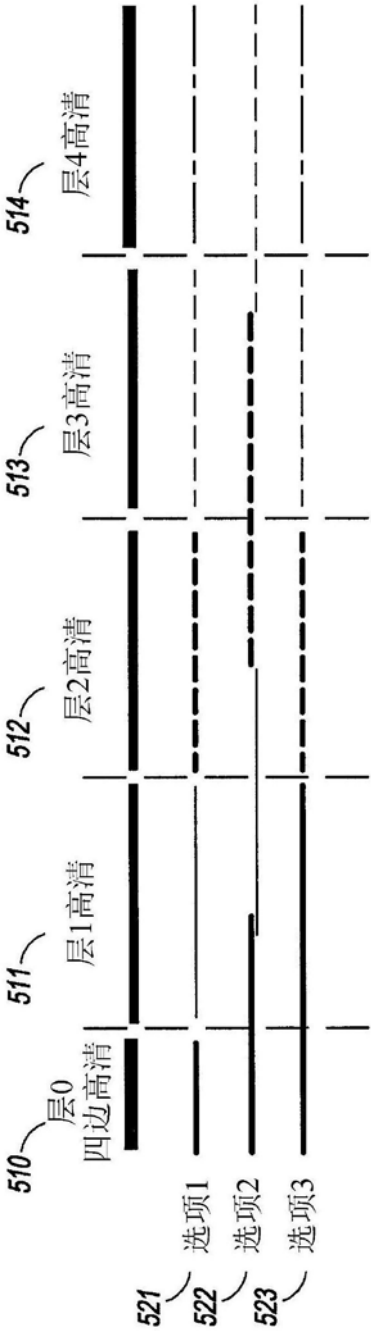


图5B

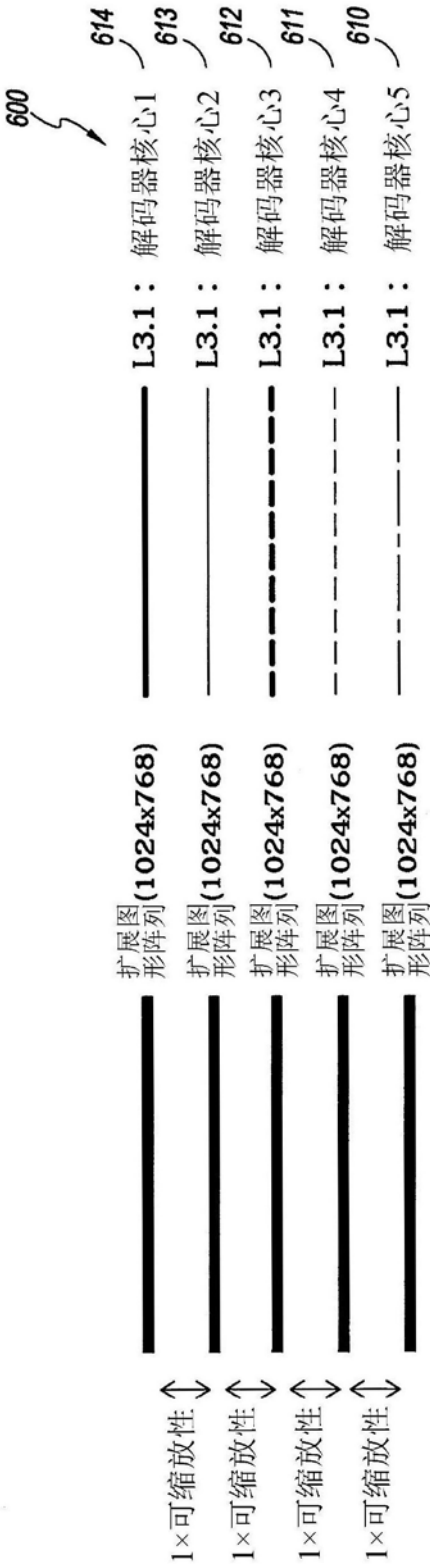


图6A

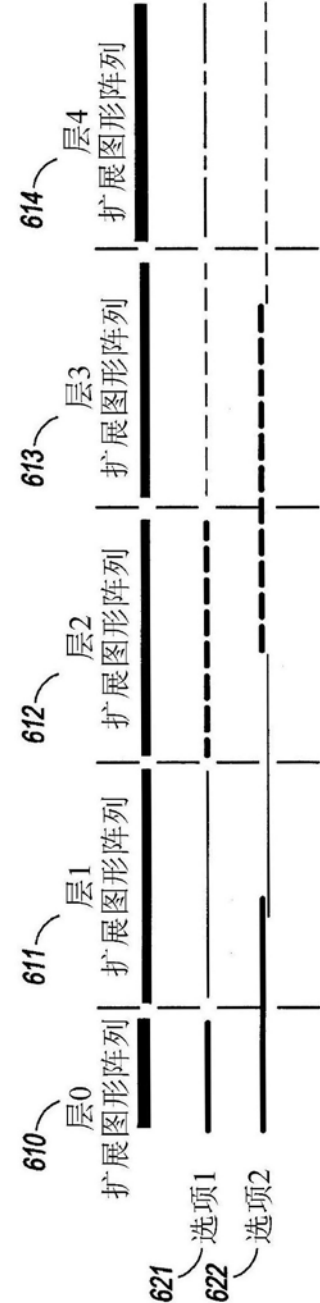


图6B