



## (12) 发明专利申请

(10) 申请公布号 CN 116366835 A

(43) 申请公布日 2023.06.30

(21) 申请号 202310275945.3

(51) Int.Cl.

(22) 申请日 2014.01.04

H04N 19/102 (2014.01)

(30) 优先权数据

H04N 19/119 (2014.01)

61/749,042 2013.01.04 US

H04N 19/12 (2014.01)

61/809,605 2013.04.08 US

H04N 19/139 (2014.01)

(62) 分案原申请数据

H04N 19/156 (2014.01)

201480012232.7 2014.01.04

H04N 19/174 (2014.01)

(71) 申请人 GE视频压缩有限责任公司

H04N 19/187 (2014.01)

地址 美国纽约

H04N 19/30 (2014.01)

(72) 发明人 卡斯滕·聚林 托马斯·席尔

H04N 19/31 (2014.01)

瓦莱里·乔治 德特勒夫·马佩

H04N 19/33 (2014.01)

罗伯特·斯库平

H04N 19/436 (2014.01)

(74) 专利代理机构 北京康信知识产权代理有限

H04N 19/44 (2014.01)

责任公司 11240

H04N 19/46 (2014.01)

专利代理人 刘彬

H04N 19/59 (2014.01)

H04N 19/70 (2014.01)

H04N 19/80 (2014.01)

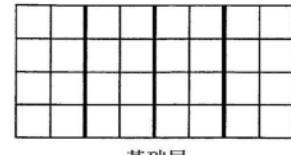
权利要求书2页 说明书61页 附图35页

## (54) 发明名称

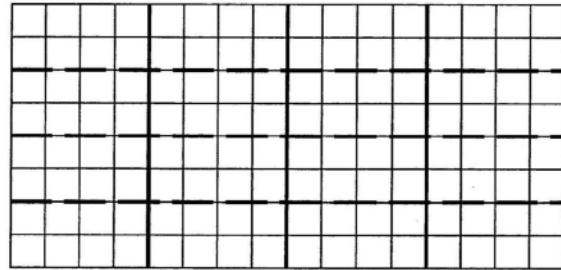
高效可伸缩编码概念

## (57) 摘要

本公开描述高效可伸缩编码概念。视频解码器，解码情景在层的层级中编码成的多层视频数据流，根据通过细分所述层的图片而得到的空间区段对多层视频数据流解码，包括：熵解码器，从多层视频数据流熵解码数据获得空间区段的残差数据和编码参数；预测器，至少基于编码参数生成与空间区段相关的预测信号；组合器，将预测信号和残差数据组合以重构图片的一部分，通过从多层视频数据流解码长期和短期语法元素确定第一层和第二层的图片的细分，长期语法元素有指示第二层图片按第二层图片一个空间区段与第一层图片的相应空间区段空间对齐方式细分的第一值，短期语法元素包括指示第二层的至少一个图片与第一层的图片的相应空间区段空间上对齐的第二值。



基础层



增强层

1. 一种视频解码器, 用于解码情景在层的层级中被编码成的多层视频数据流, 其中, 所述视频解码器被配置为根据通过细分所述层的图片而得到的空间区段对所述多层视频数据流解码, 所述视频解码器包括:

熵解码器, 被配置用于从所述多层视频数据流熵解码数据以获得关于所述层的第一层的图片或第二层的图片的各个空间区段的残差数据和编码参数;

预测器, 被配置用于至少基于所述编码参数生成与所述空间区段相关的预测信号;

组合器, 被配置用于将所述预测信号和所述残差数据组合以重构所述图片的一部分,

其中, 所述视频解码器被配置用于通过从所述多层视频数据流解码长期语法元素和短期语法元素来确定所述第一层和所述第二层的图片的细分, 其中, 所述长期语法元素具有指示所述第二层的图片按照所述第二层的所述图片的至少一个所述空间区段与所述第一层的图片的相应的一个所述空间区段空间上对齐的方式被细分的第一值, 以及

基于所述短期语法元素确定所述第一层和所述第二层的图片至所述第一层和所述第二层各自的空间区段的细分, 所述短期语法元素包括指示所述第二层的图片的空间区段不与所述第一层的图片的空间区段空间上对齐的第一值, 以及指示所述第二层的图片的图片的至少一个空间区段与所述第一层的图片的相应的一个所述空间区段空间上对齐的第二值。

2. 根据权利要求1所述的视频解码器, 其中, 所述视频解码器被配置用于:

使用帧内图片空间预测对所述第一层的图片或第二层的图片进行解码, 并且用于各个空间区段的所述帧内图片空间预测在相应空间区段的边界线处中断; 或者

通过将所述第一层的图片的空间区段并行解码且支持跨所述第一层的图片的空间区段的边界线的所述帧内图片空间预测并且服从所述第一层的图片的空间区段的解码之间的解码延迟; 以及通过将所述第二层的图片的空间区段并行解码且支持跨所述第二层的图片的空间区段的边界线的所述帧内图片空间预测并且服从所述第二层的图片的空间区段的解码之间的解码延迟, 来使用所述帧内图片空间预测对所述层的图片进行解码。

3. 根据权利要求1所述的视频解码器, 其中, 所述图片被细分成的所述空间区段包括片。

4. 根据权利要求3所述的视频解码器, 其中, 所述视频解码器被配置用于:

使用帧内图片空间预测对所述层的图片进行解码, 且用于各个片的所述帧内图片空间预测在相应片的边界线处中断。

5. 根据权利要求1所述的视频解码器, 其中, 所述第一层的图片的各个空间区段精确地由第二层的图片的n个空间区段构成, n取决于所述长期语法元素的值, 以及

如果所述长期语法元素具有第二值, 则所述视频解码器被配置用于在比预定时间段短的时间间隔内基于所述短期语法元素确定所述预定时间段内的层间偏移量。

6. 根据权利要求1所述的视频解码器, 其中, 所述视频解码器被配置为根据所述长期语法元素的所述值来决定开始或者不开始尝试对所述多层视频数据流的所述第二层进行解码。

7. 根据权利要求1所述的视频解码器, 其中, 所述视频解码器是混合式视频解码器。

8. 根据权利要求1所述的视频解码器, 其中, 所述长期语法元素具有指示在预定时间段期间所述第二层的所述图片被细分使得所述第二层的图片的空间区段之间的边界与所述

第一层的图片的空间区段的每一个边界重叠的第一值。

9. 一种视频编码器，用于将场景编码成层的层级中的多层视频数据流，使得所述多层视频数据流能够根据通过细分所述层的图片而得到的空间区段进行解码，其中，所述视频编码器包括：

预测器，被配置用于基于编码参数生成关于所述层的第一层的图片或第二层的图片的各个空间区段的预测信号；

残差确定器，被配置用于确定与所述图片的空间区段有关的残差数据；以及

熵解码器，被配置用于将所述编码参数和所述残差数据解码为所述多层视频数据流，

其中，所述视频编码器被配置用于通过将长期语法元素和短期语法元素插入所述多层视频数据流来信号通知所述第一层和所述第二层的图片的细分，

所述短期语法元素定义所述第一层和所述第二层的图片至所述第一层和所述第二层各自的空间区段的细分，所述短期语法元素包括指示所述第二层的图片的空间区段不与所述第一层的图片的空间区段空间上对齐的第一值，以及指示所述第二层的图片的至少一个空间区段与所述第一层的图片的相应的一个所述空间区段空间上对齐的第二值，以及

所述长期语法元素具有指示所述第二层的图片按照所述第二层的所述图片的至少一个所述空间区段与所述第一层的图片的相应的一个所述空间区段空间上对齐的方式被细分的第一值。

10. 根据权利要求9所述的视频编码器，其中，所述视频编码器被配置用于：

使用帧内图片空间预测对所述第一层的图片或第二层的图片进行编码，且用于各个空间区段的所述帧内图片空间预测在相应空间区段的边界线处中断；或者

通过支持跨所述第一层的图片的空间区段的边界线的所述帧内图片空间预测并且对用于将所述第一层的图片的空间区段的子集单独地进行熵编码的熵上下文概率进行初始化，或者根据该子集之间的顺序在采用处于中间适配状态的所述第一层的图片的空间区段的之前子集的所述熵上下文概率的基础上；以及通过支持跨所述第二层的图片的空间区段的边界线的所述帧内图片空间预测并且对用于将所述第二层的图片的空间区段中的子集单独地进行熵编码的所述熵上下文概率初始化，或者根据该子集之间的顺序在采用处于中间适配状态的所述第二层的图片的空间区段的之前子集的所述熵上下文概率的基础上，使用所述帧内图片空间预测并利用适配所述熵上下文概率的熵编码对所述层的图片进行编码。

## 高效可伸缩编码概念

[0001] 本申请是分案申请,其母案申请的申请号为201811477939.1,申请日为2014年1月4日,发明名称为“高效可伸缩编码概念”。

### 技术领域

[0002] 本申请涉及诸如可伸缩视频编码等可伸缩编码概念。

### 背景技术

[0003] 在本领域中,已知可伸缩编码概念。例如,在视频编码中,H.264涉及SVC扩展(可伸缩视频编码),从而允许基础层编码视频数据流附加有额外的增强层数据,以使得从诸如空间分辨率、信噪比(SNR)等不同的方面提升基础层视频质量的重构品质。通过SVC配置文件也可扩展最近定稿的HEVC标准。HEVC在许多方面不同与其前期H.264,诸如,用于并行解码/编码的适用性和低延迟传输等。至于涉及的并行编码/解码,HEVC支持WPP(波前并行处理)编码/解码以及片(tile)并行处理概念。根据WPP概念,以行式方式将各个图片分割成子流。自左至右指导每个子流内的编码顺序。子流具有其中限定的解码顺序,即,从顶部子流引导至底部子流。使用概率自适应性执行子流的熵编码。针对各个子流独立完成概率初始化,或者基于熵编码时使用的概率的初始自适配状态,使紧随的前一子流在第二CTB(编码树块)结束时从前一子流的左手边缘相应地到达特定位置。空间预测不需要受限制。即,空间预测可跨越中间连续子流之间的边界。通过这种方式,可将子流与形成波前的当前编码/解码的位置并行编码/解码,波前通过自左下方至右上方、自左至右的倾斜方式前行。根据片概念,将图片分段成片并且为了使这些片的编码/解码可能经过并行处理,抑制跨边界线的空间预测。仅可允许跨片边界线的环内滤波。为了支持低延迟处理,切片概念被延伸为:允许切片可切换,以将熵概率重新初始化,采用在处理之前子流过程中保存的熵概率,即,当前切片开始属于的子流之前的子流,并且采用连续被更新的熵概率,直至紧随的前一切片结束。通过这种措施,使得WPP和片概念更适用于低延迟处理。

[0004] 然而,更有利于具有进一步改善可伸缩编码概率的正在讨论的概念。

[0005] 因此,本发明的目标是提供进一步改善可伸缩编码概念的概念。

[0006] 通过未决独立权利要求的主题实现本目标。

### 发明内容

[0007] 本申请的第一方面涉及结合并行处理概念的可伸缩视频编码。诸如WPP和片区段等并行处理概念允许以将图片细分而成的空间区段(例如以子流、片、或者切片的形式)将视频的图片并行解码。正如空间帧内图片预测,该层间预测通过层间预测限制彼此相互依赖的解码层的并行程度。更精确地,空间帧内层预测限制单层的图片解码时的并行程度,通过不同方式解决这种问题。例如,当使用片作为空间区段时,空间帧内层预测受限于不能跨越边界线。在WPP子流的情况下,通过交错方式执行其并行处理以产生适当的倾斜处理波阵面。在层间预测的情况下,根据参考层的共置部分执行依赖层的解码。因此,依赖层的空间

区段的解码可最先开始,此时,已经将参考层的共置部分处理/解码。如同不同视图作为不同层的层间预测的情况,或者归因于自低层至高层的上采样,在允许“运动补偿”的情况下,放大了“共置部分”的区域。即,易于使使用层间预测的可伸缩解码与并行解码的视频解码器从关于这些相关依赖层的短期语法元素推导相互依赖层的并行处理的并行程度,短期语法元素限定将这些相互依赖层的图片细分成其空间区段。然而,稳定地完成该操作是比较繁琐并且在计算上比较复杂。除此之外,当执行该操作时,视频解码器不能够适当地调度并行运行的解码线程以将多层视频数据流解码。因此,根据本发明的第一方面,通过引入长期语法元素结构改善了多层视频数据流的相互依赖层的并行解码,当长期语法元素结构假设特定值时,保证了视频解码器在比短期语法元素的时间间隔更大的预定时间段内,将依赖层的图片细分成使得第二层的图片的空间区段之间的边界与第一层的空间区段的每个边界重叠。通过这种措施,视频解码器能够依赖于多层视频数据流被适当地编码的事实,以使得将相互依赖层的图片细分成空间区段,而不疏忽地降低这些相互依赖层之间可行的并行程度。更确切地,在预定时间段内,通过开发不同层中的空间区段的边界通过信号方式彼此重叠的约束,解码器可提前将空间区段的分布调度到视频解码器的并行处理线程中。尽管如此,长期语法元素结构允许关闭该保证,从而允许在其他应用情景或者使高端视频解码器基于短期仅根据短期语法元素执行并行处理调度,即,不开发关于相互依赖层的空间区段的边界之间的相对定位的任何保证。出于决定机会主义解码之目的还可使用长期语法元素。

[0008] 本申请的另一方面关于可伸缩性编码,根据其相互依赖层的图片结合相互依赖层的并行处理利用自基础层至增强层的上采样而经历层间预测。具体地,该方面关于用于执行自基础层至增强层的上采样的内插法。自然,该内插法可使基础层图片的相邻分区变得彼此依赖。即,内插法使上采样的基础层参考图片的各部分的外圆周处的内插结果与基础层图片的共置分区内的两个像素/像元以及相邻分区的像素/像元相互依赖。换言之,在被“涂黑”或者拓宽的增强层图片中预测到用作共置部分的层间预测的参考的基础层图片的区域。顺便提及,通过层间预测的内插法而导致的相互依赖性对相互依赖层的并行处理时可实现的并行程度产生不利影响。例如,根据本发明的第二方面,引入了语法元素,语法元素通知解码器:沿着基础层的分区修改内插以不混合基础层图片的相邻分区的像素/像元,基础层图片的分区及其上采样型式取决于增强层图片的空间区段或者基础层与增强层的空间区段。通过引入该语法元素,编码器能够在两种模式之间切换:如果内插法受限于使基础层图片的部分自包含,即,启动限制,则随着沿着基础层图片的分区边缘的内插质量轻微地降低,在相互依赖层的并行解码时将最大程度地增加所获得的并行程度,然而,在不限制内插法的情况下,并行程度降低且分区边缘处的内插质量增加。

[0009] 本申请的第三方面关于具有相互依赖层的并行解码的可伸缩视频编码并且通过引起长期语法元素结构试图减轻解码器执行并行处理调度的负担,即,将空间区段分布到并行处理线程中,长期语法元素结构允许解码器在比该时间间隔更大的预定时间段内确定层间偏移量或者层间延迟,在该时间间隔内,短期语法元素发出相互依赖层的图片的空间区段的大小和定位、以及这些图片的空间采样分辨率的信号。通过引入发出层间偏移量的信令的上期语法元素,视频编码器能够在两种模式之间切换:根据第一模式,编码器向解码器保证了与相互依赖层的解码之间的特定并行程度对应的特定的层间偏移量并且在该预

定时间段内相应地设置短期语法元素,以使得实际层间偏移量等于保证的一或者甚至更低。根据另一模式,不向解码器提供该保证,因此,编码器无需设置短期语法元素以满足另一白哦准,诸如,可选地,将短期语法元素适配成该预定时间段内的视频内容。因此,当服从整个预定时间段内的该计数时,在时间上共置对其的增强层图片的第一空间区段开始不面临任何冲突时,至少相对于增强层图片的第一空间区段该预定时间段内的解码,在数据流中明确发出信号的层间偏移量可以是不得不被解码的最小化的基础层空间区段的计数。

[0010] 本申请的第四方面关于可伸缩的视频编码和多层视频数据流所属的各个NAL单元的各层的信令、以及可伸缩空间内的这些层的定位、以及跨可伸缩空间的可伸缩维度的含义。为了易于通过涉及传输多层视频数据流的中间网络实体执行任务,该信息对于这些中间网络实体易于访问并且能够被容易地管理。本申请的第四方面基于发明人的法线,根据本发明,典型的应用情景证明类型指示符字段的支出,类型指示符字段概念NAL单元报头内的与各个层相关联的层指示符字段的方式,类型指示符字段被解译为:如果类型指示符字段具有第一状态,则常见信息NAL单元中的映射信息将层指示符字段的可能值映射至操作点,并且使用相应的层指示符字段和映射信息将层NAL单元与操作点相关联。同样,可从不同程度上适配各层与可伸缩的星座之间的映射并且允许实现多数可伸缩空间,然而,使管理总支出的缺点增加。如果类型指示符字段具有第二状态,则将层指示符字段分割成一个以上部分,并且通过使用这些部分的值作为可伸缩空间内的矢量的坐标而定位与相应NAL单元相关联的操作点。通过这种措施,各层与可伸缩星座之间的映射允许实现更少数目的可伸缩空间,然而,具有使网络实体的管理总支出减少的优点。在两种情况下,层指示符字段可以相同,与应用请求无关,但是,层指示符字段通过可伸缩空间引导层的NAL单元的方式可被适配于当前应用及其具体细节。自适应的优点过度补偿了类型指示符字段的额外支出的必要性。

[0011] 本申请的第五方面关于多层视频编码,即,可伸缩的视频编码,从而允许不同层使用不同的编解码器/标准。允许连续层使用不同的编解码器/标准的可能性能够支持已经存在的视频环境的延迟扩展,从而解决通过进一步的增强层而随后扩展的多层视频数据流并且鉴于此使用新的和潜在地更佳编解码器/标准。不能够理解一些增强层的编解码器/标准的网络汇集仍能够处理低层并且通过传输层解码器回馈多编解码的解码器,对于多层视频数据流的每个NAL单元,传输层解码器识别与其相关联的相同编解码器并且相应地将多层视频数据流的NAL单元移交至多标准多层解码器

[0012] 本申请的第六方面涉及多层视频编码,多层视频编码将基础层的图片和增强层的图片均细分成块的阵列。在这种情况下,通过将语法元素结构插入到多层视频数据流中可有效地发出层间偏移量的信号,语法元素结构指示用于将以基础层块为单位的基础层和增强层的图片并行解码的层间偏移量。即,本申请的第六方面基于明确传输语法元素结构的发现,语法元素结构指示以基础层块的为单位的基础层与增强层之间的层间偏移量通过镜射方式使仅被传输的数据增加,其中,与解码器通过其他方式推导用于基于其他语法元素将基础层与增强层的图片并行解码的层间偏移量的情况相比较,解码器的计算复杂度大大地减少,例如,其他语法元素显示基础层块和增强层块的块大小、基础层图片与增强层图片的采样分辨率等。当将语法元素结构实现为长期语法元素结构时,第六方面与第三方面密切相关,因此,层间偏移量向解码器指示了保证:该保证同样适用于比时间间隔更大的预定

时间段,在该时间间隔内,多层视频数据流中的短期语法元素指示另行称为必要的这些暗示,以通过相对复杂的方式将这些语法元素组合而确定层间偏移量。

[0013] 自然,可以将所有的上述方面以它们中的两个、三个、四个或者其全部的方式进行组合。

## 附图说明

[0014] 下面将参考图描述本申请的优选实施方式,在附图中:

[0015] 图1示出了用作用于实现参考下列图进一步概括的多层编码器中的任一个的示意性实施例的视频编码器;

[0016] 图2示出了显示安装至图1中的视频编码器的视频解码器的示意性框图;

[0017] 图3示出了被细分成用于WPP处理的子流的图片的示意图;

[0018] 图4示意性地示出了根据实施方式的视频解码器,采用了根据本实施方式的基础层和增强层的空间区段的层间对齐以减少解码处理;

[0019] 图5示出了将图片分别细分成代码块和片的示意图,且片由整数倍数的代码块构成并且代码块之间限定的解码顺序遵循将图片细分成片;

[0020] 图6示出了用于实现图4中的实施方式的语法实施例;

[0021] 图7示出了将一对基础层和增强层图片不同地细分成片的示意图;

[0022] 图8示出了结合图4中的实施方式适用的另一示例性语法;

[0023] 图9示出了图片的示意图,其被细分成片和内插滤波器的应用程序,以出于层间预测的缘故执行上采样;

[0024] 图10示出了多层解码器的示意性框图,该多层解码器被配置为响应多层数据流中的语法元素,以接通或者断开上采样内插分离;

[0025] 图11示出了使用自基础层至增强层的层间预测的一对基础层和增强层的示意图,从而利用上采样从基础层采样分辨率转换成增加的增强层采样分辨率;

[0026] 图12示出了显示根据图10的可切换的上采样内插分离的示意图;

[0027] 图13示出了基础层的图片与增强层的图片重叠的示意图,且基础层的图片与增强层的图片被细分成WPP子流;

[0028] 图14示出了用于实现图10中的实施方式的示例性语法;

[0029] 图15A示出了基础层的图片与增强层的图片在空间中重叠对齐的示意图,且基础层的图片与增强层的图片被细分成不同的空间区段;

[0030] 图15B示出了图15A中的基础层图片与增强层图片重叠的示意图,但是,示出了用于选择沿着其执行上采样内插分离的分区的另一种可能性;

[0031] 图16示出了根据实施方式的视频解码器的示意性框图,根据本实施方式的视频解码器响应多层视频数据流内的长期语法元素结构,以从其中的解码推导出或者不推导出关于基础层与增强层之间的层间偏移量的保证;

[0032] 图17A示出了一对基础层和增强层图片的示意图,且该对基础层和增强层图片被细分成片,以示出通过根据实施方式的图16中的长期语法元素结构而传输的层间偏移量信号通知;

[0033] 图17B示出了一对基础层和增强层图片的示意图,且该对基础层和增强层图片被

细分成用于WPP处理的子流,以对用于实现根据实施方式的图16中的长期语法元素结构的实施例进行说明;

[0034] 图17C示出了一对基础层图片和增强层图片的示意图,且该对基础层和增强层图片被细分成切片(slice),以对用于实现根据甚至进一步实施方式的图16中的长期语法元素结构的实施例进行说明;

[0035] 图18示出了被细分成用于WPP处理的子流的图片的示意图,且额外地指示了使用根据实施方式的WPP对图片进行并行解码/编码时的波前结果;

[0036] 图19示出了根据本申请的实施方式的显示最小层间解码偏移量与块大小之间的关系以及基础层与增强层之间的采样分辨率比的图表;

[0037] 图20示出了用于实现根据图16的长期语法元素结构信令的示例性语法;

[0038] 图21示出了用于实现图16中的实施方式的语法的另一实施例;

[0039] 图22示出了根据HEVC类配置的实施例的NAL单元报头的语法;

[0040] 图23示出了根据实施方式的网络实体的示意性框图,通过允许在不同的层指示符字段内插之间切换可减少根据本实施方式的可伸缩编码;

[0041] 图24示出了显示响应类型指示符字段而切换的方式的示意图;

[0042] 图25进一步详细地示出了显示根据实施方式的可切换层指示符字段内插的示意图;

[0043] 图26示出了用于实现根据图23的可切换层指示符字段内插的示出性语法;

[0044] 图27示出了显示与图26中的语法有关的层指示符字段切换的示意图;

[0045] 图28示出了被配置为简化除基础层编解码器之外的编解码器的丢弃增强层NAL单元的传输流解码器的框图;

[0046] 图29示出了与单一标准多层解码器连接的传输流解码器的框图,示出了根据实施方式的传输流解码器的行为;

[0047] 图30示出了与多标准多层解码器连接的传输流解码器以及根据实施方式的传输流解码器的行为;

[0048] 图31示出了用于实现根据进一步实施方式的可切换层指示符字段解译的语法的另一实施例;

[0049] 图32示出了显示被细分成块的任一层的图片的示意图,块指示将图片进一步细分成空间区段;

[0050] 图33示出了被细分成块和片的任一层的图片的示意图;

[0051] 图34示出了被细分成块和子流的图片的示意图;

[0052] 图35示出了根据实施方式的视频解码器的示意性框图,根据本实施方式的视频解码器被配置为使用数据流中的语法元素结构以推导出具有其中限定的光栅扫描顺序的块的单元中的预测间处理图片之间的层间偏移量;

[0053] 图36示出了显示图34中的与根据实施方式的数据流内的语法元素结构有关的视频解码器的操作的可能模式的示意图;

[0054] 图37示出了显示结合进一步实施方式的图35中的视频解码器的操作的模式的示意图,根据进一步实施方式的层间偏移量信令可在不同的显式信令类型之间切换,即,不同类型的单元中的信令;

[0055] 图38示出了根据甚至进一步实施方式的图35中的视频解码器的操作的模式的示意图,在基础层和增强层图片的并行解码过程中连续测量根据甚至进一步实施方式的层间偏移量;

[0056] 图39示出了一方面根据光栅扫描解码顺序的图片的特定块的等级(rank)与另一方面根据实施方式的行指数和列指数之间的关系;

[0057] 图40示出了用于将基础层图片和增强层图片相对规则地细分成块的不同实施例以及由这些不同的实施例产生的结果;

[0058] 图41示出了用于实现图35至图40中的任一种实施方式的语法的实施例;

[0059] 图42示出了可替代图41中的语法的另一语法实施例;

[0060] 图43示出了语法实施例,根据该语法实施例的图16和图35中的实施方式可在多层数据流的另一部分中发送信号;并且

[0061] 图44示出了与根据实施方式的多层多标准解码器连接的传输层解码器的示意性框图。

## 具体实施方式

[0062] 首先,作为概述,提出了一种关于编码器/解码器结构的实施例,该编码器/解码器结构适于任一后续提出的概念。

[0063] 图1示出了根据实施方式的编码器的整体结构。编码器10可被实现为能够以多线程方式或者不仅仅是单线程方式进行操作。即,编码器10例如可以被实现为使用多个CPU内核。换言之,编码器10可支持并行处理,但是编码器10不需要一定支持并行处理。通过单线程编码器/解码器还可产生/解码已产生的位流。然而,本申请的编码概念能够支持并行处理编码器,从而有效地应用并行处理,然而不危及压缩效率。关于并行处理能力,对于后面参考图2描述的解码器,相似声明是有效的。

[0064] 编码器10是视频编码器,但是通常编码器10还可以是图片编码器。视频14的图片12被示出为在输入16处输入编码器10。图片12示出了特定情境,即,图片内容。然而,同时,编码器10在其输入16处还即时接收与属于不同层的图片12和图片15两者有关的另一图片15。仅出于示出性之目的,示出了属于层0的图片12而图片15属于层1。图1示出了层1相对于层0可涉及更高的空间分辨率,即,可以更高数目的图片采样显示相同的情景,但是,这仅用于示出性之目的,并且可替代地,层1的图片15可具有相同的空间分辨率,且可以是不同的,例如,在相对于层0的视野方向上可以具有不同的空间分辨率,即,从不同的视点可以捕捉到图片12和图片15。

[0065] 编码器10是混合式训练编码器(breedencoder),即,通过预测器18预测图片12和图片15并且使通过残差确定器22所获得的预测残差20经历诸如DCT等光谱分析等转换以及转换/量化模块24的量化。使由此获得的转换和量化预测残差26在经过熵编码器28中经过熵编码,诸如,算术编码或者使用例如上下文自适应性的可变长度。解码器可以使用残差的重构型式(reconstructibleversion),即,通过再转换/再量化模块31恢复去量化和再转换残差信号30并且通过组合器33将去量化和再转换残差信号30与预测器18的预测信号32重新组合,从而分别产生图片12和图片15的重构34。然而,编码器10基于块操作。因此,重构信号34在块边界线处遭遇不连续性,并且因此,可以将滤波器36应用于重构信号34,以分别产

生图片12和图片15的参考图片38，预测器18基于参考图片18顺次预测不同层的编码图片。然而，如图1中的虚线所示，预测器18还可在诸如空间预测模式等其他预测模式中直接采用没有滤波器36或者中间型式的重构信号34。

[0066] 预测器18可选择不同的预测模式以预测图片12的特定块。图1中示例性地示出了图片12的一个这样的块39。可以存在时间预测模式，根据时间预测模式，将图片12分区成代表图片12的任一块的块39，基于同一层的诸如图片12'等之前编码的图片预测块39。还可存在空间预测模式，根据空间预测模式，基于同一图片12的之前编码部分（相邻块39）预测块39。图1中还示出性地示出了图片15的块41，用于代表将图片15分区的任一其他块。对于块41，预测器18可支持刚刚讨论的预测模式，即，时间预测模式和空间预测模式。此外，预测器18可提供一种层间预测模式，根据层间预测模式，基于低层的图片12的对应部分来预测块41。“对应部分”中的“对应的”应表示空间对应性，即，在图片15中预测到与块41显示相同情景部分的图片12内的部分。

[0067] 当然，预测器18的预测可不限于图片采样。预测也可应用于任何编码参数，即，预测模式、时间预测的运动矢量、多视点预测的视差矢量等。则仅可在位流40中编码残差。

[0068] 使用特定语法来编译量化残差数据26，转换系数级和其他残差数据以及编码参数，其中包括例如由预测器18确定的用于图片12的单独块39和用于图片15的单独块41的预测模式和预测参数，并且使语法元素经过熵编码器28的熵编码。通过熵编码器28所输出的、由此获得的数据流40形成由编码器10所输出的位流40。

[0069] 图2示出了适配于图1中的编码器的解码器，即，能够将位流40解码。图2中的解码器整体由参考符号50表示并且包括熵解码器、转换/去量化模块54、组合器56、滤波器58以及预测器60。熵解码器52接收位流并且执行熵解码以恢复残差数据62和编码参数64。再转换/去量化模块54使残差数据62去量化和再转换并且将由此获得的残差信号转发至组合器56。组合器56还从预测器60接收预测信号66，组合器56进而基于通过组合预测信号66和残差信号65的组合器56而确定的重构信号68来使用编码参数64形成预测信号66。该预测对最终通过预测器18所选择的预测进行镜像，即，可以使用相同的预测模式，并且针对图片12和图片15的单独块选择这些模式并且根据预测参数进行导入。如上面参考图1已经说明的，可替代地或者此外，预测器60可使用重构信号68的滤波型式或者其某种中间型式。同样，可以根据组合信号68的未滤波型式或者其某种滤波型式确定最终在解码器50的输出70处再现和输出的不同层的图片。

[0070] 根据片概念，分别将图片12细分成片80并且将图片15细分成片82，并且至少片80内的块39的预测和片82内的块41的预测作为空间预测的基础分别受限于仅使用分别与相同图片12、15的相同片有关的数据。即，块39的空间预测限于使用相同片的之前编码部分，但是，时间预测模式不受限于依赖诸如图片12'等之前编码图片的信息。同样，块41的空间预测模式受限于仅使用相同片的之前编码的数据，但是时间预测模式和层间预测模式不受限制。仅出于示出性之目的，选择了分别将图片15和图片12细分成六个片。可以分别针对图片12'、12和15、15'单独在位流40内选择细分成的片并且发送信号。每个图片12的片的数目和每个图片15的片的数目分别可以是一个、两个、三个、四个、六个等中的任一个，其中，片分区可受限于仅规则地分区成片的行和列。出于完整性之缘故，应注意，将片独立编码的方式可不受限于帧内预测或者空间预测，而是，还可包括跨片边界线的编码参数的任何预测，

并且熵编码中的上下文选择还可受限于仅取决于相同片的数据。因此,解码器能够并行地(即,以片为单位)执行刚刚提及的操作。

[0071] 可替代地或者额外地,图1和图2中的编码器和解码器能够使用WPP概念。参见图3, WPP子流100还代表了将图片12和图片15在空间上分区成WPP子流。与片和切片相反,WPP子流并不对跨WPP子流100的预测和上下文选择施加限制。WPP子流100逐行进行扩展,诸如,跨LCU(最大编码单元)101的行等,即,可在位流中单独传输的预测编码模式的最大可能块,并且为了能够支持并行处理,仅仅做出了与熵编码有关的一种妥协。具体地,沿着示例性地从顶部引导至底部的WPP子流100来定义顺序102,并且对于每个WPP子流100,除了顺序102中的第一WPP子流之外未完全重置对符号阿拉伯字母的概率评估,即,熵概率,而是采用或者设置为等于熵编码/解码之后产生的概率。前一WPP子流直接到达其第二LCU之前,如线104所示,对于位于图片12和图片15的相同侧的每个WPP子流,按照LCU顺序或者子流的解码器顺序分别从诸如由箭头106所指示的左手边开始,并且在LCU行方向上引导至另一侧。因此,通过分别服从相同图片12和相同图片15的WPP子流的序列之间的某种编码延迟,可以将这些WPP子流100并行解码/编码,以使得将相应图片12、15并行(即,同时)编码/解码的部分形成某种波前108,波前108在图片上以自左至右的倾斜方式移动。

[0072] 简要指出,顺序102还限定LCU之中的光栅扫描顺序,从而从顶部至底部逐行从顶部左侧LCU 101引导至底部右侧LCU。每个WPP子流均可对应于一个LCU行。简要地,返回参考片,后者还可受限于与LCU边界对齐。只要涉及子流内部的两个切片之间的边界,则不受约束于LCU边界,可将子流分段成一个或者多个切片。然而,在从子流的一个切片转换成子流的下一切片的情况下,采用熵概率。如果是片,则可将全部片汇总成一个切片,或者可将一个片分段成一个或者多个切片,只要涉及片内部的两个切片之间的边界,则不需要再次受约束于LCU边界。如果是片,则改变LCU之间的顺序,首先,在按照片顺序进行至下一片之前,按照光栅扫描顺序中的片顺序横断片。

[0073] 如至目前所述,可将图片12分区成片或者WPP子流,同样,也可将图片15分区成片或者WPP子流。理论上,可以选择WPP子流分区/概念用于图片12和图片15中的一个,而选择片分区/概念用于这两个图片中的另一个。可替代地,可以对位流施加限制,根据该限制,各层的概念类型(即,片或者WPP子流)必须相同。关于空间区段的另一实施例包括切片。出于传输之目的,切片用于将位流40分割成区段。切片被打包置于NAL单元中,NAL单元是用于传输的最小实体。每个切片均可独立地编码/解码。即,正如上下文选择等,禁止跨切片边界线的任何预测。总而言之,这些是空间区段的三种实施例:切片、片以及WPP子流。此外,可以组合使用所有这三种平行概念:片、WPP子流以及切片,即,可以将图片12或者图片15分割成片,其中,每个片被分割成多个WPP子流。此外,切片可用于将位流分区成多个NAL单元,例如但不限于在片或者WPPP边界线处被分区。如果使用片或者WPP子流并且额外地使用切片将图片12、图片15分区并且切片分区与另一WPP/片分区存在偏差,则应将空间区段限定为图片12、15的最小独立可解码段。可替代地,可以利用图片(12或者15)内的概念组合和/或如果必须使不同使用的概念之间的边界对齐,则可对位流施加限制。

[0074] 在讨论本申请的上述所述概念之前,再次参考图1和图2,应注意,图1和图2中的编码器和解码器的框结构仅用于示出性之目的并且结构还可以不同。

[0075] 根据被称为“片边界线的对齐”的第一方面,长期语法元素结构用于发送信号,以

保证在诸如一系列图片扩展的时间段等预定时间段内,第二层的图片15被细分成使得第二层的图片的空间区段82之间的边界84与第一层的空间区段80之间的每个边界86重叠。在比预定时间段更短的时间间隔(诸如,以各个图片为单位等)内,即,在图片间距间隔内,解码器仍基于多层视频数据流40的短期语法元素周期性地确定将第一层的图片12实际细分成空间区段80并且将第二层的图片15实际细分成空间区段82,但是,关于对齐的知识早已帮助规划并行处理工作负荷分配。例如,图1中的实线84代表了其中片边界线84在空间上与层0的片边界线86完全对齐的实施例。然而,上述保证还允许层1的片分区比层0的片分区更细,以使得层1的片分区进一步包括在空间上与层0的片边界线86中的任一个不重叠的额外的片边界线。在任何情况下,关于层1与层0之间的片注册的知识有助于解码器在同时并行处理的空间区段之间分配可用的工作负荷或者处理功率。没有长期语法元素结构,解码器将不得不在更短的时间间隔内执行工作负荷分配,即,每个图片,从而浪费执行工作负荷分配的计算机功率。另一方面是“机会主义解码”:具有多个CPU内核的解码器可利用关于层的并行性的知识来决定尝试或者不尝试将更高复杂性的层(即,更高空间分辨率或者更高数目的层)解码。通过利用同一解码器的所有内核可以将超过单核的能力的位流解码。如果配置文件和层次指示符(level indicator)不包含关于最小并行性的指示,则该信息特别有用。

[0076] 为了更清晰地理解本申请中刚刚概括的方面,参考图4,图4示出了参考图2的规定可以实现的视频解码器600。如上所述,即,解码器600被配置为将多层视频数据流解码成情景,该情景利用如上所述的从第一层0至第二层1的层间预测被编码在层的层级中。视频解码器支持以该层的图片被细分而成的空间区段来对多层视频数据流并行解码,例如,空间区段是片、WPP子流等。换言之,视频解码器能够对多层视频数据流并行解码,如此,视频解码器600以空间区段为单位对层0的图片12和层1和图片15进行操作。

[0077] 如上面所概括的,例如,空间区段可以是片,而视频解码器600被配置为使用帧内图片空间预测将层0的图片12和层1的图片15解码,视频解码器600使对于每个片的帧内图片空间预测在其片边界线处中断。诸如,个别地,对于与每个图片12和15有关的时间帧604,即,对于属于特定时间帧604的每对图片12和15,通过基于短期的短期语法元素方式在数据流40内发送信号:将图片12和图片15细分成片,诸如,以时间间隔为单位。如上所述,即,将图片12和图片15细分成片可受限于仅矩形规则地细分成片的行和列。因此,短期语法元素602将单独设置用于两个层的每个图片12和每个图片15的片细分的行数和列数。在对入站(inbound)的多层视频数据流40进行解码时,视频解码器600被配置为应用空间预测,潜在地时间预测。可选地,视频解码器600对每个片单独进行熵解码。如果在对每个片进行解码的过程中使用概率自适应性,则视频解码器600对每个片的熵概率单独进行初始化,以使得对片进行并行熵解码。除空间预测和可选地时间预测之外,只要涉及层1的图片15的片的解码,则视频解码器600支持层间预测。如上所述,层间预测可涉及解码层1中所涉及的不同参数:层间预测可预测层1的预测残差,诸如,转换系数、解码层1中使用的预测模式、解码层1中使用的预测参数、增强层1和图片的采样等。例如,在层0和层1涉及同一情景的不同视野的情况下,通过控制层间预测的视差矢量预测参数,层间预测基于层0的图片12的已解码部分(直接(在中心处)共置部分或者在空间上略微偏离于直接共置位置的部分)预测层1的图片15的片内的部分。

[0078] 如使用图4中的参考符号606所示,视频解码器600响应数据流40的长期语法元素结构,以不同程度地处理紧随长期语法元素结构606之后的预定时间段608,预定时间段608包括若干个时间间隔,即,多个时间帧604,在多个时间帧604内,短期语法元素602单独发送信号将图片细分成片。应注意,608可涉及以任何方式引起严重重新初始化的SPS的范围(=时间段)和SPS变化。只要此处提及长期特征,则对于与其他方面有关的所有实施方式,刚刚提及的注释也有效。具体地,如果长期语法元素结构606假设了第一可能值集合之外的值,则视频解码器600将该情形解译为保证:在预定时间段内,层1的图片15被细分成使得图片15的片之间的边界与层0的图片12的片的每个边界重叠。在这种情况下,视频解码器600仍对短期语法元素602进行检测,以在预定时间段608的时间间隔602内确定将图片12和图片15细分成其片,但是,通过比较时间对齐的每对图片12与图片15,视频解码器600可依赖于事实并且认识到图片12的基础层片的边界与图片15的增强层片的边界完全重叠,即,图片15的片细分局部对应于或者代表了将图片12细分成片的空间再细化。如上所述,通过相应地并行调度图片12的片和图片15的片在预定时间段608内的并行处理,即,将时间上对齐的一对图片12和图片15并行解码,视频解码器600可利用该信令的优势,即,长期语法元素结构606假设了第一可能值集合之外的值。例如,如果长期语法元素结构假设了第一可能值集合之外的值,则视频解码器600可获知如下:对于层0的特定图片12,按照照片顺序位于图片12的片之中第一片与时间对齐的增强层图片15的相应片局部一致,或者与按照照片顺序的增强层图片15的片之中的时间对齐增强层图片15的第一片完全局部重叠。因此,至少在没有视差/运动补偿的层间预测的情况下,因为刚刚提及的保证向视频解码器600指示:层间预测所需的基础层图片12的共置部分可用于增强层图片15的整个第一片,所以视频解码器600可在完成时间对齐的基础层图片12的第一片的解码时尽快开始对增强层图片15的第一片进行解码。因此,通过视频解码器600可以认识/确定基础层图片12与增强层图片15之间的层间偏移量或者并行性程度等于基础层图片12的一个片。如果层间预测包含具有非零垂直分量的视差矢量和/或具有水平分量的视差矢量,水平分量将基础层图片内的对应部分朝向右侧移位,可略微增加偏移量,其中,可按照自图片12、15的左上角至其右下角的行式光栅扫描顺序引导片之间的片顺序。

[0079] 然而,如果长期语法元素结构假设了第二可能值集合之外的值,第二可能值集合明显不同于第一可能值集合,视频解码器600并不利用任何保证的优势,而是基于使用短期语法元素602的短期计划和调度图片12和图片15的片的并行解码,且潜在地,针对时间对齐配对的图片12和图片15中的至少一些,计划和调度基础层和增强层的片的并行解码。然而,在这种情况下,视频解码器600基于短期(麻烦的程序)在层0与层1之间的并行解码中确定最小层间偏移量或者层间空间处理偏移量,即,层0与层1之间的并行性程度。至少对于短期语法元素的可能值的集合的子集,第二层的图片的空间区段之间存在不与第一层的空间区段的边界中的任一个重叠的边界。但是,根据短期语法元素的可能值的集合的进一步子集,第二层的图片的空间区段之间存在与第一层的空间区段的每个边界重叠的边界。如果长期语法元素表示基础层与增强层之间的片边界线对齐,则仅使用后一子集。

[0080] 此外或者可替代地,视频解码器600可使用或者采用下列事实:长期语法元素结构假设了第一可能值集合之外的值,以执行试验(即,尝试执行)将层1完全解码,且如果长期语法元素结构606假设了第二可能值集合之外的值,则抑制执行该试验。在这种情况下,尤

其对于电池驱动设备,在及时(即,实时)对增强层1进行解码的结果或者成功率具有投机性的情况下,则节省了宝贵的计算功率。值得一提的是,还可基于下面参考第四方面体积的层次指示符选择抑制(refraining)。

[0081] 尽管上面使用片作为空间区段示例性地描述了图4,然而,清晰的是,视频解码器600可结合诸如子流或者切片等其他空间区段利用长期语法元素结构和由此发出的保证等优势。在前者的情况下,视频解码器600将使用帧内图片空间预测将层的图片12和图片15解码,且将第一层12的图片的空间区段并行解码,并且支持跨第一层的图片的空间区段的边界线的帧内图片空间预测,并且服从这些空间区段(即,子流)的解码之间的解码延迟。如上所述,子流可对应于相应图片的水平条带,即,垂直地细分相应的图片。在对每个子流进行解码时,视频解码器600可使用通常自左引导至右的解码顺序,并且图片的子流之间限定的解码顺序可从顶部引导至底部。根据使用的典型空间预测概念,从当前图片的顶部相邻的已解码部分和左手边已解码部分执行空间预测,服从紧随其后的后续子流之间的特定解码延迟,由此,允许子流的并行解码。例如,可以在LCU的单元内测量解码延迟。可以在层0的图片12以及层1的图片15中完成此操作。因此,在将视频数据流解码时的并行性可个别地包含图片12和图片15内的两种并行性,但是,还可包含属于一个时间帧604的不同层的图片12和图片15的子流的并行解码。只要涉及子流的可选熵解码,则并行性可包含相应子流的解码过程中的熵概率的自适应性。可使每个图片12或者图片15的子流顺序中第一子流经过熵概率的单独初始化,而与其他子流无关。根据同一图片的解码顺序中的紧随的前一子流的即时自适配熵概率,诸如,通过采用在将到达距相应前一子流的左手侧的特定距离的紧随前一子流解码过程中所采用的熵概率,诸如,在将紧随的前一子流的两个LCU解码之后等,可以使任一后续的子流经过熵概率初始化。

[0082] 即使在波前并行处理子流的情况下,视频解码器600也能够利用长期语法元素结构606的优势:如果通过该语法元素结构606发出保证信号,则视频解码器600可依赖于下列事实:基础层图片12的连续/相邻子流之间的所有边界在预定时间段608内与时间对齐的增强层图片15的相邻/连续子流之间的相应边界重叠。即,基础层子流与时间对齐的增强层图片15的相应增强层子流局部一致,或者基础层子流精确地对应于时间对齐的增强层图片的两个或者多个子流。因此,如果保证适用,则解码器600获知在完成基础层图片12的第一子流的解码之后,可尽快开始时间对齐的增强层图片15的第一子流的解码。

[0083] 如上所述,不同于片细分,短期语法元素602可被选择成使得短期语法元素602限定与将这些图片细分成诸如LCU等一些编码块有关的图片12和图片15中的子流的位置。因此,子流可以是一行或者多行编码块的统称。如同片细分的情况,时间间隔604可以是如下:短期语法元素602基于每个图片(即,基于每个图片帧604)发出将图片12和图片15细分成子流的信号。如果长期语法元素结构606不提供该保证,则然而,视频解码器600可尝试将共同事件帧的不同层的子流并行解码,但是,为了实现此操作,视频解码器600需要对短期语法元素602进行检查。

[0084] 如使用片作为空间区段的情况,视频解码器600可根据通过长期语法元素结构606假设的值进行将增强层1解码的投机性试验。

[0085] 应注意,如在图1中所示,只要涉及视频编码器方面,则可以实现将长期语法元素结构606插入和设置到数据流40中并且决定是否授予对解码器600保证的相应视频编码器。

如果授予,则在对应的预定时间段608内将短期语法元素602设置成服从边界对齐保证的设置时,编码概率限制。如果不授予,则编码器保持在时间段608内按照其习惯设置短期语法元素602的自由。当使用片作为空间区段时,编码器服从下列限制:空间预测不能跨边界线,并且按照自包含方式对每个片执行图片12和图片15的片的可选熵编码。例如,对于与其他片不相互依赖的片(对于每个片)将熵概率重新初始化。如果是子流,则对相应图片12、15的任一第一子流重新执行子流的熵概率初始化,即,与任何其他子流无关,并且只要涉及任一第二子流和随后的子流,则在适应达到紧随的前一子流的中间位置时采用熵概率。在无关子流的交叉的任何限制的情况下,执行空间预测。

[0086] 通过参考图4描述的方式,即,通过下面描述的方式,可以将对齐概念引入当前设想扩展的HEVC标准。至今,下面即将提出的描述还应被解释为基于有关上面参考图4提出的描述的可能实现细节。

[0087] HEVC允许经垂直边界线和水平边界线的栅格将编码基础层图片的CTB划分成被称之为片的矩形区域并且可被单独处理(但环内滤波除外)。在片边界线处,可以关闭环内滤波,以使其完全不依赖。

[0088] 如同在图片边界线处,解析与预测相互依赖性在片边界线处断裂,其中,如果相应地配置,则环内滤波器可跨边界线,以减少边界线伪差。因此,个别片的处理并不完全依赖于图片内的其他片或者取决于滤波配置的广泛范围。安装的限制在于:片的所有CTB应属于同一切片或者切片的所有CTB应属于同一片。从图1中可以看出,片迫使CTB扫描顺序遵从片顺序,即,在继续属于第二片(例如,右上片)的CTB之前,遍历属于第一片(例如,左上片)的所有CTB。通过CTB的数目和大小将结构限定在构成图片内的栅格的每个片行和每个片列内。结构可以基于每个帧而改变或者在编码视频序列内保持不变。

[0089] 图5示出了将图片内的CTB示例性地划分成九个片。粗黑线代表边界线并且编号代表CTB的扫描顺序,也反映了片顺序。

[0090] 只要将覆盖基础层位流中的对应图像区域的所有片解码,则可将HEVC扩展的增强层片解码。

[0091] 下面部分描述了允许使用图4中的概念更容易地访问基础层信息的约束条件、发信号和解码过程变形。

[0092] 有关片层次并行性的最简单情况是基础层和增强层中的片边界线对齐。对于SNR可伸缩性,指边界线精确地位于同一位置。对于空间可伸缩性,指对于属于同一片的每两个增强层像元,对应的基础层像元也属于同一片,并且对于属于同一片的每两个基础层像元,对应的增强层像元也属于同一片。

[0093] HEVC的特征在于,与图4中的602对应的短期信令通过使用下面给出的图片参数集合:来自[1]的column\_width\_minus1[i]和row\_height\_minus1[i]基于每个图片指示图片内的片的维度和结构。图6示出了示例性的语法。

[0094] HEVC的进一步特征在于,保证有关HEVC编码视频序列的特定设置的限制信令,例如,指示单层HEVC编码视频序列中的固定片结构(下面给出的VUI语法中的cp.tiles\_fixed\_structure\_flag)。可伸缩编码视频序列中的片的进一步限制有益于解码器初始化和操作。为了允许解码器在完成基础层片之后,开始对与积极层片相关联的增强层图像区域进行解码,不需要强制实施完全对齐。尤其在空间可伸缩性方面,可能有助于允许增强层

而非基础层中的更多片解码。例如,就两种空间可伸缩性因素而言,与对应的基础层图像区域相比较,增强层图像区域中包含了四倍数目的像元。因此,对于每个基础层片,可能有助于允许增强层中的四个片解码。见图7,图7示出了具有空间可伸缩性的对齐片边界线的实施例。基础层和增强层中的所有垂直边界线皆对齐。在增强层中,使用额外的片(水平边界线),以允许在每个增强层中使用与基础层片分区时相同数目的像元的并行性。

[0095] 因此,我们通过下列方式限定片边界线对齐:仅每个基础层边界线具有增强层中的对应边界线,但是,不可能反过来。精确地,这就是指,对于属于同一片的每两个增强层像元,对应的基础层像元也属于同一片。

[0096] 信令606有助于将并行解码器环境初始化,否则,信令606将不得不通过对多个参数集合进行解析而采集信息。而且,例如,就位流限制形式而言,图4中的概念保证了所述限制对于完整的编码视频序列是有效的。

[0097] 如果基础层的片边界线是增强层的片边界线的子集,则可能实施方式允许将基础层片边界线的信令保存在增强层中。

[0098] 精确地,可以在解码器易于访问的位流中发出关于片对齐的信息的信号。

[0099] 在具体实施方式中,通过使用增强层SPS的VUI参数中的标记可以实现信令,如图8中给出的:

[0100] 当存在时,等于1的tiles\_fixed\_structure\_flag指示各个图片参数集合(即,编码视频序列中激活的)具有语法元素num\_tile\_columns\_minus1、num\_tile\_rows\_minus1、uniform\_spacing\_flag、column\_width\_minus1[i]、row\_height\_minus1[i]、以及loop\_filter\_across\_tiles\_enabled\_flag的相同值。等于0的tiles\_fixed\_structure\_flag指示不同的图片参数集合中的片语法元素可或者不可具有相同的值。当不存在tiles\_fixed\_structure\_flag语法元素时,推断其等于0。

[0101] 应注意,等于1的tiles\_fixed\_structure\_flag的信令向解码器保证了:编码视频序列中的每个图片均具有以相同方式分布了相同数目的片,在多线程解码的情况下,这可能有助于工作负载分配。

[0102] tile\_boundaries\_aligned\_flag对应于图4中的结构606。如果tile\_boundaries\_aligned\_flag等于1,则其指示对应的基础层图片的所有片边界线具有给定增强层中的对应片边界线。等于0的tile\_boundaries\_aligned\_flag指示对应的基础层与给定的增强层之间的片配置不存在限制。

[0103] 应注意,长期语法元素结构可保证,在预定时间段内,第二层的图片15被细分成的空间区段82的最小数目(例如,图片序列)是第一层的图片12被细分成的空间区段80的最小数目的n倍,或者,图片12的每个空间区段均精确由时间对齐的图片15的n个空间区段组成,且n取决于长期语法元素结构的值。在图7中的情况下,n等于3。解码器仍在比预定时间段更短的时间间隔内基于多层视频数据流40的短期语法元素周期性地确定将第一层的图片12实际细分成空间区段80并且第二层的图片15实际细分成空间区段82,但是,再次,解码器可以采用该保证来更为有效地执行工作负载分配。另一方面是“机会主义解码”:具有多个CPU内核的解码器可采用保证作为关于层的并行性的暗示,并且据此决定将更高复杂性的层(例如,更高空间分辨率或者更高数目的层)解码。通过利用同一解码器的所有内核可以将超过单核能力的位流解码。如果配置文件和层次指示符不包含关于最小并行性的指示,则

该信息特别有帮助。

[0104] 下文中讨论并且提出的第二方面关于被称为“受限制的层间上采样”的概念:如果是空间可伸缩性的多层视频,使用位流中的语法元素(例如,示例性地, `independent_tile_upsampling_idc`)来调整上采样滤波器36。如果在跨空间区段边界线86的层0中执行上采样滤波,则由于将上采样滤波组合,所以相对于层0的空间区段80的编码/解码,层1的空间区段82的并行解码/编码所满足的延迟增加,并且由此使得相互依赖,层0的相邻空间区段的信息用作层1的块41的层间预测所使用的预测参考38。例如,见图9,示出了重叠型式的图片12和图片15,且两个图片根据空间对应性注册彼此并且彼此的维度一致,即,显示相同情景部分的部分彼此重叠。示例性地示出了分别被分割成6个空间区段(诸如片)和12个空间区段的图片12和图片15。示出性地示出了滤波核200在图片12的左上片上移动,以获得其上采样型式,上采样型式用作层间预测图片15的片内的任一块(在空间上与左上片重叠)的基础。在诸如202等一些中间实例中,核200与图片12的相邻片重叠。因此,位于上采样型式的位置202的核200的中间的采样值取决于图片12的左上片的采样以及图片12的其右侧片的采样。如果图片12的上采样型式用作层间预测的基础,则层的区段的并行处理时的层间偏移量增加。因此,限制可有助于使不同层之间的并行量增加并且相应地使整体编码延迟减少。自然,语法元素还可以是对图片序列有效的长期语法元素。通过下列方式中的一种可以实现该限制:在重叠位置202处对核200的重叠部分进行滤波,例如,具有采样值位于核200的非虚线部分内的中心趋势;使用线性函数或者其他函数将非虚线部分外推成虚线部分等。

[0105] 为了使之后的方面更为清晰,参考图10,图10示出了接收对应于层0中的图片12和层1中的图片15的不同空间层中的空间可伸缩位流40(图片被编码成空间可伸缩位流40)的解码器610。至少对于这些空间层中的一个,解码器610被配置为将空间区段中的该空间层解码。与上述所述温度,这些空间区段可以是片或者位流或者切片。同样,解码器610可被配置为使图片12或者图片15的该空间区段经过并行解码。即,基础层图片12可被细分成诸如片和/或位流和/或切片等空间区段,和/或增强层图片15可被细分成片和/或位流和/或切片。只要涉及关于并行处理的细节,则参考上面图4中的描述,其已经被转移至图10中的解码器610。即,例如,在对基础层图片12进行解码时,如果基础层12和15是分层视频的一部分,则解码器610使用空间预测和可选地时间预测。如果是片,则空间预测受限于不能跨片边界线并且空间预测适用于对片单独完全执行的熵解码(如果使用熵解码)。空间预测适用于增强层图片15,并且额外支持层间预测。如上面已经描述的,层间预测不仅关于增强层的预测参数,基于在将基础层解码时所使用的对应预测参数来预测增强层的预测参数,而且还关于从基础层图片的重构采样推导出的预测,基础层图片相对于目前使用层间预测进行预测的增强层图片15的该部分位于共置部分。然而,由于位流40可以是空间上可伸缩的位流,所以通过解码器610可以对构成增强层图片15的当前处理部分的层间预测的基础的基础层图片12的任一共置部分进行上采样,以说明图片15相对于图片12的更高空间分辨率。例如,见图11,图11中使用了参考符号612表示增强层图片15的当前预测部分。使用参考符号614表示基础层图片12中的共置部分。由于增强层图片15的更高空间分辨率,所以部分612(以圆点表示)内的采样位置的数目被示出为比部分614(同样,以圆点表示)内的采样更高。因此,解码器610使用内插法对基础层图片12的部分614的已重构型式进行上采样。如

此,图10中的解码器610相应空间上可伸缩的位流40中的语法元素616。

[0106] 具体地,参考图12更为详细地说明了刚刚提及的响应性。图12示出了基础层图片12内的部分614(即,其非上采样型式的参考部分)以及使用参考符号618表示的其对应上采样型式。如刚刚提及的,例如,通过将部分618中的相应上采样的采样值复制到增强层15的部分612中,通过内插法620可以从基础层图片12获得后面用于层间预测的型式618。然而,内插法620取决于刚刚提及的语法元素616。内插法620根据语法元素616改变的方式分别与沿着基础层图片12的分区的相邻分区之间的边界的区域及其上采样型式有关。具体地,分区取决于图片12和图片15中的至少一个被细分成的上述所述空间区段。在图11中,使用虚线622示出了基础层图片12内的分区。例如,如下面更为详细概括的,分区622可对应于图片12和图片15的空间区段的边界的重叠的逻辑AND或者逻辑OR组合、或者在空间上与由增强层图片15的空间区段限定的分区一致。在任何情况下,解码器610根据语法元素616执行内插法620,与分区622无关或者考虑分区622。当考虑分区时,解码器610执行内插法620,以使得上采样部分618内的所有采样仅起源于、或者取决于、或者受分区622的分区中的一个的采样的影响,而与分区622的任何其他分区无关。例如,如果分区622是图片12和图片15的空间区段的边界的局部AND组合或者其局部OR组合,则指内插部分618的所有采样仅起源于基础层图片12的一个分区。然而,如果语法元素616引导解码器610对分区622不敏感,则可能发生内插部分618内的不同采样起源于分区622的相邻分区。

[0107] 例如,612示出了下列情况:使用滤波核200来执行内插法620,并且为了获得图12中被圈起的内插采样,核624与分区622的两个相邻分区之间的边界重叠。在这种情况下,解码器610响应语法元素616,以正常填充滤波核624,即,通过将滤波核624完全施加到基础层图片12的相应包含采样上,或者使用回退规则(fallbackrule)(如图12中使用的哈希所示)通过填充伸出至相邻分区中(分别不包括部分614和部分618)的滤波核624的片段,根据回退规则,填充与基础层图片12的基层采样不相互依赖的相应片段。例如,使用滤波核624的非哈希部分的采样值的某种平均测量或者通过某一外推法填充滤波核624的哈希部分。另一可替代地,使用诸如0等预定值填充与哈希部分重叠的基础层图片12的采样值。通常,解码器610可处理使包含部分614与相邻分区(如同图片12自身的外边缘)分离的部分的分区边界,并且可使用例如在图片12的外圆周附近或者出执行上采样/内插法时所使用的内插法620中的相同回退规则。

[0108] 根据本申请的一种实施方式,分区622被选择为使得与基础层图片细分成空间区段相一致,与增强层图片任意细分成空间区段无关。因此,因为诸如基础层图片12的部分614等部分在执行增强层图片15的部分612的层间预测之前不需要解码器610将相邻分区/空间区段解码,所以通过解码器610可使基础层图片12与增强层图片15之间的层间偏移量减少。

[0109] 可替代地,解码器610可被配置为确定分区622与图片15细分成空间区段局部一致。另一可替代地,解码器610可被配置为选择仅由图片12和图片15的空间区段的边界组成的分区622,分区622在空间上一致,即,以对应于图片12和图片15的边界的局部AND,换言之,仅图片15的这些边界被细分成与分区622的分区之间的边界对应的空间区段形式,即,在空间上对应于基础层图片12的相应边界被细分成空间区段。

[0110] 语法元素616可引导解码器610不仅忽视内插法620中的分区622,而且还对选择分

区622的不同方式进行区分,这也是可行的,并且下面将更为详细的概括。例如,见图9,其中,示例性地使用片作为空间区段。例如,如果语法元素616向解码器610发出单独执行用于分区622的内插法620的信号,因为分区622与增强层图片15被细化地细分成片一致,所以解码器610可使用基础层图片12的边界作为分区622的边界。因此,为了开始将增强层图片15的最高片行中的第二片解码,解码器610不需要等待完成对基础层图片12的最高行中的第二片的解码,因为“内插法分离”禁止任意混合基础层图片12的最高行中的前两个片的重构采样。如果采样将增强层图片15完全细分成片,为了确定分区622,还在图9中的虚线处执行内插法分离,并且解码器610能够甚至更早地开始解码增强层图片15的左上方的片,即,解码器610尽快管理基础层图片12的第一片的相应共置子部分的解码。鉴于此,应注意,甚至在对片进行解码时,解码器610可使用某一解码顺序,例如,该解码顺序包含按照行式方式自相应片的左上角引导至相应片的右下角的某种光栅扫描顺序。

[0111] 即,根据第二方面,编码器形成位流40能够通过语法元素616在两种模式之间选择:如果设置语法元素616并且将语法元素616插入到位流40中,则引导解码器610对分区622不敏感,由于更佳的内插法,可以实现更佳的层间预测,但是,在将图片12和图片15并行解码时可获得的并行程度降低,即,服从的最小层间偏移量增加。在另一模式下,语法元素616引导解码器610在执行内插法620时出于层间预测之目的而考虑分区622,并且因此,层间预测质量降低,有利于在将图片12和图片15并行解码时分别使并行程度增加并且使最小层间解码偏移量减少。

[0112] 尽管本申请的第二方面的描述主要集中于片细分或者片并行处理概念,然而,清晰的是,当使用WPP位流时,使用语法元素616来控制内插法620也是有利的。例如,见图13,图13示出了下列情况:基础层图片12被示例性地细分成两个位流,其中,增强层图片15的共置部分各自被细分成两个位流。如果响应语法元素616施加内插法分离,则解码器610可开始将增强层图片15的第一(即,最高)位流解码,一旦解码器610将基础层图片12的第一位流解码,则足够远地覆盖图片15的第一增强层位流的响应左上边部分,并且因为内插法分离使得任一层间预测与在空间上与第二增强层位流重叠的图片12的基础层位流的任一重构部分无关,所以甚至对于与图片15的第二位流边界重叠的图片15的第一增强层位流的这些部分,也同样适用。

[0113] 在下面提供了一种实现上面概括的可切换限制的层间上采样的详细实施方式,应注意,例如,正如图4中的情况,如果图片12和图片15是时间对齐配对的图片视频,语法元素616可发出信号或者在每个时间帧内切换限制。而且,再次,应注意,根据本申请的实施方式的解码器可与上面参考图4以及图10提供的描述和功能一致。鉴于此,应注意,上面参考图4提供的关于短期语法元素以及图片12和图片15的空间区段的位置和定位的信令应被视为同样适用于参考图10至图13描述的实施方式。最后,应注意,如果图10中的解码器是对层0和层1中的图片进行解码的图片解码器,则第二方面也是有利的。时间分量是可选的。

[0114] 如下完成在HEVC中实现受限制的层间上采样。就空间可伸缩性而言,使用上采样基础层图片预测增强层图片。在该过程中,使用对应基础层图像区域的多个像元值(通常,在水平方向和垂直方向上)计算增强层的每个像元位置的预测值。如果使用来自不同基础层片的像元,则不可以仅从与增强层片覆盖相同图像区域的基础层片信息将增强层片解码。就受限制的层间上采样作为位流限制发出信号保证了解码器使编码视频序列的所有参

数集合中发出的空间分区信号服从该限制并且由此简化了并行层间解码器的初始化和操作。

[0115] 图10中的概念可以被实现为不允许使用与增强层片对应的基础层片中未包含的关于上采样的相邻像元信息的机制。在位流中,就是是否允许解码器使用位于所有增强层片边界处的基础层的对应图像区域之外的像元发出信号。

[0116] 可替代地,在位流中,仅针对与基础层片边界线对应的增强层片边界线,就是是否允许解码器使用位于增强层片边界处的基础层的对应图像区域之外的像元发出信号。

[0117] 在具体实施方式中,执行片边界线处的基础层的上采样,如同基础层位于图片边界上上,此处,不可获得任何相邻像元。

[0118] 在具体实施方式中,如图14中给出的,通过使用增强层的图片参数集合中的标记可以实现信令。

[0119] `independent_tile_upsampling_idc`对应于语法元素612,不等于0的`independent_tile_upsampling_idc`限制上采样滤波器不能跨越片边界线。如果`independent_tile_upsampling_idc`等于2,则位于与增强层片对应的图片区域之外的任何基础层采样不可用于上采样。如果`independent_tile_upsampling_idc`等于1,则该限制仅适用于与基础层片边界线对齐的增强层片边界线。等于0的`independent_tile_upsampling_idc`并不默示该限制。

[0120] 最后,图15A示例性地示出了通过空间对应方式重叠的两个片分区的图片12和图片15,以示出图14中的语法元素612的实施例:等于2的`independent_tile_upsampling_idc`限制上采样滤波器不能跨越任何增强层片边界线。见单虚点线400。如果`independent_tile_upsampling_idc`等于1,则该限制仅适用于与基础层片边界线对齐的增强层片边界线。见双虚点线402,等于0的`independent_tile_upsampling_idc`并不默示该限制。

[0121] 作为模式`independent_tile_upsampling_idc=2`的可替代模式,或者作为诸如`independent_tile_upsampling_idc=3`等模式的附加模式,上采样滤波器可受限于不能跨越任何片边界线,即,不能跨越这些基础层、或者亦不能跨越这些增强层。见图15B中的线404。

[0122] 即,如上面参考图9进行说明的,根据该模式,在边界400、402、或者404处,处理上采样滤波器。

[0123] 在转至本申请的下一方面之外,应注意,例如,将在图2的预测器60中执行上面讨论的内插法620,以获得层间预测结果。类似编码器端(诸如,预测器18内),因为编码器在编码端执行相同的预测,所以根据语法元素616的设置执行信息620。例如,在编码端,可以根据应用情景执行是否设置语法元素的决策。例如,如果延迟程度是较高的重要因素,则可将语法元素设置为限制层间上采样,并且在其他应用情景中,具有更佳的预测并且使压缩比增加可能更为重要,以使得将语法元素612设置成不限制层间上采样而更具有建议性。

[0124] 刚刚提及的连续层的空间区段的编码之间的最小编码延迟或者偏移量是下一方面的主题以及可被称为“层解码延迟指示”。解码器能够基于短期语法元素确定图片15的空间区段相对于图片12的空间区段的编码之间的最小解码延迟或者偏移量,但是,根据下一概念,使用上期语法元素结构在预定时间段内提前发出层间延迟或者偏移量的信号。再次,这有助于解码器在位流40的并行解码内执行工作负荷分配。作为针对“延迟”或者“偏移量”

的措施,可以使用空间区段,即,可以在空间区段的单元(片、切片、或者WPP的CTB行)内表达该偏移量。

[0125] 为了更为详细地描述之后的方面,参考在很大程度上与图4相一致的图16。因此,图16中使用了相同的参考符号并且参考被称为这些共同元素的元素(如果可行的话),上面参考图4提出的描述同样适用于图16。还应提及的是,除了下面设定的功能之外,图16中所示的视频解码器640可整合图4中参考长期语法元素606所描述的功能。目前,本申请中的图16的该方面还使用了长期语法元素结构,即,642,长期语法元素结构642也被插入到位流40中以参考或者涉及预定时间段608。换言之,尽管视频解码器640可响应图4中的语法元素结构606以及语法元素结构642,然而,仅下面进一步概括的后者功能对于图16中的解码器640比较重要,其中,与语法元素结构606有关的功能和位流40中的语法元素结构606的存在性对于视频解码器640是可选的。然而,上面参考视频解码器600设定的描述同样适用于视频解码器640。即,视频解码器640能够将多层视频数据流40解码,即,使用自第一层(层0)至第二层(层1)的层间预测在层的层级中编码的情景。视频解码器40支持将多层视频数据流并行解码成空间区段,即,通过按照时间重叠方式顺次遍历空间区段而将层的图片分区成空间区段,且所遍历的第一层的图片的空间区段相对于所遍历的第二层的图片的空间区段之间存在层间偏移量。应略微注意,这表示空间区段可以是片、子流或者切片,但是,甚至将刚刚提及的区段单元混合也是可行的。在将切片概念与片概念和/或子流概念组合时,实际上,“空间区段”的定义可以不同。

[0126] 在任何情况下,关于共同时间帧604的图片12和图片15,图16中的视频解码器640能够一方面将图片12的空间区段与另一方面的图片15的空间区段并行(即,通过时间重叠方式)解码。自然,因此,由于层间预测,视频解码器640需要服从两个层之间的某一最小解码偏移量,图片15中的增强层1的当前解码部分必须属于层0的时间对齐图片12的已解码部分。

[0127] 如果是图16,视频解码器640使用长期语法元素结构642预先确定预定时间段608的层间偏移量。

[0128] 结合图16中的实施方式,层间偏移量是关于一方面的图12的第一空间区段的“距离”与另一方面的时间对齐图片15的标量测量。优选地,测量空间“距离”。而且为了变得有意义,基于长期语法元素结构642确定的层间偏移量应对于图片12的第一空间区段的整个解码过程有效。即,用于层间预测的图片12中的所有必须参考部分可用于图片15的整个第一空间区段的解码,前提是之前已经将基础层图片12的第一“层间偏移量”空间区段解码。

[0129] 如上所述,图片15内的“当前解码部分”以特定的预定方式对图片15进行遍历,即,如果使用片并行处理,则按照上述所述片顺序,并且在使用位流的WPP概念的情况下通过倾斜波前形式。同样适用于基础层图片12的空间区段。在使图片15的第一空间区段经过初期解码之前,层间偏移量确定已经被处理的图片12的遍历部分。

[0130] 为了更为详细地对此进行描述,参考图17A和图17B。图17A更为详细地描述了从长期语法元素结构642确定的与片概念有关的层间偏移量,而图17B更为详细地描述了基于长期语法元素结构642确定的与WPP有关的层间偏移量。后面将结合图17C示出使用长期语法元素结构642的层间偏移量信令概念不受限于所使用片和/或WPP概念。更确切地,通过定义仅将图片细分成按照自包含方式可解码的切片(即,“在切片内(in-slice)”完全执行熵解

码和空间帧内图片预测或者熵解码和空间帧内图片预测与相邻的切片无关)使得基于长期语法元素642对层间偏移量进行解译是可行的。

[0131] 图17A示出了都被细分成片的两个时间对齐的图片12和图片15。从上面提出的片概念的描述可以变得清晰,通常,对其进行解码的图片12或者图片15的片内部存在任意的固定顺序。更确切地,可按照任何顺序对片进行解码。然而,结合图16中的实施方式,至少相对于基础层图片12的片限定片顺序。片顺序644被限定为自片的规则布置的左上端片引导至右下端片的行式光栅扫描顺序。根据图17A中的实施方式,通过长期语法元素结构642发出信号的层间偏移量指示按照基础层图片12的片顺序644已经被解码的片的数目,以允许解码器640开始将增强层图片15的第一片解码。为了确定增强层图片15的片之间的“第一片”,可将增强层图片15的第一片固定地限定为增强层图片15的做上端片。自第一片开始至增强层图片15,视频解码器640可采用用于根据图片12的片细分对图片15的增强层片进行遍历的片顺序。在图17A的情况下,例如,将图片12细分成片包括两行和三列片,其中,将图片15细分成四行片与两列片。在这种情况下,有利于使解码器640选择用于对增强层片进行遍历的片顺序,以首先对前两行的左手边片进行遍历,然后,对前两行的右手边片进行遍历,并且然后,如箭头646所示,针对增强层图片15的低行片重复该遍历。然而,根据针对本文中所描述的所有方面有效的可替代实施例,图片15的增强层片之间的片解码顺序被固定并且与基础层图片细分成片无关。总之,如果仅使用信号层间编码偏移量作为用于开始/着手对增强层图片进行解码的触发器,记录器没有必要。图17A使用虚线示出了图片12中局部对应于增强层图片15的第一片的位置的定位。从图17A中可以变得清晰,在图17A的示例性情况下,通过长期语法元素结构642确定的层间偏移量可以是“2”,因为在解码器640开始将增强层图片15的第一片解码之前,不得不将图片12的前两个片解码。在这种情况下,仅基础层图片12中可以使用层间预测所需的共置位置。

[0132] 即,在图17A的情况下,视频解码器640将从长期语法元素结构642确定图片12的基础层片相对于将图片15的增强层片的第一片解码的遍历644之间的层间偏移量是两个基础层片:在视频解码器640按照照片顺序646开始对图片15的增强层片的第一片解码之前,不得不等待按照照片顺序644将前两个基础层片解码。

[0133] 图17B涉及将两个时间对齐的图片12和图片15细分成子流的示例性情况,即,在图12的情况下,细分成两个子流,并且在图15的情况下,细分成四个子流。例如,子流可与将图片12和图片15通过上述所述方式规则地细分成编码块的行和列一致,即,通过这种方式使得各个子流对应于这样的编码块的一个行。在任何情况下,如上所述,由于WPP处理,在图片12的子流与图片15的子流之间分别限定了解码顺序,解码顺序648和解码顺序650从顶部引导至底部。与图17A中的情况相似,解码器640被配置为从长期语法元素结构642确定在开始解码图片15的第一子流之前已被解码的引导子流的数目。在这种情况下,长期语法元素结构642将发出层间偏移量为1的信号,因为将基础层图片12的第一子流完全解码足以提供对增强层图片15的第一子流的任何层间预测的必备基础。

[0134] 图17C示出了两个都被细分成切片的一对时间对齐的图片12和图片15。再次,分别在图片12的切片之间与图片15的切片之间限定切片顺序或者解码顺序,顺序652和顺序654从顶部引导至底部。在图17C的示例性情况中,一方面的图片12内的切片与另一方面方面的图片15内的切片之间的边界局部对应于彼此。因此,根据通过自基础层图片12至增强层图片15

的层间预测而引入的“模糊”，长期语法元素结构642将发出信号：层间预测偏移量等于1或者等于2。具体地，例如，由于放大了用于图片15的层间预测部分的图片12中的对应共置参考部分，例如，如上面参考图9所讨论的，由于视差补偿矢量或者由于上采样内插滤波核，在解码器640能够开始将增强层图片15的第一切片解码之前，可以已经按照切片顺序652将图片12的前两个切片解码。然而，例如，由于图片12与图片15之间的采样分辨率等于彼此并且图片12和图片15与同一视野有关，以至于没有发生任何视差补偿，如果关闭或者不应用层间预测的模糊选项，则通过编码器可将长期语法元素结构设置成等于1，以使得在将基础层图片12的第一切片完全解码之前，解码器640可尽快开始将增强层图片15的第一切片解码。

[0135] 因此，图16至图17C的描述了示出使用长期语法元素结构642有助于编码器支持解码器调度时间对齐的图片12和图片15的任意的并行解码，即，通过基于关于层间偏移量的长期语法元素结构通知解码器，层间偏移量对整个预定时间段608有效并且与基础层图片12的空间区段的数目有关，在开始时间对齐的图片15的第一空间区段之外，已经将基础层图片12的空间区段解码。应注意，视频解码器640能够基于短期语法元素602的检测/评估和进一步关于与层间预测有关的潜在选项的语法匀速确定通过长期语法元素结构642（甚至通过其自身）发出信号的层间偏移量，在自基础层至增强层的层间预测中，这些选项可接通或者关闭刚刚概括的模糊。然而，视频解码器640不得不对多种语法元素进行检测，以推导出由长期语法元素结构642提供的相同信息，并且视频解码器640仅能够基于短期、而非提前推导出关于长预定时间段608的相同信息。

[0136] 与方面1和方面2相似，随后在下文中描述将延迟指示方面引入HEVC的可能方式。

[0137] 首先，参考图18，描述了目前如何在HEVC中实现WPP。即，该描述还构成关于任一上述实施方式中的WPP处理的可选实现方式的基础。

[0138] 在基础层中，波前并行处理允许编码树块(CTB)行的并行处理。跨CTB行的预测相互依赖性不断裂。关于熵编码，从图18中可以看出，WPP改变相应上层CTB行中的左上侧CTB的CABAC相互依赖性。一旦完成对应的右上侧CTB的熵解码，则可开始随后行中的CTB的熵编码。

[0139] 在增强层中，一旦包含对应图像区域的CTB被完全解码并且可用，则可开始CTB的解码。

[0140] 图16中的解码延迟或者偏移量只是概念可在信令中结束，该信令促使解码器对利用并行性的片、WPP、或者切片的分层位流的初始化和操作。

[0141] 当使用空间可伸缩性时，仅当覆盖对应图像区域的基础层CTB可用时，才可开始增强层CTB的解码。当利用空间可伸缩性使WPP与分层位流并行时，层在图片大小的方面可不同，例如，在针对HEVC[1]的可伸缩性扩展以及额外地最大CTB大小提出的已发布调用中规定层之间的图片大小缩放因子1.5和2，例如，HEVC的主要配置文件中支持16、32、以及64的图像采样。关于质量可伸缩性，图片大小缩放因子通常不变，但是，层之间的最大CTB大小仍可不同。

[0142] 各层的最大CTB大小与图片大小缩放因子之间的比影响层解码延迟，即，相对于基础层CTB行的解码，可以将增强层的第一CTB行之前偏移的CTB行解码。图19报告了使用图片大小缩放因子和CTB大小的示例性参数值以及CTB行方面的引入层解码延迟覆盖两个层中的对应图像区域的CTB的比率。

[0143] 对于层之间的质量可伸缩性,通常,层之间的图片大小缩放因子等于1,而相应层的最大CTB大小仍可不同并且影响层解码延迟。

[0144] 语法元素结构642提供了位流中的解码器暗示,即,当使跨空间增强层的WPP处理并行时,发出关于独立空间或者质量增强层的层解码延迟的信号。

[0145] 实施方式使用相应层之间的图片大小缩放因子和最大CTB大小比例因子来确定信号层解码延迟。

[0146] 根据独立基础层位流与独立增强层位流之间的可伸缩性的类型,影响层解码延迟的因子可不同。

[0147] 就所视角可伸缩性而言,层代表相机视图并且使用层间预测的机制从各个角度在相机视图之间执行预测。该预测利用运动补偿机制,以在相机设置中补偿相机的不同位置。在这种情况下,与空间可伸缩性或者质量可伸缩性的情况相比较,通过垂直方向上的最大或者实际运动矢量进一步限制层解码延迟。

[0148] 语法元素结构642描述了位流中的解码器暗示,即,当使跨多酚相机视图的WPP处理并行时,发出关于独立相机视图的层解码延迟的信号。

[0149] 实施方式使用图片大小缩放因子、最大CTB大小比例因子、以及相应层在垂直方向上之间的最大运动矢量长度来确定信号层解码延迟。

[0150] 在与增强层相关联的VUI语法或者SPS中、或者在VPS扩展语法内的编译中,当使用WPP时,就空间区段(即,CTB的行)方面而言,实施方式发出关于空间可伸缩性、质量可伸缩性、或者多视角可伸缩性的层解码延迟的信号。

[0151] 片与切片

[0152] 使用类似片或者切片的其他分区技术的并行处理还可获益于基于将图片划分成空间区段(即,片或者切片)指示解码延迟的位流内的暗示。增强层解码处理需要基础层的信息(例如,重构图像数据)。

[0153] 语法元素结构642描述了位流内的解码器暗示,即,发出关于片和/或切片的层解码延迟的信号。

[0154] 本发明的可能实施方式根据编码视频序列中所使用的并行技术的类型而使用空间区段作为单元来表达被引入的层处理延迟。

[0155] 图20中的语法针对并行工具WPP、片以及切片给出了关于增强层SPS的VUI参数中的min\_spatial\_segments\_delay的指示的示例性实施方式(关于语法元素结构642的实施例)。

[0156] min\_spatial\_segment\_delay描述了通过编码相互依赖性引入的当前层相对于对应的基础层在空间区段方面的解码延迟。

[0157] 根据min\_spatial\_segment\_delay的值,应用如下:

[0158] 如果min\_spatial\_segment\_delay等于0,则不发出关于层的解码之间的最小延迟的限制的信号

[0159] 否则,(min\_spatial\_segment\_delay不等于0),规定了仅下列条件中的一种为真的位流一致性:

[0160] • 在编码视频序列内被激活的各个图片参数集合中,tiles\_enabled\_flag等于0并且entropy\_coding\_sync\_enabled\_flag等于0(即,视频序列中既不使用片,亦不使用

WPP),并且当按照位流顺序将基础层的第一min\_spatial\_segment\_delay切片完全解码时,用于按照位顺序对当前层的第一切片的解码处理的所有基础层资源是可用的。

[0161] • 在编码视频序列内被激活的各个图片参数集合中,tiles\_enabled\_flag等于1并且entropy\_coding\_sync\_enabled\_flag等于0(即,在视频序列中使用片),并且当将覆盖相同图像区域的第一min\_spatial\_segment\_delay片完全解码时,用于按照位流顺序对当前层的第一片的解码处理的所有基础层资源是可用的。

[0162] • 在编码视频序列内被激活的各个图片参数集合中,tiles\_enabled\_flag等于0并且entropy\_coding\_sync\_enabled\_flag等于1(即,在编码视频序列中使用WPP),并且当完成基础层的第一min\_spatial\_segment\_delayCTB行时,用于当前层中的第一CTB行的解码处理的所有基础层资源是可用的。

[0163] 另一示例性实施方式是如[4]中报道的并且如图21所示被扩展的VPS扩展语法中的min\_spatial\_segments\_delay的指示。

[0164] min\_spatial\_segment\_delay描述了通过编码相互依赖性引入的层[i]相对于对应的基础层在空间区段方面的解码延迟。

[0165] 根据min\_spatial\_segment\_delay的值,应用如下:如果min\_spatial\_segment\_delay等于0,则发出针对层的解码之间的最小延迟的限制的信号

[0166] 否则(min\_spatial\_segment\_delay不等于0),规定仅下列条件中的一种为真的位流一致性:

[0167] • 在编码视频序列内被激活的各个图片参数集合中,tiles\_enabled\_flag等于0并且entropy\_coding\_sync\_enabled\_flag等于0(即,在视频序列中既不使用片,亦不使用WPP),并且当按照位顺序将基础层的第一min\_spatial\_segment\_delay切片完全解码时,用于按照位顺序对当前层的第一切片的解码处理的所有基础层资源是可用的。

[0168] • 在编码视频序列内被激活的各个图片参数集合中,tiles\_enabled\_flag等于1并且entropy\_coding\_sync\_enabled\_flag等于0(即,在视频序列中使用片),并且当将覆盖相同图像区域的第一min\_spatial\_segment\_delay片完全解码时,用于按照位顺序对当前层的第一片的解码处理的所有基础层资源是可用的。

[0169] • 在编码视频序列内被激活的各个图片参数集合中,tiles\_enabled\_flag等于0并且entropy\_coding\_sync\_enabled\_flag等于1(即,在编码视频序列中使用WPP),并且当完成基础层的第一min\_spatial\_segment\_delayCTB行时,用于当前层中的第一CTB行的解码处理的所有基础层资源是可用的。

[0170] 上面已经描述了由编码器和解码器支持的各种预测模式、以及对预测模式施加的限制、以及与熵编码/解码的上下文推导,从而能够支持诸如片和/或WPP概念等并行处理概念。上面还提到了,编码器和解码器可基于块操作。例如,基于块选择上面说明的预测模式,即,以比图片自身更细化的粒度。在继续描述本申请的另一方面之前,将对切片、片、WPP位流、以及刚刚提及的块之间的关系进行说明。

[0171] 图32示出了这样一种图片:可以是诸如层12的层0的图片或者层1的图片(诸如图片15)。该图片被规则地细分成块90的阵列。有时,这些块90被称为最大编码块(LCB)、最大编码单元(LCU)、编码树块(CTB)等。将图片细分成块90可形成一种基底(base)或者粗粒度,通过这种基底或者粗粒度执行上述所述预测和残差编码,并且可单独针对层0和层1,通过

编码器发出信号并且设置该粗粒度(即,块90的大小)。例如,可以使用诸如四叉树细分等多树并且在数据流内发出信号,以分别将各个块90细分成预测块、残差块、和/或编码块。具体地,编码块可以是块90的递归式多树细分的叶块并且可以编码块的粒度发出一些预测有关的决策(诸如,预测模式)的信号,并且预测块和残差块可以是编码块中单独的递归式多树细分的叶块,例如,诸如运动矢量(如果是时间帧间预测)和视差矢量(如果是层间预测)等预测参数通过预测块的粒度被编码,并且预测残差通过残差块的粒度被编码。

[0172] 可在块90之中限定光栅扫描编码/解码顺序92。编码/解码顺序92出于空间预测之目的限制相邻部分的可用性:仅图片中根据编码/解码顺序92的部分前进至诸如块90或者其某一更小的块等当前位置,与当前被预测的语法元素有关的当前位置可用于当前图片内的空间预测。编码/解码顺序92在各个层内对图片的所有块90进行遍历,然后,按照图片编码/解码顺序继续对相应层中的下一图片的块进行遍历,图片编码/解码顺序并不一定必须遵循图片的时间再现顺序。在各个块90中,将编码/解码顺序92细化成诸如编码块等更小块之间的扫描。

[0173] 关于刚刚概括的块90和更小的块,按照刚刚提及的编码/解码顺序92将各个图片进一步细分成一个或者多个切片。图32中示例性地示出的切片94a和切片94b相应地无隙覆盖相应的图片。一个图片中的连续切片94a与94b之间的边界或者接触面96可以或者不可以与相邻块90的边界对齐。更精确的,图32中的右手边示出的一个图片内的连续切片94a与94b可在诸如编码块等更小块的边界处彼此接壤,即,块90中的一个的细分的叶块。

[0174] 因为图片的切片94a与94b可形成最小单元,在最小单元中,图片被编码成的数据流的该部分可被打包成包,即,NAL单元。例如,上面描述了切片的另外的可能性,即,关于跨切片边界线的预测和熵上下文而确定的切片的限制。具有该限制的切片可被称为“正常”切片。如下面更为详细概括的,除正常切片之外,还可存在“依赖切片”。

[0175] 如果将片分区概念用于图片,则可改变在块的阵列90之间限定的编码/解码顺序92。图33中示出了这种情况,其中,示例性地示出了将图片分区成四个片82a至82d。如图33所示,片自身被限定为以块90为单位将图片规则地细分。即,片82a至82d中的每个均由 $n \times m$ 阵列的块90构成,且单独针对每行片设置n,并且单独针对每列片设置m。遵循编码/解码顺序92,在进行至下一片82b等之前,首先,按照光栅扫描顺序对第一片中的块90进行扫描,其中,按照光栅扫描顺序对片82a至82d自身进行扫描。

[0176] 根据WPP流分区概念,按照编码/解码顺序92将一行或者多行块90的单元中的图片细分成WPP子流98a至98d。如图34所示,例如,每个WPP子流均可覆盖一个完整行的块90。

[0177] 然而,还可混合片概念与WPP子流概念。在这种情况下,例如,每个WPP子流均可覆盖每个片内的一行块90。

[0178] 甚至可以共同使用图片的切片分区与片分区和/或WPP子流分区。关于片,根据编码/解码顺序92,图片被细分成的一个或者多个切片中的每个均可以仅由一个完成的片构成、或者一个以上完成的片构成或者仅一个片的子部分构成。鉴于此,构成用于并行性的最小单元的切片可包括一方面的正常切片和另一方面上的依赖切片:而正常切片对预测施加上述所述限制和熵上下文进行推导,依赖切片则不施加该限制。自图片边界处开始的依赖切片采用从紧随的前一行的块90中的熵解码块90产生的熵上下文,编码/解码顺序92按照行式顺次展开,并且从某一其他地方开始的依赖切片可采用因对紧随的前一切片至其结束进

行熵编码/解码而产生的熵编码上下文。通过这种方式, WPP子流98a至98b中的每个均可由一个或者多个依赖切片构成。

[0179] 即,在块90之间限定的编码/解码顺序92线性地从相应图片的第一侧(此处,示例性地,左侧)引导至相对侧(示例性地,右侧),然后,在向下/底部方向上步至块90的下一行。因此,主要布置在诸如当前块90等当前编码/解码部分的左侧和顶部的当前图片的已编码/解码部分是可用的。由于预测中断和跨边界线的熵上下文推导,可以并行处理一个图片的片。甚至,一个图片的片的编码/解码可以同时开始。限制起源于上面提及的环内滤波,在这种情况下限制允许跨边界线。进而,按照自顶部至底部的交错方式执行开始WPP子流的编码/解码。在块90中测量连续的WPP子流之间的帧内图片延迟是两个块90。

[0180] 然而,甚至有利于使图片12与图片15的编码/解码并行,即,不同层之间的时隙。显然,相对于基础层的编码/解码,不得不延迟依赖层的图片15的编码/解码,以保证已经存在可用的基础层的“空间对应”部分。甚至,如果不单独使用图片12和图片15中的任一个内的编码/解码的任何并行性,则这些理念是有效的。甚至,如果使用一个切片来分别覆盖整个图片12和图片15,且不使用任何片和任何WPP子流处理,则可使图片12和图片15的编码/解码并行。下面描述的信令(即,第六方面)是表达层的解码/编码延迟的可能性,甚至,在这种情况下,对层的任一图片使用片或者WPP处理,或者与对层的任一图片使用片或WPP处理的情况无关。

[0181] 从与连续层的编码之间的最小编码延迟有关的上述信息变得清晰,显而易见,解码器能够基于短期语法元素确定最小解码延迟。然而,如果使用上述所述长期语法元素,以提前发出关于预定时间段内的层间时间延迟的信号,解码器可计划在将来使用所提供的保证并且可在位流40的并行解码内更容易地执行工作负荷分配。

[0182] 下面描述的本申请的方面(即,本申请的第六方面)从某种程度上与方面3有关,在于涉及任一层间偏移量的明确信令。然而,相对于本申请的第六方面,明确发出层间偏移量的信令的语法元素结构不需要基于长期发出与短期语法元素(从短期语法元素可另行推导出层间偏移量)有关的信号。更确切地,本申请的第六方面采用另一发现:当描述图17A至图17C时变得清晰的是,如果按照期间限定的光栅扫描解码顺序将基础层图片和增强层图片细分成块,则通过测量基础层块的单元内的层间偏移量的测量可以有效明确地发出基础层与增强层之间的层间偏移量的信号。结合下面进一步描述的方面,明确发出层间偏移量信号的单元内的基础层块不受限于空间区段。更明确地,鉴于此,可以使用其他编码块。因此,当参考图34描述关于本申请的第六方面的实施方式时,在很大程度上共同使用参考图16至图17C所使用的参考符号,并且上面参考后者特征提出的描述应同样适用于下面进一步描述的实施方式,以便面不必要的重复。除此之外,参考图32和图33的描述,因为这些图形成一方面的编码块与另一方面空间区段之间的可能共存。

[0183] 因此,图35示出了被配置为接收多层视频数据流40的视频解码器720,利用从第一层的部分至第二层的共置部分的层间预测将情景编码在层的层级中,与上述图相似,示例性地示出了作为代表性实施例的并置的层0与层1。图35中示例性地示出了两个层的两个时间对齐图片12和图片15。基础层0的图片12被细分成一系列的第一块722,并且增强层1的图片被细分成一系列第二块724。在块722之间限定光栅扫描解码顺序726,同样,在块724之间限定光栅扫描解码顺序728。

[0184] 视频数据流40包括语法元素结构,语法元素结构指示用于通过时间重叠方式顺次遍历第一块722和第二块724来将图片12和图片15并行解码的层间偏移量以及在基础层块722的单元中测量的遍历的第一块722与第二块724之间的层间偏移量。视频解码器720被配置为响应于语法元素结构730。具体地,视频解码器从后者确定层间偏移量。

[0185] 与图16中的实施方式相符的是,语法元素结构730可指示层间偏移量作为成功执行比短期时间间隔长的预定时间段内的时间对齐图片12和15的并行解码的保证,在预定时间段内,可选的语法元素信号图片12和15分别被细分成块722和块724。然而,这不是强制性的。更确切地,在诸如每对时间对齐的基础层图片12和增强层图片15等不同范围的数据流40内可以实现通过语法元素结构730明确发出层间偏移量的信号,例如,即,在与关于大小以及细分成块722和块724的信令作用相同的间隔内。

[0186] 与上述实施方式进一步相符,解码器720可使用通过语法元素结构730明确发出信号的层片偏移量作为分别将图片12和图片15并行解码时相对于第一块722的遍历开始第二块724的遍历时的偏移量的测量。换言之,视频解码器720可被配置为从语法元素结构730推导出计数值,被配置为在仅在达到通过语法元素结构730明确发出最小计数信号的已解码块722的计数之后,在允许按照解码顺序728开始增强层图片15的块724的序列的解码的同时,按照光栅扫描解码顺序726计算基础层图片12的已解码块722。因此,视频解码器720不需要对视频数据流40的任意高复杂度和分布的部分进行检测,否则,能够使视频解码器720通过其他方式计算一方面的开始解码块722与另一方面块724之间的实际最小层间偏移量。

[0187] 然而,有趣的是,根据图35中的实施方式,块722和块724不一定必须代表尤其专门经过任何并行处理的空间区段。更确切地,块722和块724可以是以分别被编码成视频数据流40的图片12和图片15的内容为单位的常见编码块。例如,块722和块724可以是树根块,图片12和图片15被规则地(即,按照行和列)细分,然后,树根块通过上面参考如32描述的递归式多树方式被进一步单独地细分,例如,代表图片12和图片15的树根块的产生叶块以编码块为单位被细分成编码块,在视频数据流40内发出参考图15在空间预测、时间预测以及层间预测之中选择的预测模式的信号。

[0188] 为了更为详细地对图35中的实施方式进行说明,参考图36。如图36中所示,视频解码器720可使用计数器732来计算基础层图片12的已解码块722的数目,且按照光栅扫描解码顺序726从图片12的第一块722开始计数。解码器720的比较器734将通过计数器732输出的稳定增加的计数与从视频数据流40获得的语法元素结构730的明确信号值相比较。如果该计数满足与由语法元素结构730指示的值的预定关系,诸如,一旦计数器732的计数达到或者等于由语法元素结构730指示的值,则比较器激活或者使增强层图片15开始解码,即,使按照光栅扫描解码顺序728增强层图片15的第一块724开始解码。

[0189] 如下面参考语法元素结构730的语法将更为详细地描述的,将图35中的语法元素结构与图16中的语法元素结构及其目标统一是可行的。即如下。语法元素结构730可具有一组可能值,即,可能值集合。如参考图16示例性地提及的,可能值的非明确的层间偏移量集合之外的值可使视频解码器720丢弃语法元素结构730的值并且不指定图片12和图片15的并行解码或者基于短期语法元素602确定任意层间偏移量。如果语法元素结构730假设了第二组可能值之外的值,这将导致视频解码器720执行相对于图36已经概括的动作,例如,根

据该动作,语法元素结构730的值将明确发出基础层块722的单元内的层间偏移量的信号。然而,当通过语法元素结构730假设时,可能存在语法元素结构730的可能值的另一子集,从而致使视频解码器720行使如上参考图16描述的动作:通过将后者解译为测量(可以但不一定必须分别整数倍的一些由块722和块724构成)空间区段的单元的层间偏移量而确定基于语法元素结构730的值解码的基础层图片12与增强层图片15之间的层间偏移量。

[0190] 参考图37示出了将图35中的实施方式与图16中的实施方式组合的刚刚提及的可能性。如图37所示,视频解码器可对语法元素结构730进行检测,以确定语法元素结构730是否具有可能值742的集合之外的第一子集736、第二子集738或者第三子集740中的值。根据该调查或者检查结果,解码器720不能在语法元素结构730之外推导出任何保证并且从语法元素结构不能推导出关于层片偏移量的任何明确信令,或者从语法元素结构730执行层间偏移量的推导,即,在空间区段的单元内或者块的单元内执行该推导。如果是第二子集738,则不发生任何推导/任何保证,如果是子集736,则在空间区段的单元内发生层间偏移量的推导,并且如果语法元素730假设第第三子集740之外的值,则在块的单元内推导层间偏移量。在下面进一步概括的特定语法实施例中,语法元素结构包括两个标记,即,ctb\_delay\_enabled\_flag和min\_spatial\_segment\_delay,其中,ctp\_delay\_enabled\_flag=0且min\_spatial\_segment\_delay≠0对应于子集736的情况,min\_spatial\_segment\_delay=0对应于第二子集738,并且ctp\_delay\_enabled\_flag=1且min\_spatial\_segment\_delay≠0对应于第三子集740。

[0191] 最后,参考图38,示出了解码器720可被配置为对通过语法元素结构730发出信号的任何层间偏移量进行解译,该层间偏移量不仅仅是相对于开始增强层图片15的第一块或者空间区段的解码的层间偏移量,而且是连续的层间偏移量,当服从连续层间偏移量时,分别产生图片12与图片15的无冲突并行解码。然而,如图38所示,计数器732仍计算基础层图片12的已解码块722的数目,其中,附加计数器744同样按照解码顺序728计算增强层图片15的已解码块724,其中,减法器746形成两个计数之间的差,即,s与t-1,即,计算s-t+1。通过比较器734将该差值与从语法元素结构730推导出的层间偏移量值相比较,并且一旦两个值(即,推导的层间偏移量与计数之间的内插)具有预定的关系,诸如,差值等于或者超过被推导出的层间偏移量,则开始按照增强层块724之间的解码顺序728将块t解码。通过这种方式,设置一方面的基础层图片12的解码块722与另一方面增强层图片15的块724的原因之间的连续检验。

[0192] 显而易见,根据图38的连续检验也可适用于空间区段。更一般地,也已经将图38和图36的描述转移至空间区段并且这种陈述也适用于图16中的实施方式,其中,语法元素结构642可用作图36和图38中以730示出的相关语法元素结构。换言之,至少当使用片作为空间区段时,也存在其间限定的光栅扫描解码顺序,以使得可将相对于编码块讨论的图36和图38中的改变容易地转移至片的遍历及其解码。

[0193] 简要总结第六方面及与其有关描述的实施方式,语法元素结构730可通过视频编码器被插入到位流中,以向解码器提供如何控制相对于彼此的基础层图片与增强层图片的并行解码的明确暗示。通过语法元素结构明确发出信号的层间偏移量可被级或者或者去活。如果被激活,则该指示可以在诸如CTB等块的单元内,或者通过更为精确的信令发出信号的块单元和空间区段的单元中的一个内。由于使用一方面的基础层块与另一方面增强

层块之间的光栅扫描顺序,例如,两个都从顶部至底部逐行自每个图片12/15的左上角通过行式方式引导至右下角,明确发出信号的层间偏移量可仅被解译为用于开始/着手解码增强层图片15的第一块的“触发器”,或者被解译为一方面的基础层图片12的当前解码块与另一方面的增强层图片15的当前解码块之间的连续“安全距离”,即,用于确定对增强层图片15的各个块进行解码的触发器。可将参考第六方面提出的描述转移至第三方面的描述和实施方式,在于,至少只要涉及作为空间区段的片的描述,则涉及信号层间偏移量的服从性的解译和检查,可以使用图36和图38中的实施方式,即,可通过与图36和图38中的描述对应的方式控制按照光栅扫描片解码顺序对基础层图片和增强层图片内的解码片的遍历。

[0194] 因此,可以使用“延迟”空间区段作为一种测量,即,可以在空间区段的单元(片、切片或者CTB行)内表达该延迟,或者可以在块90的单元内测量延迟/偏移量。

[0195] 高效视频编码(HEVC)标准可被扩展至符合下面的第六方面。具体地,如果参考数据可用,则允许各个层(或者视图)的并行解码。通过诸如片等并行工具的粒度确定基础层编码树块(CTB)的解码与对应的独立增强层CTB的解码之间的最小延迟(具体地,层解码延迟),波前、或者切片、或者运动补偿矢量是适用的(例如,在立体或者多视点译视频编码中)。

[0196] 图20示出了通过增强层序列参数设定语法(SequenceParameterSetSyntax)和语法元素min\_spatial\_segment\_delay的语义实现的层解码延迟指示。

[0197] min\_spatial\_segment\_delay描述了通过编码相互依赖性引入的当前层相对于对应的基础层在空间区段方面的解码延迟。

[0198] 如下可以实现基于多个垂直和水平CTB允许可选地表达独立相对的基础层与独立的增强层之间的层解码延迟的HEVC高阶语法中的机制,与潜在的并行技术无关。

[0199] 可以使用指数(例如,ctb\_delay\_enabled\_flag)的标记发出信号:将层解码延迟(如利用第二语法元素发出的信号)表达为编码图片中的具体CTB地址。

[0200] 从图39中可以看出并且根据下列条件,光栅扫描顺序中的CTB地址清晰地限定了图片内用于表达该延迟的水平和垂直位置。

[0201] CTB坐标=(CTB地址%PicWidthInCTBs, CTB地址/PicWidthInCTBs)

[0202] PicWidthInCTBs描述了CTB单元内的图片的宽度。

[0203] 图39示出如下。图片内的CTB地址(例如,7)限定水平CTB列和垂直CTB行,例如,元组(2,1)。

[0204] 如果启用标记,则将当前独立层中的CTB解码时,另一语法元素的值(cp.min\_spatial\_segment\_delay)被解译成相对基础层图片中的共置CTB的CTB地址的偏移量。

[0205] 如图40所示和下列描述过程中的细节,根据两个相应层中的CTB的大小和两个相应层的图片的宽度可以计算共置的CTB。

[0206] 图40包含示出了相应的基础层与增强层的CTB大小与图片大小的各种设置的自左至右的三种实施例,与图片比例无关。基础层图片中的粗体划线框标记增强层CTB的大小中的图像区域和相应基础层的CTB布局中的其共置图像区域。

[0207] 在增强层Sequence Parameter Set Syntax和由图41中的语法元素min\_spatial\_segment\_delay的语义中给出了该可选的CTB基础层解码延迟指示的实施方式。

[0208] 等于1的ctb\_based\_delay\_enabled\_flag指示在CTB单元中给出通过min\_

spatial\_segment\_delay发出信号的延迟。ctb\_based\_delay\_enabled\_flag指示在CTB单元中未给出min\_spatial\_segment\_delay。

[0209] min\_spatial\_segment\_delay描述了通过编码相互依赖性引入的当前层相对于对应的基础层在空间区段方面的解码延迟。

[0210] 根据min\_spatial\_segment\_delay的值,应用如下:如果min\_spatial\_segment\_delay等于0,则不发出关于层的解码之间的最小延迟的任何限制的信号。

[0211] 否则(min\_spatial\_segment\_delay不等于0),并且如果ctb\_based\_delay\_enabled\_flag等于1,则规定下列条件为真的位流一致性:

[0212] • 且CtbSize<sub>A</sub>、PicWidthInCtbs<sub>A</sub>、以及ctbAddrRs<sub>A</sub>是基础层A的CtbSizeY和PicWidthInCtbsY以及按照光栅扫描顺序的基础层A中的Ctb的CtbAddress、CtbSize<sub>B</sub>、PicWidthInCtbs<sub>B</sub>、以及ctbAddrRs<sub>B</sub>是独立层/视图B的CtbSizeY和PicWidthInCtbsY以及按照光栅扫描顺序的独立层B的Ctb的CtbAddress,并且CtbScalingFactor<sub>BA</sub>、CtbRow<sub>BA</sub>(ctbAddrRs)以及CtbCol<sub>BA</sub>(ctbAddrRs)

[0213] 被确定如下:

[0214] CtbScalingFactor<sub>BA</sub> = (PicWidthInCtbs<sub>A</sub>/PicWidthInCtbs<sub>B</sub>)

[0215] CtbRow<sub>BA</sub>(ctbAddrRs) =

[0216] Ceil((Floor(ctbAddrRs/PicWidthInCtbs<sub>B</sub>)+1)\*CtbScalingFactor<sub>BA</sub>) - 1

[0217] CtbCol<sub>BA</sub>(ctbAddrRs) =

[0218] Ceil(((ctbAddrRs % PicWidthInCtbs<sub>B</sub>)+1)\*CtbScalingFactor<sub>BA</sub>) - 1

[0219] 当使用当前增强层/视图B的ctbAddrRs<sub>B</sub>对CTB进行解码时,当具有等于PicWidthInCtbs<sub>A</sub>\*CtbRow<sub>BA</sub>(ctbAddrRs<sub>B</sub>)+CtbCol<sub>BA</sub>(ctbAddrRs<sub>B</sub>)+min\_spatial\_segment\_delay的基础层CTB被完全解码时,所有必要的基础层资源是可用的。

[0220] 否则(min\_spatial\_segment\_delay不等于0并且ctb\_based\_delay\_enabled等于0),规定了仅下列条件中的一种为真的位流一致性:

[0221] • 在编码视频序列内被激活的各个图片参数集合中,tiles\_enabled\_flag等于0并且entropy\_coding\_sync\_enabled\_flag等于0(即,在视频序列中既不使用片,亦不使用WPP),并且当按照位流顺序将基础层的第一min\_spatial\_segment\_delay切片完全解码时,按照位流顺序的用于当前层的第一切片的解码处理的所有基础层资源是可用的。

[0222] • 在编码视频序列内被激活的每个图片参数集合中,tiles\_enabled\_flag等于1并且entropy\_coding\_sync\_enabled\_flag等于0(即,在视频序列中使用片),并且当将覆盖相同图像区域的第一min\_spatial\_segment\_delay片完全解码时,按照位流顺序的用于当前层的第一片的解码处理的所有基础层资源是可用的。

[0223] • 在编码视频序列内被激活的每个图片参数集合中,tiles\_enabled\_flag等于0并且entropy\_coding\_sync\_enabled\_flag等于1(即,在编码视频序列中使用WPP),并且当完成基础层的第一min\_spatial\_segment\_delayCTB行时,用于当前层中的第一CTB行的解码处理的所有基础层资源是可用的。

[0224] 可替代地,如之前的实施方式,可以发出信号:将层间偏移量作为ctb\_based\_delay\_enabled\_flag的某种最坏情况延迟、而非第一切片/片/CTB行的启动延迟。最坏情况延迟给出了保证:在将独立图片的空间区段解码的过程中,当将加上信号偏移量的共置空

间区段完全解码时,所有必须的相应基础层资源是可用的。

[0225] 在图42中示出了针对语法的实施方式。

[0226] `min_spatial_segment_delay`描述了通过编码相互依赖性引入的当前层相对于对应的基础层在空间区段方面的解码延迟。

[0227] 根据`min_spatial_segment_delay`的值应用如下:如果`min_spatial_segment_delay`等于0,则发出关于层的解码之间的最小延迟的任何限制的信号。

[0228] 否则(`min_spatial_segment_delay`不等于0),规定了仅下列条件中的一种为真的位流一致性:

[0229] • 在编码视频序列内被激活的各个图片参数集合中,`tiles_enabled_flag`等于0并且`entropy_coding_sync_enabled_flag`等于0(即,在视频序列中既不使用片也不使用WPP),并且当将按照位流顺序位于包含相对于当前层中的切片A的相同图像区域的至少一部分的最后切片区段B之后的第一切片区段C(按照位流顺序紧随基础层中的(`min_spatial_segment_delay`-1)切片之后)完全解码时,按照位流顺序用于当前层中的任一切片区段A的解码处理的所有基础层资源是可用的。

[0230] • 在编码视频序列内被激活的各个图片参数集合中,`tiles_enabled_flag`等于1并且`entropy_coding_sync_enabled_flag`等于0(即,在视频序列中使用片),并且当将按照位流顺序位于包含相对于片A的相同图像区域的至少一部分的最后片B之后的第一片C(按照位流顺序紧随(`min_spatial_segment_delay`-1)片之后)完全解码时,按照位流顺序的用于当前层中的任一片A的解码处理的所有基础层资源是可用的。

[0231] • 在编码视频序列内被激活的各个图片参数集合中,`tiles_enabled_flag`等于0并且`entropy_coding_sync_enabled_flag`等于1(即,在编码视频序列中使用WPP),并且当将按照位流顺序位于覆盖相对于增强层的CTB行A的相同图像区域的至少一部分的基础层的最后CTB行B之后的第一CTB行C(紧随(`min_spatial_segment_delay`-1)CTB行之后)完全解码时,用于当前层中的任一CTB行A的解码处理的所有基础层资源是可用的。

[0232] 图片或者`min_spatial_segment_delay`的基于`sub-coded-video-sequence`的信令也是可行的。如图20中给出的,就相关联的NAL单元而言,SEI消息的范围比时间域中的编码视频序列更小并且通过位流中的SEI消息的位置或者通过指数限定SEI消息的范围。在图43的`Layer_decoding_delay_SEI`中给出了一种实施方式。

[0233] 可以改变相对于前面描述的实施方式的语法,以反映SEI消息的范围及其语法元素。

[0234] 可以略微修改上述明确的实施方式。在上述实施例中,语法元素结构包括`min_spatial_segment_delay`和`ctb_based_delay_enabled_flag`,且`min_spatial_segment_delay`根据`ctb_based_delay_enabled_flag`通过一维方式或者标量方式按照空间区段/CTB解码顺序测量空间区段或者CTB的单元内的层间编码偏移量。但是,由于基础层图片的CTB的数目通常阿玉诸如基础层图片的片或者位流等空间区段的数目,所以,在略微不同的实施方式中,如果是指示基于CTB的层间偏移量指示的`ctb_based_delay_enabled_flag`,则不能仅基于`min_spatial_segment_delay`确定层间偏移量,而是,更确切地,在这种情况下,后面的语法元素被解译为指示基础层图片的CTB在水维度上的定位,解码器的完整解码可被用作触发器,以开始增强层图片的解码。自然,可替代地,`min_spatial_segment_delay`可被

解译成沿着垂直维度指示基础层图片的CTB。根据ctb\_based\_delay\_enabled\_flag,即,如果其指示基于CTB的指示,则在作用于刚刚提及的触发器的基础层图片的CTB在其他维度上的定位的数据流中发送进一步的语法元素。

[0235] 即,可以使用用于信令的下列语法片段,即,可被用作语法元素结构:

[0236] 指数i和j可指示基础层与增强层的层id。

[0237]	min_spatial_segment_offset_plus1[i][j] if (min_spatial_segment_offset_plus1[i][j]>0) { ctu_based_offset_enabled_flag[i][j] if (ctu_based_offset_enabled_flag[i][j]) min_horizontal_ctu_offset_plus1[i][j]	ue(v) u(1) ue(v)
--------	---	------------------------

[0238] 上述语法元素的语义可被描述成如下:

[0239] 如下规定,在第i层的第j个直接参考层的每个图片(即,不用于对第i层的任一图片进行解码的层间预测)中,min\_spatial\_segment\_offset\_plus1[i][j]自身或者与min\_horizontal\_ctu\_offset\_plus1[i][j]共同指示空间区域。min\_spatial\_segment\_offset\_plus1[i][j]的值应在0至refPicWidthInCtbsY[i][j]\*refPicHeightInCtbsY[i][j](包含)的范围内。当不存在时,推测min\_spatial\_segment\_offset\_plus1[i][j]的值等于0。

[0240] 在CTU单元内、在第j层的第j个直接参考层的每个图片(即,不用于第i层的任一图片的解码的层间预测)中,由min\_spatial\_segment\_offset\_plus1[i][j]和min\_horizontal\_ctu\_offset\_plus1[i][j]共同指示的等于1的ctu\_based\_offset\_enabled\_flag[i][j]规定了空间区域。在切片区段的单元内、或者在CTU行内、在第i层的第j个直接参考层的每个图片(即,不用用户第i层中的任一图片的解码的层间预测)中,仅由min\_spatial\_segment\_offset\_plus1[i]指示的等于0的ctu\_based\_offset\_enabled\_flag[i][j]规定了空间区域。当不存在时,推测ctu\_based\_offset\_enabled\_flag[i]的值等于0。

[0241] 如下规定,在第i层的第j个直接参考层的每个图片(即,不用于第i层中的任一图片的解码的层间预测)中,当ctu\_based\_offset\_enabled\_flag[i][j]等于1时,min\_horizontal\_ctu\_offset\_plus1[i][j]与min\_spatial\_segment\_offset\_plus1[i][j]共同指示空间区域。min\_horizontal\_ctu\_offset\_plus1[i][j]的值应在0至refPicWidthInCtbsY[i][j](包含)的范围内。

[0242] 当ctu\_based\_offset\_enabled\_flag[i][j]等于1时,推导变量minHorizontalCtbOffset[i][j]如下:

[0243]  $minHorizontalCtbOffset[i][j] = (min_horizontal_ctu_offset_plus1[i][j] > 0) ? (min_horizontal_ctu_offset_plus1[i][j] - 1) : (refPicWidthInCtbsY[i][j] - 1)$

[0244] 将变量curPicWidthInSamples\_L[i]、curPicHeightInSamples\_L[i]、curCtbLog2SizeY[i]、curPicWidthInCtbsY[i]、以及curPicHeightInCtbsY[i]设置成分别等于第i层的PicWidthInSamples\_L、PicHeightInSamples\_L、CtbLog2SizeY、PicWidthInCtbsY以及PicHeightInCtbs。

[0245] 将变量refPicWidthInSamples\_L[i][j]、refPicHeightInSamples\_L[i][j]、refCtbLog2SizeY[i][j]、refPicWidthInCtbsY[i][j]以及refPicHeightInCtbsY[i][j]设置成分别等于第i层中的第j直接参考层的PicWidthInSamples\_L、PicHeightInSamples\_L、

CtbLog2SizeY、PicWidthInCtbsY以及PicHeightInCtbsY。

[0246] 将变量curScaledRefLayerLeftOffset[i][j]、curScaledRefLayerTopOffset[i][j]、curScaledRefLayerRightOffset[i][j]以及curScaledRefLayerBottomOffset[i][j]设置成分别等于第i层的第j直接参考层的scaled\_ref\_layer\_left\_offset[j]<<1、scaled\_ref\_layer\_top\_offset[j]<<1、scaled\_ref\_layer\_right\_offset[j]<<1、scaled\_ref\_layer\_bottom\_offset[j]<<1。

[0247] colCtbAddr[i][j]表示CTU的共置CTU的光栅扫描地址,且光栅扫描地址等于第j层的图片中的ctbAddr,在第i层的第j直接参考层的图片中,推导变量colCtbAddr[i][j]如下:

[0248] 变量(xP,yP)指定CTU的左上亮度采样的定位,且光栅扫描地址等于与第i层的图片中的左上亮度采样有关的ctbAddr。

[0249]  $xP = (ctbAddr \% curPicWidthInCtbsY[i]) \ll curCtbLog2SizeY$   $yP = (ctbAddr / curPicWidthInCtbsY[i]) \ll curCtbLog2SizeY$

[0250] 推导变量scaleFactorX[i][j]和scaleFactorY[i][j]如下:

[0251] curScaledRefLayerPicWidthInSamplesL[i][j]=curPicWidthInSamplesL[i]-curScaledRefLayerLeftOffset[i][j]-curScaledRefLayerRightOffset[i][j]

[0252] curScaledRefLayerPicHeightInSamplesL[i][j]=curPicHeightInSamplesL[i]-curScaledRefLayerTopOffset[i][j]-curScaledRefLayerBottomOffset[i][j]

[0253] scaleFactorX[i][j]=((refPicWidthInSamplesL[i][j]<<16)+(curScaledRefLayerPicWidthInSamplesL[i][j]>>1))/curScaledRefLayerPicWidthInSamplesL[i][j]

[0254] scaleFactorY[i][j]=((refPicHeightInSamplesL[i][j]<<16)+(curScaledRefLayerPicHeightInSamplesL[i][j]>>1))/curScaledRefLayerPicHeightInSamplesL[i][j]

[0255] 变量(xCol[i][j],yColxCol[i][j])规定了第i层中的亮度采样定位(xP,yP)的第j直接参考层的图片中的共置亮度采样定位,推导变量(xCol[i][j],yColxCol[i][j])如下,:

[0256]  $xCol[i][j] = Clip3(0, (refPicWidthInSamplesL[i][j] - 1), ((xP - curScaledRefLayerLeftOffset[i][j]) * scaleFactorX[i][j] + (1<<15)) >> 16)$

[0257]  $yCol[i][j] = Clip3(0, (refPicHeightInSamplesL[i][j] - 1), ((yP - curScaledRefLayerTopOffset[i][j]) * scaleFactorY[i][j] + (1<<15)) >> 16)$

[0258] 推导变量colCtbAddr[i][j]如下:

[0259]  $xColCtb[i][j] = xCol[i][j] >> refCtbLog2SizeY[i][j]$

[0260]  $yColCtb[i][j] = yCol[i][j] >> refCtbLog2SizeY[i][j]$

[0261]  $colCtbAddr[i][j] = xColCtb[i][j] + (yColCtb[i][j] * refPicWidthInCtbsY[i][j])$

[0262] 当min\_spatial\_segment\_offset\_plus1[i][j]大于0时,规定应适用下列条件的位流一致性:

[0263] 如果ctu\_based\_offset\_enabled\_flag[i][j]等于0,则仅使用下列中的一种:

[0264] • 在被分配至(referredtoby)第i层的第j直接参考层中的图片的每个PSS中,tiles\_enabled\_flag等于0并且entropy\_coding\_sync\_enabled\_flag等于0,并且适用如

下：

[0265] • 使切片区段A成为第i层的图片的任一切片区段并且使ctbAddr成为切片区段A中的最后CTU的光栅扫描地址。使切片区段B成为与切片A属于相同访问单元的切片区段(属于第i层的第j直接参考层)并且包含具有光栅扫描地址colCtbAddr[i][j]的CTU。使切片区段C成为与切片区段B位于同一图片内的切片区段并且按照解码顺序紧随切片区段B之后，并且按照解码顺序,min\_spatial\_segment\_offset\_plus1[i]-1切片区段存在于切片区段B与该切片区段之间。当切片区段C存在时,切片区段A的语法元素受约束,以使得切片区段C中的任何采样或语法元素值、或者按照解码顺序紧随C之后的任何切片区段或相同图片可用于切片区段A的任何采样的解码处理时的层间预测。

[0266] • 在被分配至第i层的第j直接参考层中的图片的每个PPS中,tiles\_enabled\_flag等于1并且entropy\_coding\_sync\_enabled\_flag等于0,并且适用如下:

[0267] • 使片A成为第i层的任一图片picA中的任一片并且使ctbAddr成为片A中的最后CTU的光栅扫描地址。使片B成为图片picB中与picA属于相同访问单元并且术语第i层的第j直接参考层的片并且包含具有光栅扫描地址colCtbAddr[i][j]的CTU。使片C成为也位于picB中并且按照解码顺序紧随片B之后的片,并且按照解码顺序,min\_spatial\_segment\_offset\_plus1[i]-1片存在于片B与该片之间。当存在切片区段C时,片A的语法元素受到约束,以使得片C中的任何采样或语法元素值、或者按照解码顺序紧随C之后的任一片或相同图片可用于片A内的任何采样的解码处理时的层间预测。

[0268] • 在被分配至第i层的第j直接参考层中的图片的各个PPS中,tiles\_enabled\_flag等于0并且entropy\_coding\_sync\_enabled\_flag等于1,并且适用如下:

[0269] • 使CTU行A成为第i层的任一图片picA中的任一CTU行并且使ctbAddr成为CTU行A中的最后CTU的光栅扫描地址。使CTU行B成为图片picB中与picA属于同一访问单元并且属于第i层的第j直接参考层的CTU行并且包含具有光栅扫描地址colCtbAddr[i][j]的CTU。使CTU行C成为也位于picB中并且按照紧随CTU行B之后的CTU行,并且按照解码顺序,min\_spatial\_segment\_offset\_plus1[i]-1CTU行存在于CTU行B与该CTU行之间。当存在CTU行C时,CTU行A的语法元素受约束,以使得CTU行C中的任何采样或语法元素、或者按照解码顺序紧随C之后的相同图片的行可用于CTU行A内的任何采样的解码处理时的层间预测。

[0270] 否则(ctu\_based\_offset\_enabled\_flag[i][j]等于1),并且适用如下:

[0271] • 推导变量refCtbAddr[i][j]如下:

[0272]  $xOffset[i][j] = ((xColCtb[i][j] + minHorizontalCtbOffset[i][j]) > (refPicWidthInCt_bsY[i][j] - 1)) ? (refPicWidthInCtbsY[i][j] - 1 - xColCtb[i][j]) : (minHorizontalCtbOffset[i][j])$

[0273]  $yOffset[i][j] = (min_spatial_segment_offset_plus1[i][j] - 1) * refPicWidthInCtbsY[i][j]$

[0274]  $refCtbAddr[i][j] = colCtbAddr[i][j] + xOffset[i][j] + yOffset[i][j]$

[0275] • 使CTUA成为第i层的任一图片picA中的任一CTU,并且使ctbAddr成为CTUA的光栅扫描地址ctbAddr。使CTUB成为该图片中与picA属于同一访问单元并且属于第i层的第j直接参考层的CTU并且具有比refCtbAddr[i][j]更大的光栅扫描地址。当CTUB存在时,CTUA的语法元素受约束,使得CTUB中的任何采样或者语法元素值被用于CTUA内的任何采样的解

码处理时的层间预测。

[0276] 即,总结刚刚提出的实施方式,使用标记通过可切换方式可完成层间偏移量的基于CTB的指示,标记在一方面的基于CTB的指示与另一方面层间偏移量的基于空间区段的指示之间切换,如此,基于CTB的可替代指示可使用无条件发送的语法元素,即,与基于CTB或者基于空间区段的层间偏移量指示无关,即,如果发出关于空间区段式指示的信号,则为指示空间区段的单元中的层间偏移量的语法元素,该语法元素作为基础层图片中的“触发器CTB”的定位的诸如水平分量或者垂直分量等一个分量。根据在基于CTB的层间偏移量指示与基于空间区段的层间偏移量指示之间切换的语法元素,则不再发送进一步的语法元素。具体地,如果发出基于CTB指示的信号,则发送进一步的语法元素。在这种情况下,后面的语法元素指示“触发器CTB”的定位的缺少维度,因此,解码器可使用两种语法元素按照行和列在基础层图片的CTB的规则布置中识别“触发器CTB”,并且一旦将CTB解码,则解码器可开始将增强层图片解码。尽管如此,通过使用第一语法元素的代表性状态中的一种(即,min\_spatial\_segment\_delay)可完整地切换任一层间偏移量的指示。由于CTB之间限定的预定解码顺序,所以在基于CTB的层间偏移量指示情况中,解码器仍可将触发器CTB的位置的水平分量和垂直分量转换成基础层图片的CTB的数目,在开始增强层图片的第一CTB的解码之前,不得不将基础层图片完全解码,以使得解码器可使用图36中的实施方式来控制层间偏移量的基于CTB指示的服从性。

[0277] 在下文中,将更为详细地描述本发明的另外的方面。第四方面涉及下列问题:优选地,从编码器接收位流的所有分区网络实体能够容易区分在多层次数据流中传递的各个层。例如,中间网络实体可能对从进一步传输中排除诸如关于超过某一分辨率阈值的采样分辨率的层等特定信息层感兴趣。下列描述提供了关于HEVC的设想扩展的当前情形的概述。

[0278] HEVC[1]的视频参数集合(VPS)提供访问编码位流的高阶并且包含对处理中间设备或者端设备中的位流比较重要的信息。即将到来的HEVC的可伸缩和多视图扩展将进一步得益于提供指定用于可伸缩位流的语法的VPS扩展。VPS扩展的主要任务之一是提供一种对NAL单元报头中的nuh\_reserved\_zero\_6bits进行解译的统一解决方法.nuh\_reserved\_zero\_6bits旨在被重新加签为layer\_id并且用作可伸缩视频编码情景中的普通层标识符。在下表中给出了NAL单元报头中的layer\_id语法元素以及在[1]中和如图2所示给出的NAL单元报头。

[0279] 在设计过程[5]中考虑了两种常见解决方法。最后,在VPS型式中发出将NAL单元的报头中的单个标识符的值映射至潜在的多个可伸缩标识符的信号的解决方法。其次,在VPS扩展中发出将NAL的报头中的单个标识符的各个位(或者位块)分配至特定可伸缩的标识符的解决方法。

[0280] 在[4]中报告的当前VPS扩展语法的设计使用映射解决方法,但是,已经包含两种解决方法所必需的所有语法元素,即,两种语法元素指示可伸缩性的类型(cp.scalability\_map)和每个可伸缩性维度的层数(cp.dimension\_id\_len\_minus1)。

[0281] 映射解决方法将进一步的语法元素引入至VPS扩展语法,即,则为被编码成u(v)的可伸缩标识符的实际值,可选地,如果编码器选择通过非连续形式稀疏地分配layer\_id的值,则为layer\_id。

[0282] 在许多非常类似的可伸缩情景中,例如,两个或者三个空间层、两个或者三个视图

等,不一定必须使用NAL单元报头中的6位层标识符的所有63个值。对于这些可伸缩的情景,与基于映射的解决方法相比较,将NAL单元报头中的层标识符的各个位分配至具体的可伸缩维度的解决方法具有两种优点。

[0283] • 就可伸缩性维度标识符而言,NAL单元报头中的层标识符值的解译不需要间接法或者查找法。

[0284] • 不需要发送映射解决方法所需的VPS扩展语法元素,而VPS扩展语法元素是用于可伸缩信令的VSP扩展位的重要部分。

[0285] • 中间设备不需要存储用于通过各个视频位流的映射表。

[0286] 根据上面所描述的第四方面的概念,HEVC的高阶语法中可能存在指示是否使用映射解决方法或者分区解决方法的暗示。

[0287] 根据一种实施方式,依据暗示,映射有关语法元素的一部分 (cp.vps\_nuh\_layer\_id\_present\_flag、layer\_id\_in\_nuh[i] 以及 dimension\_id[i][j]) 被发送或者被漏掉,并且发出关于可伸缩性类型 (cp.scalability\_mask) 和每种可伸缩性的层的量 (cp.dimension\_id\_len\_minus1) 的语法元素的信号并且根据暗示对语法元素进行解译,该暗示为关于NAL单元报头中的可伸缩标识符的分区或者映射的信息。

[0288] 参考图23提出与本发明的第四方面的概念对应或者采用本发明的第四方面的概念的实施方式。图23示出了网络实体,网络实体可以是如上面已经讨论的任一种的视频解码器,或者可以是在编码器与解码器之间产生的中间网络实体。通常,使用参考符号680表示网络实体。网络实体用于处理诸如上述所述数据流40中的任一种等多层视频数据流682。在网络实体680是视频解码器的情况下,该处理将涉及多层视频数据流628的解码。在中间网络实体的情况下,例如,该处理可涉及视频数据流的转发。

[0289] 在层中将情景编码成多层视频数据流,以使得在各个层中,将该情景编码成通过可伸缩维度扩展的可伸缩性空间的不同操作点,其中,多层视频数据流由第一NAL单元构成,第一NAL单元中的每个均与一个层相关联,并且第二NAL单元散布在第一NAL单元内并且在多层视频数据流中呈现常见信息。换言之,第一NAL单元684可携带视频的图片中的一个或者多个切片,且“图片”对应于多层视频数据流682的任一层。

[0290] 在上述实施方式中,为易于描述,仅讨论了两个层。自然,层的数目可以大于二并且甚至关于层参与之前任一层的信息分类可因层不同而不同。除第一NA两单元684之外,示出了散布在NAL单元684之间的NAL单元686,但是,与第一NAL单元684相比较,通过单独的信道可执行其发送。第二NAL单元通过下面更为详细设定的方式在多层视频数据流中呈现常见信息。

[0291] 为了更为详细地描述一方面的第一NAL单元与另一方面数据流682的层的设置之间的关联性,参考图24。图24示出了代表全部第一NAL单元684的第一NAL单元684。报头684内包括层指示符字段690。除报头688之外,NAL单元684包括关于上面参考其他图讨论的切片数据的有效载荷数据692,即,关于使用层间预测而编码的视频内容的数据。图24还示出了层的设置(即,694)。具体地,图24中所示的层的设置694应表示全部可能层,通过NAL单元684中的层指示符字段690可代表并且区分彼此。即,一方面的设置694与层指示符字段690的可能值之间的相关联函数应被假设成一一映射。在图24中,使用小圆圈示例性地示出了设置694中的各个层,每个小圆圈均具有其上标记的特定编号。尽管这些标记编号建议了

设置694的层之间限定的顺序,然而,应注意,下列讨论将揭示并不仅基于层指示符字段690推导设置694中的层的布置或者分类方式。更确切地,鉴于此,网络实体680需要对散布的第二NAL单元686内的类型指示符字段696进行检测。然而,后面将对此进行描述。

[0292] 换言之,直至目前,图24中的设置694的各个元素仅代表NAL单元684中的层指示符字段690的一种可能状态。通过层指示符字段690可以区分设置694中的层,但是,对于没有第二NAL单元686提供的额外信息的网络实体680,这些层之间的语义含义和顺序并不能变得清晰,然而,实际上,设置694的层形成树的节点,且树之间的分支对应于特定的可伸缩性维度或者轴。例如,一个层是基础层并且对应于树的根。每个分支连接两层,即,树的两个节点,描述了特定层如何贡献于另一层,即,使用层间预测额外提供信息分类,信息分类对应于可伸缩性维度并且可以是例如空间分辨率增加、SNR增加等。例如,出于简化之缘故,图24示出了通过与空间可伸缩性和SNR可伸缩性对应的两个可伸缩性维度700和702扩展的二维可伸缩性空间698。图24中示出了层的示例性树机器通过层698的扩展。图24还示出了数据流682中并不可以呈现设置694的全部可能层。例如,在图24的示例性情况中,实际上,数据流682中仅使用五个层。

[0293] 例如,图22示出了层指示符字段可具有6个位,从而区分设置694的 $2^6=64$ 中可能状态或者可能层。通过下面更为详细描述的方式经由第二NAL单元686可以调整自设置694的这些可能值或层至可伸缩性空间698中的操作点的映射。图24中使用参考符号704表示该映射。“操作点”应表示可伸缩空间698内的设置694内的至少实际呈现层的位置。例如,可伸缩空间698的原点可以与基础层或者树根相关联,而沿着可伸缩空间698的轴700和轴702中的任一个的每个分支均可具有固定长度1。因此,指向可伸缩空间698中的操作点的矢量可具有整数值坐标。

[0294] 目前,简要总结所提出的描述,多层视频数据流682提供关于多个层中的视频内容或者情景的信息。层布置在树中,且每个层均经由分支连接至树。从形成树根的基础层开始,紧随的下一层贡献于关于某种信息的视频内容信息的重构型式,可被解译为可伸缩性维度。因此,每个层是树根层或者通过分支的特定路径连接至后者,并且需要属于沿着该路径布置的层的NAL单元684老重构相应层的视频内容。自然,优选地,如果执行映射704,使得相对于树根引导相应分支的末梢端的任何“贡献”层具有层指示符字段690的值,即,比位于相应分支的近端处的层的层指示符字段的值更高。

[0295] 图25更为详细地示出了类型指示符字段696的含义。图25示出了作为固定位长度的字段的层指示符字段690。在任何情况下,字段690的长度与类型指示符字段696的值无关。然而,如果类型指示符字段具有第一状态,则层指示符字段690被视为整体,即,其全部n个位被共同处理,以区分其可能值。优选地,如果类型指示符字段696假设第一状态,则通过网络实体680从层指示符字段690推导出n位整数。在类型指示符字段696假设第一状态的情况下,网络实体680通过使用第二NAL单元686内传递的映射信息执行m位字段690的可能值到操作点的映射704。如图25所示,例如,映射信息708包括分配层指示符字段690的每个实际使用的可能值的表,矢量指向相应可能值的相关联操作点。即,如果类型指示符字段696假设第一状态,则网络实体680可从第二NAL单元686推导出映射信息708并且可对每个层指示符字段690执行查找映射信息或者标708,以查找相关联的矢量,以在空间698中定位相应的/相关联的操作点。与映射信息708中的可能值相关联的矢量的维度的数目p可以

为默认设置或者可以在诸如第二NAL单元686等数据流内发出信号。之后,示出了可以发出下列信息的信号,以传递关于映射信息708的信息:`vps_max_layers_minus1`可确定实际使用的M位字段690的可能值的数目M。`num_dimensions_minus1`可限定维数。使用语法元素`scalability_mask`可以发送后面两个变量。然后,经由一个面的配对可能值(即,`layer_id_in_nuh`)与p维矢量(即,`dimension_id[i][j]`)可以发出表自身的信号。映射704则引导至p维矢量,即,经由映射信息708被映射至相应矢量710中的一个矢量,矢量710指向空间698内与具有层指示符字段690的NAL单元的层相关联的操作点

[0296] 然而,如果类型指示符字段696假设第二状态,则不同程度地执行映射704。具体地,在这种情况下,通过将层指示符字段690分割成一个以上部分执行该映射,即,通过将m位字段690的序列分割成n个连续子序列的位。将由此获得的每个部分用作n维矢量712的坐标 $x_1 \cdots x_n$ ,反之,坐标 $x_1 \cdots x_n$ 指向可伸缩空间698内与包括层指示符字段690的NAL单元的层相关联的操作点。例如,通过形成字段690的前(最重要) $m_1$ 个位之外的第一部分、字段690的随后 $m_2$ 个(次重要)位之外的第二部分等、直至具有 $m_1 + \cdots + m_n = m$ 的第n部分,将层指示符字段690的m个位分割成n个部分。每个部分的位可直接被视为整数值表示法。

[0297] 在类型指示符字段假设第二状态的情况下,维数n可以是默认设置或者经由数据流发出信号。如果在基于相同的语法元素推导出n之后,使用本文中所描述的具体实施方式来确定p,如果类型指示符字段696假设第一状态,即,基于`scalability_mask`。还经由语法元素`dimension_id_len_minus1`示例性地发出信号:层指示符字段690被细分成的部分的位长度(即, $m_1, \cdots, m_n$ ),但是,再次,通过默认(无显式发送)可实现该细分。

[0298] 应注意,可被解析的诸如`scalability_mask`等相同语法结构(与类型指示符字段696无关)可指示可伸缩空间的维度的数目和语义含义。然而,例如因为在映射情况(类型指示符字段的第一状态)中可用的可伸缩空间的维度的最大数目比分量解译情况(类型指示符字段的第一状态)中可用的最大数目更高,所以相对于映射情况,如果类型指示符字段696指示层指示符字段的分量式解译,则该语法元素的允许状态可能受限于语法元素的可使用可能状态的适当子集。编码器仅相应地服从这种限制。

[0299] 关于图23至图25中的实施方式的示例性使用情况如下:

[0300] 大规模多方会议

[0301] 在基于RTP的大规模会话服务情景中,例如,网络会议,在获知相应视频位流的参数集合的多点控制单元(MCU)中适配多方之间的传输视频。各方提供拇指位流和具有讲话者的增强空间分辨率(例如,720p和4K)的两个位流。MCU就关于向哪一方提供哪些流做出决策。因此,可伸缩性参数的易于解析明显减轻了MCU的负担。与可伸缩性信令的基于映射的解决方法相比较,基于分区的解决方法需要更少的计算和内存资源。

[0302] 传输系统

[0303] 在诸如传RTP或者MPEG2-TS等输系统中,与基于映射的解决方法相比较,可伸缩性相关的编解码器信息到相应元素的映射可获益于诸如分区等较不复杂的位保存机制。传输系统不可能采用映射解决方法形式的可伸缩性信令,而是,更确切地,解决映射间接并且针对各个可伸缩性维度产生专用可伸缩性标识符,进而在诸如分区解决方法中明确发出可伸缩性维度的信号。

[0304] 图26中的实施例提出了一种作为VPS扩展语法中的语法标记的可能实施方式,该

实施方式允许基于如[4]中报道的VPS扩展语法在用于HEVC的扩展中的可伸缩性信令的基于映射解决方法与基于分区解决方法之间切换。

[0305] 具有等于1的值的dedicated\_scalability\_ids\_flag指示根据dimension\_id\_len\_minus1[ ]的值分割NAL单元报头中的layer\_id字段的位并且layer\_id字段的位属于在scalability\_mask中发布信号的相应可伸缩性维度。具有等于1的值的dedicated\_scalability\_ids\_flag指示不发出关于语法元素vps\_nuh\_layer\_id\_present\_flag、layer\_id\_in\_nuh[i]、以及dimension\_id[i][j]的信号。具有等于1的值的dedicated\_scalability\_ids\_flag指示描述对应NAL单元的可伸缩标识符的变量的推导仅使用NAL单元报头(cp.layer\_id)中的可伸缩标识符的相应相关联位,例如,如下形式:

[0306] DependencyId=layer\_id&&0x07

[0307] QualityId=layer\_id&&0x38

[0308] 给定在NAL单元报头中发出关于layer\_id的信号,具有等于0的值的dedicated\_scalability\_ids\_flag指示发出关于语法元素vps\_nuh\_layer\_id\_present\_flag、layer\_id\_in\_nuh[i]、以及dimension\_id[i][j]的信号,并且NAL单元报头中的layer\_id的位与具体的可伸缩性维度不相关联,而是被映射至VPS扩展中的可伸缩标识符。具有等于0的值的dedicated\_scalability\_ids\_flag指示描述对应NAL单元的可伸缩标识符的变量的推导使用语法元素dimension\_id[i][j],例如,如下形式:

[0309] if(layer\_id==layer\_id\_in\_nuh[0]) {

[0310] DependencyId=dimension\_id[0][0]

[0311] QualityId=dimension\_id[0][1]

[0312] 当dedicated\_scalability\_ids\_flag等于0时,dimension\_id\_len\_minus1[i]指示dimension\_id[i][j]的位的长度。当dedicated\_scalability\_ids\_flag等于1时,dimension\_id\_len\_minus1[i]指示:

[0313] 如由scalability\_mask指示的,与第i可伸缩性维度相关联的NAL单元报头中的layer\_id的位数。

[0314] 图27中的流程图进一步示出了本发明的可能实施方式。直接经由掩模位复制(maskedbitcopy)从layer\_id的位推导可伸缩标识符或者经由与layer\_id的具体值的相关联性在VPS中发出可伸缩标识符的信号。

[0315] 图31中示出了另一语法实施例。此处,通过“splitting\_flag”发出类型指示符字段的信号,而层指示符字段被称为nuh\_layer\_id。根据“splitting\_flag”,使用映射概念或者分区概念从nuh\_layer\_id推导可伸缩空间内的相应NAL单元层的操作点。通过等于1的splitting\_flag示例性地发出分区概念的信号。然后,通过位掩模复制可以从NAL单元报头中的nuh\_layer\_id语法元素推导可伸缩标识符,即,关于可伸缩空间的可伸缩维度的矢量分量。关于图25中的矢量712的第i个分量(第i可伸缩性维度)的相应位掩模被限定如下。具体地,等于1的splitting\_flag指示不存在dimension\_id[i][j]语法元素(即,不存在任何映射信息708)并且NAL单元报头中的nuh\_layer\_id值的二进制标识符被分割成一系列NumScalabilityType,即,n个区段,具有根据dimension\_id\_len\_minus1[j]的值的以位表示的长度的x<sub>1...n</sub>,并且从字段690的NumScalabilityType区段推断dimension\_id[LayerIdxInVps[nuh\_layer\_id]][j]的值,即,矢量712的分量x<sub>1...n</sub>。通过一系列标记

scalability\_mask\_flag发出关于可伸缩性空间的可伸缩轴的语法含义和数目的信号,从而向此处示例性地固定数目的预定可伸缩类型中的每个指示相应的可伸缩性类型是否属于可伸缩空间698的可伸缩性维度中的任一个。具体地,网络实体680能够根据图31中的for循环从标记scalability\_mask\_flag的序列中推导可伸缩性空间,即,可伸缩轴NumScalabilityType的语法含义和数目:

[0316] for(i=0,NumScalabilityTypes=0;i<16;i++) {  
 [0317] scalability\_mask\_flag[i]  
 [0318] NumScalabilityTypes+=scalability\_mask\_flag[i]  
 [0319] }

[0320] 其中,等于1的scalability\_mask\_flag[i]指示存在第i可伸缩性维度,并且等于0的scalability\_mask\_flag[i]指示不存在第i可伸缩性维度。此处,i=1可表示多视角(视角可伸缩性),i=2可表示空间/SNR可伸缩性,并且i=0可表示深度映射信息的添加。其他可伸缩性维度类型也可存在,自然,刚刚概括的实施例仅是示出性的。假设nuh\_layer\_id的长度示例性地为6,则可如下完成分区或者掩模复制:

[0321] 变量dimBitOffset[0]被设置成等于0,并且对于1至NumScalabilityTypes-1(包含)范围内的j, dimBitOffset[j]的推导如下:

[0322] 
$$\text{dimBitOffset}[j] = \sum_{\text{dimIdx}=0}^{j-1} (\text{dimension_id_len_minus1}[\text{dimIdx}] + 1)$$

[0323] dimension\_id\_len\_minus1[NumScalabilityTypes-1]的值被推测等于5-dimBitOffset[NumScalabilityTypes-1]。

[0324] dimBitOffset[NumScalabilityTypes]的值被设置为等于6。

[0325] 规定了位流一致性,即,当NumScalabilityTypes大于0时, dimBitOffset[NumScalabilityTypes-1]应小于6。

[0326] 对于从0至NumScalabilityTypes-1(包含)的j,推测dimension\_id[i][j]等于((nuh\_layer\_id&((1<<dimBitOffset[j+1])-1))>>dimBitOffset[j])。

[0327] 总之,下列语法元素涉及将各个第一NAL单元与其可伸缩性空间内的操作点相关联:

[0328] 1)nuh\_layer\_id,即,层指示符字段690

[0329] 2)scalability\_mask\_flag的序列,即,显示空间698的可伸缩轴700、702的数目和含义的信息,以及由此字段的部分 $x_i$ 的数目n

[0330] 3)每个轴的字段690的各部分 $x_i$ 的dimension\_id\_len\_minus1,即,位长度(只有一个除外,因为可推测剩余的一个包括字段690的所有其余位706)

[0331] 4)此外,根据图31中的可选实施方式发送vps\_max\_layers\_minus1,该语法元素指示可能 $\sum_i 2^m$ 一些中的已使用/实际层的数目以及分区层指示符字段的vps\_max\_layers\_minus1可能值的序列,即,layer\_id\_in\_nuh[i]的序列,从而限定可能的操作点之间的排列顺序。

[0332] 如果splitting\_flag等于0,则使用映射概念。因此,使用下列信息发送映射信息708:

[0333] 1)scalability\_mask\_flag的序列以及由此表708中的M矢量的分量的数目p,即,

显示空间698的可伸缩轴700、702的数目和含义的信息

[0334] 2) 表708中的矢量dimension\_id[i][j]的各个分量 $x_j$ 的dimension\_id\_len\_minus1, 即, 位长度, 即, 空间698的轴中的每一个

[0335] 3) 可选地, layer\_id\_in\_nuh[i]用作M个矢量dimension\_id[i][j]的列表的索引

[0336] 4) 可选地, 尽管图31中未示出, 然而, 发送vps\_max\_layers\_minus1, 从而指示可能的 $\sum_i 2^{\text{dimension\_id\_len\_minus1}[i]}$ 一些中的使用/实际层的数目M

[0337] 因此, 如果splitting\_flag等于0, 则通过没有明确信令的分区固有地推导矢量710, 即, 没有转而被推测的信令dimension\_id[i][j]。

[0338] 因此, 根据第四方面, 即, 位流40的概念“NAL单元层标识符的可切换解译”可包括NAL单元, 即, VPSNAL单元, VPSNAL单元包括类型指示符字段300, 经由类型指示符字段300, 可以在映射概念与位分割概念之间切换, 以对“正常”NAL单元中的层指示符字段302进行解译。因此, 在两种模式下使用字段302的相同位位置, 但是, 在模式之间发生位流变化时发送发出位的解译和解译指令信息的信号, 即, 映射信息或者分割和语义信息。尽管这使类型指示符字段的额外传输成为必要, 然而, 该概念一起引导更为有效的位流传输作为映射概念的优点, 并且可根须需要采用位分割概念, 因为诸如根据层数等, 两者概念可从不同的程度适用于不同的多层次数据。

[0339] 本申请的第五方面关于多标准的多层次视频解码器接口。此处设定的概念描述了传输层解码器以及传输层解码器(类似MPEG传输流或者RTP)与支持不同层中的不同视频编码标准的可伸缩视频解码器之间的接口(例如, 基础层中的H.264/AVC和增强层中的HEVC)。

[0340] 可伸缩的视频位流由各层构成: 其由一个基础层与一个或者多个增强层构成, 一个基础层包含独立的解码视频信号, 并且一个或者多个增强层可仅结合基础层(以及潜在地其他增强层)被解码, 并且提供更高的时间分辨率(时间伸缩性)、空间分辨率(空间伸缩性)、质量(SNR伸缩性)、更高的位深度(位深度伸缩性)、视频信号或者其他相机视图(多视角伸缩性)

[0341] 如H.264/AVCSVC的现有可伸缩视频编码标准限定相同标准中的基础层和增强层。通过可伸缩位流具有与非可伸缩位流相同的基础格式的方式设计基础层与增强层。如果将可伸缩位流输入到非可伸缩解码器中, 则仍可以找出包类型并且丢弃未知的包。

[0342] HEVC是允许使用基础层的不同视频编码标准(例如, H.264/AVC)的第一视频编码标准。两种标准的包格式是不同的, 因此, 基础层解码器不能理解增强层包。另一方面, 增强层解码器可理解增强层包格式, 但是, 不能理解基础层包格式。

[0343] 在音频/视频系统中, 使用传输层将若干个音频流与视频流合成并且提供如定时和流类型等元数据。

[0344] 在现有的多层次传输层解码器中, 将基础层和增强层的访问单元多路复用成单一视频数据流(例如, H.264/AVC的Annex B Byte stream)。将该视频流输入至视频解码器中。

[0345] 在所有情况下, 如果基础层与增强层使用不同的视频编码标准, 则不能将基础层包与增强层包合成单个位流。

[0346] 根据第五方面的实施方式, 传输层解码器区分下列情况:

[0347] 1. 读取输出的视频解码器仅可将基础层解码

[0348] 2. 读取输出的视频解码器可将基础层与增强层解码, 并且使用相同的视频编码标

准将基础层与增强层编码

[0349] 3. 读取输出的视频解码器可将基础层与增强层解码, 并且使用不同的视频编码标准将基础层与增强层编码

[0350] 在情况1中, 传输层解码器使用下列行为:

[0351] 仅从传输层中提出包含基础层的包并且将包输入至视频编码标准中规定的格式的单标准单层视频解码器中。

[0352] 具体实施方式是传输流解码器通过仅选择具有被分配的流类型“与ITU-T Rec.H.264|ISO/IEC 14496-10的Annex A中定义的一个或者多个配置文件相符的AVC视频流或者AVC视频子位流”的流从MPEG-2传输流中提取H.264/AVC NAL单元并且将H.264/AVC NAL单元输入至H.264/AVC specification的Annex B中定义的字节流格式的H.264/AVC视频解码器中。在传输流解码器中丢弃属于其流类型不等于流类型“与ITU-T Rec.H.264|ISO/IEC 14496-10的Annex A中定义的一个或者多个配置文件相符的AVC视频流或者AVC视频子位流”的流的NAL单元。图28给出了具体实施方式的例证。

[0353] 在情况2中, 传输层解码器使用下列行为:

[0354] 从传输层提取来自基础层和增强层的包并且将包输入至视频编码标准中规定的格式的单标准多层视频解码器中。

[0355] 具体实施方式可通过选择具有被分配的流类型“与ITU-T Rec.H.264|ISO/IEC 14496-10的Annex A中定义的一个或者多个配置文件相符的AVC视频流或者AVC视频子位流”的基础层流并且还选择具有被分配的流类型“ITU-T Rec.H.264|ISO/IEC 14496-10的Annex G中定义的一个或者多个配置文件相符的AVC视频流的SVC视频子位流”的一个或者多个增强层而从MPEG-2传输流中提取H.264/AVC NAL单元。不同层的NAL单元被多路复用成H.264/AVC specification的Annex B中定义的字节流格式并且被输入至H.264/AVC SVC视频解码器中。图29给出了具体实施方式的例证。

[0356] 在情况3中, 传输层解码器使用下列行为:

[0357] 从传输层提取来自基础层和增强层的包。通过下列部分中描述的方式之一在多标准多层视频解码器中处理这些包。

[0358] 接口A

[0359] 如果增强层标准包格式允许携带基础层包, 则通过增强层格式封装基础层包。这指将报头添加到由增强层标准视频解码器可以理解的各个基础层包中并且允许增强层视频解码器将包识别为不同视频编码标准的基础层。

[0360] 在本发明的具体实施方式中, 使用H.264/AVCNAL单元作为HEVC NAL单元的有效载荷, 即, 在H.264/AVC NAL单元报头的前面添加HEVC NAL单元报头。使用HEVC NAL单元报头中的字段(例如, nal\_unit\_type)将有效载荷识别为H.264/AVC的NAL单元。位流可被输入至HEVC Annex B字节流格式的视频解码器中。图30给出了具体实施方式的例证。

[0361] 接口B

[0362] 可伸缩位流的各个层使用不同的信道。通过信道选择在解码器中确定视频编码标准。

[0363] 在本发明的具体实施方式中, 在两个单独的信道中交换两个层。第一信道仅用于H.264/AVC基础层包(或者AnnexB字节流), 而第二信道仅用于HEVC增强层包。

[0364] 接口C

[0365] 指示视频编码标准的类型的元数据字段与从传输流解码器传递至多标准多层视频解码器的各个包相关联。通过相同的方式可发出例如类似定时的其他元数据的信号。

[0366] 在具体实施方式中,通过相关联的元数据字段将各个基础层NAL单元识别为H.264/AVC NAL单元并且通过相关联的元数据字段将各个增强层NAL单元识别为HEVC NAL单元。

[0367] 因此,可将第五方面的概念描述为“多标准多层视频解码器接口”,从而就如何合成不同编解码器的位流的方式提供可能性。

[0368] 因此,根据本申请的第五方面,传输层解码器可被配置成如下面参考图44概括的。通常,使用参考符号770表示图44中所示的传输层解码器。传输层解码器770被配置为使入站多层视频数据流40经过在层中编码的情景,从而通过多标准多层解码器772进行解码且多标准多层解码器772连接传输层解码器770的输出接口。多层视频数据流40由上面已经参考本申请的各个其他方面概括的NAL单元构成,其描述可转移至图44的实施方式。各个NAL单元均与一个层相关联。各层与不同的编解码器(即,不同的标准)相关联。对于各个层,使用相同编解码器将与相应层相关联的NAL单元编码,即,与相应层相关联的一个。

[0369] 对于各个NAL单元,传输层解码器770被配置为识别与其相关联的相同编解码器并且将多层视频数据流40的NAL单元移交至多标准多层解码器,多标准多层解码器使用与不同编解码器相关联的层之间的层间预测,以将多层视频数据流解码。

[0370] 如上所述,各个NAL单元可通过上面参考本申请的第四方面已经概括的特定层指示符字段与多层视频数据流40中的一个层相关联,NAL单元中的一些或者全部可携带内容相关数据,即,一个或者多个切片。通过收集关于特定集合的层的所有NAL单元,解码器772利用由该集合的层授予的信息量将被编码成数据流40的视频内容或者情景解码。关于层相关性、具有一个以上可伸缩性维度的选项等,参考本申请的第四方面的描述。

[0371] 多层多标准解码器772能够处理不同的编解码器/标准。上面已经提出了关于不同标准的实施例,即,H.264、HEVC,但是,也可混合其他标准。不同的编解码器/标准并不受限于混合式编解码器。更确切地,还可使用混合的不同种类的编解码器。多层多标准解码器772所使用的层间预测可涉及不同层中使用的预测参数或者可涉及各个时间对齐层的图片采样。上面已经参考其他方面和实施方式对此进行了描述。

[0372] 传输层解码器770可被配置为对属于编码器的各层的NAL单元执行移交,多层多标准解码器772仅能够处理该移交。即,由传输层解码器770执行的移交可取决于与各个NAL单元相关联的编解码器的传输层解码器770的标识。具体地,传输层解码器770可对各个NAL单元执行如下:

[0373] • 例如,通过检测NAL单元的NAL单元报头中的层指示符字段而识别与当前被检测的NAL单元相关联的层。

[0374] • 一方面,基于数据流40的各层之间的相关联性与传输层解码器770基于数据流40的相应高阶语法的检测推导相同的编解码器/标准的相关联性,传输层解码器40确定当前检测的NAL单元是否满足下列两种标准:NAL单元层属于被转发至解码器772的层的子集,通过可伸缩空间中的当前检测的NAL单元层的操作点确定该子集,以及关于允许将可伸缩空间内的哪些操作点转发至多层多标准解码器772并且不能将可伸缩空间内的哪些操作点

转发至多层多标准解码器772的外部指令。此外,传输层解码器770检查当前检测的NAL单元层编解码器是否属于多层多标准解码器772能够处理的编解码器/标准的集合。

[0375] • 如果检查显示通过当前检测的NAL单元满足这两种标准,则传输层解码器770将当前NAL单元转发至解码器772,以用于解码。

[0376] 对于传输层解码器770,存在确定一方面的数据流40中包含的各层之间的上述所述相关联性以及另一方面以上述所述相关联性为基础的编解码器/标准的不同可能性。例如,如上所述,参考“接口B”,可以使用不同的信道传递数据流40(即,一个信道中的一种编解码器/标准的各层的NAL单元)以及根据另一信道的另一编解码器/标准编码的各层的NAL单元。通过这种措施,传输层解码器770能够通过区分各个信道而推导一方面的各层之间的刚刚提及关联性以及另一方面的编解码器/标准。例如,对于数据流40的各个NAL单元,传输层解码器770确定相应NAL单元到达的信道,以识别与相应NAL单元或者相应NAL单元的层相关联的编解码器/标准。

[0377] 此外或者可替代地,传输层解码器770可通过根据相应编解码器/标准的方式将属于不同编解码器/标准的各层的NAL单元转发至多层多标准解码器772,以使得通过一个信道将属于一个编解码器/标准的各层的NAL单元发送至解码器772,并且通过另一信道将不同编解码器/标准的各层的NAL单元转发至多层多标准解码器772。

[0378] 基础传输层可提供“不同的信道”。即,为易于理解,通过区分由图44中未示出的基础传输层提供的不同信道标识符可以实现不同信道之间的区分。

[0379] 将数据流40移交至多层多标准解码器772的另一种可能性在于传输层解码器770使用具有NAL单元类型指示符的预定编解码器的NAL单元报头封装被识别为与层(与不同于预定编解码器的任一编解码器相关联)相关联的的NAL单元,NAL单元类型指示符字段被设置成指示相应层的编解码器的状态。这指如下:例如,预定编解码器可以是数据流40的任一增强层的任一编解码器。例如,基础层编解码器(即,与数据流40的基础层相关联的编解码器)可以不同于预定的编解码器(例如,HEVC)。因此,当使数据流40经过多层多标准解码器722时,传输层解码器770可将数据流40转换成与预定编解码器相符的数据流。鉴于此,传输层解码器770使用预定编解码器的NAL单元报头封装属于不使用预定编解码器被编码的层的每个NAL单元并且将NAL单元报头内的NAL单元类型指示符设置成指示相应实际层的编解码器的状态。例如,基础层NAL单元是使用HEVC和NAL单元报头(具有被设置成指示H.264的状态的NAL单元类型指示符)被相应地封装的基础层的H.264NAL单元。因此,多层多标准772将接收与数据流相符的HEVC。

[0380] 自然,如参考接口C描述的,可替代地,传输层解码器770可以为入站数据流40的每个NAL单元提供指示与相应NAL单元相关联的层与其相关联的编解码器的元数据。因此,将通过由此扩展的方式将数据流的40个NAL单元转发至解码器772。

[0381] 使用刚刚描述的可替代方法,易于通过进一步层扩展被编码成数据流的内容,然而,使用诸如新编码器等另一编码器将进一步进行层编码,并且不需要修改编解码器40的现有部分。多层多标准解码器进而能够处理新编解码器(即,新添加的一个),能够使用利用新编解码器编码的各层处理现有已混合的数据流。

[0382] 因此,上面提出了关于HEVC可伸缩位流的并行/低延迟视频编码概念。

[0383] 高效率视频编码(HEVC)标准[1]最初的特征在于:允许在编码器端和解码器端上

并行处理的两种专用并行工具：片和波前并行处理 (WPP)。当与特征不是各个图片内的并行处理的HEVC编码视频相比较时，这些工具允许图片内的并行，目标在于提高处理时间，同时，使编码效率损失最小化。

[0384] 在可伸缩[2]的HEVC位流或者多视图[3]HEVC位流中，使用解码的基础层或者基本视图图片来预测增强层或者相关的视图图片。在上述描述中，使用术语“层”还可同时覆盖视图的概念。

[0385] 上述实施方式描述了能够在完成相关联的基础层图片的解码之前而开始增强层图片的解码的可伸缩视频解码器。基于每个层中使用的高阶并行工具完成图片区域解码。基础层解码器和增强层解码器可彼此并行操作并且还与实际层并行。基础层与增强层之间的每个层内的并行量可不同。此外，描述了规定适当设置具体位流的并行解码环境的信令。

[0386] 作为一般性提示，应注意如下：上述实施方式描述了解码器并且相应地根据各个方面设计编码器。应注意，只要涉及这些方面，则其全部可能具有共同点：解码器和编码器支持WPP和/或片并行处理并且相应地描述与其有关的细节，这些细节应被视为同时适用于任何其他方面及其对应的描述，以引导这些其他方面的新实施方式或者补足这些其他方面的实施方式的描述，与关于使用术语“部分”、“空间区段”等(而非关于更为一般性表示的图片的并行处理区段的片/子流)描述的相应方面(相应描述被转移至该相应方面)无关。同样适用于关于设定细分图片的可能方式的编码/预测参数和描述的细节：全部方面可被实现为产生解码器/编码器，解码器/编码器通过确定LCU/CTB的单元内的片和/或子流使用LCU/CTB的细分。因此，通过任一这些方面，LCU/CTB通过使用上面参考各方面及其实施方式的子集而描述的递归式多树细分可被进一步细分成编码块。此外或者可替代地，从参考这些方面描述的切片与子流/片之间的关系，所有方面的实施方式可采用切片概念。

[0387] 尽管在装置的上下文中已经描述了一些方面，然而，令人清晰的是，这些方面还代表对对应方法的描述，其中，块或者设备对应于方法步骤或者方法步骤的特征。类似地，在方法步骤的上下文中描述的各方面还代表了对应块的描述或者对应装置的项或者特征。通过(或者使用)硬件装置可执行方法步骤中的一些或者全部，例如，微处理器、可编程计算机、或者电子电路。在一些实施方式中，通过该装置可执行一个或者多个最为重要的方法步骤中的一些。

[0388] 根据一种特定的实现方式规定，可以在硬件或者在软件中实现本发明的实施方式。使用数字存储介质可以执行该实现方式，例如，软盘、DVD、蓝光光盘、CD、ROM、PROM、EPRO、EEPROM或者闪存，数字存储介质具有存储在其上的电可读控制信号，与可编程计算机系统协作(或者能够协作)，以使得执行相应的方法。因此，数字存储介质可以是计算机可读的。

[0389] 根据本发明的一些实施方式包括具有电可读控制信号的数据载体，数据载体能够与可编程的计算机系统协作，以使得执行此处所描述的方法中的一种。

[0390] 通常，本发明的实施方式可被实现为具有程序代码的计算机程序产品，当计算机程序产品在计算机上运行时，程序代码可操作为执行方法中的一种。例如，程序代码可被存储在机器可读载体上。

[0391] 其他实施方式包括存储在机器可读载体上的用于执行本文中所描述的方法中的一种的计算机程序。

[0392] 因此,换言之,本发明方法的实施方式是计算机程序,当计算机程序在计算机上运行时,计算机程序具有用于执行本文中所描述的方法中的一种的程序代码。

[0393] 因此,本发明方法的进一步实施方式是数据载体(或者数字存储介质、或者计算机可读介质),数据载体包括其上记录的用于执行本文中所描述的方法中的一种的计算机程序,数据载体、数字存储介质、或者记录介质通常是易失性和/或非易失性的。

[0394] 因此,本发明方法的进一步实施方式是代表用于执行本文中所描述的方法中的一种的计算机程序的数据流或者信号序列。例如,数据流或者信号序列可被配置为经由数据通信连接(例如,经由互联网)被传输。

[0395] 进一步的实施方式包括被配置为或者被适配成执行本文中所描述的方法中的一种的处理装置(例如,计算机)或者可编程逻辑设备。

[0396] 进一步实施方式包括计算机,该计算机具有安装在其上的用于执行本文中所描述的方法中的一种的计算机程序。

[0397] 根据本发明的进一步实施方式包括被配置为将用于执行本文中所描述的方法中的一种的计算机程序(例如,通过电或者光学)传输至接收器的装置或者系统。例如,接收器可以是计算机、移动设备、内存设备等。例如,该装置或者系统可包括用于将计算机程序传输至接收器的文件服务器。

[0398] 在一些实施方式中,可以使用可编程的逻辑设备(例如,场可编程门阵列)执行本文中所描述的方法的功能中的一些或者全部。在一些实施方式中,场可编程门阵列可与微处理器协作,以执行本文中所描述的方法中的一种。总之,优选地,通过任一硬件装置执行方法。

[0399] 使用硬件装置、或者使用计算机、或者使用硬件装置与计算机的组合可以实现本文中所描述的装置。

[0400] 使用硬件装置、或者使用计算机、或者使用硬件装置与计算机的组合可以执行本文中所描述的方法。

[0401] 上述所述实施方式仅示出了本发明的原理。应当理解的是,本文中所描述的布置的变形和改造以及细节对本领域技术人员显而易见,因此,其旨在仅受悬而未决的专利权利要求的范围的限制,而不受通过此处的实施方式的描述和说明而提出的具体细节的限制。

[0402] 根据本发明,可以包括以下各项:

[0403] 1. 一种视频解码器,用于将多层视频数据流(40)解码成使用从第一层至第二层的层间预测而被编码在层的层级中的情景,所述视频解码器支持在将所述层的图片(12,15)被细分而成的空间区段(80)中对所述多层视频数据流并行解码,其中,所述解码器被配置为:

[0404] 检测所述多层视频数据流的长期语法元素结构(606;例如, tile\_boundaries\_aligned\_flag)以

[0405] 解译所述长期语法元素结构,所述长期语法元素结构假设第一可能值集合之外的值(例如, tile\_boundaries\_aligned\_flag=1)作为对在预定时间段(608)内将所述第二层的图片(15)细分,使得所述第二层的图片的空间区段之间的边界与所述第一层的图片(12)的空间区段的每一个边界重叠,并且在比所述预定时间段短的时间间隔(604)内基于所述

多层视频数据流的短期语法元素(602;例如, column\_width\_minus1[i]和column\_width\_minus1[i])周期性地确定将所述第一层的图片和所述第二层的图片细分成空间区段的保证;并且

[0406] 如果所述长期语法元素结构假设了第二可能值集合之外的值(例如, tile\_boundaries\_aligned\_flag=0),则在比所述预定时间段短的时间间隔内从所述多层视频数据流的所述短期语法元素周期性地确定将所述层的图片细分成空间区段,使得至少对于所述短期语法元素的第一可能值,所述第二层的图片的空间区段之间存在不与所述第一层的空间区段的边界中的任一个重叠的边界,并且至少对于所述短期语法元素的第二可能值,所述第二层的图片的空间区段之间存在与所述第一层的空间区段的每一个边界重叠的边界。

[0407] 2.根据项1所述的视频解码器,其中,所述视频解码器被配置为:

[0408] 使用帧内图片空间预测对所述层的图片进行解码,并且用于各个空间区段的所述帧内图片空间预测在相应空间区段的边界线处中断;或者

[0409] 通过将所述第一层的图片的空间区段并行解码且支持跨所述第一层的图片的空间区段的边界线的所述帧内图片空间预测并且服从所述第一层的图片的空间区段的解码之间的解码延迟;以及通过将所述第二层的图片的空间区段并行解码且支持跨所述第二层的图片的空间区段的边界线的所述帧内图片空间预测并且服从所述第二层的图片的空间区段的解码之间的解码延迟,来使用所述帧内图片空间预测对所述层的图片进行解码。

[0410] 3.根据项1或2所述的视频解码器,支持在将所述层的图片细分而成的片中对所述多层视频数据流进行片并行解码,其中,所述解码器被配置为:

[0411] 解译所述长期语法元素,所述长期语法元素假设所述第一可能值集合之外的值作为对在预定时间段内将所述第二层的图片细分,使得所述第二层的图片的片之间的边界与所述第一层的片的每一个边界重叠,并且在比所述预定时间段短的时间间隔内基于所述短期语法元素周期性地确定所述第二层相对于所述第一层的图片的细分的片细化的保证;以及

[0412] 如果所述长期语法元素假设了所述第二可能值集合之外的值,则在比所述预定时间段短的时间间隔内从所述多层视频数据流的所述短期语法元素周期性地确定将所述层的图片细分成片,使得至少对于所述短期语法元素的第一可能值,所述第二层的图片的片之间存在不与所述第一层的片的边界中的任一个重叠的边界,并且至少对于所述短期语法元素的第二可能值,所述第二层的图片的片之间存在与所述第一层的片的每一个边界重叠的边界。

[0413] 4.根据项3所述的视频解码器,其中,所述视频解码器被配置为:

[0414] 使用所述帧内图片空间预测对所述层的图片进行解码,且用于各个片的所述帧内图片空间预测在相应片的边界线处中断。

[0415] 5.根据项1或2所述的视频解码器,其中,所述解码器被配置为:

[0416] 解译所述长期语法元素结构,所述长期语法元素结构假设第一可能值集合之外的值作为对在预定时间段内将所述第二层的图片细分,使得所述第一层的图片的各个空间区段精确地由所述第二层的图片的n个空间区段构成的保证,并且n取决于所述长期语法元素结构的值;并且

[0417] 如果将所述长期语法元素设置成第二可能值集合之外的值,则在比所述预定时间段短的时间间隔内基于所述多层视频数据流的所述短期语法元素周期性地确定所述预定时间段内的层间偏移量。

[0418] 6. 根据项1或2所述的视频解码器,其中,所述解码器被配置为根据是否假设了所述第一可能值之外的值的所述长期语法元素结构来决定开始或者不开始尝试对所述多层视频数据流的所述第二层进行解码。

[0419] 7. 根据项1至6中任一项所述的视频解码器,其中,所述视频解码器是混合式视频解码器。

[0420] 8. 一种视频编码器,用于利用从第一层至第二层的层间预测在层的层级中将场景编码成多层视频数据流,使得所述多层视频数据流能够在将所述层的图片细分而成的空间区段中被并行解码,其中,所述编码器被配置为:

[0421] 将长期语法元素结构(606)和短期语法元素(602)插入到所述多层视频数据流中,所述短期语法元素限定在时间间隔内将所述第一层的图片和所述第二层的图片细分成空间区段;并且

[0422] 在将所述长期语法元素结构设定为以下的值之间进行切换:

[0423] 第一可能值集合之外的值,且在比所述时间间隔大的预定时间段(608)内将所述短期语法元素设置成可能设置的集合之外的适当子集,所述适当子集被选择使得在所述预定时间段内将所述第二层的图片细分,使得所述第二层的图片的空间区段之间的边界与所述第一层的空间区段的每一个边界重叠;或者

[0424] 第二可能值集合之外的值,且在所述预定时间段内将所述短期语法元素设置成所述可能设置的集合中的任一个,所述可能设置的集合包含至少一种设置和至少另一种设置,根据所述至少一种设置,所述第二层的图片的空间区段之间存在不与所述第一层的空间区段的边界中的任一个重叠的边界,并且根据所述至少另一种设置,所述第二层的图片的空间区段之间存在与所述第一层的空间区段的每一个边界重叠的边界。

[0425] 9. 根据项8所述的视频编码器,其中,所述视频编码器被配置为:

[0426] 使用帧内图片空间预测对所述层的图片进行编码,且用于各个空间区段的所述帧内图片空间预测在相应空间区段的边界线处中断;或者

[0427] 通过支持跨所述第一层的图片的空间区段的边界线的所述帧内图片空间预测并且对用于将所述第一层的图片的空间区段的子集单独地进行熵编码的熵上下文概率进行初始化,或者根据该子集之间的顺序在采用处于中间适配状态的所述第一层的图片的空间区段的之前子集的所述熵上下文概率的基础上;以及通过支持跨所述第二层的图片的空间区段的边界线的所述帧内图片空间预测并且对用于将所述第二层的图片的空间区段中的子集单独地进行熵编码的所述熵上下文概率初始化,或者根据该子集之间的顺序在采用处于中间适配状态的所述第二层的图片的空间区段的之前子集的所述熵上下文概率的基础上,使用所述帧内图片空间预测并利用适配所述熵上下文概率的熵编码对所述层的图片进行编码。

[0428] 10. 根据项8或9所述的视频编码器,其中,所述空间区段是片,并且

[0429] 所述编码器被配置为:

[0430] 如果将所述长期语法元素结构设置为:

[0431] 所述第一可能值集合之外的值(tile\_boundaries\_aligned\_flag=1),则在比所述时间间隔大的预定时间段内将所述短期语法元素设置成可能设置的集合之外的适当子集,所述适当子集被选择使得在所述预定时间段内使被细分成片的所述第二层的图片与被细分成片的所述第一层的图片相一致,或者使被细分成片的所述第一层的图片细化;

[0432] 或者

[0433] 第二可能值集合之外的值(tile\_boundaries\_aligned\_flag=0),则在所述预定时间段内将所述短期语法元素设置成所述可能设置的集合中的任一个,使得至少对于所述预定时间段内的一个时间间隔,将所述短期语法元素设置成所述可能设置的集合的第一可能值,根据所述第一可能值,所述第二层的图片的片之间存在不与所述第一层的片的边界中的任一个重叠的边界;并且至少对于所述预定时间段内的另一时间间隔,将所述短期语法元素设置为所述可能设置的集合中的第二可能值,根据所述第二可能值,所述第二层的图片的片之间存在与所述第一层的片的每一个边界重叠的边界。

[0434] 11.根据项10所述的视频编码器,其中,所述视频编码器被配置为利用所述帧内图片空间预测对所述层的图片进行编码,且用于各个片的所述帧内图片空间预测在相应片的边界线处中断。

[0435] 12.根据项8或9所述的视频编码器,其中,所述编码器被配置为:

[0436] 如果将所述长期语法元素结构设置为:

[0437] 所述第一可能值集合之外的值,则使用比所述时间间隔大的预定时间段将所述短期语法元素设置为所述可能设置的集合之外的适当子集,所述适当子集被选择使得在所述预定时间段内,所述第一层的图片的各个空间区段精确地由所述第二层的图片的n个空间区段构成,且n取决于所述长期语法元素结构的值。

[0438] 13.一种用于将空间可伸缩位流(40)解码成图片的解码器,所述图片被编码在不同的空间层中,并且对于所述空间层中的至少一个来说所述图片被编码在第一空间区段中,其中,所述解码器被配置为:

[0439] 对第一空间层的图片(12)进行上采样以获得上采样参考图片,并且使用所述上采样参考图片预测第二空间层的图片(15),其中,

[0440] 所述解码器响应于所述空间可伸缩位流中的语法元素(616;例如,independent\_tile\_upsampling\_idc)以根据所述语法元素插入(620)

[0441] 所述第一空间层的图片,

[0442] 使得依赖于所述第一空间区段的所述上采样参考图片的分区(622)的任一部分独立于与所述分区的其他部分中的任一个覆盖的所述第一空间层的图片的部分;或者

[0443] 使得所述上采样参考图片的所述分区(622)的任一部分依赖于由空间上与相应部分相邻的所述分区的另一部分覆盖的所述第一空间层的图片的部分。

[0444] 14.根据项13所述的解码器,其中,所述解码器被配置为并行解码所述不同的空间层。

[0445] 15.根据项13或14所述的解码器,其中,所述解码器响应于所述空间可伸缩位流中的所述语法元素(616)以根据所述语法元素插入(620)所述第一空间层的图片,

[0446] 使得由所述第一空间区段中的任一个在空间上覆盖的所述上采样参考图片中的任一部分独立于与由所述第一空间区段中的另一个覆盖的所述第一空间层的图片中的部

分;或者

[0447] 使得由所述第一空间区段中的任一个在空间上覆盖的所述上采样参考图片的任一部分依赖于由所述第一空间区段中与相应空间区段相邻的另一空间区段覆盖的所述第一空间区段的图片的部分。

[0448] 16.根据项13至15中任一项所述的解码器,其中,所述空间可伸缩位流具有在所述第一空间区段中被编码成所述空间可伸缩位流的所述第二空间层的图片。

[0449] 17.根据项13至16中任一项所述的解码器,其中,所述解码器被配置为使用帧内图片空间预测执行解码,

[0450] 并且用于各个第一空间区段的所述帧内图片空间预测在相应的第一空间区段的边界线处中断;或者

[0451] 并且支持跨所述第一空间区段的边界线的所述帧内图片空间预测,使用熵上下文概率的自适应性将各个第一空间区段熵解码并且将所述第一空间区段的所述熵上下文概率初始化,所述第一空间区段的所述熵上下文概率独立于任一其他的第一空间区段,或者根据所述第一空间区段之中的顺序适配之前的第一空间区段的所述熵上下文概率以适配达到所述之前的第一空间区段的中间位置。

[0452] 18.根据项13或14所述的解码器,其中,所述空间可伸缩位流具有在所述第一空间区段中被编码成所述空间可伸缩位流的所述第一空间层的图片,其中,所述空间可伸缩位流具有在第二空间区段中被编码成所述空间可伸缩位流的所述第二空间层的图片,其中,所述解码器响应于所述空间可伸缩位流中的所述语法元素(606)以根据所述语法元素内插所述第一空间层的图片

[0453] (例如,independent\_tile\_upsampling\_idc=2),使得由第二片中的任一个在空间上覆盖的所述上采样参考图片的任一部分独立于由所述第二片中的另一个在空间上覆盖的所述第一空间层的图片中的部分;或者

[0454] (例如,independent\_tile\_upsampling\_idc=1),使得由第一片和所述第二片的空间共置边界线在空间上限定的所述上采样参考图片的任一部分独立于与由所述第一片和所述第二片的空间共置边界线在空间上限定的由所述上采样参考图片的另一部分覆盖的所述第一空间层的图片的部分;或者

[0455] (例如,independent\_tile\_upsampling\_idc=0),使得由所述第二片中的任一个在空间上覆盖的所述上采样参考图片中的任一部分依赖于与由与相应片相邻的另一第二片覆盖的所述第一空间层的图片中的部分。

[0456] 19.根据项13至18中任一项所述的解码器,其中,所述解码器被配置为使用回退规则填充在内插所述第一空间层的图片中使用的滤波核的片段,所述滤波核伸出至所述第一空间层的图片的任一部分中、被所述分区的其他分区中的任一个覆盖,根据所述回退规则,独立于所述滤波核伸出至其中的所述第一空间层的所述图片的所述相应部分填充所述片段,以实现独立性。

[0457] 20.根据项19所述的解码器,其中,所述解码器被配置为还在填充伸出所述第一空间层的图片的外边界线的所述滤波核的片段时使用所述回退规则。

[0458] 21.根据项13至20中任一项所述的解码器,其中,所述解码器是视频解码器并且所述解码器被配置为以每个图片或者每个图片序列为基础响应于所述空间可伸缩位流中的

所述语法元素(606)。

[0459] 22.根据项13至21中任一项所述的解码器,其中,所述空间可伸缩位流具有在所述第一空间区段中被编码成所述空间可伸缩位流的所述第一空间层的图片,其中,所述空间可伸缩位流具有在第二空间区段中被编码成所述空间可伸缩位流的所述第二空间层的图片,其中,

[0460] 所述分区的边界对应于所述第一空间区段与所述第二空间区段的边界或者所述第二空间区段的边界的空间重叠的逻辑与,其中,所述解码器根据所述语法元素响应于所述空间可伸缩位流中的所述语法元素(606)以填充在所述第一空间层的图片的内插中使用的滤波核的片段,所述滤波核使用回退规则从一个分区伸出至所述分区的相邻分区中,根据所述回退规则独立于所述滤波核伸出至其中的所述第一空间层的图片的相应部分、或者使用所述滤波核伸出至其中的所述第一空间层的图片的相应部分填充所述片段。

[0461] 23.根据项13至22中任一项所述的解码器,其中,所述解码器被配置为使用依赖于所述语法元素(606)的层间偏移量将所述第一层和所述第二层并行解码。

[0462] 24.根据项13或23所述的解码器,其中,所述解码器被配置为根据所述语法元素放置所述分区的边界,以对应于所述第一空间区段和所述第二空间区段的边界或者所述第二空间区段的边界的空间重叠的逻辑与。

[0463] 25.一种编码器,用于在不同空间层中将图片编码成空间可伸缩位流并且对于所述空间层中的至少一个在第一空间区段中对所述图片进行编码,其中,所述编码器被配置为:

[0464] 对第一空间层的图片进行上采样以获得上采样参考图片并且使用所述上采样参考图片预测第二空间层的图片,其中,所述编码器被配置为根据语法元素设置并且将语法元素(606)插入到所述空间可伸缩位流中并且内插所述第一空间层的图片,

[0465] 使得依赖于所述第一空间区段的所述上采样参考图片的分区的任一部分独立于由所述分区的其他部分中的任一个覆盖的所述第一空间层的图片的部分;或者

[0466] 使得所述上采样参考图片的所述分区的任一部分依赖于由在空间上与相应部分相邻的所述分区的另一部分覆盖的所述第一空间层的图片的部分。

[0467] 26.根据项25所述的解码器,其中,所述编码器被配置为根据所述语法元素设置并且将所述语法元素插入到所述空间可伸缩位流中并且内插所述第一空间层的图片,

[0468] 使得由所述第一空间区段中的任一个空间覆盖的所述上采样参考图片的任一部分独立于由所述第一空间区段的任意其它部分覆盖的所述第一空间层的图片的部分;或者

[0469] 使得由所述第一空间区段中的任一个空间覆盖的所述上采样参考图片的任一部分依赖于由与相应的第一空间区段相邻的所述第一空间区段中的另一空间区段覆盖的所述第一空间层的图片的部分。

[0470] 27.根据项25或26所述的编码器,其中,所述编码器被配置为将所述第一空间层的图片编码成所述第一空间区段中的所述空间可伸缩位流。

[0471] 28.根据项27所述的编码器,其中,所述编码器被配置为使用帧内图片空间预测对所述第一空间层的图片进行编码,且用于各个第一空间区段的所述帧内图片空间预测在相应第一空间区段的边界线处中断。

[0472] 29.根据项27或28所述的编码器,其中,所述编码器被配置为将所述第二空间层的

图片在第二空间区段中编码成所述空间可伸缩位流,其中,所述编码器被配置为根据所述语法元素设置并且将所述语法元素插入到所述空间可伸缩位流中并且内插所述第一空间层的图片,

[0473] 使得由所述第二空间区段中的任一个在空间上覆盖的所述上采样参考图片的任一部分独立于由所述第二空间区段中的另一个在空间上覆盖的所述第一空间层的图片的部分;或者

[0474] 使得由所述第一空间区段和所述第二空间区段的空间共置边界线在空间上限定的所述上采样参考图片的任一部分独立于由所述第一空间区段和所述第二空间区段的空间共置边界线在空间上限定的、由所述上采样参考图片的另一分区覆盖的所述第一空间层的图片的部分;或者

[0475] 使得由所述第二空间区段中的任一个在空间上覆盖的所述上采样参考图片的任一部分依赖于由与相应空间区段相邻的另一第二空间区段覆盖的所述第一空间层的图片的部分。

[0476] 30.根据项29所述的编码器,其中,所述编码器被配置为使用帧内图片空间预测对所述第二空间层的图片编码,且用于各个第二空间区段的所述帧内图片空间预测在相应第二空间区段的边界线处中断。

[0477] 31.根据项25至30中任一项所述的编码器,其中,所述编码器被配置为在内插所述第一空间层的图片时使用滤波核的填充片段以实现与所述第一空间层的图片中的任一部分的独立性,所述滤波核使用回退规则伸出至所述第一空间层的图片的任一部分中,根据所述回退规则,独立于所述滤波核伸出至其中的所述第一空间层的图片的相应部分填充所述片段。

[0478] 32.根据项31所述的编码器,其中,所述编码器被配置为还在填充伸出所述第一空间层的图片的外边界线的所述滤波核的片段时使用所述回退规则。

[0479] 33.根据项25至32中任一项所述的编码器,其中,所述编码器是视频编码器并且所述编码器被配置为以每个图片或者每个图片序列为基础设置并且将所述语法元素插入到所述空间可伸缩位流中。

[0480] 34.一种视频解码器,用于将多层视频数据流解码成使用从第一层至第二层的层间预测而被编码在层的层级中的情景,所述视频解码器利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历之间的层间延迟,以时间重叠方式顺次遍历空间区段支持在将所述层的图片分区而成的空间区段中并行解码所述多层视频数据流,所述视频解码器被配置为:

[0481] 检测所述多层视频数据流的长期语法元素结构(642;例如,min\_spatial\_segment\_delay)使得:

[0482] 如果将所述长期语法元素结构(例如,min\_spatial\_segment\_delay≠0)设置成可能值的第一集合中的值,则使用所述长期语法元素的值预先确定预定时间段内的层间偏移量,并且在比所述预定时间段短的时间间隔内基于所述多层视频数据流的所述短期语法元素(602)分别周期性地确定所述第一层的图片的空间区段和所述第二层的图片的空间区段的大小和定位以及所述第一层的图片与所述第二层的图片的空间采样分辨率;

[0483] 如果将所述长期语法元素设置成与可能值的所述第一集合不相交的可能值的第

二集合(例如,min\_spatial\_segment\_delay=0)中的值,则基于所述多层视频数据流的所述短期语法元素在比所述预定时间段短的时间间隔内周期性地确定所述预定时间段内的层间偏移量。

[0484] 35.根据项34所述的视频解码器,其中,所述视频解码器被配置为利用相同图片的紧随的连续子流的遍历之间的帧内图片子流间延迟以及所述第一层的图片的子流的遍历相对于所述第二层的图片的子流的遍历的层间偏移量,通过以时间重叠方式顺次遍历子流的波前并行处理将所述多层视频数据流并行解码成将所述层的图片分区而成的并且由所述层的图片被规则细分而成的块的行构成的子流。

[0485] 36.根据项35所述的视频解码器,其中,所述视频解码器被配置为:

[0486] 将所述子流并行解码且支持跨所述子流的边界线的所述帧内图片空间预测。

[0487] 37.根据项34所述的视频解码器,其中,所述视频解码器被配置为将所述多层视频数据流解码成将层的图片分区而成的片,且在所述第一层和所述第二层的图片的每一个内以所述片之间的片顺序对所述片进行遍历,并且利用所述第一层的图片的片的遍历相对于所述第二层的图片的片的遍历的层间偏移量将所述第一层的图片的紧随的连续片和所述第二层的图片的的紧随的连续片来对所述层的图片分区而成的连续片并行解码。

[0488] 38.根据项37所述的视频解码器,其中,所述视频解码器被配置为:

[0489] 使用帧内图片空间预测将所述第一层的图片和所述第二层的图片解码,且用于各个片的所述帧内图片空间预测在相应片的边界线处中断。

[0490] 39.根据项34至38中任一项所述的视频解码器,其中,所述视频解码器被配置为在通过使用所述长期语法元素的值作为用于在所述第一层的图片的空间区段的单元中的所述层间偏移量的测量值来确定所述层间偏移量时使用所述长期语法元素的值。

[0491] 40.根据项34至39中任一项所述的视频解码器,其中,所述视频解码器被配置为在通过使用所述长期语法元素的值作为所述第一层的图片的空间区段的数目来确定所述层间偏移量时使用所述长期语法元素的值,通过该值解码所述第二层的图片的第一空间区段相对于开始所述第一层的图片的解码和遍历被延迟。

[0492] 41.一种视频编码器,用于使用从第一层至第二层的层间预测在层的层级中被编码的情景编码成多层视频数据流,使得能够利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历之间的层间偏移量,通过以时间重叠方式顺次遍历空间区段在将所述层的图片分区而成的空间区段中对所述多层视频数据流并行解码,所述视频编码器被配置为:

[0493] 设置长期语法元素结构(min\_spatial\_segment\_delay)和短期语法元素并且将所述长期语法元素结构和所述短期语法元素插入到所述多层视频数据流中,所述多层视频数据流的所述短期语法元素分别周期性地限定所述第一层的图片的空间区段和所述第二层的图片的空间区段在时间间隔内的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率;

[0494] 其中,所述编码器被配置为在下列设置之间切换:

[0495] 将所述长期语法元素结构(min\_spatial\_segment\_delay≠0)设置为可能值的第一集合中的值,且该值在比所述时间间隔大的预定时间段内发出所述层间偏移量的信令,且在所述预定时间段内将所述短期语法元素设置成可能设置的集合之外的适当子集,所述

适当子集被选择使得在所述预定时间段内利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历之间的、小于或者等于通过所述长期语法元素发出信号的所述层间偏移量的实际层间偏移量,所述第一层的图片的空间区段和所述第二层的图片的空间区段的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率分别能够通过以时间重叠方式顺次遍历空间区段来支持所述多层视频数据流的解码;

[0496] 将所述长期语法元素设置成与所述可能值的第一集合不相交的可能值的第二集合中的值( $\text{min\_spatial\_segment\_delay}=0$ ),且在所述预定时间段内将所述短期语法元素设置成可能设置的集合中的任一个,所述可能设置的集合包含至少一种设置和至少另一种设置,根据所述至少一种设置,利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历之间的、小于或者等于通过所述长期语法元素发出信号的所述层间偏移量的实际层间偏移量,所述第一层的图片的空间区段和所述第二层的图片的空间区段的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率分别禁用通过以时间重叠方式顺次遍历空间区段而将所述多层视频数据流解码,并且根据所述至少另一种设置,利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历之间的、小于或者等于通过所述长期语法元素发出信号的所述层间偏移量的实际层间偏移量,所述第一层的图片的空间区段和所述第二层的图片的空间区段的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率分别能够通过以时间重叠方式顺次遍历空间区段来将所述多层视频数据流解码。

[0497] 42.根据项41所述的视频编码器,其中,所述视频编码器被配置为执行编码,使得所述空间区段是所述层的图片被分区而成的子流,并且

[0498] 所述子流由所述层的图片被规则细分成的块的行构成,通过这种方式,利用相同图片的中间连续子流的遍历之间的帧内图片子流间延迟以及所述第一层的图片的子流的遍历相对于所述第二层的图片的子流的遍历之间的所述层间偏移量,允许通过以时间重叠方式顺次遍历所述子流而使用波前并行处理将所述多层视频数据流并行解码成所述子流。

[0499] 43.根据项42所述的视频编码器,其中,所述视频编码器被配置为通过支持跨所述子流的边界线的所述帧内图片空间预测并且将熵上下文概率初始化,所述熵上下文概率用于将所述子流单独地进行熵编码,或者在根据所述子流之间的顺序采用处于中间自适配状态的前一子流的所述熵上下文概率的基础上使用所述帧内图片空间预测并利用适配的熵上下文概率的熵编码对所述子流进行编码。

[0500] 44.根据项41所述的视频编码器,其中,所述视频编码器被配置为执行编码,使得所述空间区段是所述层的图片被分区而成的片,通过这种方式允许通过在所述第一层和所述第二层的图片中的每个内以所述片之间的片顺序对所述片进行遍历来将所述多层视频数据流解码成所述片,并且利用所述第一层的图片的片的遍历相对于所述第二层的所述图片的片的遍历的所述层间偏移量将所述第一层的图片的紧随的连续片与所述第二层的图片的紧随的连续片并行解码。

[0501] 45.根据项44所述的视频编码器,其中,所述视频编码器被配置为使用帧内图片空间预测将所述第一层的图片与所述第二层的图片编码,

[0502] 且用于各个片的所述帧内图片空间预测在相应片的边界线处中断。

[0503] 46.根据项41至45中任一项所述的视频编码器,其中,所述视频编码器被配置使得

所述长期语法元素的值限定在所述第一层的图片的空间区段的单元中的所述层间偏移量的测量。

[0504] 47. 根据项41至46中任一项所述的视频解码器, 其中, 所述视频编码器被配置为设置所述长期语法元素的值以发出所述第一层的图片的空间区段的数目的信号, 通过该值使所述第二层的所述图片的第一空间区段的解码相对于开始所述第一层的所述图片的解码和遍历被延迟。

[0505] 48. 一种用于将多层视频数据流处理成情景的网络实体, 所述情景被编码在层中, 使得在各个层中, 所述情景被编码在被可伸缩维度横跨的可伸缩空间的不同操作点中, 其中, 所述多层视频数据流由第一NAL单元和第二NAL单元构成, 所述第一NAL单元中的每个与所述层中的一个相关联, 并且所述第二NAL单元散布在所述第一NAL

[0506] 单元内并且呈现关于所述多层视频数据流的整体信息, 所述网络实体被配置为:

[0507] 检测所述第二NAL单元中的类型指示符字段(609; 例如, `dedicated_scalability_ids_flag`) ;

[0508] 如果所述类型指示符字段具有第一状态(例如, `dedicated_scalability_ids_flag=0`), 则从所述第二NAL单元读取将第一NAL单元报头中的层指示符字段(例如, `layer_id`)的可能值映射至所述操作点的映射信息(例如, `layer_id_in_nuh[i]`, `dimension_id[i][j]`)并且通过所述层指示符字段和所述映射信息使所述第一NAL单元与所述第一NAL单元中的所述操作点相关联;

[0509] 如果所述类型指示符字段具有第二状态(`dedicated_scalability_ids_flag=1`), 则通过将所述第一NAL单元中的所述层指示符字段分割成一个以上的部分并且通过使用该部分的值作为所述可伸缩空间内的矢量的坐标来定位所述第一NAL单元的所述操作点使所述第一NAL单元与所述操作点相关联。

[0510] 49. 根据项48所述的网络实体, 其中, 所述网络实体被配置为:

[0511] 如果所述类型指示符字段具有所述第二状态(`dedicated_scalability_ids_flag=1`), 则根据所述第二NAL单元中的语法元素(`dimension_id_len_minus1`)通过将所述第一NAL单元中的所述层指示符字段划分成一个以上的部分, 通过使用该部分的值作为所述可伸缩空间内的矢量的坐标来定位所述第一NAL单元的所述操作点, 并且根据所述第二NAL单元中的另外的语法元素(`scalability_mask`)在语义上确定所述可伸缩维度使所述第一NAL [0512] 单元与所述操作点相关联。

[0513] 50. 根据项48或49所述的网络实体, 其中, 所述网络实体被配置为:

[0514] 如果所述类型指示符字段具有所述第一状态(`dedicated_scalability_ids_flag=0`), 则从所述第二NAL单元中的另外的语法元素(`scalability_mask`)确定所述可伸缩维度的数目p和语义含义并且通过从所述第二NAL单元读取p维矢量的列表(708)

[0515] 使所述层指示符字段的可能值与所述操作点相关联。

[0516] 51. 根据项50所述的网络实体, 其中, 所述网络实体被配置为如果所述类型指示符字段具有所述第二状态, 则跳过从所述第二NAL单元读取所述列表。

[0517] 52. 根据项49或51中任一项所述的网络实体, 其中, 所述网络实体被配置为在不考虑具有所述第一状态或者所述第二状态的所述类型指示符字段的情况下从所述第二NAL单元读取所述另外的语法元素;

[0518] 并且使得在不考虑具有所述第一状态或者所述第二状态的所述类型指示符字段的情况下,所述层指示符字段的大小相同。

[0519] 53.根据项48或52中任一项所述的网络实体,其中,所述网络实体包括视频解码器。

[0520] 54.一种用于在层中将情景编码成多层视频数据流的视频编码器,使得在各个层中,所述情景被编码在被可伸缩维度横跨的可伸缩空间的不同操作点中,其中,所述多层视频数据流由第一NAL单元和第二NAL单元构成,所述第一NAL单元中的每个与所述层中的一个相关联,并且所述第二NAL单元散布在所述第一NAL单元内并且呈现关于所述多层视频数据流的整体信息,所述视频编码器被配置为:

[0521] 将类型指示符字段插入到所述第二NAL单元中并且在下列设

[0522] 置之间切换:

[0523] 设置所述类型指示符字段使得所述类型指示符字段具有第一状态,且将第一NAL单元报头中的层指示符字段的可能值映射至所述操作点的映射信息插入到所述第二NAL单元中并且设置所述第一NAL单元中的所述层指示符字段使得通过所述映射信息使所述第

[0524] 一NAL单元的操作点与相应的层指示符字段相关联;

[0525] 设置所述类型指示符字段使得所述类型指示符字段具有第二状态 (dedicated\_scalability\_ids\_flag=1),且通过将所述第一NAL单元中的所述层指示符字段分割成一个以上的部分并且设置所述一个以上部分使得该部分的值对应于所述可伸缩空间内的矢量的坐标从而指向与相应的第一NAL单元相关联的操作点,来设置所述第一NAL

[0526] 单元中的所述层指示符字段。

[0527] 55.根据项54所述的视频编码器,其中,所述视频编码器被配置为:

[0528] 当设置所述类型指示符字段使得所述类型指示符字段具有所述第二状态时,设置并且将语法元素插入到所述第二NAL单元中,所述语法元素限定将所述第一NAL单元中的所述类型指示符字段划分成所述一个以上部分,并且设置且将另外的语法元素插入到所述第二NAL单元中,所述另外的语法元素在语义上限定所述可伸缩维度。

[0529] 56.一种将多层视频数据流转换成在层中编码的情景,使得在各个层中,所述情景被编码在被可伸缩维度横跨的可伸缩空间的不同操作点中,其中,所述多层视频数据流由第一NAL单元和第二NAL单元构成,所述第一NAL单元中的每个与所述层中的一个相关联,并且

[0530] 所述第二NAL单元散布在所述第一NAL单元内并且呈现关于所述多层视频数据流的整体信息,其中,根据下列条件在所述第二NAL

[0531] 单元中呈现类型指示符字段(696;例如,dedicated\_scalability\_ids\_flag);

[0532] 如果所述类型指示符字段具有第一状态(例如,dedicated\_scalability\_ids\_flag=0),则所述第二NAL单元中的映射信息将第一NAL单元报头中的层指示符字段(例如,layer\_id)的可能值映射至所述操作点;

[0533] 如果所述类型指示符字段具有第二状态 (dedicated\_scalability\_ids\_flag=1),则将所述第一NAL单元中的所述层指示符字段分割成一个以上的部分,且所述第一NAL单元的所述操作点被该部分的值限定为所述可伸缩空间内的矢量的坐标。

[0534] 57.一种传输层解码器,用于使多层视频数据流转换成在层中被编码的情景以通

过多标准多层解码器被解码,其中,所述多层视频数据流由NAL单元构成,所述NAL单元中的每个与所述层中的一个相关联,其中,所述层与不同的编解码器相关联,使得对于每个层使用与相应层相关联的所述编解码器对与相应层相关联的所述NAL单元进行编码,所述传输层解码器被配置为:

[0535] 对于每个NAL单元,识别与NAL单元相关联的相同编解码器,并且将所述多层视频数据流的所述NAL单元移交至所述多标准多层解码器,所述多标准多层解码器使用与不同编解码器相关联的层之间的层间预测对所述多层视频数据流进行解码。

[0536] 58.根据项57所述的视频解码器,进一步被配置为:

[0537] 使用具有被设置成指示相应层的所述编解码器的状态的NAL

[0538] 单元类型指示符的预定编解码器的NAL单元报头封装被识别为与不同于所述预定编解码器的任一编解码器相关联的层的NAL单元。

[0539] 59.根据项57或58所述的视频解码器,进一步被配置为:

[0540] 根据所述NAL单元分别到达的信道执行识别。

[0541] 60.根据项57或59中任一项所述的视频解码器,进一步被配置为执行移交,使得在不同的信道上将与不同的编解码器相关联的所述NAL

[0542] 单元移交至所述多标准多层解码器。

[0543] 61.根据项57或60所述的视频解码器,进一步被配置为:

[0544] 向各个NAL单元提供指示与相应NAL单元相关联的层相关联的所述编解码器的元数据。

[0545] 62.一种视频解码器,用于将多层视频数据流解码成情景,使用从第一层的部分至第二层的共置部分的层间预测将所述情景编码在层的层级中,其中,将所述第一层的图片细分成第一块的阵列并且将所述第二层的图片细分成第二块的阵列,其中,在所述第一块与所述第二块中分别限定了光栅扫描解码顺序,所述视频解码器被配置为:

[0546] 利用所述第一层的图片的第一块的遍历相对于所述第二层的图片的第二块的空间遍历之间的以所述第一块为单位测量的所述层间偏移量,根据所述多层视频数据流的语法元素结构(例如,ctb\_delay\_enabled\_flag,min\_spatial\_segment\_delay)以时间重叠方式顺次遍历所述第一块和所述第二块来确定用于将所述第一层的图片与所述第二层的图片并行解码的层间偏移量。

[0547] 63.根据项62所述的视频解码器,其中,所述语法元素结构是长期语法元素结构,并且所述视频解码器被配置为:

[0548] 在预定时间段内预先执行确定;并且

[0549] 在比所述预定时间段短的时间间隔内基于所述多层视频数据流的短期语法元素分别周期性地确定所述第一层的图片的所述第一块和所述第二层的图片的所述第二块的大小和定位以及所述第一层的图片与所述第二层的图片的空间采样分辨率。

[0550] 64.根据项63所述的视频解码器,其中,所述视频解码器利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历的所述层间偏移量,通过以时间重叠方式顺次遍历空间区段来支持在沿着所述光栅扫描解码顺序按顺序布置的空间区段将所述多层视频数据流并行解码成被分区的所述层的图片,所述视频解码器被配置为:

[0551] 检测所述多层视频数据流的所述长期语法元素结构使得:

[0552] 如果将所述长期语法元素结构(例如,ctb\_delay\_enabled\_flag=0,min\_spatial\_segment\_delay $\neq$ 0)设置成可能值的第一集合中的值,则使用所述长期语法元素结构的值预先确定预定时间段内的所述层间偏移量,从而以所述第一层的图片的空间区段为单位测量所述层间偏移量,并且在比所述预定时间段短的时间间隔内基于所述多层视频数据流的短期语法元素分别周期性地确定所述第一层的图片的空间区段与所述第二层的图片的空间区段的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率;

[0553] 如果将所述长期语法元素结构设置成与所述可能值的第一集合不相交的可能值的第二集合的值(例如,min\_spatial\_segment\_delay=0),则在比所述预定时间段短的时间间隔内基于所述多层视频数据流的所述短期语法元素周期性地确定所述预定时间段内的所述层间偏移量;以及

[0554] 如果将所述长期语法元素设置成与可能值的所述第一集合和所述第二集合不相交的可能值的第三集合的值(例如,ctb\_delay\_enabled\_flag=1,min\_spatial\_segment\_delay $\neq$ 0),则执行以所述第一块为单位确定所述层间偏移量并且分别周期性地确定所述第一层的图片的所述第一块和所述第二层的图片的所述第二块的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率。

[0555] 65.根据项63或64所述的视频解码器,其中,所述视频解码器利用相同图片的中间连续子流的遍历之间的帧内图片子流间延迟以及所述第一层的图片的子流的遍历相对于所述第二层的图片的子流的遍历之间的层间偏移量,以时间重叠方式顺次遍历子流来使用波前并行处理支持在将所述层的图片分区而成的并且由所述第一块和所述第二块的行构成的子流中对所述多层视频数据流并行解码。

[0556] 66.根据项64至65中任一项所述的视频解码器,其中,所述长期语法元素结构包括单元标记(例如,示例性地,ctb\_delay\_enabled\_flag)

[0557] 和延迟指示符(例如,示例性地,min\_spatial\_segment\_delay),其中,

[0558] 在检测所述长期语法元素结构时,所述视频解码器被配置为:

[0559] 检测所述延迟指示符以确定是否将所述延迟指示符设置为零;如果将所述延迟指示符设置为零,则确定将所述长期语法元素结构的值设置为所述第二集合中的值;并且

[0560] 如果将所述延迟指示符设置为非零值,则使用所述非零值确定所述长期语法元素结构的值,并且如果所述单元标记为零,则确定将所述长期语法元素结构的值设置成所述第一集合的值,并且如果所述单元标记是一,则确定将所述长期语法元素结构的值设置为所述第三集合中的值。

[0561] 67.根据项65至66中任一项所述的视频解码器,所述视频解码器被配置为在开始对所述第一层和所述第二层并行解码时依赖于所述层间偏移量。

[0562] 68.根据项62至67中任一项所述的视频解码器,所述视频解码器被配置为检查所述第一层的数目为s的空间区段或者编码块是否被完全解码,且数目s唯一地取决于所述语法元素结构,并且除了在所述检查显示在所述第一层的空间区段或者编码块之中至少已被完全解码了s个的情况下,在所述第一层的解码期间推迟开始所述第二层的解码。

[0563] 69.根据项62至68中任一项所述的视频解码器,所述视频解码器被配置为在开始并且完全处理所述第一层和所述第二层的并行解码时依赖于所述层间偏移量。

[0564] 70. 根据项62至69中任一项所述的视频解码器, 所述视频解码器被配置为检查所述第一层的数目为s个的空间区段或者编码块是否被完全解码, 且数目为s唯一地取决于所述语法元素结构以及所述第二层的数目为t-1个的已解码的空间区段或者编码块, 并且在除了所述检查显示在所述第一层的空间区段或者编码块之中至少已完全解码s个的情况下, 在所述第一层的解码期间推迟开始所述第二层的第t

[0565] 一个空间区段或者编码块的解码。

[0566] 71. 一种用于将多层视频数据流(40)解码成使用从第一层至第二层的层间预测而被编码在层的层级中的情景的方法, 视频解码器支持在将所述层的图片(12, 15)被细分而成的空间区段(80)中对所述多层视频数据流并行解码, 其中, 所述方法包括:

[0567] 检测所述多层视频数据流的长期语法元素结构(606; 例如, tile\_boundaries\_aligned\_flag)以:

[0568] 解译所述长期语法元素结构, 所述长期语法元素结构假设第一可能值集合之外的值(例如, tile\_boundaries\_aligned\_flag=1)作为对在预定时间段(608)内将所述第二层的图片(15)细分, 使得所述第二层的图片的空间区段之间的边界与所述第一层的图片(12)的空间区段的每一个边界重叠, 并且在比所述预定时间段短的时间间隔(604)内基于所述多层视频数据流的短期语法元素(602; 例如, column\_width\_minus1[i]和column\_width\_minus1[i])周期性地确定将所述第一层的图片和所述第二层的图片细分成空间区段的保证; 并且

[0569] 如果所述长期语法元素结构假设了第二可能值集合之外的值(例如, tile\_boundaries\_aligned\_flag=0), 则在比所述预定时间段短的时间间隔内从所述多层视频数据流的所述短期语法元素周期性地确定将所述层的图片细分成空间区段, 使得至少对于所述短期语法元素的第一可能值, 所述第二层的图片的空间区段之间存在不与所述第一层的空间区段的边界中的任一个重叠的边界, 并且至少对于所述短期语法元素的第二可能值, 所述第二层的图片的空间区段之间存在与所述第一层的空间区段的每一个边界重叠的边界。

[0570] 72. 一种用于利用从第一层至第二层的层间预测在层的层级中将场景编码成多层视频数据流的方法, 使得所述多层视频数据流能够在将所述层的图片细分而成的空间区段中被并行解码, 其中, 所述方法包括:

[0571] 将长期语法元素结构(606)和短期语法元素(602)插入到所述多层视频数据流中, 所述短期语法元素限定在时间间隔内将所述第一层的图片和所述第二层的图片细分成空间区段; 并且

[0572] 在将所述长期语法元素结构设定为以下的值之间进行切换:

[0573] 第一可能值集合之外的值, 且在比所述时间间隔大的预定时间段(608)内将所述短期语法元素设置成可能设置的集合之外的适当子集, 所述适当子集被选择使得在所述预定时间段内将所述第二层的图片细分, 使得所述第二层的图片的空间区段之间的边界与所述第一层的空间区段的每一个边界重叠; 或者

[0574] 第二可能值集合之外的值, 且在所述预定时间段内将所述短期语法元素设置成所述可能设置的集合中的任一个, 所述可能设置的集合包含至少一种设置和至少另一种设置, 根据所述至少一种设置, 所述第二层的图片的空间区段之间存在不与所述第一层的空

间区段的边界中的任一个重叠的边界，并且根据所述至少另一种设置，所述第二层的图片的空间区段之间存在与所述第一层的空间区段的每一个边界重叠的边界。

[0575] 73. 一种用于将空间可伸缩位流(40)解码成图片的方法，所述图片被编码在不同的空间层中，并且对于所述空间层中的至少一个，在第一空间区段中对所述图片进行编码，其中，所述方法包括：

[0576] 对第一空间层的图片(12)进行上采样以获得上采样参考图片，并且使用所述上采样参考图片预测第二空间层的图片(15)，其中，

[0577] 用于解码的所述方法响应于所述空间可伸缩位流中的语法元素(616；例如，`independent_tile_upsampling_idc`)以根据所述语法元素内插(620)所述第一空间层的图片，

[0578] 使得依赖于所述第一空间区段的所述上采样参考图片的分区(622)的任一部分独立于由所述分区的其他部分中的任一个覆盖的所述第一空间层的图片的部分；或者使得所述上采样参考图片的所述分区(622)的任一部分依赖于由在空间上与相应部分相邻的所述分区的另一部分覆盖的所述第一空间层的图片的部分。

[0579] 74. 一种用于将不同空间层中的图片编码成空间可伸缩位流的方法，并且对于所述空间层中的至少一个，在第一空间区段中对所述图片进行编码，其中，所述方法包括：

[0580] 对第一空间层的图片进行上采样以获得上采样参考图片并且使用所述上采样参考图片预测第二空间层的图片，其中，所述方法包括设置语法元素语法元素并且将所述语法元素(606)插入到所述空间可伸缩位流中，并且根据所述语法元素内插所述第一空间层的图片，

[0581] 使得依赖于所述第一空间区段的所述上采样参考图片的分区的任一部分独立于由所述分区的其他部分中的任一个覆盖的所述第一空间层的图片的部分；或者

[0582] 使得所述上采样参考图片的所述分区的任一部分依赖于由在空间上与相应部分相邻的所述分区中的另一分区覆盖的所述第一空间层的图片的部分。

[0583] 75. 一种用于将多层视频数据流解码成使用从第一层至第二层的层间预测而被编码在层的层级中的情景的方法，视频解码器支持利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历的层间延迟，通过以时间重叠方式顺次遍历空间区段来支持将所述多层视频数据流并行解码成将所述层的图片细分而成的空间区段，所述方法包括：

[0584] 检测所述多层视频数据流的长期语法元素结构(642；例如，`min_spatial_segment_delay`)使得：

[0585] 如果将所述长期语法元素结构(例如，`min_spatial_segment_delay` $\neq 0$ )设置成可能值的第一集合中的值，则使用所述长期语法元素的值预先确定预定时间段内的层间偏移量，并且在比所述预定时间段短的时间间隔内基于所述多层视频数据流的短期语法元素(602)分别周期性地确定所述第一层的图片的空间区段与所述第二层的图片的空间区段的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率；

[0586] 如果将所述长期语法元素设置成与所述可能值的第一集合不相交的可能值的第二集合的值(例如，`min_spatial_segment_delay` $= 0$ )，则在比所述预定时间段短的时间间隔内基于所述多层视频数据流的所述短期语法元素周期性地确定所述预定时间段内的层

间偏移量。

[0587] 76. 一种用于将使用从第一层至第二层的层间预测在层的层级中将情景编码成多层视频数据流的方法,使得能够利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历的层间偏移量,通过以时间重叠方式顺次遍历空间区段在将所述层的图片分区而成的空间区段中解码所述多层视频数据流,所述方法包括:

[0588] 将长期语法元素结构(min\_spatial\_segment\_delay)和短期语法元素插入并且设置到所述多层视频数据流中,所述多层视频数据流的所述短期语法元素在时间间隔内分别基于周期限定所述第一层的图片的空间区段与所述第二层的图片的空间区段的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率;

[0589] 其中,所述方法包括在下列设置之间切换:

[0590] 将所述长期语法元素结构(min\_spatial\_segment\_delay $\neq$ 0)设置成可能值的第一集合中的值,且该值在比所述时间间隔大的预定时间段内发出所述层间偏移量的信号,且在所述预定时间段内将所述短期语法元素设置成可能设置的集合之外的适当子集,所述适当子集被选择使得在所述预定时间段内,利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历的、小于或者等于由所述长期语法元素发出信号的所述层间偏移量的实际层间偏移量,所述第一层的图片的空间区段和所述第二层的图片的空间区段的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率通过所述时间重叠方式顺次遍历所述空间区段而分别启用所述多层视频数据流的解码;

[0591] 将所述长期语法元素设置成与可能值的所述第一集合不相交的可能值的第二集合(min\_spatial\_segment\_delay=0)中的值,且在所述预定时间段内,将所述短期语法元素设置成可能设置的集合中的任一个,可能设置的集合包含至少一种设置和至少另一种设置,根据所述至少一种设置,利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历的、小于或者等于由所述长期语法元素发出信号的所述层间偏移量的实际层间偏移量,所述第一层的图片的空间区段与所述第二层的图片的空间区段的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率通过以时间重叠方式顺次遍历空间区段而分别禁用所述多层视频数据流的解码,并且根据所述至少另一种设置,利用所述第一层的图片的空间区段的遍历相对于所述第二层的图片的空间区段的遍历的、小于或者等于由所述长期语法元素发出信号的所述层间偏移量的实际层间偏移量,所述第一层的图片的空间区段与所述第二层的图片的空间区段的大小和定位以及所述第一层的图片和所述第二层的图片的空间采样分辨率通过以时间重叠方式顺次遍历空间区段而分别启用所述多层视频数据流的解码。

[0592] 77. 一种用于将多层视频数据流处理成在层中被编码情景的方法,使得在各个层中,在由可伸缩维度跨越的可伸缩空间的不同操作点内编码所述情景,其中,所述多层视频数据流由第一NAL单元和第二NAL单元构成,所述第一NAL单元中的每个与所述层中的一个相关联,并且所述第二NAL单元散布在所述第一NAL单元内并且呈现关于所述多层视频数据流的整体信息,所述方法包括:

[0593] 检测所述第二NAL单元中的类型指示符字段(696;例如,dedicated\_scalability\_ids\_flag);

[0594] 如果所述类型指示符字段具有第一状态(例如,dedicated\_scalability\_ids\_

`flag=0`), 则从所述第二NAL单元中读取将第一NAL单元报头中的层指示符字段(例如, `layer_id`)的可能值映射至所述操作点的映射信息(例如, `layer_id_in_nuh[i]`, `dimension_id[i][j]`)并且通过所述层指示符字段和所述映射信息使所述第一NAL单元与所述第一NAL单元内的所述操作点相关联;

[0595] 如果所述类型指示符字段具有第二状态(`dedicated_scalability_ids_flag=1`), 则通过将所述第一NAL单元中的所述层指示符字段分割成一个以上的部分使所述第一NAL单元与所述操作点相关联, 并且通过使用该部分的值作为所述可伸缩空间内的矢量的坐标来定位所述第一NAL单元的所述操作点。

[0596] 78. 一种用于在层中将情景编码成多层视频数据流的方法, 使得在各个层中, 所述情景被编码在在由可伸缩维度跨越的可伸缩空间的不同操作点中, 其中, 所述多层视频数据流由第一NAL单元和第二NAL

[0597] 单元构成, 所述第一NAL单元中的每个与所述层中的一个相关联, 并且所述第二NAL单元散布在所述第一NAL单元内并且呈现关于

[0598] 所述多层视频数据流的整体信息, 所述方法包括:

[0599] 将类型指示符字段插入到所述第二NAL单元中并且在以下设

[0600] 置之间进行切换:

[0601] 设置所述类型指示符字段使得所述类型指示符字段具有第一状态, 且将第一NAL单元报头中的层指示符字段的可能值映射至所述操作点的映射信息插入到所述第二NAL单元中并且设置所述第一NAL单元中的所述层指示符字段, 使得通过所述映射信息使所述第一NAL单元的操作点与相应层的指示符字段相关联;

[0602] 设置所述类型指示符字段使得所述类型指示符字段具有第二状态(`dedicated_scalability_ids_flag=1`), 且通过将所述第一NAL单元中的所述层指示符字段分割成一个以上的部分来设置所述第一NAL单元中的所述层指示符字段, 并且设置所述一个以上的部分使得该部分的值对应于所述可伸缩空间内的矢量的坐标, 从而指向与相应的第一NAL单元相关联的操作点。

[0603] 79. 一种用于使多层视频数据流转换成被编码在层中的情景的方法, 以通过多标准多层解码器进行解码, 其中, 所述多层视频数据流由NAL

[0604] 单元构成, 所述NAL单元中的每个与所述层中的一个相关联, 其中,

[0605] 所述层与不同的编解码器相关联使得对于各个层, 使用与相应层相关联的编解码器将与相应层相关联的所述NAL单元编码, 所述方法

[0606] 包括:

[0607] 对于每个NAL单元, 识别与NAL单元相关联的相同编解码器;

[0608] 并且

[0609] 将所述多层视频数据流的所述NAL单元移交至所述多标准多层解码器, 所述多标准多层解码器使用与不同编解码器相关联的层之间的层间预测来对所述多层视频数据流进行解码。

[0610] 80. 一种用于将多层视频数据流解码成情景的方法, 使用从第一层的部分至第二层的共置部分的层间预测在层的层级中编码所述情景, 其中, 将所述第一层的图片细分成第一块的阵列并且将所述第二层的图片细分成第二块的阵列, 其中, 由所述第一块和所述

第二块分别限定光栅扫描解码顺序,所述方法包括:

[0611] 利用在所述第一块的单元内测量的、第一层的图片的第一块的遍历相对于所述第二层的图片的第二块的空间遍历的所述层间偏移量,根据所述多层视频数据流的语法元素结构(例如,ctb\_delay\_enabled\_flag,min\_spatial\_segment\_delay)通过以时间重叠方式顺次遍历所述第一块和所述第二块来确定用于将所述第一层的图片和所述第二层的图片并行解码的层间偏移量。

[0612] 81.一种具有程序代码的计算机程序,当在计算机上运行所述计算机程序时,所述程序代码用于执行根据项71至80中任一项所述的方法。

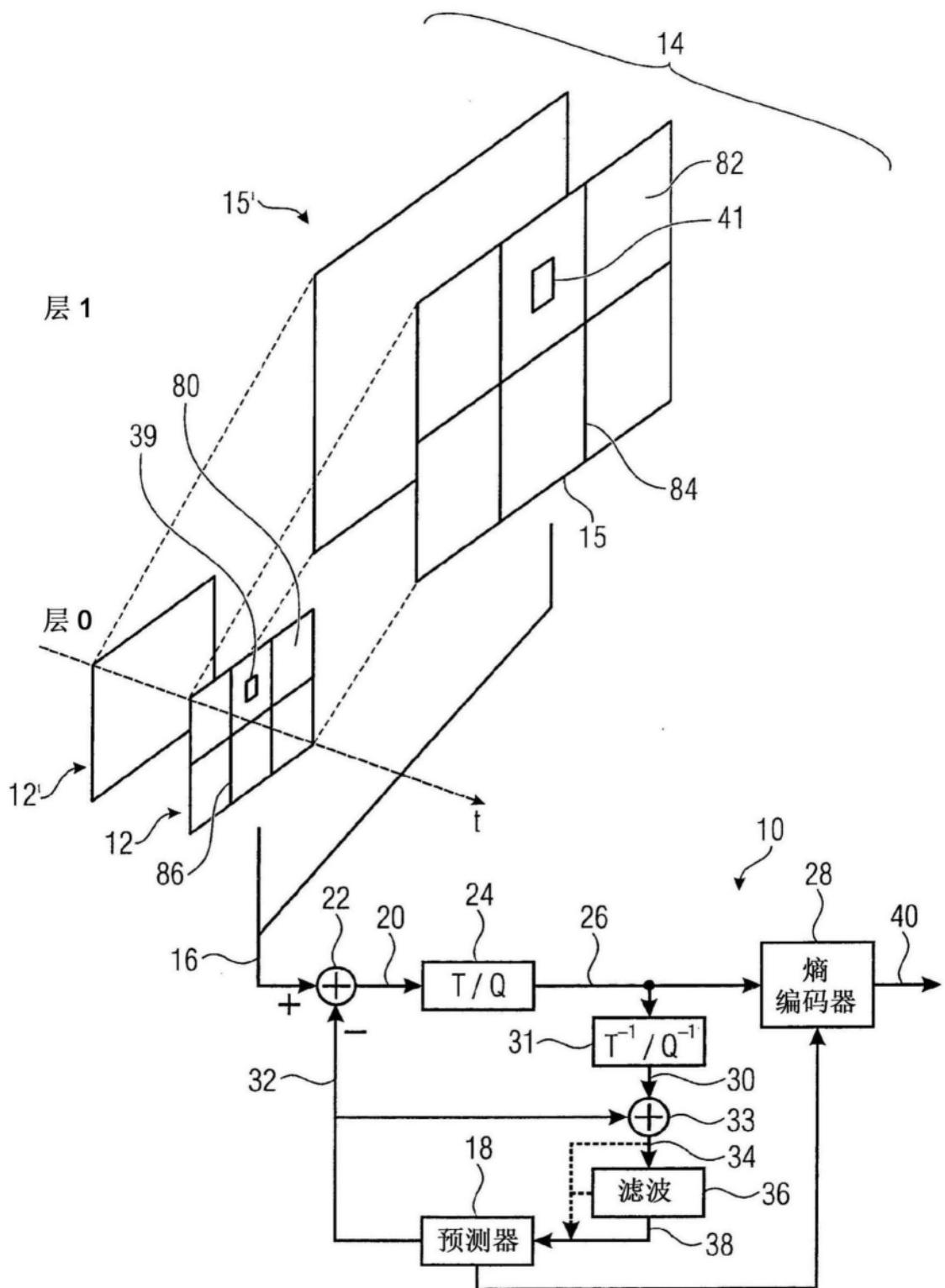


图1

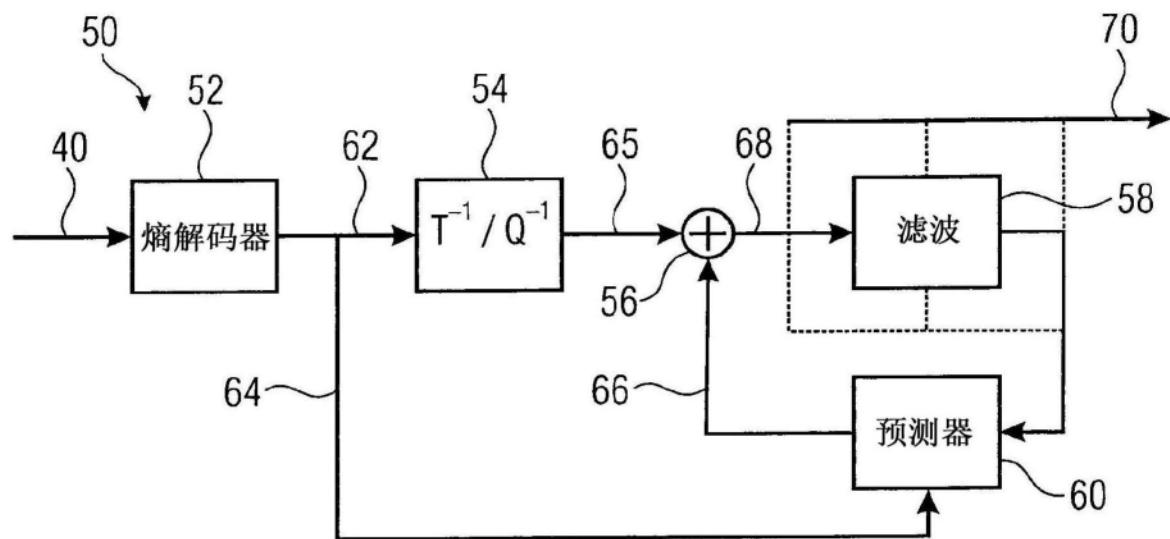


图2

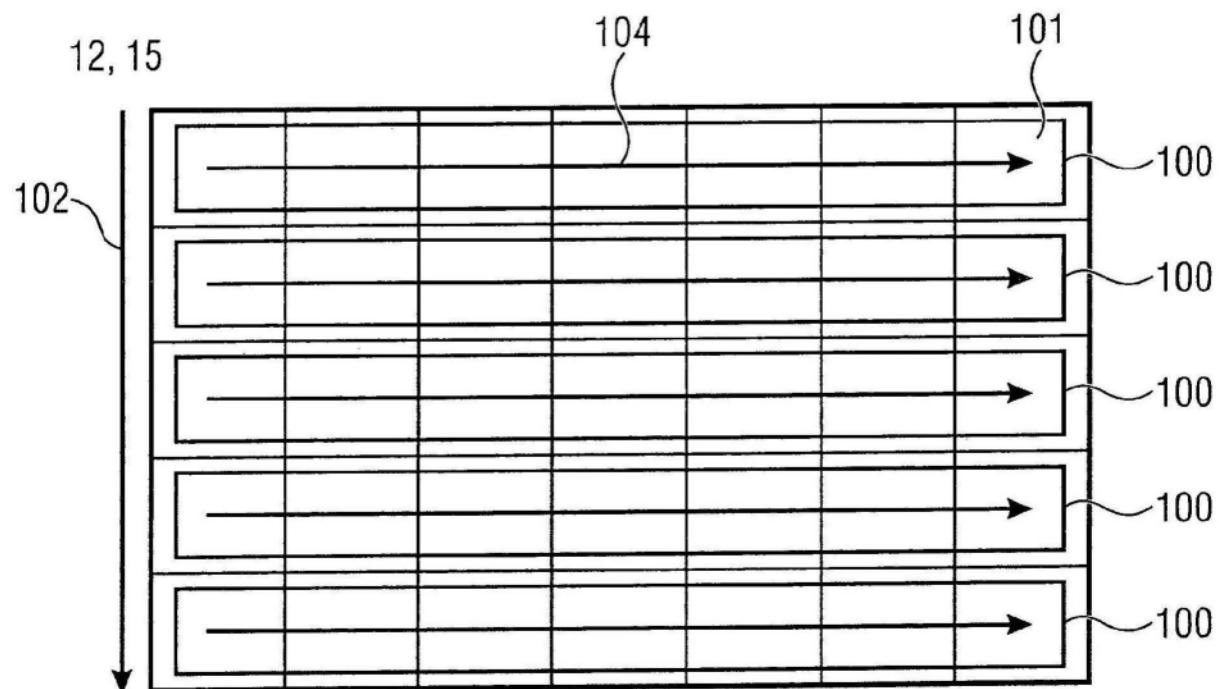


图3

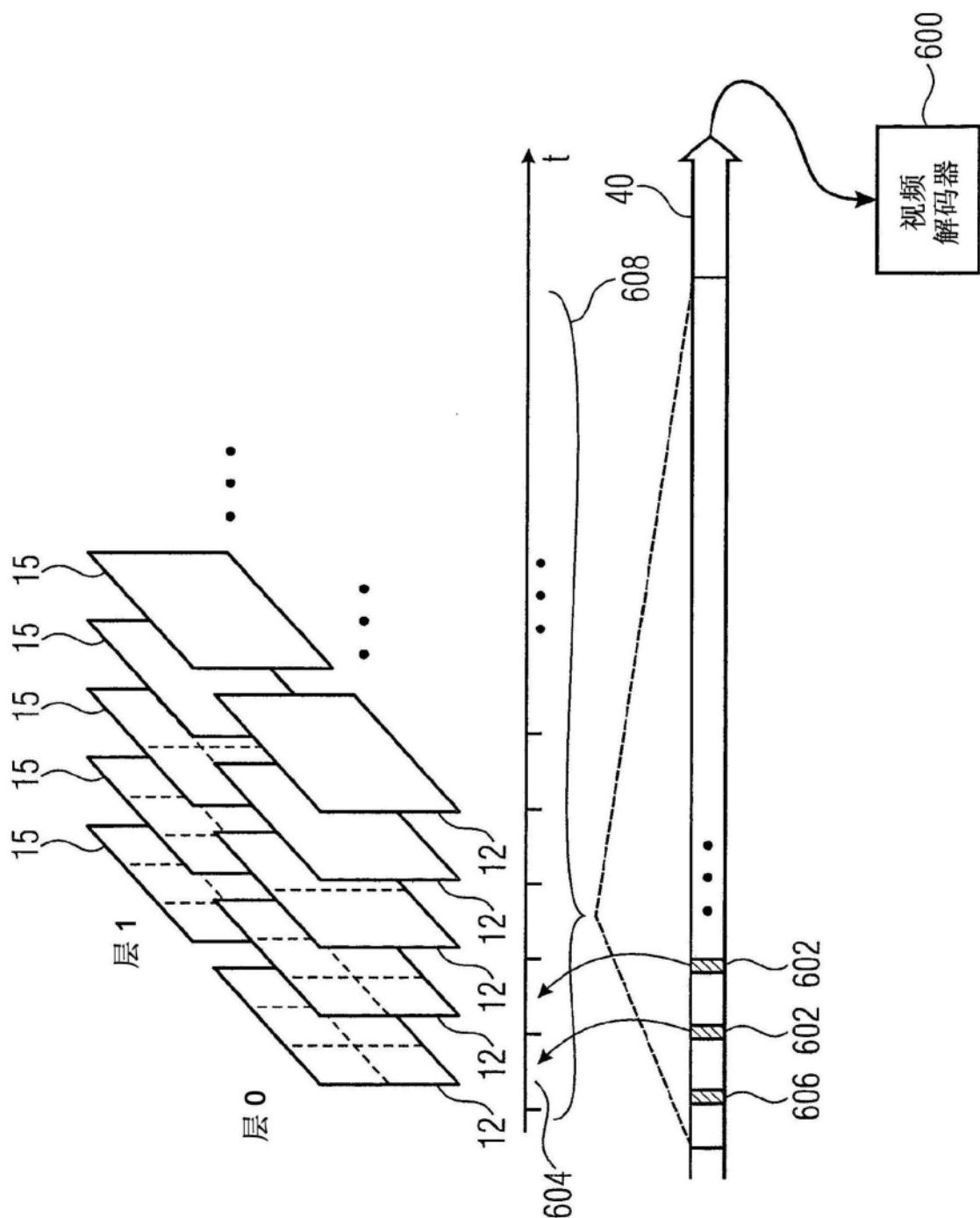


图4

0	1	2	3	12	13	14	21	22	23	24
4	5	6	7	15	16	17	25	26	27	28
8	9	10	11	18	19	20	29	30	31	32
33	34	35	36	41	42	43	47	48	49	50
37	38	39	40	44	45	46	51	52	53	54
55	56	57	58	63	64	65	69	70	71	72
59	60	61	62	66	67	68	73	74	75	76

图5

pic_parameter_set_rbsp( ) {	描述符
pic_parameter_set_id	ue(v)
...	
if( tiles_enabled_flag ) {	
num_tile_columns_minus1	ue(v)
num_tile_rows_minus1	ue(v)
uniform_spacing_flag	u(1)
if( !uniform_spacing_flag ) {	
for( i = 0; i < num_tile_columns_minus1; i++ )	
column_width_minus1[ i ]	ue(v)
for( i = 0; i < num_tile_rows_minus1; i++ )	
row_height_minus1[ i ]	ue(v)
}	
loop_filter_across_tiles_enabled_flag	u(1)
}	
...	
}	

图6

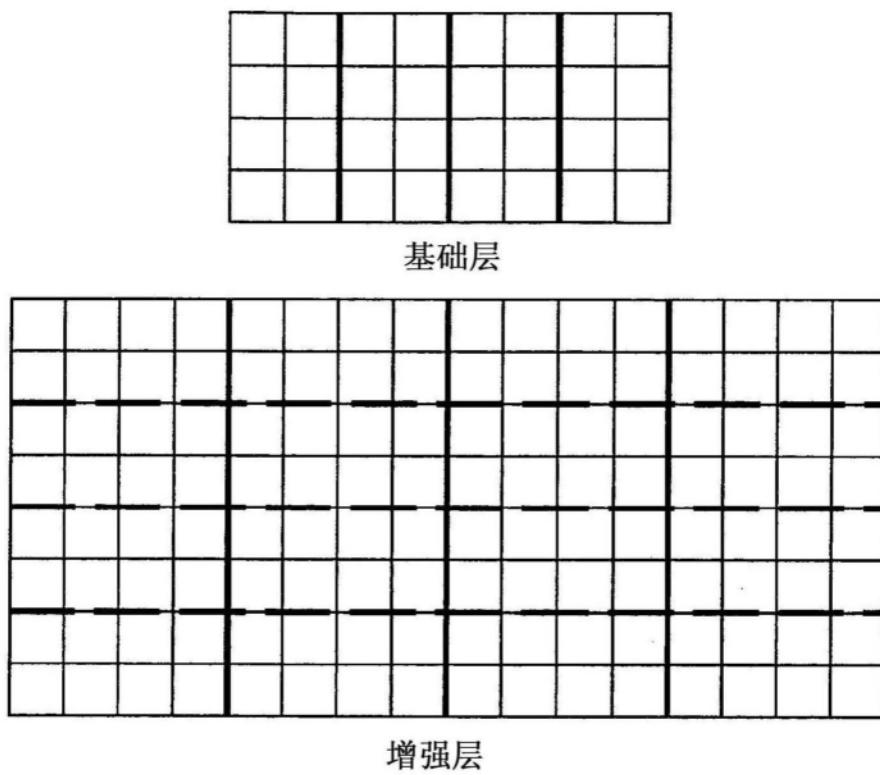


图7

vui_parameters( ) {	描述符
aspect_ratio_info_present_flag	u(1)
...	
if( bitstream_restriction_flag ) {	
tiles_fixed_structure_flag	u(1)
tile_boundaries_aligned_flag	u(1)
motion_vectors_over_pic_boundaries_flag	u(1)
restricted_ref_pic_lists_flag	u(1)
min_spatial_segmentation_idc	u(8)
max_bytes_per_pic_denom	ue(v)
max_bits_per_minCU_denom	ue(v)
log2_max_mv_length_horizontal	ue(v)
log2_max_mv_length_vertical	ue(v)
}	
}	

图8

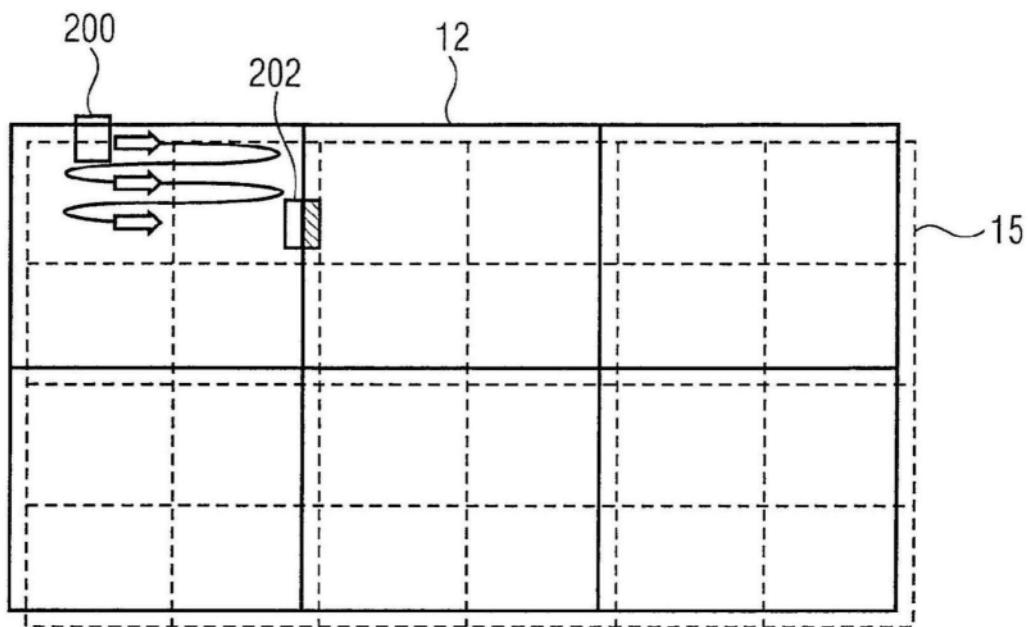


图9

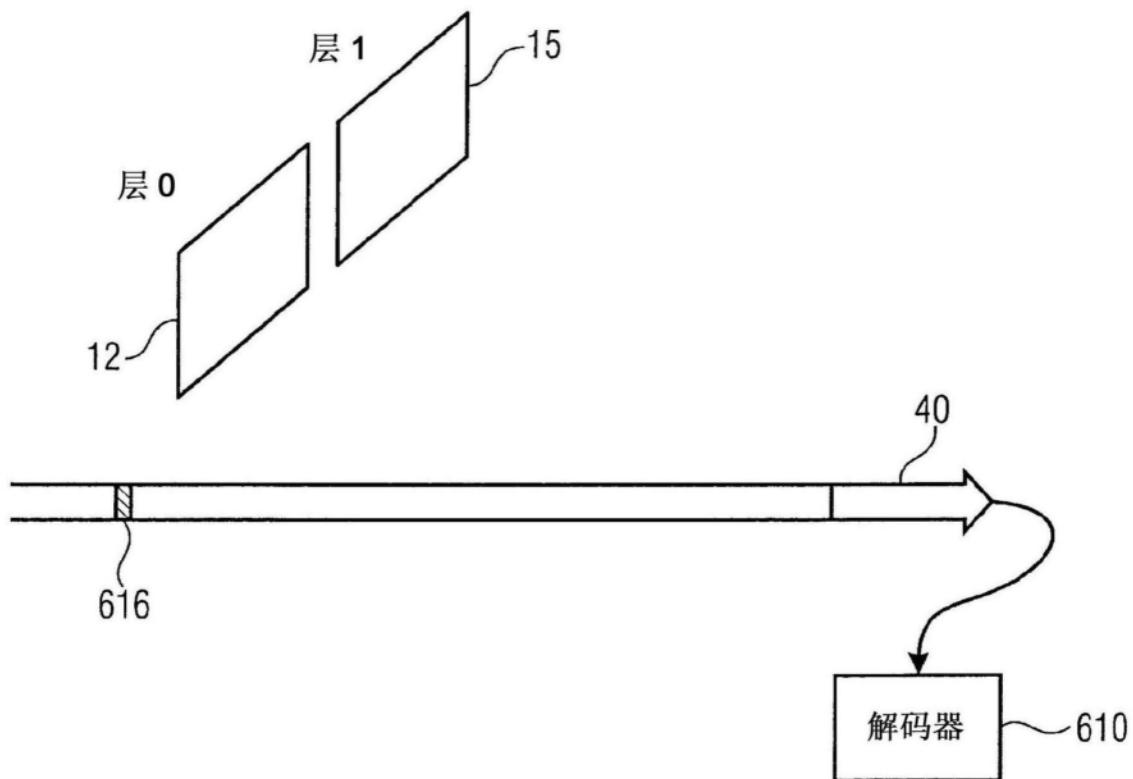


图10

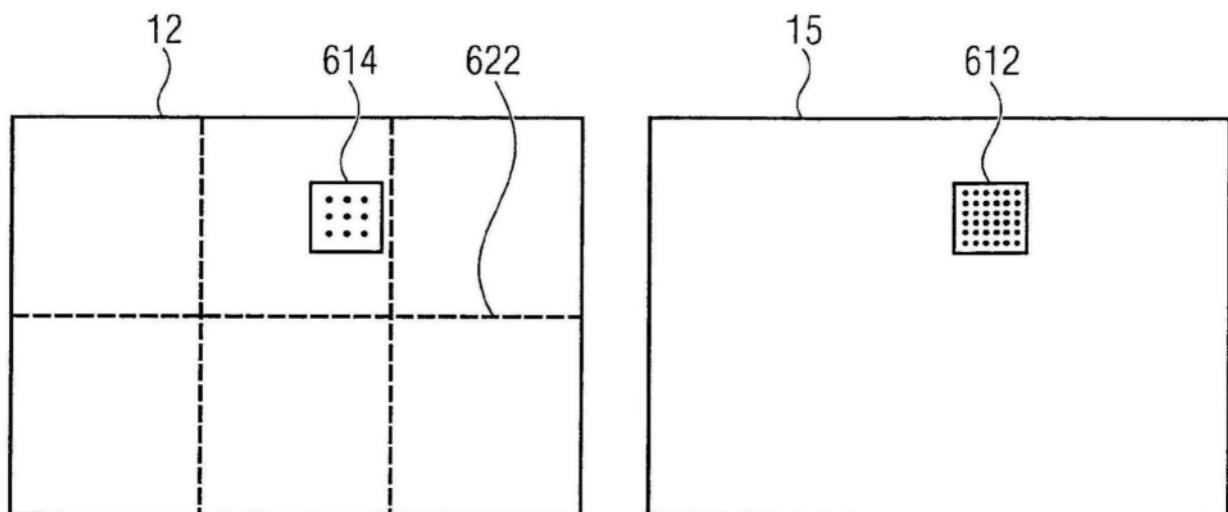


图11

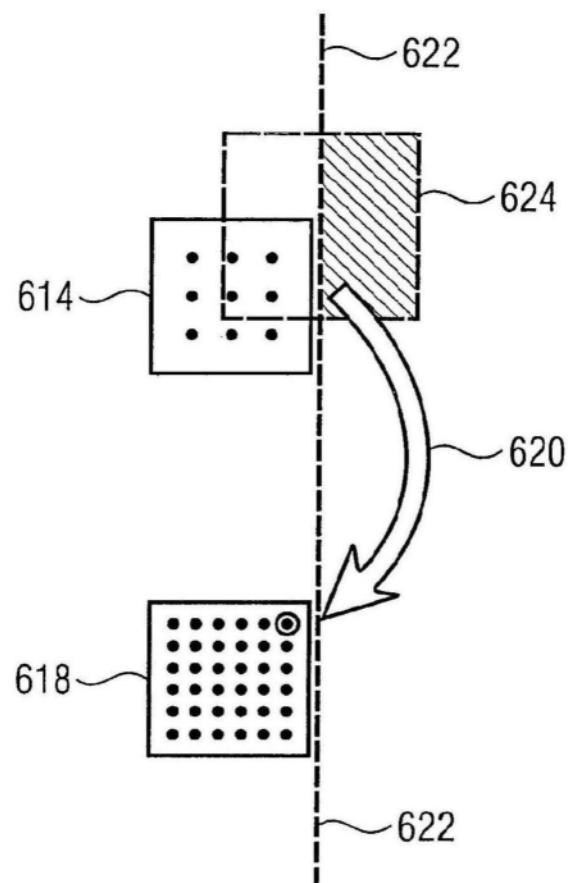


图12

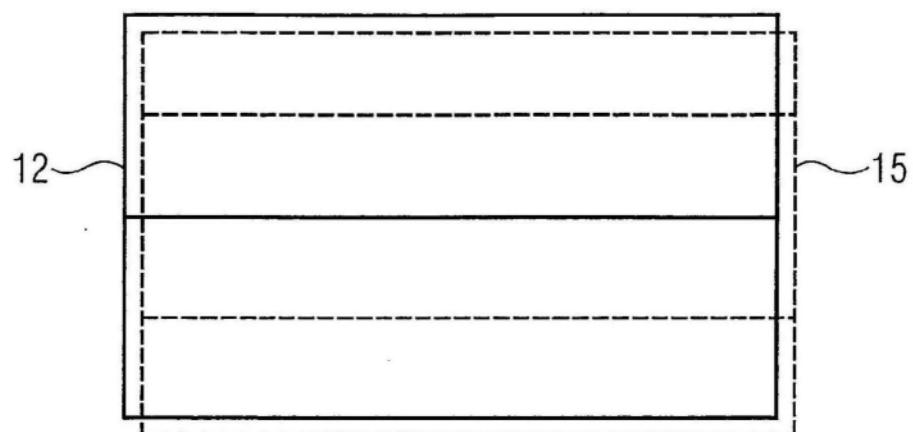


图13

描述符
pic_parameter_set_rbsp( ) {
pic_parameter_set_id
...
if( tiles_enabled_flag ) {
num_tile_columns_minus1
num_tile_rows_minus1
uniform_spacing_flag
if( !uniform_spacing_flag ) {
for( i = 0; i < num_tile_columns_minus1; i++ )
column_width_minus1[ i ]
for( i = 0; i < num_tile_rows_minus1; i++ )
row_height_minus1[ i ]
}
loop_filter_across_tiles_enabled_flag
independent_tile_upsampling_idc
}
...
}

图14

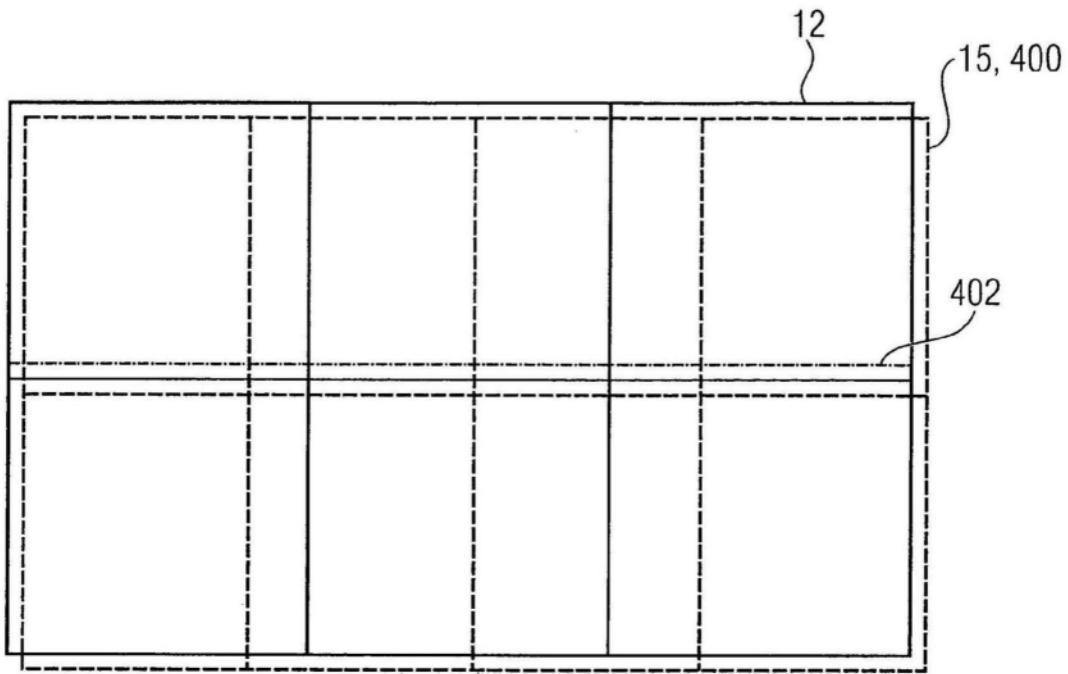


图15A

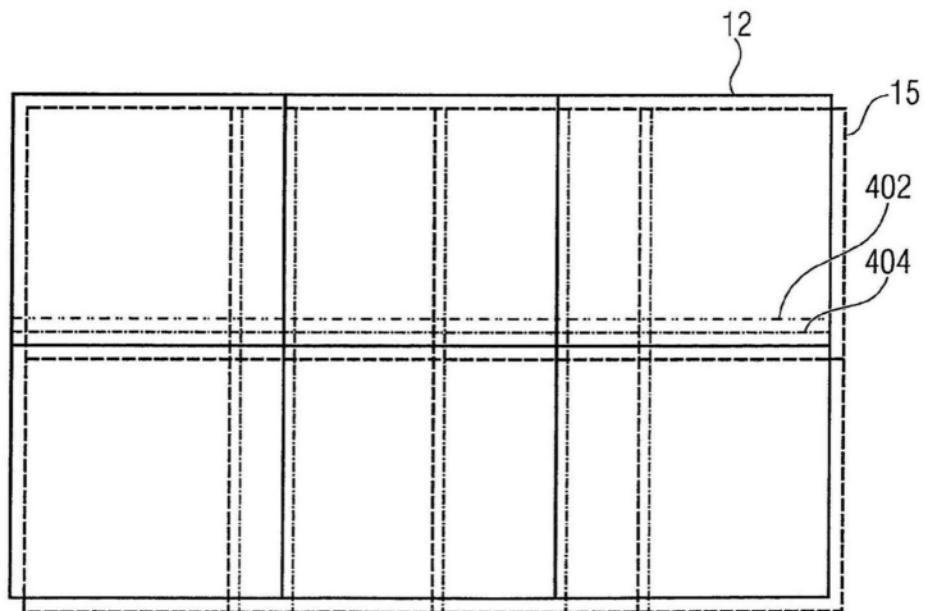


图15B

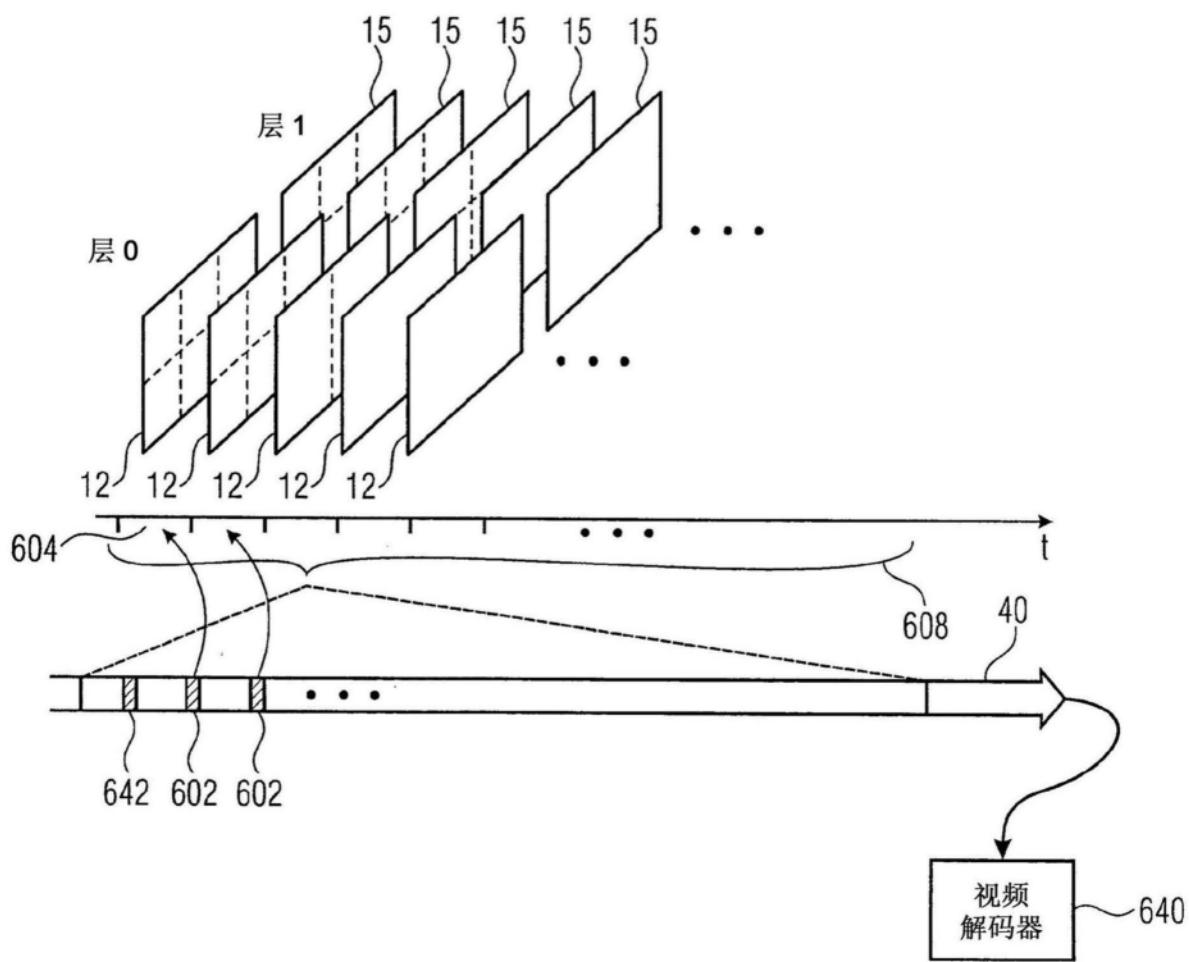


图16

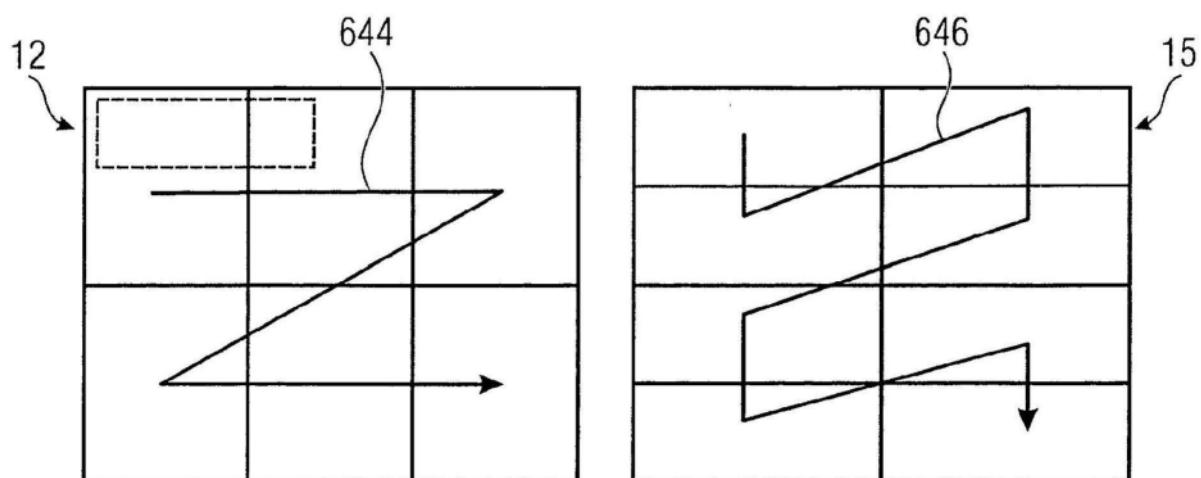


图17A

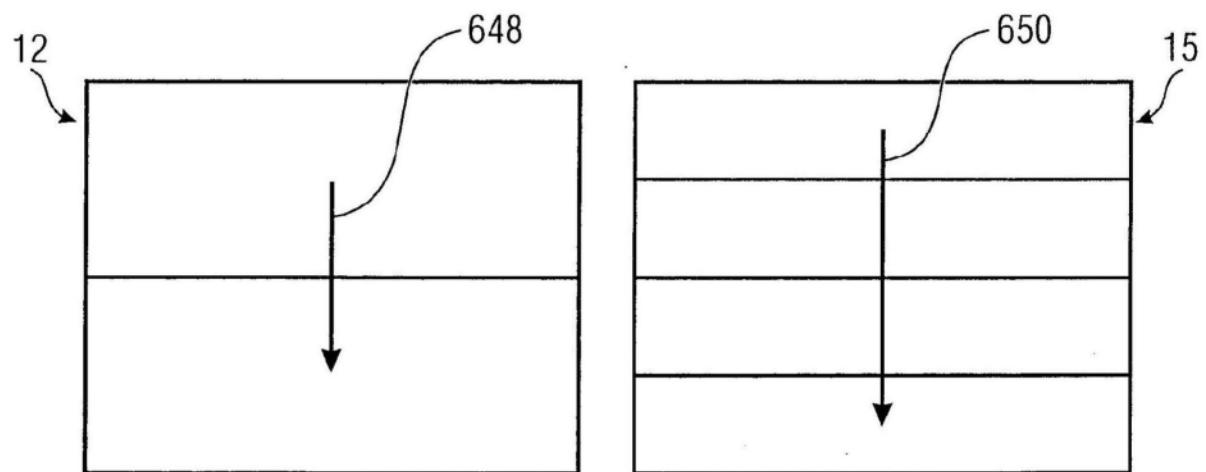


图17B

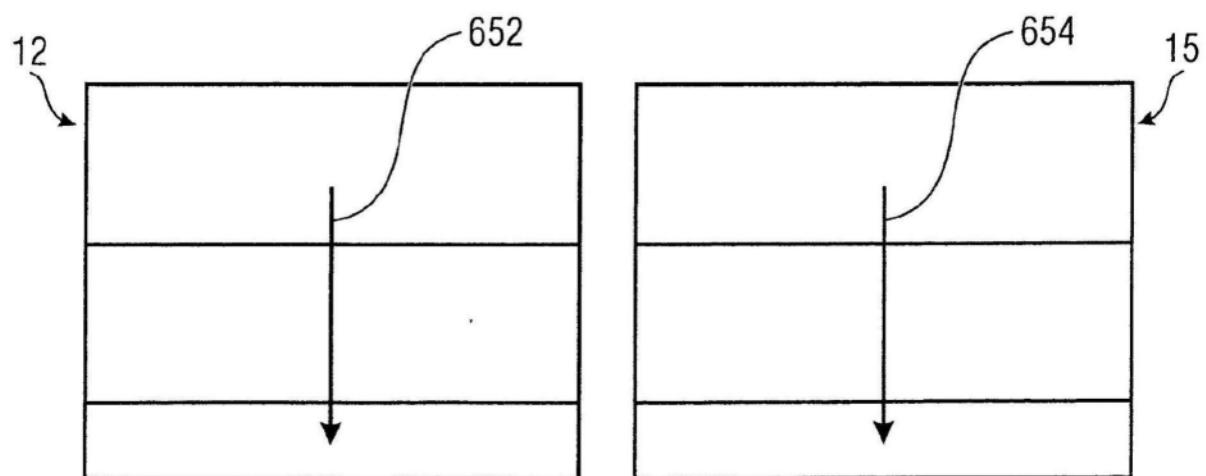


图17C

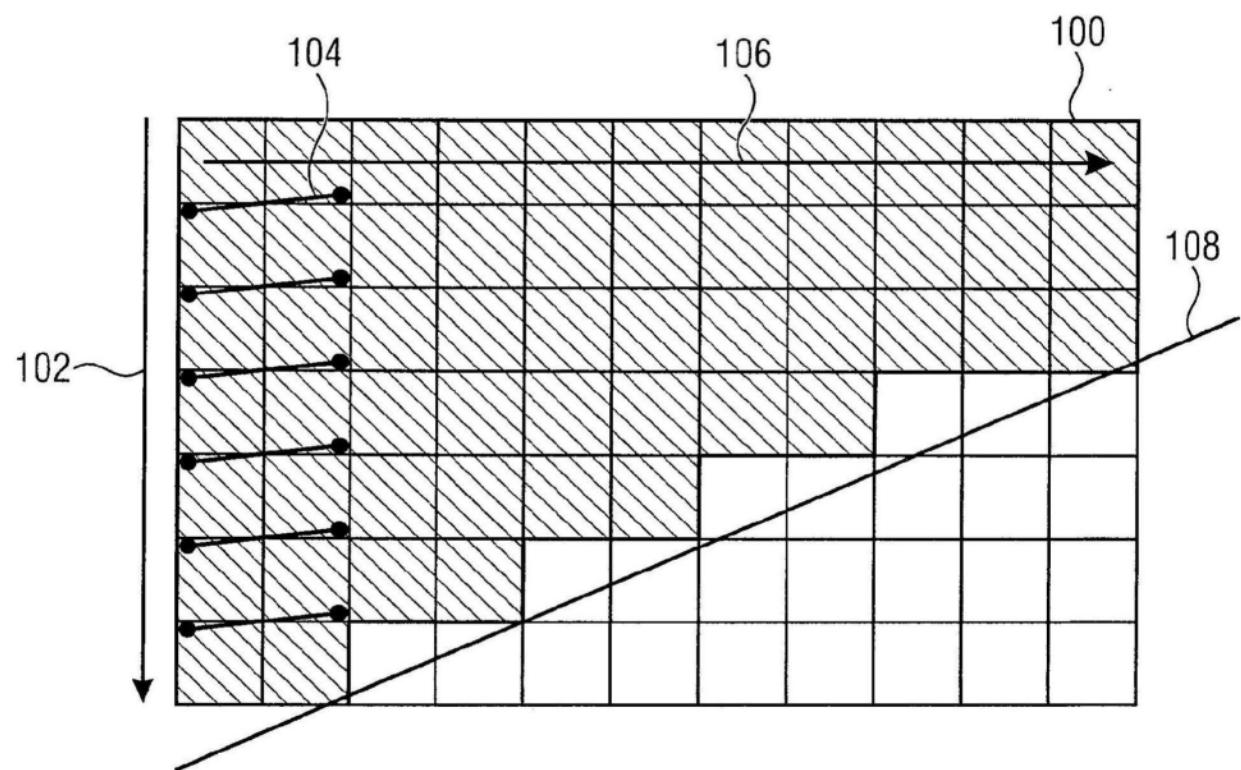
12.15

图18

两个层中的 CTB 大小		覆盖对应图像区域的 CTB 行		层解码延迟 (可以将层 (n+1) 的 第一 CTB 行之前的 CTB 行偏移解码)	
图片大小 缩放因子	基础层最大 CTB 大小	增强层最大 CTB 大小	基础层 CTB 行	增强层 CTB 行	在 CTB 行中的 层解码延迟
1.5	64	16	1	6	1
1.5	64	32	1	3	1
1.5	64	64	2	3	1
1.5	32	64	4	3	2
1.5	16	64	8	3	3
2	64	16	1	8	1
2	64	32	1	4	1
2	64	64	1	2	1
2	32	64	1	1	1
2	16	64	2	1	2

图19

vui_parameters( ) {	描述符
aspect_ratio_info_present_flag	u(1)
...	
if( bit stream_restriction_flag ) {	
tiles_fixed_structure_flag	u(1)
motion_vectors_over_pic_boundaries_flag	u(1)
restricted_ref_pic_lists_flag	u(1)
min_spatial_segmentation_idc	u(8)
min_spatial_segment_delay	ue(v)
max_bytes_per_pic_denom	ue(v)
max_bits_per_mincu_denom	ue(v)
log2_max_mv_length_horizontal	ue(v)
log2_max_mv_length_vertical	ue(v)
}	
}	

图20

vps_extension( ) {	描述符
while( !byte_aligned( ) )	
...	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
profile_tier_level( 1, vps_max_sub_layers_minus1 )	
min_spatial_segment_delay[ i ]	ue(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// layer dependency	
num_direct_ref_layers[ i ]	u(6)
for( j = 0; j < num_direct_ref_layers[ i ]; j++ )	
ref_layer_id[ i ][ j ]	u(6)
}	
}	

图21

nal_unit_header( ) {	描述符
forbidden_zero_bit	f(1)
nal_unit_type	u(6)
layer_id	u(6)
nuh_temporal_id_plus1	u(3)
}	

图22

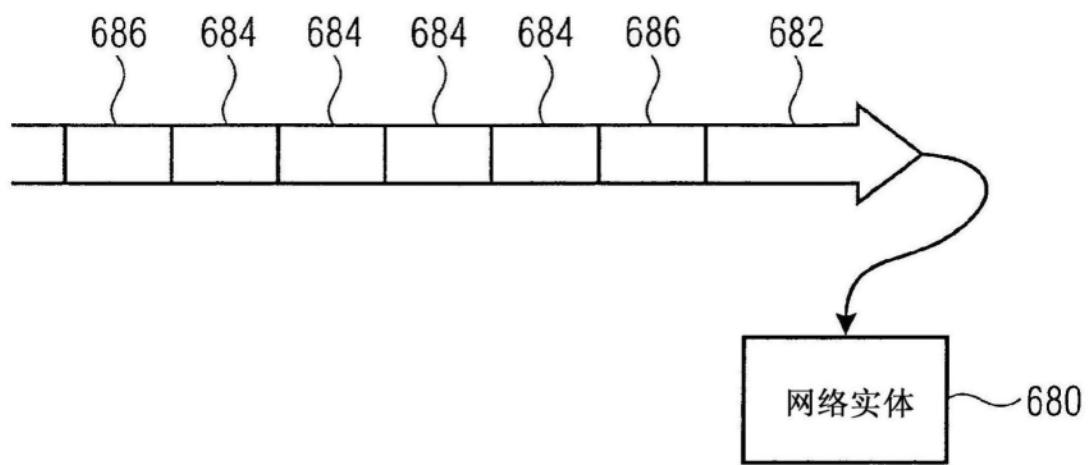


图23

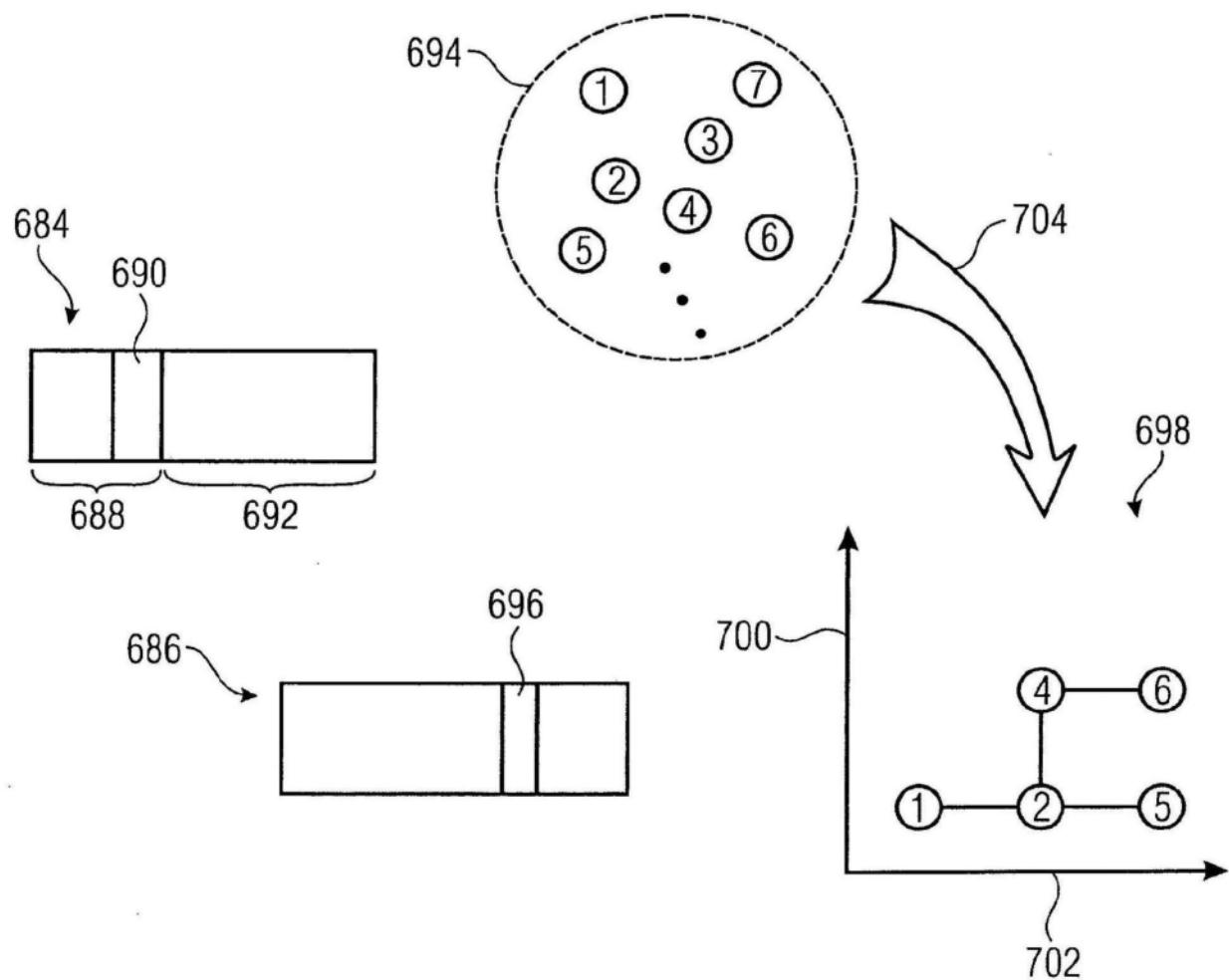


图24

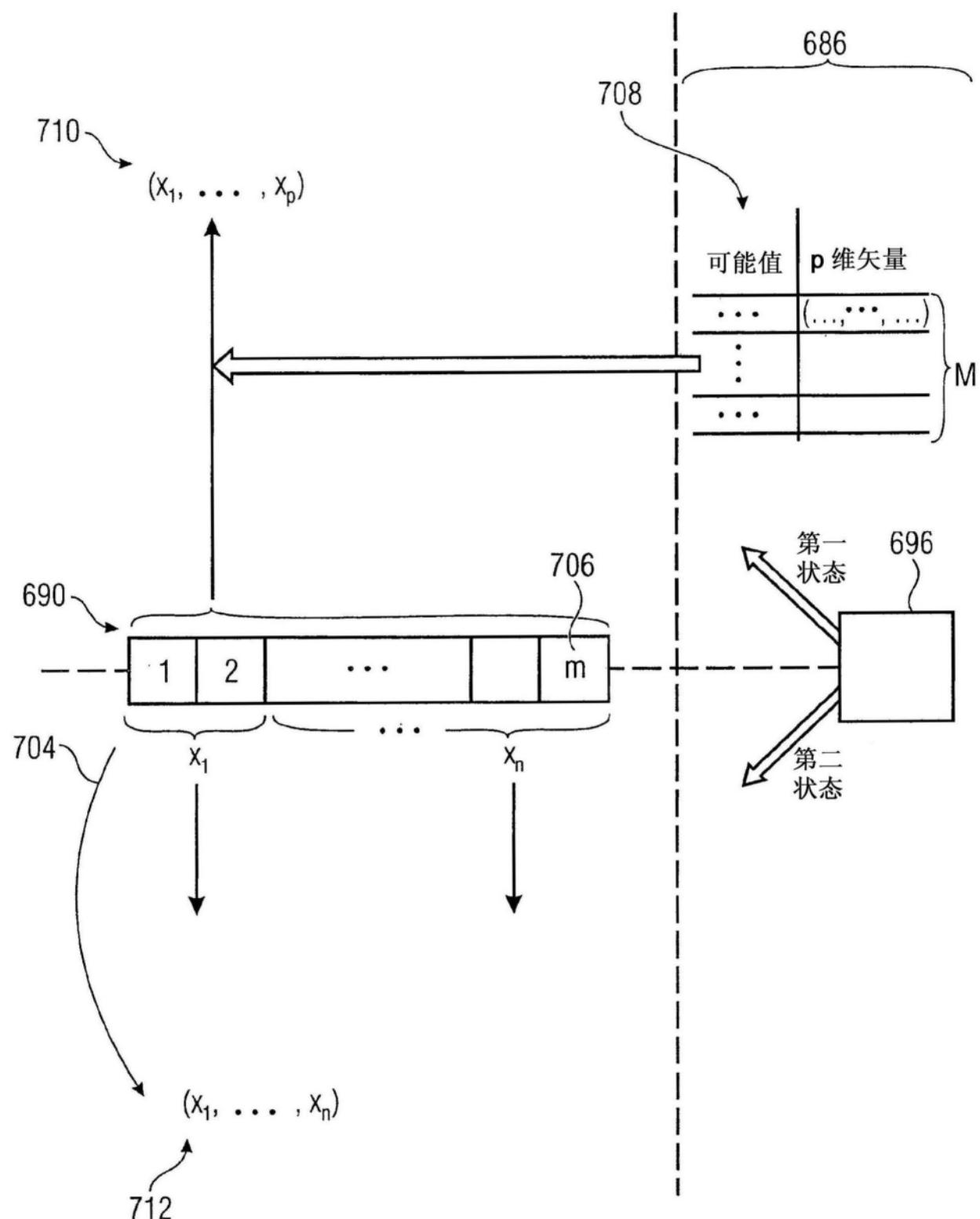


图25

vps_extension( ) {	描述符
...	
dedicated_scalability_ids_flag	u(1)
scalability_mask	u(16)
for( i = 0; i < NumScalabilityTypes; i++ ) {	
dimension_id_len_minus1[ i ]	u(3)
}	
if ( !dedicated_scalability_ids_flag ) {	
vps_nuh_layer_id_present_flag	u(1)
// layer specific information	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// mapping of layer ID to scalability dimension IDs	
if( vps_nuh_layer_id_present_flag )	
layer_id_in_nuh[ i ]	u(6)
for( j = 0; j <= num_dimensions_minus1; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
}	
...	
}	

图26

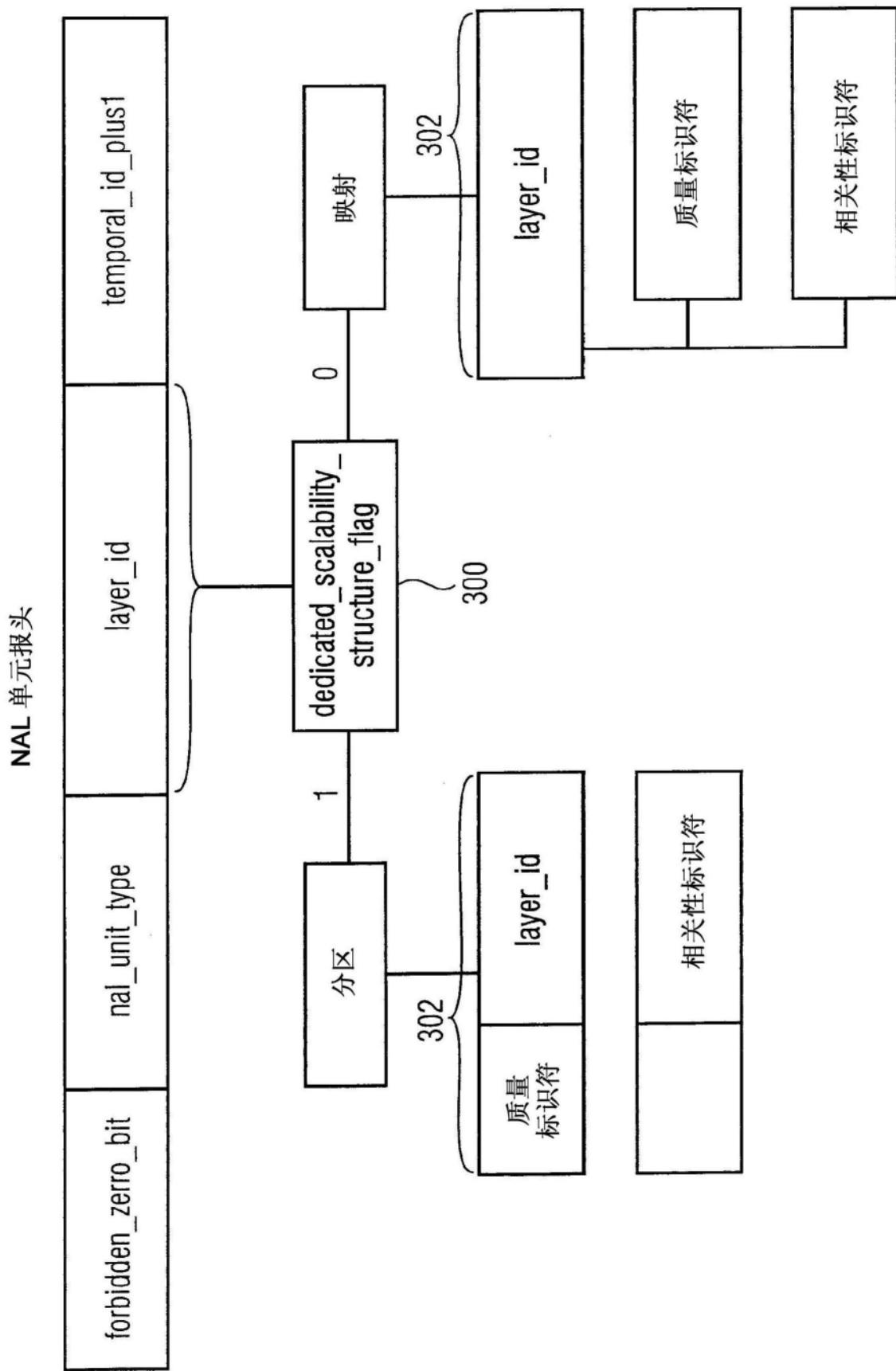


图27

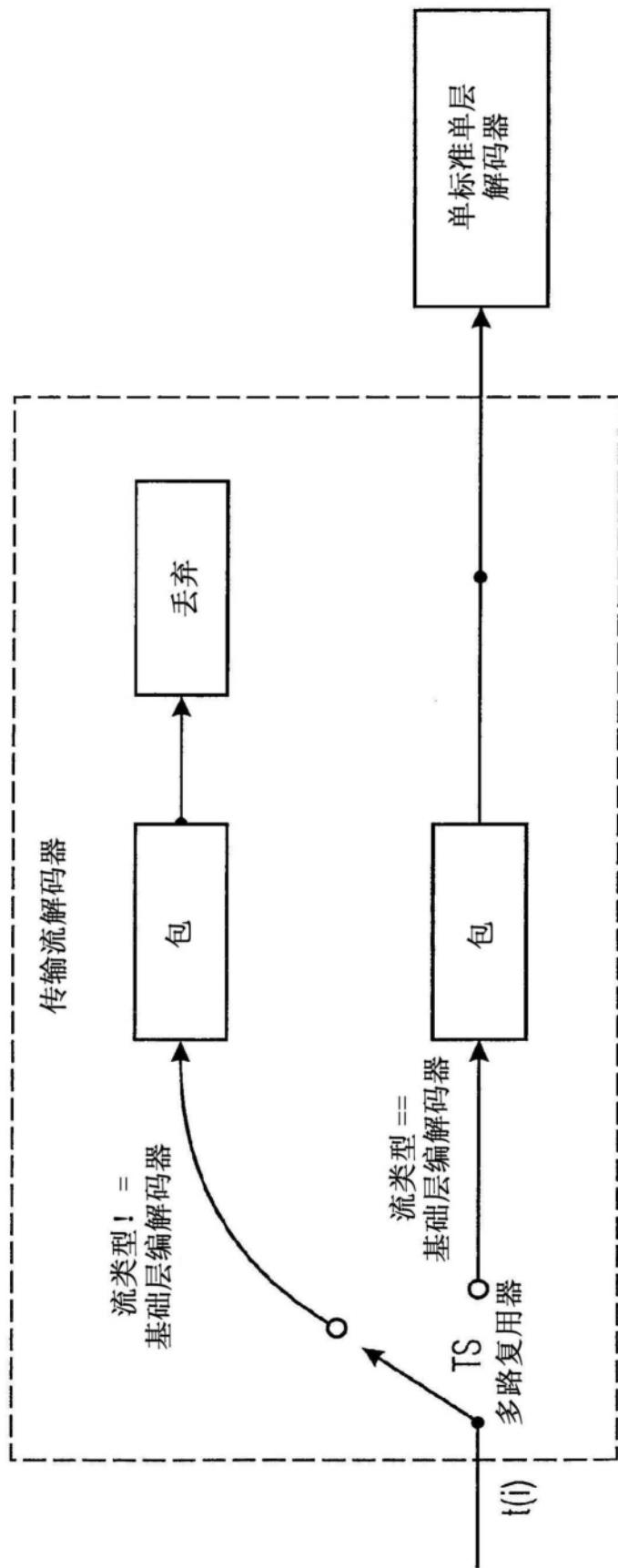


图28

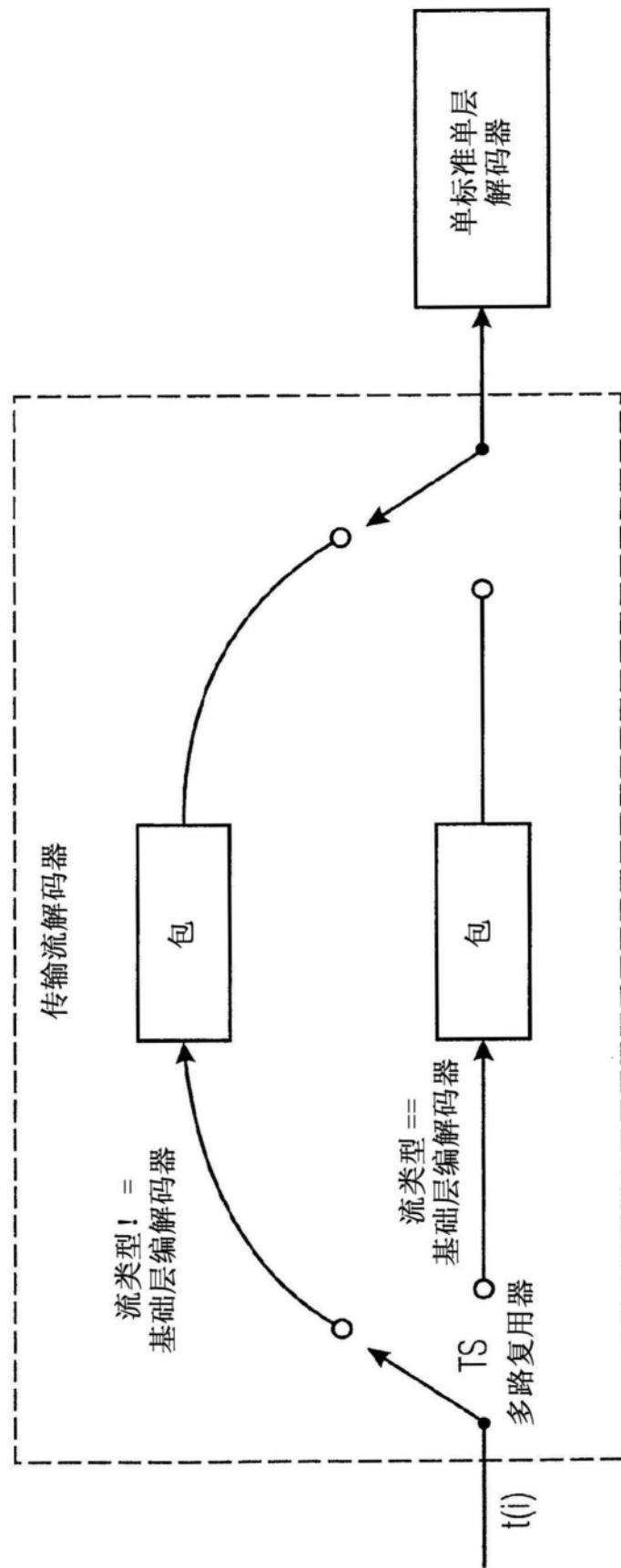


图29

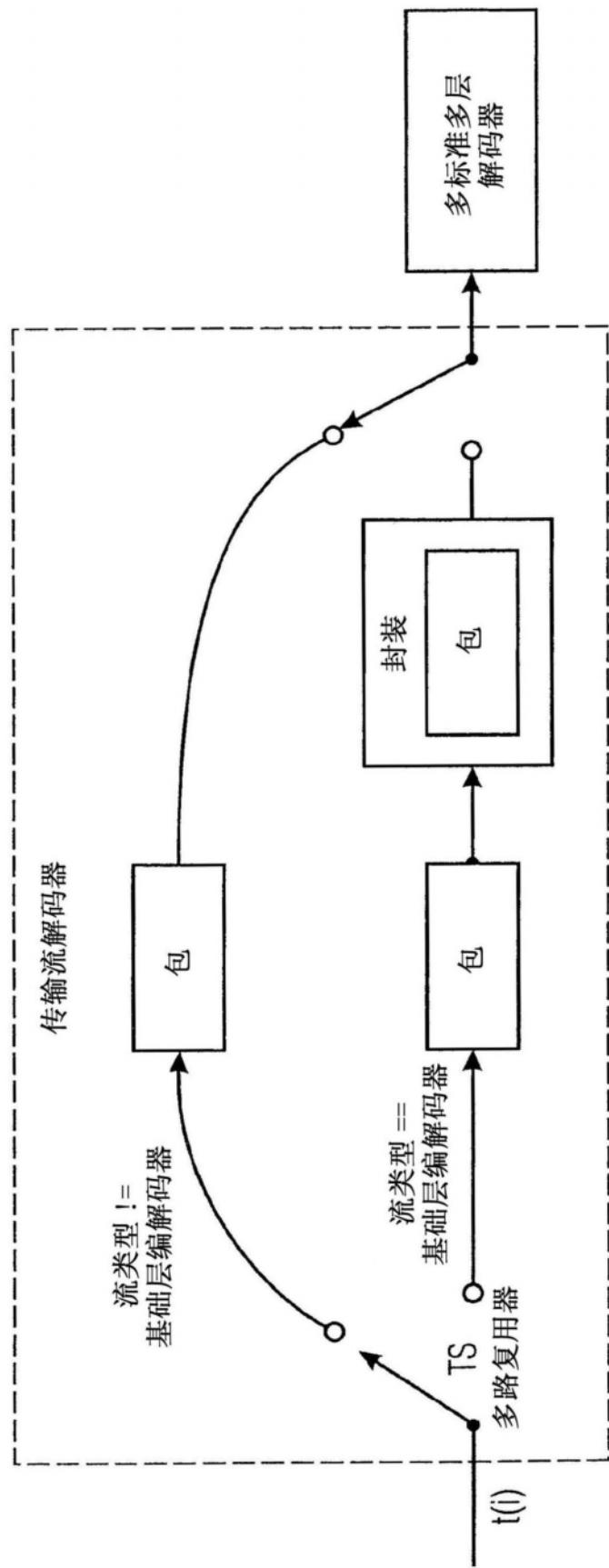


图30

vps_extension( ) {		描述符
avc_base_layer_flag		u(1)
vps_vui_offset		u(16)
splitting_flag		u(1)
for( i = 0, NumScalabilityTypes = 0; i < 16; i++ ) {		
scalability_mask_flag[ i ]		u(1)
NumScalabilityTypes += scalability_mask_flag[ i ]		
}		
for( j = 0; j < ( NumScalabilityTypes - splitting_flag); j++ )		
dimension_id_len_minus1[ j ]		u(3)
vps_nuh_layer_id_present_flag		u(1)
for( i = 1; i <= vps_max_layers_minus1; i++ ) {		
if ( vps_nuh_layer_id_present_flag )		
layer_id_in_nuh[ i ]		u(6)
if ( !splitting_flag )		
for( j = 0; j < NumScalabilityTypes; j++ )		
dimension_id[ i ][ j ]		u(v)
}		
⋮		
⋮		
⋮		
⋮		

图31

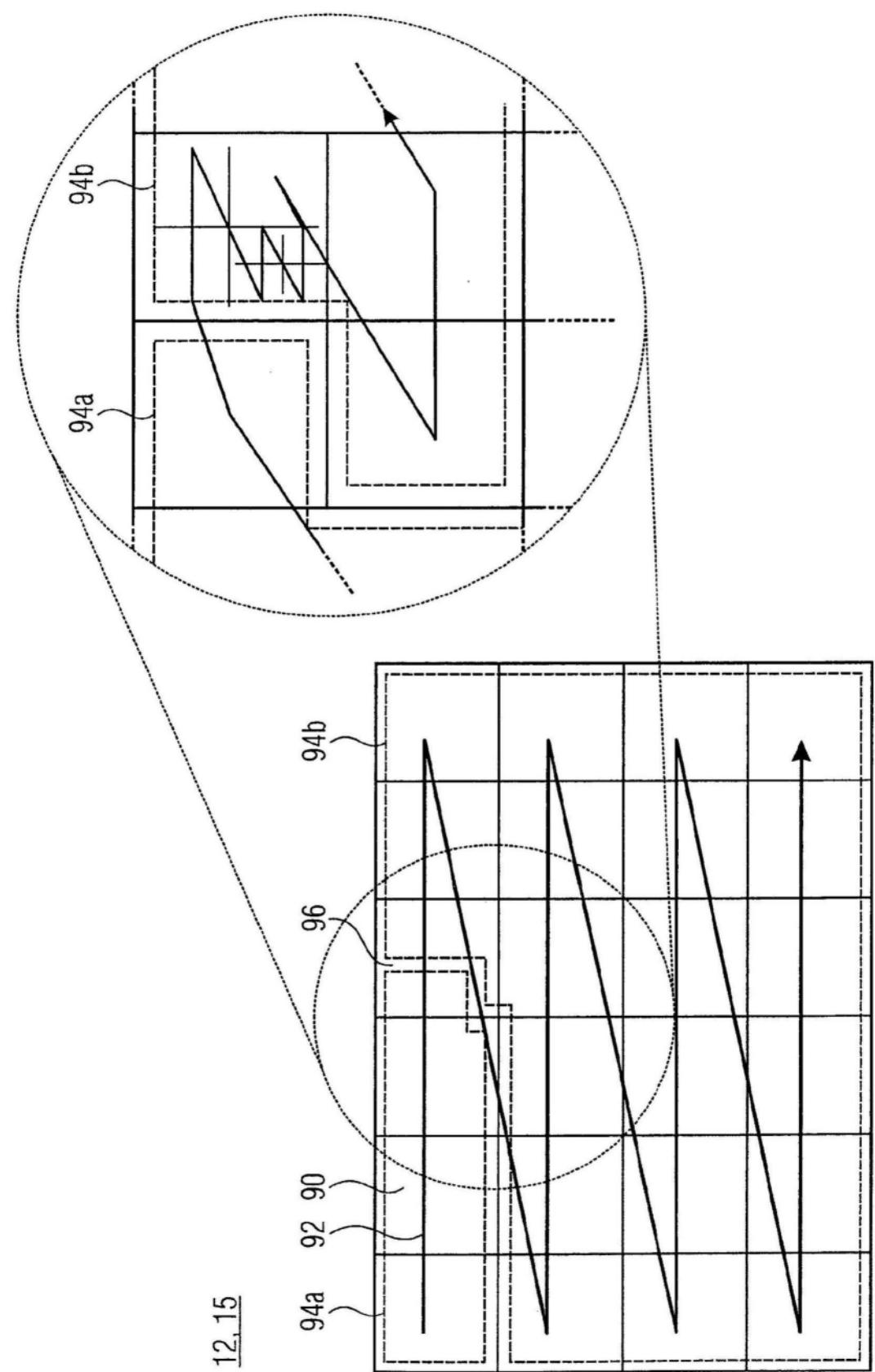


图32

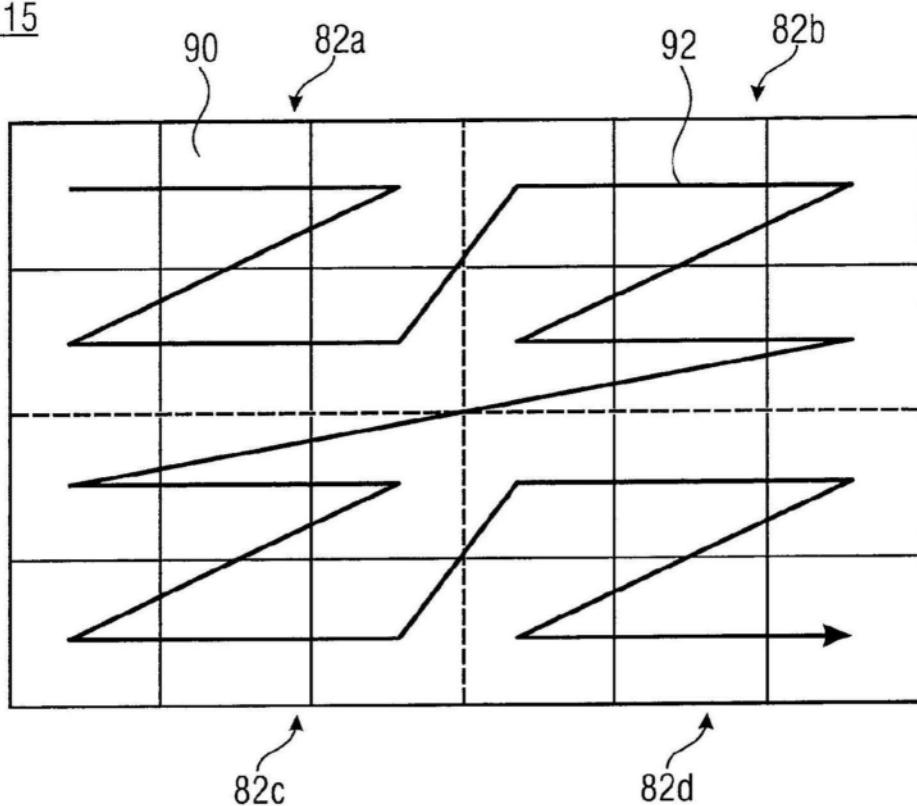
12, 15

图33

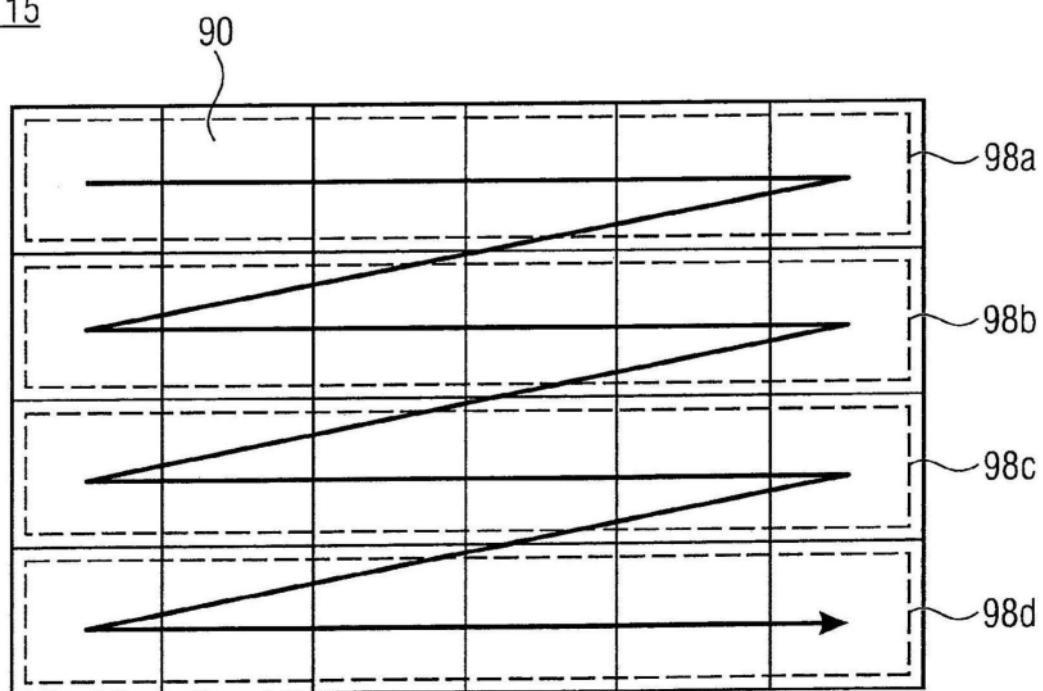
12, 15

图34

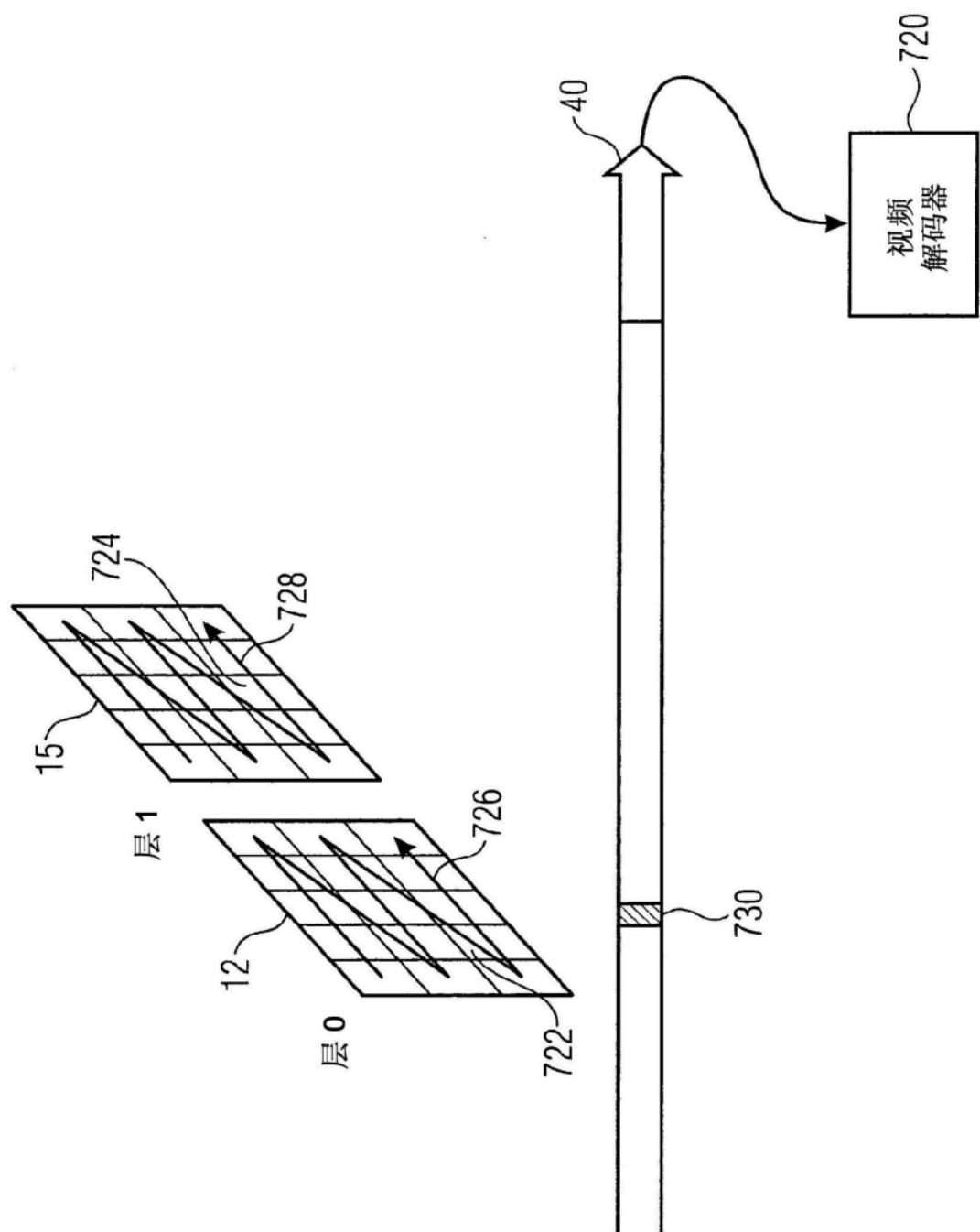


图35

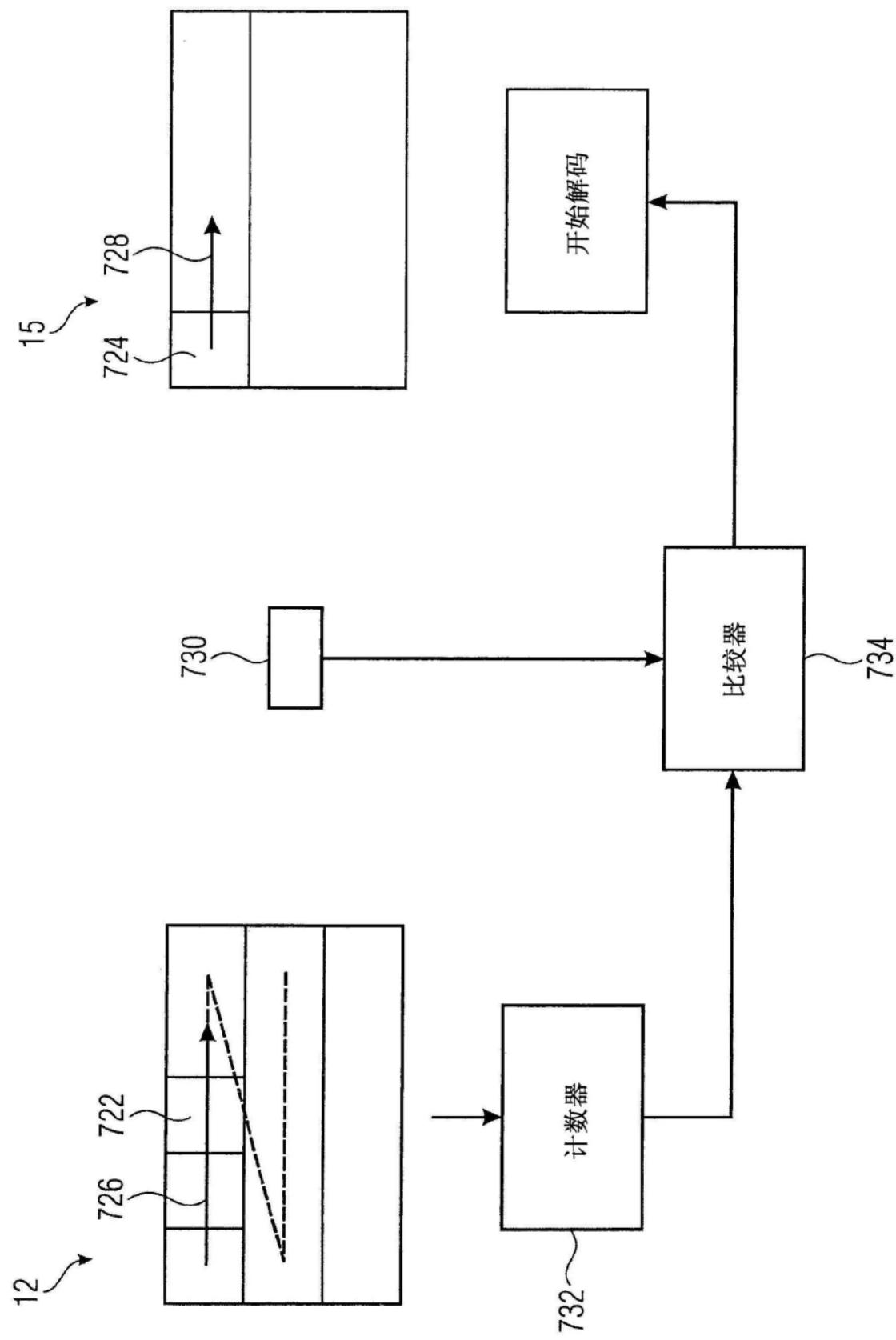


图36

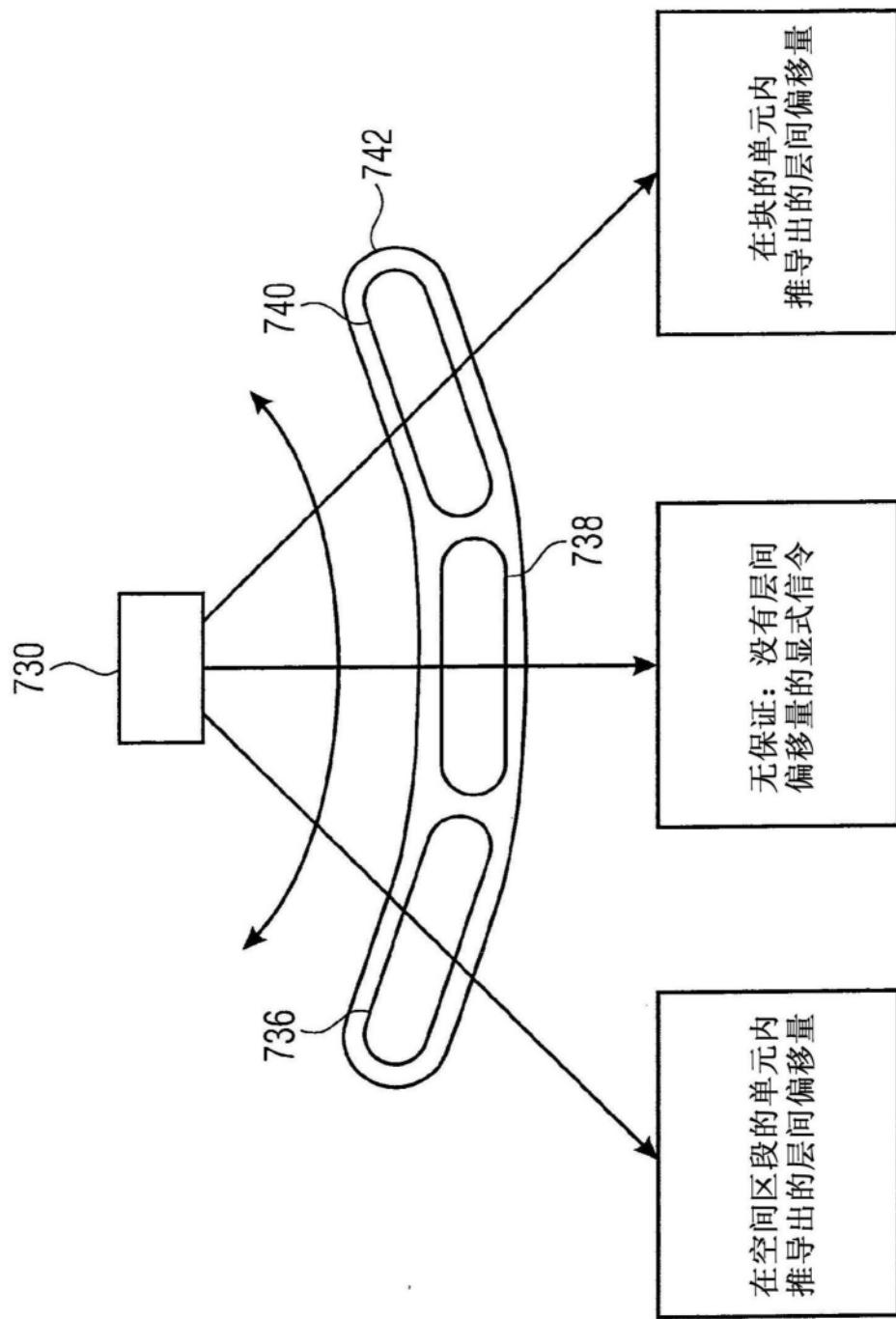


图37

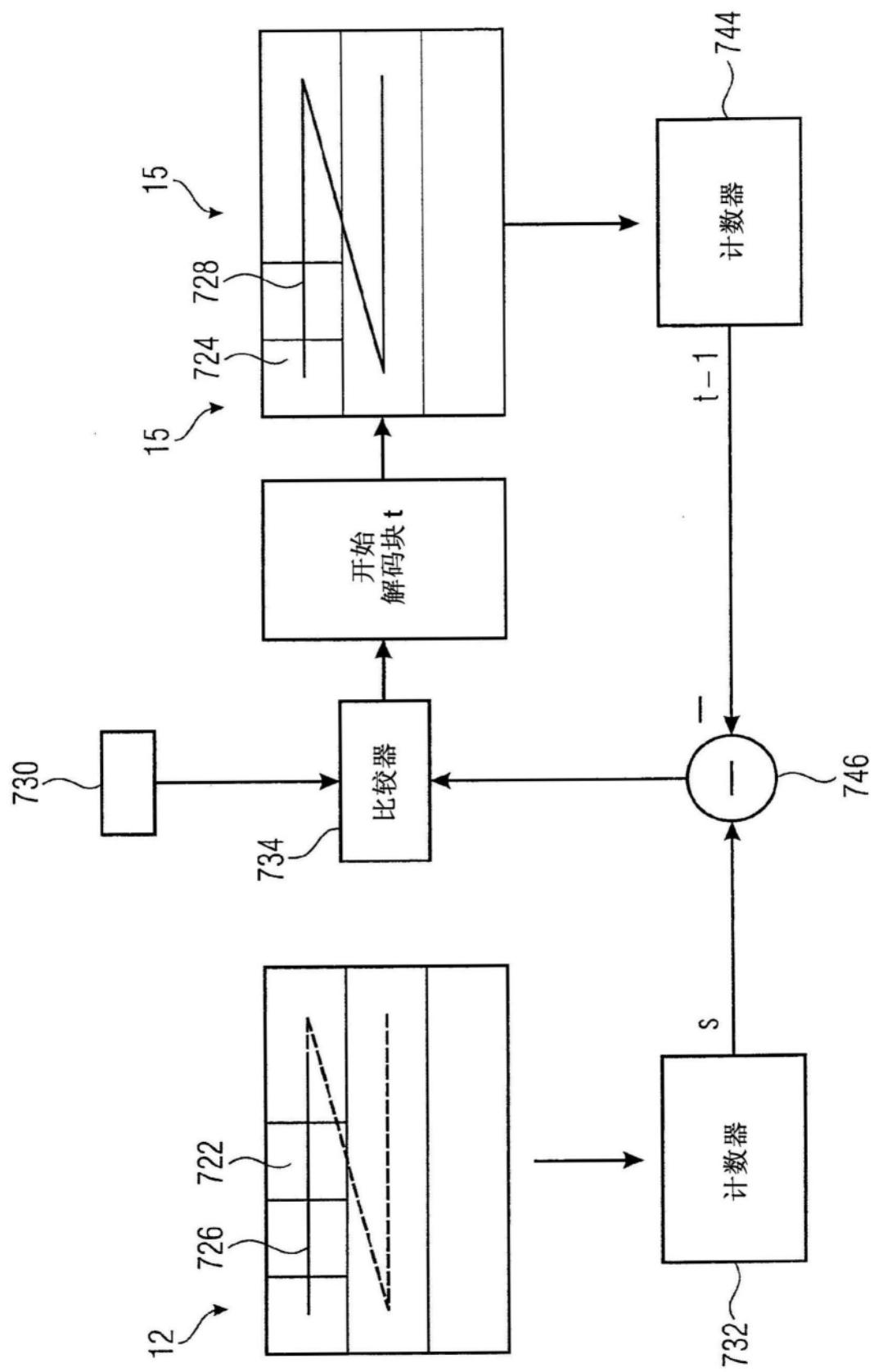


图38

## CTB 列 (CTB 地址)

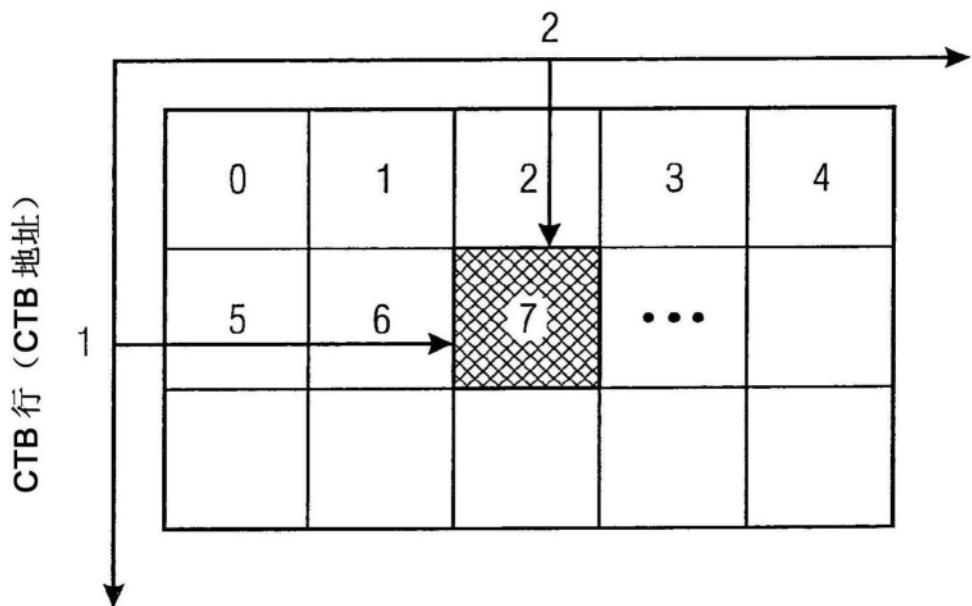


图39

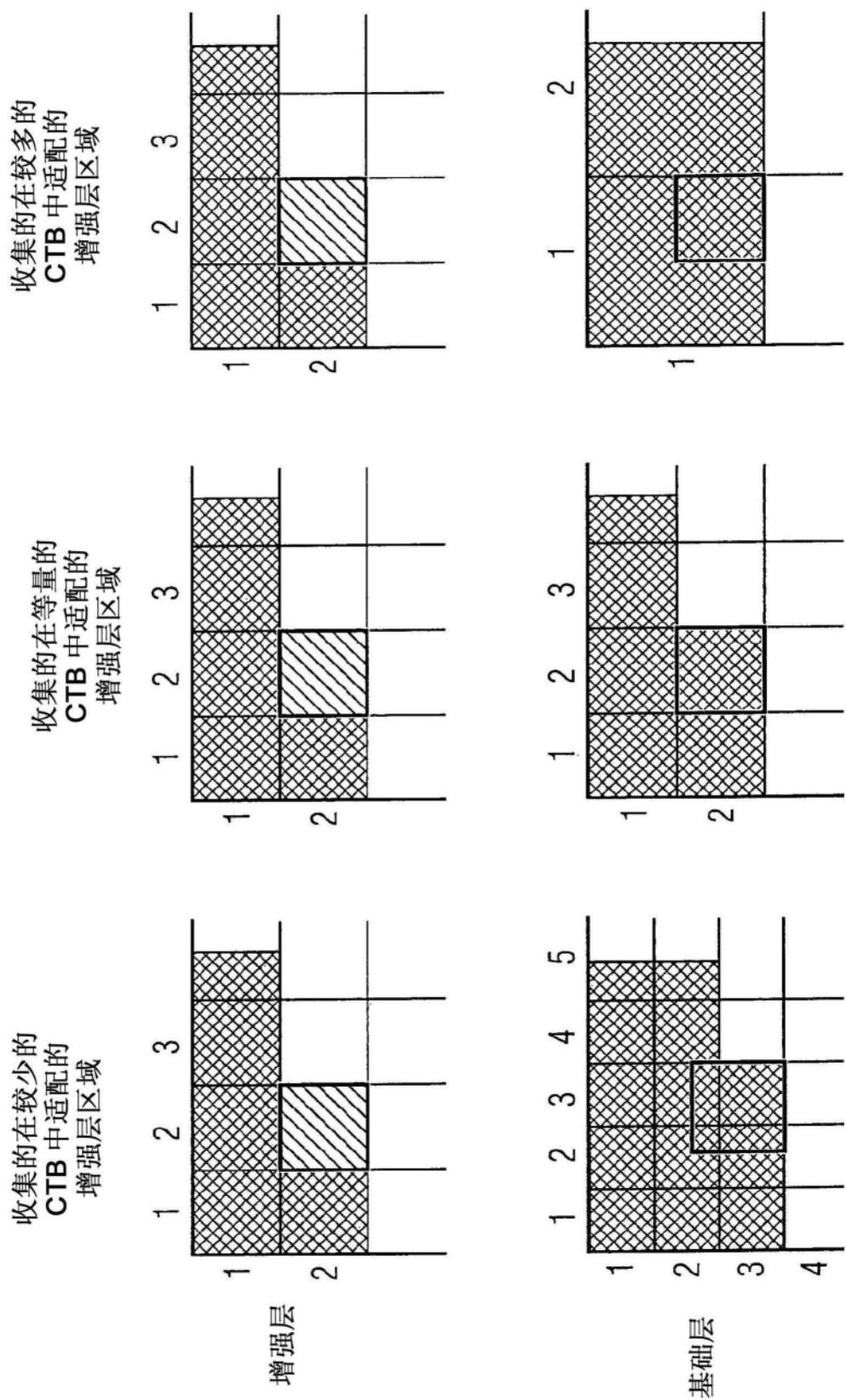


图40

vui_parameters( ) {		描述符
	aspect_ratio_info_present_flag	u(1)
...		
	if( bit stream_restriction_flag ) {	
	tiles_fixed_structure_flag	u(1)
	motion_vectors_over_pic_boundaries_flag	u(1)
	restricted_ref_pic_lists_flag	u(1)
	min_spatial_segmentation_idc	u(8)
	ctb_based_delay_enabled_flag	u(1)
	min_spatial_segment_delay	ue(v)
	max_bytes_per_pic_denom	ue(v)
	max_bits_per_mincu_denom	ue(v)
	log2_max_mv_length_horizontal	ue(v)
	log2_max_mv_length_vertical	ue(v)
}		
}		

图41

vui_parameters( ) {		描述符
	aspect_ratio_info_present_flag	u(1)
...		
	if( bit stream_restriction_flag ) {	
	tiles_fixed_structure_flag	u(1)
	motion_vectors_over_pic_boundaries_flag	u(1)
	restricted_ref_pic_lists_flag	u(1)
	min_spatial_segmentation_idc	u(8)
	<del>min_spatial_segment_delay</del>	<del>ue(v)</del>
	max_bytes_per_pic_denom	ue(v)
	max_bits_per_mincu_denom	ue(v)
	log2_max_mv_length_horizontal	ue(v)
	log2_max_mv_length_vertical	ue(v)
}		
...		
}		

图42

Layer_decoding SEI( payloadSize ) {	描述符
ctb_based_delay_enabled_flag	u(1)
min_spatial_segment_delay	ue(v)
}	

图43

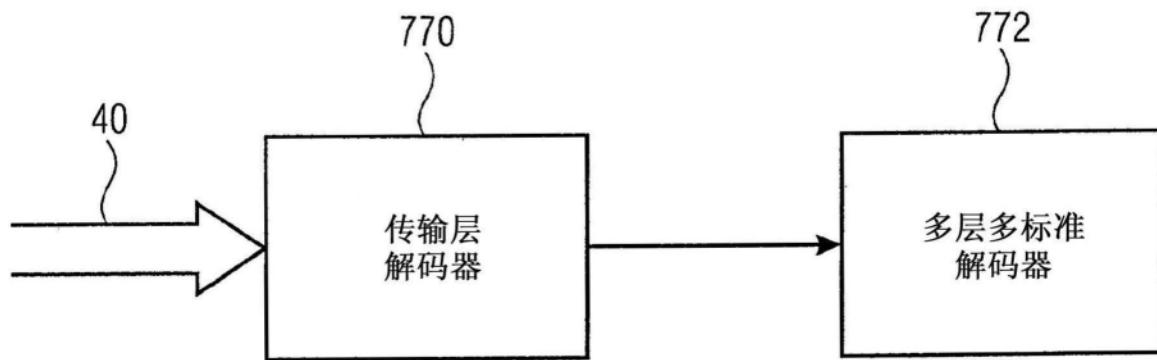


图44