(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117063465 A
(43)申请公布日 2023.11.14

(21)申请号 202180096356.8
(22)申请日 2021.12.10
(30)优先权数据
2021-056391 2021.03.30 JP
(85)PCT国际申请进入国家阶段日
2023.09.25
(86)PCT国际申请的申请数据
PCT/JP2021/045639 2021.12.10
(87)PCT国际申请的公布数据
WO2022/209032 JA 2022.10.06
(71)申请人 日本电气株式会社
地址 日本东京都
(72)发明人 德满健太 蝶野庆一
(74)专利代理机构 中科专利商标代理有限责任公司 11021
专利代理师 庄锦军
(51)Int.Cl.
H04N 19/105(2006.01)

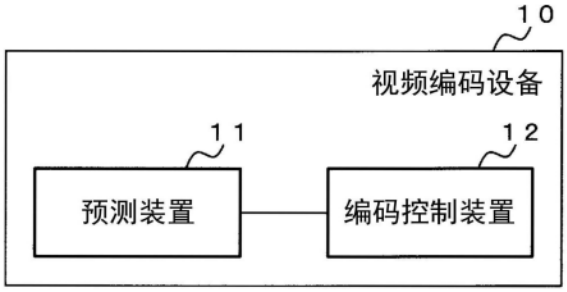
权利要求书2页 说明书11页 附图10页

(54)发明名称

视频编码设备、视频解码设备、视频编码方法、视频解码方法和视频系统

(57)摘要

视频编码设备10包括预测装置11和编码控制装置12，预测装置11用于使用帧内预测或帧间预测来执行预测过程，编码控制装置12用于控制预测装置11，使得在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下，当参考较低层中的图片对图片进行编码时，预测装置11使用在显示顺序中最接近要被编码的图片的图片作为参考图片。



CN 117063465 A

1.一种使用包括多级结构的SOP结构生成比特流的视频编码设备,包括:

预测装置,用于使用帧内预测或帧间预测来执行预测过程,以及

编码控制装置,用于控制所述预测装置,使得在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下,当参考所述较低层中的图片对图片进行编码时,所述预测装置使用在所述显示顺序中最接近要被编码的图片的图片作为参考图片。

2.根据权利要求1所述的视频编码设备,其中,

当使用包括16个帧的L4结构对除了120/P之外的视频格式的视频信号进行编码时,所述编码控制装置控制所述预测装置,使得在所述预定条件下,所述预测装置使用所述较低层中在所述显示顺序中最接近要被编码的图片的图片作为所述参考图片。

3.根据权利要求1所述的视频编码设备,其中,

当使用包括32个帧的L5结构对120/P的视频格式的视频信号进行编码时,所述编码控制装置控制所述预测装置,使得在所述预定条件下,当在基本层中参考所述较低层中的图片对图片进行编码时,所述预测装置使用在所述显示顺序中最接近要被编码的图片的图片作为所述参考图片。

4.根据权利要求1至3中任一项所述的视频编码设备,其中,

所述预测装置基于VVC标准执行所述预测过程。

5.一种输入使用包括多级结构的SOP结构生成的比特流并且执行解码过程的视频解码设备,包括:

预测装置,用于使用帧内预测或帧间预测来执行预测过程,

其中,在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下,当参考所述较低层中的图片对图片进行编码时,所述预测装置使用在所述显示顺序中最接近要被编码的图片的图片作为参考图片。

6.一种用于使用包括多级结构的SOP结构生成比特流的视频编码方法,包括:

使用帧内预测或帧间预测来执行预测过程,以及

控制所述预测过程,使得在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下,当参考所述较低层中的图片对图片进行编码时,使用在所述显示顺序中最接近要被编码的图片的图片作为参考图片。

7.一种用于输入使用包括多级结构的SOP结构生成的比特流并且执行解码过程的视频解码方法,包括:

使用帧内预测或帧间预测来执行预测过程,

其中,在所述预测过程中,在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下,当参考所述较低层中的图片对图片进行编码时,使用在所述显示顺序中最接近要被编码的图片的图片作为参考图片。

8.一种用于使用包括多级结构的SOP结构生成比特流的视频编码程序,使得计算机执行以下操作:

使用帧内预测或帧间预测来执行预测过程,以及

控制所述预测过程,使得在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下,当参考所述较低层中的图片对

图片进行编码时，使用在所述显示顺序中最接近要被编码的图片的图片作为参考图片。

9.一种用于输入使用包括多级结构的SOP结构生成的比特流并且执行解码过程的视频解码程序，使得计算机执行以下操作：

使用帧内预测或帧间预测来执行预测过程，

其中，在所述预测过程中，在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下，当参考较低层中的图片对图片进行编码时，使用在所述显示顺序中最接近要被编码的图片的图片作为参考图片。

10.一种视频系统，包括：

根据权利要求1至4中任一项所述的视频编码设备，

音频编码部分，对音频信号进行编码，以及

复用部分，复用并且输出来自所述视频编码设备的比特流和来自所述音频编码部分的比特流。

视频编码设备、视频解码设备、视频编码方法、视频解码方法和视频系统

技术领域

[0001] 本发明涉及视频编码设备、视频解码设备、视频编码方法、视频解码方法和视频系统。

背景技术

[0002] 例如,在视频内容分发系统中,发射器基于H.264/AVC(高级视频编码)标准或HEVC(高效视频编码)标准对视频信号进行编码,并且接收器执行解码过程以再现视频信号。

[0003] 非专利文献1引入了SOP(图片结构)的概念。SOP是在执行时间可伸缩编码的情况下描述每个AU(访问单元)的编码顺序和参考关系的单元。时间可伸缩编码是能够从多个帧的视频中部分地提取一个帧的编码。一个GOP(图片组)包括一个或多个SOP。

[0004] 非专利文献1指定了适用于除了120/P(逐行)之外的视频格式的SOP结构和适用于120/P视频格式的SOP结构。

[0005] 图12是示出了应用于除了120/P之外的视频格式的SOP结构的说明图。图12示出了当应用重新排序(图片重新排序)时的SOP结构。

[0006] 图12所示的SOP结构包括L=0的结构、L=1的结构、L=2的结构和L=3的结构。如图12所示,Lx(x=0,1,2,3)是以下结构。时间ID是指示每个AU的层的标识信息。L=x的结构在下文中被称为Lx结构。时间ID为m(大于或等于0的自然数)的AU有时被称为时间ID为m的图片。

[0007] =L0结构:仅由时间ID为0的一个或多个图片组成的SOP结构(即,SOP中包括的图片的行数(层数)为1。换句话说,指示最大时间ID的L为0。)

[0008] -L1结构:由时间ID为0的一个或多个图片和时间ID为1的一个或多个图片组成的SOP结构(即,SOP中包括的图片的层数为2。换句话说,指示最大时间ID的L为1。)

[0009] -L2结构:由时间ID为0的一个或多个图片、时间ID为1的一个或多个图片、以及时间ID为2的一个或多个图片组成的SOP结构(即,SOP中包括的图片的层数为3。换句话说,指示最大时间ID的L为2。)

[0010] -L3结构:由时间ID为0的一个或多个图片、时间ID为1的一个或多个图片、时间ID为2的一个或多个图片、以及时间ID为3的一个或多个图片组成的SOP结构(即,SOP中包括的图片的层数为4。换句话说,指示最大时间ID的L为3。)

[0011] 图13是示出了应用于120/P视频格式的SOP结构的说明图。图13所示的SOP结构对应于时间可伸缩编码。

[0012] 图13所示的SOP结构包括L0结构、L1结构、L2结构、L3结构和L4结构。如图13所示,Lx(x=0,1,2,3,4)结构是以下结构。

[0013] -L0结构:仅由时间ID为0的一个或多个图片组成的SOP结构(即,SOP中包括的图片的层数为1。换句话说,指示最大时间ID的L为0。)

[0014] -L1结构:由时间ID为0的一个或多个图片和时间ID为M的一个或多个图片组成的

SOP结构(即,SOP中包括的图片的层数为2。换句话说,指示最大时间ID的L为1(或M)。)

[0015] -L2结构:由时间ID为0的一个或多个图片、时间ID为1的一个或多个图片、以及时间ID为M的一个或多个图片组成的SOP结构(即,SOP中包括的图片的层数为3。换句话说,指示最大时间ID的L为2(或M)。)

[0016] -L3结构:由时间ID为0的一个或多个图片、时间ID为1的一个或多个图片、时间ID为2的一个或多个图片、以及时间ID为M的一个或多个图片组成的SOP结构(即,SOP中包括的图片的层数为4。换句话说,指示最大时间ID的L为3(或M)。)

[0017] -L4结构:由时间ID为0的一个或多个图片、时间ID为1的一个或多个图片、时间ID为2的一个或多个图片、时间ID为3的一个或多个图片、以及时间ID为M的一个或多个图片组成的SOP结构(即,SOP中包括的图片的层数为5。换句话说,指示最大时间ID的L为4(或M)。)

[0018] 非专利文献2公开了一种称为VVC(通用视频编码)的视频编码方法。VVC也被称为ITU-T H.266。在VVC中,编码树单元(CTU)的最大尺寸从HEVC标准的64×64像素(在下文中简称为64×64)扩展到128×128。

[0019] 在非专利文献2中描述的视频编码方法中,数字化视频的每一帧被划分成编码树单元(CTU),并且每个CTU被编码。

[0020] 每个CTU通过要被编码的四叉树(QT)结构或多类型树(MMT)结构被划分成编码单元(CU)。在使用四叉树结构的划分中,CTU在水平和竖直方向上被均等地划分。在使用多类型树结构的划分中,CTU在水平或竖直方向上被划分成两个或三个块。

[0021] 每个CU都是预测编码的。预测编码包括帧内预测和运动补偿预测。基于频率变换对每个CU的预测误差进行变换编码。运动补偿预测是从显示时间不同于要被编码的帧的显示时间的重构图像(参考图片)生成预测图像的预测。在下文中,运动补偿预测也被称为帧间预测。

[0022] 基于运动补偿预测编码的CU被称为帧间CU。仅用帧内CU编码的帧被称为I帧(或I图片)。不仅用帧内CU而且用帧间CU编码的帧被称为P帧(或P图片)。不仅使用一个参考图片而且同时使用两个参考图片进行块的帧间预测的通过帧间CU编码的帧被称为B帧(或B图片)。使用一个参考图片的帧间预测被称为单向预测,而同时使用两个参考图片的帧间预测被称为双向预测。

[0023] 当在同等图像质量下比较时,与基于HEVC标准的编码量相比,基于VVC标准的编码量预计减少30%-50%。

[0024] 引用列表

[0025] 非专利文献

[0026] NPL1:ARIB(无线电工业和商业协会)标准STD-B323.3版,2015年7月3日,无线电工业和商业协会

[0027] NPL2:Benjamin Bross等人的《多功能视频编码(草案10)》(Versatile Video Coding(Draft 10)),JVET-S2001-v7,ITU-T SG16WP3和ISO/IEC JTC1/SC29/WG11联合视频专家组(JVET)第19次会议:电话会议,2020年6月22日至7月1日

发明内容

[0028] 技术问题

[0029] 在图12和图13中，箭头指示要被编码的图片的参考图片。换句话说，箭头指示参考关系。例如，在图12所示的示例中，在L2结构中，显示顺序中的第五AU的图片是指显示顺序中的第三AU的图片和第七AU的图片。在L3结构中，显示顺序中的第四AU的图片是指显示顺序中的第三AU的图片和第五AU的图片。

[0030] 这样，通过从参考图片预测获得的图片可以被用作另一参考图片。在图12和图13所示的分层结构中，多个较低层图片(具有较小时间ID值的层:较浅层)可以被参考用于预测较高层(具有较大时间ID值的层:较深层)中的图片。这种参考关系减少了预测残差。

[0031] 如上所述，与基于HEVC标准的编码相比，基于VVC标准的编码的代码量减少了。当使用非专利文献1中指定的SOP结构时，层越高，图片之间的相关性越高，但是即使在L4结构中，层数也为0、1、2、3和M中的五层。因此，当在基于VVC标准进行编码时使用SOP结构时，编码效率(压缩效率)可能不如预期的高。

[0032] 在应用于60/P视频格式的SOP结构中，层数为0、1、2和3中的四层，即使在L3结构中也是如此，如图12所示。为了增加层数，可以转用图13所示的应用于120/P视频格式的L4结构。然而，L4结构是适于时间可伸缩编码的结构。因此，如果L4结构照原样应用于60/P视频格式，则属于最大时间ID(L4)的层的要被编码的图片和参考图片(属于时间ID＝0至3的层的图片)之间的间隔变大。因此，压缩效率降低。

[0033] 要被编码的图片和参考图片之间的大间隔意味着图片的显示顺序的差异大。换句话说，要被编码的图像和参考图片之间的大间隔意味着图像在时间轴上相距很远。图像之间的间隔在下文中被称为帧间隔。

[0034] 本发明的目的是提供当使用SOP结构执行编码时不会降低压缩效率的视频编码设备、视频解码设备、视频编码方法、视频解码方法和视频系统。

[0035] 问题的解决方案

[0036] 根据本发明的视频编码设备是一种使用包括多级结构的SOP结构生成比特流的视频编码设备，并且包括预测装置和编码控制装置，预测装置用于使用帧内预测或帧间预测执行预测过程，编码控制装置用于控制预测装置，使得在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下，当参考较低层中的图片对图片进行编码时，预测装置使用在显示顺序中最接近要被编码的图片的图片作为参考图片。

[0037] 根据本发明的视频解码设备是一种输入使用包括多级结构的SOP结构生成的比特流并且执行解码过程的视频解码设备，并且包括用于使用帧内预测或帧间预测来执行预测过程的预测装置，其中，在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下，当参考较低层中的图片对图片进行编码时，预测装置使用在显示顺序中最接近要被编码的图片的图片作为参考图片。

[0038] 根据本发明的视频编码方法是一种用于使用包括多级结构的SOP结构来生成比特流的视频编码方法，并且包括使用帧内预测或帧间预测来执行预测过程，以及控制预测过程，使得在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下，当参考较低层中的图片对图片进行编码时，使用在显示顺序中最接近要被编码的图片的图片作为参考图片。

[0039] 根据本发明的视频解码方法是一种用于输入使用包括多级结构的SOP结构生成的

比特流并且执行解码过程的视频解码方法，并且包括使用帧内预测或帧间预测来执行预测过程，其中，在预测过程中，在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下，当参考较低层中的图片对图片进行编码时，使用在显示顺序中最接近要被编码的图片的图片作为参考图片。

[0040] 根据本发明的视频编码程序是一种用于使用包括多级结构的SOP结构生成比特流的视频编码程序，并且使得计算机执行以下操作：使用帧内预测或帧间预测来执行预测过程，以及控制预测过程，使得在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下，当参考较低层中的图片对图片进行编码时，使用在显示顺序中最接近要被编码的图片的图片作为参考图片。

[0041] 根据本发明的视频解码程序是一种用于输入使用包括多级结构的SOP结构生成的比特流并且执行解码过程的视频解码程序，并且使得计算机执行以下操作：使用帧内预测或帧间预测来执行预测过程，其中，在预测过程中，在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下，当参考较低层中的图片对图片进行编码时，使用在显示顺序中最接近要被编码的图片的图片作为参考图片。

[0042] 根据本发明的视频系统包括：上述视频编码设备；音频编码部分，对音频信号进行编码；以及复用部分，复用并且输出来自视频编码设备的比特流和来自音频编码部分的比特流。

[0043] 发明的有益效果

[0044] 根据本发明，当使用SOP结构执行编码时，压缩效率不会降低。

附图说明

[0045] 图1描绘了示出在示例实施例中应用于除了120/P之外的视频格式的SOP结构的说明图。

[0046] 图2描绘了示出在示例实施例中应用于120/P的视频格式的SOP结构中的L5结构的说明图。

[0047] 图3描绘了示出视频编码设备的示例实施例的配置示例的框图。

[0048] 图4描绘了示出视频编码设备的操作的流程图。

[0049] 图5描绘了示出视频解码设备的示例实施例的配置示例的框图。

[0050] 图6描绘了示出视频解码设备的操作的流程图。

[0051] 图7描绘了示出视频系统的示例的框图。

[0052] 图8描绘了示出视频系统的另一示例的框图。

[0053] 图9描绘了示出具有CPU的计算机的示例的框图。

[0054] 图10描绘了示出视频编码设备的主要部分的框图。

[0055] 图11描绘了示出视频解码设备的主要部分的框图。

[0056] 图12描绘了示出应用于除了120/P之外的视频格式的SOP结构的说明图。

[0057] 图13描绘了示出应用于120/P的视频格式的SOP结构中的L5结构的说明图。

具体实施方式

[0058] 在下文中，将参照附图说明视频编码设备的示例实施例。

[0059] 图1是示出在示例实施例中应用于除了120/P之外的视频格式的SOP结构的说明图。在图1所示的SOP结构中，L0、L1、L2和L3结构中的每一层中的图片的布置和参考关系与图12所示的相同。在示例实施例中，除了120/P之外的视频格式的示例是没有应用时间可伸缩编码的60P视频格式。

[0060] 在图1所示的SOP结构中，L4结构被添加到图12所示的L0、L1、L2和L3结构中。L4结构包括16个图片。图1所示的L4结构中的每一层中的图片的布置和参考关系与图13所示的不同。在图13所示的L4结构中，为了实现时间可伸缩编码，以奇数连续编码顺序(也是解码顺序)对放置在最大时间ID(L4)的层中的每个图片进行编码。然而，在图1所示的L4结构中，每个图片的布置没有这样的限制。

[0061] 在图1所示的L4结构中，在编码顺序不与每层的显示顺序相反(在较早显示顺序的图片之前不对较晚显示顺序的图片进行编码)并且较低层(具有较小时间ID值的层)中的图片不参考较高层(具有较大时间ID值的层)中的图片的条件下，参考较低层的图片要被编码的图片使用在显示顺序中最接近要被编码的图片的图片作为参考图片。编码顺序不与显示顺序相反的事实意味着B_a的图片相对于Bb的图片总是显示顺序中的较早图片(a＜b)。此外，参考图片必须在要被编码的图片之前被编码。

[0062] 例如，对于时间ID为4的层，参考由显示顺序为1的B_3指示的图片和由显示顺序为3的B_2指示的图片，对由显示顺序为2的B_5指示的图片进行编码。由B_3指示的图片和由B_2指示的图片在显示顺序中最接近较低层(在该示例中，具有时间ID 0到3的多个层)中由B_5指示的图片。

[0063] 对于时间ID为3的层，参考由显示顺序为3的B_2指示的图片和由显示顺序为7的B_1指示的图片，对由显示顺序为5的B_6指示的图片进行编码。由显示顺序为3的B_2指示的图片和由显示顺序为7的B_1指示的图片在显示顺序中最接近较低层(在该示例中，具有时间ID 0到2的多个层)中由B_6指示的图片。

[0064] L4结构中的编码顺序不限于图1所示的编码顺序。在编码顺序不与每层的显示顺序相反并且较低层中的图片不参考较高层中的图片的条件下，当参考较低层图片要被编码的图片可以使用在显示顺序中最接近要被编码的图片的图片作为参考图片时，可以使用其他编码顺序。

[0065] 在该示例实施例中，由于在显示顺序中最接近要被编码的图片的图片是参考图片，因此与应用于图13所示的120/P视频格式的SOP结构中的L4结构相比，要被编码的图片和参考图片之间的帧间隔较小。因此，图片的压缩效率比使用图13所示的L4结构时更高。

[0066] 图2是示出应用于120/P视频格式的SOP结构中的L5结构的说明图。L5结构包括32个图片。在该示例实施例中，应用于120/P视频格式的SOP结构也是实现时间可伸缩编码的结构。尽管在图2中省略了，但L0、L1、L2和L4结构与图13所示的相同。换句话说，L5结构被添加到图13所示的L0、L1、L2、L3和L4结构。

[0067] 在图2所示的L5结构中，具有时间ID＝0至4的层中的图片的编码顺序和参考关系与图1所示的L4结构中的相同。然而，由于图2所示的L5结构包括32个图片，因此图2中指示解码顺序(也是编码顺序)的数量是图1中指示解码顺序的数量的两倍。

[0068] 当使用应用于120/P视频格式的SOP结构中的L5结构时，属于基本层的图片(在这种情况下，具有时间ID＝0至4的层中的图片)的压缩效率更高，正如当使用图1所示的示例实施例中的L4结构时，每层(具有时间ID＝0至4的层中的每个图片)的压缩效率更高。

[0069] L5结构中的编码顺序不限于图2所示的编码顺序。如果参考较低层的图片要被编码的图片可以在预定条件下使用在显示顺序中最接近要被编码的图片的图片作为时间ID＝0到4的层的参考图片，则可以使用其他编码顺序。预定条件是编码顺序不与每层的显示顺序相反，并且较低层中的图片不参考较高层中的图片。

[0070] 将L4结构添加到应用于60/P视频格式的SOP结构以及将L5结构添加到应用于120/P视频格式的SOP结构提高了图片压缩的效率。这是因为，由于一般来说，在包括较高层(60/P视频格式的L4，120/P视频格式的LM)的结构中，较大量的代码被分配给解码顺序中的更频繁参考的第0图片(通常是I或P图片)，解码顺序中的第0图片的出现频率变得相对较低。

[0071] 示例实施例1。

[0072] 图3是示出视频编码设备的示例实施例的配置示例的框图。图3所示的视频编码设备包括排序单元101、减法器102、变换器/量化器103、逆量化器/逆变换器104、熵编码器105、加法器106、缓冲器107、预测器108、编码控制器109和复用器110。图1所示的视频编码设备基于VVC标准执行编码过程，但是也可以基于诸如HEVC标准、H.264/AVC标准和MPEG(运动图像专家组)-2标准之类的其他标准执行编码过程。

[0073] 排序单元101是存储按照显示顺序输入的视频信号中的每个图像(图片)的存储器。在该示例实施例中，假设每个图片按照输入顺序存储。换句话说，假设按照显示顺序输入的每个图片从存储器中的最小地址开始存储。然而，按照显示顺序输入的图片也可以按照编码顺序存储在排序单元101中。换句话说，可以从存储器中的最小地址开始，按照编码顺序存储每个输入图片。

[0074] 不管采用哪种存储器存储方法(显示顺序或编码顺序)，都按照编码顺序从排序单元101读取每个图片。在下文中，图片有时被称为输入视频信号。

[0075] 当使用SOP结构执行编码时，利用排序单元101。当不使用SOP结构时，输入视频信号中的每个图片被原样提供给减法器102。

[0076] 减法器102从由排序单元101读取的输入视频信号(具体地，像素值)中减去预测信号，以生成预测误差信号。预测误差信号也被称为预测残差或预测残差信号。

[0077] 变换器/量化器103对预测误差信号进行频率变换。此外，变换器/量化器103量化频率变换后的预测误差信号(变换系数)。在下文中，量化的变换系数被称为变换量化值。

[0078] 熵编码器105对预测参数和变换量化值进行熵编码。预测参数是与CTU(编码树单元)和块预测相关的信息，诸如预测模式(帧内预测、帧间预测)、帧内预测块大小、帧内预测方向、帧间预测块大小和运动矢量。

[0079] 复用器110复用由熵编码器105提供的熵编码数据和来自编码控制器109的数据(编码信息等)以将它们作为比特流输出。

[0080] 预测器108生成输入视频信号的预测信号。预测器108基于帧内预测或帧间预测生成预测信号。也就是说，对于作为编码单位的每个块(单位)，预测器108使用帧内预测或帧间预测来生成预测信号。

[0081] 逆量化器/逆变换器104对变换量化值进行逆量化，以恢复变换系数。此外，逆量化

器/逆变换器104对逆量化变换系数进行逆频率变换,以恢复预测误差信号。加法器106将恢复的预测误差信号和预测信号相加,以生成重构图像。重构图像被提供给缓冲器107。缓冲器107存储重构图像。缓冲器107对应于用于存储帧内预测的参考块的块存储器和用于存储帧间预测的参考图片的帧存储器。

[0082] 编码控制器109从视频编码设备外部输入编码信息。编码信息包括所使用的编码方法(VVC标准、HEVC标准、H.264/AVC标准、MPEG-2)、测试序列信息(60/P、120/P等)、可伸缩编码可用性等。编码控制器109基于编码信息控制视频编码设备中的每个块。

[0083] 接下来,将参考图4中的流程图说明当视频编码设备使用SOP结构执行编码时的操作。

[0084] 首先,按照显示顺序输入的视频信号中的每个图片被存储在排序单元101中(步骤S101)。

[0085] 排序单元101根据编码控制器109的指令按照编码顺序将图片顺序地输出到减法器102(步骤S102)。

[0086] 当外部指定遵循60/P视频格式时,编码控制器109进行控制,使得按照图1所示的解码顺序(也是编码顺序)从排序单元101读取图片。即,当每个图片按照显示顺序存储在排序单元101中时,编码控制器109进行控制,使得从存储器的要被读取的地址输出要被编码的图片。当每个图片按照编码顺序存储在排序单元101中时,编码控制器109指示排序单元101,使得当图片被写入排序单元101时,图片被存储在存储器的对应于编码顺序的地址处。

[0087] 当外部指定遵循120/P视频格式时,编码控制器109进行控制,使得按照图12所示的解码顺序(在L0到L4结构的情况下)或者按照图2所示的解码顺序(在L5结构的情况下)从排序单元101读取图片。

[0088] 作为示例,编码控制器109可以根据视频的场景的情况来确定使用L0至L4结构中的哪一个(在60/P的情况下)或者L0至L5结构中的哪一个(在120/P的情况下)。例如,编码控制器109确定对构成整个屏幕移动不太多的场景图像的图像(图片)使用具有较小x值的Lx结构,并且对构成整个屏幕快速移动的场景图像的图像使用具有较大x值的Lx结构。在这种情况下,编码控制器109中包括预先检测图像中的运动程度的功能。

[0089] 预测器108基于帧内预测或帧间预测生成输入视频信号的预测信号(步骤S103)。另外,减法器102生成预测误差信号(步骤S103)。

[0090] 编码控制器109指示预测器108根据图1所示的图片参考关系(在60/P的情况下)或图12和图2所示的图片参考关系(在120/P的情况下)来执行编码。

[0091] 变换器/量化器103对预测误差信号进行频率变换,以生成变换系数(步骤S104)。此外,变换器/量化器103用量化步长对变换系数进行量化,以生成变换量化值(步骤S105)。变换量化值被输入到逆量化器/逆变换器104和熵编码器105。

[0092] 逆量化器/逆变换器104对变换量化值进行逆量化,并且对逆量化的变换量化值进行逆频率变换(步骤S106)。熵编码器105对变换量化值进行熵编码(例如,算术编码)以生成熵编码数据(步骤S107)。

[0093] 对包括SOP的所有图片执行步骤S102至S107的过程(步骤S108)。

[0094] 复用器110复用由熵编码器105提供的熵编码数据和来自编码控制器109的数据(编码信息等)以将它们作为比特流输出。

[0095] 在该示例实施例中，当使用应用于除了120/P之外的视频格式的SOP结构中的L4结构时，在编码顺序不与L4结构中的每层的显示顺序相反并且较低层中的图片不参考较高层中的图片的条件下，编码控制器109进行控制，使得在参考较低层中的图片对图片进行编码时，预测器108使用在显示顺序中最接近要被编码的图片的图片作为参考图片。这种控制提高了SOP中的每个图片的压缩效率。在该示例实施例中，为了实现这种控制，编码控制器109使得排序单元101按照图1所示的编码顺序输出图片。

[0096] 当使用应用于120/P视频格式的SOP结构中的L5结构时，在L5结构中具有时间ID＝0到4的每层的编码顺序不与显示顺序相反并且较低层中的图片不参考较高层中的图片的条件下，编码控制器109控制预测器108，使得参考较低层的图片要被编码的图片使用在显示顺序中最接近要被编码的图片的图片作为参考图片。这种控制增加了属于基本层的图片(在这种情况下，具有时间ID＝0到4的层中的图片)的压缩效率。在该示例实施例中，为了实现这种控制，编码控制器109使得排序单元101按照图12所示的解码顺序(在L0至L4结构的情况下)或者根据图2所示的解码顺序(在L5结构的情况下)输出图片。

[0097] 示例实施例2。

[0098] 图5是示出了视频解码设备的示例实施例的配置示例的框图。图5所示的视频解码设备包括解复用器201、熵解码器202、逆量化器/逆变换器203、加法器204、预测器205、缓冲器206、解码控制器207和排序单元208。例如，图5所示的视频解码设备可以从图3所示的视频编码设备接收比特流。

[0099] 解复用器201解复用输入比特流并且提取熵编码数据。它还将比特流中包括的编码信息等输出到解码控制器207。

[0100] 熵解码器202对熵编码数据进行熵解码。熵解码器202将熵解码的变换量化值提供给逆量化器/逆变换器203。熵解码器202还将比特流中包括的预测参数提供给预测器205。熵解码器202将比特流中包括的编码信息提供给解码控制器207。

[0101] 逆量化器/逆变换器203对变换后的量化值进行逆量化。此外，逆量化器/逆变换器203对逆量化的频率变换系数进行逆频率量化。

[0102] 预测器205基于预测参数生成每个子块的预测信号。由逆量化器/逆变换器203进行逆频率变换的预测误差信号通过加法器204与由预测器205提供的预测信号相加，然后作为重构图像提供给缓冲器206。缓冲器206存储重构图像。

[0103] 存储在缓冲器206中的重构图像被传送到排序单元208。排序单元208是存储按照解码顺序输入的视频信号中的每个图像(图片)的存储器。在该示例实施例中，假设每个图片按照解码顺序存储。换句话说，假设按照解码顺序输入的每个图片从存储器中的最小地址开始存储。然而，按照解码顺序输入的每个图片可以按照显示顺序存储在排序单元101中。换句话说，可以从存储器中的最小地址开始，按照显示顺序存储每个输入图片。

[0104] 不管采用哪种存储器存储方法(解码顺序或显示顺序)，都按照显示顺序从排序单元208读取每个图片。

[0105] 接下来，将参考图6中的流程图说明当使用SOP结构执行解码时视频解码设备的操作。

[0106] 熵解码器202对比特流中包括的熵编码数据进行熵解码(步骤S201)。

[0107] 逆量化器/逆变换器203按照量化步长对变换量化值进行逆量化(步骤S202)。此

外,逆量化器/逆变换器203对逆量化的频率变换系数进行逆频率变换(步骤S203)。

[0108] 预测器205使用存储在缓冲器206中的重构图像生成预测信号(步骤S204)。加法器204将由预测器205提供的预测信号与由逆量化器/逆变换器203逆频率变换的预测误差信号相加,以生成重构图像(步骤S204)。重构图像存储在缓冲器206中。

[0109] 将存储在缓冲器206中的重构图像传送到排序单元208(步骤S205)。

[0110] 对包括SOP的所有图片执行步骤S201至S205的过程(步骤S206)。

[0111] 排序单元208根据解码控制器207的输出指令按照显示顺序输出每个图像(步骤S207)。

[0112] 当编码信息指定遵循60/P视频格式时,解码控制器207进行控制,使得按照图1所示的显示顺序从排序单元101读取图片。即,当每个图片按照解码顺序存储在排序单元208中时,解码控制器207进行控制,使得从存储器的要被读取的地址(存储根据显示顺序要被读取的图片的地址)输出图片。当每个图片按照显示顺序存储在排序单元101中时,解码控制器207指示排序单元208,使得当图片被写入排序单元208时,图片根据显示顺序被存储在存储器的地址处。

[0113] 当外部指定遵循120/P视频格式时,解码控制器207进行控制,使得按照图12所示的显示顺序(在L0至L4结构的情况下)或者按照图2所示的显示顺序(在L5结构的情况下)从排序单元208读取图片。

[0114] 当接收基于由第一示例实施例的视频编码设备使用SOP结构编码的编码数据的比特流时,视频解码设备可以针对每个图片以高压缩效率从编码数据再生视频。

[0115] 也就是说,该示例实施例的视频解码设备可以从视频编码设备接收比特流,该视频编码设备被配置为使得在编码顺序不与显示顺序相反并且较低层中的图片不参考较高层中的图片的条件下,其预测装置在参考较低层的图片对图片进行编码时,使用在显示顺序中最接近要被编码的图片的图片作为参考图片。当接收到这样的比特流时,在该示例实施例的视频解码设备中,在编码顺序不与显示顺序相反并且较低层中的图片不参考较高层中的图片的条件下,当参考较低层的图片对图片进行编码时,预测器205可以使用在显示顺序中最接近要被编码的图片的图片作为参考图片。

[0116] 示例实施例3。

[0117] 图7是描绘了视频系统的示例的框图。图7所示的视频系统是以下系统,其中上述示例实施例中的视频编码设备100和视频解码设备200通过无线传输路径或有线传输路径300连接。

[0118] 视频编码设备100的配置和操作与图3所示的视频编码设备的配置和操作相同。视频解码设备200的配置和操作与图5所示的视频解码设备的配置和操作相同。

[0119] 示例实施例4。

[0120] 图8是描绘了视频系统的另一个示例的框图。图8所示的视频系统包括音频编码部分401、视频编码部分402和复用部分403。

[0121] 音频编码部分401基于例如在ARIB STD-B32标准中定义的MPEG-4AAC(高级音频编码)标准或MPEG-4ALS(音频无损编码)标准,对包括视频和音频的数据(内容)中的音频信号进行编码,以生成并输出音频比特流。

[0122] 视频编码部分402例如如图3所示配置,并且生成并输出视频比特流。

[0123] 例如,复用部分403通过基于ARIB STD-B32标准复用音频比特流、视频比特流和其他信息来生成并输出比特流。

[0124] 尽管可以通过硬件来配置上述示例实施例,但是它们可以通过计算机程序来实现。

[0125] 图9是示出了具有CPU(中央处理单元)的计算机的示例的框图。计算机在图3所示的视频编码设备和图5所示的视频解码设备的每一个中实现。CPU 1000根据存储在存储设备1001中的程序(视频编码程序或视频解码程序)执行处理,以实现上述示例实施例中的功能。

[0126] 即,当在图3所示的视频编码设备中实现计算机时,CPU 1000实现视频编码设备中的减法器102、变换器/量化器103、逆量化器/逆变换器104、熵编码器105、加法器106、预测器108、编码控制器109和复用器110的功能。

[0127] 当在图5所示的视频解码设备中实现计算机时,CPU 1000实现视频解码设备中的解复用器201、熵解码器202、逆量化器/逆变换器203、加法器204、预测器205和解码控制器207的功能。

[0128] 存储设备1001例如是非暂时性计算机可读介质。非暂时性计算机可读介质是各种类型的有形存储介质之一。非暂时性计算机可读介质的具体示例包括磁存储介质(例如,硬盘)、CD-ROM(光盘只读存储器)、CD-R(可记录光盘)、CD-R/W(可重写光盘)和半导体存储器(例如,掩模ROM、PROM(可编程ROM)、EPROM(可擦除PROM)、闪存ROM)。

[0129] 程序可以存储在各种类型的暂时性计算机可读介质中。通过例如有线或无线通信信道(即,通过电信号、光信号或电磁波)向临时计算机可读介质提供程序。

[0130] 存储器1002是例如由RAM(随机存取存储器)实现的存储装置,并且当CPU 1000执行处理时临时存储数据。可以假设保存在存储设备1001或临时计算机可读介质中的程序被传送到存储器1002,并且CPU 1000基于存储器1002中的程序执行处理。

[0131] 存储器1002可以用作图3所示的排序单元101和缓冲器107。存储器1002也可以用作图5所示的缓冲器206和排序单元208。

[0132] 图10是示出了视频编码设备的主要部分的框图。图10所示的视频编码设备10是使用SOP结构生成比特流的视频编码设备,该SOP结构包括多级(例如,L0-L4、L0-L5)结构(例如,L0-L4结构、L0-L5结构),并且包括预测装置11(在示例实施例中,由预测器108实现)和编码控制装置12(在示例实施例中,由编码控制器109实现),预测装置用于使用帧内预测或帧间预测执行预测过程,编码控制装置用于控制预测装置11,使得在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层中的图片不参考较高层中的图片的预定条件下,当参考较低层中的图片对图片进行编码时,预测装置11使用在显示顺序中最接近要被编码的图片的图片作为参考图片。

[0133] 图11是示出了视频解码设备的主要部分的框图。图11所示的视频解码设备20是输入使用包括多级(例如,L0-L4、L0-L5)结构(例如,L0-L4结构、L0-L5结构)的SOP结构生成的比特流并且执行解码过程的视频解码设备,并且包括预测装置21(在示例实施例中,由预测器205实现),该预测装置21用于使用帧内预测或帧间预测来执行预测过程,其中,在较后显示顺序的图片不在较前显示顺序的图片之前被编码并且较低层(具有较小时间ID值的层)中的图片不参考较高层(具有较大时间ID值的层)中的图片的预定条件下,当参考较低层中

的图片对图片进行编码时，预测装置21使用在显示顺序中最接近要被编码的图片的图片作为参考图片。

[0134] 附图标记列表

[0135] 11 预测装置

[0136] 12 编码控制装置

[0137] 21 预测装置

[0138] 10、100 视频编码设备

[0139] 101 排序单元

[0140] 102 减法器

[0141] 103 变换器/量化器

[0142] 104 逆量化器/逆变换器

[0143] 105 熵编码器

[0144] 106 加法器

[0145] 107 缓冲器

[0146] 108 预测器

[0147] 109 编码控制器

[0148] 110 复用器

[0149] 20、200 视频解码设备

[0150] 201 解复用器

[0151] 202 熵解码器

[0152] 203 逆量化器/逆变换器

[0153] 204 加法器

[0154] 205 预测器

[0155] 206 缓冲器

[0156] 207 解码控制部分

[0157] 208 排序单元

[0158] 401 音频编码部分

[0159] 402 视频编码部分

[0160] 403 复用部分

[0161] 1000 CPU

[0162] 1001存储设备

[0163] 1002存储器。

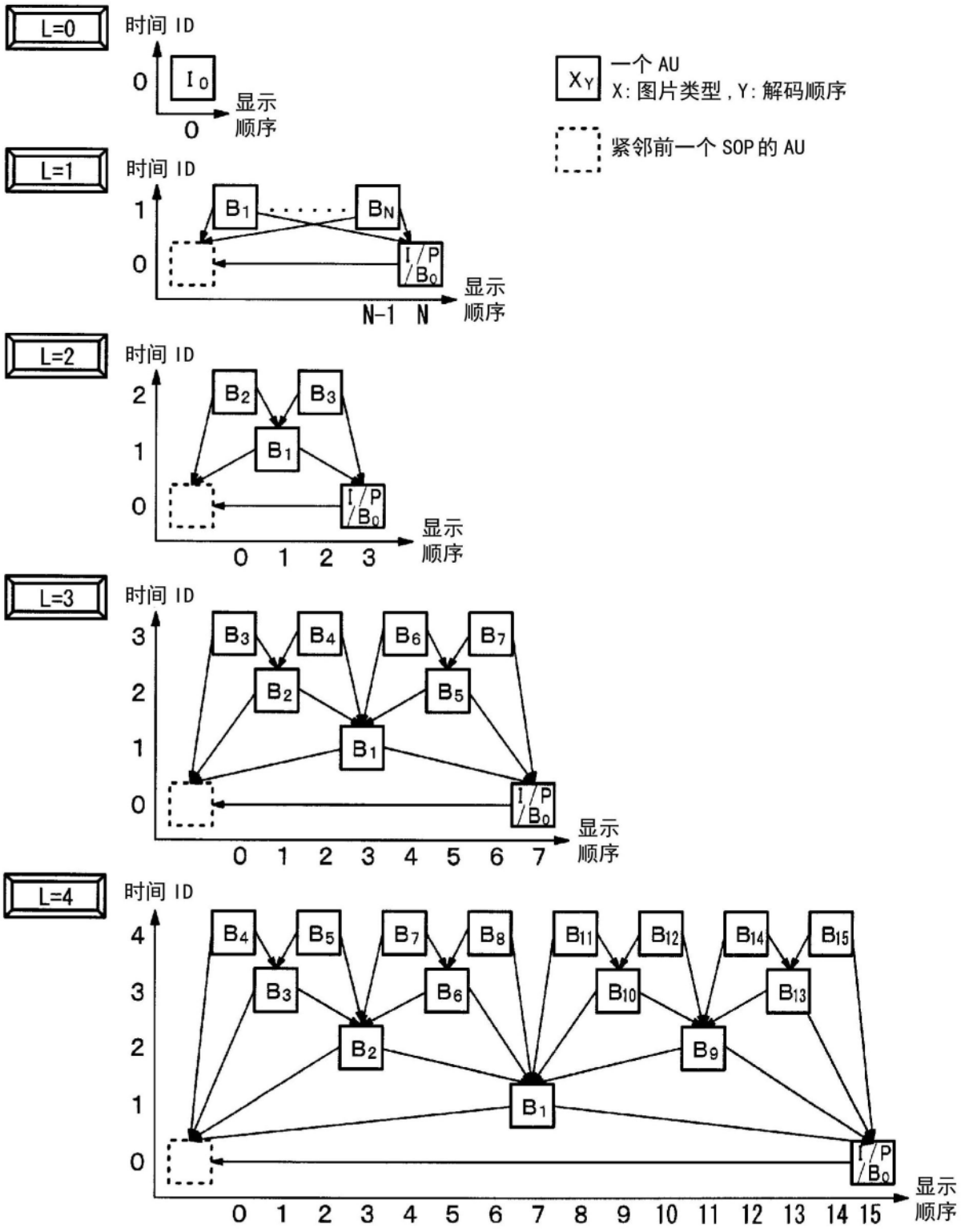

L=0
时间 ID
0
I_0
显示顺序
0
X_Y
一个 AU
X: 图片类型 , Y: 解码顺序
紧邻前一个 SOP 的 AU
L=1
时间 ID
1
0
B_1
B_N
I/P/B_0
显示顺序
N-1 N
L=2
时间 ID
2
1
0
B_2
B_3
B_1
I/P/B_0
显示顺序
0 1 2 3
L=3
时间 ID
3
2
1
0
B_3
B_4
B_6
B_7
B_2
B_5
B_1
I/P/B_0
显示顺序
0 1 2 3 4 5 6 7
L=4
时间 ID
4
3
2
1
0
B_4
B_5
B_7
B_8
B_{11}
B_{12}
B_{14}
B_{15}
B_3
B_6
B_{10}
B_{13}
B_2
B_9
B_1
I/P/B_0
显示顺序
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

图1

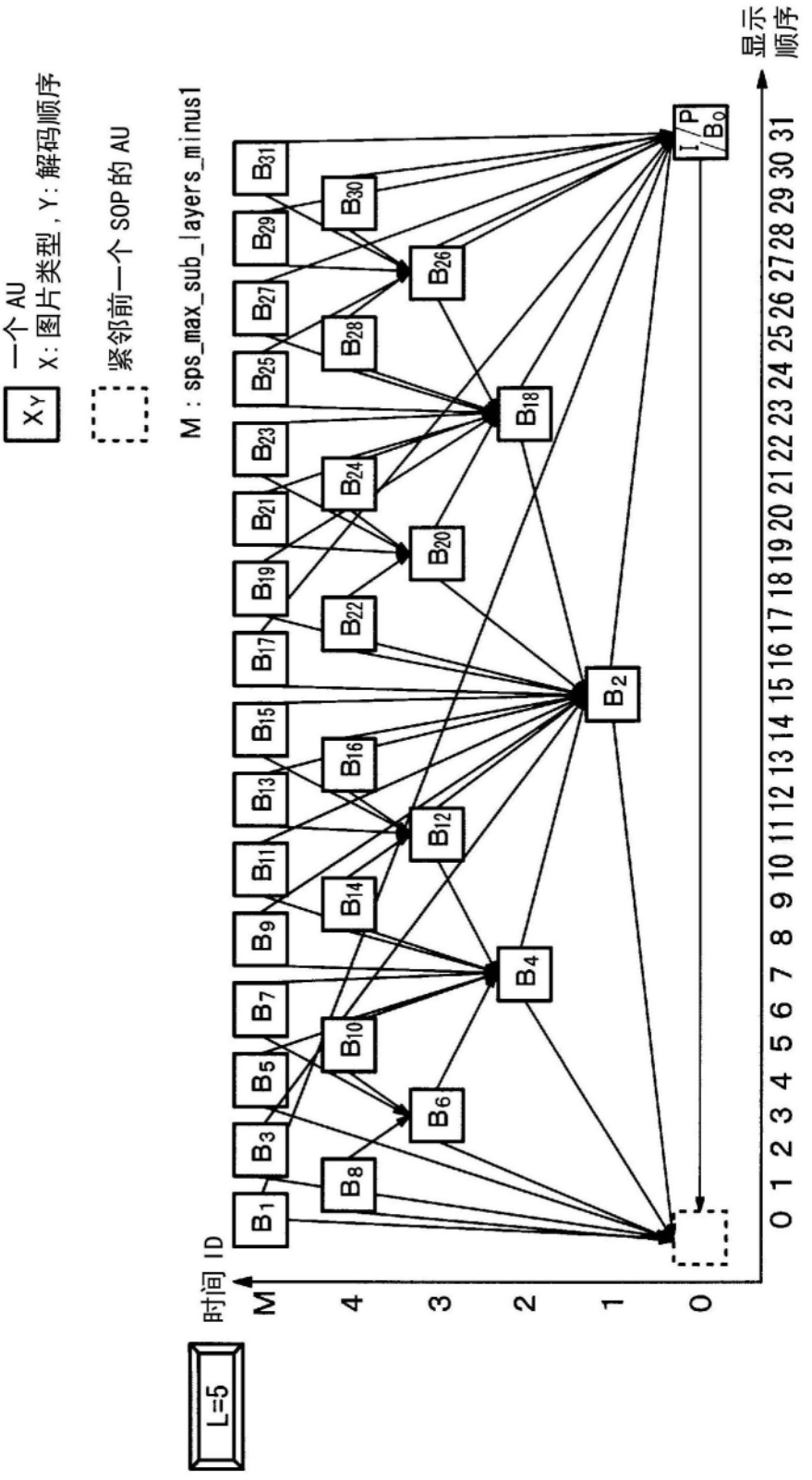

XY 一个 AU
X: 图片类型，Y: 解码顺序
紧邻前一个 SOP 的 AU
M : sps_max_sub_layers_minus1
时间 ID
M
4
3
2
1
0
L=5
显示顺序
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
I / P / B0
B1 B3 B5 B7 B9 B11 B13 B15 B17 B19 B21 B23 B25 B27 B29 B31
B8 B6 B10 B4 B14 B12 B16 B2 B22 B20 B24 B18 B28 B26 B30

图2

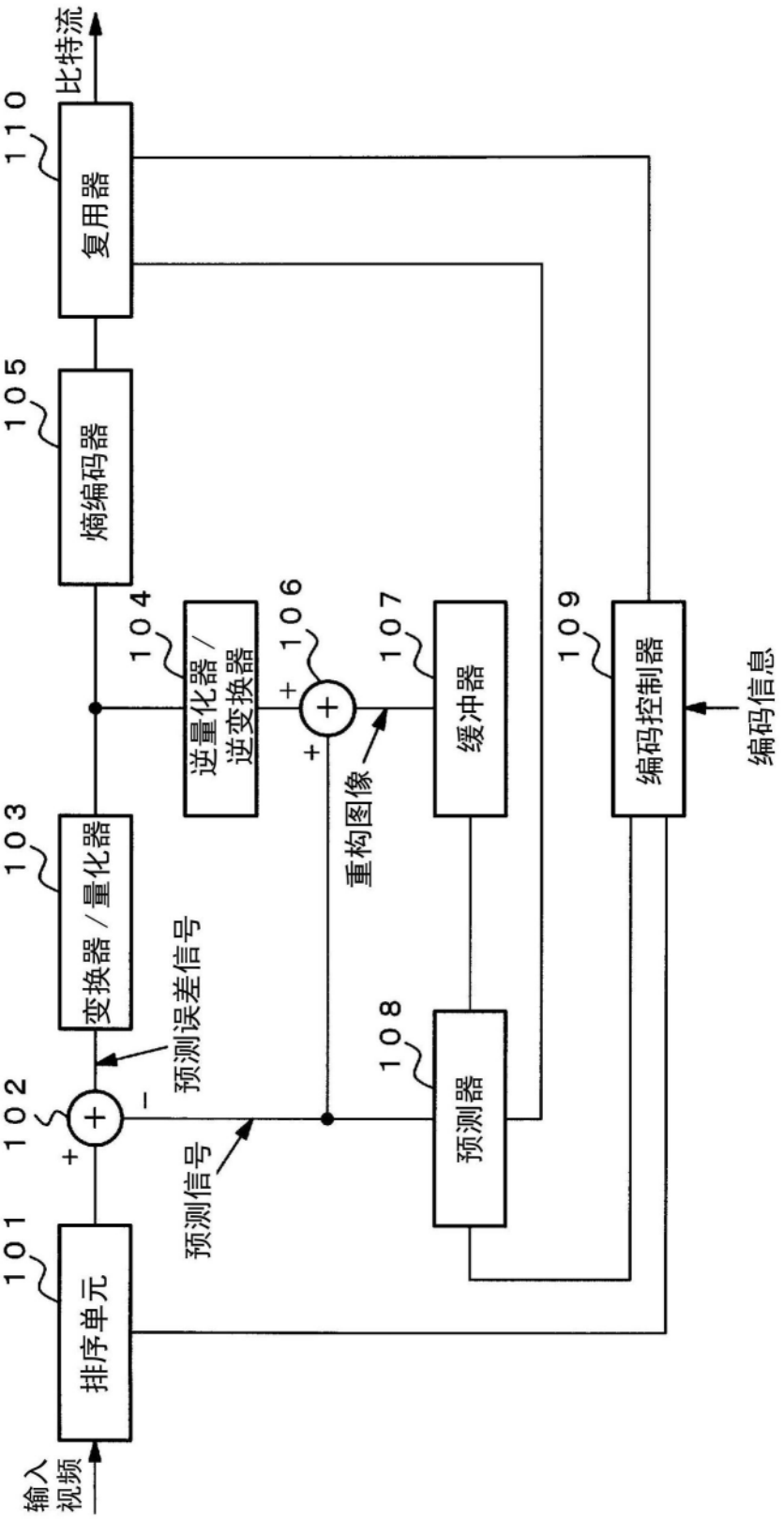

比特流
110
复用器
105
熵编码器
104
逆量化器 / 逆变换器
106
107
缓冲器
109
编码控制器
编码信息
103
变换器 / 量化器
预测误差信号
重构图像
102
108
预测器
预测信号
101
排序单元
输入视频

图3

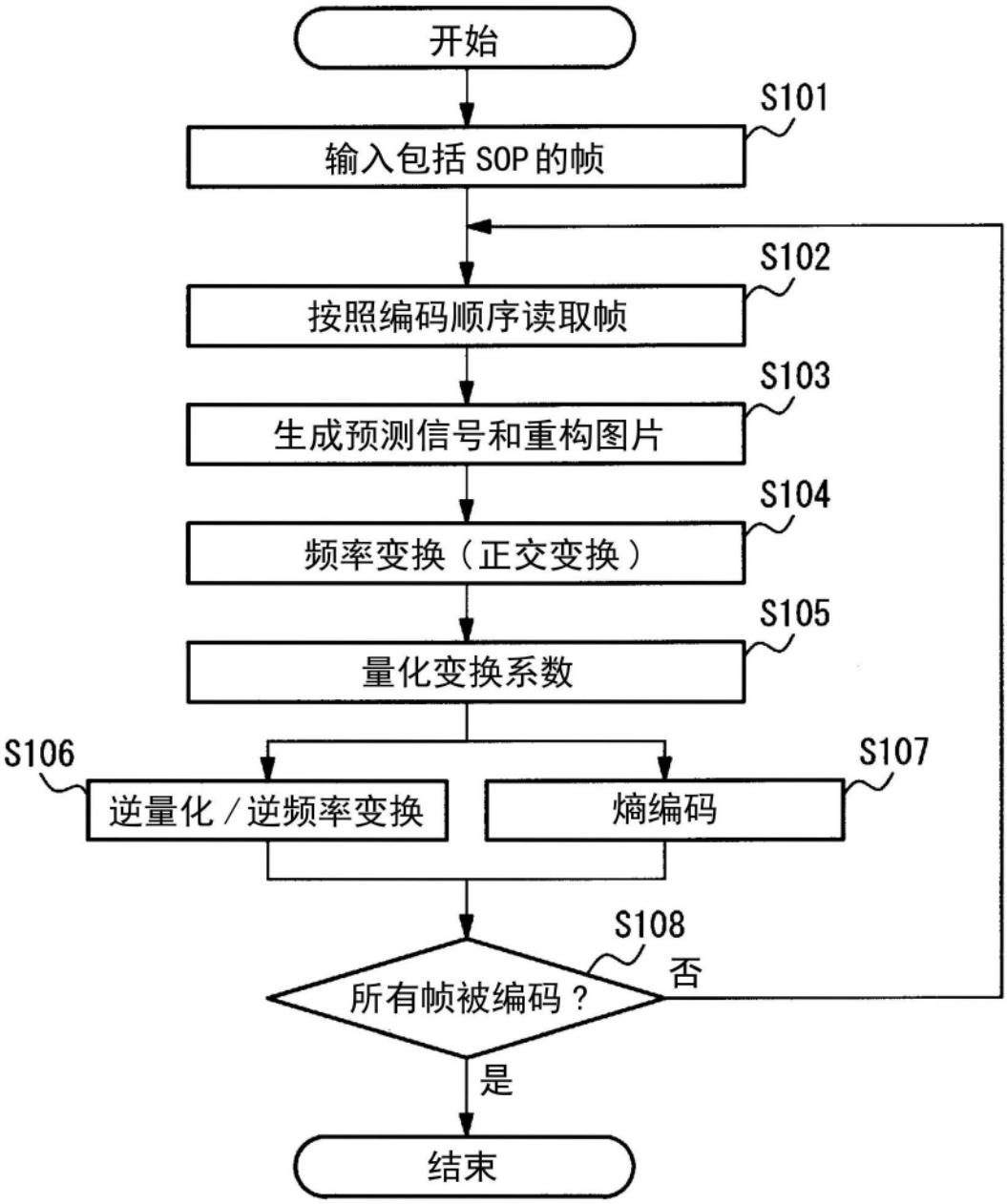

开始
S101
输入包括SOP的帧
S102
按照编码顺序读取帧
S103
生成预测信号和重构图片
S104
频率变换（正交变换）
S105
量化变换系数
S106
逆量化／逆频率变换
S107
熵编码
S108
所有帧被编码？
否
是
结束

图4

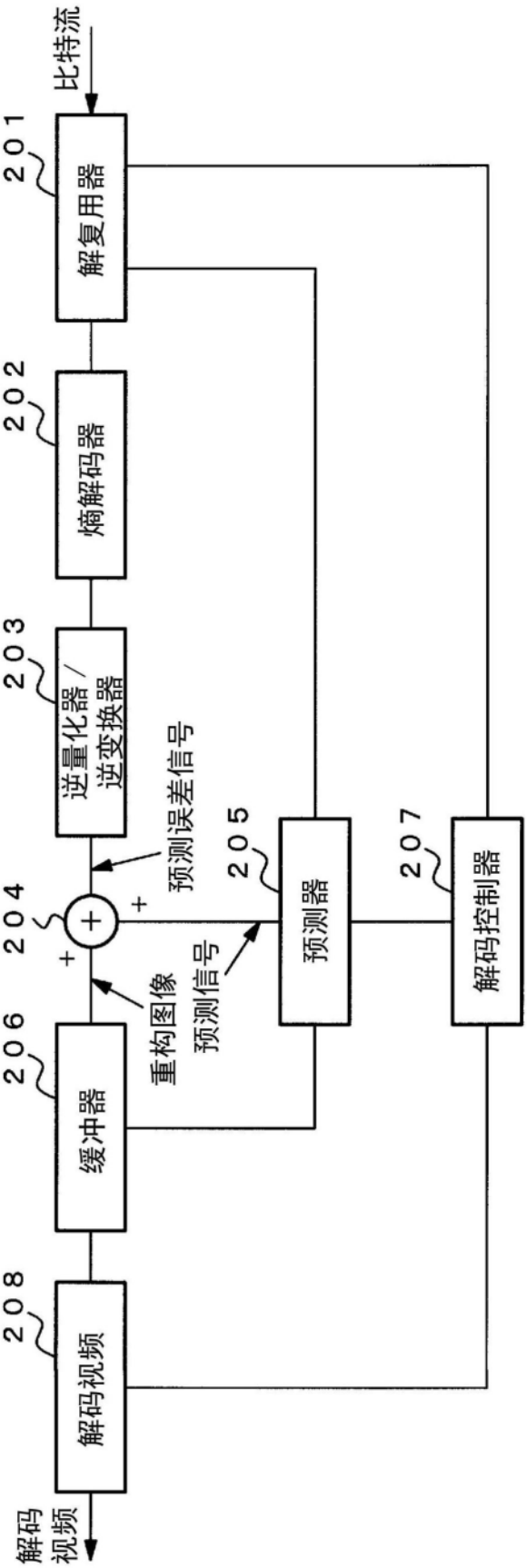

比特流
201
解复用器
202
熵解码器
203
逆量化器/
逆变换器
预测误差信号
+
204
+
205
预测器
207
解码控制器
重构图像
预测信号
206
缓冲器
208
解码视频
解码
视频

图5

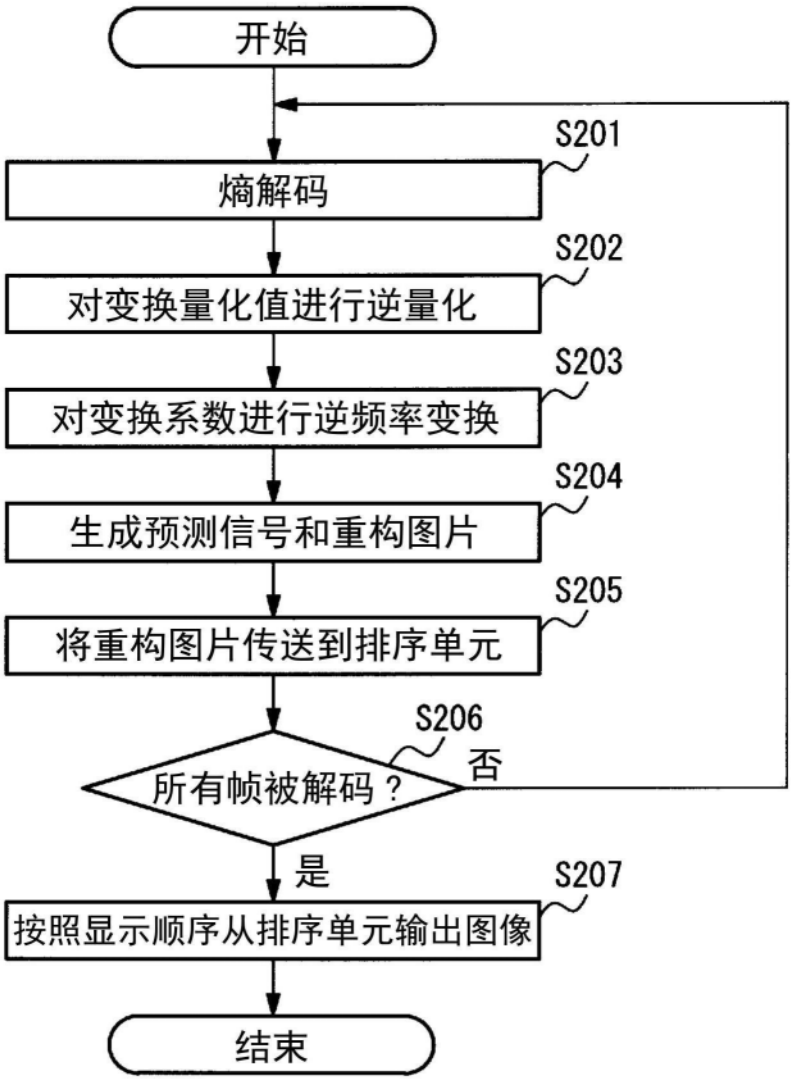

开始
S201
熵解码
S202
对变换量化值进行逆量化
S203
对变换系数进行逆频率变换
S204
生成预测信号和重构图片
S205
将重构图片传送到排序单元
S206
所有帧被解码？
否
是
S207
按照显示顺序从排序单元输出图像
结束

图6

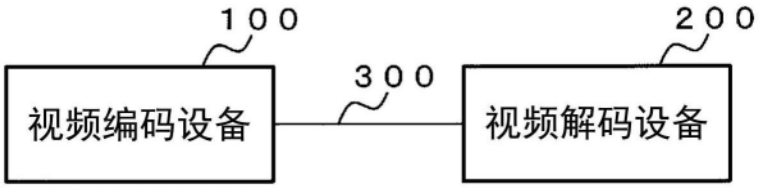

100
200
300
视频编码设备
视频解码设备

图7

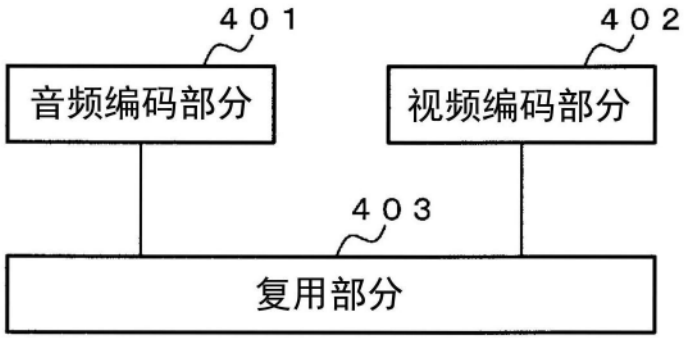

401
402
音频编码部分
视频编码部分
403
复用部分

图8

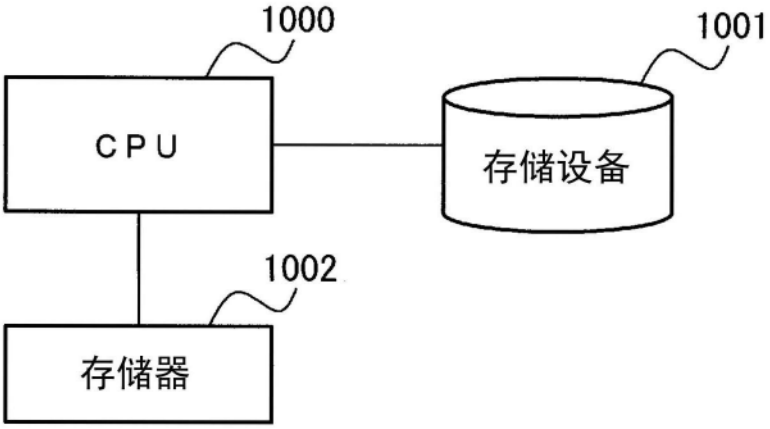

1000
1001
CPU
存储设备
1002
存储器

图9

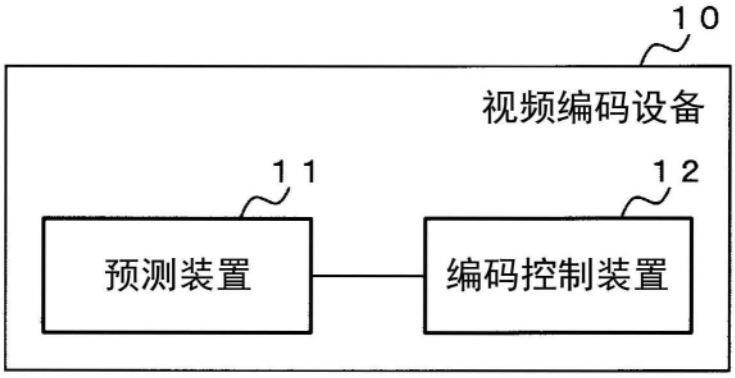

10
视频编码设备
11
12
预测装置
编码控制装置

图10

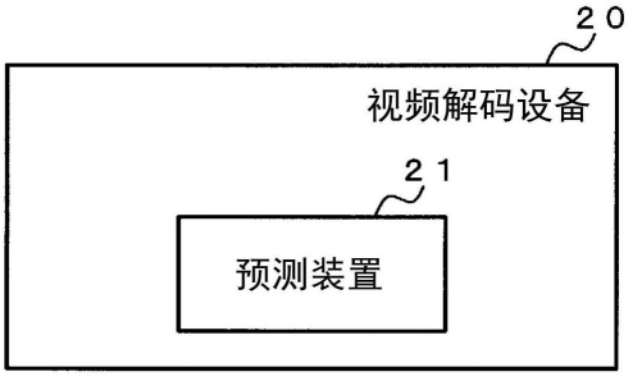


图11

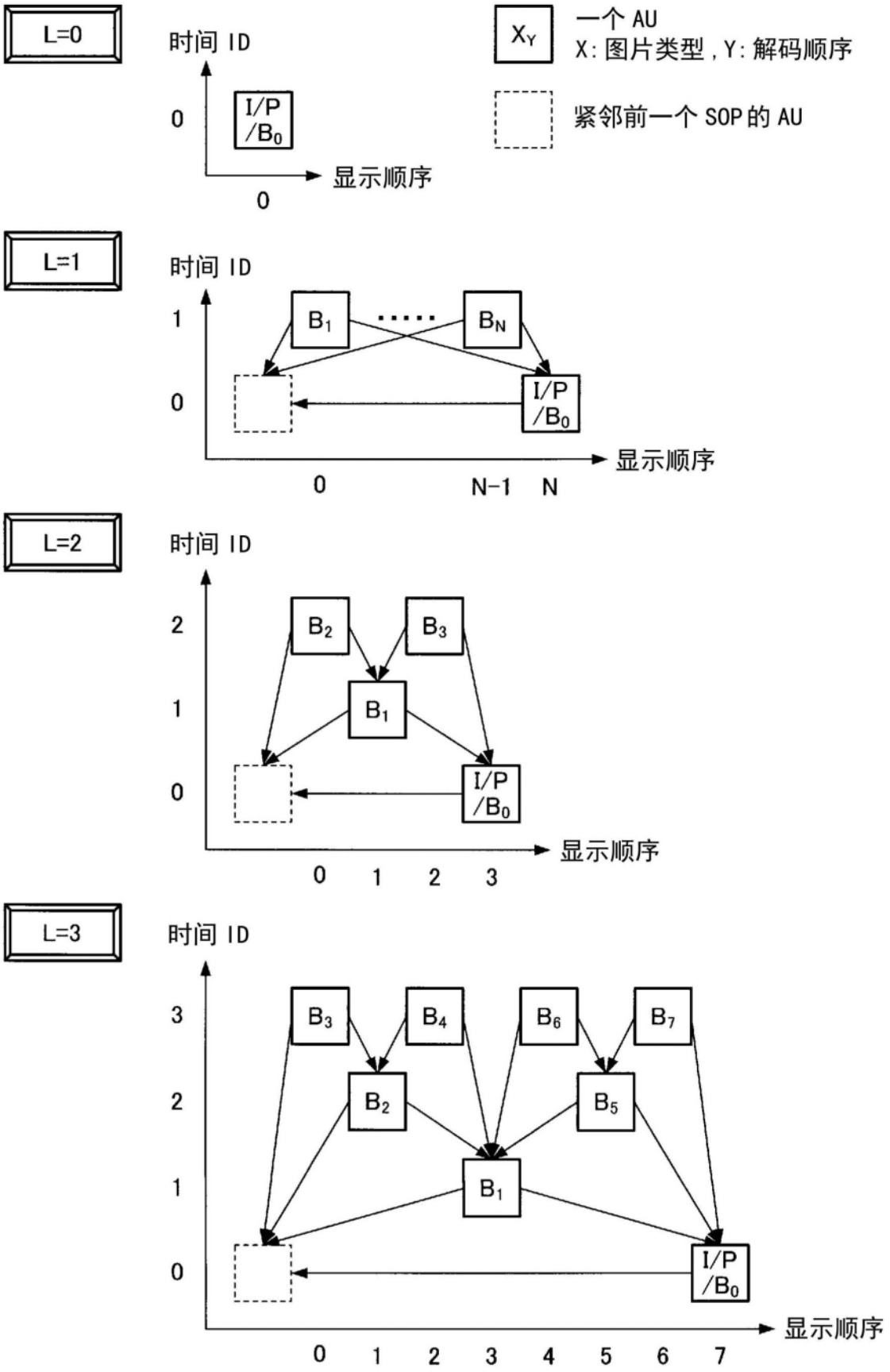

L=0
时间 ID
0
I/P
/B0
显示顺序
0
XY
一个 AU
X: 图片类型，Y: 解码顺序
紧邻前一个 SOP 的 AU
L=1
时间 ID
1
0
B1
BN
I/P
/B0
显示顺序
0
N-1
N
L=2
时间 ID
2
1
0
B2
B3
B1
I/P
/B0
显示顺序
0
1
2
3
L=3
时间 ID
3
2
1
0
B3
B4
B6
B7
B2
B5
B1
I/P
/B0
显示顺序
0
1
2
3
4
5
6
7

图12

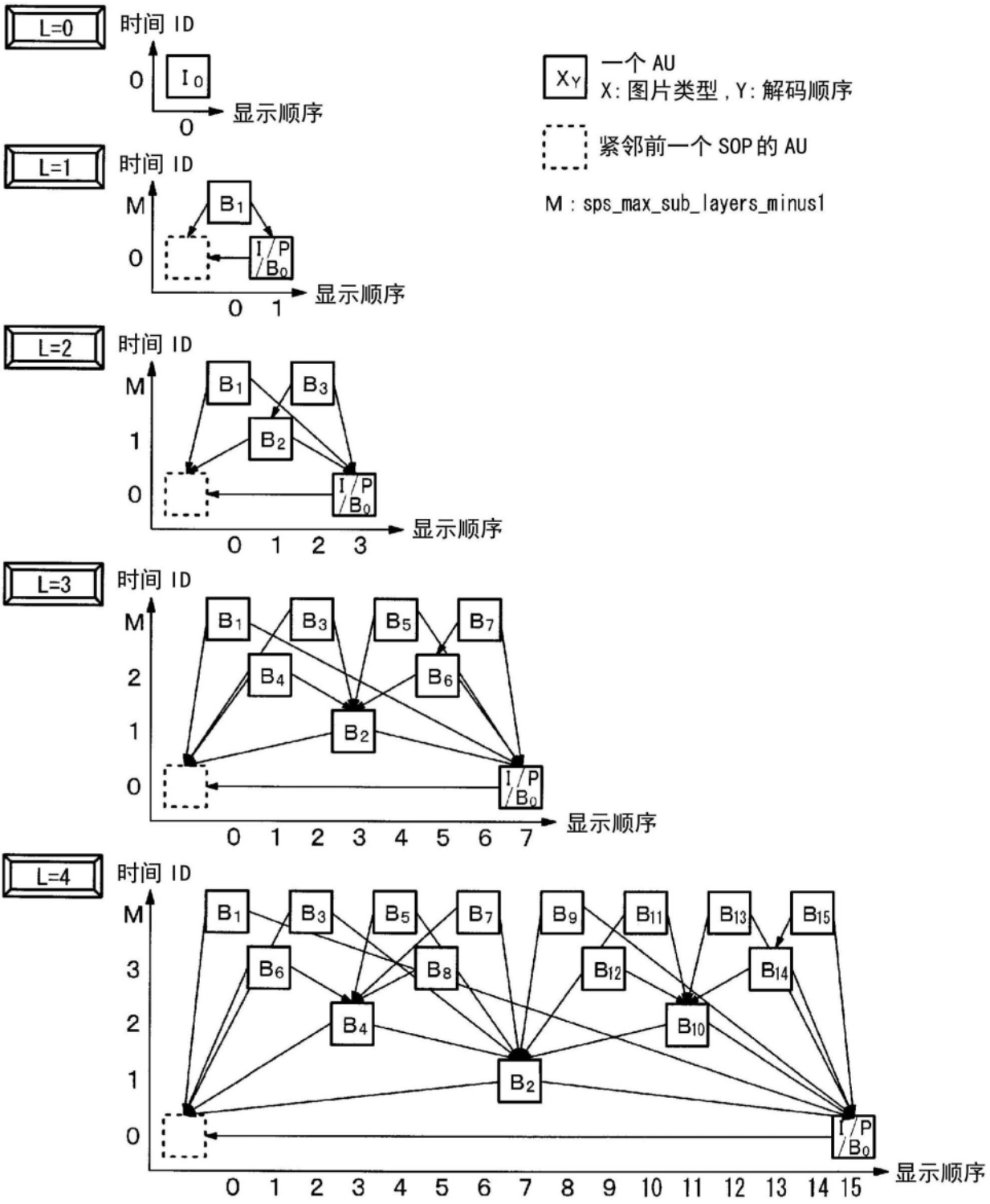

L=0
时间 ID
0
I0
0
显示顺序
L=1
时间 ID
M
0
B1
I/P/B0
0 1
显示顺序
L=2
时间 ID
M
1
0
B1
B3
B2
I/P/B0
0 1 2 3
显示顺序
L=3
时间 ID
M
2
1
0
B1
B3
B5
B7
B4
B6
B2
I/P/B0
0 1 2 3 4 5 6 7
显示顺序
L=4
时间 ID
M
3
2
1
0
B1
B3
B5
B7
B9
B11
B13
B15
B6
B8
B12
B14
B4
B10
B2
I/P/B0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
显示顺序
XY
一个 AU
X: 图片类型 , Y: 解码顺序
紧邻前一个 SOP 的 AU
M : sps_max_sub_layers_minus1

图13