



## (12)发明专利申请

(10)申请公布号 CN 107950025 A

(43)申请公布日 2018.04.20

(21)申请号 201680050362.9

(22)申请日 2016.06.27

(30)优先权数据

15306049.6 2015.06.30 EP

(85)PCT国际申请进入国家阶段日

2018.02.28

(86)PCT国际申请的申请数据

PCT/EP2016/064868 2016.06.27

(87)PCT国际申请的公布数据

W02017/001344 EN 2017.01.05

(71)申请人 汤姆逊许可公司

地址 法国伊西莱穆利诺

(72)发明人 多米尼克·托罗

迈克尔·勒·庞迪 罗南·伯特德

马丁·阿兰

(74)专利代理机构 中科专利商标代理有限责任  
公司 11021

代理人 倪斌

(51)Int.Cl.

H04N 19/11(2006.01)

H04N 19/157(2006.01)

H04N 19/176(2006.01)

H04N 19/593(2006.01)

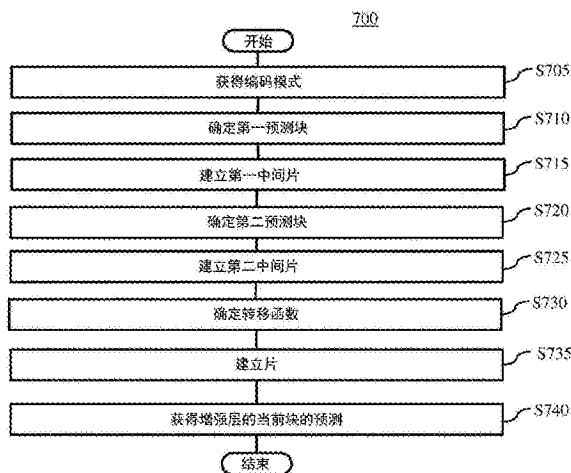
权利要求书3页 说明书19页 附图10页

### (54)发明名称

用于确定增强层的当前块的预测的方法和装置

### (57)摘要

一种方法包括:建立(S715)低动态范围的第一中间片;建立(S725)高动态范围的第二中间片;通过在变换域中对基础层的变换后的初始片应用转移函数,并且随后对所得到的片应用逆变换以返回至像素域,来建立(S735)片;通过从片中提取块,来预测(S740)增强层的当前块的预测;以及对增强层的当前块与增强层的当前块的预测之间的残差进行编码。



1. 一种方法,包括:

利用基础层的同位置块的相邻像素和第一预测块来建立 (S715) 低动态范围的第一中间片,所述第一预测块是利用所述基础层的编码模式根据所述基础层的所述同位置块的所述相邻像素而预测的;

利用增强层的当前块的相邻像素和第二预测块来建立 (S725) 高动态范围的第二中间片,所述第二预测块是利用所述编码模式根据所述增强层的当前块的所述相邻像素而预测的;

通过在变换域中对所述基础层的变换后的初始片应用转移函数,并且随后对所得到的片应用逆变换以返回至像素域,来建立 (S735) 片,其中,所述转移函数被确定 (S730) 为在所述变换域中将所述第一中间片变换到所述第二中间片;

通过从所述片中提取块来预测 (S740) 所述增强层的当前块的预测,在所述片中提取的块与所述第二中间片中的所述增强层的当前块是同位置的;以及

对所述增强层的当前块与所述增强层的当前块的预测之间的残差进行编码。

2. 根据权利要求1所述的方法,其中,所述基础层是使用专用于低动态范围视频的色调映射算子来进行色调映射的。

3. 根据权利要求1所述的方法,其中,当所述基础层的所述同位置块的第一编码模式可用于所述增强层的当前块时,将所述第一编码模式用作所述编码模式。

4. 根据权利要求1所述的方法,其中,当所述基础层的所述同位置块的第一编码模式不可用于所述增强层的当前块时,通过从可能的编码模式中选择最适当的编码模式来获得所述编码模式。

5. 根据权利要求4所述的方法,其中,选择最适当的编码模式是通过以下操作来执行的:利用所述增强层的可能的编码模式中的每一个,选择使得所述基础层的所述同位置块与所述基础层的所述同位置块的虚拟预测之间的差最小化的编码模式。

6. 根据权利要求1所述的方法,其中,如果所述增强层的当前块的大小与所述基础层的上采样同位置块的大小相同,则将所述基础层的所述同位置块的第一编码模式用作所述编码模式。

7. 根据权利要求1所述的方法,其中,通过考虑所述基础层和所述增强层中的重构误差与所述基础层和所述增强层的编码成本之间的折中,来选择所述基础层的所述同位置块的第一编码模式。

8. 一种装置 (400),包括:

第一中间片创建单元 (428),被配置为:利用基础层的编码模式根据所述基础层的同位置块的相邻像素来预测第一预测块;以及,利用所述基础层的所述同位置块的所述相邻像素和所述第一预测块来建立低动态范围的第一中间片;

第二中间片创建单元 (412),被配置为:利用所述编码模式根据增强层的当前块的相邻像素来预测第二预测块;以及,利用所述增强层的当前块的所述相邻像素和所述第二预测块来建立高动态范围的第二中间片;

单元 (411),用于:确定转移函数,以在变换域中将第一中间片变换为第二中间片;通过在所述变换域中对所述基础层的变换后的初始片应用所述转移函数,并且随后对所得到的片应用逆变换来返回至像素域,来建立片;以及,通过从所述片中提取块来预测所述增强层

的当前块的预测,在所述片中提取的块与所述第二中间片中的所述增强层的当前块是同位置的;以及

编码器(404),用于对所述增强层的当前块与所述增强层的当前块的预测之间的残差进行编码。

9.根据权利要求8所述的装置,其中,使用专用于低动态范围视频的色调映射算子来对所述基础层进行色调映射。

10.根据权利要求8所述的装置(400),其中,当所述基础层的所述同位置块的第一编码模式可用于所述增强层的当前块时,将所述第一编码模式用作所述编码模式。

11.根据权利要求8所述的装置(500),其中,当所述基础层的所述同位置块的第一编码模式不可用于所述增强层的当前块时,从可能的编码模式中选择最适当的编码模式。

12.根据权利要求11所述的装置(500),其中,通过以下操作来选择最适当的编码模式:利用所述增强层的可能的编码模式中的每一个,选择使得所述基础层的所述同位置块与所述基础层的所述同位置块的虚拟预测之间的差最小化的编码模式。

13.根据权利要求8所述的装置(400),其中,如果所述增强层的当前块的大小与所述基础层的上采样同位置块的大小相同,则将所述基础层的所述同位置块的第一编码模式用作所述编码模式。

14.根据权利要求8所述的装置(600),其中,通过考虑所述基础层和所述增强层中的重构误差与所述基础层和所述增强层的编码成本之间的折中,来选择所述基础层的所述同位置块的第一编码模式。

15.一种方法,包括:

对残留预测误差进行解码;

利用基础层的同位置块的相邻像素和第一预测块来建立(S715)低动态范围的第一中间片,所述第一预测块是利用所述基础层的编码模式根据所述基础层的所述同位置块的所述相邻像素而预测的;

利用增强层的当前块的相邻像素和第二预测块来建立(S725)高动态范围的第二中间片,所述第二预测块是利用所述编码模式根据所述增强层的当前块的所述相邻像素而预测的;

通过在变换域中对所述基础层的变换后的初始片应用转移函数,并且随后对所得到的片应用逆变换以返回至像素域,来建立(S735)片,其中,所述转移函数用于在所述变换域中将所述第一中间片变换到所述第二中间片;

通过从所述片中提取块来预测(S740)所述增强层的当前块的预测,在所述片中提取的块与所述第二中间片中的所述增强层的当前块是同位置的;以及

通过将所述预测误差添加到所述增强层的当前块的预测来重构所述增强层的块。

16.一种装置(450),包括:

解码器(451),用于对残留预测误差进行解码;

第一中间片创建单元(475),被配置为利用基础层的同位置块的相邻像素和第一预测块来建立低动态范围的第一中间片,所述第一预测块是利用所述基础层的编码模式根据所述基础层的所述同位置块的所述相邻像素而预测的;

第二中间片创建单元(455),被配置为利用增强层的当前块的相邻像素和第二预测块

来建立 (S725) 高动态范围的第二中间片, 所述第二预测块是利用所述编码模式根据所述增强层的当前块的所述相邻像素而预测的; 以及

单元 (457), 用于通过在变换域中对所述基础层的变换后的初始片应用转移函数, 并且随后对所得到的片应用逆变换来返回至像素域, 来建立片, 其中, 所述转移函数用于在所述变换域中将所述第一中间片变换为所述第二中间片; 以及通过从所述片中提取块来预测所述增强层的当前块的预测, 在所述片中提取的块与所述第二中间片中的所述增强层的当前块是同位置的; 以及

单元 (453), 用于将所述预测误差添加到所述增强层的当前块的预测, 以重构所述增强层的块。

## 用于确定增强层的当前块的预测的方法和装置

### 技术领域

[0001] 本公开涉及一种用于确定增强层的当前块的预测的方法和装置。

### 背景技术

[0002] 在图像处理领域中,色调映射算子(在下文中可以称为“TMO”)是已知的。在对自然环境中的实际对象进行成像时,实际对象的动态范围远高于成像设备(比如,相机)可以成像的或显示器可以显示的动态范围。为了以自然的方式在这样的显示器上显示实际对象,TMO用于将高动态范围(在下文中可以称为“HDR”)图像转换为低动态范围(在下文中可以称为“LDR”)图像,同时保持良好的可见状况。

[0003] 一般而言,TMO被直接应用于HDR信号以获得LDR图像,并且该图像可以显示在经典的LDR显示器上。存在各种各样的TMO,并且许多TMO是非线性算子。

[0004] 关于与LDR/HDR视频压缩相关的技术,作为一种可能方案提出了使用全局TMO/iTMO(逆色调映射操作),如Z.Mai,H.Mansour,R.Mantiuk,P.Nasiopoulos,R.Ward和W.Heidrich的文章“On-the-fly tone mapping for backward-compatible high dynamic range image/video compression,”ISCAS,2010中所说明的。

[0005] 在这篇文章中,为了最小化总量化误差,考虑了浮点数据的分布。该算法通过以下步骤来描述(这里使用的变量在图1中示出)。

[0006] 步骤1:计算亮度值的对数。因此,对于亮度为L的每个像素,以下步骤基于值 $l = \log_{10}(L)$ 。(l仍然是浮点数格式)。

[0007] 步骤2:通过将元阵(bin)大小固定为 $\delta = 0.1$ 来计算l值的直方图。例如,图像序列中的所有像素均可用于建立直方图。因此,对于每个元阵k( $k = 1..N$ ),像素属于该元阵的概率 $p_k$ 是已知的。将值 $l_k = \delta \cdot k$ 指派给元阵。

[0008] 步骤3:根据下述公式(1)描述的模型来为每个元阵K计算斜率值:

$$[0009] \quad s_k = \frac{v_{max} \cdot p_k^{1/3}}{\delta \cdot \sum_{k=1}^N p_k^{1/3}} \quad (1)$$

[0010] 其中, $v_{max}$ 是所考虑的整数表示的最大值(如果数据被量化为n比特整数,则 $v_{max} = 2^n - 1$ )。

[0011] 为了避免在逆方程(5中的逆色调映射)中出现除零的风险,如果 $s_k = 0$ ,则 $s_k$ 可被替代地设置为非零最小值 $\epsilon$ 。

[0012] 步骤4:已知N个斜率值,可以定义全局色调映射曲线。对于 $[1, N]$ 中的每个k,将满足 $l_k < l \leq l_{k+1}$ 的浮点数 $l$ 映射到由下述公式(2)定义的整数值v:

$$[0013] \quad v = (l - l_k) \cdot s_k + v_k \quad (2)$$

[0014] 其中,值 $v_k$ 是通过 $v_{k+1} = \delta \cdot s_k + v_k$ (并且 $v_1 = 0$ )根据 $s_k$ 来定义的。

[0015] 然后,对值v进行四舍五入,以获得在区间 $[0, 2^n - 1]$ 中的整数。

[0016] 步骤5:为了执行逆色调映射,必须向解码器发送参数 $s_k$ ( $k = 1..N$ )。对于色调映射

图像中的具有值 $v$ 的给定像素,首先,必须找到满足 $v_k \leq v < v_{k+1}$ 的值 $k$ 。

[0017] 于是,逆方程被表示为下述公式(3):

$$[0018] \quad l_{dec} = l_k + \frac{(v-v_k)}{s_k} \quad (3)$$

[0019] 这里,解码后的像素值被实现为 $L_{dec} = 10^{l_{dec}}$

[0020] 此外,为了应用逆色调映射(iTMO),解码器必须知道图1中的曲线。

[0021] 这里的术语“解码”对应于解量化操作,其不同于视频编码器/解码器中的术语“解码”不同。

[0022] 另一种可能方案是使用局部色调映射算子,如M.Grundland等人的文章“Non linear multiresolution blending”,Machine Graphis&vision International Journal Volume 15 Issue 3 Feb 2006和Zhe Wendy Wang;Jiefu Zhai;Tao Zhang;Llach,Joan的文章“Interactive tone mapping for High Dynamic Range video”.ICASSP 2010中所公开的。例如,可以基于如下参考文献的公开内容来使用TMO拉普拉斯金字塔:Peter J.Burt Edward H.Adelson的文章“The Laplacian Pyramid as a compact image code,”IEEE Transactions on Communications,vol.COM-31,no.4,April 1983;Burt P.J.的文章“The Pyramid as Structure for Efficient Computation.Multiresolution Image Processing and Analysis”Springer-Verlag,6-35;以及,Zhai jiedu,Joan Llach的专利申请“Zone-based tone mapping”WO 2011/002505 A1。TMO的效率在于从HDR图像中提取不同的中间LDR图像,其中各中间LDR图像与不同的曝光相对应。因此,当(原始HDR图像的)亮区域饱和时,过度曝光的LDR图像在暗区域中包含精细细节。相反,当将暗区域裁剪掉时,曝光不足的LDR图像在亮区域中包含精细细节。

[0023] 之后,将每个LDR图像分解成 $n$ 个层级的拉普拉斯金字塔,同时最高层级专用于最低分辨率,而其它层级提供(梯度的)不同谱带。因此,在这个阶段,每个LDR图像对应于一个拉普拉斯金字塔,并且此外,还可以注意到,只有在没有四舍五入误算的情况下,每个LDR图像才能通过使用逆分解或逆“瓦解”根据其拉普拉斯金字塔来重建。

[0024] 最后,在融合中间LDR图像集的不同金字塔层级的情况下实现色调映射,并且瓦解所得到的混合金字塔以给出最终的LDR图像。

[0025] 事实上,对不同谱带(或金字塔层级)的梯度(gradient)进行融合是非线性处理。这种算法类型的优点在于色调映射的有效结果,但是有时会导致很多众所周知的渲染错误(例如,光晕伪影)。以上参考资料提供了关于该技术的更多细节。

[0026] 事实上,因为这种色调映射是非线性的,因此难以实现LDR的逆色调映射使得在SNR(信噪比)或空间视频可伸缩性方面给出针对HDR层的当前块的可接受预测。

[0027] 此外,W02010/018137公开了一种用于修改参考图像的参考块的方法、一种用于在参考块帮助下对图像的块进行编码或解码的方法及对应设备以及携带在修改的参考B的帮助下编码而成的块的存储介质或信号。在现有技术中,根据相邻平均值来估计转移函数,并且该函数用于校正图像间预测。然而,在W02010/018137中,该方法限于平均值,从而给出当前块和同位置(collocated)块的第一近似值。

## 发明内容

[0028] 根据本公开的实施例,提供了这样一种方法,包括:利用基础层的同位置块的相邻像素和第一预测块来建立低动态范围的第一中间片,所述第一预测块是利用所述基础层的编码模式根据所述基础层的所述同位置块的所述相邻像素而预测的;利用增强层的当前块的相邻像素和第二预测块来建立高动态范围的第二中间片,所述第二预测块是利用所述编码模式根据所述增强层的当前块的所述相邻像素而预测的;通过在变换域中对所述基础层的变换后的初始片应用转移函数并且随后对所得到的片应用逆变换以返回至像素域,来建立片,其中,确定所述转移函数以在所述变换域中将所述第一中间片变换成所述第二中间片;通过从所述片中提取块来预测所述增强层的当前块的预测,在所述片中提取的块与所述第二中间片中的所述增强层的当前块是同位置的;以及对所述增强层的当前块与所述增强层的当前块的预测之间的残差进行编码。

[0029] 根据本公开的实施例,提供了一种装置,包括:第一中间片创建单元,被配置为:利用基础层的编码模式根据所述基础层的同位置块的相邻像素来预测第一预测块;以及,利用所述基础层的所述同位置块的所述相邻像素和所述第一预测块来建立低动态范围的第一中间片;第二中间片创建单元,被配置为:利用所述编码模式根据增强层的当前块的相邻像素来预测第二预测块;以及,利用所述增强层的当前块的所述相邻像素和所述第二预测块来建立高动态范围的第二中间片;用于执行以下操作的单元:确定转移函数,以在变换域中将第一中间片变换为第二中间片;通过在所述变换域中对所述基础层的变换后的初始片应用所述转移函数并且随后对所得到的片应用逆变换来返回至像素域,来建立片;以及,通过从所述片中提取块来预测所述增强层的当前块的预测,在所述片中提取的块与所述第二中间片中的所述增强层的当前块是同位置的;以及编码器,用于对所述增强层的当前块与所述增强层的当前块的预测之间的残差进行编码。

[0030] 根据本公开的另一实施例,提供了一种方法,包括:对残留预测误差进行解码;利用基础层的同位置块的相邻像素和第一预测块来建立低动态范围的第一中间片,所述第一预测块是利用所述基础层的编码模式根据所述基础层的所述同位置块的所述相邻像素而预测的;利用增强层的当前块的相邻像素和第二预测块来建立高动态范围的第二中间片,所述第二预测块是利用所述编码模式根据所述增强层的当前块的所述相邻像素而预测的;通过在变换域中对所述基础层的变换后的初始片应用转移函数并且随后对所得到的片应用逆变换以返回至像素域,来建立片,其中,所述转移函数用于在所述变换域中将所述第一中间片变换成所述第二中间片;通过从所述片中提取块来预测所述增强层的当前块的预测,在所述片中提取的块与所述第二中间片中的所述增强层的当前块是同位置的;以及通过将所述预测误差添加到所述增强层的当前块的预测来重构所述增强层的块。

[0031] 根据本公开的又一实施例,提供了一种装置,包括:解码器,用于对残留预测误差进行解码;第一中间片创建单元,被配置为利用基础层的同位置块的相邻像素和第一预测块来建立低动态范围的第一中间片,所述第一预测块是利用所述基础层的编码模式根据所述基础层的所述同位置块的所述相邻像素而预测的;第二中间片创建单元,被配置为利用增强层的当前块的相邻像素和第二预测块来建立(S725)高动态范围的第二中间片,所述第二预测块是利用所述编码模式根据所述增强层的当前块的所述相邻像素而预测的;以及用于执行以下操作的单元:通过在变换域中对所述基础层的变换后的初始片应用转移函数并且随后对所得到的片应用逆变换来返回至像素域,来建立片,其中,所述转移函数用于在所

述变换域中将所述第一中间片变换为所述第二中间片,以及通过从所述片中提取块来预测所述增强层的当前块的预测,在所述片中提取的块与所述第二中间片中的所述增强层的当前块是同位置的;以及用于执行以下操作的单元:将所述预测误差添加到所述增强层的当前块的预测,以重构所述增强层的块。

[0032] 当结合附图来阅读以下详细描述,本公开的其它目的、特征和优点将变得更加清楚。

## 附图说明

[0033] 图1是浮点值 $l = \log_{10}(L)$ 的直方图及其关联的基于斜率 $sk$ 的色调映射曲线;

[0034] 图2A和图2B是重构的基础层的图像和要被编码的增强层的当前块的图像;

[0035] 图3A至图3J是示出了在H.264标准中规定的帧内 $4 \times 4$ 预测的示例的图;

[0036] 图4A和图4B是示出了第一实施例的用于确定增强层的当前块的预测的装置的框图,其中图4A是编码器侧,而图4B是解码器侧;

[0037] 图5A和图5B是示出了本公开实施例的第二实施例的用于确定增强层的当前块的预测的装置的配置的框图,其中图5A是编码器侧,而图5B是解码器侧;

[0038] 图6是示出了本公开第四实施例的用于确定增强层的当前块的预测的装置的配置的框图;以及

[0039] 图7是示出了根据本公开实施例的用于确定增强层的当前块的预测的示例性方法的流程图。

## 具体实施方式

[0040] 下文参考附图给出了对本公开的实施例的描述。

[0041] 本公开的实施例旨在改进逆色调映射操作(在下文中可以称为“iTMO”)的处理以及以全局或局部(非线性)方式使用的先前的TMO,显然是在基础层信号仍然可用的话的情况下。

[0042] 该思想例如涉及编码有专用于LDR视频编码的且使用给定TMO的第一色调映射基础层 $l_b$ 和专用于HDR视频编码的第二增强层 $l_e$ 的HDRSNR可伸缩视频。在这种情况下(SNR可伸缩)下,对于增强层的(要被编码的)当前块 $b_e$ ,应该找到从基础层 $b_b$ (同位置块)提取的预测块,并且该块必须通过逆色调映射来处理。

[0043] 为了实现块 $b_b$ 的逆色调映射,应当估计变换 $T_{be}$ 函数,以允许将片(patch) $p'_b$ (由虚拟块 $b'_b$ ( $b_b$ 的同源)及其相邻物组成)的像素变换到当前片 $p'_e$ (由虚拟块 $b'_e$ ( $b_e$ 的同源)及其相邻物组成)。

[0044] 一旦确定了 $T_{be}$ ,变换 $T_{be}$ 函数可被应用于片 $p_b$ (由块 $b_b$ 及其相邻物组成)从而给出片 $p_b^T$ ,最终,最后步骤属于在片 $p_b^T$ 中提取与当前块位置相同的块 $\tilde{b}_e$ 。这里,块 $\tilde{b}_e$ 与块 $b_e$ 的预测相对应。

[0045] 这里,应该注意的是,在估计变换 $T_{be}$ 之前,需要基础层的同位置块 $b_b$ 的编码模式,或者需要基于基础层、根据( $l_b$ 的)重构图像在(增强层的编码器的)可用编码模式集合中提取预测模式。



[0046] 还应重点注意,以上说明的全部处理步骤也在解码器侧以及编码器侧实现。

[0047] [原理]

[0048] 为了说明在本公开实施例中提出的方法,下文给出基于SNR可伸缩性的示例。在这种情况下(SNR可伸缩性)下,应当针对增强层的(要被编码的)当前块 $b_e$ 找到从基础层 $b_b$ (同位置块)提取的预测块,并且该预测块必须通过逆色调映射进行处理。

[0049] 图2A和图2B分开示出了重构的基础层的图像和要被编码的当前块的图像。

[0050] 图2B中所示的与增强层 $l_e$ 的当前图像相关的标记如下:

[0051] • 要预测的增强层的当前块(未知)是: $X_u^B$

[0052] • 当前块的已知重构(或解码)的相邻物(或模板): $X_k^T$

[0053] • 当前片是: $X = \begin{bmatrix} X_k^T \\ X_u^B \end{bmatrix}$  (4)

[0054] • 索引 $k$ 和 $u$ 分别指示“已知”和“未知”。

[0055] 图2A中所示的与基础层 $l_b$ 的图像相关的标记如下:

[0056] • 基础层的同位置块(已知)(其与要预测的增强层的当前块实质是同位置的)是: $Y_k^B$

[0057] • 当前块的已知重构(或解码)的相邻物(或模板)是: $Y_k^T$

[0058] • 同位置片(X的同位置的)是:

[0059]  $Y = \begin{bmatrix} Y_k^T \\ Y_k^B \end{bmatrix}$  (5)

[0060] 目标是根据块 $Y_k^B$ 来确定针对当前块 $X_u^B$ 的预测块。实际上,将估计在片Y和X之间的变换,这种变换与一种逆色调映射相对应。

[0061] 显然,在视频压缩的上下文下,块 $X_u^B$ 是不可得的(记住:解码器将实现相同的处理),但是存在可以提供当前块 $X_u^B$ 的第一近似值(更精确的预测)的很多可能的预测模式。

这里,当前块 $X_u^B$ 及其相邻物 $X_k^T$ 的第一近似值构成片X的中间片 $X'$ 。

[0062] 之后,使用块 $X_u^B$ 的第一近似值,以便找到变换函数 $\text{Trf}(l_b \rightarrow l_e)$ ,该变换函数允许将X的中间片变换为Y的中间片(分别被标记为 $X'$ 和 $Y'$ ),并且该变换最终应用于初始片Y,从而允许提供确定性的预测块。

[0063] [第一实施例]

[0064] 参考图3A至图3J和图4来给出用于确定增强层的当前块的预测的方法和装置的第一实施例的描述。

[0065] 更具体地,本公开的第一实施例是关于SNR可伸缩性,也就是说,LDR基础层和HDR

增强层之间的相同空间分辨率。此外,在第一实施例中,已经用增强层的编码器的帧内编码模式之一(例如,MPEG-4AVC/H.264中定义的并且在文件ISO/IEC 14496-10中描述的H.264标准的帧内模式)对当前块 $X_u^B$ 的同位置块 $Y_k^B$ 进行了编码。

[0066] 利用块 $Y_k^B$ 的索引m的编码模式并且利用相邻像素 $Y_k^T$ ,可以重构预测块 $Y_{prd,m}^B$ 。

[0067] 图3A至图3J是示出了在H.264标准中规定的帧内 $4 \times 4$ 预测的图。如图3A至图3J所示,H.264标准中提供了N个(这里,在H264的情况下,N=9)不同的帧内模式预测。

[0068] 在H.264中,帧内 $4 \times 4$ 和帧内 $8 \times 8$ 预测与要基于相邻重构像素而被编码的当前块的像素的空间估计相对应。H.264标准规定了不同的方向预测模式,以便详细说明像素预测。在宏块(MB)的 $4 \times 4$ 和 $8 \times 8$ 块大小上定义了九(9)个帧内预测模式。如图3所示,这些模式中的八(8)个模式由对要预测的当前块周围(从左列和顶行开始)的像素的一维方向外推组成。帧内预测模式2(DC模式)将预测块像素定义为可用周围像素的平均值。

[0069] 在帧内 $4 \times 4$ 的示例中,如图3A至图3J所示那样来建立预测。

[0070] 例如,如图3C所示,在模式1(水平)中,利用(左列的)重构像素J来预测像素e、f、g和h。

[0071] 此外,如图3G所示,在模式5中,作为第一示例,通过 $(Q+A+1)/2$ 来预测“a”。类似地,作为第二示例,通过 $(A+2B+C+2)/4$ 来预测“g”和“p”。

[0072] 这里,返回到上面讨论的问题,优选地建立当前块 $X_u^B$ 的预测,以实现利用与在基础层中使用的相同的索引为m的预测模式和当前相邻物 $X_k^T$ 提供预测块 $X_{prd,m}^B$ 的目的。

[0073] 这里,两个中间片 $X'$ 和 $Y'$ 可以如下面的公式(6)和(7)那样构成。

[0074] 当前中间片 $X'$  :

$$[0075] \quad X' = \begin{bmatrix} X_k^T \\ X_{prd,m}^B \end{bmatrix} \quad (6)$$

[0076] 基础层的中间片 $Y'$  :

$$[0077] \quad Y' = \begin{bmatrix} Y_k^T \\ Y_{prd,m}^B \end{bmatrix} \quad (7)$$

[0078] 在变换域(TF)中计算 $Y'$ 和 $X'$ 之间的期望变换Trf,并且该变换可以是阿达玛(Hadamard)、离散余弦变换(DCT)、离散正弦变换(DST)、或傅里叶变换等。提供以下公式(8)和(9)。

$$[0079] \quad T_{X'} = \text{TF}(X') \quad (8)$$

$$[0080] \quad T_{Y'} = \text{TF}(Y') \quad (9)$$

[0081] 公式 $\text{TF}(Y')$ 与片 $Y'$ 的2D变换“TF”(例如,DCT)相对应。

[0082] 下一步是计算转移函数Trf,该转移函数允许将 $T_{Y'}$ 变换为 $T_{X'}$ ,其中,将以下公式(10)和(11)应用于每对系数。

[0083] If  
 [0084]  $(\text{abs}(T_{X'}(u,v)) > \text{th} \text{ and } \text{abs}(T_{Y'}(u,v)) > \text{th})$   
 [0085] then  
 [0086]  $\text{Trf}(u,v) = T_{X'}(u,v) / T_Y(u,v) \quad (10)$   
 [0087] else  
 [0088]  $\text{Trf}(u,v) = 0 \quad (11)$   
 [0089] end if

[0090] 这里,  $u$ 和 $v$ 是 $T_{X'}$ 、 $T_{Y'}$ 和 $\text{Trf}$ 的系数的转移变换坐标, 并且 $\text{th}$ 是给定值的阈值, 其避免了 $\text{Trf}$ 转移函数中的奇点。例如, 在H.264或HEVC标准压缩的上下文下,  $\text{th}$ 可以等于1。在B.Bross, W.J.Han, G.J.Sullivan, J.R.Ohm, T.Wiegand JCTVC-K1003, "High Efficiency Video Coding (HEVC) text specification draft 9," Oct 2012中描述了HEVC(高效率视频编码)。

[0091] 将函数 $\text{Trf}$ 应用于基础层 $Y$ 的初始片的变换(TF), 所得结果在逆变换( $\text{TF}^{-1}$ )之后给出片 $Y''$ 。如公式(12)至(14)所示, 片 $Y''$ 由模板 $Y''^T$ 和块 $Y_m''^B$ 组成。

$$[0092] \quad Y'' = \begin{bmatrix} Y''^T \\ Y_m''^B \end{bmatrix} \quad (12)$$

[0093] 其中 $Y'' = \text{TF}^{-1}(T_Y)$  (13)

[0094] 并且 $T_Y = \text{TF}(Y) \cdot \text{Trf}$  (14)

[0095] 公式 $\text{TF}(Y) \cdot \text{Trf}$ 对应于将转移函数 $\text{Trf}$ 应用于基础层的初始片 $Y$ 的变换片 $T_Y$ 的分量, 并且该应用是针对(坐标 $u$ 和 $v$ )的每个变换分量来执行的, 如公式(15)所示。

[0096]  $T_Y(u,v) = T_Y(u,v) \cdot \text{Trf}(u,v) \quad (15)$

[0097] 最后, 当前块 $X_u^B$ 的预测属于在从片 $Y''$ 提取块 $Y_m''^B$ , 并且标记 $m$ 指示预测块是在基础层的帧内模式索引 $m$ 的帮助下建立的。

[0098] 图4A和图4B是示出了第一实施例的用于确定增强层的当前块的预测的装置的框图。图4A和图4B中也示出了帧内SNR可伸缩性的这种描述的原理。

[0099] 参考图4A和图4B, 描述局部层间LDR HDR预测。

[0100] 为了使描述清楚, 特别是对于解码器, 我们描述了SNR可伸缩视频编码(SVC)方案:

[0101] (1) 首先是基础层

[0102] (2) 其次是增强层

[0103] 在图4A所示的编码器(或编码装置)侧, 以及在图4B所示的解码器侧, 已知: 所提方案关注于层间( $b1 \rightarrow e1$ )预测。

[0104] 在编码器侧和解码器侧, 由于我们的层间预测模式使用帧内模式( $m$ ), 因此仅描述了使用帧内模式( $m$ )的帧内图像预测模式。因此众所周知, 预测单元(使用给定的RDO(速率失真优化)准则)的功能在于根据以下方面来确定最佳预测模式:

[0105] (1) 在基础层层级处的帧内和帧间图像预测

[0106] (2) 在增强层层级处的帧内、帧间图像和层间预测(我们的新预测模式)

[0107] 索引的含义

[0108] k: 已知

[0109] u: 未知

[0110] B: 块

[0111] T: 块的相邻物(在视频压缩域中通常称为“模板”)

[0112] Pred: 预测

[0113] m: 来自N个可用模式的帧内编码模式的索引

[0114] 参考图2A和图2B, Y、X、Y'、X' 和Y”是由块和模板组成的片。

[0115] 图4A中的编码器侧(单元400):

[0116] 使用TMO 406来对原始块401b<sub>e</sub>进行色调映射,其给出原始色调映射块b<sub>bc</sub>。

[0117] 基础层(b1)

[0118] 我们考虑编码原始基础层块b<sub>bc</sub>。

[0119] a) 利用参考帧缓冲器426中存储的原始块b<sub>bc</sub>和(先前解码的)图像,运动估计器(运动估计单元)429利用给定的运动矢量(时间预测单元)找到最佳帧间图像预测块,并且时间预测(Temp Pred Pred)单元430给出时间预测块。根据可用帧内预测模式(在图3中示出,在H264的情况下)和相邻重构(或解码)的像素,空间预测(Sp Pred)单元428给出帧内预测块。

[0120] b) 如果模式决定处理(单元425)选择帧内图像预测模式(从N个帧内可用模式中选择索引为m的帧内图像预测模式),则利用原始块b<sub>bc</sub>和预测块 $\hat{b}_k^B$  ( $\hat{Y}_{pred,m}^B$ )之间的差(通过组合器421)来计算残差预测rb。

[0121] c) 之后,通过T Q单元422将残差预测rb变换和量化为r<sub>bq</sub>,并且最终由熵编码器单元423进行熵编码并在比特流基础层中发送。

[0122] d) 通过下述方式局部重建解码的块:(利用组合器427)将由 $T^{-1}Q^{-1}$ 单元424逆变换和解量化后的预测误差块r<sub>bdq</sub>添加到预测块 $\hat{b}_k^B$ ,给出重构(基础层)块。

[0123] e) 将重构(或解码)的帧存储在(b1)参考帧缓冲器426中。

[0124] 增强层(e1)

[0125] 我们可以注意到,增强层的编码器的结构与基础层的编码器相似,例如,就编码模式决定、时间预测和参考帧缓冲器方面而言,单元407、408、409和413具有与基础层的编码器的相应单元425、426、429和430相同的功能。

[0126] 现在考虑编码原始增强层块b<sub>e</sub>。

[0127] f) 对于增强层的块,如果基础层的同位置块是以帧内图像模式被编码的,则我们考虑该同位置块的(索引为m的)帧内模式(图7所示的方法700的S705)。

[0128] g) 利用基础层的(索引为m的)这种帧内模式,我们确定:

[0129] o 利用b1空间预测(Sp pred)单元428,确定或重用基础层层级的帧内预测块( $\hat{b}_k^B$ )  $\hat{Y}_{pred,m}^B$  (S710,图7),

[0130] o 然后,根据公式(7)利用同位置块( $\hat{Y}_k^B$ )的相邻物( $\hat{Y}_k^T$ )和预测块 $\hat{Y}_{pred,m}^B$ 来建立第一中间片Y' (S715,图7)

[0131] h) 类似地,利用基础层的(索引为m的)这种帧内模式,我们确定:

[0132] o (利用e1空间预测 (Sp pred) 单元412;S720,图7) 确定在增强层层级处的中间帧内预测块 $\hat{Y}_{pred,m}^B$ ;

[0133] o并且,然后根据公式 (6) 利用当前块 ( $b_e$ ) 的相邻物( $\hat{X}_k^B$ )和中间预测块 $\hat{Y}_{pred,m}^B$ 来建立第二中间片 $X'$  (S725,图7)

[0134] i) 在变换域 (例如,DCT) 中,我们使用公式 (8) 至 (11) 来确定从片 $Y'$  片到 $X'$  的转移函数Trf (S730,图7)。

[0135] j) 现在,我们考虑根据公式 (5) 由同位置块( $\hat{Y}_k^B$ )及其相邻物 $\hat{Y}_k^B$ 组成的基础层Y的初始 (解码) 片 (图7中的S735-S740)

[0136] 1.我们对片Y应用变换 (例如,DCT) :TF (Y)

[0137] 2.现在,将Trf函数应用于变换域,比如: $T_{Y''} = TF(Y) \cdot Trf$

[0138] 3.对 $T_{Y''}$ 计算逆变换 (例如,DCT-1),给出 $Y'' = TF^{-1}(T_{Y''})$ ,其中所得到的片是如公式 (12) 那样组成的。

[0139] 4.最后,从片 $Y''$  提取与块 $\hat{Y}_m^B$ 相对应的预测。

[0140] 从f到j的所有步骤是在图4A中的“预测e1/b1 (Trf)”单元411中实现的。

[0141] k) (T Q单元403) 将 (使用组合器402) 计算的增强层块 $b_e$ 和在步骤f到j处计算的层间预测( $\hat{Y}_m^B$ )之间的误差残量 $r_e$ 变换和量化为 $r_{eq}$ ,并且 $r_{eq}$ 由熵编码器单元404进行熵编码并且在增强层比特流中发送

[0142] l) 最后,通过 (用组合器410) 将由 $T^{-1}Q^{-1}$ 单元405逆变换和解量化的预测误差块 $r_{edq}$ 添加到预测 $\hat{Y}_m^B$ ,来局部地重建解码的块,并且将重构 (解码) 的图像存储在 (e1) 参考帧缓冲器408中。

[0143] 图4B中的解码器侧 (单元450) :

[0144] 基础层 (b1)

[0145] a) 根据b1比特流,对于给定块,熵解码器 (熵解码器单元) 471对量化误差预测 $r_{bq}$ 和相关联的索引为m的编码帧内模式进行解码

[0146] b) 通过 $T^{-1}Q^{-1}$ 单元472将残差预测 $r_{bq}$ 解量化和逆变换为 $r_{bdq}$ ,

[0147] c) 在帧内模式m的帮助下,“空间预测 (Sp Pred)”单元475和“预测”单元474利用解码的相邻像素来给出图像内预测块 $\hat{X}_k^B$ 或 $\hat{Y}_{pred,m}^B$ 。

[0148] d) 通过下述方式布局重建解码的块: (利用组合器473) 将解码和解量化后的预测错误块 $r_{bdq}$ 添加到预测块 $\hat{X}_k^B$  (或 $\hat{Y}_{pred,m}^B$ ),给出基础层的重构块。

[0149] e) 将重构 (或解码) 的帧存储在参考帧缓冲器476中, (使用运动补偿单元477) 将解码的帧用于下一个 (b1) 帧内图像预测和帧间预测。

[0150] 增强层 (e1)

[0151] f) 来自e1比特流,对于给定块,熵解码器451对量化的误差预测 $r_{eq}$ 进行解码。

[0152] g) 由 $T^{-1}Q^{-1}$ 单元452对残差预测 $r_{eq}$ 进行解量化和逆变换,并且输出 $r_{edq}$ 。

[0153] h) 如果用于待解码的块的编码模式与我们的层间模式相对应,则我们考虑基础层的同位置块的 (索引为m的) 帧内模式。

[0154] i) 利用基础层的 (索引为m的) 这种帧内模式, 我们确定:

[0155] o (利用b1空间预测 (Sp pred) 单元475) 确定或重用基础层层级的帧内预测块

$$(\tilde{b}_2) \quad y_{pred,m}^B,$$

[0156] o 然后, 根据公式 (7) 利用同位置块  $(Y_k^B)$  的相邻物  $(Y_k^I)$  和预测块  $y_{pred,m}^B$  建立第一中间片  $Y'$

[0157] h) 类似地, 利用基础层的 (索引为m的) 这种帧内模式, 我们确定:

[0158] o 利用e1空间预测 (Sp pred) 单元455, 确定增强层层级的中间帧内预测块  $x_{pred,m}^S$ ;

[0159] o 并且, 然后根据公式 (6) 利用当前块  $(b_e)$  的相邻物  $(X_k^I)$  和中间预测块  $x_{pred,m}^S$  来建立第二中间片  $X'$ 。

[0160] k) 在变换域 (例如, DCT) 中, 我们使用公式 (8) 至 (11) 来确定从片  $Y'$  到片  $X'$  的转移函数  $Trf$ 。

[0161] l) 现在, 我们考虑根据公式 (5) 由同位置块  $(Y_k^I)$  及其相邻物  $Y_k^I$  组成的基础层  $Y$  的初始 (解码) 片。

[0162] 1. 我们对片  $Y$  应用变换 (例如, DCT):  $TF(Y)$

[0163] 2. 现在, 将  $Trf$  函数应用于变换域, 比如:  $T_{Y''} = TF(Y) \cdot Trf$

[0164] 3. 对  $T_{Y''}$  计算逆变换 (例如,  $DCT^{-1}$ ), 给出  $Y'' = TF^{-1}(T_{Y''})$ , 其中所得到的片是如下组成的。

$$[0165] \quad Y'' = \begin{bmatrix} Y^{I'} \\ Y_m^{I'B} \end{bmatrix} \quad (12)$$

[0166] 4. 最后, 从片  $Y''$  提取与块  $Y_m^{I'B}$  相对应的预测。

[0167] 从h到l的所有步骤是在“预测e1/b1 (Trf)”单元457中实现的, 我们可以注意到, 步骤h到l与 (第一实施例的) 编码器的步骤f到j严格相同; 显然, 如果e1编码器通过e1编码器407的模式决定选择该层间预测模式的话。

[0168] m) 通过下述方式来建立e1解码块: (利用组合器453) 将解码和解量化的预测误差块  $redq$  添加到 (经由预测块454) 的预测块  $y_m^{I'B}$ , 给出重构的 (e1) 块。

[0169] n) 将重构 (或解码) 的图像存储在 (e1) 参考帧缓冲器456中, (使用运动补偿单元458) 将解码的帧用于下一个 (e1) 帧内图像预测和帧间预测。

[0170] 如上所述, 第一实施例的装置可以配置为如图4A和图4B所示的那样, 可以通过该装置来执行第一实施例的方法。

[0171] 根据该用于确定增强层的当前块的预测的方法和装置, 通过利用基础层的同位置块的编码模式, 可以容易且精确地获得增强层的当前块的预测。

[0172] [第二实施例]

[0173] 在第一实施例中, 基础层的帧内预测模式可以用于目标以具有当前块和同位置块的第一近似值, 并且接下来的步骤与利用公式 (8) 至 (14) 详述的算法相对应。

[0174] 在第二实施例中, 下面给出了更复杂的情况的描述, 在该情况中, 用于对基础层和增强层进行编码的编码器算法彼此不同, 使得预测模式不兼容。简单的示例可以与利用 JPEG2000 (例如, 其在如下文献中描述: The JPEG-2000 Still Image Compression

Standard, ISO/IEC JTC Standard, 1/SC29/WG1, 2005; 以及Jasper Software Reference Manual (Version 1.900.0), ISO/IEC JTC, Standard 1/SC29/WG1, 2005) 编码的基础层和利用H.264编码的增强层相对应。在这种情况下, 第一实施例不适用, 因为帧内模式m在(例如, JPEG2000) 基础层中不可用。

[0175] 为了解决这个问题, 对基础层的像素测试(在增强层的编码器中可用的) 预测模式, 以检查那些解码的像素明显是可用的, 并且最终根据给定的准则来选择最佳帧内模式。

[0176] 通过以下公式(16)和(17)示出增强层和基础层的当前片和同位置片。

[0177] 当前片是:

$$[0178] \quad X = \begin{bmatrix} X_k^T \\ X_u^B \end{bmatrix} \quad (16)$$

[0179] 同位置片(X的同位置的)是:

$$[0180] \quad Y = \begin{bmatrix} Y_k^T \\ Y_k^B \end{bmatrix} \quad (17)$$

[0181] (索引为m的) 最佳帧内模式的选择是根据n个可能的帧内模式(例如, 它们与图3所示的模式相对应) 的集合 $S = \{m_0, \dots, m_{n-1}\}$  来实现的。为此, 根据索引为j的给定模式利用(同位置块 $Y_k^B$ 的) 虚拟预测 $Y_{prd,j}^B$ 来计算虚拟预测误差, 并且在块 $Y_k^B$ 和虚拟预测 $Y_{prd,j}^B$ 之间的虚拟预测误差 $ER_j$ 如公式(18)所示。

$$[0182] \quad ER_j = \sum_{p \in Y_k^B} \left( Y_k^B(p) - Y_{prd,j}^B(p) \right)^2 \quad (18)$$

[0183] 这里, p与预测块 $Y_k^B$ 和虚拟预测块 $Y_{prd,j}^B$ 中的像素的坐标相对应;  $Y_k^B(p)$ 是预测块 $Y_k^B$ 的像素值; 并且 $Y_{prd,j}^B(p)$ 是根据索引为j的帧内模式的虚拟预测块的像素值。

[0184] 通过如以下公式(19)所示的那样将来自n个可用帧内模式预测的虚拟预测误差最小化来给出最佳虚拟预测模式。

$$[0185] \quad J_{mode} = \underset{j}{\operatorname{Argmin}} \{ ER_j \} \quad (19)$$

[0186] 这里, 需要注意的是, 用于通过公式(18)计算虚拟预测误差的度量不限于平方差之和(SSE), 其它度量也是可能的: 绝对差之和(SAD)、绝对阿达玛变换差之和(SATD)。

[0187] 获得适用于同位置块 $Y_k^B$ 的虚拟预测 $Y_{prd,J_{mode}}^B$ , 然后使用相同模式( $J_{mode}$ )来计算专用于增强层的当前块( $X_u^B$ )的虚拟预测( $X_{prd,J_{mode}}^B$ )。

[0188] 如以下公式(20)和(21)那样来提供新的中间片。

[0189] 当前中间片 $X'$ :

$$[0190] \quad X' = \begin{bmatrix} X_k^T \\ X_{prdJ_{mode}}^B \end{bmatrix} \quad (20)$$

[0191] 基础层的中间片Y' :

$$[0192] \quad Y' = \begin{bmatrix} Y_k^T \\ Y_{prdJ_{mode}}^B \end{bmatrix} \quad (21)$$

[0193] 现在,当获得了中间虚拟预测块 $Y_{prdJ_{mode}}^B$ 和 $X_{prdJ_{mode}}^B$ 时,使用转移函数Trf从基础层找到当前块的(确定的)预测的处理与通过先前的公式(8)和(9)给出的处理类似。

[0194] 具有转移函数Trf,将该函数应用于片Y,在逆变换之后,给出从其中提取期望预测的片Y",如公式(22)所示。

$$[0195] \quad Y'' = \begin{bmatrix} Y''^T \\ Y_{J_{mode}}^B \end{bmatrix} \quad (22)$$

[0196] 在公式(22)中,当前块的预测是 $Y_{J_{mode}}^B$ 。这里,该处理与通过使用公式(13)、(14)和(15)以及此处的虚拟模式 $J_{mode}$ 而用于公式(12)的处理类似。

[0197] 图5A和图5B示出了对帧内SNR可伸缩性的这种描述的原理。图5是示出了本公开第二实施例的用于确定增强层的当前块的预测的装置的配置的框图。

[0198] 图5A中的编码器侧(单元500) :

[0199] 使用TMO 506对由块 $b_e$  501组成的原始HDR图像 $im_{e1}$ 进行色调映射,其给出原始色调映射图像 $im_{b1}$ 。

[0200] 基础层(b1)

[0201] 我们考虑编码原始基础层图像 $im_{b1}$ 。利用给定的视频编码器531,图像由编码器531编码,并由局部在环(in loop)解码器532局部解码。将局部解码的图像存储在“重构图像缓冲器”533中。在基础层比特流中发送所得到的编码图像。

[0202] 增强层(e1)

[0203] 我们现在考虑编码原始增强层块 $b_e$ 。

[0204] a) 对于增强层的当前块,我们考虑(索引为m的)增强层编码器帧内模式的所有可用的帧内编码模式,

[0205] o我们(利用公式(19)和“ $J_{mode} = \text{Argmin}_j \{FR_j\}$ ”单元542)(根据给定的准则(公式(19),以及增强层解码器的解码模式))从该同位置块的相邻像素中找到专用于(基础层)的同位置块的(索引为 $J_{mode}$ 的)最佳预测模式。

[0206] b) 利用增强层的(索引为 $J_{mode}$ 的)这种帧内模式,我们确定:

[0207] o(利用b1空间预测(Sp pred)单元541)确定基础层层级的帧内预测块 $Y_{prdJ_{mode}}^B$ ;



[0208] o然后,根据公式(21)利用同位置块( $y_k^B$ )的相邻物( $y_k^I$ )和预测块 $Y_{prd/jmode}^B$ 建立第一中间片 $Y'$ 。

[0209] c)类似地,利用基础层的(Jmode索引的)这种帧内模式,我们确定:

[0210] o(利用e1空间预测(Sp Pred)单元512)确定增强层层级的中间帧内预测块 $X_{prd/jmode}^B$ ;

[0211] o并且,然后根据公式(20)利用当前块( $b_e$ )的相邻物( $x_k^I$ )和中间预测块 $X_{prd/jmode}^B$ 来建立第二中间片 $X'$ 。

[0212] d)在变换域(例如,DCT)中,我们使用公式(8)至(11)来确定从片 $Y'$ 到片 $X'$ 的转移函数Trf。

[0213] e)现在,我们考虑根据公式(5)由同位置块( $y_k^B$ )及其相邻物 $y_k^I$ 组成的基础层Y的初始(解码)片。

[0214] 1.我们对片Y应用变换(例如,DCT): $TF(Y)$

[0215] 2.现在,将Trf函数应用于变换域,比如: $T_{Y''} = TF(Y) \cdot Trf$

[0216] 3.对 $T_{Y''}$ 计算逆变换(例如, $DCT^{-1}$ ),给出 $Y'' = TF^{-1}(T_{Y''})$ ,其中所得到的片是如公式(22)那样组成的。

[0217] 4.最后,从片 $Y''$ 提取与块 $Y_{mode}^B$ 相对应的预测。

[0218] 从b到e的所有步骤是在“预测e1/b1(Trf)”单元511中实现的。

[0219] f)通过T,Q单元503将(使用组合器502计算出的)在增强层块 $b_e$ 和在步骤a到e处计算的层间预测( $y_{mode}^B$ )之间的误差残量 $r_s$ 变换和量化为 $re_q$ ,并且 $re_q$ 由熵编码器504进行熵编码,并且在增强层比特流中发送。

[0220] g)最后,通过(用组合器514)将由 $T^{-1}Q^{-1}$ 单元505逆变换和解量化的预测误差块 $re_{dq}$ 添加到预测 $y_{mode}^B$ ,来局部地重建解码的块,并且将重构(解码)的图像存储在(e1)参考帧缓冲器508中。

[0221] 关于其它单元507和509,该功能分别专用于经典编码模式决定和用于图像间预测的运动估计。

[0222] 图5B中的解码器侧(单元550):

[0223] 基础层(b1)

[0224] 根据b1比特流,利用解码器584对基础层序列进行解码。重构图像缓冲器582存储用于层间预测的解码帧。

[0225] 增强层(e1)

[0226] a)根据e1比特流,对于给定块,熵解码器551对量化的误差预测 $re_q$ 进行解码。

[0227] b)通过 $T^{-1}Q^{-1}$ 单元552将残差预测 $re_q$ 解量化和逆变换以产生 $re_{dq}$ 。

[0228] c)如果用于解码的块的编码模式与我们的层间模式相对应,则需要基础层的同位置块的(索引为Jmode的)帧内模式。

[0229] o对于HDR层的当前块,我们考虑(索引为Jmode的)增强层编码器帧内模式的所有可用的帧内编码模式,

[0230] o(利用公式(19)和“ $Jmode = \text{Argmin}_j \{ER_j\}$ ”单元581)(根据给定的准则(公式

(19), 以及增强层解码器的解码模式)) 从该同位置块的相邻像素中找到专用于(基础层)的同位置块的(索引为Jmode的)最佳预测模式。

[0231] d) 利用增强层的(索引为Jmode的)这种帧内模式, 我们确定:

[0232] o 利用b1空间预测(b1 Sp Pred)单元583, 确定基础层层级的帧内预测块 $\hat{Y}_{pred, Jmode}^B$ ;

[0233] o 然后, 根据公式(21)利用同位置块( $Y_k^B$ )的相邻物( $Y_k^B$ )和预测块 $\hat{Y}_{pred, Jmode}^B$ 建立第一中间片Y'

[0234] e) 类似地, 利用基础层的(索引为Jmode的)这种帧内模式, 我们确定:

[0235] o 利用e1空间预测(Sp Pred)单元555, 确定增强层层级的中间帧内预测块 $\hat{X}_{pred, Jmode}^B$ ;

[0236] o 并且, 然后根据公式(20)利用当前块( $b_e$ )的相邻物( $Y_k^B$ )和中间预测块 $\hat{X}_{pred, Jmode}^B$ 来建立第二中间片X'。

[0237] f) 在变换域(例如, DCT)中, 我们使用公式(8)至(11)来确定从片Y'到片X'的转移函数Trf。

[0238] g) 现在, 我们考虑根据公式(5)由同位置块( $Y_k^B$ )及其相邻物 $Y_k^B$ 组成的基础层Y的初始(解码)片。

[0239] 1. 我们对片Y应用变换(例如, DCT):  $TF(Y)$

[0240] 2. 现在, 将Trf函数应用于变换域, 比如:  $T_{Y''} = TF(Y) \cdot Trf$

[0241] 3. 对 $T_{Y''}$ 计算逆变换(例如,  $DCT^{-1}$ ), 给出 $Y'' = TF^{-1}(T_{Y''})$ , 其中所得到的片是如公式(22)那样组成的。

[0242] 4. 最后, 从片Y''提取与块 $Y_{pred, Jmode}^B$ 相对应的预测。

[0243] 从c到g的所有步骤是在“预测e1/b1(Trf)”单元557中实现的, 我们可以注意到, 步骤d到h与(第而实施例的)编码器的步骤b到e严格相同; 显然, 如果e1编码器通过e1编码器(单元507)的模式决定选择该层间预测模式的话。

[0244] h) 通过下述方式来建立e1解码块: (使用组合器553)将(经由单元552)解码和解量化的预测误差块 $r_{edq}$ 添加到(经由预测单元554和单元557)的预测块 $\hat{Y}_{pred, Jmode}^B$ , 给出重构的(e1)块。

[0245] i) 将重构(或解码)的图像存储在(e1)参考帧缓冲器556中, 使用运动补偿单元558将解码的帧用于下一个(e1)帧内图像预测和帧间预测。

[0246] 根据该用于确定增强层的当前块的预测的方法和装置, 即使当基础层的编码模式不同于增强层的编码模式时, 也选择适当的层间编码模式, 然后可以获得当前块的预测。

[0247] [第三实施例]

[0248] 下面给出了本公开第三实施例的用于确定增强层的当前块的预测的方法和装置的描述。

[0249] 在空间可伸缩性方面, 基础层(1e)和增强层(1b)的空间分辨率彼此不同, 但是关于基础层的预测模式的可用性, 却存在不同的可能性。

[0250] 更具体地, 下面给出了其中空间可伸缩性处于相同视频编码标准(与第一实施例类似)的情况的描述。

[0251] 如果当前块( $X_u^B$ )的大小与基础层的块( $Y_k^B$ )的同位置上采样(up-sampled)相同, 则可以利用基础层的预测模式m, 并且在第一实施例中说明的处理可以应用于这种情况。例

如(在空间可伸缩性 $N \times N \rightarrow 2N \times 2N$ 的情况下),给定的 $8 \times 8$ 当前块在基础层中具有 $4 \times 4$ 同位置块。于是,帧内模式 $m$ 与用于编码( $1_b$ 层的)该 $4 \times 4$ 块的帧内编码模式相对应,并且 $8 \times 8$ 预测块 $Y_{prd,m}^B$ 可以是基础层的上采样预测( $4 \times 4 \rightarrow 8 \times 8$ ),或者可以利用相同的编码模式 $m$ 对基础层的上采样图像计算预测 $Y_{prd,m}^B$ 。作为第一实施例,当获得基础层和增强层中间预测块时,就建立基础层和增强层中间组块。之后,根据两个中间片,使用公式8至11来估算转移函数。最后,将转移函数应用于基础层的上采样和变换(ex DCT)的片,如第一实施例那样提取层间预测。

[0252] 相反,如果当前块( $X_u^B$ )的大小与基础层的块( $Y_k^B$ )的上采样不同,则编码模式 $m$ 事实上不可用。在这种情况下,可以使用在第二实施例中说明的原理。换言之,必须估计上采样的基础层中的最佳编码模式 $m$ , (专用于层间预测的) 其余处理与第二实施例相同;已知:所估计出的转移函数(Trf)应用于上采样和变换(ex DCT)的基础层片。

[0253] [第四实施例]

[0254] 下面给出了本公开第四实施例的用于确定增强层的当前块的预测的方法和装置的描述。

[0255] 基于LDR/HDR可伸缩视频编码,本公开的第四实施例为基础层的块提供了编码模式选择算法,以便重用所选择的模式,以利用第一实施例中提供的技术来建立预测( $1_b \rightarrow 1_e$ )。在基础层层级处选择编码模式可能导致两层层级处的固有失真。

[0256] 这里,RDO(速率失真优化)技术用于解决LDR和HDR的失真以及当前HDR和同位置的LDR块的编码成本,并且RDO准则给出了在重构误差以及基础层和增强层的编码成本之间提供了最佳折中的预测模式。为此,用于这两层的经典RDO准则如以下公式(23)和(24)那样提供。

$$[0257] \quad \text{LDR: } \text{Cst}_{bl} = \text{Dist}_{bl} + \lambda_{bl} \cdot B_{bl}^{\text{cst}} \quad (23)$$

$$[0258] \quad \text{HDR: } \text{Cst}_{el} = \text{Dist}_{el} + \lambda_{el} \cdot B_{el}^{\text{cst}} \quad (24)$$

[0259] 术语 $B_{bl}^{\text{cst}}$ 和 $B_{el}^{\text{cst}}$ 分别由基础层和增强层的预测的误差残量的DCT系数的编码成本以及块( $B_{bl}^{\text{cst}}$ 和 $B_{el}^{\text{cst}}$ ) (其允许在解码器侧重构预测)的头部中包含的语法元素(块大小、编码模式等等)组成。

[0260] 考虑基础层的块 $Y_{or}^B$  (作为原始块)的示例,对在逆量化和逆变换(例如, $\text{DCT}^{-1}$ )之后的预测的误差残量的系数进行量化,将该残差添加到预测提供了重构(或解码)的块( $Y_{dec}^B$ )。利用原始块 $Y_{or}^B$ 和解码的块 $Y_{dec}^B$ ,与该块相关联的基础层失真如以下公式(25)那样被提供。

$$[0261] \quad \text{Dist}_{bl} = \sum_{p \in Y_{or}^B} \left( Y_{or}^B(p) - Y_{dec}^B(p) \right)^2 \quad (25)$$

[0262] 在RDO准则中,使用众所周知的参数 $\lambda_{bl}$ 以便给出最佳的折中速率失真。在该示例

中,如以下公式 (26) 那样来提供N个可能模式中的最佳模式。

$$[0263] \quad J_{\text{mode}}^{\text{bl}} = \underset{j}{\text{Argmin}}\{Cst_{\text{bl}}^j\} \quad (26)$$

[0264] 可以将公式 (23) 和 (24) 重新书写成其它形式,如公式 (27) 和 (28) 所示。

$$[0265] \quad \text{LDR}: Cst'_{\text{bl}} = \frac{\text{Dist}_{\text{bl}}}{\lambda_{\text{bl}}} + B_{\text{bl}}^{\text{cst}} \quad (27)$$

$$[0266] \quad \text{HDR}: Cst'_{\text{el}} = \frac{\text{Dist}_{\text{el}}}{\lambda_{\text{el}}} + B_{\text{el}}^{\text{cst}} \quad (28)$$

[0267] 公式 (27) 和 (28) 可以利用混合参数 $\alpha$ 进行混合,这允许基础层和增强层之间的全局折中,如以下公式 (29) 所示。

$$[0268] \quad Cst' = \left( \frac{\text{Dist}_{\text{bl}}}{\lambda_{\text{bl}}} + B_{\text{bl}}^{\text{cst}} \right) \cdot (1 - \alpha) + \left( \frac{\text{Dist}_{\text{el}}}{\lambda_{\text{el}}} + B_{\text{el}}^{\text{cst}} \right) \cdot \alpha \quad (29)$$

[0269] 其中

$$[0270] \quad 0 \leq \alpha \leq 1$$

[0271] (根据公式 (29) 的) 最佳模式给出了基础层的模式,其经由基础层的N个编码模式中的一个来产生最小全局成本 $Cst'$ ,如以下公式 (30) 所示。

$$[0272] \quad J_{\text{mode}}^{\text{bl}} = \underset{j}{\text{Argmin}}\{Cst'_j\} \quad (30)$$

[0273] 根据该公式 (30),注意到以下事项。

[0274] 如果 $\alpha=0$ ,则情况与在第一实施例中提出的算法相对应,其中,可以使用基础层的(索引为m的)编码模式,来经由转移函数Trf建立层间预测( $\text{bl} \rightarrow \text{el}$ ),并且最终利用  $m = J_{\text{mode}}^{\text{bl}}$  来提供层间预测  $Y''_{\text{m}}^{\text{B}}$ 。

[0275] 相反,如果 $\alpha=1$ ,则编码模式的选择主要集中在增强层,并且存在基础层包含大量视觉伪影的风险。

[0276] 如果 $\alpha=0.5$ ,则两个层之间的折中是必要的。在这种情况下,重要的是要注意:基础层的编码模式的选择实际上不仅基于在基础层层级处的影响还基于在增强层层级处的影响,更确切地基于:

[0277] • 基础层编码模式的选择对基础层的影响;以及

[0278] • 基于先前的基础层编码模式使用在第一实施例中说明的整个处理对增强层(即层间预测)的影响。

[0279] 图6示出了第四实施例的用于确定增强层的当前块的预测的装置的框图。

[0280] 参考图6,描述了局部层间预测。因此,为了描述,仅描述使用帧内模式(m)的帧内图像预测模式,因为我们的层间预测模式使用帧内模式(m)。

[0281] 注意,因为在第四实施例中的相关联的解码器与第一实施例中的相同并且与图4b所示的解码器相对应,因此仅描述编码器侧。

[0282] 图6中的编码器侧(单元600):

[0283] 使用TMO 606来色调映射原始块601b<sub>bc</sub>,其给出原始色调映射块b<sub>e</sub>。

[0284] 注意,在第四实施例的层间预测的特定情况下,不使用单元625和607(与基础层和

增强层的编码模式决定单元相对应)。在该情况下,单元642替代单元625和607,事实上,单元642使用公式30选择最佳帧内模式 $J_{mode}^{bl}$ ,并且向单元625和607发送该模式( $J_{mode}^{bl}$ )。

[0285] 单元642中的基础层帧内编码模式选择( $J_{mode}^{bl}$ )

[0286] 对于给定的混合参数 $\alpha$ (其允许基础层和增强层之间的全局折中如下面的公式(29)所示),并且对于每N个可用的帧内预测模式(图3所示,在H264的情况下),我们对编码模式操作N次迭代:

[0287] 关于索引m的N个帧内模式的循环{

[0288] a) 利用基础层的相邻重构(或解码)像素、空间预测和帧内编码模式m(m是索引), (Sp Pred) 单元658给出帧内基础层预测块

[0289] b) 利用增强层的相邻重构(或解码)像素、空间预测和相同的m帧内编码模式, (Sp Pred) 单元612给出中间帧内增强层预测块

[0290] o单元611建立由帧内基础层相邻物和步骤(a)的预测块组成的基础层的片

[0291] o单元611建立由帧内增强层相邻物和步骤(b)的预测块组成的增强层的片

[0292] o在变换域(例如,DCT)中,(在单元611中)使用公式(8)至(11)来确定从片Y'到片X'的转移函数Trf。

[0293] o仍然在单元611中,

[0294] 考虑根据公式(5)由同位置块( $Y_{\text{pos}}^{\text{pos}}$ )及其相邻物 $Y_{\text{pos}}^{\text{pos}}$ 组成的基础层Y的初始(解码)片

[0295] 对片Y应用变换(例如,DCT):  $TF(Y)$

[0296] 将Trf函数应用于变换域,比如:  $T_{Y''} = TF(Y) \cdot Trf$

[0297] 对 $T_{Y''}$ 应用逆变换(例如,  $DCT^{-1}$ ),给出 $Y'' = TF^{-1}(T_{Y''})$ ,其中所得到的片是如公式(12)那样组成的。

[0298] 从片Y''中提取与块 $Y_{\text{pos}}^{\text{pos}}$ 相对应的预测

[0299] c) 在单元642中,选择(根据公式(29)的)最佳模式,该最佳模式经由N个编码模式中的一个来产生最小全局成本 $C_{st}'$ (公式(30))

[0300] }关于索引m的N个帧内模式的循环结束

[0301] 最后,将最佳帧内 $J_{mode}^{bl}$ 发送给基础层空间预测单元658和决定单元607,并且发送给增强层单元611。

[0302] 当找到 $J_{mode}^{bl}$ 时,处理的其余部分与第一实施例的编码器的描述类似,其中已知基础层帧内模式索引 $m = J_{mode}^{bl}$ 。

[0303] 基础层(b1)

[0304] 我们考虑编码原始基础层块 $b_{bc}$

[0305] d) 利用参考帧缓冲器626中存储的原始块 $b_{bc}$ 和(先前解码的)图像,运动估计器(运动估计单元)629利用给定的运动矢量(时间预测单元)找到最佳帧间图像预测块,并且时间预测(Temp Pred Pred)单元630给出时间预测块。

[0306] e) 如果模式决定处理(单元625)选择(索引为 $m = J_{mode}^{bl}$ 的)帧内图像预测模式,则

利用原始块 $b_{bc}$ 和预测块 $\hat{Y}_m^B$  ( $\hat{Y}_{pred,m}^B$ )之间的差(通过组合器621)来计算残差预测 $r_b$ 。

[0307] f) 之后,通过TQ单元622将残差预测 $r_b$ 变换和量化为 $r_{bq}$ ,并且最终由熵编码器单元623进行熵编码并在比特流基础层中发送。

[0308] g) 通过(利用组合器657)将由 $T^{-1}Q^{-1}$ 单元624逆变换和解量化后的预测误差块 $r_{bdq}$ 添加到预测块 $\hat{Y}_m^B$ ,给出重构(基础层)块,来局部重建解码的块

[0309] h) 将重构(或解码)的帧存储在(b1)参考帧缓冲器626中。

[0310] 增强层(e1)

[0311] 我们可以注意到,增强层的编码器的结构与基础层的编码器相似,例如,就编码模式决定、时间预测和参考帧缓冲器方面而言,单元607、608、609和613具有与基础层的编码器的各个单元625、626、629和630相同的功能。

[0312] 我们现在考虑编码原始增强层块 $b_e$ 。

[0313] i) 对于增强层的块,如果基础层的同位置块是以帧内图像模式被编码的,则我们考虑该同位置块的(索引为 $m$ 的,其中 $m = J_{mode}^{bl}$ )帧内模式。

[0314] j) 利用基础层的(索引为 $m$ 的)这个帧内模式,我们确定:

[0315] o 利用b1空间预测(Sp pred)单元658,确定或重用基础层层级的帧内预测块 $\hat{Y}_{pred,m}^B$ ,

[0316] o 然后,根据公式(7)利用同位置块( $\hat{Y}_m^B$ )的相邻物( $\hat{Y}_k^B$ )和预测块 $\hat{Y}_{pred,m}^B$ 建立第一中间片 $Y'$

[0317] k) 类似地,利用基础层的(索引为 $m$ 的)这种帧内模式,我们确定:

[0318] o (利用e1空间预测(Sp Pred)单元612)确定增强层层级的中间帧内预测块 $\hat{X}_{pred,m}^B$ ,

[0319] o 并且,然后根据公式(6)利用当前块( $b_e$ )的相邻物( $\hat{X}_k^B$ )和中间预测块 $\hat{X}_{pred,m}^B$ 来建立第二中间片 $X'$

[0320] l) 在变换域(例如,DCT)中,我们使用公式(8)至(11)来确定从片 $Y'$ 到片 $X'$ 的转移函数Trf。

[0321] m) 现在,我们考虑根据公式(5)由同位置块( $\hat{Y}_m^B$ )及其相邻物 $\hat{Y}_k^B$ 组成的基础层 $Y$ 的初始(解码)片。

[0322] 5. 我们对片 $Y$ 应用变换(例如,DCT):  $TF(Y)$

[0323] 6. 现在,将Trf函数应用于变换域,比如:  $T_{Y''} = TF(Y) \cdot Trf$

[0324] 7. 对 $T_{Y''}$ 计算逆变换(例如, $DCT^{-1}$ ),给出 $Y'' = TF^{-1}(T_{Y''})$ ,其中所得到的片是如公式(12)那样组成的。

[0325] 8. 最后,从片 $Y''$ 提取与块 $\hat{Y}_m^B$ 相对应的预测。

[0326] 从j到m的所有步骤是在“预测e1/b1(Trf)”单元611中实现的。

[0327] n) (T Q单元603)将(使用组合器602)计算的在增强层块 $b_e$ 和在步骤j到m处计算的层间预测( $\hat{Y}_m^{B'}$ )之间的误差残量 $r_s$ 变换和量化为 $r_{eq}$ ,并且 $r_{eq}$ 由熵编码器单元604进行熵编码并且在增强层比特流中发送

[0328] o) 最后,通过(用组合器610)将由 $T^{-1}Q^{-1}$ 单元605逆变换和解量化的预测误差块 $r_{edq}$ 添加到预测 $\hat{Y}_m^{B'}$ ,来局部地重建解码的块,并且将重构(解码)的图像存储在(e1)参考帧缓冲器608中。

[0329] 如上所述,本公开的实施例涉及对于两层使用相同或不同的编码器来进行编码的SNR和空间可伸缩LDR/HDR视频编码。可以利用任何色调映射算子根据HDR视频来实现LDR视频,所述色调算子为:全局或局部的、线性或非线性的。在实施例的可伸缩解决方案中,层间预测是实时实现的,而无需附加的特定元数据。

[0330] 本公开的实施例涉及编码器和解码器两者。本公开的实施例应用于总体公开的解码处理,并且根据本公开的实施例,所述解码是可检测的。

[0331] 本公开的实施例可以应用于图像和视频压缩。特别地,本公开的实施例可以被提交给ITU-T或MPEG标准化组,来作为专用于LDR/HDR视频内容的存档和分发的新一代编码器的开发的一部分。

[0332] 这里记载的所有示例和条件语言旨在用于教导目的,以帮助读者理解本发明人为改进现有技术而贡献的公开内容和构思,并且应被解释为不限于这些具体记载的示例和条件,说明书中这些示例的组织也不涉及示出本公开的优势或劣势。

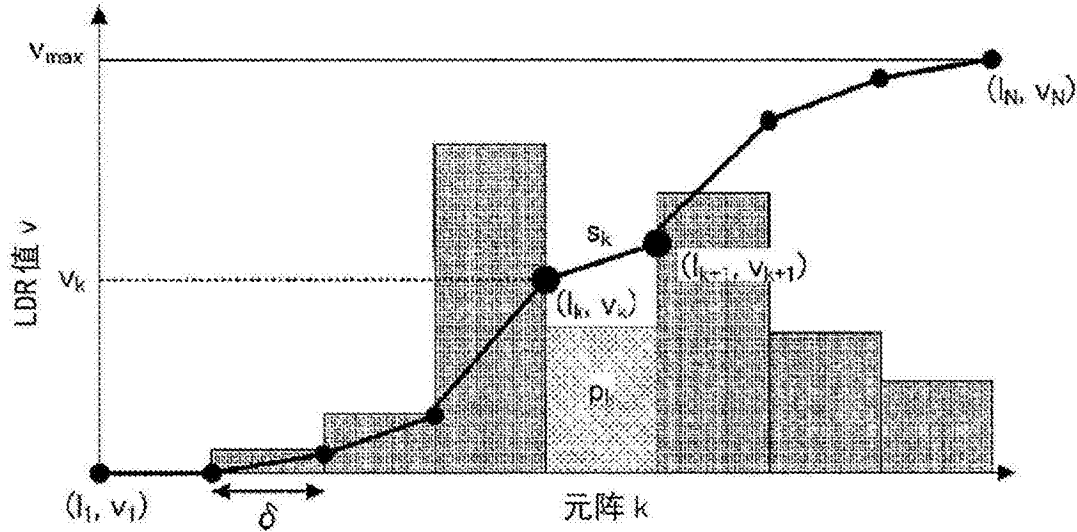


图1

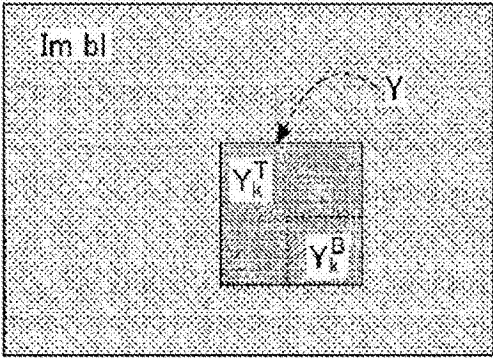


图2A

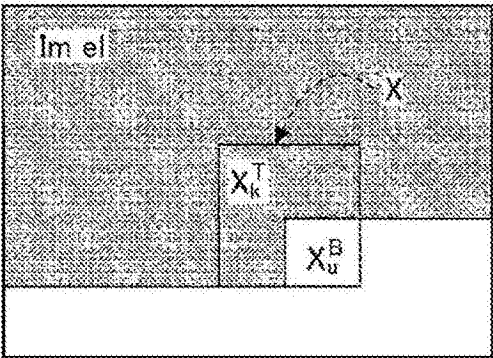


图2B



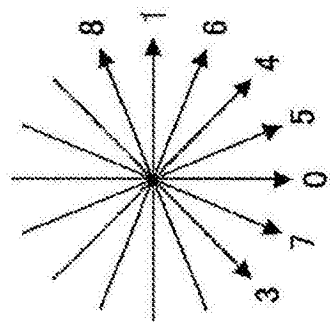


图3A

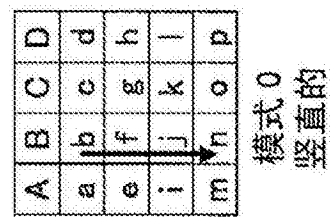


图3B

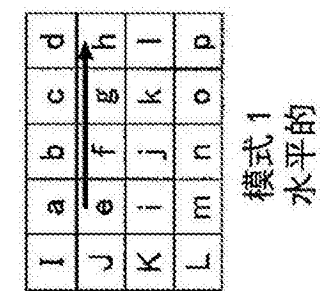


图3C

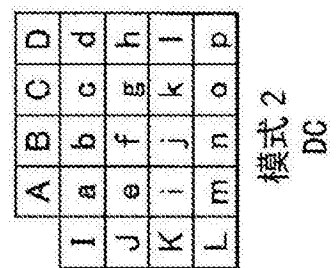


图3D

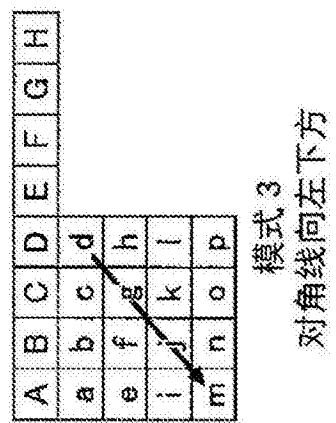


图3E

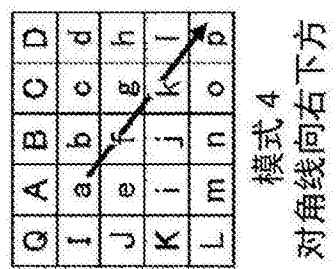


图3F

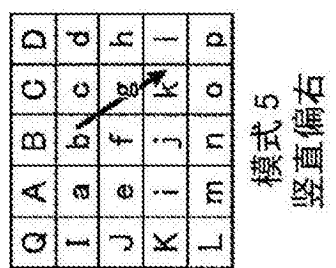


图3G

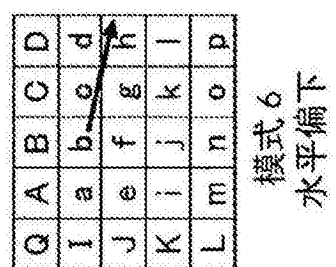


图3H

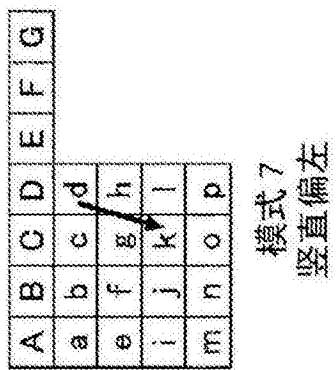


图3I

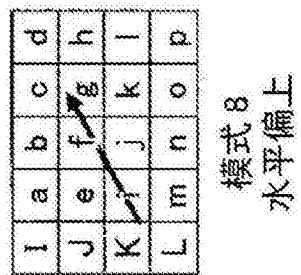


图3J

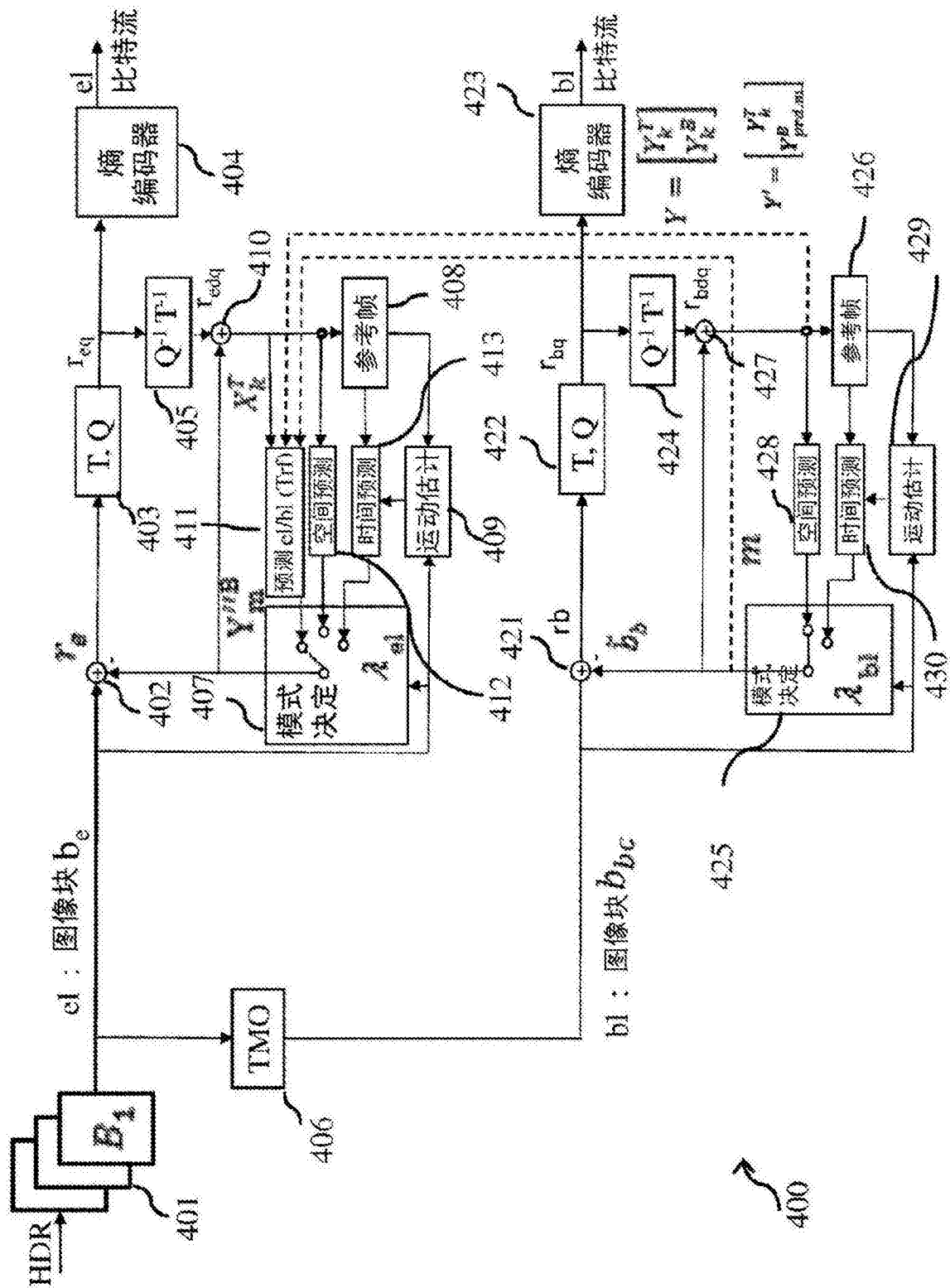


图4A

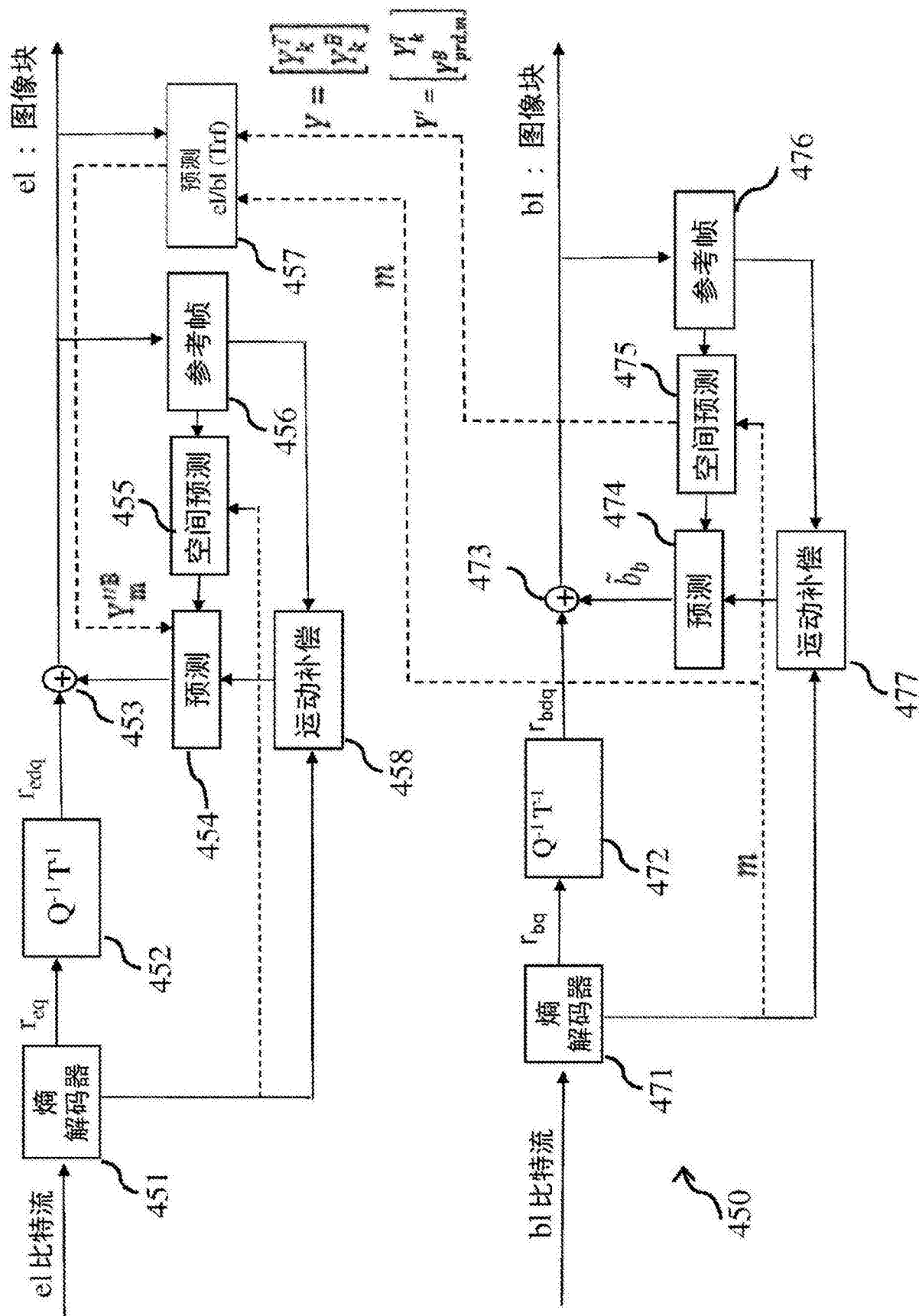


图4B

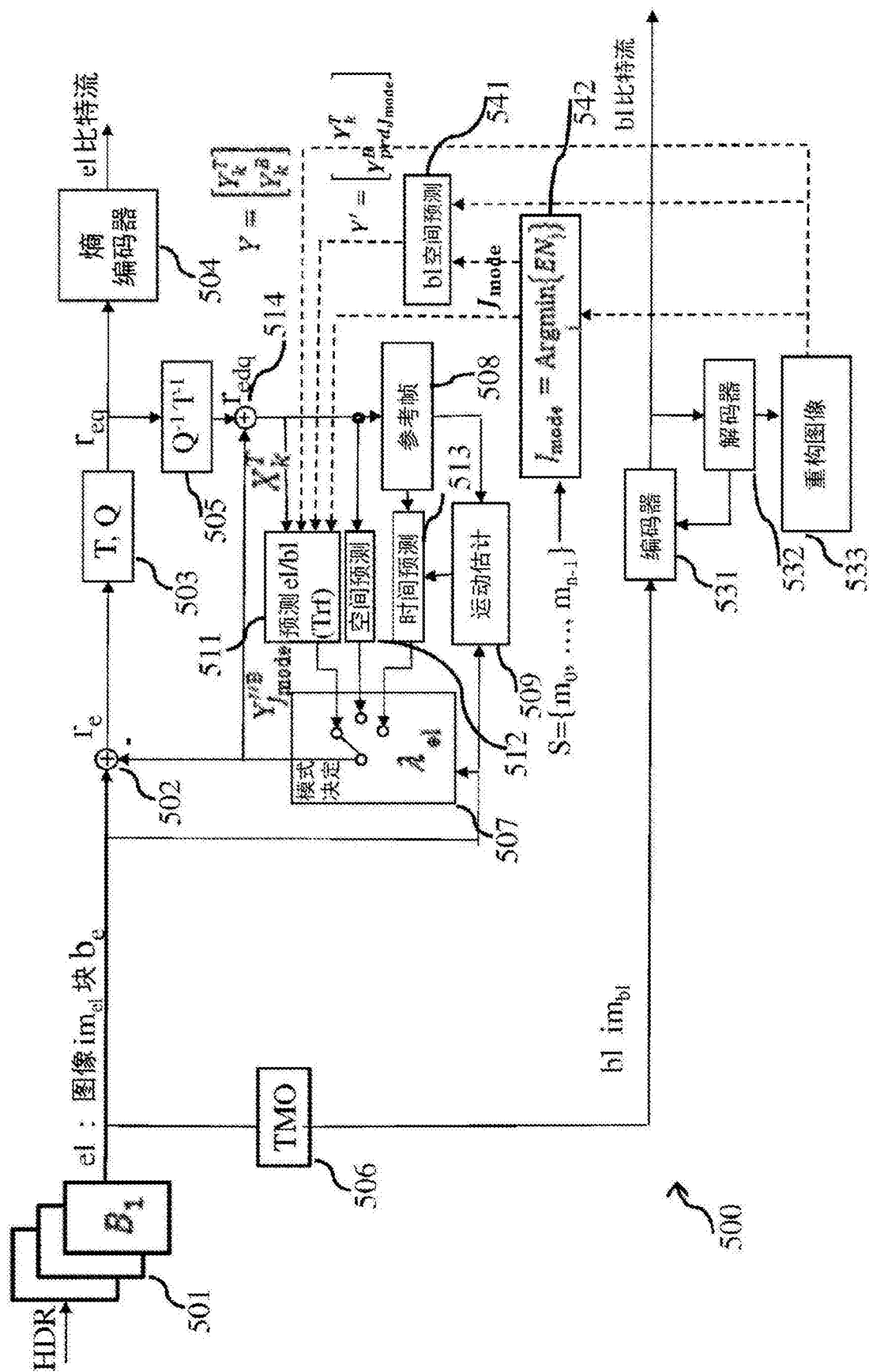


图5A

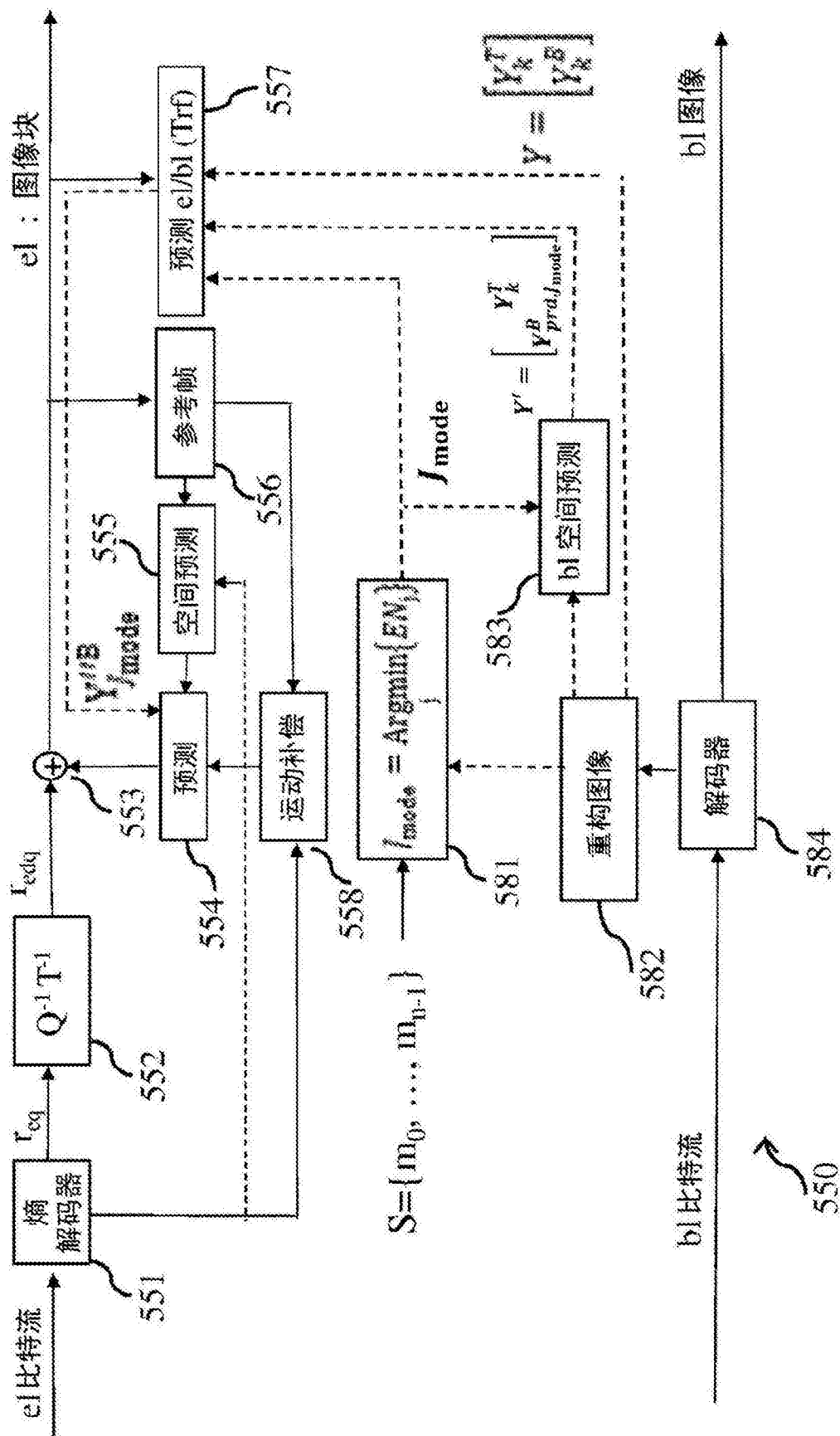
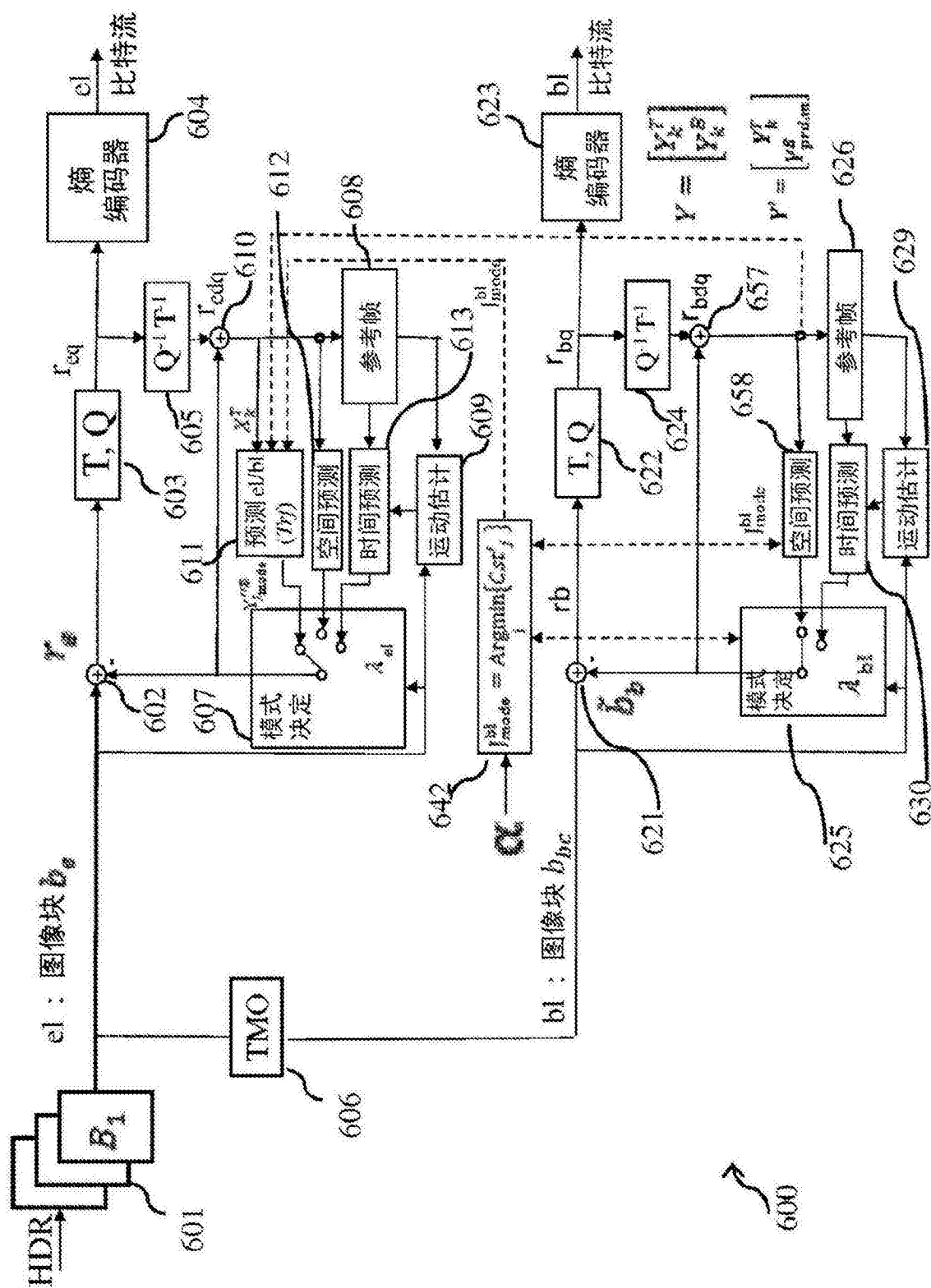


图5B





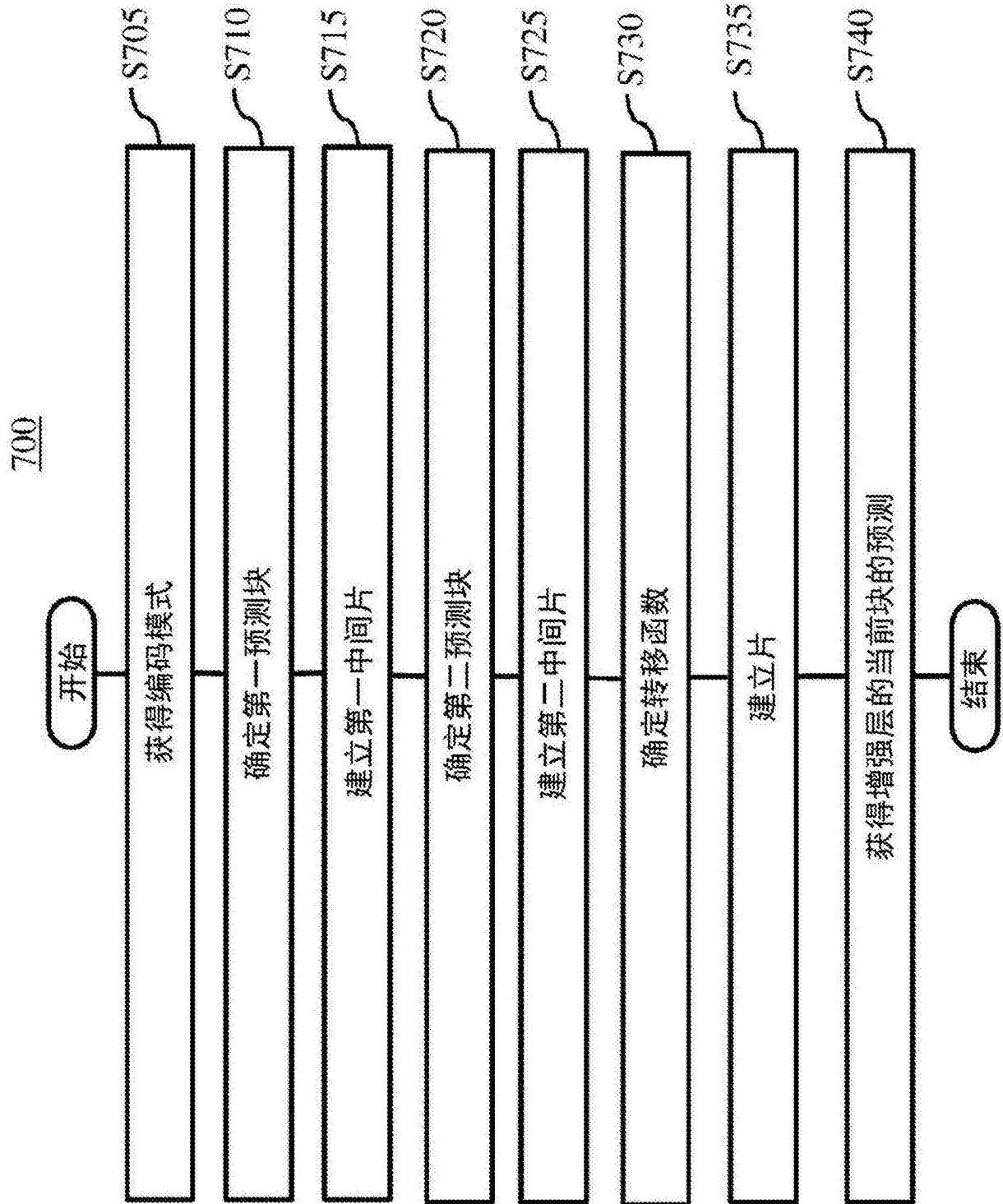


图7