引言

## HEVC关键技术

### HEVC编码整体框架

HEVC视频编码层沿用以往视频编码标准的混合编码方式，其整体编码框架如图所示，其输入为原始视频序列，输出为符合HEVC标注你的比特流。简要的编码过程如下。

（1）将每一帧图像划分为大小不同的图像块单元，并将相应的块划分信息加入到码流中，传到解码器。

（2）对每个单元进行帧内或帧间预测，原始像素值和预测值相减形成该单元的残差；若为帧间预测，则进行运动估计和运动补偿，对需要用到的重建图像要提前进行去方块滤波和自适应采样值补偿（ASO）滤波。

（3）对每个单元的残差进行整数变换（近似离散余弦变换和正弦变换），对得到的变换系参数继续进行量化和扫描。

（4）对量化后的变化系数、预测信息、模式信息、运动信息和头信息等进行熵编码，新城压缩的视频码流（语法元素）输出。



### HEVC编码结构

在HEVC中，一个很重要的革新之处就是其图像分块方式。HEVC针对预测和变换编码而对图像采取了一种基于四叉树的划分方式。其基本划分层次如图所示。



视频序列由若干时间连续的图像组成，在对这些图像进行处理时，HEVC首先将其分为若干图像组（Group Of Picture，GOP），其大小可以通过配置文件进行设置。每一组GOP由多帧图像组成，每一帧图像即为HEVC中四叉树划分的基本单位，每一帧图像经过划分，形成覆盖全帧的多个同样尺寸的编码树块（Coding Tree Block，CTB）。CTB还可以进一步分割为更小的编码块（Coding Block，CB）。CB时HEVC中进行视频编码算法的基本单位，它还可以划分为预测块（Prediction Block，PB）和变换块（Transform Block，TB）。接下来依次介绍它们的划分方式。

1. 编码树块（CTB）和编码树单元（CTU）的划分

HEVC将一帧编码图像划分为同意大小、相邻但不重叠的2N×2N样点的编码快CTB，这一点类似H.264/AVC中的宏块（Macro Block，MB）。CTB的尺寸可以为16×16、32×32或64×64。同一位置的亮度CTB和2块色度CTB，以及相应的语法元素和所包含的CU形成一个CTU。

CTU可以按照四叉树结构分解为若干方型编码单元（CU）,统一层次的CU必须时统一尺寸的4个方块，最多可以有4层，即8×8、16×16、32×32或64×64。如果不分解，则一个CTU仅包含一个CU，此时亮度CTU的尺寸就是亮度CB的最大尺寸。每个CU包含一个亮度编码块（CB）以及两个色度编码块（CB）。同时还包含相应的语法元素，如预测模式（帧内、帧间）、PU划分、从属的PU和TU信息等。CU时决定进行帧内预测还是帧间预测的单元，因此整个CU只会有一种预测模式。

CTB与CB之间的关系可以用一棵四叉树表示。CTB为树的根节点，往下还可以划分为更小的方块，由于分一层划分都是等分为四块，因此在树结构上形成四个分支，最后构成四叉树。以64×64尺寸的CTB为例，其划分方式如图（a）所示，对应四叉树结构如图（b）所示，扫描顺序按图（a）中索引顺序。





1. 预测单元的划分

预测单元PU是HEVC进行预测运算的基本单元，只能定义在不能继续划分的最底层CU中。CU决定了其包含的所有PU的预测方式，包括帧内预测和帧间预测。一个CU可以划分为一个或多个预测单元PU。CU到PU的划分仅允许一层。若CU的大小为2N×2N，当金星帧内预测，则可选的PU模式有2N×2N和N×N两种；如果进行帧间预测，则划分方式除了2N×2N、N×N、2N×N、N×2N以外，还有4种非对称划分方式。除此之外，帧间预测的跳过模式中只允许2N×2N的划分。

PU具体的划分方式如下图所示。



1. 变换单元的划分

变换单元是进行变换和量化操作的基本单元，和PU划分类似，他也是在CU的基础上进行划分的，但它不受所在CU的预测单元（PU）划分的限制。TU的划分方式也是一种四叉树划分。HEVC中TU的尺寸可以为4×4、8×8或32×32，这是因为最大DCT变换运算的尺寸为32×32。变换单元的最大尺寸以及四叉树的层级可以根据不同的应用进行相应的配置，对实时性或复杂度较低的应用可以通过增加四叉树的深度来提高编码效率。如图所示是一个深度为3的TU四叉树划分。



### HEVC帧内预测编码

帧内预测编码利用了视频序列每一帧内的空间相关性，利用当前图像中已经编码的部分对当前正在编码的块进行像素值预测。这一思路基于同一幅图像中存在的平坦区域和纹理区域，利用不同位置上像素块的相关性，从而达到去除空间域的冗余，达到压缩编码的目的。

HEVC的帧内预测模式相比于H.264/AVC，很重要的一个变化就是添加了更加细分的预测角度，大大增加了预测方向。如图所示，HEVC中共有33种角度预测模式，联通直流和平面模式，共35种预测模式。



对于每种预测模式，帧间预测主要分为三步：参考像素的准备，参考像素的平滑滤波，以及根据参考像素计算出预测像素的预测值。下面分别对帧内预测的每个过程进行简要阐述。

1. 参考像素的准备

与H.264/AVC不同，HEVC中规定当前预测块的左下方和右上方的像素也可以作为参考像素块，这是因为HEVC需要支持不同大小的编码快，所以常出现左下方和右上方像素块已经完成编码的情况。如图所示，设参考像素坐标为p[x][y]，其中x为横坐标，y为纵坐标，则左边参考像素的坐标为x=-1，y=-1，-2,…,-nTbS×2+1，上方的参考像素坐标为x=0，1，…,nTbS×2-1，y=-1。确定像素坐标后，HEVC根据坐标的可用性采用三种不同的填充方式。如果所有参考像素p[x][y]都可用，则p[x][y]直接取重建图像中对应像素的值；如果所有p[x][y]都被标记为不可用，则所有p[x][y]都填充为，其中bitDepth为比特深度，比如在常见的比特深度为8时，p[x][y]填充为128；如果至少一个p[x][y]被标记为不可用，HEVC根据相应规则将其临近可用的像素值填充进p[x][y]。



1. 参考像素的平滑滤波

为了提高帧内预测的效率，HEVC对8×8或更大的帧内PU的部分参考像素在预测前进行平滑滤波处理，从而减少噪声对预测的影响。对于不同的模式，选取需要进行平滑滤波的块也有所不同。对于4×4预测块的参考像素集，不需要进行平滑滤波。对于其他尺寸，如果是直流预测模式（DC），也不需要平滑滤波，如果是平面预测模式（Planner），则需要平滑滤波。对于角度模式，当参考像素及尺寸为8×8，只需要对3个对角方向的模式，即2、18、34模式进行平滑滤波。对16×16预测块的参考像素，出9、10、11和25、26、27外，其余模式都需要平滑滤波。对32×32预测块，有常规平滑滤波和强滤波两种处理方式。其中常规平滑滤波可通过3抽头滤波器来实现，强平滑滤波则用一个双线性插值滤波器实现[1]。

1. 根据参考像素计算预测块

完成像素值填充和平滑滤波之后，HEVC针对不同的预测模式，根据参考像素值计算出预测块的像素值。其种主要分为角度类预测和非角度类预测算法，非角度类预测又包含直流和平面预测两种。当预测模式为平面模式时，预测块p由像素由a、b、c、d个参考像素来决定，具体如图所示。



记，则预测值的计算公式为:

当预测模式为直流模式时，首先对所有像素块计算怕平均值dcVal作为基本直流量。计算公式为：

当预测模式为角度模式时，左侧角度模式预测值和上方角度模式预测值的计算方法类似。从（预测模式示意图）种可以看出，对于编号为18～34的模式 ，各个模式的角度分布是不均匀的，越是靠近垂直和水平两个方向，分布越密集。这以设计是因为自然界中大量的景物是垂直或水平分布的。在角度预测模式种，编号2～18称为水平方向类预测模式，18～34称为垂直方向类预测模式。一水平方向类的上方角度预测模式为例，介绍帧内预测值的计算方式。对角度模式modelIntra，尺寸为nTbS的块中某一位置的像素计算可分为三步。

首先找出该预测快的参考像素集，的值根据预测模式角度参数intraPredAngle和相应的反角度参数invAngle确定。具体如下：

然后设置不同模式的索引Idx和权重Fact，具体计算方法如下：

最后计算预测值predSample,。具体计算方法如下，其中预测值predSample记作pS：

### 变换与量化

HEVC采用混合编码框架，首先对视频数据进行预测，然后对预测结果的残差进行整数变换，再对变换后的系数进行量化，从而提取低频信息，达到提高编码压缩效率的目的。下面依次介绍基本的量化和变换算法。

1. 二维离散（DCT）矩阵与离散变换

离散变换的基本思想是将一个实函数对称延拓成一个实偶函数，再对偶函数进行离散傅里叶变换（DFT），其物理意义与傅里叶变换基本一致。由于DCT的正变换与逆变换核都是相同的，并且相比于DFT而言没有复杂的复数运算，因此在图像领域使用广泛。此处以4×4矩阵的变换为例，在二维DCT变换中，N×N的信号矩阵f，经过变换后形成矩阵F，其正变换和逆变换公式如下：

其中C为二维DCT变换矩阵，对于C的第u行第v个元素记为，则该元素可表示为：

其中x,u=0,1,,N-1。

1. 二维离散正弦(DST)矩阵与离散正弦变换

## CUDA关键技术