

编者按:随着视频信息的日益普及,视频流量的飞速增长,以“智能”为标志的视频(图像)分析技术和以“高效”为标志的视频压缩技术近年来取得了长足的进展。新一代视频编码国际标准 HEVC(High Efficiency Video Coding)在ITU-T的 VCEG 和 ISO/IEC 的 MPEG 通力合作下已经开发成功。智能视频分析正处于发展的起步阶段,但是目标十分明确,即在视频的采集、存储、检索、分析和传输中向人的智能化方向逼近。为此,本刊特别邀请了本校活跃在“视频分析与编码新技术”前沿的部分专家,撰写了新一代视频编码标准、自适应背景更新及运动目标检测算法、基于颜色纹理联合特征直方图的自适应 Meanshift 跟踪算法、多频点视频编码框架及关键技术、多媒体通信业务流识别与分类方法综述、基于边界平滑性准则的立体图像误码掩盖算法等八篇相关学术论文,以飨读者。

新一代视频编码标准——HEVC

朱秀昌,李欣,陈杰

(南京邮电大学 江苏省图像处理与图像通信重点实验室,江苏 南京 210003)

摘要:在2003年制定的H.264/AVC视频编码标准获得巨大的成功后,新一代视频编码国际标准 HEVC(High Efficiency Video Coding)在ITU-T的 VCEG 和 ISO/IEC 的 MPEG 通力合作下已经开发成功。HEVC提供了多项先进的视频编码技术。尽管 HEVC 的视频编码层结构仍然是常见的基于块运动补偿的混合视频编码模式,但是和先前的标准相比具有多处重要改进。文中对 HEVC 标准的技术的主要特点和性能进行了综述。

关键词:H.264/AVC; HEVC; 视频编码标准; JCT-VC; 视频压缩

中图分类号:TN919.8 **文献标志码:**A **文章编号:**1673-5439(2013)03-0001-11

Next Generation Video Coding Standard—HEVC

ZHU Xiu-chang, LI Xin, CHEN Jie

(Jiangsu Province Key Lab on Image Processing & Image Communication, Nanjing University of Posts and Telecommunications,
Nanjing 210003, China)

Abstract: After the profound success of defining H.264/AVC video coding standard in 2003, the next video coding international standard HEVC has been developed and standardized collaboratively by both the ITU-T VCEG and ISO/IEC MPEG organizations. HEVC represents a number of advances in video coding technology. Its video coding layer design is based on conventional block-based motion compensated hybrid video coding concepts, but with some important differences relative to prior standards. This paper provides an overview of the main technical features and characteristics of the HEVC standard.

Key words: H.264/AVC; HEVC; video coding standards; JCT-VC; video compression

收稿日期:2013-04-24;修回日期:2013-05-06

基金项目:国家自然科学基金(61071091)和“信息与通信工程”江苏高校优势学科建设工程资助项目

通讯作者:朱秀昌 电话:(025)83492402 E-mail:zhuxc@njupt.edu.cn

3种编码结构(Coding Structures)^[8],即帧内编码、低延时编码和随机访问编码。在帧内(All Intra)编码结构中,每一帧图像都是按帧内方式进行空间域预测编码,不使用时间参考帧。在低时延(Low Delay)编码结构中,只有第一帧图像按照帧内方式进行编码,并成为IDR(Instantaneous Decoding Refresh)帧,随后的各帧都作为一般的P帧和B帧(generalized P and B picture,GPB)进行编码。这主要是为交互式实时通信设计的。随机访问(Random Access)编码结构主要由分批B帧构成(Hierarchical B Structure),周期性地(大约每隔1秒)插入一度随机访问(Clean Random Access,CRA)帧。这些具有一定密度HEVC的新定义的CRA帧,成为编码视频流中的随机访问点(Random Access Point,RAP)。所谓随机访问点是指对这些帧的解码可以独立进行,不需要参考比特流中前面已经解码的图像帧。这种对随机访问方式有力地支持了信道转换、搜索以及动态流媒体服务等应用。

1.2 条和片的划分

条(Slice)的划分在H.264/AVC中就有,HEVC也允许将图像帧划分为若干条,条是HEVC中独立的编码区域^[9]。图2显示了一帧图像中3个条的划分情况,条的划分以CTU或更小的CU为界。为了支持并行计算和差错控制,某一个条可以划分为更小的条,谓之“熵条”(Entropy Slice,ES)。每个ES都可独立地进行熵解码,而毋须参考其它的ES。如在多核的并行处理中,就可以安排每个核单独处理一个ES。在HEVC的码流中,NAL层比特流的格式符合H.264/AVC的Annex B,但是在NAL头信息增加了1Byte的HEVC标注信息。每个条编码为一个NAL单元,其容量小于等于最大网络传输单元(Maximum Transmission Unit,MTU)容量。

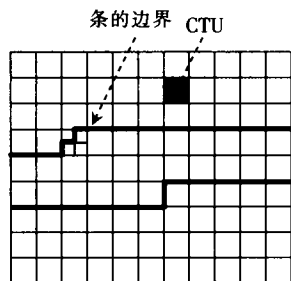


图2 图像中条划分的示例

在HEVC中新引入了片(Tile)的划分,用水平和垂直的若干条边界将图像帧划分为多个矩形区域,每个区域就是一个片,每一个片包含整数个CTU,片之间也可以互相独立,以此实现并行处理。

图3显示了一例图像中12个片的划分情况。片划分时并不要求水平或垂直边界均匀分布,可根据并行计算和差错控制的要求灵活掌握。在编码时,图像中的片也是按扫描顺序进行处理,每个片中的CTU也是按扫描顺序进行。在HEVC中,允许条和片在同一图像帧中同时使用,既可以一个条中包含若干片,又可以一个片中包含若干条。

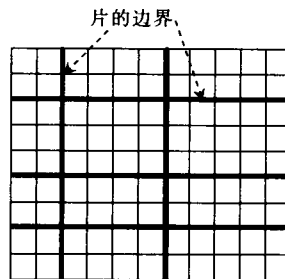


图3 图像中片划分的示例

1.3 四叉树单元划分

HEVC依然采用分块(Block)编码方式,但块的尺寸是可以自适应改变的。在HEVC中包含3类单元,即编码单元(Coding Unit,CU),预测单元(Prediction Unit,PU)和变换单元(Transform Unit,TU)。

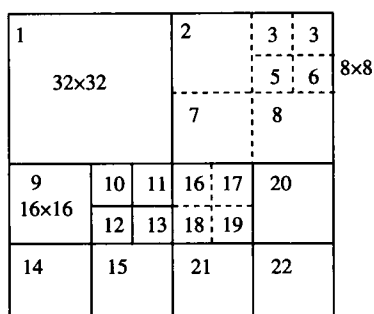
(1) 编码树单元。类似于H.264/AVC中的宏块(MB),HEVC将编码帧分为若干编码树块(Coding Tree Blocks,CTB),是进行预测、变换、量化和熵编码等处理的基本单元,其尺寸或者所包含的像素数可以是 16×16 、 32×32 或 64×64 。同一位置的亮度CTB和两块色度CTB,再加上相应的语法元素(Syntax Elements)形成一编码树单元(Coding Tree Units,CTU)。由于目前HEVC只支持4:2:0的彩色亚采样,所以色度块CTB的像素数为同等亮度块的1/4,即水平和垂直方向都是亮度块的一半。

(2) 编码单元。CTU又可以按照四叉树(quadtree)结构分解为若干方形编码单元(Coding Units,CU),同一层次的CU必须是同一尺寸的4个方块,最多可有4层分解,即 64×64 、 32×32 、 16×16 和 8×8 。如果不分解,则这个CTU仅包含一个CU。每个CU包含一块亮度编码块(Coding Blocks,CB)、两个色度CB以及相应的语法元素,可见CU的“树根”在CTU。CU是决定进行帧内预测还是帧间预测的单元,也就是说,整个CU只能是一种预测模式,不是帧内就是帧间。

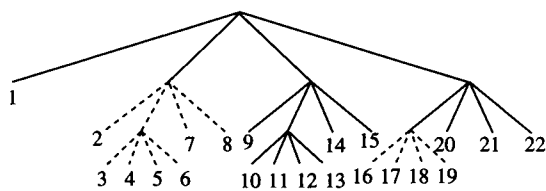
CU还可以按照四叉树层次分解(或不分解)为更小的预测单元(Prediction Units,PU)和变换单元(TB,Transform Units)。自然,每个PU包含亮度、色度预测块(Prediction Blocks,PB)和相应的语法元

素, PB 的尺寸可从 4×4 直至 64×64 。每个 TU 包含亮度、色度变换块(Transform Blocks, TB) 和相应的语法元素, TB 的尺寸从 4×4 到 32×32 。显然, TB 和 PB 在几何位置上有可能是重合的。

图 4 所示为某一 CTB 的划分为 CB 及其对应的二叉树结构的示例, 其中实线表示 CB 的界线, 虚线表示 TB 的界线。CB 的大小和图像的特性是自适应的, 在图像比较平缓区域, 选择比较大的 CB, 而在图像边缘或纹理复杂的区域, 选择比较小的 CU, 有利于提高编码效率。



(a) 一个 CTU 的划分



(b) 对应的二叉树结构

图 4 64×64 CTU 的二叉树划分示例

(3) 预测单元。预测单元 PU 是进行预测运算的基本单元, 包括帧内和帧间预测两类。一个编码单元 CU 可以包含一个或者多个预测单元, CU 到 PU 仅允许一层划分, 最小的 PU 为 4×4 。划分可以是对称的, 也可以是不对称的。如图 5 所示, 一个 $2N \times 2N$ (N 可以是 4、8、16、32) 的 CU 可划分为 8 种包含 PU 方式。帧间预测时可以在这 8 种方式中任意选择, 在帧间预测的跳过(skip)模式中只允许选择 $2N \times 2N$ 这种模式, 帧内预测时只允许选择 $2N \times 2N$ 或 $N \times N$ 方式。

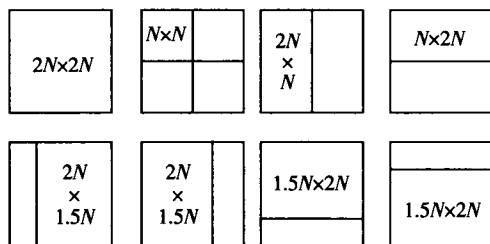


图 5 $2N \times 2N$ 的 CU 的 8 种 PU 划分方式

(4) 变换单元。变换单元 TU 是进行变换和量化的基本单元, 是在 CU 的基础上划分的, 但受到所

在的预测单元 PU 的限制。如 PU 是正方形, 则 TU 也必须是正方形, 尺寸为 4×4 、 8×8 、 16×16 、 32×32 。帧内编码模式中变换单元的尺寸需小于或者等于预测单元, 而帧间编码模式中变换单元的可以大于预测单元, 但是不能超过编码单元 CU。一个编码单元中多个变换单元也是按照二叉树的结构排列, 即每向下一层划分为 4 个小的正方形。如 PU 为非正方形时, TU 也需非正方形, 其大小为 32×8 、 8×32 、 16×4 、 4×16 , 可用于亮度分量, 其中只有 32×8 、 8×32 可用于色度分量。

1.4 图像格式

(1) 取样结构。HEVC 原则上支持 ITU601^[10] 建议规定的 $4:4:4$ 、 $4:2:2$ 和 $4:2:0$ 三种彩色图像取样结构。 $4:4:4$ 取样结构中两个色差信号(R-Y 和 B-Y)的取样频率和亮度信号(Y)取样频率一致, 几何位置重合, 这样亮度信号与两个色差信号的取样点数之比为 $4:4:4$ 。 $4:2:2$ 取样结构中两个色差信号的取样频率均为亮度信号取样频率的一半, 这样亮度信号与两个色差信号的取样点数之比为 $4:2:2$ 。 $4:2:0$ 取样结构中两个色差信号的取样频率也为亮度信号取样频率的一半, 但每一行只传送两种色差信号中的一种, 两种色差信号每行交替传送。3 种方式中, 比较常用的为 $4:2:0$ 的格式, 也是 HEVC 目前支持的主要格式。

(2) 比特深度。像素值的最大比特数表示图像灰度分辨率, 又称为比特深度(bit depth), 其值通常为 8 比特, 即灰度误差不会超过 $1/256$ 。这对于人眼的观察已经足够了, 但是对于编码处理而言, 有时显得精度不够, 往往会带来的计算误差和累积误差。为此, HEVC 提供了一种对编码器内外有别的比特深度表示方式, 即内部比特深度增加 (IBDI, Internal Bit Depth Increase) 方式。这种方式允许在编码器内采用高于 8 比特的像素深度 (如 10 比特) 进行编码运算, 而编码器外部, 即编码器的输出图像仍然是 8 比特深度。

(3) 隔行处理。随着数字视频技术的迅速发展, 视频的隔行扫描(interlaced scanning)方式日渐式微, 因此不同于 H. 264/AVC, HEVC 没有提供专用于隔行视频的工具, 而是将隔行视频的一帧看作两个独立的场, 对各个场数据分别进行编码, 简化了编码器的实现。

1.5 档次、等级和水平

为了提供应用的灵活性, HEVC 设置了编码的不同的档次(profile)、等级(tier)和水平(level)。档

次规定了一套用于产生不同用途码流的编码工具或算法,共有3个主档次,即常规8比特像素精度的Main,支持10比特精度的Main 10和支持静止图像的Main Still Picture。Main档次的主要技术指标包括:(1)像素的比特深度限制为8比特;(2)色度亚采样限制为4:2:0格式;(3)解码图像的缓存容量限制为6幅图像,即该档次的最大图像缓存容量;(4)允许选择波前(wavefront)和片(tiles)划分方式,但不是必需的,如果使用,在该级别必须有所限制:片的尺寸至少高为64像素,宽为256像素。

HEVC设置了1,2,2.1,3,3.1,4,4.1,5,5.1,5.2,6,6.1,6.2等13个水平,一个水平实际上就是一套对编码比特流的一系列编码参数的限制,如支持4:2:0格式视频,定义的图像分辨率从176×144(QCIF)到7680×4320(8K×4K),限定最大输出码率等。如果说一个解码器具备解某一水平码流的能力,则意味着该解码器具有解码这一水平以及低于这一水平所有码流的能力。

对同一水平,按照最大码率和缓存容量要求的不同,HEVC设置了两档等级(Tier),定义为高等级(High Tier)和主等级(Main Tier)。主等级可用于大多数场合,涵盖13个水平,要求码率较低;高等级可用于特殊要求或苛刻要求的场合,包括4和4以上的8个水平,允许码率较高,在同一水平大约高3~4倍。

1.6 语法和语义(Syntax and Semantics)

为了和现已广泛使用的H.264/AVC编码器尽量兼容,HEVC编码器也使用H.264/AVC的NAL单元语法。为了传输全局参数(如视频序列的分辨率、彩色格式、最大参考帧数、起始QP值等),采用H.264/AVC的SPS(Sequence Parameter Set)和PPS(Picture Parameter Set)语法和语义。HEVC的条(slice)的头信息的语法和语义同H.264/AVC的语法和语义非常接近,只是增加了一些必要的新的编码工具。

2 帧内预测

2.1 预测模式

HEVC的帧内预测技术^[11]的原理也和H.264/AVC类似,采用基于块的多方向帧内预测方式来消除图像的空间相关性,但是比H.264/AVC预测方向更细、更灵活。HEVC为亮度信息定义了33种不同的帧内预测方向,连同平面(Planar)和直流(DC)

模式,总共35种帧内预测模式。DC模式和H.264/AVC类似,平面模式的预测值由水平和垂直两个方向上已知像素的线性内插产生,特别适用于图像的平坦区域。

图6显示了HEVC亮度预测方向的定义,35种预测模式都有相应的编号。预测方向并没有用几何角度来表示,而是用像素的个数或格数来表示的。帧内预测单元(PU)的尺寸从4×4到32×32,对于4×4单元,采用0~16及34共18种模式;对于8×8、16×16和32×32单元,采用所有的35种模式。

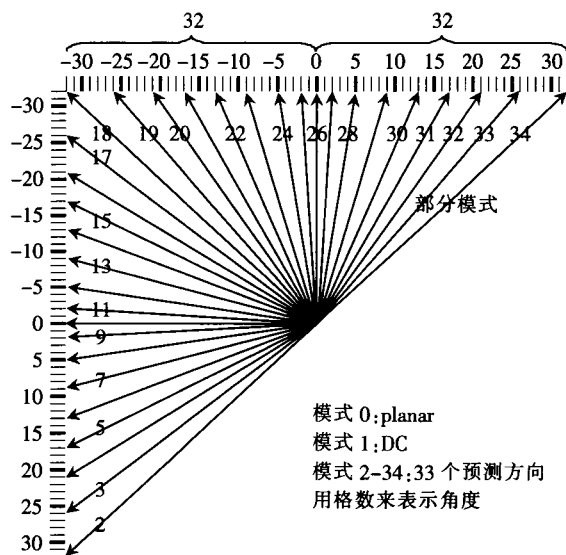


图6 帧内预测模式和角度

(1) 平面(Planar)模式。Planar预测模式跟DC模式一样,适合纹理比较平滑的区域。DC模式对当前块的所有像素使用同一个预测值,即它们的平均值。而Planar模式对当前块中每个像素使用不同的预测值,如图7所示。对当前块中的每一个像素,都使用水平和垂直两个方向的线性插值的平均作为当前像素的预测值,用式(1)表示,(x,y)点的亮度预测值为:

$$p(x,y) = \left[\frac{(N-1-x) \cdot b + (x+1) \cdot a + (N-1-y) \cdot d + (y+1) \cdot c + N}{2N} \right] \quad (1)$$

其中,N为PU的宽度, $p(0,y)=b$ 表示当前块左边、且与当前要预测的像素在同一行的参考像素值, $p(x,0)=d$ 表示当前块上方、且与当前要预测的像素同一列的参考像素值; $p(x,N)=c$ 表示和 $p(x,y)$ 在同一列的最下面的一个像素值,其值等于预测块左边参考列最下面的一个像素值; $p(N,y)=a$ 表示和 $p(x,y)$ 在同一行最左边的一个像素值,其值等于预测块上边参考行最右边的一个像素。从公式可以

看出,经 Planar 模式预测后的预测像素将呈现出从左上方到右下方渐变的趋势。

由于在彩色视频中同一位置的色度信号和相应的亮度信号的特性大多类似,因此在视频处理中,常常对色度信号采用和相应的亮度信号相同的方法。HEVC 也不例外,在帧内编码中色度信号也采用相应亮度信号所采用的预测模式。这样,对于色度信息,HEVC 有 5 种帧内预测模式:Planar 模式(0)、垂直模式(1)、水平模式(2)、DC 模式(3)和导出(derived)模式(4),即相应的亮度模式。

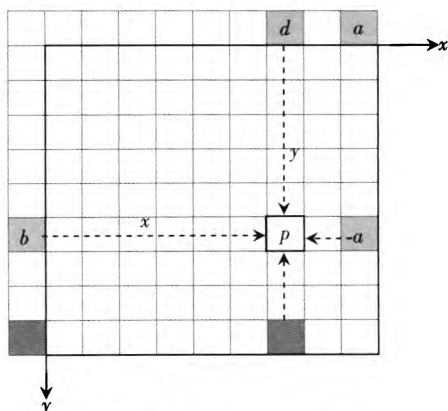


图 7 Planar 模式示意图

2.2 平滑预处理

为了提高帧内预测的效率,HEVC 对 8×8 或更大的 PU 进行平滑滤波预处理。平滑滤波器实际上是一个简单的一维有限冲激响应低通滤波器,滤波系数为 $(1/4, 1/2, 1/4)$,用它对处于预测方向上的参考像素进行滤波。对参考像素适当的平滑处理,可以减少噪声对预测的影响,提高了预测的精度和效率,尤其是大的 PU 其效果更好。

3 变换和量化

3.1 离散余弦和正弦变换

HEVC 的变换运算^[12]和 H. 264 类似,也是一种对预测残差进行近似 DCT 的整数变换。HEVC 在一个编码单元(CU)内进行变换运算时,可以将 CU 按照编码树层次细分,从 32×32 直至 4×4 的小块。例如一个 16×16 的 CU 可以用一个 16×16 的变换单元(TU)进行变换,或者 4 个 8×8 的 TU 进行变换。其中任意一个 8×8 的 TU 还可以进一步分为 4 个 4×4 的 TU 进行变换。变换运算的顺序和 H. 264/AVC 不同,变换时首先进行列运算,然后再进行行运算。HEVC 的整数变换的基矢量具有相同

的能量,不需要对它们进行调整或补偿,而且对 DCT 的近似性要比 H. 264/AVC 好。

HEVC 还支持 4×4 的离散正弦变换(Discrete Sine Transform, DST),仅将它用于对 4×4 块的帧内预测残差的编码。在帧内预测块中,那些接近预测参考像素的像素,如左上边界的像素将获得比那些远离参考像素的像素预测得更精确,预测误差较小,而远离边界的像素预测残差则比较大。DST 对编码这一类的残差效果比较好。这是因为不同 DST 基函数在起始处很小,往后逐步增大,和块内预测残差变化的趋势比较吻合,而 DCT 基函数在起始处大,往后逐步衰减。

3.2 率失真优化的量化

HEVC 的量化机理和 H. 264/AVC 基本相同,是在整数 DCT 变换时一并完成的。量化处理实际上就是用量化步长除以变换系数,得到变换系数更简单的表示。

量化是压缩编码产生失真的主要根源,因此选择恰当的量化步长,使失真和码率之间达到最好的平衡就成了量化环节的关键问题。HEVC 中的量化步长是由量化参数(QP)的标记的,共有 52 个等级(0 ~ 51),每一个 QP 对应一个实际的量化步长。QP 的值越大表示量化越粗,将产生的码率越低,当然带来的失真也会越大。HEVC 采用了率失真优化的量化(Rate Distortion Optimized Quantization, RDOQ)技术,在给定码率的情况下选择最优的量化参数使重建图像的失真最小。

量化操作是在变换单元 TU 中分别对亮度和色度分量进行的。在 TU 中所有的变换系数都是按照一个特定的量化参数 QP 统一进行量化和反量化的。HEVC 的 RDOQ 可比 H. 264/AVC 提高编码效率 5% 左右(亮度),当然带来的负面影响是计算复杂度的增加。

3.3 HEVC 的扫描方法

在 HEVC 中,将 TU 中量化后的系数通过特定的扫描方式形成一维数据,然后对它进行熵编码,系数的扫描是以 4×4 的“系数区”(Coefficient Region)为单位进行的,如图 8(a)所示。一个大的 TU 可分为若干个 4×4 的系数组。如果 TU 是帧内预测模式,则可使用水平、垂直和对角 3 种方式扫描。扫描方式的选择和帧内预测的方向有关,如果是接近水平方向的预测,则采用垂直扫描方式;如果是接近垂直方向的预测,则采用水平扫描方式;其它方向都采用对角方式。

如果 TU 是帧间预测模式,则只使用对角扫描方式。在帧间预测块中,由于可以使用矩形的变换,扫描方式要随之改变,如图 8(b)所示。在对角扫描方式中,若干系数组也是按照某种顺序扫描的,但必须在先前系数组的所有系数被扫描以后才可扫描当前系数组。

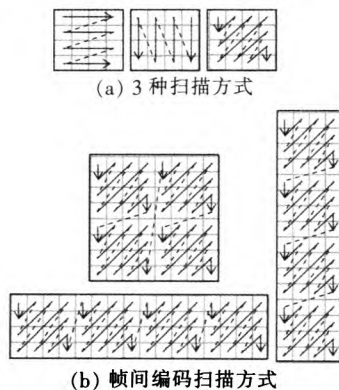


图 8 TU 的扫描方式

4 熵编码

4.1 自适应算术编码

输入视频在变换和量化以后,还需对量化、扫描后的变换系数进行熵编码(Entropy Coding)^[13],以获得进一步的信息压缩。在熵编码中常见的有较为简单的变长编码(如 Huffman 编码)和效率较高的算术编码两大类。如果将编码方式和编码的内容联系起来,则可获得更高的编码效率,这就是常见的基于上下文的变长编码(Context-Adaptive Variable Length Coding, CAVLC)和基于上下文自适应二进制算术编码(Context-Adaptive Binary Arithmetic Coding, CABAC)。这两类熵编码都是高效、无损的熵编码方法,尤其是在高码率的情况下更是如此,此时量化参数(QP)比较小,码流中变换系数占绝大部分。当然其计算量也较之常规的变长编码、算术编码要高。HEVC 只采用了 H. 264/AVC 非常类似的 CABAC 算法进行熵编码。

4.2 波前并行处理

考虑到高清、超高清视频编码的巨大运算量,在 HEVC 提供了基于条和基于片的便于并行编码和解码处理的机制^[14-15]。然而,这样又会引起编码性能的降低,因为这些条和片是独立预测的,打破了穿越边界的预测相关性,每个条或片的用于熵编码的统计必须从头开始。为了避免这个问题,HEVC 提出了一种波前并行处理(Wavefront

Parallel Processing, WPP)的熵编码技术,在熵编码时不需要打破预测的连贯性,可尽可能多地利用上下文信息。

在编码的波前并行处理时,一旦当前 CTU 行上边的两块 CTU 已经处理完了,立即用一个新的并行进程(通常是一个新的线程)开始处理这新的一行 CTU,其过程如图 9 所示。之所以需要两块 CTU,是因为帧内预测和运动矢量预测需要依赖于来之当前 CTU 上边和右上边的 CTU 的数据。WPP 熵编码初始化的参数可来自于从当前行上边的两块完全编码的 CTU 获得的信息,这样就可新的编码线程中使用尽可能多的上下文信息。

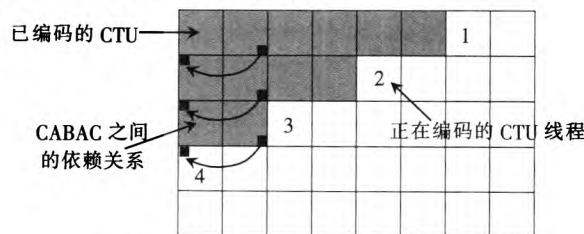


图 9 4 个线程的 WPP 处理示意图

5 帧间预测

帧间预测技术利用图像时域中邻近帧之间的相关性,使用先前已编码重建帧作为参考帧,通过运动估计和运动补偿对当前图像进行预测,从而除去视频信息中的时间冗余。HEVC 的帧间预测编码^[16]总体上和 H. 264/AVC 相似,但进行了如下几点改进,现简述如下。

5.1 可变 PU 尺寸的运动补偿

如前所述,每个 CTU 都可以递归地分解为更小的方形 CU,这些帧间编码的小 CU 还可划分一次成为更小的预测单元 PU。CU 可以使用对称的和非对称运动划分(Asymmetric Motion Partitions, AMP),允许非对称地划分 64×64 到 16×16 的 CU 为更小的 PU,PU 可以是方形的,也可以是矩形的。如前面图 5 所示,左上角的一个 $2N \times 2N$ 的 CU 可以划分为右边的 7 种情况。每个采用帧间预测方式编码的 PU 都有一套运动参数(Motion Parameters),包括运动矢量、参考帧索引和参考表标志。因为 AMP 允许 PU 在运动估计和运动补偿中更精确地符合图像中运动目标的形状,而不需要进一步的细分来解决,因此可以提高编码效率。

5.2 运动估计的精度

(1) 亮度样点内插。和 H. 264/AVC 类似,

HEVC 亮度分量的运动估计精度为 1/4 像素。为了获得非整数样点的数值,不同的非整数位置亮度的插值滤波器的系数是不同的,分别采用一维 8 抽头和 7 抽头的内插滤波器产生 1/2 和 1/4 像素的亮度值。用内插点周围的整像素样点值产生小数样点值的情况可见图 10。

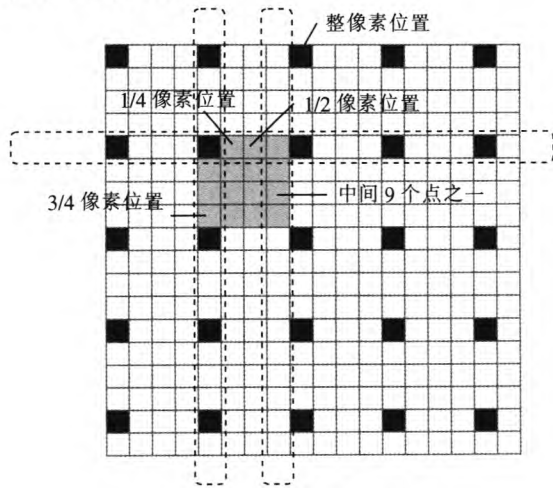


图 10 子像素内插方式

和整像素点在同一水平线上的水平方向插值滤波器系数为:

1/4 内插点	-1	+4	-10	+58	+17	-5	+1	
2/4 内插点	-1	+4	-11	+40	+40	-11	+4	-1
3/4 内插点	+1	-5	+17	+58	-10	+4	-1	

和整像素点在同一垂直线上的内插点用垂直方向的一维滤波器,系数和水平方向一样。处于中间的 9 个点沿用上述的垂直方向 8 抽头、7 抽头滤波器,利用刚才内插出来的非整数像素值,滤波器系数值仍然和前面一样。

(2) 色度样点内插。对于一般的 4:2:0 数字视频,色度整数样点的距离比亮度大一倍,要达到和亮度同样的插值密度,其插值精度需为 1/8 色度像素。色度分量的预测值由一维 4 抽头内插滤波器用类似亮度的方法得到。

和整像素点在同一水平方向的内插点用水平方向的 4 抽头滤波器系数为:

1/8 内插点	-2 + 58 + 10 - 2
2/8 内插点	-4 + 54 + 16 - 2
3/8 内插点	-6 + 46 + 28 - 4
4/8 内插点	-4 + 36 + 36 - 4
7/8 内插点	-2 + 10 + 58 - 2
6/8 内插点	-2 + 16 + 54 - 4
5/8 内插点	-4 + 28 + 46 - 6

和整像素点在同一垂直方向的内插点用垂直方

向的 4 抽头滤波器,系数和水平方向一样。处于中间的 7×7=49 个点只用垂直方向 4 抽头滤波器,利用刚才内插出来的非整数像素值,滤波器系数值仍然和前面一样。

5.3 运动参数的编码模式

HEVC 对运动参数的编码有 3 种模式,简称之为 Inter 模式、Skip 模式和 Merge 模式。前两种模式和 H.264/AVC 类似,Merge 模式是 HEVC 新引入的一种“运动合并”(Motion Merge)技术。这样,对于每个帧间编码的 PU,可以在下列 3 种编码模式中选择:(1) 采用对运动参数直接编码(如对运动矢量进行预测编码)的 Inter 模式,(2) 采用运动合并技术的 Merge 模式,(3) 采用改进的 Skip 模式。

Inter 模式需要传送当前编码 PU 的运动矢量(MV)来完成基于运动补偿的帧间预测。在 HEVC 中,运动矢量可采取空间或时间预测编码的方式(H.264/AVC 仅采用空间预测)^[17-18]。而 Skip 模式和 Merge 模式都无需传送 MV 信息,只传送候选 PU 块的索引信息,由解码端采用运动推理(Motion Inference)方法(如拷贝)来获得运动信息,推理依据来自空域邻近块或时域邻近块(候选块)的 MV。在 Skip 模式中连残差信息也省略了。

(1) Merge 模式。在以往标准的帧间预测中,对以每个帧间预测块都要传输一套运动参数。为了进一步提高编码效率,可对邻近的 PU 块进行合并(Merge)处理,形成 Merge 帧间预测编码模式,即将毗邻的运动参数相同或相近的几个 PU 块合并起来形成一个所谓的“小区”(Region),只要为每个小区传输一次运动参数,不必为每个 PU 块分别传输运动参数。

在运动合并模式中,HEVC 为当前正在编码的 PU 创建一个先前已编码的邻近 PU(称为“候选”PU)列表。其中候选 PU 既可以是在编 PU 的空间邻近,也可以是时间邻近。编码器按照最小率失真(RD)准则选取并标注这个运动合并列表中哪一个候选 PU 被选中,继而使得当前编码的 PU 在解码时只要拷贝中选 PU 的运动参数。显然,运动合并方式不需要为当前 PU 去编码运动矢量,取而代之的是只需要编码运动合并列表中候选者的简单索引信息。

在 Merge 模式中,编码的 PU 在运动矢量预测(Motion Vector Prediction, MVP)候选集中选择运动矢量预测块。MVP 候选集包括 5 个空域 MVP 和 2 个时域 MVP,如图 11 所示。5 个空域 MVP 分别为

当前 PU 左下角的预测块 A_0 和 A_1 , 左上角的预测块 B_2 , 右上角的块 B_0 和 B_1 ; 2 个时域 MVP 分别是先前已编码帧中和 PU 相同位置的预测块 T_1 和右下角的 T_0 。

如果某一块编码 PU 为 Merge 模式, 则需要标注一个 MVP 索引来指定在 MVP 候选集中哪一个 MVP 被用作这个将要被合并 (Merged) 的块。为了遵循运动信息共享的原则, 每个合并的 PU 都可以再使用这个所选 MVP 的 MV、预测方向和参考图像索引。需要注意, 如果所选的 MVP 是时域的 MVP, 参考图像索引总是指向第一幅参考图像。

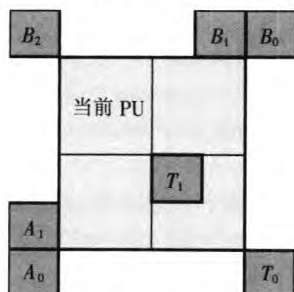


图 11 Merge 模式的候选 MVP 集

HEVC 的编码器采用率失真优化 (Rate-distortion optimization, RDO) 的决策机制, 为 Merge 模式在给定的候选集中选择最终的 MVP, 并将所选择的 MVP 的索引 (Index) 传送到解码器。Merge 模式可以比 H. 264/AVC 的直接模式 (Direct Mode) 减少 3% 左右的码率。

实际上, Merge 模式的这种 MVP 候选集的设置和选择方法在 HEVC 中称之为高级运动矢量预测 (Advanced Motion Vector Prediction, AMVP), 它也同样适用于 Skip 模式和 Inter 模式的最佳 MVP 块的选择。

(2) Skip 模式。在 HEVC 新的 Skip 模式中, 编码器只需要编码运动合并候选者的索引信息, 以保证解码时 Skip PU 可以拷贝中选者的运动参数作为自己的运动参数。这样, 对那些帧间变化很少, 或者运动静止的区域编码只需很少的比特。

6 环路滤波

环路滤波 (Loop Filtering) 位于编码器预测环路中的反量化/反变换单元之后、重建的运动补偿预测参考帧之前。因而, 环路滤波是预测环路的一部分, 属于环内处理, 而不是环外的后处理。环路滤波的目标就是消除编码过程中预测、变换和量化等环节

引入的失真。由于滤波是在预测环路内进行的, 减少了失真, 存储后为运动补偿预测提供了较高质量的参考帧。HEVC 除了采用和 H. 264/AVC 类似环内去方块滤波 (DeBlocking Filter, DBF) 外, 还增加了新的样值自适应偏移 (Sample Adaptive Offset, SAO) 环内滤波工具, 形成了 2 个按顺序进行的处理: 去方块滤波和 SAO 滤波^[19]。

6.1 去方块滤波

方块效应是由于采用图像分块压缩方法所形成的一种图像失真, 尤其在块的边界处更为惹眼。为了消除这类失真, 提高重建视频的主观和客观质量, H. 264/AVC 在方块的边界按照“边界强度”进行自适应低通滤波, 又称去方块滤波。HEVC 也使用了类似的环内去方块滤波来减轻各种单元边界 (如 CU、PU、TU 等) 的块效应。HEVC 为了减少复杂性, 利于简化硬件设计和并行处理, 不对 4×4 的块边界滤波, 且仅定义了 3 个边界强度等级 (0、1 和 2), 滤波边界附近的像素, 省却了对非边界处像素的处理。在滤波前, 对于每一个边界需要判决: 是否需要去方块滤波? 如果需要, 还要判决到底是进行强滤波还是弱滤波。判决是根据穿越边界像素的梯度值以及由此块的量化参数 QP 导出的门限值共同决定的。HEVC 的去方块滤波对需要进行滤波的各类边界统一进行, 先对整个图像的所有垂直边界进行水平方向滤波, 然后再对所有的水平边界进行垂直方向滤波。

6.2 样值自适应偏移

样值自适应偏移 (SAO) 是 HEVC 的一项新的编码工具, 对要重建的图像逐像素进行滤波。它先按照像素的灰度值或边缘的性质, 将像素分为不同的类型, 然后按照不同的类型为每个像素值加上一个简单的偏移值, 达到减少失真的目的, 从而提高压缩率, 减少码流。采用 SAO 后, 平均可以减少 2% ~ 6% 的码流, 而编解码器的复杂度仅增加约 2%。加偏移值有两种模式, 带偏移 (Band Offset, BO) 和边缘偏移 (Edge Offset, EO)。编码器对图像的不同区域选择施加 BO 模式或 EO 模式的偏移, 并在码流中给出相应的标识。

BO 模式将像素值从 0 到最大值分为 32 个相等的间隔——“带 (bands)”, 例如, 对 8 比特而言, 最大值是 255, 则此时的带的宽度就是 $256/32 = 8$, 每个带所包含的像素值都比较相近。如果某 CTB 的亮度值分布在 4 个相邻的带中间, 说明这原本是一个比较平坦的图像区域, 这样的区域人为失真容易

出现,则需对这些像素值施加偏移量(可正可负),使像素值的分布趋向更集中。当然这个偏移量也要作为带偏移传输到解码端。

EO 模式使用一种 3 像素结构来对所处理的像素进行分类。在该像素所处边缘(一维)的方向上、以该像素为中心再加上最近邻的 2 个像素,形成 45°、垂直、水平和 135° 四种“3 像素结构”,如图 12 所示。按照 3 像素值的不同分布,该像素可被分为峰值像素(如果其值大于相邻的 2 个像素),谷值像素(如果其值小于相邻的 2 个像素),边缘像素(如果其值等于任何一个相邻像素)和其它像素(如果不符合前 3 种),共 4 类。对 4 类不同的像素所加的偏移值也不同。



图 12 3 像素模式示意图

7 结束语

迄今为止的测试表明,在 HEVC 的多项关键技术支持下,其性能已经达到和部分超过当初设立的目标^[20]。图 13 为 Park Scene 序列编码的主观测试一例,在 JM 编码 1 Mbit/s 编码速率以上,相同质量的 HEVC 码率仅为 JM 的一半。

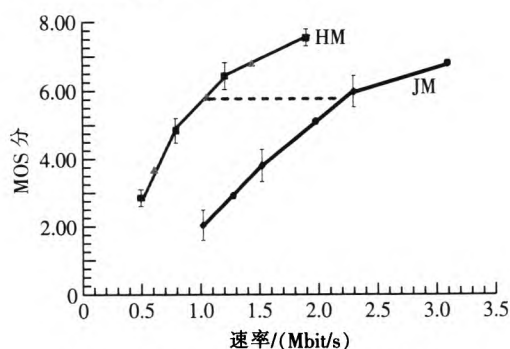


图 13 Park Scene 序列编码性能对比

HEVC 的完成是一个渐近的过程,如 2012 年 7 月的草案仅包含一个档次——主档次(main profile),其它档次和级别尚待开发完成。类似于 H.264/AVC 为可分级视频编码(SVC)和多视点视频编码(MVC)扩展定义那样,HEVC 也包含了为今后 3D、SVC 和 MVC 的扩展提供了准备。

HEVC 标准覆盖了相当广泛的数字视频应用范围,如家庭影院、数字电影、视频监控、广播电视、实

时通信、网络视频、视频会议、移动流媒体、高清存储、VOD、远程呈现(telepresence)、医疗图像、遥感图像,等等。将来还可用于 3D 视频、多视点视频、可分级视频等。可以预计,HEVC 的正式颁布,将给视频应用带来不可估量的影响。

参考文献:

- [1] Video Codec for Audiovisual Services at px64 kbit/s[S]. ITU-T Rec. H.261, version 1; Nov. 1990, version 2; 1993.
- [2] Advanced Video Coding for Generic Audio-Visual Services[S]. ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC), ITU-T and ISO/IEC JTC 1, 2003 (and subsequent editions).
- [3] BROSS B, HAN W J, SULLIVAN G J, et al. High Efficiency Video Coding (HEVC) Text Specification Draft 9[S]. document JCTVC-K1003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), 2012.
- [4] Joint Call for Proposals on Video Compression Technology[S]. ITU-T SG16/Q6 document VCEG-AM91 and ISO/IEC MPEG document N11113, ITU-T and ISO/IEC JTC 1, 2010.
- [5] BROSS B, HAN W J, OHM J R, et al. High Efficiency Video Coding (HEVC) Text Specification Draft 8[S]. document JCTVC-K1003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), 2012.
- [6] SULLIVAN G J, OHM J R, HAN Woojin, et al. Overview of the High Efficiency Video Coding (HEVC) Standard[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2012, 22(12): 1649-1668.
- [7] POURAZAD M T, DOUTRE C, AZIMI M. HEVC: The New Gold Standard for Video Compression How Does HEVC Compare with H.264/AVC[J]. IEEE Consumer Electronics Magazine, 2012, 1(3): 36-46.
- [8] HANG Hsuehming, PENG Wenhsiao, CHAN Chiahshin, et al. Towards the Next Video Standard: High Efficiency Video Coding. Proceedings of the Second APSIP[C]. // Biopolis Singapore. 2010: 609-618.
- [9] KIM Ilkoo, MIM Junghye, LEE Tammy, et al. Block Partitioning Structure in the HEVC Standard[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2012, 22(12): 1697-1706.
- [10] Studio encoding parameters of digital television for 4:3 and wide-screen 16:9 aspect ratios[S]. Recommendation ITU-R BT. 601, version 6, 2007.
- [11] LAINEMA J, BOSSEN F, HAN Woojin, et al. Intra Coding of the HEVC Standard[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2012, 22(12): 1792-1801.
- [12] SOLE J, JOSHI R, NGUYEN N, et al. Transform Coefficient Coding in HEVC[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2012, 22(12): 1765-1777.
- [13] SZE V, BUDAGAVI M. High Throughput CABAC Entropy Coding in HEVC[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2012, 22(12): 1778-1791.

- [14] CHI C C, ALVAREZ-MESA M, JUURLINK B, et al. Parallel Scalability and Efficiency of HEVC Parallelization Approaches [J]. IEEE Transaction on Circuits and Systems for Video Technology, 2012, 22(12): 1827 – 1838.
- [15] ALVAREZ-MESAL M, CHI C C, JUURLINK B. Parallel Video Decoding in the Emerging HEVC Standard [C] // IEEE International Conference on Acoustics, Speech and Signal Processing (IC-ASSP). Kyoto, Japan, 2012: 1545 – 1548.
- [16] LIN Jianliang, CHEN Yiwon, TSAI Yupao, et al. Motion Vector Coding Techniques for HEVC [C] // Multimedia Signal Processing (MMSP). 2011: 1 – 6.
- [17] HELLE P, OUDIN S, BROSS B. Block Merging for Quadtree-Based Partitioning in HEVC [J]. IEEE Transaction on Circuits and Systems for Video Technology, 2012, 22(12): 1720 – 1731.
- [18] JCT-VC. Subversion Repository for the HEVC Test Model Version HM8. 0 [EB/OL]. <https://hevc.hhi.fraunhofer.de/svn/svn-HEVCSoftware/tags/HM-8.0/>.
- [19] NORKIN A, FULDSETH A, NARROSKHE M. HEVC Deblocking Filter [J]. IEEE Transaction on Circuits and Systems for Video Technology, 2012, 22(12): 1746 – 1654.
- [20] OHM J R, SULLIVAN G J, SCHWARZ H. Comparison of the Coding Efficiency of Video Coding Standards-Including High Efficiency Video Coding (HEVC) [J]. IEEE Transaction on Circuits and Systems for Video Technology, 2012, 22(12): 1 – 15.

作者简介:



朱秀昌(1947 –),男,江苏丹徒人。南京邮电大学通信与信息工程学院教授,博士生导师。长期从事图像通信方面的科研和教学工作。1993年获国务院特殊津贴。曾先后获原邮电部科技进步一等奖和国家科技进步三等奖。

李欣(1981 –),女,安徽巢湖人。南京邮电大学通信与信息工程学院讲师,博士研究生。研究方向为图像处理与多媒体通信。

陈杰(1983 –),女,山东龙口人。南京邮电大学通信与信息工程学院博士研究生。研究方向为图像处理与多媒体通信。

(本文责任编辑:胡长贵)

声 明

为适应我国信息化建设的需要,扩大作者学术交流渠道,实现期刊编辑、出版工作的网络化,本刊已加入《中国学术期刊(光盘版)》、《中国期刊网》全文数据库、《万方数据——数字化期刊群》和《中文科技期刊数据库》,并已许可《中国学术期刊(光盘版)》电子杂志社在中国知网及其系列数据库产品中,以数字化方式复制、汇编、发行、信息网络传播本刊全文,作者著作权使用费随本刊稿酬一次性给付。如不同意将文章编入相关数据库,请在来稿时声明,本刊将做适当处理。

本刊编辑部