

高效视频编码

沈燕飞^{1),2)} 李锦涛¹⁾ 朱珍民¹⁾ 张勇东¹⁾

¹⁾(中国科学院计算技术研究所 北京 100190)

²⁾(中国科学院大学 北京 100049)

摘要 近年来随着视频编码技术的发展以及高清视频需求的快速增长,国际运动图像专家组 ISO-IEC/MPEG 和国际电信联盟的视频编码专家组 ITU-T/VCEG 于 2010 年成立了视频编码联合工作组 JCT-VC,开始针对高分辨率视频应用制定新一代视频编码标准,即高效视频编码 HEVC(High Efficiency Video Coding),2012 年底形成了国际标准草案,并开发了相应的测试模型 HM,进行性能评估和算法测试.该文将对标准制定过程中涉及到的关键技术进行详细介绍,主要包括:基于二叉树结构的编码块组织形式、先进的运动矢量预测和合并方法、多方向的帧内预测模式、基于 DCT 的分像素插值技术和内容自适应的算术编码技术等,最后对相关编码工具的编码效率、计算复杂度进行了分析,并且指出了未来的研究方向.

关键词 视频编码;高清视频;HEVC;H.264 视频编码;视频通信
中图法分类号 TP391 **DOI号** 10.3724/SP.J.1016.2013.02340

High Efficiency Video Coding

SHEN Yan-Fei^{1),2)} LI Jin-Tao¹⁾ ZHU Zhen-Min¹⁾ ZHANG Yong-Dong¹⁾

¹⁾(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

²⁾(University of Chinese Academy of Sciences, Beijing 100049)

Abstract In recent years, with the development of video coding technology and the increasing popularity of high definition (HD) video content, the ISO/IEC Moving Picture Experts Group (MPEG) and the ITU-T Video Coding Experts Group (VCEG) formed the joint collaborative team on video coding (JCV-VC) in 2010 which aims to develop the next-generation video coding standard for HD video application, called High Efficiency Video Coding (HEVC). So far the test model HM has been used for performance evaluation and algorithm test. The final international standard of HEVC has been completed and published in the end of 2012 based on the working plan of JCV-VC. In this paper, the key technologies about HEVC will be surveyed, specifically including the quad-tree structure of coding unit, prediction unit and transform unit, the advanced motion vector prediction and merging technology, angular intra prediction, DCT-based fractional pixel interpolation filter and context adaptive arithmetic coding etc. Finally the coding performance of some coding tools and complexity of HEVC are analyzed in detail.

Keywords video coding; high definition video; HEVC; H.264 video coding; video communication

1 引言

随着网络通信技术的不断发展和终端处理性能

的不断提升,多媒体业务的支持需求日益增长,据调查显示,目前网络视频数据量占据了互联网 90% 的流量,即使在移动互联网中,视频数据也占了 50% 以上.另外,由于人们对视频质量的需求也在不断提

收稿日期:2011-12-12;最终修改稿收到日期:2013-08-09. 本课题得到国家自然科学基金(61001123,61070109)和广东省产学研结合项目(2012B091100212)资助. 沈燕飞,男,1976年生,博士研究生,助理研究员,主要研究方向为视频图像编解码、视频处理多媒体通信等. E-mail: syf@ict.ac.cn. 李锦涛,男,1962年生,研究员,主要研究领域为多媒体技术、虚拟现实技术以及普适计算等. 朱珍民,男,1962年生,教授,主要研究领域为嵌入式技术、多媒体技术等. 张勇东,男,1973年生,研究员,主要研究领域为多媒体内容分析、数字信号处理和视频编解码技术等.

升,包括时间和空间分辨率的不断提升、从二维视频到立体视频的转换等,导致视频数据量的不断增长。因此对视频压缩编码技术的要求也越来越高,视频压缩编码技术也成为目前多媒体技术研究领域中最热门的技术之一。

近 20 年来,国际电联通信标准化组 ITU-T、国际标准化组织和国际电子委员会的移动图像专家组 ISO/IEC MPEG 以及中国的音视频编解码工作组 AVS 都成功制定了一系列视频通信标准,包括 H.261、H.263、H.264、MPEG-1、MPEG-2、MPEG-4 和 AVS 等等。其中 2003~2009 年制定并不断完善的 H.264/MPEG AVC 视频编码标准,由于高效的压缩效率以及网络适应性等方面的明显优势,逐步成为视频应用领域的主流标准。

但随着数字视频应用的快速发展,视频空间分辨率正在从标清(640×480 , 720×576)向高清($720P$ 、 $1080P$)全面升级,甚至在一些视频应用领域中出现了 $4K \times 2K$ 、 $8K \times 4K$ 的数字视频格式;视频时间分辨率,即帧率,也在由 30 帧/s 向 60 帧/s 转换;视频编码器芯片化的趋势也日益明显,并行化需求日益加剧;存储空间和网络带宽是视频应用中最为关键的资源等等。面对这些新的应用需求,传统的视频编解码技术出现了以下性能限制:

(1) 由于空间分辨率的增加,每个大小为 16×16 宏块所承载的信息量大大减少,而相邻块之间的低频系数相关性却大大增加,因此出现大量的块间冗余;另外,由于分辨率的增加,帧间编码宏块的运动矢量幅值也大大增加,其统计特征不再呈现集中化分布趋势,因此指数哥伦布编码或变长编码模式的编码效率将逐渐下降。

(2) 随着图像分辨率的增加,宏块个数也出现爆发式的增长,用于表示每个编码宏块的模型参数,包括预测模式、运动矢量、参考帧索引以及量化步长等辅助信息(Side Information)等所占数据量过多,而用于编码预测残差变换系数的数据明显减少。

(3) 传统视频编码技术大都采用串行编解码模式,从左至右、从上至下逐个宏块进行处理,特别对于 H.264/MPEG AVC 中的熵编码技术,即基于上下文的 CABAC 和 CAVLC 技术,以及运动矢量预测技术和环路滤波技术,宏块之间数据耦合性大,难以满足并行处理的要求,而对于 GPU/DSP/FPGA/ASIC 等并行化程度非常高的处理器来说,这种串行化处理模式已经成为制约运算性能提高的瓶颈之一。

(4) 传统视频编码技术中,帧内/帧间预测块和

预测残差变换块的大小相对固定,难以适应高清视频图像纹理特征的多样化分布特点,导致预测性能与预测残差系数编码效率均出现不同程度的下降。

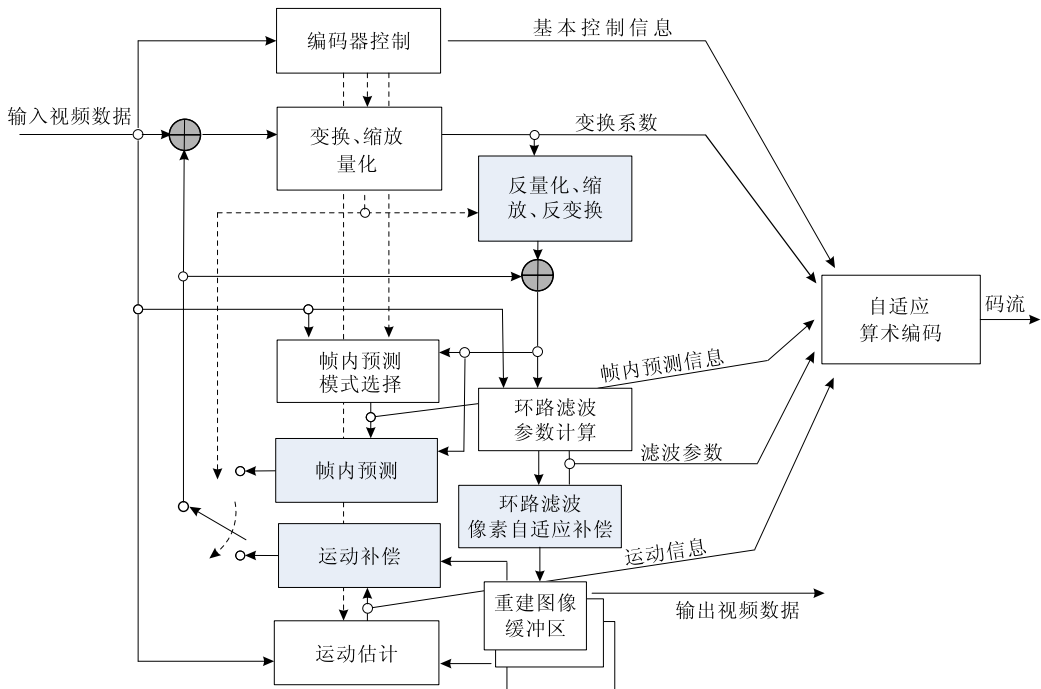
(5) 由于处理器性能的不不断提升和芯片设计规模的不断扩大,传统视频编解码的运算复杂度不再是视频应用的主要瓶颈,可以采用相对复杂的预测模式选择和率失真优化技术来提高压缩编码效率。

为了适应以上技术发展趋势及需求变化,国际运动图像专家组 ISO-IEC/MPEG 和国际电信联盟的视频编码专家组 ITU-T/VCEG 于 2010 年成立了视频编码联合工作组 JCT-VC^[1-2],在全球范围内征集视频编码技术提案^[3],在 H.264/AVC^[4]基础上开始制定新一代高效视频编码标准 HEVC^[5-6],核心目标是针对高分辨率视频图像,在 H.264/AVC 高级档次 HP(High Profile)基础上将压缩效率再提高一倍^[7-8],即在保证相同重建视频图像质量的前提下,视频编码率减少 50% 以上。JCT-VC 工作组共收到 27 个技术提案,并随后根据文献[1]中制定的测试方法做了技术评估和测试,遴选出最优的 5 个技术提案^[9-15],然后形成了相应的测试模型 TMuC(Test Model under Consideration)^[16],其中涉及到的共性关键技术文献[10-11]进行了描述。根据测试结果,2010 年 10 月形成了第 1 个测试模型 HM1,随后对其中的关键技术进行反复评估与验证,并于 2012 年 7 月发布了公开国际标准草案。2013 年 1 月份发布了最终国际标准草案,在 H.264/AVC 标准 2~4 倍的复杂度基础上,将压缩效率提升了 1 倍以上^[17-18]。

本文首先对 HEVC 编码框架及其中的关键技术进行介绍,然后详细描述了其中的高级编码工具,最后对编码性能和计算复杂度进行了测试评估并指出了其中的研究热点。

2 HEVC 编码框架

HEVC 仍然沿用了 H.261 就开始采用的基于运动补偿的混合编码框架^[6],即帧间和帧内预测用于消除时间域和空间域的相关性;对预测残差进行离散余弦变换和量化以消除空间相关性;自适应熵编码消除统计冗余;环路滤波用于消除量化噪声,但 HEVC 在图像编码单元组织、内容自适应算术编码、多方向帧内预测、先进运动矢量预测和合并、分像素运动估计与补偿、自适应像素补偿和环路滤波等方面都做了很大的改进,HEVC 编码框架如图 1 所示,图中阴影部分为编码和解码共用模块。

图 1 HEVC 编码框架^[6]

HEVC 的主要技术特征如下:

(1) 树形编码块结构

在 HEVC 中,基本编码单元为编码树单元 CTU(Coding Tree Unit),大小为 16×16 、 32×32 或者 64×64 ,每个编码树单元由亮度编码树块与对应的色度编码树块组成^[19-20].每个编码树块可以进一步分成编码块 CB(Coding Block),根据预测与变换的率失真性能,可以对编码块进一步的分成预测块 PB(Prediction Block)和变换块 TB(Transform Block).编码块、预测块和变换块的分离使视频编码更加灵活,各种不同的划分组合更能适应视频图像的纹理特征,有助于编码效率的提高.

(2) 四叉树残差变换块结构

残差四叉树变换(Residual Quad-tree Transform)属于一种自适应的变换技术,是对 H. 264/MPEG-AVC 中自适应块变换 ABT(Adaptive Block-size Transform)技术的延伸和扩展^[21].它允许变换块的大小能够根据预测残差的特征进行自适应的选择,在残差能量集中、图像细节保留以及图像三者之间最优均衡.

(3) 像素自适应补偿技术

像素自适应补偿 ASO(Adaptive Sample Offset)位于编解码环路内,通过对环路滤波后的重建图像进行分类,然后对每一类图像像素选择一种补偿方式,以减少源图像与重构图像之间的失真,补偿方式分为区间补偿和边缘补偿两类^[22].

(4) 并行设计技术

为了适应多核处理器的发展趋势,HEVC 中引入了片层(Slice)、条带(Slice)以及波形并行处理 WPP(Wavefront Parallel Processing)的思想^[6,18].其中片层方法与条带方法属于粗粒度的并行方法,片层方法将图像划分成非重叠的矩形区域,每个矩形区域独立的编码和解码;条带方法将图像分成多个条带分别编码和解码,每个条带可以进一步切分成可以独立编码和解码的熵条带(Entropy Slice),然而条带方法更多地应用于抗误码设计与满足最大传输单元 MTU(Maximum Transmission Unit)的需求.波形并行处理方法属于细粒度的并行方法,类似于流水线设计思想,在完成上一行中第二个编码单元编解码时,即可开始下一行第一个编码单元的编解码处理.

(5) 自适应熵编码技术

在 HEVC 中对所有的语法元素和预测残差变换系数都采用自适应算法编码技术 CABAC 进行熵编码.算法原理与 H. 264/AVC 中的 CABAC 一致,主要由二值化、模型选择和算术编码 3 个部分组成,但是模型数减少了很多,H. 264/AVC 中有 299 个模型,而 HEVC 中只有 154 个模型,另外为了提高 CABAC 的数据吞吐率,模型之间的相关性也进行了大规模简化^[23].

除此之外,HEVC 还针对特定应用和特定编码模式定义了一些编码工具,如应用于屏幕图像编码的无变换(Transform Skip)编码技术、应用于帧内

预测残差编码的离散正弦变换 DST 等等^[6].

HEVC 编码性能的提升是所有编码工具共同作

用的结果,并非某单项技术所为. HEVC 与 H. 264/AVC 中的关键技术对比如表 1 所示.

表 1 HEVC 与 H. 264/AVC 关键技术对比表

编码工具	H. 264/AVC	HEVC
编码单元大小	$4 \times 4 \sim 16 \times 16$	$4 \times 4 \sim 64 \times 64$
分像素插值滤波器	亮度分量半像素 $\{1, -5, 20, 20, -5, 1\}$ 亮度 $1/4$ 像素 $\{1, 1\}$	亮度分量半像素 $\{-1, 4, -11, 40, 40, -11, 4, 1\}$ 亮度 $1/4$ 像素 $\{-1, 4, -10, 58, 17, -5, 1\}$
运动矢量预测方法	空域中值预测	空域+时域预测 运动矢量合并技术
亮度帧内预测	4×4 块和 8×8 块, 9 种预测模式 16×16 块, 4 种预测模式	34 种方向预测 平面预测、DC 预测
色度帧内预测	四种预测模式	与亮度相同
变换	4×4 DCT 和 8×8 DCT	$4 \times 4/8 \times 8/16 \times 16/32 \times 32$ DCT 4×4 DST
环路滤波	4×4 和 8×8 块边缘滤波	4×4 边缘不进行环路滤波

针对不同的目标应用和码流兼容性测试, HEVC 还定义了 3 个档次 (Profile), 即 Main、Main10 和 Main 静态图像档次, 目前所有档次都只支持 $4:2:0$ 格式的视频数据编码, 除 Main10 支持 10 bit 颜色深度的视频数据外, 其它两种只支持 8 bit 深度. 根据编解码图像的分辨率以及编解码器的实际处理能力, HEVC 还定义了 13 个应用级 (Level), 支持的图像分辨率从 QCIF (176×144) 到 $8K \times 4K$, 最大码流达到 80 Mbps.

3 HEVC 中的高级编码工具

3.1 灵活的编码单元、预测单元和变换单元

在以前的视频标准中^[4, 24-28], 编码单元都是以 16×16 大小的宏块 MB (Macroblock) 作为最基本单

元, 不同的编码标准中对宏块的划分模式略有差别, 如 MPEG-2 中只能采用 16×16 或者 8×8 块进行帧间预测编码, 而 H. 264 中为了更好地适应不同边缘块形状, 可以对宏块进行更加复杂的划分, 如 $16 \times 8, 8 \times 16, 8 \times 8, 8 \times 4, 4 \times 8$ 和 4×4 等. 但是由于视频应用的变化, 高清视频越来越普及, 研究表明使用更大的编码单元能够进一步提高高清视频的编码效率^[29-32]. 因此, 在 HEVC 中采用了更加灵活的编码单元组织形式, 编码单元的大小从 8×8 到 64×64 , 并且采用递归的编码树 CTB (Coding Tree Block) 方式进行表示^[31]. 例如, 当编码单元大小为 64×64 和递归深度为 4 时, 编码单元划分形式如图 2(a) 所示, 组织形式如图 2(d) 所示. 如果编码单元的大小为 16×16 , 递归深度为 2 时, 则编码单元的组织形式就与 H. 264 一样.

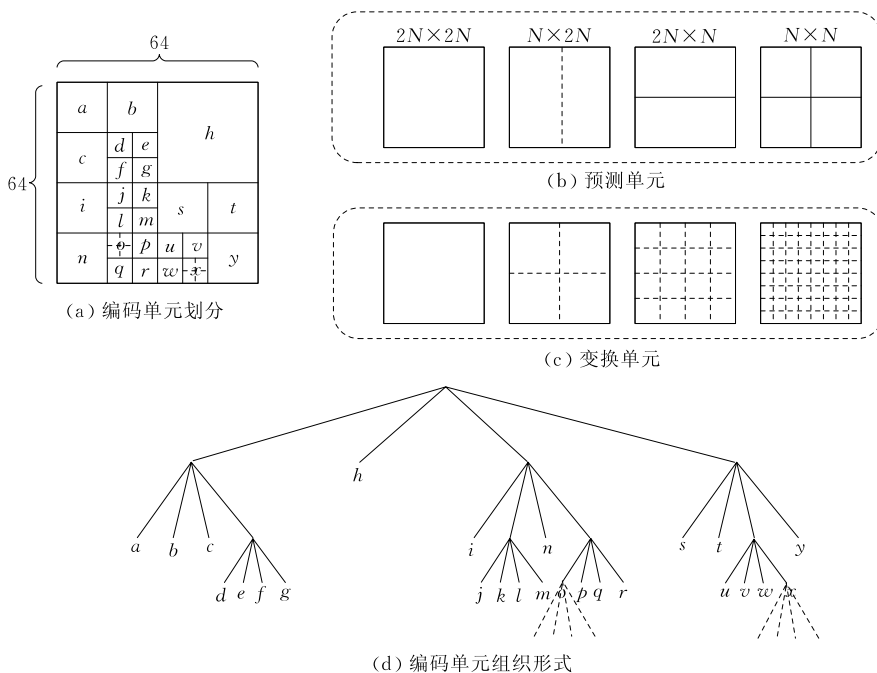


图 2 编码单元、预测单元和变换单元

每个编码单元的叶子节点可以划分成一个或多个预测单元,如图 2(a)中的叶子节点 h 可以采用 4 种不同形状的预测单元进行帧间预测编码,如图 2(b)所示,但帧内预测都是基于正方形块进行的.对预测残差的变换同样可以采用不同大小的块形状,划分模式如图 2(c)所示.每个变换单元可以包含多个预测单元,但不能超过一个编码单元的大小,变换单元的大小从 32×32 到 4×4 ,而一个预测单元也可以使用多个变换单元.

3.2 帧内预测

帧内预测算法是充分利用图像的空间相关性,用周围重建像素值对当前编码块进行预测,提高编码效率.

3.2.1 亮度分量帧内预测

在 HEVC 中共定义了 34 种帧内预测方向^[33-35],预测方向角度由当前预测单元 PU 的最下面一行或最左边一列像素与上面参考像素行或者左边参考像素列之间的位移确定,如图 3 所示,预测角度分别为 $\pm[0, 2, 5, 9, 13, 17, 21, 26, 32]/32$,这样能够更加精确地预测当前编码块的纹理特性.

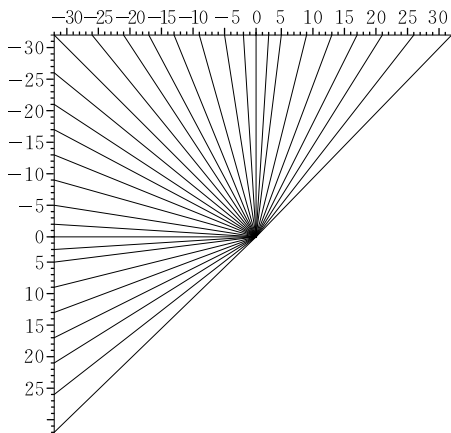


图 3 HEVC 中帧内预测方向

在使用参考像素对当前预测单元进行预测的过程中,首先需要根据预测方向和当前预测像素的坐标位置计算参考像素的位置,如果参考像素是整像素位置,则直接进行预测,否则要对参考像素进行插值处理,计算出分像素位置处的参考像素值.以 8×8 预测块和预测角度为“ -32 ”为例,如图 4 所示.假设当前预测像素的坐标位置为 $P(x, y)$,预测角度为 $PredAng$,则参考像素的位置坐标计算如下:

$$Idx = (256 \times 32 \times (P + 1) / PredAng) \gg 8 \quad (1)$$

$$Ifact = (256 \times 32 \times (P + 1) / PredAng) \% 256 \quad (2)$$

其中, Idx 和 $Ifact$ 分别为参考像素的整像素和分像素坐标.为了减少除法运算,可以将 $256 \times 32 / PredAng$ 做成表格 $[4096, 1638, 910, 630, 482, 390, 315, 256]$ 进行查表处理.

由于不同位置的预测像素需要逐个判断使用水平参考像素还是垂直参考像素,为了减少计算复杂度,可以使用投影的方法,在图 4 中将垂直参考像素投影成水平方向,如图 5 所示.

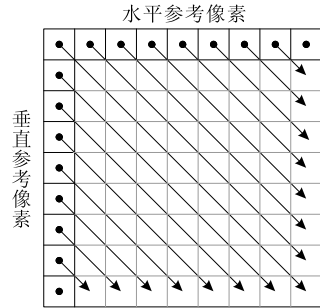


图 4 预测角度为 -32 的帧内预测示意图

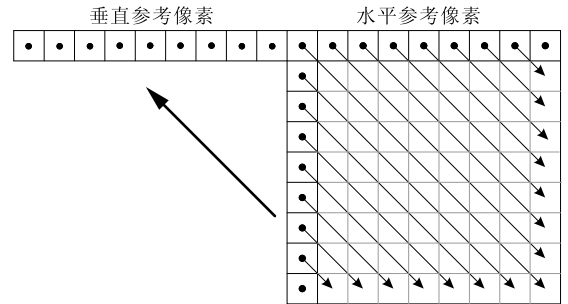


图 5 帧内预测中的投影过程

另外,对于图像中平滑区域的帧内预测,HEVC 中使用平面预测模式(Planar Mode)来替代原先的 DC 预测模式,在相同客观质量的条件下,非常有助于图像主观质量的提高^[36-37],预测过程如图 6 所示,首先右下角的像素 Z 由图中重建像素 L 和 T 平均得到,然后利用重建像素 L 和 T 与 Z 进行线性插值计算出最下面一行和最右边一列像素值,如图 6 中像素 M 和 N ,然后利用插值得到的像素进行双线性插值计算其它预测像素值,如图 6 中像素 P .

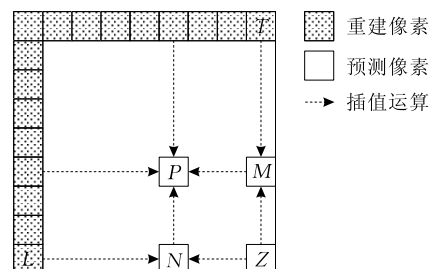


图 6 平面预测模式

3.2.2 色度分量帧内预测

对于色差分量,所有的预测单元 PU 都可以使用对应亮度预测单元的预测模式,但是预测模式的排序和编码略有差别,除此之外,色差分量还可以用“基于颜色通道相关性”进行色差分量预测^[38-40].因为在视频编码中,色差分量和亮度分量可能有相同的纹理信息,因此可以使用通道之间的相关性和线性模型,将对应预测单元的重建亮度分量像素值作为参考对色差分量进行预测,如图 7 所示.

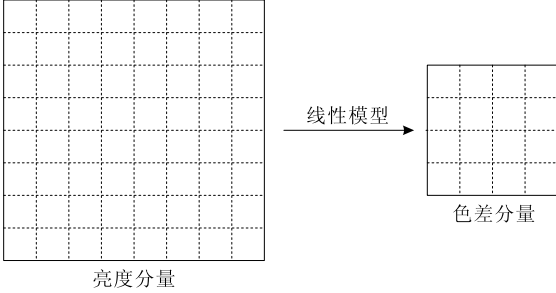


图 7 使用亮度分量重建值预测色差分量

预测方式如式(3)所示:

$$Pred_C[x, y] = \alpha \times Rec'_L[x, y] + \beta \quad (3)$$

其中, $Pred_C$ 表示色差分量预测值, Rec'_L 表示对应块的重建亮度值, 在 4:2:0 格式的 YUV 数据中, 由于亮度分量的分辨率是色差分量的 4 倍, 因此在使用这种模式进行色差帧内预测之前, 首先对重建亮度块数据分别进行水平方向和垂直方向的下采样处理, 如式(4)所示:

$$Rec'_L[x, y] = (Rec_L[2x, 2y] + Rec_L[2x, 2y+1]) \gg 1 \quad (4)$$

参数 α 和 β 使用最小线性二乘法计算得到

$\alpha =$

$$\frac{I \times \sum_{i=0}^I Rec_C(i) \times Rec'_L(i) - \sum_{i=0}^I Rec_C(i) \times \sum_{i=0}^I Rec'_L(i)}{I \times \sum_{i=0}^I Rec'_L(i) \times Rec'_L(i) - \left(\sum_{i=0}^I Rec'_L(i) \right)^2} = \frac{A_1}{A_2} \quad (5)$$

$$\beta = \frac{\sum_{i=0}^I Rec_C(i) - \alpha \times \sum_{i=0}^I Rec'_L(i)}{I} \quad (6)$$

这里 $Rec_C(i)$ 和 $Rec'_L(i)$ 分别表示重建的色差分量和下采样的重建亮度分量, 如图 8 中灰色位置处的像素所示. 对于 $N \times N$ 的色差块, 如果相邻的上边块和左边块都可用, 则用于计算的像素总数 I 为 $2N$, 如果只有左边或者上面一个相邻块可用, 则用于计算的像素总数 I 为 N .

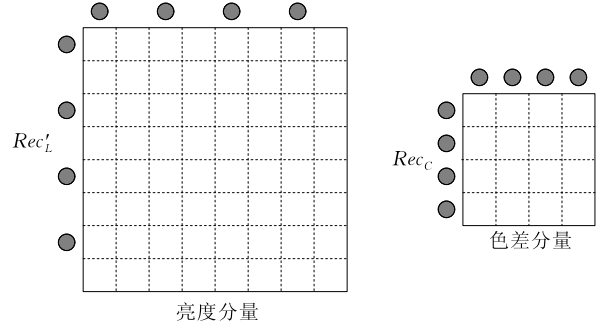


图 8 亮度分量和色差分量重建像素位置

3.3 帧间预测

在帧间预测方面, 和传统的视频编码标准一样, HEVC 采用的是基于运动补偿的预测编码技术, 消除视频数据时间域冗余, 但不同的是采用了基于 DCT 的分像素插值滤波器^[41-42] 和先进的运动矢量预测与合并技术^[43-45].

3.3.1 基于 DCT 的分像素插值滤波器

在运动补偿过程中, 如果运算矢量指向整像素位置, 则根据运动矢量和参考帧索引, 直接进行运动补偿操作; 如果运动矢量指向分像素位置, 则需要对分像素位置像素值的插值处理. 在 HEVC 标准制定的过程中, 提出了许多插值算法, 包括自适应插值滤波器、高阶插值滤波器等^[46], 其中基于 DCT 的插值滤波器有更好的频率响应, 并且响应形成的脉冲平稳, 波纹较小, 非常适合低通滤波器的图像分像素插值, 其设计原理及插值系数的计算过程如下^[47]:

经典 DCT 前向变换与后向变换的定义如式(7)和式(8)所示.

$$X(k) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} c_k x(n) \cos\left(\frac{(n+0.5)\pi k}{N}\right) \quad (7)$$

$$x(n) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} c_k X(k) \cos\left(\frac{(n+0.5)\pi k}{N}\right) \quad (8)$$

其中 c_k 的定义如式(9)所示.

$$c_k = \begin{cases} 1/\sqrt{2}, & k = 0 \\ 1, & \text{其它} \end{cases} \quad (9)$$

将前向变换式(7)代入后向变换式(8)即可得到分像素插值滤波器系数求解的形式化公式(10)^[48], α 为分像素的坐标位置, N 为滤波器的阶数, 例如对于半像素位置的 8 阶滤波器, 则 $\alpha = 1/2, N = 8$.

$$x(i+\alpha) = \sum_{m=0}^{N-1} x(m) \frac{2}{N} \sum_{k=0}^{N-1} \left(c_k^2 \cos\left(\frac{(m+0.5)\pi k}{N}\right) \times \cos\left(\frac{(i+\alpha+0.5)\pi k}{N}\right) \right) \quad (10)$$

在 HEVC 中, 运动矢量是 1/4 像素精度, 根据计算复杂度与性能的综合评估, 半像素亮度分量插

值滤波器为 8 阶, 1/4 亮度分量插值滤波器为 7 阶, 色差分量插值滤波器为 4 阶。

假设 $p(a)$ 表示分像素位置处的像素值, 已知 $2M-1$ 个整像素点的值, 即 $i=0, 1, \dots, 2M-1$, 如图 9 所示, 插值公式如式(11)所示. 在插值过程中, 对使用到的相邻分像素值由最近的已知整像素值替代.

$$x(\alpha) = \sum_{m=-M+1}^M x(m) \frac{2}{M} \sum_{k=0}^{2M-1} \left(c_k^2 \cos\left(\frac{(2m-1+2M)\pi k}{4M}\right) \right) \times \cos\left(\frac{(2\alpha-1+2M)\pi k}{4M}\right) \quad (11)$$

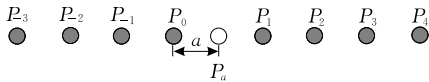


图 9 分像素位置示意图

相应的滤波器的系数如式(12)

$$f_m(\alpha) = \frac{1}{M} \sum_{k=0}^{2M-1} \left(c_k^2 \cos\left(\frac{(2m-1+2M)\pi k}{4M}\right) \right) \times \cos\left(\frac{(2\alpha-1+2M)\pi k}{4M}\right) \quad (12)$$

为了增强滤波器的平滑性能, 需要对式(12)中的滤波器系数进行一次平滑操作^[49], 如式(13)所示.

$$f_m(\alpha) = \frac{1}{M} \cos\left(\pi \frac{m-\alpha}{N-1}\right) \sum_{k=0}^{2M-1} \left(c_k^2 \cos\left(\frac{(2m-1+2M)\pi k}{4M}\right) \right) \times \cos\left(\frac{(2\alpha-1+2M)\pi k}{4M}\right) \quad (13)$$

根据上面等式, 可以计算出任意阶数 M 和任意分像素位置 a 的插值滤波器系数 $f_m(a)$. 高效的插值运算都是无乘法运算, 可以通过对系数做一些调整和整形化处理, 这样插值过程只需要加法和移位.

在 HEVC 中插值滤波器采用的是非级联形式, 即任意分像素位置的像素值可以直接用整像素值运算得到. 在如图 10 所示的像素位置分布图中, A、B、C 和 D 是整像素位置, 水平位置分像素 a 、 b 和 c 使用水平方向插值滤波器, 垂直位置分像素 d 、 h 和 l 使用垂直方向插值滤波器, 对于图中其余位置分像素, 先使用垂直方向插值进行块扩充, 然后再进行水平方向插值, 插值滤波器系数如表 2 所示, 两个 1/4 像素位置的滤波器系数是反对称关系. 表 3 中也列出了亮度各个分像素位置插值运算所需要的加法、乘法以及需要访问的像素次数, 并且与 H. 264/AVC 中的分像素插值运算做了对比. 从表 3 中的数据可以看出, HEVC 分像素插值滤波的计算复杂度明显大于 H. 264/AVC, 主要原因是 HEVC 中插值滤波器的阶数大于 H. 264/AVC. 当然在实际编码器设计的时候, 乘法一般都是使用移位来实现的.

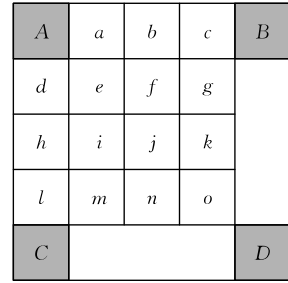


图 10 1/4 像素位置示意图

表 2 亮度分量分像素插值滤波器系数

分像素位置	滤波器系数
1/4	{-1, 4, -10, 58, 17, -5, 1, 0}
2/4	{-1, 4, -11, 40, 40, -11, 4, -1}
3/4	{0, 1, -5, 17, 58, -10, 4, -1}

表 3 HEVC 与 H. 264/AVC 中亮度分像素插值滤波器的复杂度对比

像素位置	HEVC 复杂度			H. 264/AVC 复杂度		
	访问像素	乘法	加法	访问像素	乘法	加法
a	7	7	6	6	6	6
b	8	8	7	6	6	5
c	7	7	6	6	6	6
d	7	7	6	6	6	6
e	49	56	48	11	12	11
f	56	63	55	36	42	36
g	49	56	48	11	12	11
h	8	8	7	6	6	5
i	56	63	55	36	42	36
j	64	72	63	36	42	35
k	56	63	55	36	42	36
l	7	7	6	6	6	6
m	49	56	48	11	12	11
n	56	63	55	36	42	36
o	49	56	48	11	12	11
平均	33	37	32	16	18	16

由于亮度分量采用 1/4 像素精度的运动矢量, 因此对应色差分量的运动矢量精度为 1/8 像素, 同样也是采用基于 DCT 的插值滤波器, 滤波器系数如表 4 所示.

表 4 色差分量分像素插值滤波器系数

分像素位置	滤波器系数
1/8	{-2, 58, 10, -2}
2/8	{-4, 54, 16, -2}
3/8	{-6, 46, 28, -4}
4/8	{-4, 36, 36, -4}
5/8	{-4, 28, 46, -6}
6/8	{-2, 16, 54, -4}
7/8	{-2, 10, 58, -2}

3.3.2 先进运动矢量预测和合并技术

在传统视频编码技术中, 为了提高运动矢量的编码效率, 编码端和解码端使用相同的预测模型对当前编码运动矢量进行预测, 即使用空域中相邻的

左上块、左边块和右上块的运动矢量中值作为预测运动矢量, 然后对运动矢量预测误差进行编码. 也有一些使用基于竞争的算法在构建的运动矢量集中选择最优的运动矢量作为当前块的预测运动矢量^[33], 并且将运动矢量索引值发送到解码端, 然而对于较大编码块来说, 构建的预测运动矢量集较大, 选择过于复杂并不利于优化编码.

在先进运动矢量预测技术中^[50-51], 候选运动矢量集由来自左边、上边空域相邻块以及参考帧中对应位置的时域相邻块运动矢量构成, 如图 11 所示, 包括当前编码块上面相邻块 B_0, B_1 与 B_2 , 左边相邻块 A_0 与 A_1 以及时域对应位置的相邻块 C_0 与 C_1 . 如果时域相邻块 C_0 的运动矢量不可用, 则直接使用当前预测块对应位置的中间块 C_1 的运动矢量作为候选运动矢量^[45, 52]. 候选运动矢量集中运动矢量个数最多为两个, 如果候选运动矢量小于两个, 则使用零运动矢量填充. 如果有两个运动矢量值相同, 则只采用一个作为候选运动矢量, 这也称为消除冗余运动矢量技术.

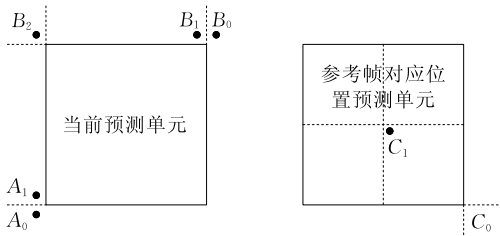
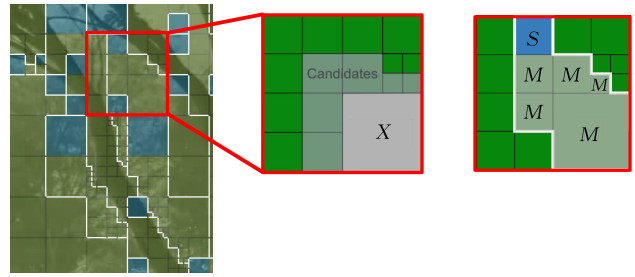


图 11 空域和时域预测运动矢量位置示意图

在 HEVC 中, 预测单元可以进行多次循环迭代划分, 并以二叉树形式进行表示, 有可能导致编码块过于细分, 需要编码大量的辅助信息, 从而导致编码效率的下降^[53-54], 因此需要对具有相同运动参数(运动矢量、参考帧索引和预测方向等)的块节点进行合并处理, 从而减少运动参数的编码和传输. 如图 12 所示, 假设当前编码块为 X , 与上述先进运动矢量预测技术一样, 首先利用空域与时域相邻块构建候选运动矢量列表, 然后对具有相同运动参数的相邻块进行合并, 共享相同的参考帧索引和运动矢量信息, 从而减少传输的数据量, 如图 12(c) 所示, 所有 M 块的运动矢量信息都与 S 块相同. 这种合并模式虽然在概念上与 H. 264/AVC 中的 Direct 模式和 Skip 模式相似, 但是 Direct 模式只是从相邻块中计算出当前块的运动矢量, 不需要传输任何信息, 而合并模式则是创建了一个具有相同运动信息参数区域, 对每个编码块需要将选择的候选运动矢量索引及其参考图像列表信息传输给解码端, 因此适应性更强, 编码效率更高.



(a) 图像的一个编码单元 (b) 当前编码块 X 及其相邻块 (c) 合并的预测单元及其相邻块

图 12 运动矢量合并

3.4 变换与量化

在 HEVC 中, 变换块的大小支持 4×4 到 32×32 , DCT 变换矩阵近似正交矩阵^[21, 55], 同时具备 DCT 变换的对称和反对称性质. 变换过程中, 所有中间值都限制在 16 位精度范围内, 所使用的乘法操作也是 16 位, 而对帧内预测残差, 则采用模式自适应的 DCT/DST^[56-57].

3.4.1 DCT 变换与量化

为了有效实现 DCT 变换, $2N$ 点的 DCT 都会重用 N 点的 DCT 过程, 也就是说, 较小 DCT 的变换系数都是较大 DCT 变换系数的子集, 如图 13 中的 8×8 和 4×4 DCT 系数所示, 4×4 DCT 系数是 8×8 DCT 系数的子集, 以此类推, 8×8 DCT 系数是 16×16 DCT 系数的子集, 16×16 DCT 系数是 32×32 DCT 系数的子集, 而且在量化和反量化过程中, 对于不同的变换基和不同大小的变换都使用统一的量化归一化参数进行量化. 在硬件实现和芯片设计中, 这些基本运算单元都可以进行有效复用.

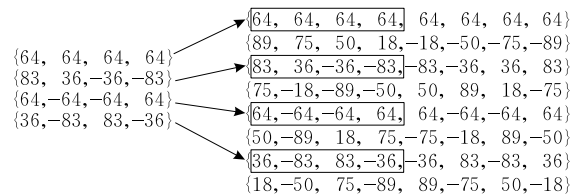


图 13 4×4 和 8×8 DCT 系数

DCT 变换过程由独立的水平变换和垂直变换两部分矩阵乘法组成, 为了保证 16 位精度的实现, 水平变换之后, 需要对变换系数进行右移位处理, 然后再进行垂直变换和量化. 在使用矩阵乘法实现 $N \times N$ 点的 DCT 变换时, 需要 N^3 个乘法和 $N^2 \times (N-1)$ 次加法运算. 由于 DCT 变换的对称性和反对称性, 可以用蝶形运算来实现快速 DCT 变换.

3.4.2 DST 变换与量化

对于 HEVC 中的帧内预测算法, 大致可以分为 3 类: 来自单边参考像素的方向性预测、来自双边参

考像素的方向性预测和无方向性的 DC 或者 Planar 预测模式. 在单边预测模式中, 距离参考像素越远, 预测误差越大, 如图 14 所示, 左边像素为参考像素, 右边像素为预测像素, 那么左边预测像素的预测准确性高于右边预测像素的预测精度, 离散正弦变换 DST 能够很好地适应这种预测误差的统计特性, 因此在在对残差进行变换编码的过程中, 根据使用的预测模式进行自适应的选择 DCT/DST 能够有效地提高编码效率^[58].

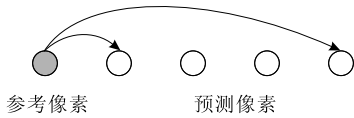


图 14 参考像素与预测像素的位置关系

3.5 熵编码

在 HEVC 中, 最初有两种熵编码方案, 即应用于低复杂度配置方案中的 LCEC 编码 (Low Complexity Entropy Coding) 和高效配置方案中的内容自适应二进制算术编码 CABAC (Context Adaptive Binary Arithmetic Coding)^[59], 后来考虑到编码效率和实现复杂度等原因, 在最终的 HEVC 编码方案中, 只保留了 CABAC 一种编码方案. CABAC 算法原理与 H. 264/AVC 中的编码方案类似, 首先对需要编码的语法元素和量化系数进行二值化, 然后选择合适的上下文模型进行概率估计, 最后再进行二进制算术编码^[23,60]. 但是 H. 264/AVC 中 CABAC 的数据相关性及其处理速度是视频编解码设计中的瓶颈, 特别对于解码器. 具体表现在以下几个方面:

(1) 概率模型和算术编码中的范围划分都需要递归运算.

(2) 上下文概率模型选择与编码语法元素相关性太强, 当前解码的语法元素概率模型由前面已解码的语法元素确定, 不利于并行编码和解码.

(3) 对于二值化后的语法元素, 上下文概率模型选择还与比特位置有关, 不同比特位置二进制数值的概率模型需要切换.

因此在 HEVC 标准制定过程中, 保证编码效率不变的前提下仅可能降低数据之间的相关性并增加数据处理速度, 包括: 减少概率模型的数量, 设计分组直接编码模式, 降低概率模型的相关性, 减少内存需求等等.

在对预测残差变换系数编码方面, 由于大变换

块系数统计特性比较分散, 难以使用统一的模型进行概率估计和编码, 因此首先对变换系数块划分成子块. 如图 15 所示, 将一个 8×8 变换块分成 4 个 4×4 子块, 首先对每个子块中数据是否全为零进行标注, 如果不全为零, 则进行对角、水平或者垂直方向的扫描, 将二维数据转换成一维数据进行编码传输. 这里需要注意的是, 在每个方向扫描中, 各行、各列或者各对角线数据都可以独立编码和解码. 与传统 ZIGZAG 扫描方式相比, 大大降低了编码系数之间的相关性, 有利于进行并行熵编解码, 提高数据处理速度.

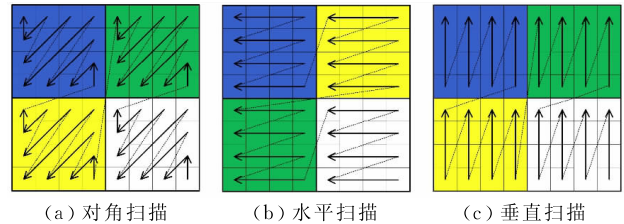


图 15 熵编码中子块的划分及扫描方式

3.6 环路滤波与自适应像素补偿

在 HEVC 中, 对重建图像进行级联去块斑滤波、像素自适应补偿和自适应环路滤波操作, 以提高解码图像的主客观质量^[61]. 其中去块斑滤波用于消除块效应^[62], 而像素自适应补偿和自适应环路滤波则用于消除量化噪声^[63].

3.6.1 去块斑滤波

类似于 H. 264/AVC 中的去块斑滤波, 以编码块为单位, 对重建图像进行块边缘滤波处理, 包括水平滤波和垂直滤波, 但不同的是用于滤波的最小块为 8×8 , 而不是 H. 264/AVC 中的 4×4 块, 有效消除了水平滤波和垂直滤波之间的相关性, 对于高分辨率视频图像在不影响图像主观质量的前提下降低了计算复杂度. 对于滤波权重的计算, 如果相邻块有帧内编码块、有非零系数或者相邻块的运动矢量差大于一个整像素, 都要进行去块斑滤波, 除此之外, 还要计算块边缘两侧像素值的差. 如图 16 所示, 计算方式如下:

$$d_i = |p_2 - 2p_1 + p_0| + |q_2 - 2q_1 + q_0| \quad (14)$$

这里 p_0 、 p_1 、 p_2 、 q_0 、 q_1 和 q_2 是位于块边缘两边的像素, 如图 16(a) 所示. 和 H. 264/AVC 不同的是, 不需要每行都计算两侧像素值的差 d_i , 例如: 对于 8×8 的块边缘, 只要计算图 16(b) 中两行像素的 d_i 值, 其余像素行滤波权重的计算都以这两行的 d_i 值为参考, 这样大大降低了计算滤波权重的复杂度.

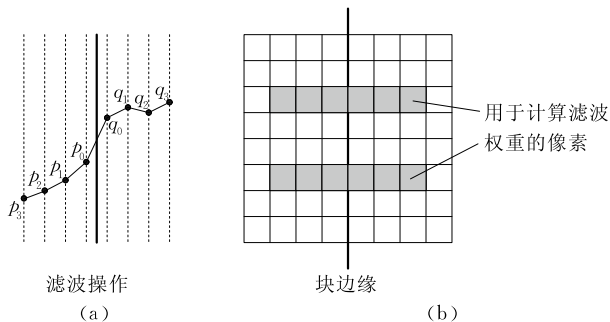


图 16 用于去块斑滤波权重计算的像素位置示意图

根据计算得到的滤波权重, 强滤波过程与 H. 264/AVC 相似, 而在弱滤波过程中, 首先按照式(15)计算 Δ 值。

$$\Delta = (13 \times (q_0 - p_0) + 4 \times (q_1 - p_1) - 5(q_2 - p_2)) / 32 \quad (15)$$

Δ 值的范围截取至 $(-t_c, t_c)$, t_c 是与量化参数 QP 成正比的参数, 量化参数越大, 滤波强度越大. 块边缘的滤波如下:

$$p_0 = p_0 + \Delta \quad (16)$$

$$q_0 = q_0 - \Delta \quad (17)$$

$$p_1 = p_1 + \Delta / 2 \quad (18)$$

$$q_1 = q_1 - \Delta / 2 \quad (19)$$

3.6.2 像素自适应补偿

像素自适应补偿技术通过对编码单元 CU 的重建像素进行分类, 并且对每类像素进行自适应的补偿以减少由于变换、量化和预测带来的块效应和振铃效应. HEVC 中使用了两种类型的像素自适应补偿技术: 边缘补偿和区间补偿. 在边缘补偿中, 首先对 CU 中的每个像素进行边缘分类, 如图 17 所示, 像素 c 是当前像素, 相邻像素为 a 和 b , 像素 c 的边缘属性根据表 5 中的条件分为 5 种类型. 在每个编码单元中, 编码器选择一种类型的边缘进行自适应补偿并把选择的边缘类型作为辅助信息传给解码端, 补偿的边缘类型和具体的补偿值由编码器率失真优化算法确定. 这里需要注意的是: 在边缘补偿中, HEVC 只允许平滑操作, 因此补偿值的符号由算法隐式确定; 另外, 亮度分量和色差分量的像素自适应补偿是互相独立的. 在区间补偿中, 将每个编码单元中的像素等分成 32 个区间(即量化步长为 8), 编码器根据率失真优化算法选择连续 5 个区间进行自适应补偿操作, 补偿区间的起始位置及每个区间的补偿值作为辅助信息传给解码端, 通过区间补偿可以有效地消除由于运动估计或量化算法导致的相移问题. 每个编码单元只能选择一种边缘补偿或者

区间补偿操作, 相关辅助信息也可由左边或者上边编码单元预测得到.

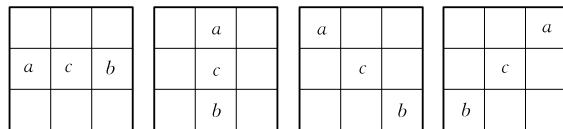


图 17 边缘检测模式

表 5 边缘类型分类表

边缘类型	分类条件
1	$c < a \&\& c < b$
2	$(c < a \&\& c = b) \parallel (c = a \&\& c < b)$
3	$(c > a \&\& c = b) \parallel (c = a \&\& c > b)$
4	$c > a \&\& c > b$
0	其它

3.6.3 自适应环路滤波

自适应环路滤波是利用维纳滤波技术对去块斑滤波后的重建图像进行处理, 提高解码图像的主客观质量. 对于亮度分量, 编码器在 Slice 层决定是否进行自适应的环路滤波, 使用的滤波器为 2 维钻石滤波器, 滤波器的大小分别为 5×5 , 7×7 和 9×9 , 如图 18 所示. 滤波像素根据其周围像素的方差选择合适的滤波器进行滤波处理. 对于色差分量, 只使用 5×5 大小的滤波器进行滤波操作, 但是滤波系数、滤波性能以及如何进行滤波仍然是编码优化的问题.

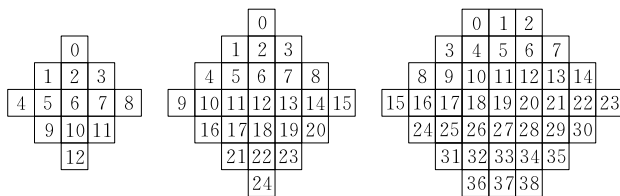


图 18 自适应环路滤波器

具体的滤波过程包括计算每个像素与周围像素的方差, 然后对像素周围 $\{-3, 3\}$ 范围内的方差进行统计分析, 进行滤波器索引值的计算, 根据滤波器索引值, 使用相应的滤波器系数和滤波器模板对重建像素进行自适应环路滤波.

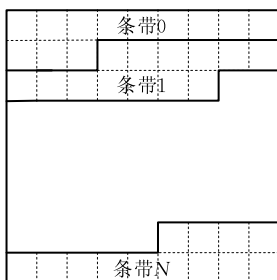
4 HEVC 计算复杂度分析

虽然 HEVC 编码技术框架仍然沿用基于空域和时域预测的混合编码框架, 但是与传统的视频编码算法相比(例如 H. 264/AVC), 采用了很多新的编码工具来提高编码效率, 包括更加灵活的块划分模式、35 种帧内预测模式、基于 DCT 的分像素插值

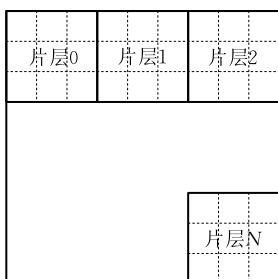
滤波器、先进的运动矢量预测算法与运动合并模式、像素自适应补偿滤波器与环路滤波器等。这些新的编码工具在提高压缩效率的同时,也增加了编解码器的计算复杂度,尤其对于编码器设计,在进行率失真优化过程中,有更多的编码选项需要进行组合优化。

在保证压缩效率的同时,JCT-VC 在制定 HEVC 标准过程中,也尽可能降低各个编码工具带来的计算复杂度的增加^[18]。在环路滤波方面,最小滤波块为 8×8 ,不仅大大减少了滤波像素的数量,而且将水平滤波和垂直滤波进行了分离,更加有利于并行化操作;在变换和量化方面,不仅数据位宽限制在 16 位范围内,而且通过对变换系数的设计,减少了运算复杂度,完成 8×8 、 16×16 和 32×32 块反变换分别需要 158、861 和 4696 个指令周期^[64];在熵编码方面,仅仅只使用 CABAC 方法,用于编码语法元素和残差系数的模型也大大减少,H. 264/AVC 中的模型数为 299 个,而 HEVC 仅仅有 154 个模型。

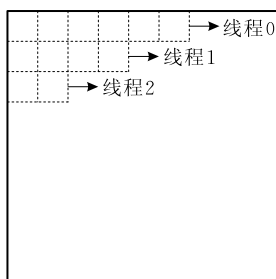
在标准设计方面,为了适应多核 CPU 技术和多线程编码技术的发展,在高层语法方面,HEVC 还提供了并行编解码机制,主要包括条带并行、片层并行和波形并行 3 种方式,如图 19 所示。在条带并行方法中,每个条带可以独立编码和解码,条带之间的分割按照顺序扫描方式进行,条带边缘的环路滤波由高层语法控制,这种条带划分方式对误码有较



(a) 条带并行



(b) 片层并行



(c) 波形并行

图 19 HEVC 中的并行模式

强的鲁棒性;在片层并行方法中,编码图像被分割成大小相同的图像片,每个图像片同样可以进行单独的编码和解码;在波形并行方法中,每个线程负责一行编码单元的编码或者解码,但是下一行编码单元的处理必须延迟前一行两个编码单元,在每行编码单元开始时,必须对 CABAC 进行重新初始化,其它与正常编码完全一样,这种波形并行方法对压缩效率损失最少。

为了进行 HEVC 的计算复杂度分析,使用文献 [65] 中的通用测试条件对 JCT-VC 的 HEVC 参考实现软件 HM6.0 进行测试,硬件平台为 P9700 (双核 2.4 GHz 主频),使用下述 3 种典型配置对 B 类和 C 类 (见表 6) 视频测试序列进行编码,包括帧内编码模式 (AI)、随机访问模式 (RA) 以及低延时模式 (LB),量化参数设置为 22 和 37,分别对应于高带宽和低带宽码流。实际编码时间见表 7,尽管 HM 只是用于 HEVC 的功能验证,不是用于做编码性能测试,但是在一定程度上也反应了其计算复杂度,其中率失真优化、分像素插值、变换与量化占了 70% 以上的计算复杂度。根据目前的评估,HEVC 的编码复杂度至少是 H. 264/AVC 的 4~8 倍,因此如何在保证压缩效率的前提下降低编码计算复杂度是今后一段时间的研究热点。

表 6 测试序列参数

类别	序列名称	分辨率	帧率
B	Kimono	1920×1080	24
	ParkScene		24
	Cactus		50
	Basketball Drive		50
	BQTerrace		60
C	Basketball Drill	832×480	50
	BQMall		60
	PartyScene		50
	RaceHorses		30

表 7 编码时间

序列	实际编码时间/s					
	AI22	AI37	RA22	RA37	LB22	LB37
Kimono	5444	4741	19175	14635	28283	21522
ParkScene	6105	4814	17741	13852	26833	19972
Cactus	12710	9852	40608	29156	55644	39181
Basketball	10883	9406	47207	34058	66554	46094
Drive	14180	11845	55054	34829	74883	46046
Basketball	2421	1954	8281	6085	11353	8240
Drill	2627	2390	9354	7453	13098	10003
BQMall	2422	2175	8926	6356	12832	8629
PartyScene	1512	1189	6694	4439	9386	6295

在实际应用中,尤其对于数字电视、流媒体等一些非实时视频应用,更加关注的是视频解码的计算复杂度。目前 HM 参考软件中解码器的计算复杂度

与 H. 264/AVC 的参考软件 JM 基本相当. 在文献[65-66]中, 基于 ARM 平台和 X86 平台, 对优化后的解码器计算复杂度进行了详细分析, 在 ARM (1GHz CortexA9) 平台上, 实时解码 C 类视频码流 (2 Mbps, 30 帧/s) 的各模块平均计算复杂度如图 20. 从图 20 中可以看出, 解码器中各模块计算复杂度的分布与 H. 264/AVC 解码器基本相当.

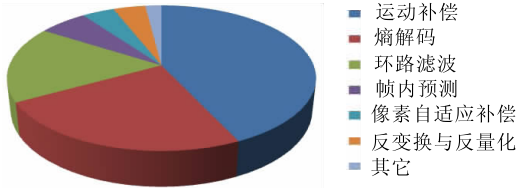


图 20 HEVC 解码器各模块的性能分析

5 HEVC 编码性能分析

为了验证 HEVC 的编码性能, 视频编码联合工作组 JCT-VC 在文献[48]中定义了通用测试条件、测试场景和相应的编码器配置文件信息, 率失真性能差的计算采用 Bjøntegaard 测量方法^[67], 综合峰值信噪比 PSNR 值由亮度分量与色差分量的 PSNR 值加权得到, 在文献[17]中, 对 HEVC 中各个编码工具对编码性能的提升做了详细的测试和分析. 根据测试结果可知, 采用大尺寸的编码单元 CU 以及自适应预测单元 PU 和变换单元 TU 对性能的提升最大, 特别对于高分辨率视频序列效果最明显, 最好情况可以提高压缩效率 39.2%.

由于 H. 264/AVC 是目前压缩性能最好、应用最广的视频编码标准, 因此在与传统视频编码标准做性能对比的过程中, 仅使用 H. 264/AVC 高级档次 (High Profile) 编码作为比较对象, 参考实现软件版本为 JM18.30^[68], HEVC 的测试版本为 HM8.0^[69], 测试序列采用通用测试条件中规定的 A 类至 F 类, HEVC 编码模式采用全帧内编码模式 (AI)、随机访问模式 (RA) 以及低延时模式 (LB) 3 种, 控制图像质量的量化参数分别设置为 22, 27, 32 和 37, 3 种模式的测试结果如表 8~表 10 所示. 从表中的数据可知, 在 LB 编码模式下, HEVC 获得的编码性能提高最大, 达到 35.1%, 另外, 对于高分辨率视频序列压缩性能的提升大于低分辨率视频序列, 如表中 A 类与 C 类性能的提升一般都大于其它类. 另外, 实验数据还表明在低码率下 HEVC 性能的提升大于高码率.

表 8 AI 编码模式下 HEVC 与 H. 264/AVC 的编码性能对比

测试序列	Y/%	U/%	V/%	YUV/%
A 类	-23.7	-21.6	-20.4	-23.4
B 类	-22.8	-22.5	-22.1	-22.8
C 类	-19.9	-21.1	-21.3	-20.3
D 类	-16.7	-16.9	-17.6	-16.9
E 类	-29.1	-27.4	-27.1	-28.8
F 类	-22.8	-19.5	-20.1	-22.7
平均值	-22.2	-21.3	-21.2	-22.2

表 9 RA 编码模式下 HEVC 与 H. 264/AVC 的编码性能对比

测试序列	Y/%	U/%	V/%	YUV/%
A 类	-36.7	-37.8	-39.4	-36.9
B 类	-39.7	-38.0	-37.2	-39.5
C 类	-30.0	-32.2	-32.2	-30.4
D 类	-27.2	-31.5	-33.2	-28.2
F 类	-25.6	-25.7	-27.0	-26.0
平均值	-32.2	-33.3	-34.0	-32.5

表 10 LB 编码模式下 HEVC 与 H. 264/AVC 的编码性能对比

测试序列	Y/%	U/%	V/%	YUV/%
B 类	-42.1	-34.7	-35.3	-41.1
C 类	-32.3	-32.3	-32.9	-32.5
D 类	-29.1	-31.5	-33.2	-29.7
E 类	-43.7	-39.5	-38.2	-42.9
F 类	-28.8	-30.5	-33.2	-29.9
平均值	-35.1	-33.5	-34.4	-35.1

6 未来研究方向

HEVC 是联合视频专家组 JCT-VC 为解决高清及超高清视频数据的存储、传输而制定的具有更高压缩效率的新一代视频编码标准. 尽管仍然基于传统的混合视频编码框架, 但对该框架中每个模块都进行了技术创新, 包括灵活的块划分模式、基于竞争的运动视频预测、基于 DCT 的分像素插值滤波器、高效的自适应算术编码以及波形并行处理技术等等. 这些新技术使得 HEVC 编码效率比 H. 264/AVC 提高了一倍. 但是这些编码选项的灵活组合使得 HEVC 的编码复杂度急剧增加, 阻碍了 HEVC 标准的快速推广和应用. 因此研究快速、高效的编码率失真优化算法、码率控制算法、主观质量优化算法应该是未来一段时期的研究热点.

基于对 HEVC 参考模型 HM^[69] 的测试分析可知, 灵活的块划分模型, 包括 CU、PU 和 TU, 对视频编码的率失真性能提高最为明显, 但其带来的计算复杂度也最大, 因此 HEVC 编码器如何根据相邻块上下文属性以及当前编码块的纹理特征进行快速的 CU、PU 和 TU 划分, 具有非常重要的研究价值. Shen 和 Alvarez-Mesa 等人在文献[70-71]中提出了

基于贝叶斯分类和支持向量基分类的编码单元 CU、变换单元 TU 以及预测单元 PU 的快速选择算法,在保证率失真性能的前提下大大提高了编码速度.当然还可以通过相邻单元之间的相关性、不同块层次之间的相关性来进一步降低计算复杂度.

为了适应多核处理器的发展趋势,HEVC 中引入了片层、条带以及波形并行处理 WPP 的思想,然而目前已有的并行视频编码方法粒度较粗,并行度不高,负载不均衡,无法充分挖掘多核处理器的计算能力,影响编码效率.因此研究适用于多核处理器的高并行度视频编码方法,为视频编码发展提供持续的计算能力保证,具有重要意义^[72-74].目前这方面的研究刚刚起步,研究点主要包括并行运动估计算法、并行帧内预测算法和并行熵编码算法等等.

在视频编码算法方面,尽管 HEVC 于 2013 年 1 月已经正式发布,但是随着应用需求的不断变化,HEVC 还在不断扩展,主要研究包括:(1) 三维视频、立体视频和多视角视频的编码技术^[75-76],不仅对视频数据进行压缩,还要对深度信息进行编码;(2) 分层 HEVC 编码技术^[77-78],用于提高 HEVC 对网络的自适应能力,目前主要包括空域和质量域可分层视频编码;(3) 将人类对视频图像的认知模型融入到视频编码算法中,对视频图像中感兴趣的区域分配较多的码率以提高视频图像的主观质量,具体包括感兴趣区域的检测技术和动态码率分配技术等;(4) 优化视频图像的主客观质量评价方法,即视频编码中率失真优化目标,目前主要以峰值信噪比以及结构相似性指标作为评价标准,难以精确模拟人类视觉系统的功能,还有待于进一步提高和改进.

7 HEVC 的应用前景

HEVC 标准是在 H.264 标准的基础上发展起来的,结合 H.264 在视频应用领域的主流地位可以预见 HEVC 标准在未来广大的发展前景.尽管 HEVC 的计算复杂度非常高,但是随着多核处理器性能的提高以及芯片处理能力越来越强,算法复杂性对应用的影响越来越小.相反,在实时通信应用以及 IPTV 应用中,业务的不断扩展和需求的增加使得有限的带宽资源逐渐成为瓶颈,高压缩率的视频编码是解决这一难题的有效技术手段,这也为 HEVC 在 IP 流媒体服务领域的应用奠定了坚实的基础.另外,由于视频编码软件开发周期短,平台灵活,因此基于 HEVC 的网络视频应用不久将大规模出现.

8 结束语

HEVC 是由视频编码联合工作组 JCT-VC 主导制定的新一代视频编码标准,尽管仍然采用的是混合视频编码框架,但是集成了很多先进的视频编码工具,包括 35 种预测模式的帧内预测、树形块编码组织形式、先进的运动矢量预测和运动合并算法、分像素插值算法以及去块斑滤波、像素自适应补偿、环路滤波等等,大大提高了视频压缩效率,使得 HEVC 比 H.264/AVC 标准的压缩性能更加高效.与 H.264/AVC 相比,在客观性能上,即重建图像的质量在相同的平均 PSNR 下,HEVC 的编码码率可以节省 40%;在主观性能上,即重建图像在相同的主观质量下,HEVC 的编码码率可以节省 50% 以上,这都说明了 HEVC 标准的优越.但需要指出的是,HEVC 的高效是以巨大的计算复杂度为代价的.从前面的分析可以看出,HEVC 标准中的新技术,给编码器优化留下了更多的编码参数和优化空间.只有合理有效地选择编码参数,才能使 HEVC 标准发挥其高效的压缩性能.

针对 HEVC 的研究,特别是优化算法的研究,随着标准的形成已经越来越受到学术界和工业界的重视.如何合理有效地发挥 HEVC 的压缩效率,是 HEVC 能否在工业界成功应用的至关重要的问题,也是学术界研究的热点问题. HEVC 的编码优化算法,主要集中在以下几个方面: HEVC 的码率控制算法、HEVC 低复杂度编码优化算法、HEVC 的主观质量优化算法.

致 谢 在本文撰写和修改过程中,审稿专家和编辑老师提出了许多宝贵意见和建议,在此表示衷心感谢!

参 考 文 献

- [1] ITU-T Q6/16 Visual Coding and ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Audio, Joint Call for Proposals on Video Compression Technology, ISO/IEC JTC1/SC29/WG11/N11113, Kyoto, Japan, 2010
- [2] Joint Call for Proposals on Video Compression Technology, ITU-T Q6/16, Document VCEG-AM91, Kyoto, Japan, 2010
- [3] Vision, Applications and Requirements for High-Performance video coding. ISO/IEC JTC1/SC29/WG11, MPEG09/N11096, 2009
- [4] Advanced Video Coding for Generic Audiovisual Services, Information Technology-Coding of Audio-Visual Objects,

- Part 10: Advanced Video Coding, document ITU-T Rec. H. 264, ISO/IEC 14496-10; 2009, 2010
- [5] Bross B, Han W-J, Ohm J-R, et al. High efficiency video coding (HEVC) text specification draft 8. Joint Collaborative Team on Video Coding (JCTVC-J1003), 2012
- [6] Sullivan G J, Ohm J-R, Han W-J, Wiegand T. Overview of the high efficiency video coding (HEVC) standard. *IEEE Transactions on Circuits Systems Video Technology*, 2012, 22(12): 1648-1667
- [7] Ohm J-R, Sullivan G, et al. Special section on the joint call for proposals on high efficiency video coding standardization. *IEEE Transactions on Circuits Systems Video Technol*, 2010, 20(12): 1661-1666
- [8] Ostermann J, Narroschke M. Draft requirements for next-generation video coding project. ITU-T Q. 6/SG16, VCEG-AL96, 2009
- [9] Sullivan G J, Ohm J-R. Meeting report of the first meeting of the joint collaborative team on video coding (JCT-VC). Joint Collaborative Team on Video Coding, document JCTVC-A200.doc, Dresden, Germany, 2010
- [10] Amonou I, et al. Description of video coding technology proposal by france telecom, NTT, NTT DOCOMO, Panasonic and Technicolor, ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 document JCTVCA114, Joint Collaborative Team on Video Coding (JCT-VC), Dresden, Germany, 2010
- [11] Winken M, et al. Video coding technology proposal by fraunhofer HHI, ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 document JCTVC-A116, Joint Collaborative Team on Video Coding (JCT-VC), Dresden, Germany, 2010
- [12] Ugur K, et al. Video coding technology proposal by tandberg, Nokia, and Ericsson. ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 document JCTVC-A119, Joint Collaborative Team on Video Coding (JCTVC), Dresden, Germany, 2010
- [13] Karczewicz M, et al. Video coding technology proposal by qualcomm. ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 document JCTVCA121, Joint Collaborative Team on Video Coding (JCT-VC), Dresden, Germany, 2010
- [14] Han W-J, et al. Video coding technology proposal by samsung (and BBC). ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 document JCTVC-A124, Joint Collaborative Team on Video Coding (JCT-VC), Dresden, Germany, 2010
- [15] Davies T, et al. Video coding technology proposal by samsung (and BBC). ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 document JCTVC-A125, Joint Collaborative Team on Video Coding (JCT-VC), Dresden, Germany, 2010
- [16] Test Model under Consideration, ISO/IEC JTC1/SC29/WG11, JCTVCB205, 2010
- [17] Ohm J-R, Sullivan G J, Schwarz H, et al. Comparison of the coding efficiency of video coding standards- including high efficiency video coding (HEVC). *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 22(12): 1669-1684
- [18] Bossen F, Bross B, Sühring K, Flynn D. HEVC complexity and implementation analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 22(12): 1685-1696
- [19] Kim I-K, Min J, Lee T, et al. Block partitioning structure in the HEVC Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 22(12): 1697-1706
- [20] Yuan Yuan, Kim I-K, Zheng Xiao-Zhen, et al. Quadtree based nonsquare block structure for inter frame coding in high efficiency video coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 22(12): 1707-1719
- [21] Sole J, Joshi R L, Nguyen N, et al. Transform coefficient coding in HEVC. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 22(12): 1765-1777
- [22] Fu Chih-Ming, Alshina E, Alshin A, et al. Sample adaptive offset in the HEVC standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 22(12): 1755-1764
- [23] Sze V, Budagavi M. High throughput CABAC entropy coding in HEVC. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 22(12): 1778-1791
- [24] Video Codec for Audiovisual Services at px64 kbit/s, ITU-T Rec. H. 261, Version 1; Nov. 1990, version 2; 1993
- [25] Video Coding for Low Bit Rate Communication, ITU-T Rec. H. 263, 1995
- [26] Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mbit/s—Part 2: Video, ISO/IEC 11172-2 (MPEG-1), ISO/IEC JTC 1, 1993
- [27] Coding of Audio-Visual Objects—Part 2: Visual, ISO/IEC 14496-2 (MPEG-4 Visual version 1), ISO/IEC JTC 1, 1999
- [28] Generic Coding of Moving Pictures and Associated Audio Information—Part 2: Video, ITU-T Rec. H. 262 and ISO/IEC 13818-2 (MPEG 2 Video), ITU-T and ISO/IEC JTC 1, 1994
- [29] Ma S, Kuo C-C J. High-definition video coding with super-macroblocks//*Proceedings of the SPIE*, part 1, San Jose, USA, 2007, 6508: 650816
- [30] Chen P, Ye Y, Karczewicz M. Video coding using extended block sizes. ITU-T Q. 6/SG16, document C123, 2009
- [31] Marpe D, Schwarz H, Bosse S, et al. Video compression using nested quadtree structures, leaf merging, and improved techniques for motion representation and entropy coding. *IEEE Transactions on Circuits Systems Video Technology*, 2010, 2(10): 1676-1687
- [32] Naito S, Koike A. Efficient coding scheme for super high definition video based on extending H. 264 high profile//*Proceedings of the SPIE*, Visual Communications and Image Processing. San Jose, USA, 2006, 6077: 607727-1-607727-8
- [33] Lainema J, Bossen F, Han W, et al. Intra coding of the HEVC standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 22(12): 1792-1801
- [34] Min Jung-Hye, Lee Sunil, Kim I-K, et al. Unification of the directional intra prediction methods in TMuC. ISO/IEC JTC1/SC29/WG11, JCTVC-B100, 2010

- [35] Lei Shawmin, et al. Description of core experiment 14: Intra mode coding. Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-D614r2, 4th Meeting: Daegu, Korea, 2011
- [36] Kanumuri S, Tan T K, Bossen F. Enhancements to intra coding. Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-D235, 4th Meeting: Daegu, Korea, 2011
- [37] Lainema J, Ugur K, Bici O. Planar intra coding for improved subjective video quality. Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-D326, 4th Meeting: Daegu, Korea, 2011
- [38] Kim J S, Park S W, et al. New intra chroma prediction using inter-channel correlation. ITU-T/ISO JCT-VC, JCTVC-B021, Geneva, Switzerland, 2010
- [39] Lee S H, Cho N I. Intra prediction method based on the linear relationship between the channels for YUV 4 : 2 : 0 intra coding//Proceedings of the IEEE International Conference of Image Processing. Cairo, Egypt, 2009: 1037-1040
- [40] Chen Jian-Le, Seregin V. Chroma intra prediction by reconstructed luma samples. ISO/IEC JTC1/SC29/WG11, JCTVC-C206, 2010
- [41] Tseng C C, Lee S L. Design of fractional delay FIR filter using discrete cosine transform//Proceedings of the APC-CAS. Macao, China, 2008: 858-861
- [42] Wedi T, Musmann H G. Motion- and aliasing-compensated prediction for hybrid video coding. IEEE Transactions on Circuits Systems Video Technology, 2003, 13(7): 577-586
- [43] Jung J, Laroche G. Competition-based scheme for motion vector selection and coding. VCEG-AC06, 2006
- [44] Kim S D, Ra J B. An efficient motion vector coding scheme based on minimum bitrate prediction. IEEE Transactions on Image Processing, 1999, 8(8): 1117-1120
- [45] Helle P, Oudin S, Bross B, et al. Block merging for quadtree-based partitioning in HEVC. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1720-1731
- [46] Lakshman H, Bross B, Schwarz H, Wiegand T. Fractional sample motion compensation using generalized interpolation//Proceedings of the PCS. Nagoya, Japan, 2010: 530-533
- [47] Lv Hao, Wang Rong-Gang, Xie Xiao-Dong, et al. A comparison of fractional-pel interpolation filters in HEVC and H. 264/AVC//Proceedings of the IEEE International Conference on Visual Communications and Image Processing (VCIP). San Diego, USA, 2012: 1-6
- [48] Tseng Chien-Cheng, Lee Su-Ling, Design of fractional delay FIR filter using discrete cosine transform//Proceedings of the IEEE Asia Pacific Conference on Circuits and Systems. Macao, China, 2008: 858-861
- [49] Alshin A, Alshina E. CE3: DCT derived interpolation filter test by Samsung, JCTVC-F247. Torino, Italy, 2011
- [50] Jung J, Bross B, Chen Pei-Song, Han Woo-Jin. Description of Core Experiment 9 (CE9): MV Coding and Skip/Merge operations, JCTVC D609. Daegu, Korea, 2011
- [51] Jung J, Bross B, Chen Pei-Song, Han Woo-Jin. CE9: Summary report of core experiment 9 on MV coding and skip/merge operations, JCTVC E029. Geneva, Switzerland, 2011
- [52] Lin Jian-Liang, Tsai Yu-Pao, Huang Yu-Wen, Lei Shawmin. Improved advanced motion vector prediction. ISO/IEC JTC1/SC29/WG11, JCTVC-D125, 2011
- [53] Shukla R, Dragotti P L, Do M N, Vetterli M. Rate-distortion optimized tree-structured compression algorithms for piecewise polynomial images. IEEE Transactions on Image Processing, 2005, 14(3): 343-359
- [54] De Forni R, Taubman D.S. On the benefits of leaf merging in quad-tree motion models//Proceedings of the IEEE International Conference Image Proceedings. Gena, Italy, 2005: 858-861
- [55] Fuldseth A, Bjøntegaard G, Sadafale M, Budagavi M. Transform design for HEVC with 16 bit intermediate data representation. ISO/IEC JTC1/SC29/WG11, JCTVC-E243, 2011
- [56] Saxena A, Fernandes F. CE7: Mode-dependent DCT/DST for intra prediction in video coding. JCTVC-D033. Daegu, Korea, 2011
- [57] Han J, Saxena A, Rose K. Towards jointly optimal spatial prediction and adaptive transform in video/image coding//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Dallas, USA, 2010: 726-729
- [58] Saxena A, Fernandes F C. CE7: Mode-dependent DCT/DST without $4 * 4$ full matrix multiplication for intra prediction. ISO/IEC JTC1/SC29/WG11, JCTVC-E125, 2011
- [59] Ugur K, Andersson K, et al. High performance, low complexity video coding and the emerging HEVC standard. IEEE Transactions on Circuits Systems Video Technology, 2010, 20(12): 1688-1697
- [60] Marpe D, Schwarz H, Wiegand T. Context-based adaptive binary arithmetic coding in the H. 264/AVC video compression standard. IEEE Transactions on Circuits Systems Video Technology, 2003, 13(7): 620-636
- [61] Norkin A, Bjøntegaard G, Fuldseth A, et al. HEVC deblocking filter. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1746-1754
- [62] List P, Joch A, Lainema J, et al. Adaptive deblocking filter. IEEE Transactions on Circuits Systems Video Technology, 2003, 13(7): 614-619
- [63] Huang Y-W, Fu C-M, Chen C-Y, et al. In-loop adaptive restoration. ISO/IEC JTC1/SC29/WG11, JCTVC-B077, 2010
- [64] Bossen F. On software complexity, document JCTVC-G757, JCT-VC. Geneva, Switzerland, 2011
- [65] Bossen F. Common test conditions and software reference configurations. Joint Collaborative Team on Video Coding (JCTVC-G1200). 2011
- [66] Bossen F. On software complexity: Decoding 720p content on a tablet, document JCTVC-J0128, JCT-VC, Stockholm, Sweden, 2012

- [67] Bjøntegaard G. Calculation of average PSNR differences between RD curves, document VCEG-M33, ITU-T SG 16/Q 6, Austin, TX, 2001
- [68] H. 264/AVC Reference Software [Online]. Available: <http://iphome.hhi.de/suehring/tml/>
- [69] HEVC Reference Software HM8.0 [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-8.0/
- [70] Shen Xiao-Lin, Yu Lu, Chen Jie. Fast coding unit size selection for HEVC based on Bayesian decision rule, picture coding symposium (PCS). Krakow, Poland, 2012; 453-456
- [71] Shen Xiao-Lin, Yu Lu. CU splitting early termination based on weighted SVM. EURASIP Journal on Image and Video Processing, 2013, 10(4): 28-38
- [72] Alvarez-Mesa M, Chi C C, Juurlink B, et al. Parallel video decoding in the emerging HEVC standard//Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Kyoto, Japan, 2012; 1545-1548
- [73] Chi Ching Chi, Mesa M A, Juurlink B H H, et al. Improving the parallelization efficiency of HEVC decoding//Proceedings of the ICIP. Florida, USA, 2012; 213-216
- [74] Chi C C, Juurlink B H H, Meenderink C. Evaluation of parallel H. 264 decoding strategies for the Cell Broadband Engine//Proceedings of the ICS. Arizona, USA, 2010; 105-114
- [75] Schwarz H, Bartnik C, Bosse S, et al. 3D video coding using advanced prediction, depth modeling, and encoder control methods//Proceedings of the IEEE International Conference on Image Processing. Krakow, Poland, 2012; 1-4
- [76] Tech G, Schwarz H, Müller K, Wiegand T. Effects of synthesized View Distortion based 3D Video Coding on the Quality of interpolated and extrapolated Views//Proceedings of the IEEE International Conference on Multimedia and Exposition. Melbourne, Australia, 2012; 20-31
- [77] ISO/IEC JTC-1/SC29/WG11 w12957, Joint Call for Proposals on Scalable Video Coding Extensions of High Efficiency Video Coding (HEVC), 2012
- [78] Segall A. BoG report on HEVC scalable extensions. JCT-VC document, JCTVC-K0354, Shanghai, China, 2012



SHEN Yan-Fei, born in 1976, Ph. D. candidate, associate professor. His research interests include video coding technology, video processing and multimedia communication.

LI Jin-Tao, born in 1962, professor, Ph. D. supervisor. His main research interests include multimedia technology, virtual reality and pervasive computing.

ZHU Zhen-Min, born in 1962, professor. His main research interests include embedded system technology and pervasive computing.

ZHANG Yong-Dong, born in 1973, professor, Ph. D. supervisor. His main research interests include signal processing, video coding algorithm and video content analysis.

Background

In recent years, with the development of video coding technology and the increasing popularity of high definition (HD) video content, the ISO/IEC Moving Picture Experts Group (MPEG) and the ITU-T Video Coding Experts Group (VCEG) formed the joint collaborative team on video coding (JCV-VC) in 2010. The JCT-VC group is aiming to develop next generation video coding standard, called HEVC. There are two distinct goals for the HEVC standard, a low complexity operation point with significant coding efficiency improvement suitable for mobile and real-time communication to enable emerging high resolution video services, and a high performance operating point with more substantial improvement in coding efficiency with additional complexity. So far the test model (HM) that is under intensive development has been used for performance evaluation and algorithm test. The final international standard of HEVC will be completed and published in the end of 2012 based on the working plan of JCV-VC.

In this paper, the key technology about HEVC will be surveyed, specifically including the quad-tree structure of coding unit, prediction unit and transform unit, the advanced motion vector prediction and merging technology, angular intra prediction, DCT-based fractional pixel interpolation filter and context adaptive binary arithmetic coding (CABAC) etc.

This work is supported by National Natural Science Foundation of China (Nos. 61001123, 61070109), which aims to improve video coding performance for mobile multimedia communication, and Guangdong province and Chinese Academy of Sciences Collaborative Research Project (No. 2012B091100212). In these years, the authors participated in the development of audio video coding standard of china (AVS) and proposed five patents included in AVS, at the same time, the author authored or co-authored over 30 papers in respectable journals and international conferences, including ICASSP, ICME, ICIP and IEEE transactions.