

【本文献信息】蔡晓霞,崔岩松,邓中亮,等.下一代视频编码标准关键技术[J].电视技术,2012,36(2).

下一代视频编码标准关键技术

蔡晓霞,崔岩松,邓中亮,常志峰

(北京邮电大学 电子工程学院,北京 100876)

【摘要】新一代视频压缩标准(High Efficient Video Coding, HEVC)是视频压缩领域继H.264/AVC之后的又一重大突破,主要面向高清电视(HDTV)以及视频编解码系统,提供从QVGA至1 080p以至超高清电视(7 680×4 320)不同级别的视频应用。首先从HEVC基本结构出发,较全面地介绍了HEVC在预测估计、DCT变换、滤波补偿以及熵编码方面采用的关键技术,最后介绍了HEVC当前的研究热点及其最新发展方向。

【关键词】视频编码;HEVC;二叉树编码、变换结构;环路滤波;熵编码

【中图分类号】TP37

【文献标识码】A

Model of Next-generation Video Standard and Relative Key Technologies

CAI Xiaoxia, CUI Yansong, DENG Zhongliang, CHANG Zhifeng

(School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)

【Abstract】As the successor to H.264/AVC, the latest video compression standard under development(High Efficiency Video Coding, HEVC) targets at next-generation HDTV displays and video compression systems which provide flexible, high efficient applications with resolutions from QVGA (320×240) up to 1 080p and Ultra HDTV (7 680×4 320). The encode architectures and some of the key technologies used in the new model are introduced in this paper. Those key technologies involve with prediction, transformation, in-loop filter and entropy. Besides that, the paper also reviews the latest improving issues and the future of HEVC.

【Key words】video coding; HEVC; quadtree coding and transforming structure; in-loop filter; entropy coding

1 HEVC背景与发展

H.264视频编码标准使得视频压缩效率提高到了一个新的水平。自该标准发布以来,H.264以其高效的压缩效率,良好的网络亲和性以及优越的稳健性等优点迅速得到了广大用户的认同。然而,随着终端处理能力以及人们对多媒体体验要求的不断提高,高清、3D、无线移动已经成为视频应用的主流趋势。而现有的H.264编码标准的压缩效率仍然不足以应对高清、超高清视频应用,需要更为高效的编码压缩方案。与此同时,近年许多新型有效的技术不断涌现,使得新标准的定制成为可能。为此国际电联组织(ITU-T)和移动视频专家组(MPEG)成立了视频编码联合小组(Joint Collaborative Team on Video Coding, JCT-VC)^[1],将新标准的定制正式提上日程。

2010年4月JCT-VC第一次会议在德国德累斯顿召开,所收到的27个提案从增加编码复杂度、提高压缩效率,或者从保证编码质量、降低编码复杂度的角度出发^[2],讨论如何在H.264/AVC高级档次的基础上进一步提高编码性能。新一代视频压缩标准主要面向高

清电视(HDTV)以及视频捕获系统的应用,提供从QVGA至1 080p以至超高清电视(7 680×4 320)不同级别的视频应用。其核心目标在于:在H.264/AVC High Profile的基础上,压缩效率提高1倍,即在保证相同视频图像质量的前提下,视频流的码率减少50%^[1]。

2 HEVC编码框架及其关键技术

HEVC依然沿用自H.263就开始采用的混合编码框架^[3],如帧内预测和基于运动补偿的帧间预测,残差的二维变换、环路滤波、熵编码等。在此混合编码框架下,HEVC进行了大量的技术创新,其中具有代表性的技术方案有:基于大尺寸二叉树块的分割结构和残差编码结构,多角度帧内预测技术,运动估计融合技术,高精度运动补偿技术,自适应环路滤波技术以及基于语义的熵编码技术。下文将对这个技术方案进行介绍。

2.1 基于二叉树结构的编码分割

为了提高高清、超高清视频的压缩编码效率,HEVC提出了超大尺寸二叉树编码结构,使用编码单元(Coding Unit, CU),预测单元(Prediction Unit, PU)和变换单元(Transform unit, TU)3个概念描述整个编码过程。其中

基金项目:中央高校基本科研业务费专项资金资助项目(2011RC0304)

CU类似于H.264/AVC中的宏块或子宏块,每个CU均为 $2N \times 2N$ 的像素块(N 为2的幂次方),是HEVC编码的基本单元,目前可变范围为 64×64 至 8×8 。图像首先以最大编码单元(LCU,如 64×64 块)为单位进行编码,在LCU内部按照四叉树结构进行子块划分,直至成为最小编码单元(SCU,如 8×8 块)为止,如图1所示。

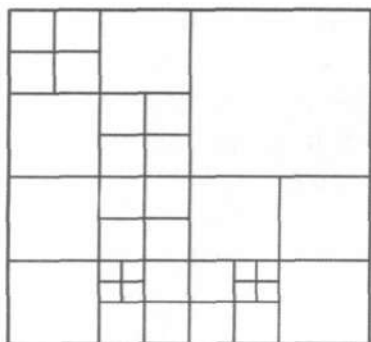


图1 LCU的树形结构示意图

对于每个CU,HEVC使用预测单元(PU)来实现该CU单元的预测过程。PU尺寸受限于其所属的CU,可以是方块(如 $2N \times 2N, N \times N$),也可以为矩形(如 $2N \times N, N \times 2N$),现有HM模型的对称PU分割类型如图2所示^[4]。

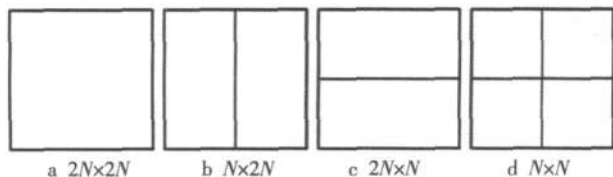


图2 4种PU分割类型

此外,一种新的不对称运动分割预测(Asymmetric Motion Partition, AMP)方案也已经被JCT所接受^[5],这也是HEVC与H.264在分块预测技术中最为不同之处。所谓AMP,即将编码单元分为两个尺寸大小不一致的预测块,其中一个PU单元的宽/长为CU单元的1/4,另一个PU对应的宽/长为CU单元的3/4,如图3所示。这种预测方式考虑了大尺寸可能的纹理分布,可以有效提高大尺寸块的预测效率。

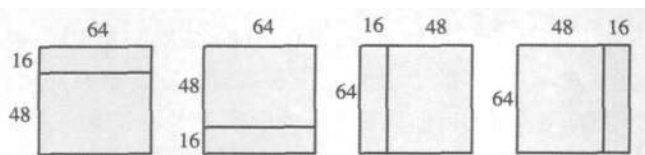


图3 64×64 CU所支持的4种AMP分割形态

2.2 HEVC的变换结构

HEVC突破了原有的变换尺寸限制,可支持 4×4 至 32×32 的编码变换,以变换单元(TU)为基本单元进行变换和量化。为提高大尺寸编码单元的编码效率,DCT变

换同样采用四叉树型的变换结构。图4为编码单元、变换单元的四叉树结构关系图示例,其中虚线为变换单元四叉树分割,实线为编码单元四叉树分割,编号为各编码单元的编码顺序。采用Z型编码顺序的好处为:对于当前编码单元,其上方块、左方块以及左上方块预测信息(如果存在)总是可以获得。

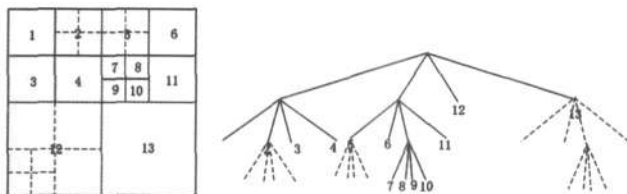


图4 编码单元与变换单元四叉树结构示意图

配合不对称预测单元以及矩形预测单元,新的HM4.0模型还采纳了相应的矩形四叉树TU结构^[6],突破了方块变换的限制。图5展示了3级矩形四叉树变换水平TU结构,同理可有垂直分割结构。

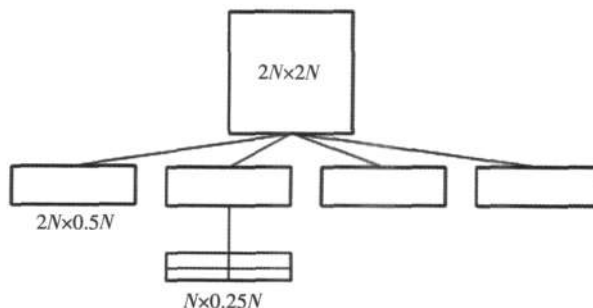


图5 矩形变换四叉树结构

尽管TU的模板发生了变化,但其变换核并没有发生实质性的变化。现有的关于不对称变换所使用的变换核是由方形变换核剪裁得到的。通常, $n \times m$ 的变换系数矩阵的计算公式为

$$C_{n \times m} = T_m \times B_{n \times m} \times T_n^T$$

式中: $B_{n \times m}$ 为 $n \times m$ 的像素块, T_m, T_n 分别为 $m \times m, n \times n$ 的变换核, $C_{n \times m}$ 为 $B_{n \times m}$ 的变换系数。

测试结果表明,非正方形四叉树更适合矩形PU和AMP变换,可节省大约0.3%的比特,同时增加2%左右的编码复杂度,对解码几乎没有影响^[7]。

采用大尺寸树形编码结构有利于支持大尺寸图像编码。当感兴趣区域一致时,一个大的CU可以用较少的标识代表整个区域,这比用几个小的块分别标识更合理。其次,任意LUC尺寸可以使编解码器对不同的内容、应用和设备达到最优化。对于目标应用,通过选择合适的LCU尺寸和最大分级深度,使编解码器具有更好的适应能力。LCU和SCU尺寸范围可被定义到档次和级别部分以匹配需求。

3 HEVC 预测编码技术

HEVC的帧间、帧内预测的基本框架与H.264基本相同:采用相邻块重构像素对当前块进行帧内预测,从相邻块的运动矢量中选择预测运动矢量,支持多参考帧预测等。同时,HEVC采用了如多角度预测,高精度运动补偿等多种技术,使得预测精度大大提高。

3.1 多角度帧内预测

HEVC的帧内预测将原有的8种预测方向扩展至33种,增加了帧内预测的精细度。另外,帧内预测模式保留了DC预测,并对Planar预测方法进行了改进。目前HM模型中共包含了35种预测模式,如图6所示。但由于受到编码复杂度限制,编码模型对4×4和64×64尺寸的PU所能使用的预测模式进行了限制。

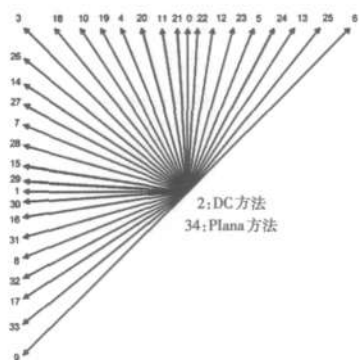


图6 35种帧内预测方式

原有的HM模型中色度分量帧内预测采用了5种预测模式,分别为水平、垂直、DC预测、亮度模式以及对角模式。JCT-VC第五次会议后增加了以基于亮度的色度帧内预测^[8],以取代对角预测模式。在该预测模式下,色度分量使用亮度分量的值进行线性预测,相关系数根据重构图像特性进行计算。该方案在色度分量上取得了8%左右的性能增益,而编码复杂度基本不变。

然而,尽管现有的帧内预测技术已对PU预测方向有所限制,但编码的复杂度仍然很高。不少研究人员提出了快速帧内预测算法,以进一步降低编码的复杂度^[9]。

3.2 帧间预测技术

3.2.1 广义B预测技术

在高效预测模式下,HEVC仍然采用H.264中的等级B预测方式,同时还增加了广义B(Generalized P and B picture, GPB)预测方式取代低时延应用场景中的P预测方式。GPB预测结构^[10]是指对传统P帧采取类似于B帧的双向预测方式进行预测。在这种预测方式下,前向和后向参考列表中的参考图像都必须为当前图像之前的图像,且两者为同一图像。对P帧采取B帧的运动预测方

式增加了运动估计的准确度,提高了编码效率,同时也有利于编码流程的统一。

3.2.2 高精度运动补偿技术

HEVC的编码器内部增加了像素比特深度,最大可支持12 bit的解码图像输出,提高了解码图像的信息精度。同时,HM模型采取了高精度的双向运动补偿技术^[11],即无论最终输出图像比特深度是否增加,在双向运动补偿过程都将使用14 bit的精度进行相关计算。

3.2.3 运动融合技术和自适应运动矢量预测技术

运动融合技术(Merge)将以往的跳过预测模式(Skip Mode)和直接预测模式(Direct Mode)的概念进行了整合。采用融合模式时,当前PU块的运动信息(包括运动矢量、参考索引、预测模式)都可以通过相邻PU的运动信息推导得到。编码时当前PU块只需要传送融合标记(Merge Flag)以及融合索引(Merge Index),无需传送其运动信息^[12]。

自适应运动矢量预测技术(Adaptive Motion Vector Prediction, AMVP)为一般的帧间预测PU服务,通过相邻空域相邻PU以及时域相邻PU的运动矢量信息构造出一个预测运动矢量候选列表,PU遍历运动矢量候选列表选择最佳的预测运动矢量。利用AMVP技术可充分发掘时域相关性和空域相关性。

值得一提的是,无论是运动融合技术还是自适应运动矢量预测技术,两者在候选运动矢量列表的设计上都进行了精心考量,以保证运动估计的高效性以及解码的稳健性。在早期的HM模型中,两种预测方式所使用的候选运动矢量列表是相互独立的;在JCT第6次会议结束后,新的HM模型中将两者的参考列表构造进行了统一^[13],Merge将采用与AMVP相同的方式构造候选运动矢量列表,进行运动信息的推导。

4 环路滤波

1个完整的HEVC的环路滤波过程包括3个环节:去块滤波,自适应样点补偿(Sample Adaptive Offset, SAO),自适应环路滤波(Adaptive Loop Filter, ALF)。去块滤波在H.264的去块滤波技术基础上发展而来,但为了降低复杂度,目前的HM模型取消了对4×4块的去块滤波^[14]。自适应样点补偿和自适应环路滤波均为HEVC的采用的新技术。

4.1 自适应样点补偿

自适应样点补偿是一个自适应选择过程,在去块滤波后进行。若使用SAO技术,重构图像将按照递归的方式分裂成4个子区域^[15],每个子区域将根据其图像

像素特征选择一种像素补偿方式,以减少源图像与重构图像之间的失真。目前自适应样点补偿方式分为带状补偿(Band Offset, BO)和边缘补偿(Edge Offset, EO)两大类。

带状补偿将像素值强度等级划分为若干个条带,每个条带内的像素拥有相同的补偿值。进行补偿时根据重构像素点所处的条带,选择相应的带状补偿值进行补偿。现有的HM模型将像素值强度从0到最大值划分为32个等级,如图7所示^[16]。同时这32个等级条带还分为两类,第一类是位于中间的16个条带,剩余的16个条带是第二类。编码时只将一类条带的补偿信息写入片头;另一类条带信息则不传送。这样的方式编码将具有较小补偿值的一类条带忽略不计,从而节省了编码比特数。



图7 32级像素值条带分割示意图

边缘补偿主要用于对图像的轮廓进行补偿。它将当前像素点值与相邻的2个像素值进行对比,用于比较的2个相邻像素可以在图8中所示的4种模板中选择,从而得到该像素点的类型:局部最大、局部最小或者图像边缘。解码端根据码流中标示的像素点的类型信息进行相应的补偿校正。

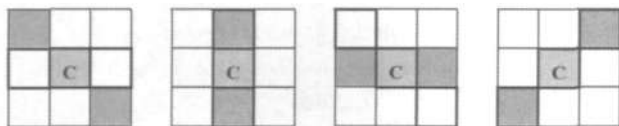


图8 4种边缘样点补偿模板

4.2 自适应环路滤波

自适应环路滤波(ALF)在SAO或者去块滤波后进行,目的是为了进一步减少重构图像与源图像之间的失真。ALF采用二维维纳滤波器,滤波系数根据局部特性进行自适应计算^[17]。对于亮度分量,采用CU为单位的二叉树ALF结构^[18]。滤波使用5×5,7×7和9×9三种大小的二维钻石型模板,如图9所示。滤波器计算每个4×4块的Laplacian系数值,并根据该值将所有4×4块分成16类,分别对应16种滤波器^[19],每种滤波器的滤波系数通过自适应维纳滤波器进行计算。

除上述基于像素的ALF分类外,提案^[20]还提出了基于区域的ALF分类。此时,每帧将被划分为16个区域,每个区域可以包含多个LCU。每个区域使用同一种滤波器,滤波器系数同样可以自适应训练得到。

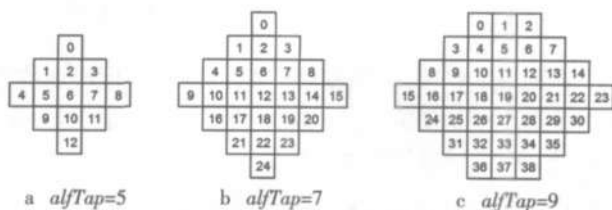


图9 3种ALF滤波模板

对于色度分量,滤波的选择过程却简单许多。首先,色度分量的滤波决策只需要在图像层级上进行。其次,滤波时色度分量统一使用5×5矩形滤波模板,不需要通过Laplacian系数来选择滤波器类型。

5 熵编码

CABAC是H.264的两种熵编码方案之一。现有的CABAC编码器采用串行处理的方式,解码端需要足够高频率的计算能力方能实时地对高码率的码流进行解码,直接导致解码功耗和实现复杂度的增加。为了解决CABAC的吞吐能力问题,JCT提出了熵编码模型并行化的要求。所收集的提案大致从3个角度提出了并行化CABAC解决方案:基于比特的并行CABAC^[21-23],基于语法元素的并行^[21]CABAC和基于片的并行CABAC。最后,基于语法元素的并行CABAC编码方案(即SBAC)被HM模型所采纳。目前,HM可支持上下文自适应变长编码(CAVLC)和基于语法元素的上下文自适应二进制算术编码(SBAC),分别用于低复杂度的编码场合和高效的编码场合。

SBAC的目的在于为具有不同统计模型的句法元素提供高效的编码方式。在SBAC中^[24],句法元素被分成N个类别,每个类别并行地维护着自己的上下文概率模型及其更新状态,每个类别的句法元素可对应一个或者多个概率表。因此,当各个类别所处理的比特量较均衡时,与原有串行编码器相比,并行编码器的吞吐量将提高N倍。然而实际运用中,各个类别的句法元素比特数不可能均衡,因此编码器吞吐量的提升将小于N倍。目前,HM中每一个句法元素都对应着一个或者多个概率模型,不同句法元素间的初始概率模型可能相同,并且可为每一个语法元素的每一位设计其选择概率模型的规则,以便为编码器提供最准确的概率估计。总体来说,SBAC的编码过程与原有的CABAC编码过程大致相同,都包括语法元素值二进制化、上下文概率模型选择、概率估计与上下文概率模型更新、二进制算术编码4个部分。具体SBAC的句法元素分类办法原则及其概率模型选择办法可进一步参考文献[21,24-27]。

6 总结

目前, HEVC的基本编码框架已经确定, 但许多技术细节仍在不断地研究中。专家组的主要力量集中在进一步提高 HEVC 编码效率以及降低其复杂度上。但除了提高 HEVC 编码效率以及降低编码复杂度的提案以外, 许多研究人员已经开始研究 HEVC 的可伸缩编码和多视点编码方案, 相关研究工作正在有计划地展开。与 H.264 High Profile 的编码性能相比, 目前 HEVC 已经取得了 40%左右的压缩性能提升, 而编码复杂度也达到了 150%左右, 不同测试场景的编码复杂度和性能提升程度有较大的差异。降低编码复杂度仍然是 HEVC 发展的一项重要议题。2011 年 7 月 22 日, 第 6 届 JCT 会议结束, 本次会议总共提出了 700 多项提案, 这些丰富的研究成果正极大地推动着 HEVC 前进的脚步。而 HEVC 的发展与完善必将极大地推动高清、超高清视频的应用步伐, 为人类献上更丰富的视觉盛宴。

参考文献:

- [1] ITU-T SG 16. Joint collaborative team on video coding[EB/OL]. [2011-07-20]. <http://www.itu.int/en/ITU-T/studygroups/com16/video/Pages/jctvc.aspx>.
- [2] SULLIVAN G, OHM J. Meeting report of the first meeting of the joint collaborative team on video coding, JCTVC-A200[R].[S.L.]: JCT-VC, 2010.
- [3] WIEGAND T, OHM J, SULLIVAN G, et al. Special section on the joint call for proposals on high efficiency video coding (HEVC) standardization[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(12): 1661-1666.
- [4] SUZUKI Y, TAN T, CHIEN W, et al. Extension of uni-prediction simplification in B slices, JCTVC-D421[R].[S.L.]: JCT-VC, 2011.
- [5] BOSSEN F. Common test conditions and software reference configurations, JCTVC-E700[R].[S.L.]: JCT-VC, 2011.
- [6] YUAN Y, ZHENG X, PENG X, et al. CE2: Non-square quadtree transform for symmetric and asymmetric motion partition, JCTVC-F412[R].[S.L.]: JCT-VC, 2011.
- [7] YUAN Y, ZHENG X, PENG X. Asymmetric motion partition with obmc and non-square TU, JCTVC-E376[R].[S.L.]: JCT-VC, 2011.
- [8] CHEN J, SEREGIN V, HAN W, et al. CE6.a.4: Chroma intra prediction by reconstructed luma samples, JCTVC-E266[R].[S.L.]: JCT-VC, 2011.
- [9] BROSS B, KIRCHHOFFER H, SCHWARZ H, et al. Fast intra encoding for fixed maximum depth of transform quadtree, JCTVC-C311[R].[S.L.]: JCT-VC, 2010.
- [10] MCCANN K, SEKIGUCHI S, BROSS B, et al. HEVC Test Model 3 (HM3) Encoder Description, JCTVC-E602[R].[S.L.]: JCT-VC, 2011.
- [11] UGUR K, LAINEMA J, HALLAPURO A. High precision bi-directional averaging, JCTVC-D321[R].[S.L.]: JCT-VC, 2011.
- [12] TAN T, HAN W, BROSS B, et al. BoG report of CE9: Motion vector coding, JCTVC-D441[R].[S.L.]: JCT-VC, 2011.
- [13] HUANG Y, BROSS B, ZHOU M, et al. CE9: Summary report of core experiment on MV coding and skip/merge operations, JCTVC-F029[R].[S.L.]: JCT-VC, 2011.
- [14] WIEGAND T, BROSS B, HAN W, et al. WD3: Working draft 3 of high-efficiency video coding, JCTVC-E603[R].[S.L.]: JCT-VC, 2011.
- [15] FU C, CHEN C, HUANG Y, et al. TE10 subtest 3: Quadtree-based adaptive offset, JCTVC-C147[R].[S.L.]: JCT-VC, 2010.
- [16] FU C, CHEN C, HUANG Y, et al. CE8 Subset3: Picture quadtree adaptive offset, JCTVC-D122[R].[S.L.]: JCT-VC, 2011.
- [17] EKSTROM M. Realizable wiener filtering in two dimensions[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1982, 30(1): 31-40.
- [18] CHUJOH T, WADA N, YASUDA G. Quadtree-based Adaptive Loop Filter, ITU-T Q.6/SG16 C181[R].[S.L.]: JCT-VC, 2009.
- [19] CHONG S, KARCZEWICZ M, CHEN C, et al. CE8 Subtest 2: Block based adaptive loop filter, JCTVC-E323[R].[S.L.]: JCT-VC, 2011.
- [20] CHEN C, FU C, TSAI C, et al. CE8 Subtest 2: Adaptation between pixel-based and region-based filter selection, JCTVC-E046[R].[S.L.]: JCT-VC, 2011.
- [21] BUDAGAVI M, SZE V, DEMIRCI M, et al. Video coding technology proposal by Texas Instruments, JCTVC-A101[R].[S.L.]: JCT-VC, 2010.
- [22] WINKEN M, BOBE S, BROSS B, et al. Video coding technology proposal by fraunhofer HHI, JCTVC-A116[R].[S.L.]: JCT-VC, 2010.
- [23] HE D, KORODI G, MARTIN-COCHER G, et al. Video coding technology proposal by research in motion, JCTVC-A120[R].[S.L.]: JCT-VC, 2010.
- [24] BUDAGAVI M, DEMIRCI M. Parallel context processing techniques for high coding efficiency entropy coding in HEVC, JCTVC-B088[R].[S.L.]: JCT-VC, 2010.
- [25] SZE V, BUDAGAVI M. Parallelization of HHI_TRANSFORM_CODING, JCTVC-C227[R].[S.L.]: JCT-VC, 2010.
- [26] AUYEUNG C, LIU W. Parallel processing friendly simplified context selection of significance map, JCTVC-D260[R].[S.L.]: JCT-VC, 2011.
- [27] SZE V, BUDAGAVI M, CHANDRAKASAN A. Massively Parallel CABAC, ITU-T SGI 6/Q.6 VCEG-AL21[R].[S.L.]: JCT-VC, 2009.

作者简介:

蔡晓霞(1985-), 女, 硕士生, 从事视频编码快速算法方面的研究;

崔岩松(1975-), 副教授, 从事视频编码、传输以及嵌入式系统设计方面的研究;

邓中亮(1965-), 教授, 博士生导师, 从事卫星导航、定位及多媒体技术方面的研究;

常志峰(1984-), 女, 硕士生, 从事视频编码技术中的码率控制技术方面的研究。

责任编辑: 时 雯

收稿日期: 2011-09-07