

基于运动特征的 HEVC 快速帧间预测算法

胡锦涛, 滕国伟, 成益龙, 晏 轲, 李国平, 赵海武, 王国中
(上海大学 通信与信息工程学院, 上海 200072)

摘要: 高效视频编码 (high efficiency video coding, HEVC) 帧间预测方法相比于其他视频编码标准具有更高的编码效率, 但同时也带来了更高的编码复杂度. 为了加速 HEVC 帧间预测过程, 提出了一种快速帧间预测算法. 该算法首先判断当前编码单元 (coding unit, CU) 的运动特征, 然后根据不同的运动特征采取不同的优化措施以减少帧间预测时间. 基于 HEVC 校验模型 (HM6.0) 的实验结果表明, 该算法在基本保持传统算法性能的基础上, 编码时间平均可减少 53.33%.

关键词: 高效视频编码; 帧间预测; 运动特征; 快速算法

中图分类号: TN 919.81

文献标志码: A

文章编号: 1007-2861(2013)03-0245-05

Fast Inter-Frame Prediction Algorithm for HEVC Based on Motion Features

HU Jin-wen, TENG Guo-wei, CHENG Yi-long, YAN Ke, LI Guo-ping,
ZHAO Hai-wu, WANG Guo-zhong

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China)

Abstract: Compared with other video coding standards, the inter-frame prediction of HEVC provides higher coding efficiency, leading to higher coding complexity however. This paper presents a new inter-frame prediction algorithm for HEVC to speed up the process. The method is based on motion features of coding unit (CU). It takes different optimization measures according to different motion features to reduce the inter-frame prediction time. Experiments on HM6.0 show that the proposed algorithm can reduce encoding time by about 53.33% for the same performance of the original algorithm.

Key words: high efficiency video coding (HEVC); inter-frame prediction; motion features; fast algorithm

随着因特网技术和移动通信技术的飞速发展, 超高清、3D和移动无线通信等新要求已被提出, 而已有的 MPEG-2, MPEG-4 和 H.264 等视频压缩标准难以满足上述要求. 为此, 视频编码专家组 (video coding experts group, VCEG) 和运动图像专家组 (moving picture experts group, MPEG) 组成了一个联合视频编码小组 (joint collaborative team on video coding, JCT-VC), 并开始了下一代视频编码标准的制订. 2010 年 4 月, JCT-VC 第一次会议在德国 Dresden 召开, 确定了新一代视频编码标准的名称——

HEVC, 并建立了测试模型 (TMuC), 其核心目标是在 H.264/AVC High Profile 的基础上, 使压缩效率提高一倍. 2013 年 1 月, HEVC 正式成为国际标准^[1].

HEVC 仍沿用传统的时空混合编码框架, 但在此基础上进行了大量的创新, 其中具有代表性的技术包括基于大尺寸四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^[2]. HEVC 采用了更加灵活的编码结构, 包括编码单元 (coding unit, CU)、预测单元 (predict

收稿日期: 2013-05-09

基金项目: 国家自然科学基金资助项目(61271212); 新型显示教育部重点实验室开放课题资助项目(P201104); 上海市科委重点资助项目(12511502502)

通信作者: 滕国伟(1975—), 男, 高级工程师, 博士, 研究方向为先进音视频压缩技术、新一代视频压缩技术等. E-mail: tenggw@shu.edu.cn

unit, PU)和变换单元 (transform unit, TU). CU 类似于 H.264 中宏块的概念, 最大可达 64×64 ; PU 是进行预测的基本单元; TU 是进行变换和量化的基本单元. 这三个单元的分离有助于它们更优化地完成各自的功能. 上述技术极大地提高了 HEVC 的编码效率, 同时也带来了极高的计算复杂度, 不利于 HEVC 在实时性要求较高的场合下应用, 其中帧间预测是 HEVC 中最为重要的技术之一, 也是导致 HEVC 计算复杂度急剧升高的最主要原因.

对编码模式的快速选择是帧间预测优化技术研究中的关键内容. 本研究利用 CU 的运动特征, 加速帧间预测模式判断, 在保证编码质量的同时, 显著提高了编码速度.

1 HEVC 帧间复杂度分析及优化

由于要考虑高清以及超高清视频的编码, HEVC 帧间预测采用了更大的编码尺寸, 最大可达 64×64 , 并以四叉树形式对 CU 进行递归处理, 如图 1 所示. 对于每层深度的 CU, 采用 Skip, $2N \times 2N$, $N \times 2N$, $2N \times N$ 及 $N \times N$ 等模式进行预测编码(非对称的预测暂不考虑). 通常对于 1 个 CU, 会有 1 个 Skip 模式编码块、1 个 $2N \times 2N$ 、2 个 $N \times 2N$ 和 2 个 $2N \times N$ 帧间模式编码块(共 6 个)进行率失真代价值计算. 对于处于最大编码深度的 CU, 还额外有 4 个 $N \times N$ 块需要计算.

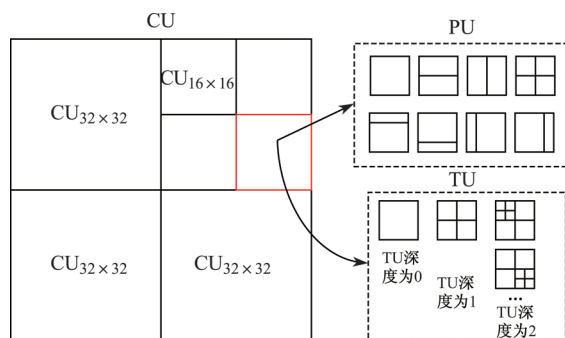


图 1 各单元间的关系

Fig. 1 Relationship between each unit

以编码尺寸为 64×64 , 最大编码深度为 3 的最大编码单元 (largest coding unit, LCU) 为例, 其编码过程如下: 当编码深度为 0 时, 需要进行 6 个预测块的计算; 当编码深度为 1 时, 需要进行 $4 \times 6 = 24$ 个预测块的计算; 当编码深度为 2 时, 需要进行 $4 \times 4 \times 6 = 96$ 个预测块的计算; 当编码深度为 3 时, 需要进行 $4 \times 4 \times 4 \times 6 + 256 = 640$ 个预测块的计算, 总计算次数为 $6 + 24 + 96 + 640 = 766$ 次. 以编码一幅大小为 1920×1080 的视频帧为例, 所需进行率失真代价值计算的帧间块预测次数约为 $766 \times 1920 \times 1080 \div$

$64 \div 64 \approx 387\,787$ 次. 由此可见, HEVC 的帧间预测因 CU/PU 的引入而导致计算量急剧增加. 因此, 有必要对快速帧间模式选择算法进行优化.

目前, 已有的 HEVC 快速帧间模式选择算法借鉴了 H.264 快速帧间模式选择算法, 主要有如下 3 类: ① 根据图像的运动特征, 对运动平缓的块使用大尺寸的预测模式, 对运动剧烈的块使用小尺寸的预测模式^[3]; ② 根据预测模式的时空相关性, 即利用时空相邻块的预测模式, 预测当前帧间模式^[4]; ③ 根据预测模式选择的统计规律, 跳过了某些极少被选为最佳模式的预测模式^[5]. 文献 [6] 通过对图像运动剧烈程度分级来进行模式选择, 但是该方法对运动剧烈的视频序列不甚精确, 码率增加较大, 为 3.51%. 文献 [7] 利用前后两帧相同位置处像素的绝对误差均值来判定运动剧烈区和运动平缓区, 以确定预测模式. 该方法仅在减少 CU 划分次数方面进行了优化, 若能与减少预测模式代价计算的方法相结合, 可进一步优化此方法. 文献 [8] 利用时空误差均值 (average of spatio-temporal error, ASTE) 判定当前宏块与周围宏块以及相邻共同位置处宏块的运动相关性, 并结合运动的剧烈程度来进行模式选择, 但该方法中的阈值是一个固定值, 不适用于不同纹理以及不同量化参数 (quantization parameter, QP) 的情况. 文献 [9-10] 也是通过前后帧的变化情况来判断运动剧烈程度, 以快速判定编码模式, 并且取得了较高的编码效率. 由此可见, 根据当前编码块的运动特征进行模式快速选择是一种切实有效的方法. 因此, 本研究将基于 HEVC 编码单元的运动特征来加快其帧间模式选择的速度.

2 快速帧间预测算法

2.1 CU 的运动特征衡量模型

在视频编码过程中, 对静态的、平坦的图像往往采用大尺寸块进行编码, 以减少运动向量、参考帧索引和预测模式的编码比特数; 对动态的、复杂的图像往往利用小尺寸块进行编码, 以提高编码后的视频质量. 判断 CU 的运动特征是整个算法的关键, 其中最简单有效的方式是利用相邻帧相同位置块的亮度像素值差来衡量运动变化程度. 文献 [9] 提出, 在 H.264 中利用前后编码帧相同位置处像素的绝对差值和 (sum of absolute difference, SAD) 来表征当前块的运动剧烈程度. SAD 的计算公式如下:

$$\text{SAD}_{16 \times 16} = \sum_{i=1}^{16} \sum_{j=1}^{16} |x_{i,j} - y_{i,j}|, \quad (1)$$

式中, $x_{i,j}$ 为当前帧 (i, j) 的像素灰度值, $y_{i,j}$ 为参考帧 (i, j) 的像素灰度值. 当 $\text{SAD}_{16 \times 16}$ 小于一个阈值

时, 表示其运动比较平坦, 此时仅需对 16×16 , 16×8 和 8×16 块进行率失真代价值计算; 反之, 则需对其他更小子块进行率失真代价值计算. 由于该模型比较简单, 文献 [10] 对其进行了改进, 利用 $SAD_{16 \times 16}$ 来衡量图像的运动剧烈程度, 即

$$SAD_{16 \times 16} = \sum_{l=0}^3 \sum_{k=0}^3 sad_{4 \times 4}^{l,k}, \quad (2)$$

式中, $sad_{4 \times 4} = E(sad_{4 \times 4}^2) - E^2(sad_{4 \times 4})$. 通过该模型可提前判定 Skip 模式或跳过某些尺寸的计算. 实验结果表明, 与文献 [9] 模型相比, 文献 [10] 的模型能够更准确地表征当前 CU 的运动剧烈程度, 所需编码时间更少.

本研究采用类似的思想, 首先判断当前 CU 的运动特征, 并选取相应的编码策略: 对运动剧烈的编码块采用小尺寸块进行编码, 对运动静止或者平缓的编码块采用大尺寸块进行编码. 与上述两种模型不同的是, 本研究将采用当前深度 CU 与时域相邻的已编码图像共同位置处 CU 的亮度像素值的差值平方和均值 (average sum of square difference, ASSD) 来表示当前 CU 的运动特征. ASSD 的计算公式如下:

$$ASSD = \frac{1}{2N \times 2N} \cdot \sum_{i=0}^{2N-1} \sum_{j=0}^{2N-1} (CU_{Cur(i,j)} - CU_{Col(i,j)})^2, \quad (3)$$

式中, $2N$ 表示 CU 的尺寸, $CU_{Cur(i,j)}$ 表示当前 CU 第 i 行第 j 列的亮度像素值, $CU_{Col(i,j)}$ 表示前一帧共同位置处 CU 第 i 行第 j 列的亮度像素值.

2.2 快速选择算法

基于上述分析, 本研究根据 CU 运动特征衡量指标将视频帧分为 3 个区域: 运动剧烈区、运动平缓区和运动适中区, 并通过定义提前终止阈值 $ThStop$ 和跳过阈值 $ThSkip$ 来区分这 3 个区域, 即

$$\text{当前图像} = \begin{cases} \text{运动剧烈区, } ASSD_{Cur} > ThSkip, \\ \text{运动平缓区, } ASSD_{Cur} < ThStop, \\ \text{运动适中区, 其他.} \end{cases} \quad (4)$$

为了避免 $ThSkip$ 和 $ThStop$ 的相近而导致对当前编码块的运动特征的误判, 设定 $ThSkip > 2 \times ThStop$, 其中 $ThStop = a \times Ave(ASSD_X)$, $a=0.8$ 为系数 (可根据实际编码要求进行调整), $Ave(ASSD_X)$ 为已编码的最佳深度为 X 的 CU 的 ASSD 均值; $ThSkip = b \times Ave(ASSD_{notX})$, $b=1.2$ 为系数 (可根据实际编码要求进行调整), $Ave(ASSD_{notX})$ 为已编码的

最佳深度大于 X 的 CU 的 ASSD 均值. $ASSD_{notX}$ 和 $ASSD_X$ 的具体计算公式如下:

$$ASSD_{notX} = \frac{1}{TotalNum} \left(\sum_{i=0}^{Num_{X+1}-1} ASSD_{X+1} + \sum_{i=0}^{Num_{X+2}-1} ASSD_{X+2} + \cdots + \sum_{i=0}^{Num_{MaxDepth}-1} ASSD_{MaxDepth} \right), \quad (5)$$

式中, $TotalNum$ 表示编码深度大于 X 的 CU 的总个数, 即

$$TotalNum = Num_{X+1} + Num_{X+2} + \cdots + Num_{MaxDepth}, \quad (6)$$

$Num_{X+1}, Num_{X+2}, \cdots, Num_{MaxDepth}$ 表示编码深度为 $X+1, X+2, \cdots, MaxDepth$ 的 CU 的个数, $MaxDepth$ 表示最大编码深度;

$$ASSD_X = \frac{1}{2N \times 2N} \cdot \sum_{i=0}^{2N-1} \sum_{j=0}^{2N-1} (CU_{Cur(i,j)} - CU_{Col(i,j)})^2, \quad (7)$$

式中, $2N$ 表示深度为 X 的 CU 的尺寸.

本算法可根据 CU 的运动特征对其深度进行自适应选择, 详细过程如图 2 所示, 主要步骤如下: ① 计算当前 CU 的 ASSD, 并根据已编码 CU 的信息, 判断当前 CU 的运动特征; ② 若当前 CU 处于运动剧烈区且其深度小于 $MaxDepth$, 则跳过当前 CU 所有预测模式的计算, 进入下一深度 CU 的编码; 若当前 CU 处于运动剧烈区且编码深度等于 $MaxDepth$, 则需要对 Skip, $2N \times 2N$, $N \times 2N$, $2N \times N$ 以及 $N \times N$ 等模式进行预测, 并计算率失真代价值; ③ 若当前 CU 处于运动适中区且编码深度小于 $MaxDepth$, 则对除 Skip 和 $N \times N$ 模式之外的其他模式进行率失真代价值的计算; 若当前 CU 处于运动适中区且编码深度等于 $MaxDepth$, 则跳过当前 Skip 模式, 对 $2N \times 2N$, $N \times 2N$, $2N \times N$ 以及 $N \times N$ 模式进行率失真代价值的计算; ④ 若当前 CU 处于运动平缓区, 则不再继续划分.

3 实验结果及分析

本算法测试了分辨率为 416×240 , 832×480 , 1280×720 , 1920×1080 的视频序列共 12 个, 测试平台为 HM6.0, 编码器采用低复杂度配置, 低延时编码方案, 最大 CU 深度为 3. 算法实验结果如表 1 所示,

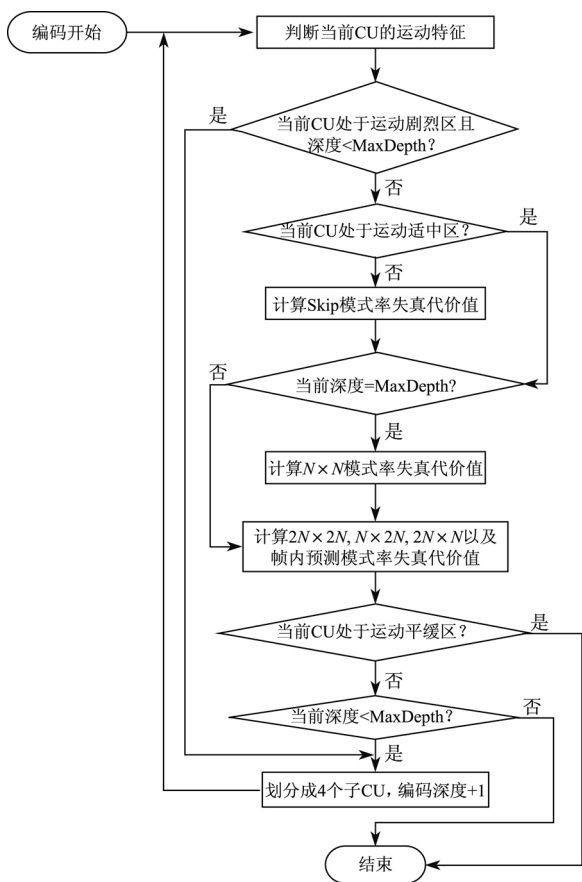


图 2 基于 CU 运动特征的快速算法流程

Fig. 2 Flow chart of the fast algorithm based on motion features of CU

其中 QP 为量化参数, ΔPSNR 为峰值信噪比 (peak signal to noise ratio, PSNR) 的差值, ΔBR 为比特率 (bit ratio, BR) 的增加幅度百分比, ΔTime 为编码时间的增加幅度百分比.

由表 1 可见, 本算法平均减少了 53.33% 的帧间预测编码时间, 亮度像素的 PSNR 值平均下降 0.07 dB, 输出码率平均增加 1.09%. 这一结果说明, 本算法可在保证编码效率的同时, 大幅减少编码时间, 其中对于某些视频序列 (如 Vidyo1 和 Vidyo4) 帧间编码的加速效果尤为显著, 但是对于某些视频序列 (如 RaceHorses 和 PartyScene) 加速效果略差. 经统计数据分析发现, Vidyo1 和 Vidyo4 等视频序列的每个深度的 ASSD_X 与 $\text{ASSD}_{\text{not } X}$ 的差值较大, 亦即大多数的 CU 处于运动剧烈区和运动平缓区. 本算法能够较为容易和准确地判断出当前 CU 的运动特征类型, 并采用跳过当前所有预测模式的编码以及停止划分的策略, 大大减少预测时间. 对于 RaceHorses 和 PartyScene 等视频序列, 编码过程中的 ASSD_X 与 $\text{ASSD}_{\text{not } X}$ 相差并不是很大, 即多数 CU 处于运动适中区. 对此, 本算法仅跳过了 Skip 模式的编码, 节省的预测编码时间有限, 平均只分别减少了 32.20% 和 33.69% 的帧间预测编码时间.

4 结束语

本研究在分析已有快速帧间预测算法的基础上, 提出了利用 CU 运动特征的快速帧间预测模式选择算

表 1 基于 CU 运动特征的快速算法结果

Table 1 Fast algorithm results based on motion features of CU

分辨率	序列	$\Delta\text{PSNR}/\text{dB}$			$\Delta\text{BR}/\%$			$\Delta\text{Time}/\%$		
		QP=27	QP=32	QP=37	QP=27	QP=32	QP=37	QP=27	QP=32	QP=37
416×240	BasketballPass	-0.06	-0.06	-0.06	1.70	1.67	2.45	-45.70	-49.83	-55.77
	BQSquare	-0.11	-0.11	-0.13	-0.48	-0.20	-1.19	-33.74	-42.44	-48.99
	RaceHorses	-0.10	-0.05	-0.06	4.84	3.16	2.15	-27.17	-32.71	-37.33
832×480	BasketballDrill	-0.07	-0.03	-0.05	1.96	1.19	1.24	-49.50	-48.70	-58.20
	BQMall	-0.06	-0.08	-0.06	2.37	3.40	1.85	-50.73	-58.59	-61.17
	PartyScene	-0.06	-0.08	0.09	0.55	0.15	0.05	-29.29	-32.78	-39.00
1 280×720	Vidyo1	-0.08	-0.08	-0.08	1.06	0.53	0.28	-68.21	-70.92	-72.93
	Vidyo3	-0.10	-0.10	-0.06	1.23	1.51	-0.22	-61.14	-65.89	-67.14
	Vidyo4	-0.04	-0.05	-0.04	0.35	0.39	1.13	-66.49	-67.46	-68.75
1 920×1 080	BQTerrace	-0.06	-0.06	-0.03	-0.55	-0.75	-0.95	-48.42	-53.78	-56.61
	Cactus	-0.06	-0.07	-0.07	2.76	2.36	1.38	-54.63	-58.34	-62.13
	ParkScene	-0.11	-0.11	-0.08	1.53	0.61	-0.13	-52.96	-58.91	-63.52
平均		-0.08	-0.07	-0.05	1.44	1.17	0.67	-49.00	-53.36	-57.63

法. 该算法通过对处于不同运动区域的 CU 采取停止划分和跳过编码的策略来提高编码速度. 实验结果表明: 与 HEVC 的校验模型 HM6.0 相比, 该算法可平均减少 53.33% 的编码时间, 而编码效率几乎没有变化, 有利于 HEVC 在 PC 平台上实现实时编码以及降低硬件成本.

参考文献:

- [1] SULLIVAN G J, OHM J R. Meeting report of the first meeting of the joint collaborative team on video coding (JCT-VC) [R]. Dresden: JCT-VC, 2010: 1-57.
- [2] POURAZAD M T, DOUTRE C, AZIMI M, et al. 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.
- [3] AHMAD A, KHAN N, MASUD S. Efficient block size selection in H.264 video coding standard [J]. Electronics Letters, 2004, 40(1): 19-21.
- [4] FENG B, ZHU G X, LIU W Y. Fast adaptive inter-prediction mode decision method for H.264 based on spatial correlation [C]// IEEE International Symposium on Circuits and Systems. 2006: 1804-1807.
- [5] XU L D, LIN X G. Fast mode decision for inter frames in H.264 AVC [C]// IEEE International Symposium on Communications and Information Technology. 2005: 433-436.
- [6] 曾庚卓, 马跃, 张伟. 一种基于 H.264 的快速帧间预测模式选择算法 [J]. 计算机系统与应用, 2010, 19(10): 80-84.
- [7] JING X, CHAU L P. Fast approach for H.264 inter mode decision [J]. Electronics Letters, 2012, 48(17): 1050-1052.
- [8] YIN M, WANG H Y. An improvement fast inter mode selection for H.264 joint with spatio-temporal correlation [C]// International Conference on Wireless Communications, Networking and Mobile Computing. 2005: 1237-1240.
- [9] JING X, CHAU L P. An efficient inter mode decision approach for H.264 video coding [C]// IEEE International Conference on Multimedia and Expo. 2004: 1111-1114.
- [10] LEE Y M, LIN Y Y. A fast intermode decision for H.264 video coding [C]// 8th International Conference on Signal Processing. 2006: 1-4.