

H.264/AVC 快速模式选择算法的研究与改进

宋学瑞, 段青青, 蔡子裕

SONG Xue-rui, DUAN Qing-qing, CAI Zi-yu

中南大学 信息科学与工程学院, 长沙 410083

School of Information Science and Engineering, Central South University, Changsha 410083, China

SONG Xue-rui, DUAN Qing-qing, CAI Zi-yu. Analysis and improvement of fast mode decision for H.264/AVC. Computer Engineering and Applications, 2008, 44(23): 68-70.

Abstract: H.264/AVC achieves high coding efficiency due to utilizing the full mode decision based on Rate-Distortion Optimization technique (RDO), which brings about tremendous computational complexity. To overcome it, JVT reference software proposes a new fast mode decision algorithm. In this paper, this new algorithm is introduced in detail. Some improved schemes are proposed based on the study and analysis of the advantages of the NFMD, such as early predict the mode of current macroblock, introduce timing terminate criterion, fast intra mode selection determination. Experimental results show that the improved strategies can reduce about 35%~55% of encoding time and 45%~60% of motion estimation time, while the R-D performance and bit rate are almost same as original algorithm.

Key words: H.264/AVC; inter-frame prediction; mode decision; spatio-temporal correlation

摘要: 针对 H.264 运用基于率失真优化的全模式搜索技术, 带来的模式选择异常复杂的计算, JVT 提出了一种新的快速模式选择算法(简化为 NFMD)。介绍了该新算法, 对其综合性能的优越性进行了初步研究和分析, 并提出了改进策略, 如早期预测当前宏块编码模式、引入适时终止模式选择策略、快速帧内模式选择判断。实验结果表明, 改进的策略能减少 NFMD 算法 35%~55% 的编码时间和 45%~60% 的运动估计时间, 而码率和率失真性能的变化可以忽略不计。

关键词: H.264/AVC; 帧间预测; 模式选择; 时空域相关性

DOI: 10.3778/j.issn.1002-8331.2008.23.021 文章编号: 1002-8331(2008)23-0068-03 文献标识码: A 中图分类号: TP391

1 引言

H.264 与以往的视频压缩标准相比具有更好的压缩性能, 其编码效率的提高源于帧间预测应用了复杂模式选择技术和变尺寸的块状运动估计(VSBME)。H.264 支持 7 种模式的帧间预测, 对于 P 片还可以使用 SKIP 模式, 帧内预测 2 种模式划分, 编码器计算每一宏块各种模式的 RDcost (Rate Distortion cost), 然后选择代价最小的模式作为最佳的宏块编码模式。图 1 展示了对某一特定模式计算其 RDcost 的过程。对已给定的宏块需要进行所有模式的 RDcost 计算, 计算 RDcost 需要重复的进行运动估计、变换/反变换、变长编码、量化/反量化等复杂过程, 这些计算毫无疑问会使编码速度减慢。

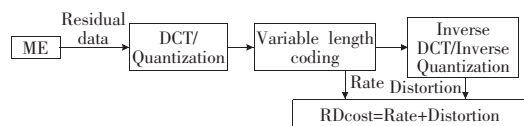


图 1 RDcost 的计算过程

如果能根据某些特征预先判断出最优模式或者预测出可能的模式范围, 就可以减少不必要的模式搜索和 RDcost 计算,

提高编码速度, 降低编码复杂度。考虑到这些特征, 视频联合小组(JVT)提出了一些快速模式选择算法: 基于边缘直方图的快速帧内预测模式选择算法(FMDintra)、依据帧内预测得到的边缘映射信息的快速帧间模式选择算法(FMDinter)等。最新的 JVT 参考软件融合了以上两种算法, 增加了两项早期终止模式选择的判断, 提出了一种新的快速模式选择算法^[4]。仿真结果表明, 该算法相对于全模式搜索算法能在率失真性能几乎没有变化的情况下, 能平均降低 30% 的编码时间和减少 72% 的 RD-cost 运算次数。

2 新快速模式选择算法研究与分析

2.1 采用 RDO 技术的全模式选择过程

H.264 编码标准中支持其中不同块大小的帧间编码模式。每个宏块按照 16×16、16×8、8×16、8×8 进行分割, 如果选择 8×8 块模式, 还要按 8×8、8×4、4×8、4×4 进行亚分割, 完整的全模式选择算法流程如下:

(1) 对当前宏块的模式 16×16、16×8、8×16 进行运动估计并计算 RDcost;

作者简介: 宋学瑞(1957-), 男, 副教授, 主要研究领域为视频压缩编码及通信、大规模集成电路; 段青青(1984-), 女, 硕士, 主要研究领域为视频压缩编码及多媒体; 蔡子裕(1983-), 男, 硕士, 主要研究领域为多媒体通信、嵌入式系统。

收稿日期: 2007-10-23

修回日期: 2008-01-15

(2)对每个 8×8 模块的模式 $8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4$ 计算 RDcost, 选择 RDcost 最小的模式作为 8×8 块的最佳模式;

(3)重复执行第(2)步,直到计算完 4 个 8×8 块,即得到 8×8 模式的 RDcost;

(4)为 SKIP 模式计算预测运动矢量和 RDcost;

(5)计算 intra 模式的 RDcost, intra 模式包括 I4MB, I16MB, high profile 中还有 I8MB 和 IPCM;

(6)从 $16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, \text{SKIP}, \text{intra}$ 模式中选择 RDcost 最小的模式作为帧间宏块编码模式。

2.2 新快速模式选择算法研究与分析

统计数据表明 P 帧中编码宏块运用 SKIP 模式和 inter16 \times 16 模式的比率最大,尤其是在运动相对平缓的视频序列中。而且采用 SKIP 模式不需要传送像素残差量化值,运动矢量及参考索引等信息,解码时只需将前一帧对应位置重建宏块信息复制到当前块即可,所以一旦确定当前宏块的最佳编码模式为 SKIP 模式,则可以省略后续的率失真计算、运动补偿等一系列复杂工作。同时发现采用 intra 模式编码的宏块所占比率很小,平均小于 3%^[2],即便是在运动剧烈的视频序列中最多达到 9%^[2],而且 intra 模式预测比较复杂,对图像的质量影响很小。因此对给定的宏块提前进行 SKIP 模式的判断和是否进行帧内模式搜索的判断对提高编码速度是非常有效的。

基于这样的理论,JVT 在 2.1 节的全模式搜索算法中提出了两点关键的预测技术:早期 SKIP 模式判断和选择性的 intra 模式检测,形成了 NFMD。

(1)早期 SKIP 模式判断。算法首先检测是否为 SKIP 模式,SKIP 模式判决是基于 16×16 块进行的,其过程为:先对宏块进行运动估计和参考帧预测,然后计算其 RDcost,当其同时满足以下条件时可判定宏块最佳编码模式为 SKIP 模式,从而终止其他模式的检测:①运动补偿块尺寸是 16×16 ;②最佳参考帧是已重建的得前一帧;③运动矢量为预测运动矢量;④量化后的变换系数均为零。

(2)选择性的 intra 模式检测。在进行完各个 inter 模式检测后,算法利用由 inter 预测得到的编码当前宏块所需的平均最小比特数 AR 和宏块平均边缘差值 ABE 来预先检验是否使用 intra 模式进行编码,如果 AR 足够小或者 $AR < K \times ABE$,则无需再进行 intra 模式的检测,选择当前 RDcost 最小的模式作为最佳编码模式即可。AR 和 ABE 定义如式(1)和式(2)所示:

$$AR = 1/384 (\text{Bitrate}) * \lambda \quad (1)$$

式中 $\lambda = 0.85 * 2^{OPB}$, Bitrate 为由 inter 预测得到的编码当前宏块所需的最小比特数。

$$ABE = 1/32 * SBE \quad (2)$$

计算 SBE 时要根据左侧宏块 mb_left 和顶侧宏块 mb_upper 的是否可用来判断:

(1)当 mb_left 和 mb_upper 都不可用时:如果 $AR = 0$,则不用进行 intra 模式检测;否则 $ABE = 0$;

(2)当 mb_left 不可用,mb_upper 可用时:

$$SBE = \sum_{i=1}^{15} |Y_{orig}(x+i, y) - Y_{rec}(x+i, y-1)| ABE = 1/16 * SBE \quad (3)$$

(3)当 mb_left 可用,mb_upper 不可用时:

$$SBE = \sum_{i=1}^{15} |Y_{orig}(x, y+i) - Y_{rec}(x-1, y+i)| ABE = 1/16 * SBE \quad (4)$$

(4)当 mb_left 和 mb_upper 都可用时:

$$SBE = \sum_{i=1}^{15} |Y_{orig}(x, y+i) - Y_{rec}(x-1, y+i)| + \sum_{i=1}^{15} |Y_{orig}(x+i, y) - Y_{rec}(x+i, y-1)| ABE = 1/32 * SBE \quad (5)$$

式(3)~式(5)中 Y_{orig} 和 Y_{rec} 分别是编码宏块的像素及重建的相邻宏块的像素值。式中只显示宏块亮度值的 ABE,色度值的 ABE 类似。

3 基于新快速模式选择算法的分析和改进

NFMD 提出的早期 SKIP 模式判断方法只需要改变检测帧间模式的顺序,可以很方便的运动用到传统的编码结构中。选择性的 intra 模式检测技术中 AR 的估值运用了已得的 inter 模式信息,ABE 的计算运用已编码的相邻块的信息,没有增加其它的计算量。实验结果表明 NFMD 算法可以降低 72% 的 RDcost 计算量和 30% 编码时间,而且其 PSNR 最多减少 0.03 dB。

尽管 NFMD 算法已有效的提高了编码效率,但其却未考虑到视频序列以下许多特征:

(1)视频序列具有时空相关性^[3]。在空域中,当前宏块的编码模式以很大概率等同于与已编码的相邻宏块的编码模式,同时多数运动对象存在时间连续性,前一帧中相应位置的块的编码模式也与当前块相似。因此可以根据图像的这些相关信息,对当前编码模式进行预测,缩小模式选择范围,减少编码时间。

(2)大量实验数据表明,如果较大块尺寸的 RDcost 小于当前尺寸块的 RDcost 时,更小尺寸块的检测就可以排除而不会有任何率失真性能的降低^[4]。适时的引入一些早期终止模式选择的判断标准,可进一步加速编码速度。

(3)在帧内模式选择过程中,由于 intra16 采用 SATD 的模式选择准则,而 intra4 采用率失真优化技术选择最佳模式,同时 intra4 比 intra16 子块数和可选择的模式数都多,因此 intra4 比 intra16 预测过程复杂得多。帧内模式为最佳编码模式时:量化参数 QP 较小时,intra16 与 intra4 的 RDcost 相近,QP 较大时,intra16 比 intra4 的 RDcost 小的多,采用 intra16 比采用 intra4 的概率大,因此在一定的条件下跳过 intra4 模式对编码性能不会有太大影响,而对编码速度提高却比较有效。

针对以上特点,对 NFMD 算法进行了以下几点改进:

(1)对当前宏块的编码模式进行预测

本文仅选用与当前宏块相邻的已编码的上方、左方的宏块和前一帧中相应位置的宏块来预测当前块,并且只预测 SKIP 模式和 16×16 模式。设当前块为 X,相邻的左边块 A,上边块 B,前一帧当前块为 X_{-1} ,各自的编码模式分别为 $modeA, modeB, modeX_{-1}$, RDcost 分别为 $modeA_rdcost, modeB_rdcost, modeX_{-1}_rdcost$,具体预测步骤如下:

①当 A 和 B 都可得时:如果 $modeA = modeB = modeX_{-1} = mode$,选择 mode 编码模式作为当前块最佳编码模式,结束模式选择;如果 $modeA = modeB \neq modeX_{-1}$,检测空间相邻预测块的均值 $RDcost_aver = (modeA_rdcost + modeB_rdcost) / 2$,若 $RDcost_aver < modeX_{-1}_rdcost$,选择 modeA 或 modeB 作为当前块最佳编码模式,反之选择 $modeX_{-1}$,结束模式选择;如果 $modeA \neq modeB \neq modeX_{-1}$,选择 $modeA_rdcost, modeB_rdcost, modeX_{-1}_rdcost$ 中最小的模式 $mode_min$,对当前块采用 $mode_min$ 编码模式并计算其 RDcost,如果 $RDcost < \min(modeA_rdcost, modeB_rdcost, modeX_{-1}_rdcost)$

eX_{-1_rdcost}), 则选择 $mode_min$ 作为当前块的编码模式, 否则进行全模式搜索。

②当 A 可得, B 不可得时: 如果 $modeA = modeX_{-1} = mode$, 则选择 $mode$ 编码模式作为当前块最佳编码模式, 结束模式选择; 选择 $modeA_rdcost, modeX_{-1_rdcost}$ 中小的模式 $mode_min$, 对当前块采用 $mode_min$ 编码模式并计算其 $RDcost$, 如果 $RDcost < \min(modeA_rdcost, modeX_{-1_rdcost})$, 则选择 $mode_min$ 作为当前块的编码模式, 否则进行全模式搜索。

A 不可得, B 可得时与 A 可得, B 不可得时的选择方法类似。 A 和 B 都不可得时则进行全模式搜索。

(2) 引入适时终止模式选择策略

当编码 $16 \times 8, 8 \times 16$ 模式的 $RDcost$ 都大于 16×16 模式时, 可判定 16×16 模式为编码当前块的最佳模式, 结束模式选择, 否则就按顺序进行 $P8 \times 8$ 和 $Intra$ 模式检测。

当编码 $8 \times 4, 4 \times 8$ 模式的 $RDcost$ 都大于 8×8 模式时, 可终止模式搜索, 比较已编码的所有模式的 $RDcost$, 选出最优帧间模式。

(3) 简单有效的快速帧内选择

利用 $inter$ 预测和 $intra16$ 预测已得的信息判断是否可以跳过 $intra4$ 预测。由于对不同的视频序列及同一个序列的不同 QP 下编码时 $inter$ 和 $intra16$ 的 $RDcost$ 有所不同, 所以本文采取自适应的阈值的取值对 $intra4$ 预测进行判断, 当满足式 6 时则跳过 $intra4$ 预测。 ($QP < 18$ 时 T 取 1.5; $QP > 18$ 时, T 取 1.5)

$$intra16_rdcost > Inter_rdcost * T \quad (6)$$

改进后的 P 帧模式选择全过程如图 2 所示。

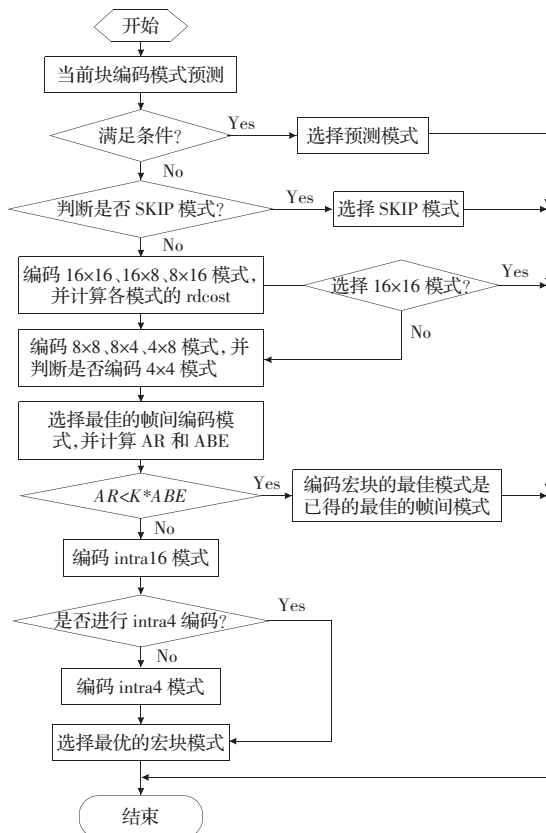


图 2 改进后的模式选择流程图

4 实验结果与分析

为了验证算法改进的有效性, 采用 JVT 参考软件 JM12.2^[5] 作为实验平台进行了大量的实验。实验使用 PIV 处理器, CPU 2.4 GHz, 256 M 内存, 在 VC++6.0 下进行, 针对不同运动类型的标准视频序列进行测试, 剧烈运动序列 Coastguard_cif, 中运动序列 Froeman_cif、Flower_cif, 小运动序列 Slient_cif。仿真的编码控制参数如表 1 所示。

表 1 仿真 JM 的编码控制参数

Name	Value	Name	Value	Name	Value
ProfileIDC	66	LevelIDC	40	EarlySkipEnable	1
UseHadamard	ON	SearchRange	32	SelectiveIntraEnable	1
NumberReferenceFrames	3	SymbolMode	CAVLC	SearchMode	2
RDOptimization	2	FrameRate	30	GOP structure	IPPP

分别利用上述的两种算法对不同视频序列的前 80 帧进行编码比较, 所得的实验结果如表 2 所示。

表 2 部分视频序列测试结果比较

视频序列	QP	Encoding Time/(%)	ME Time/(%)	Bitrate/(%)	PSNR _Y /dB	PSNR _U /dB	PSNR _V /dB
Silent_cif	16	-47.47	-62.90	0.926	0.31	0.20	0.02
	28	-43.02	-55.37	0.822	0.02	-0.01	-0.01
	36	-36.42	-46.08	0.246	0.02	-0.01	0.00
	40	-35.07	-44.68	1.711	-0.08	-0.01	-0.05
Foreman_cif	16	-46.74	-58.30	0.427	-0.05	-0.03	-0.04
	28	-49.07	-59.82	0.754	0.03	-0.04	-0.07
	36	-47.42	-55.68	0.665	0.04	-0.02	0.01
	40	-48.32	-55.93	0.709	0.03	-0.08	-0.08
Coastguard_cif	16	-55.94	-62.70	0.092	-0.08	-0.03	-0.04
	28	-48.82	-60.01	0.187	0.01	-0.05	0.02
	36	-48.09	-56.17	0.366	0.02	0.02	-0.12
Flower_cif	16	-49.11	-60.50	0.253	-0.02	-0.02	-0.03
	28	-49.55	-64.45	0.577	0.02	-0.04	-0.07
	36	-46.24	-61.14	0.829	-0.04	-0.14	-0.09
	40	-50.55	-62.36	0.645	-0.10	-0.08	0.27

表 2 给出了视频序列测试结果比较。改进的算法能平均降低 45%~60% 的运动估计时间和 35%~55% 编码时间, 对于剧烈运动的序列效果更为明显, 可见算法复杂度大大降低了。比特率会产生比较小的变化, 这是由于快速算法预测后的编码模式一些不是最佳模式, 视频编码后的比特数一般都会增加, 即快速的模式选择算法通常以提高码率为代价的。但这种较小的码率变化应用于实际视频编码、传输是基本不会造成影响。同时也可看出改进算法后 PSNR 的变化可以忽略不计, 甚至对于一些方向性运动比较强的序列 PSNR 有一定程度的增加。这说明, 改进的算法较好的保持了原算法率失真性能, 对图像质量影响很小。

5 结束语

本文对 H.264 近期所采用的快速模式选择算法进行了一定的研究和分析, 并在其原有算法基础上, 提出了进一步改进的方法。改进的算法充分考虑视频序列时域和空域的相关性, 利用了可靠的预测技术和早期停止判断标准, 大大减少了不必要模式的 $RDcost$ 计算, 降低了模式选择的复杂度。实验结果显示该算法能在保持率失真变化不大的情况下, 提高 35%~55% 的编码速度, 使得 H.264 能更容易在低码率高实时的应用系统中实现。

(下转 85 页)