

H. 264/AVC 中计算复杂度可调的快速模式选择算法

朱向军 朱善安

(浙江大学电气工程学院, 杭州 310027)

摘要 H. 264/AVC 视频编码标准采用多模式编码技术虽显著提高了编码效率,但由于该技术计算复杂度极高,因此为了降低计算复杂度,提出了一种计算复杂度可调的快速模式选择算法。该算法依据编码模式的特点和最佳模式分布规律,利用运动活性及已测试模式所得残差的量化系数来预测尚未测试模式的编码性能,并略过不必要测试的模式,通过选择恰当的阈值即可实现计算复杂度的调整。实验表明,与全模式选择算法相比,采用该算法编码速度可以提高 2~4 倍,而且重建质量和编码效率损失极小。

关键词 视频编码 H. 264/AVC 模式选择 编码效率

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2006)09-1210-07

Fast Mode Decision with Adjustable Computational-complexity for H. 264/AVC Encoder

ZHU Xiang-jun, ZHU Shan-an

(College of Electrical Engineering, Zhejiang University, Hangzhou 310027)

Abstract H. 264/AVC video coding standard employs multiple coding-modes and significantly improves coding efficiency at the cost of heavy computational load. To reduce the computation burden, a fast mode decision algorithm with adjustable computational-complexity is presented. According to the analysis of the characteristics of the coding-modes and statistical characteristics about the best mode, the proposed algorithm exploits motion activity and the quantized transform coefficients of the residue obtained by the tested coding-modes, to effectively predict the coding efficiency of the remaining modes and skip the unnecessary modes, further more thresholds used in this algorithm can be adjusted to regulate computational complexity. Experimental results show that the speed of the proposed method is about 2~4 times faster than the full-mode decision algorithm without any noticeable loss of image quality and coding efficiency.

Keywords video coding, H. 264/AVC, mode decision, coding efficiency

1 引言

H. 264/AVC 视频编码标准^[1]是 ISO/IEC MPEG(ISO 运动图像专家组)与 ITU-T VCEG(视频编码专家组)联合提出的最新视频编码标准。H. 264/AVC 沿用了以往标准采用的运动估计/补偿加变换的混合编码技术,由于其采用了块尺寸可变的多参考帧运动估计、多模式帧内预测以及上下文

自适应的算术编码等新技术,因此与以往标准相比,编码效率可以提高 1.5~2 倍^[2]。采用这些新技术的基本思想是,充分考虑到视频信号的非平稳性,通过提供多种可选的编码模式来适应不同时刻不同区域信号的特征,以获得更高的编码效率。为获得率失真最优的编码性能,可采用基于率失真优化(rate-distortion optimization, RDO)的全模式选择方法(RDO-full mode decision, RDO-FMD)^[3],通过比较所有模式下编码所得的率失真代价(rate-distortion

cost, RDC)来选择率失真代价最小的模式,但是却大幅度地增加了编码器的计算复杂度,因为这种基于RDO的全模式选择与运动估计过程占编码器计算量的80%以上。

为降低编码器的复杂度,目前已有一些快速模式选择策略^[4-9]。其基本思想是先通过分析宏块的时域和空域特征来筛选出可提供较好率失真性能候选模式,然后仅针对这些候选模式进行基于RDO的运动估计和编码,最后从中选择率失真代价最小的编码模式。根据预测可选模式时所利用信息的差异,现有的快速模式选择算法大体可以分为以下两类:一类是基于源视频信号特征分析的方法,其中包括基于边缘方向直方图的帧内预测模式筛选的方法^[4]和先利用帧差确定静止区域或亮度一致区域^[5,6],再利用边缘图检测出亮度变化较小的宏块^[5],而对此类宏块仅选择块尺寸较大的模式的方法,此类方法的性能依赖于分类的准确性和视频信号,其不仅没有考虑量化参数对模式选择的影响,且特征分析也需要额外的计算量和存储空间;另一类是基于已编码模式所得信息及各模式间相关性的方法,其中有首先比较几个代表性模式的率失真代价,然后选出可能具有较好编码性能的候选模式^[7]的方法,根据SKIP模式及最佳帧间模式的编码效果提前终止模式选择过程的方法^[8],还有根据已测试模式下所得全零系数块的特征分析,跳过不必要的编码模式的方法^[9]。由于此类算法利用了已编码模式的结果,因此对不同特征的视频信号和目标码率变化具有较好的适应性。

针对现有的快速模式选择算法对编码速度提高有限或引起编码效率降低的问题,本文提出了一种计算复杂度可调的快速模式选择算法(complexity adjustable fast mode decision, CA-FMD),该算法首先分析了各模式的特点和最佳模式分布规律,然后提出了基于运动活性分析的SKIP模式预判技术,以检测出静止区域及运动一致区域,最后利用已测试模式下的系数价值来预测后续模式的相对编码性能,并提出了计算复杂度可调的模式筛选规则,以略过不必要测试的模式。本文还通过仿真实验测试了采用该算法的编码器的编码效率和计算复杂度。

2 H.264/AVC 的 P 帧编码模式

H.264/AVC 标准^[1,2]采用的 P 帧编码模式包

括 8 种帧间编码模式和 13 种帧内预测模式。其中,帧间编码模式与运动补偿中采用的宏块(macroblock, MB)分裂方式相对应,与宏块分裂方式对应的亮度块尺寸分别为 16×16 、 16×8 、 8×16 或 8×8 ,而在 8×8 的宏块分裂方式下,还允许每个 8×8 子宏块(Sub-MB)分裂为 8×4 、 4×8 或 4×4 的块,也即形成了如图 1 所示的宏块级与子宏块级的分裂方式,此外, H.264 标准还采用了 SKIP 模式,以有效地编码大面积的静止区域和运动一致区域,帧内预测模式(INTRA)包括 16×16 与 4×4 两种模式类型,其中 16×16 模式以 16×16 亮度块为预测单位,支持 4 种预测模式,适合亮度变化平缓的宏块,而 4×4 模式则以 4×4 亮度块为预测单位,支持 9 种预测模式,适合纹理复杂的宏块。

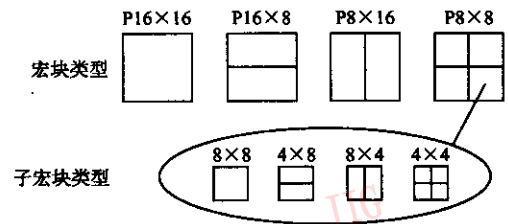
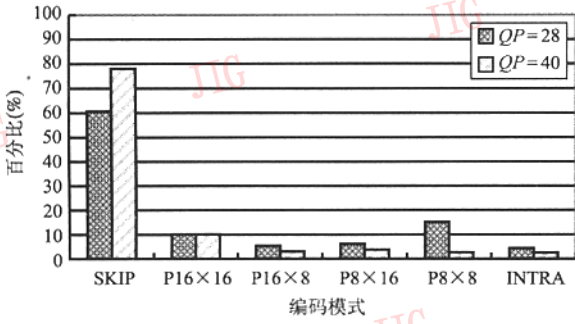


图 1 运动补偿中采用的宏块分裂方式

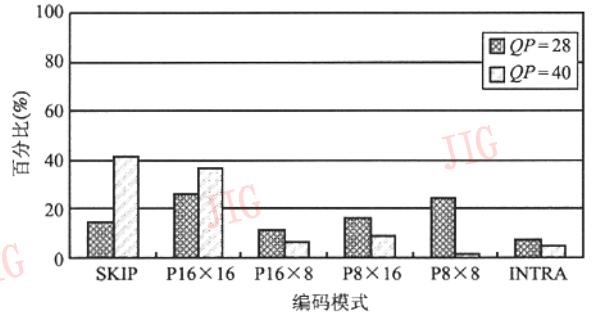
Fig. 1 Segmentation of MB for motion compensation

针对不同纹理复杂度和运动特征的序列,在不同量化参数(quantization parameter, QP)下统计了最佳编码模式选择概率(如图 2 所示,在第 5 节的实验条件下,采用 RDO-FMD 算法)。其中, News 序列包含大面积静止区域,但也具有运动和纹理都很复杂的小尺寸物体; Foreman 序列具有中等程度的全局运动,人脸区域纹理和运动复杂。由图 2 可见,两类视频序列的最佳模式分布具有以下共同特点: (1)在 P 帧中采用帧内预测模式的概率很小; (2)SKIP 与 16×16 模式所占比例较大, 8×8 模式相对较少; (3)随着量化参数增大,采用 16×8 、 8×16 与 8×8 模式为最佳模式的概率明显下降。由此可见,最佳模式的选择与信号特征密切相关,即较大块尺寸的模式更适宜纹理平坦或运动一致的区域,而较小块尺寸的模式则对具有复杂纹理和运动细节的区域更有效。

因此,基于各编码模式的特点和最佳模式的分布规律,通过分析视频信号特征或已测试模式下的编码结果来预测后续模式的编码性能,以略过不可能进一步降低率失真代价的模式,可以显著降低编码过程的计算复杂度。



(a) News



(b) Foreman

图 2 采用 RDO-FMD 算法的最佳模式选择结果

Fig. 2 Statistic results of best modes with RDO-FMD algorithm

3 宏块的运动活性与 SKIP 模式预判

在实际视频场景中,通常会存在静止背景或瞬时静止区域及运动一致的物体,由于对此类区域内的宏块采用 SKIP 模式即可获得较好的编码效率,因此可以依据宏块的运动活性来检测出此类宏块。鉴于空域和时域相邻宏块的运动信息具有很强的相关性,因此宏块的运动活性可由相邻块的运动信息估计。

将当前编码宏块 B 在空间上的左、上和右上相邻宏块分别记为 $B_1 \sim B_3$ (若右上相邻宏块不存在,则用左上相邻宏块代替),将前一帧中相同位置的宏块记为 B_4 。若左相邻和上相邻宏块存在,且它们的最佳模式都属于模式集合 $\{SKIP, P16 \times 16, P16 \times 8, P8 \times 16\}$,则认为相邻块间的运动矢量 (motion vector, MV) 相关性较大,反映运动活性的局部运动矢量的大小 L^{MV} 与差异性 D^{MV} 定义如下:

$$L^{MV} = \max_{i=1,2,3,4} \{|v_{x_i}|, |v_{y_i}|\} \quad (1)$$

$$D^{MV} = \max_{i=1,2,3,4} \{|v_{x_i} - \bar{v}_x|, |v_{y_i} - \bar{v}_y|\} \quad (2)$$

$$\bar{v}_x = \frac{1}{4} \sum_{i=1}^4 v_{x_i}, \bar{v}_y = \frac{1}{4} \sum_{i=1}^4 v_{y_i} \quad (3)$$

其中, v_{x_i} 与 v_{y_i} 分别为第 i 个宏块 B_i 中与当前编码宏块紧邻的 8×8 子块的运动矢量的水平和垂直分量, \bar{v}_x 与 \bar{v}_y 分别为相邻块运动矢量水平和垂直分量的均值。若 L^{MV} 较小,则当前编码宏块可能属于静止区域;若 D^{MV} 较小,则当前编码宏块可能属于运动一致性强的视频物体。

基于以上分析,给出以下 SKIP 模式预判规则:若当前编码宏块满足 $L^{MV} < 1$ 或 $D^{MV} < 1$,并且在 SKIP 模式下运动补偿后残差足够小(如宏块的“系

数值 γ (下一节中将介绍)不大于 2),则跳过其他帧间模式的测试。表 1 给出了在量化参数为 30 时,用此判据预判的检测率和准确率(实验条件见第 5 节)。检测率为根据上述判据正确检测到的 SKIP 模式的宏块数与所有最佳帧间模式为 SKIP 模式的宏块数之比,准确率为满足上述判据条件的宏块中最终选择 SKIP 模式为最佳帧间模式的比例。对于具有大面积瞬时静止或运动一致区域的序列 (News、Silent 与 Paris 等),由于此方法的检测率可达到 40% 左右,而且准确率高于 95%,因此,采用 SKIP 模式预判能够显著降低计算量。对于其他运动复杂的序列,检测率在 15% 左右,准确率也不够高(80% 左右),但是由于此类序列采用 SKIP 模式的概率较低(如图 2(b)所示),因此,误判对视频重建质量影响也很小。综上所述,若宏块满足 SKIP 模式预判条件,则省略对其他帧间模式的测试是可行的。

表 1 SKIP 模式预判的准确性 (QP = 30)

Tab. 1 Accuracy of the SKIP mode early-decision rule

序列	News	Foreman	Silent	Paris	Tempete	Mobile
检测率 (%)	43.9	13.6	36.9	39.3	27.9	19.4
准确率 (%)	97.3	85.0	98.0	98.1	78.1	91.2

4 基于系数价值分析的模式筛选

帧间模式编码性能的优劣取决于运动估计的准确性,由于运动估计的有效性可以用运动补偿所得残差的能量和编码运动信息所需比特数来度量,因此,可以依据已测试的帧间模式所得的残差来评价

已测试模式是否满足编码效率要求,并预测尚未测试模式的相对编码性能。

4.1 系数价值

实践表明,运动补偿所得残差决定了重建信号的失真度,而残差的变换系数经量化后所得的系数(简称“系数”)则决定了纹理编码所需的比特数,也反映了在特定量化参数下残差能量的相对大小。参考软件 JM8.4^[10]从率失真性能角度分析了系数的价值,并定义了宏块内索引为 k 的 8×8 子块与整个宏块的“系数价值”,分别记为 $C^{(8)}$ 和 $C^{(8)}$:

$$C_k^{(8)} = \sum_{i=0}^3 \sum_{j=0}^{15} \alpha(i, j), C^{(16)} = \sum_{k=0}^3 C_k^{(8)} \quad (4)$$

$$\alpha(i, j) = \begin{cases} \infty & l > 1 \\ 3 & l = 1, \text{ 且 } N = 0 \\ 2 & l = 1, \text{ 且 } 1 \leq N \leq 2 \\ 1 & l = 1, \text{ 且 } 3 \leq N \leq 5 \\ 0 & \text{其他} \end{cases} \quad (5)$$

其中 i, k 分别为 4×4 块和 8×8 块的索引, j 为按迂回方式扫描 4×4 块系数的索引, $\alpha(i, j)$ 即为“系数价值”, l 为系数的幅值, N 为其前面连续零系数的个数。通过大量实验发现,若子块或宏块中幅值很小($l = 1$)且孤立的系数所需比特数较大,则相对于提供的率失真性能来说,其价值不大。因此,为节省比特数,可强制把这些系数置零,仅引入极小的质量损失,称其为系数阈值化^[9,10]。可见,相对于残差能量,“系数价值”更适于评价编码模式在特定码率下的编码效率。

4.2 计算量可调的模式筛选规则

根据对应的亮度块尺寸,可以把所有帧间模式分为 SKIP、 $G_{16} = \{P16 \times 16\}$ 、 $G_{MBpart} = \{P16 \times 8, P8 \times 16\}$ 、 $G_8 = \{\text{subMB}8 \times 8, \text{subMB}8 \times 4, \text{subMB}4 \times 8, \text{subMB}4 \times 4\}$ 等 4 个等级。在“Foreman”序列中,各等级模式下编码宏块的头信息(模式、运动矢量及参考帧索引等)所需比特数的统计结果为 0bits、11.7bits、22.8bits 和 56.2bits。可见,在此等级划分下,编码头信息所需比特数依次增加两倍以上,在残差能量很小时,若采用后续的较小块尺寸模式,则会显著增加比特数开销,而且对率失真代价的降低也很有限。在特定模式下,各 8×8 子块的“系数价值”的最大值记为 $C_{\max}^{(8)} = \max_{k=0,1,2,3} \{C_k^{(8)}\}$,若满足条件

$$C_{\max}^{(8)} \leq T \quad (6)$$

则可以认为此编码模式具有较好的编码效率,其中 T 为阈值。

基于以上分析,下面给出与文献[9]类似的两个模式筛选规则:

决策 I:若在 SKIP 模式下满足式(6),由仅测试 $P16 \times 16$ 模式,并从候选模式集 $\{\text{SKIP}, P16 \times 16\}$ 中选择 RDC 最小的模式作为最佳帧间模式,此时式(6)的阈值记为 T_1 ;

决策 II:若在 SKIP 模式或 $P16 \times 16$ 模式下,满足式(6),则仅测试 G_{MBpart} 级模式,而省略 G_8 级模式的测试,并在候选模式集 $\{\text{SKIP}, G_{16}, G_{MBpart}\}$ 中选择 RDC 最小的模式作为最佳帧间模式,此时式(6)的阈值记为 T_2 。

与文献[9]中的全零系数块判据不同,本文采用式(6)的模式筛选条件,通过调整阈值引入了计算量可调机制。若阈值越大,则有越多的宏块满足上述决策条件,这不仅使检测率增大,而且节省的计算量也越大,但同时误判率也会增大,从而引起一定的质量损失。此处,检测率是指 RDO-FMD 算法所得最佳帧间模式属于候选模式集合的宏块满足式(6)的比例;而误判率则为依据规则选择的最佳模式与 RDO-FMD 算法结果不一致的宏块占符合决策条件的宏块的比例。这种阈值变化对检测率、误判率及编码效率的影响,可以反映上述决策机制的特点。

大量实验显示,对运动较小的序列,由于其检测率和检测准确性均高于运动快速复杂的序列,因此,选择运动较大、较复杂的“Foreman”序列进行编码测试(量化参数为 30,其他测试条件见第 5 节)具有很好的代表性,上述模式筛选规则在不同阈值下的检测率和误判率的统计结果分别见图 3 与图 4。对于 SKIP 模式预测效果不佳($C_{\max}^{(8)}$ 值较大)的宏块,采用 $P16 \times 16$ 模式则可能会取得最佳编码性能,因此,如图 3 显示,“决策 I”的检测率最大只达 60% 左右。由于“决策 I”的误判率随阈值增大快速上升,考虑到保持编码效率比计算量的小幅度减小更重要,因此阈值 T_1 的取值限定在 4 以内。如图 4 所示,“决策 II”可以达到 85% 左右的检测率,而误判率在 10% 以内,而阈值为 10 时,在 80% 的检测率下仅有 6% 左右的误判,并且在 $T_2 \leq 10$ 时,随着阈值增大,检测率提高明显。由此可见,利用 16×16 块尺寸下的运动补偿效果来预测 G_8 等级模式的编码效率具有很高的准确性。图 5 给出了因模式筛选规则的误判而引入的率失真代价的相对增量(即相对于 RDO-FMD 所得最小 RDC 的增量)。由图 5 可见,RDC 的相对增量可以控制在 10% 以内。“决策

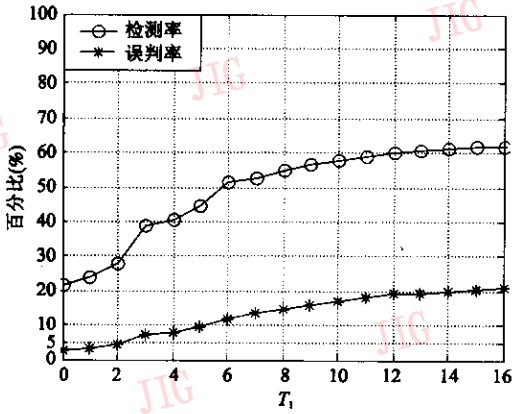


图 3 “决策 I”的统计结果(“Foreman”)

Fig. 3 Static results of “Decision I” (“Foreman”)

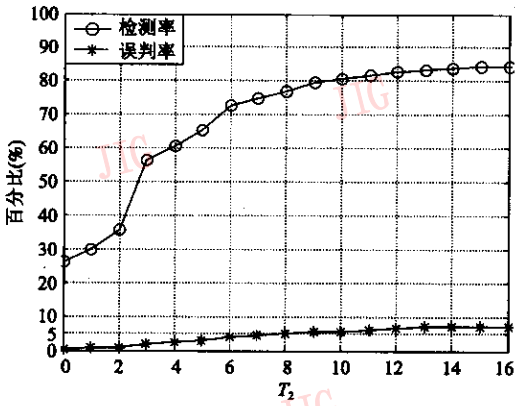


图 4 “决策 II”的统计结果(“Foreman”)

Fig. 4 Static results of “Decision II” (“Foreman”)

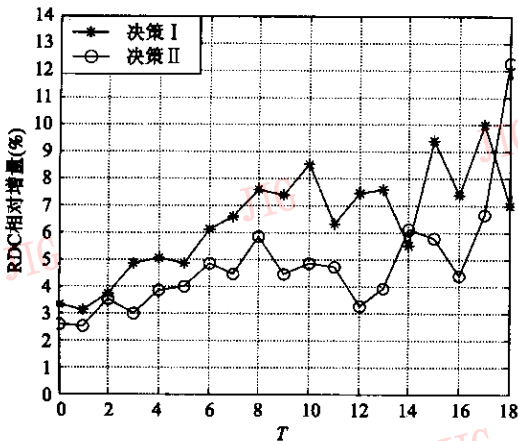


图 5 决策误判引起的编码效率损失(“Foreman”)

Fig. 5 Loss of coding efficiency caused by false decision (“Foreman”)

II”引入的编码效率损失较小,当 $T_2 \leq 10$ 时,编码效率损失在6%以内。

类似于以上两种模式筛选规则的分析,在测试 G_MBpart 级模式后,若所得最佳模式为 $P16 \times 8$ 或 $P8 \times 16$,则不仅各子块的“系数值”较小($C^{(8)} \leq 4$),而且整个宏块的“系数值”也较小($C^{(16)} \leq 5$),即跳过其他帧间模式;若子宏块在 $subMB8 \times 8$ 模式下所得的“系数值”较小($C^{(8)} \leq 4$),则可以省略 $subMB4 \times 4$ 模式的测试,以进一步降低计算量。

4.3 帧内预测模式的自适应筛选

由于帧间模式的选择结果可以反映信号的空间特征^[9],因此可结合帧内预测模式的特点来决定候选帧内模式。由于满足第3节给出的SKIP模式预判准则的宏块可能属于亮度缓变区域,因此有必要测试 $I16 \times 16$ 模式,而且由于满足筛选规则“决策I”或“决策II”的宏块已经具有较好的编码效果,而 $I4 \times 4$ 模式所需比特数又较大,采用此模式的可能性极小,因此可仅测试 $I16 \times 16$ 模式。由于最佳帧间模式为 $P8 \times 8$ 的宏块具有复杂的纹理,而采用 $I16 \times 16$ 模式的可能性又极小,因此仅测试 $I4 \times 4$ 模式。此外,若 $I16 \times 16$ 模式下宏块的“系数值”较小($C^{(16)} \leq 5$),则可省略 $I4 \times 4$ 模式的测试。

5 实验结果

为测试本文算法的性能,在JVT的参考软件JM8.4^[10]上实现了CA-FMD算法,并对其进行了实验测试,实验条件为:在基本框架(baseline profile)下采用RDO、SATD、快速运动估计^[11]及码率控制^[12],运动估计搜索范围为16pixels,1/4pixels精度,参考帧数为5,首帧为I帧(帧内编码),其余各帧都为P帧。如表2所示,测试序列的纹理和运动复杂度各不相同,格式有QCIF和CIF,帧率也有差异,并在较高和较低的目标码率下编码,以充分测试算法的性能,其中 QP_1 为I帧编码的量化参数。试验中采用以下指标来比较CA-FMD算法与基于RDO的全模式选择算法(RDO-FMD)的性能:①两种算法下亮度信号的峰值信噪比(peak signal-to-noise ratio, PSNR)之差,记为 $\Delta PSNR$,用于度量在相同目标码率下编码效率的差异;② Δt 为采用CA-FMD算法编码较采用RDO-MD算法编码相对减少的时间比例,其计算公式分别为

$$\Delta PSNR = PSNR_{CA-FMD} - PSNR_{RDO-FMD}$$

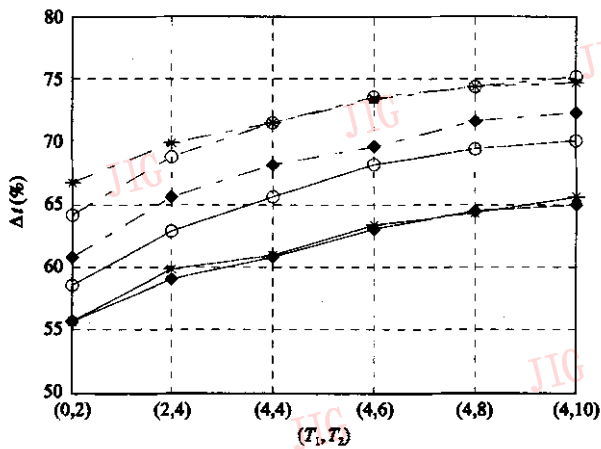
$$\Delta t = \frac{t_{RDO-FMD} - t_{CA-FMD}}{t_{RDO-FMD}} \times 100[\%]$$

表 2 测试序列
Tab. 2 Test sequence

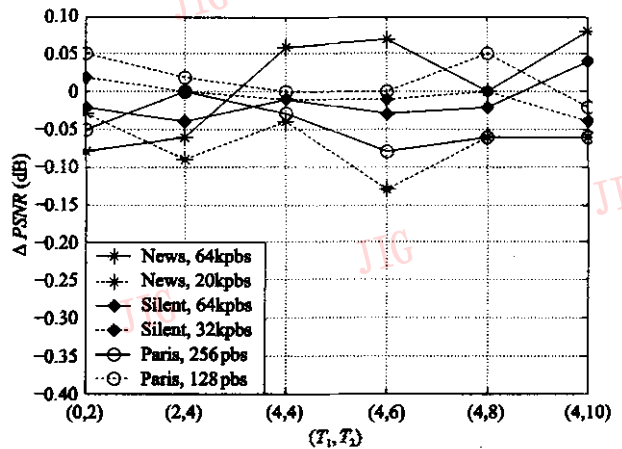
序列	格式 帧率 (fps)	帧数	码率 (kbps) $\times (QP_1)$	
			较高	较低
News, Foreman	QCIF 10	100	64 (28)	20 (36)
Silent	QCIF 15	150	64 (28)	32 (32)
Paris	CIF 15	150	256 (30)	128 (36)
Tempete, Mobile	CIF 30	100	1024 (28)	512 (32)

图 6 和图 7 分别给出了中低/高运动复杂度序列的编码时间的降低百分比及峰值信噪比损失与两个模式筛选规则的阈值组合 (T_1, T_2) 的关系。在 4.2 节中关于阈值对计算量和编码效率影响进行分

析的基础上,可以选取如下两个模式筛选规则的阈值,以实现计算量的调节: $T_1 \in \{0, 2, 4\}, T_2 = \{2, 4, 6, 8, 10\}$ 。图 6 与图 7 给出了代表性阈值组合的实验结果。实验结果表明,本文算法可以显著降低编码复杂度,同时可较好地保持编码效率。其中,对中低运动复杂度的序列,采用改算法可以降低 65% ~ 75% 的计算量;对运动复杂度高的序列,编码时间降低虽相对较小,但也可降低 50% ~ 70% 的计算量。形成以上差异的主要原因是:对具有不同信号特征的序列,较大块尺寸模式的编码效率不同,图 2 给出的最佳模式分布统计也证明了这一点,并且,SKIP 模式预判与模式筛选规则的检测率也有差异(如表 1 所示)。CA-FMD 算法引入的峰值信噪比下降平



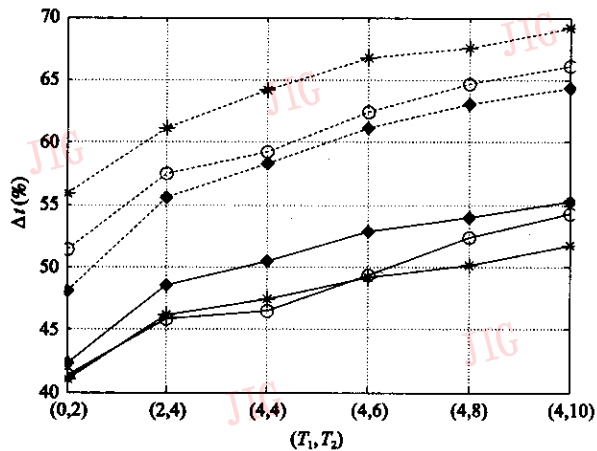
(a) 不同阈值下各测试序列 Δt 的变化曲线 (图例同图 6(b))



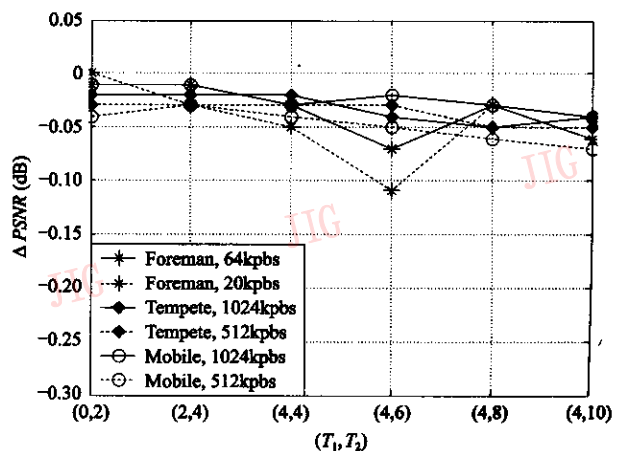
(b) 不同阈值下各测试序列 $\Delta PSNR$ 的变化曲线

图 6 在不同阈值下 CA-FMD 算法对中低运动复杂度测试序列的 Δt 及 $\Delta PSNR$

Fig. 6 Δt and $\Delta PSNR$ versus (T_1, T_2) for CA-FMD with low and medium motion complexity sequences



(a) 不同阈值下各测试序列 Δt 的变化曲线 (图例同图 7(b))



(b) 不同阈值下各测试序列 $\Delta PSNR$ 的变化曲线

图 7 在不同阈值下 CA-FMD 算法对高运动复杂度测试序列的 Δt 及 $\Delta PSNR$

Fig. 7 Δt and $\Delta PSNR$ versus (T_1, T_2) for CA-FMD with high motion complexity sequences

均小于 0.05dB,峰值信噪比最大下降在 0.15dB 以内,小于人眼视觉系统可感知的阈值(0.2dB)。通过对基于两个模式筛选规则的计算量调节结果的分析可见,在测试的阈值组合范围内,计算量具有 10% 左右的调节空间。对纹理复杂的快速运动序列,计算量的调节范围较大,其在较低码率下的调节范围可以达到 15% 以上。

至于阈值变化对图像重建质量的影响,需要指出的是,实验中,由于码率控制算法以一行宏块为单元,且所有测试条件下,实际码率的波动在目标码率的 0.5% 以内,因此可以排除实际码率对图像重建质量变化的影响。由图 7 可见,对运动复杂度高的序列,随着阈值增大,峰值信噪比大体上呈逐渐下降趋势,这主要是由于在运动细节较多的区域,CA-FMD 算法带来了一定的编码效率损失,即图 5 所示的率失真代价的增大。图 6 表明,针对中低运动复杂度的序列, $\Delta PSNR$ 随阈值的变化具有较大波动,并且对于运动细节较多的 News 序列,其波动相对更大,波动范围达到 0.15dB。对于 News 序列,在较大的阈值组合(4, 10)下,峰值信噪比却有 0.08dB 的提高。从比特数分配的角度来看,由 4.2 小节中关于不同模式下编码宏块头信息所需比特数的统计结果可见,CA-FMD 算法产生的误判实际上节省了系数价值较小(失真较小)区域的比特数,若将其用于复杂度高的区域,需要更多的比特数,则能更好地体现比特数的价值,并可能会提高峰值信噪比。

以上实验结果对实际应用中阈值的选择提供了有益的参考。对于中低运动复杂度的视频场景,由于阈值增大并不一定引起峰值信噪比的下降,因此选取较大的阈值是可行的,而对高速运动或运动复杂的序列,则可以根据编码器的运算能力适当调节阈值,但阈值增大会引起一定的图像质量损失。虽然可以依据视频序列的运动特征和编码器实现平台的运算能力以及功耗约束来动态地自适应调整阈值,但鉴于功耗-率失真分析的复杂性,在此不详论述。

6 结 论

本文针对率失真优化框架下因 H.264/AVC 编码器的复杂度太大而不易于实时应用的问题,提出了一种计算复杂度可调的快速模式选择算法。该算法首先通过估计宏块的运动活性来检测出静止及运动一致的区域,然后分析已测试模式的残差及帧间

模式的编码结果,并给出了由帧间到帧内模式及由大块尺寸模式到小块尺寸模式依序逐级筛选的机制。该算法实现简单,且不需要额外的存储空间。实验结果表明,采用 CA-FMD 算法的编码器不仅具有与原编码器相似的编码效率,而且编码速度提高了 2~4 倍。由于通过调整系数价值阈值易于调节计算复杂度与编码效率,因此能满足特定应用场景下对能耗、实时性及压缩比的要求。但是,宏块的模式选择与码率分配的联合优化及基于功耗-率失真分析的自适应阈值调整机制尚需进一步研究。

参考文献(References)

- 1 Joint video team(JVT) of ISO/IEC MPEG and ITU-T VCEG. Draft ITU-T Rec. and FDIS of Joint Video Specification[S]. H.264/ISO/IEC 14496-10 AVC, 2003.
- 2 Wiegand T, Schwarz H, Joch A, et al. Rate-constrained coder control and comparison of video coding standards[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7):688~703.
- 3 Sullivan G J, Wiegand T. Rate-distortion optimization for video compression[J]. IEEE Signal Processing Magazine, 1998, 15(6):74~90.
- 4 Pan F, Lin X, Susanto R, et al. Fast mode decision algorithm for Intra prediction in JVT, Doc. JVT-G013[A]. In: 7th JVT meeting [C], Pattaya, Thailand, March 2003.
- 5 Lim K P, Wu S, Wu D J, et al. Fast INTER mode selection, Doc. JVT-I020[A]. In: 9th JVT meeting[C], San Diego, CA, USA, September 2003.
- 6 Jing X, Chau L P. Fast approach for H.264 inter mode decision[J]. Electronics Letters, 2004, 40(17):1051~1052.
- 7 Yin P, Tourapis H C, Touapis A M, et al. Fast mode decision and motion estimation for JVT/H.264[A]. In: Proceedings of the International Conference on Image Processing[C], Barcelona, Spain, 2003:14~17.
- 8 Jeon B, Lee J. Fast mode decision for H.264, Document JVT-J033[A]. In: 10th JVT Meeting[C], Waikoloa, Hawaii, USA, December 2003.
- 9 Kim Y H, Yoo J W, Lee S W, et al. Adaptive mode decision for H.264 encode[J]. Electronics Letters, 2004, 40(19):1172~1173.
- 10 Joint video team(JVT) of ISO/IEC MPEG and ITU-T VCEG. JVT Reference Software unofficial version JM8. 4[CP/OL]. <http://iphome.hhi.de/suehring/tml/download/>.
- 11 Chen Z, Zhou P, He Y. Fast motion estimation for JVT, Document JVT-G016[A]. In: 7th JVT meeting[C], Pattaya, Thailand, March 2003.
- 12 Li Z G, Pan F, Lim K P, et al. Adaptive basic unit layer rate control for JVT, Document JVT-G012[A]. In: 7th JVT meeting[C], Pattaya, Thailand, March 2003.