

H.264/AVC 中基于编码复杂度的率失真优化算法

王晓龙, 武静, 张艳芳

(中国航空工业集团公司洛阳电光设备研究所, 河南 洛阳 471000)

摘要: 提出了一种基于复杂度的率失真联合优化算法, 该算法在 H.264/AVC 视频编码标准中率失真优化算法的基础上考虑了视频编码算法复杂度因素, 以求达到对视频编码复杂度和率失真性能联合优化的效果, 并通过引入复杂度控制因子对编码复杂度进行控制调节。实验结果表明, 该算法可以在保证率失真性能降低很小的情况下降低视频编码的复杂度, 在率失真和复杂度之间取得了很好的均衡效果。

关键词: 视频编码标准; H.264/AVC; 率失真优化; 编码复杂度

中图分类号: V271.4; TN919.8 **文献标志码:** A **文章编号:** 1671-637X(2014)10-0057-04

A Coding Complexity Based Rate-Distortion Optimization Algorithm for H.264/AVC

WANG Xiao-long, WU Jing, ZHANG Yan-fang

(Luoyang Institute of Electro-Optical Equipment, AVIC, Luoyang 471000, China)

Abstract: A complexity based joint rate-distortion optimization algorithm is presented, which taking the complexity factor of video coding into account based on the rate-distortion optimization algorithm in H.264/AVC video coding standard. Thus a better effect can be obtained on the tradeoff between rate-distortion performance and video coding complexity. Besides that, this framework introduces a complexity control factor to control the extent of complexity. Tests show that the proposed framework can reduce the complexity without much sacrifice in coding efficiency.

Key words: video coding standard; H.264/AVC; rate-distortion optimization; coding complexity

0 引言

H.264/AVC 视频编码标准是由 ITU-T/MPEG 联合组织于 2003 年推出的视频编码标准^[1], 该标准在保留传统编码标准基本框架的基础上, 引入了很多新的先进编码技术, 比如可变块尺寸运动估计^[2]以及内容自适应的二进制算术编码等编码选项^[3]。这些技术的引入使得 H.264/AVC 标准取得了很好的率失真性能, 与之前的视频编码标准相比较, 在输出同等视频编码质量的情况下节省了一半码率。但是, 压缩性能的提高是以高复杂度为代价的, 其中运动估计部分是整个编码流程中耗时最大的一个模块, 因为这部分集中了多参考帧选择、可变宏块尺寸运动估计和分数像素插值等运算量大的编码技术。高复杂度不利于视频编码算法的实时应用, 为了解决这个问题, 很多专家学者针

对不同的模块提出了各种快速算法^[4-9]。

这些算法通过使用相关性预测、提前退出等方法可以在尽可能少地降低率失真性能的情况下提高编码算法的运行速度, 并取得了很好的效果。但是, 在实际系统应用中, 由于视频编码算法要受到实时性和系统资源等方面的限制, 需要在编码效率和编码复杂度之间寻找一个均衡。为了能够在两者之间取得有效折中, 需要建立一个复杂度率失真联合优化模型。首先, 并不是所有的宏块都有必要对各种编码模式进行遍历, 然后选择合适的预测模式, 全遍历虽然可以保证编码效率最高, 但是运算复杂度代价过大, 可以对编码模式进行选择性的遍历, 尽量用较小的复杂度达到同等效果的编码效率; 其次, 即使是在进行选择性的遍历的时候, 对于不同的编码模式, 考虑到实际应用的需求和限制, 在进行率失真优化的过程中, 不能只关注编码模式的失真和码率因素, 也要考虑复杂度因素。为此, 本文建立了一个复杂度率失真联合优化模型, 综合考虑编码效率和复杂度控制来选取合适的编码模式。

1 H.264/AVC 标准中的率失真优化模型

香农在 1959 年提出了信息率失真函数 $R(D)$,即在允许一定的失真度 D 的情况下,信源输出的信息率可以压缩到 $R(D)$ 。上述定理可以表示为

$$\min\{D\}, \quad \text{s. t.} \quad R \leq R_c \quad (1)$$

式中: D 代表失真度; R 为码率; R_c 为信道传输速率。根据率失真理论,当前帧编码模式的选择问题可以归结为:在可选编码模式中选择最佳的模式,使得在限定码率 R_c 下,失真度 D 最小。

假设 $M = \{m_1, m_2, \dots, m_n\}$, 表示当前帧所有宏块, F 表示视频编码可选的编码模式集,那么当前帧的率失真优化编码就可以表示为

$$\min\{D(M, F)\}, \quad \text{s. t.} \quad R(M, F) \leq R_c \quad (2)$$

式中: $D(M, F)$ 和 $R(M, F)$ 分别表示当前帧在编码模式集 F 下的失真和编码比特数; R_c 为限定码率,意义同式(1)中的 R_c 。因为当前视频标准是基于宏块的编码结构^[1],所以当前帧的失真或码率可以由各个宏块的失真或比特数累加得到,算式为

$$D(M, F) = \sum_{i=1}^N D(m_i, F) \quad (3)$$

$$R(M, F) = \sum_{i=1}^N R(m_i, F) \quad (4)$$

采用拉格朗日乘子法可求解约束条件下极值,式(2)可转换为

$$\min\{J(M, F, \lambda)\} \quad (5)$$

其中

$$J(M, F, \lambda) = D(M, F) + \lambda \cdot R(M, F) \quad (6)$$

式中, λ 是拉格朗日算子。显然,只要参数 λ 可知,式(5)的极值就可以决定出最优编码模式 F' ,使得此时的编码失真 $D(M, F)$ 就是在限率 R_c 下的最小失真。

文献[10-11]中依照 R-D 的近似数量关系和大量的实验结果,得出参数 λ 与量化步长 Q^2 有着近似的线性关系,即

$$\lambda_{\text{mode}} = 0.85 \times Q^2 \quad (7)$$

式中, $Q = Q_p/2$, Q_p 为量化参数,该式被用于 H.263 标准。而在 H.264/AVC 中由于编码参数的改变,对此进行了改进,在 H.264/AVC 中 λ 和量化参数 Q_p 的关系如下所述:

1) 如果没有 B 帧编码,对于 I 帧和 P 帧则采用

$$\lambda_{\text{mode}} = 0.85 \times 2^{\frac{Q_p-12}{3}}; \quad (8)$$

2) 如果有 B 帧编码,对于 I 帧和 P 帧则采用

$$\lambda_{\text{mode,ip}} = 0.68 \times 2^{\frac{Q_p-12}{3}} \quad (9)$$

而 B 帧则采用

$$\lambda_{\text{mode,b}} = 0.68 \times 2^{\frac{Q_p-12}{3}} \times \max\{2.00, \min\{4.00, \frac{Q_p}{6}\}\} \quad (10)$$

另外,H.264/AVC 标准中采用了运动向量预测和多参考帧技术,为了让运动估计获得最佳运动信息,同样采用率失真优化技术。运动估计失真 D_{motion} 采用 SAD 度量,而不是编码模式决定采用的 SSD,因此,近似认为拉格朗日系数 λ_{motion} 、 λ_{mode} 二者关系^[11]为

$$\lambda_{\text{motion}} = \sqrt{\lambda_{\text{mode}}} \quad (11)$$

一旦确定了 λ_{mode} ,即可通过式(9)得到 λ_{motion} 。

2 基于编码复杂度的率失真联合优化模型

本节提出了基于编码复杂度的率失真联合优化模型,在选择编码模式的同时考虑率失真和复杂度两方面的因素。本模型的建立主要包括以下 3 个方面:1) 建立复杂度率失真联合优化方程;2) 量化衡量编码复杂度;3) 计算复杂度控制因子。下面分别进行介绍。

2.1 复杂度率失真联合优化方程

为了使视频编码算法在进行模式选择的过程中综合考虑复杂度和率失真两方面的因素,根据标准中的率失真优化方程给出一个复杂度率失真联合优化方程为

$$J'(M, F, \lambda, \lambda_c) = D(M, F) + \lambda \cdot R(M, F) + \lambda_c \cdot C(M, F) \quad (12)$$

式中: $J'(M, F, \lambda, \lambda_c)$ 为复杂度率失真联合优化代价函数; λ_c 为复杂度控制因子,也是复杂度拉格朗日乘子; $C(M, F)$ 为复杂度函数。在编码过程中遍历各种编码模式,满足 $\min\{J'(M, F, \lambda, \lambda_c)\}$ 的编码模式就是最佳选择。通过式(12)可提前淘汰复杂度代价过高的块尺寸编码模式,使得最终选择的编码模式尽可能同时满足对复杂度和率失真的需求。根据编码模式种类的不同, λ_c 的两种表达形式为 $\lambda_{c,\text{motion}}$ 和 $\lambda_{c,\text{mode}}$,在此采用 H.264/AVC 标准中的关系式: $\lambda_{\text{motion}} = \sqrt{\lambda_{\text{mode}}}$ 。由于篇幅所限,本文仅以可变块尺寸搜索编码模式为例对复杂度率失真联合优化模型进行分析,其他的编码模式可以据此类推,若无特别说明,后面提到的 λ_c 均指 $\lambda_{c,\text{mode}}$ 。到此,问题的关键就在于如何量化衡量不同编码模式的复杂度以及如何利用 λ_c 较为准确地对视频序列的复杂度进行控制,下面就这两个问题分别进行分析。

H.264/AVC 标准中引入的可变块尺寸技术共包括 7 种块尺寸,如图 1 所示: 16×16 , 16×8 , 8×16 , 8×8 , 8×4 , 4×8 和 4×4 。其中, 16×16 是一个完整宏块的尺寸, 4×4 是最小单位块的尺寸,这些模式的引入使得运动估计中的匹配块搜索更加灵活。不同块尺寸的搜索模式对应了不同的复杂度和率失真性能。

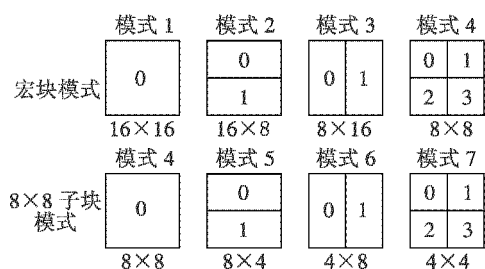


图 1 不同尺寸的编码块

Fig. 1 Coding block with different size

2.2 量化衡量编码复杂度

定义编码复杂度衡量公式为

$$C_i = \omega_i f(R_i) \quad (13)$$

式中: $i = 0, 1, \dots, 6$, 依次代表 7 种不同块尺寸的编码模式; C_i 为第 i 种编码模式的复杂度加权系数; R_i 为使用第 i 种编码模式编码产生的比特数; $f(R_i)$ 为产生比特数的线性函数, $f(R_i) = k \cdot R_i$, 其中, k 为模型参数。由文献[12]可知, 块尺寸搜索所占用的指令周期数与其尺寸大小近似成正比, 因此可得 7 种块尺寸编码模式的复杂度加权系数, 分别为: $\omega_0 = 16, \omega_1 = \omega_2 = 8, \omega_3 = 4, \omega_4 = \omega_5 = 2, \omega_6 = 1$ 。

对于其他种类的编码模式, 可用同样的方法通过计算执行该模式的指令周期数来确定不同模式的加权系数。

2.3 复杂度控制因子

最后需要一个复杂度控制因子来控制视频编码算法的复杂程度, 也就是前文提到的拉格朗日乘子 λ_c 。但是, 不同运动类型的视频序列对编码算法的复杂度要求不一样, 为了可以较为准确地调整视频算法的复杂度, 需要建立复杂度控制因子 λ_c 和算法复杂度之间的关系模型。关于这两者之间的关系, 国内外的文献中尚未有理论分析, 文献[12]中提到的模型也是基于经验统计的结果, 具体为

$$C = a(t) + \frac{b(t)}{\gamma + 1}, \gamma = \begin{cases} \lg \lambda_c + 1, & \lambda_c > 0 \\ 0, & \lambda_c = 0 \end{cases} \quad (14)$$

式中: C 为每一帧的复杂度; $a(t)$ 和 $b(t)$ 为模型参数, 在编码过程中保持更新。根据文献[12]中的实验结果, 该统计模型可以较好地描述 λ_c 和 C 之间的关系。但是这个模型没有考虑输入视频序列的特点, 而研究人员一般把视频序列分为低速率、中速率和高速率 3 种, 复杂度依次增高。为了使模型可以对不同类型视频序列的复杂度有所区别, 同时更准确地描述 λ_c 和 C 之间的关系, 本文对式(14)作了修正, 即

$$C = m[a(t) + \frac{b(t)}{\gamma + 1}], \gamma = \begin{cases} \lg \lambda_c + 1, & \lambda_c > 0 \\ 0, & \lambda_c = 0 \end{cases} \quad (15)$$

式中, m 为输入视频序列每一帧的内容复杂度, 可通过

对编码序列本帧的平均绝对差值的预测得到, 在这里采用标准中码率控制算法中使用的线性预测计算方法, 具体的计算细节参考文献[13]。

3 实验仿真

为了测试算法性能, 实验以 JM12.1 为测试平台, 选取了 3 个具有典型特性的标准视频序列, 分别是 Akiyo(慢速率), Foreman(中速率)和 Mobile(高速率) CIF 格式(352 × 288)进行实验仿真测试。实验采用的编码器的编码选项为: 帧率 25 fps, 编码帧数 150 帧, 编码格式 IPPPP..., 打开 7 种块尺寸, 参考帧为 3 帧, 开启分数像素插值, 开启快速运动搜索选项, 与校验模型 JM12.1 中的算法、文献[12]中的算法进行比较, 图 2 和图 3 分别给出了 Foreman 序列在 λ_c 取值为 8 时本算法与其他两种算法之间的性能比较情况。

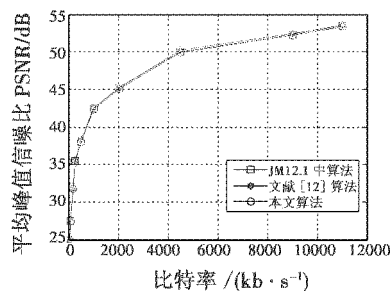


图 2 Forman 序列在 $\lambda_c = 8$ 时率失真性能比较

Fig. 2 Rate-distortion curves of different algorithms for Foreman when $\lambda_c = 8$

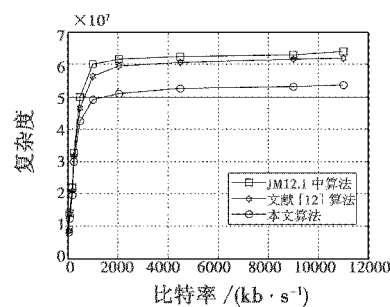


图 3 Forman 序列在 $\lambda_c = 8$ 时的 BitRate-Complexity 曲线比较

Fig. 3 BitRate-Complexity curves of different algorithms for Foreman when $\lambda_c = 8$

从图中可以看出, λ_c 的取值可以控制编码算法的复杂程度, 而且在有效控制复杂度的情况下视频序列的 R-D 性能降低极少, 对其他两个序列进行测试也可以得到类似的结果, 表 1 和表 2 分别列出了 3 个序列在不同 λ_c 值和 Q_p 值情况下的测试结果。由表 1 可以看出, 复杂度率失真联合优化方案在尽可能保持 R-D 性能不变的情况下较大幅度地降低了运算复杂度, 取得了很好的均衡效果。

表 1 本算法相比 JM12.1 中标准算法在不同 λ_c 值和 Q_p 值情况下的测试结果

Table 1 The test result of the proposed algorithm for different λ_c and Q_p compared with JM12.1

测试序列		$Q_p = 18$		$Q_p = 28$	
		$\lambda_c = 8$	$\lambda_c = 16$	$\lambda_c = 8$	$\lambda_c = 16$
Akiyo	PSNR 损失/dB	0.00018	0.00020	0.00016	0.00019
	复杂度减少/%	6.02	7.38	8.12	9.47
Foreman	PSNR 损失/dB	0.0196	0.0206	0.0349	0.0385
	复杂度减少/%	22.01	24.93	26.73	28.95
Mobile	PSNR 损失/dB	0.0296	0.0306	0.0483	0.0514
	复杂度减少/%	11.34	12.60	8.62	10.57

表 2 本算法相比文献[12]中算法在不同 λ_c 值和 Q_p 值情况下的测试结果

Table 2 The test result of the proposed algorithm for different λ_c and Q_p compared with paper [12]

测试序列		$Q_p = 18$		$Q_p = 28$	
		$\lambda_c = 8$	$\lambda_c = 16$	$\lambda_c = 8$	$\lambda_c = 16$
Akiyo	PSNR 损失/dB	0.00018	0.00020	0.00016	0.00019
	复杂度减少/%	6.02	7.38	8.12	9.47
Foreman	PSNR 损失/dB	0.0196	0.0206	0.0349	0.0385
	复杂度减少/%	22.01	24.93	26.73	28.95
Mobile	PSNR 损失/dB	0.0296	0.0306	0.0483	0.0514
	复杂度减少/%	11.34	12.60	8.62	10.57

4 结束语

本文提出了一种基于复杂度的率失真联合优化算法,用来均衡视频编码过程中编码效率和编码复杂度之间的关系。实验结果表明,该算法可以在尽可能保持编码效率不变或变化很小的情况下有效地降低视频编码算法的复杂度,并可以通过调节复杂度控制因子的大小改变对复杂度的控制,达到较好的基于复杂度的率失真联合优化效果。

参 考 文 献

- [1] WIEGAND T, SULLIVAN G, BJONTEGAARD G, et al. Overview of the H. 264/AVC video coding standard [J]. IEEE Transactions on Circuit and Systems for Video Technology, 2003, 13(7):560-576.
- [2] ITU-T Rec. H. 264/ISO/IEC 11496-10, Advanced video coding[Z]. Final Committee Draft, Document JVTF100, 2002.
- [3] 毕厚杰,王健.新一代视频压缩编码标准:H. 264/AVC [M]. 2版.北京:人民邮电出版社,2009.
BI H J, WANG J. A new generation video compression

coding standard [M]. 2nd ed. Beijing: Posts & Telecom Press, 2009.

- [4] WU D, PAN F, LIM K P, et al. Fast inter mode decision in H. 264/AVC video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(6):953-958.
- [5] CHOI W, JEON B, JEONG J. Fast motion estimation with modified diamond search for variable motion block sizes [C]//Proceedings of International Conference on Image Processing (ICIP2003), Barcelona, 2003:371-374.
- [6] MENG B, AU O. Fast intra-prediction mode selection for 4×4 blocks in H. 264 [C]//Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03), Hong Kong, 2003:389-392.
- [7] ZHU S, MA K K. A new diamond search algorithm for fast block matching motion estimation [J]. IEEE Transactions on Image Processing, 2000, 9(2):287-290.
- [8] AHMAD I, ZHENG W G, LUO J C, et al. A fast adaptive motion estimation algorithm [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(3):420-438.
- [9] MALVAR H S, HALLAPURO A, KARCZEWICZ M, et al. Low-complexity transform and quantization in H. 264/AVC [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7):598-603.
- [10] WIEGAND T, GIROD B. Lagrangian multiplier selection in hybrid video coder control [C]//Proceedings of International Conference on Image Processing (ICIP2001), Thessaloniki, 2001:542-545.
- [11] 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.
- [12] HU Y, LI Q, MA S W, et al. Joint rate-distortion-complexity optimization for H. 264 motion search [C]//IEEE International Conference on Multimedia & Expo (ICME2006), Toronto, 2006:1949-1952.
- [13] JOCH A, KOSSENTINI F, SCHWARZ H, et al. Performance comparison of video coding standards using Lagrangian coder control [C]//Proceedings of International Conference on Image Processing (ICIP2002), Rochester, 2002:501-504.