

一种改进的基于柯西模型的 H.264 码率控制方法

胡 栋 孙前锋 谢光剑

(南京邮电大学江苏省图像处理与图像通信重点实验室, 江苏南京 210003)

摘 要: H.264/AVC 编码中的码率控制是通过有效控制输出码流的码率来提高其压缩视频质量的重要技术。本文基于 H.264/AVC 中的 JVT-H017 码率控制方案提出了一种改进算法。新算法根据 H.264 中 DCT 系数的分布特征, 将柯西分布引入到码率控制模块, 用更精确的柯西率失真模型取代了原先的二次率失真模型。在此基础上, 进一步引入了一种联合 PSNR 比率和 MAD 比率进行图像复杂度预测的方法, 并依此来调整帧级比特的分配和量化参数, 克服了在出现复杂运动或场景切换时, 因视频序列相邻帧之间相关性降低而导致的 MAD 预测失准的情况。实验结果表明, 与 JVT-H017 方案及文献 [5] 中的算法比较, 新的算法不仅具有更精确的码率控制, 而且改善了输出码率的平稳性及重建图像的 PSNR。

关键词: 码率控制; 柯西分布; 峰值信噪比比率和平均绝对差值比率

中图分类号: TN **文献标识码:** A **文章编号:** 1003-0530(2011)11-1671-04

An Improved Rate Control Method for H.264 Based on Cauchy Rate Models

HU Dong SUN Qian-feng XIE Guang-jian

(The Provincial Key Lab of Image Processing and Image Communications, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: Rate control in H.264 is a key technology which improves the quality of the compressed video via controlling the bit rate of output streaming. An improved algorithm is proposed on the basis of the JVT-H017 algorithm in this paper. According to the distribution of DCT coefficients in H.264, Cauchy rate module which replaces traditional quadric rate distortion model is adopted into the rate control. Moreover, a prediction algorithm which is used to predict the image complexity via PSNR ratio and MAD ratio is also proposed in this paper, and then the frame bits and QP can be adjusted based on the predicted results. Thus, in the case of complex movement or scene change, this prediction method can overcome the inaccuracy of MAD prediction method in the JVT-H017 algorithm because of the reduced correlation in the adjacent video frames. Simulation results show that, compared with the algorithm in JVT-H017 and reference[5], the new algorithm not only achieves more accurate rate control, but also obtains more stable output bit rate and better PSNR value of reconstructed image.

Key words: Rate Control; Cauchy Distribution; Ratio of Peak Signal to Noise Ratio and Mean Absolute Difference Ratio

1 引言

在 H.264 编码中, 码率控制是决定编码效率的一个重要技术。对于 H.264 编码在码率控制过程中出现“蛋鸡悖论”的情况^{[1][2]}, 现有的解决方法大都是基于 JVT-H017 的解决方案, 即首先利用已编码区域的平均绝对差值(MAD)作为当前编码区域 MAD 的预测值, 然后结合编码单元的分配码率, 通过二次率失真模型(R-D 模型)^[3]估计量化参数, 最后使用率失真优化 RDO 获得最佳编码模式参数, 实现最佳编码, 并且更新工作参数, 用于下一时刻的编码。以往的码率控制算法中,

大多是假设 DCT 系数以高斯分布或拉普拉斯分布, 如 JVT-H017 算法就是基于拉普拉斯分布。由于 H.264 采用多种编码模式、多参考帧、整数 DCT 变换等新技术后, 其编码过程中的变换系数也呈现出新的统计特性, 需要使用更加准确的模型才能获得准确的码率控制。另外, 在实际应用中发现, 当出现复杂运动甚至场景切换的情况时, 视频序列相邻帧之间相关性会大大降低, 从而导致传统码率控制算法中 MAD 预测的失准; 同时, 由于采用固定长度的图像组(GOP)而不能有效地处理视频序列中的场景切换, 最终导致场景切换帧后续各帧编码质量严重下降。所以, 在码率控制过程中,

在复杂运动或场景变换条件下进行图像复杂度的准确预测,并依此作出比特分配的调整尤为重要。

本文内容安排如下:第二节首先简要介绍和分析目前 H.264 编码中码率控制的关键技术和存在的问题;第三节提出本文的解决思路,详细描述本文的算法改进要点,包括基于柯西(Cauchy)分布的率失真模型^[4],基于 PSNR 比率和 MAD 比率^[5]联合的图像复杂度预测方法;第四节通过实验仿真,对本文提出的码率控制算法,与 JVT-H017 和文献[5]的算法进行算法性能对比。本文最后对全文工作进行总结。

2 JVT-H017 码率控制算法简要分析

JVT-H017 提出的码率控制算法可以从 GOP 级逐步到达基本单元层。首先,可以根据传输速率、帧率等对整个 GOP 分配码流比特数,然后再从 GOP 已经分配的比特数中,再分配给每一帧图像一些比特数,最后可以利用每帧图像的比特数,再分配给每个宏块。在只考虑 GOP 采用 IPPP... 模式的情况下,整个码率控制的基本过程如下^[2]:

- (1) 分配当前编码 GOP 所占用的比特数;
- (2) 计算当前帧的目标码率;
- (3) 将整个 GOP 所占用的比特数减去包括当前帧已编码图像的比特数后的剩余比特数平均分配给当前帧未编码的基本单元;
- (4) 根据线性预测理论利用前一编码帧相应位置的 MAD 值预测当前编码帧编码基本单元的 MAD 值;
- (5) 利用二次 R-D 模型计算出相应的量化参数;
- (6) 当前编码基本单元的每一个宏块计算率失真优化(RDO)。
- (7) 重复(1),为下一个 GOP 进行编码。

在上述过程中,编码完第 j 帧后给下一帧分配的比特数 $f(n_{i,j})$ 由(1)式给出:

$$f(n_{i,j}) = \beta \cdot \frac{T_r(n_{i,j})}{N_{p,r}(j-1)} + (1-\beta) \cdot \tilde{f}(n_{i,j}) \quad (1)$$

其中, $T_r(n_{i,j})$ 是为 GOP 中未编码帧分配的比特数, $N_{p,r}(j-1)$ 表示 GOP 中未编码的 P 帧数, $\tilde{f}(n_{i,j})$ 表示编码完 j 帧时的实际缓存占用量, β 是一个取值为 0.75 的常数。

在进行 MAD 预测时,利用前一帧相应的基本单元的 MAD 值,预测出当前编码帧基本单元的 MAD 值。假设当前帧的当前编码基本单元的 MAD 值为 MAD_{cb} ,前一编码帧相应的基本单元的 MAD 值为 MAD_{pb} ,则根据线性预测模型有:

$$MAD_{cb} = a_1 \cdot MAD_{pb} + a_2 \quad (2)$$

其中 a_1 、 a_2 是预测模型的两个系数,初始化时, a_1 的

值为 1, a_2 的值为 0。每编码完一个基本单元时, a_1 、 a_2 的值根据模型得到一次更新。

根据线性预测模型得到 MAD 预测值和分配的码流比特数,利用式(3)的二次 R-D 模型解二次方程得到量化步长 Q ,进而得到量化参数 QP 。为保证图像视觉质量的平滑过渡,前后 QP 的变化不能大于 2。

$$H = c_1 * MAD * Q^{-1} + c_2 * MAD * Q^{-2} + H_{hdr} \quad (3)$$

其中 c_1 、 c_2 分别是一阶和二阶系数,它在编码完每帧图像后重新更新一次。 H 是编码一个基本单元所分配的总的比特数, H_{hdr} 是前一帧图像相应基本单元头信息所使用的比特数, MAD 表示基本单元图像复杂度,在此为 Y 分量的预测值 MAD_{pb} 。注意到(3)式的二次 R-D 模型基于对变换系数的拉普拉斯分布假设,即:

$$p(x) = \frac{\alpha}{2} e^{-\alpha|x|} \quad (4)$$

其中, α 为模型参数。

由于在(1)式中的比特分配中没有考虑根据图像内容的变化情况进行调整,因此在变化的场景下会出现控制不够准确的现象;另外,根据文献[4]的研究结果表明,采用柯西分布要比现有的拉普拉斯分布更加准确。

3 算法改进要点

基于以上两点考虑,我们进行了相应的算法改进。改进后的码率控制算法详细的内容参见[6],下面简要描述其中的基于柯西分布的率失真模型,以及基于 PSNR 比率和 MAD 比率联合的帧码率分配调整。

3.1 基于柯西分布的率失真模型

码率控制的核心是建立率失真模型,即 R-Q 关系,它建立在 DCT 系数分布的基础上。JVT-H017 码率控制算法中延用了 MPEG-4 的假设,将 DCT 系数分布近似为拉普拉斯分布,其优点是模型简单,但对于采用了多种模式的 H.264 编码而言不够准确。因此,在进行算法改进中我们采用了文献[4]的柯西分布:

$$f(x) = \frac{1}{\pi} \frac{\mu}{u^2 + x^2}; x \in R \quad (5)$$

式中, μ 为柯西分布函数的分散系数。相应的码率模型(R-Q 模型)可以简化为:

$$H(Q) = aQ^{-\alpha}; a, \alpha > 0 \quad (6)$$

将式(6)应用到宏块层码率控制上,宏块层的码率模型可表示为:

$$R_i(Q_i) = A(aQ_i^{-\alpha} + C) \quad (7)$$

式中, A 表示为宏块内的像素数, C 为宏块头信息的每像素的平均比特数, Q_i 为第 i 个宏块的量化步长。由此,帧层的码率模型可以表示为:

$$R_i(Q_i, \dots, Q_N) = \sum_{i=1}^N A(aQ_i^{-\alpha} + C) \quad (8)$$

其中, N 表示一帧中宏块的个数。

3.2 基于 PSNR 比率和 MAD 比率的联合帧码率分配

在 H.264 的码率控制算法 JVT-H017 中,码率分配时主要基于平均剩余比特和上一帧编码的比特数,如(1)式所示。由于在分配中没有考虑图像内容变化的需要,因此在出现场景变化或图像内容有剧烈变化时码率分配不够准确,降低了编码效率。在文献[5]中, Jiang 等人使用 MAD 比率作为当前帧比特分配的调整因素对(1)式中的第一项进行修改,获得了编码性能的改善。受其启发,本文进一步引入 PSNR 比率,利用 PSNR 比率与 MAD 比率来联合预测当前帧的复杂度,这样可以更好的解决物体剧烈运动或场景切换发生时所造成的比特分配失准的问题。

定义 $PSNR_j$ 为当前第 j 帧的 PSNR 值; $PSNR_{ratio,j}$ 为 PSNR 的差值比率,其值的大小反应 PSNR 的值的变化程度。

$$PSNR_{ratio,j} = \frac{PSNR_j - PSNR_{j-1}}{\frac{1}{j-1} \sum_{i=1}^{j-1} (PSNR_i - PSNR_{i-1})} \quad (9)$$

MAD 比率 MAD_{ratio} 如文献[5]定义:

$$MAD_{ratio}(i) = \frac{MADP_i}{(\sum_j MADPA_j)/(i-1)} \quad (10)$$

式中 $MADP_i$ 表示当前编码的第 i 帧预测的 MAD 值, $MADPA_j$ 表示 GOP 中前面已经编码的第 j 帧的实际 MAD 值。由式(9)、(10)两式得到联合比率:

$$F = \mu * MAD_{ratio,j} + (1-\mu) * PSNR_{ratio,j} \quad (11)$$

其中, μ 为常数,取 $\mu=0.70$ 来进行仿真实验。根据 F 的大小进行目标比特的分配调整,即相当于对(1)式的第一项乘以一个因子 k 。 k 的表达式如(12)式所示,由实验确定:

$$k = \begin{cases} 0.8-F & F < 1.1 \\ 0.88+0.3(F-1.1) & 1.1 < F < 2.0 \\ 1.15 & F \geq 2.0 \end{cases} \quad (12)$$

4 实验结果与分析

为检测本文算法的性能,我们在 JVT 的 JM8.6 上将新算法与标准算法 JVT-H017 及文献[5]中的 MAD-ratio 算法进行了比较,这里给出部分实验结果。实验序列为:简单运动类型的 missa(QCIF)和 silent(CIF)序列,复杂运动类型的 football(QCIF)和 stefan(CIF)序列。从表 1 和表 2 中可以看出,新算法与其它算法比较不仅实际输出码流更接近目标码率,而且其平均 PSNR 也相应有所提高。从表 2 中也可以发现,对于高

码率运动复杂的序列,本算法改善的效果更明显。这也从理论上得到验证,对于运动复杂的序列, MAD 预测会失准,而引入 PSNR 因子后会改善图像复杂度的预测准确性,从而提升视频的质量。

表 1 平均 PSNR 及其标准偏差和输出码率 (图像格式为 QCIF)

Tab. 1 Mean PSNR, standard deviation and output bit rate (picture format: QCIF)

两种目标码率下的测试序列 (kbps)		平均 PSNR (dB)			平均 PSNR 标准偏差 ε			输出码率 (kbps)		
		JVT-H017	本文算法	MAD-ratio	JVT-H017	本文算法	MAD-ratio	JVT-H017	本文算法	MAD-ratio
missa	128	44.42	44.45	44.43	0.67	0.59	0.63	127.72	128.09	127.78
	256	46.20	46.21	46.20	0.89	0.81	0.85	255.67	256.07	255.95
football	128	26.73	27.10	26.97	0.62	0.54	0.60	128.41	127.79	128.60
	256	29.93	30.17	29.97	1.26	1.07	1.15	256.60	256.16	256.37

表 2 平均 PSNR 及其标准偏差和输出码率 (图像格式为 CIF)

Tab. 2 Mean PSNR, standard deviation and output bit rate (picture format: CIF)

两种目标码率下的测试序列 (kbps)		平均 PSNR (dB)			平均 PSNR 标准偏差 ε			输出码率 (kbps)		
		JVT-H017	本文算法	MAD-ratio	JVT-H017	本文算法	MAD-ratio	JVT-H017	本文算法	MAD-ratio
silent	384	36.99	37.10	37.02	1.25	1.18	1.21	384.44	384.17	384.31
	768	41.13	41.34	41.20	1.19	1.11	1.16	769.00	767.98	768.34
stefan	384	30.47	30.65	30.50	0.63	0.60	0.61	384.43	384.21	384.60
	768	33.54	33.67	33.58	0.37	0.32	0.35	768.51	768.03	767.90

为了更形象的显示结果,本文选取了 CIF 格式的 stefan 和 silent 序列,分别在 384kbps 和 768kbps 目标码率下序列生成的码流的码率和目标码率的关系图以及各帧 Y 分量的 PSNR 值的走势图,如图 1 和图 2 所示。从图 1(a)和图 2(a)中可以明显看出,在高低目标码率下,本文算法的输出码流均更加平稳;而且输出视频序列的 PSNR 基本不会下降且大部分帧的 PSNR 在本文算法下都会有所提高。

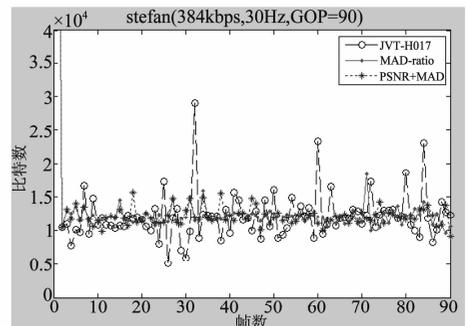


图 1 (a) 输出码率比较(stefan-CIF-384kbps)

Fig. 1 (a) comparison of bit rate(stefan-CIF-384kbps)

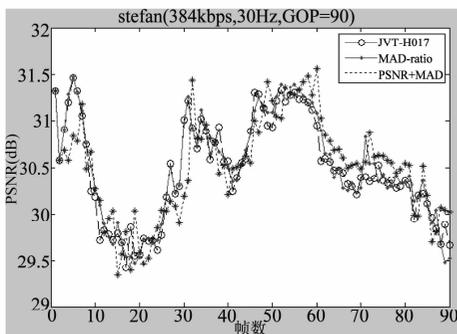


图1 (b) PSNR 比较(stefan-CIF-384kbps)

Fig. 1 (b) comparison of PSNR(stefan-CIF-384kbps)

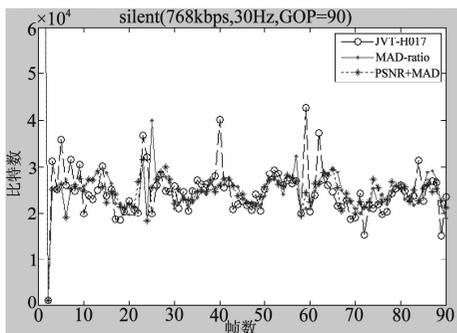


图2 (a) 输出码率比较(silent-CIF-768kbps)

Fig. 2 (a) comparison of bit rate(silent-CIF-768kbps)

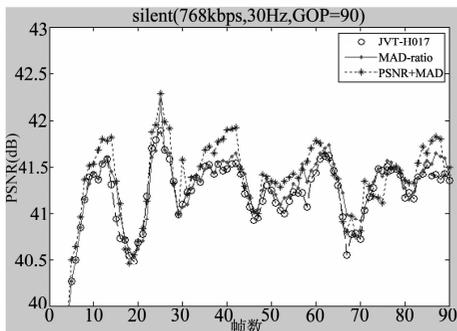


图2 (b) PSNR 比较(silent-CIF-768kbps)

Fig. 2 (b) comparison of PSNR(silent-CIF-768kbps)

5 结束语

本文基于对 H.264 码率控制关键因素的分析,提出了一种有效的改进算法。首先,新算法引入更准确的 Cauchy 模型作为 DCT 系数的统计分布,由此确定相应的 R-D 关系,调整量化参数;其次,在比特分配中,采用 PSNR 比率和 MAD 比率联合预测的算法进行帧级图像比特分配,以改善在场景切换或运动变换剧烈情况下 MAD 预测的性能。仿真实验结果显示,本文的算法码率控制更精确,而且场景变化的情况下输出码平稳性和 PSNR 上均取得了一定改善。

参考文献

- [1] Z. G. Li, F. Pan, K. P. Lim, et al. Adaptive Basic Unit Layer Rate Control for JVT. JVT-G012. 7th Meeting: Pattaya II, Thailand. March, 2003.
- [2] "Proposed draft of adaptive rate control," Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q. 6) 8th Meeting: Geneva, May 20-26, 2003 Document: JVT-H017 Filename: JVT-H017.doc.
- [3] Hung-Ju Lee, Ya-Qin Zhang. Scalable Rate Control for MPEG-4 Video. IEEE Trans. Circuits and Systems for Video Technology. 2000, 10(6):878-894.
- [4] Characi N, Altunbasak Y, and Mercurial R M. Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(8):994-1006.
- [5] Minqiang Jiang, Xiaoquan Yi and Nam Ling, "Improved Frame-Layer Rate Control for H.264 using MAD Ratio", IEEE ISCS, Vancouver, Canada, vol. 3, May. 2004, pp. 813-816.
- [6] 谢光剑, 基于统计模型的 H.264 码率控制技术的研究, 南京邮电大学硕士学位论文, 2011 年 3 月。
Xie Guangjian, A Research on H.264 Rate Control Based on Statistical Model, Master's Thesis of Nanjing University of Posts and Telecommunications, Mar. 2011. (in Chinese)

作者简介

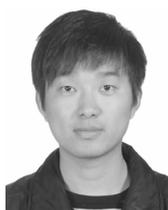
胡 栋,男,博士,南京邮电大学通信与信息工程学院教授、硕士生导师,主要研究方向为图像处理与多媒体通信。

E-mail: hud@njupt.edu.cn



孙前锋,男,硕士,南京邮电大学通信与信息工程学院学生,主要研究方向为图像处理与多媒体通信。

E-mail: Y001090719@njupt.edu.cn



谢光剑,男,硕士,南京邮电大学通信与信息工程学院学生,主要研究方向为图像处理与多媒体通信。

E-mail: Y080823@njupt.edu.cn