文章编号:1001-9081(2010)02-0330-03

一种有效的初始量化参数选择算法

缪广忠,林其伟

(华侨大学 信息科学与工程学院,福建 泉州 362021)

(mgz@hqu.edu.cn)

摘要:在网络环境中,码率控制是影响数字视频质量的重要机制之一。针对提案 JVT-G012 算法中 I 帧量化参数选择不合理造成的视频图像质量 (PSNR) 低、波动大的问题,提出了一种基于 R-Q 模型的 I 帧量化参数选择算法。首先,通过曲线拟合得到 I 帧的 R-Q 模型;其次,利用已编码帧的信息更新 I 帧相对于 P 帧的权重,为 I 帧分配合理的目标比特数;最后,在该目标比特下由 R-Q 模型计算得到 I 帧量化参数。实验结果表明:与 H.264/AVC 原算法相比,改进的算法在准确控制码率的同时,视频图像平均质量得到提高且更加平稳。

关键词:码率控制;JVT-G012;R-Q 模型;量化参数;峰值信噪比

中图分类号: TN919.81;TP391.41 **文献标志码:** A

Effective selection algorithm of initial quantization parameters

MIAO Guang-zhong, LIN Qi-wei

(College of Information Science and Engineering, Huaqiao University, Quanzhou Fujian 362021, China)

Abstract: Rate control is one of the most important mechanisms affecting digital video quality in network environments. As for the problem for low-quality and fluctuation of video images in JVT-G012 algorithm, caused by unreasonable choice of I-frame quantization parameter, this paper put forward a selection algorithm of I-frame quantization parameter, based on the R-Q model. Firstly, R-Q model of I-frame was got by curve fitting; secondly, the information of encoded frames was used to update I-frame weight relative to P-frame, and a reasonable target-bit was allocated for I-frame; finally, I-frame quantization parameter was calculated under the target-bit by R-Q model. Simulation results show that compared to the H.264/AVC standard algorithm, the improved algorithm accurately controls the bit-rate while improving the average PSNR of testing sequences, and makes the quality of video images more stable and smooth.

Key words: rate control; JVT-G012; R-Q model; quantization parameter; Peak Signal-to-Noise Ratio (PSNR)

0 引言

作为新一代视频编码标准,H.264/AVC 在吸收以往各种编码方案优点的前提下,采用了多参考帧、1/4 像素运动估计、整数变换、RDO 等多项技术,使编码效率得到大幅度提高。但与此同时,码流的变速率特性给固定带宽和时延条件下的视频传输带来了更大的困难。为了在给定码率下高效地传输视频码流数据,并在解码端获得高质量的重建图像,码率控制成为编码器不可缺少的一部分,它已经成为衡量编码器性能的重要指标之一。虽然各类视频编码标准没有对码率控制做出明确的规定,但是都提出了各自的方案。例如 MPEG-2 的 TMS^[1],H.263 的 TMN8^[2],MPEG-4 的 VM8^[3]等。而适用于 H.264/AVC 的码率控制算法主要有 JVT-F086 和 JVT-G012^[4]。在 JVT-G012 中采用了 MPEG-4 的二次 R-Q 模型,可以根据信源特征及时地更新模型参数,预测误差较小,同时利用线性 MAD 模型预测当前未编码图像块的复杂度解决了蛋鸡悖论问题。和 JVT-F086 相比,JVT-G012 仅需要进行一次编码处理,计算复杂度较低,编码效率较高。同时引入基本单元概念,将码率控制分为 GOP 层、帧层、基本单元层三层。这种分层的码率控制结构使得我们能够在码率控制精度和 PSNR 值波动之间取得一个较好的折中。

JVT-G012 作为目前主流的经典算法,存在着一些不足^[5]。首先,算法没有充分利用视频图像的时空相关性,使线性预测模型不够准确。其次,剩余比特平均分配给未编码帧,没有考虑当

前帧的编码复杂度,等等。针对这些不足,很多学者提出了相应的改进算法。本文针对 I 帧量化参数选择不合理而造成的峰值信噪比波动过大的问题,提出一种基于 R-Q 模型的 I 帧量化参数选择算法,明显减小了视频图像质量的波动。

1 JVT-G012 的 I 帧量化参数选取

第一个 GOP 的 I 帧和第一个 P 帧采用预定义的量化参数 QP_0 来编码,在没有指定的情况下通过下式选取:

$$QP_0 = \begin{cases} 40, & bpp \leq l1 \\ 30, & l1 < bpp \leq l2 \\ 20, & l2 < bpp \leq l3 \\ 10, & bpp > l3 \end{cases} \quad (1)$$

$$bpp = \frac{R}{f \times N_{\text{pixel}}} \quad (2)$$

其中: R 是目标比特率, f 是帧率, N_{pixel} 是一帧图像的像素个数, bpp 是每像素比特数, $l1, l2, l3$ 是经验常数。后续 GOP 的量化参数 QP_{st} 通过下式计算得到:

$$QP_{st} = \frac{Sum_{PQP}}{N_p} - \frac{8T_r(n_{i-1,N_{GOP}})}{T_r(n_{i,0})} - \min\left\{2, \frac{N_{GOP}}{15}\right\} \quad (3)$$

其中: N_p 是前一个 GOP 中 P 帧的总数, Sum_{PQP} 是 P 帧量化参数的总和。

实验证明,通过上述得到的 QP_0 和 QP_{st} 不能充分反映不同场景下的特点,比如在一个运动剧烈和一个运动十分平缓的场景下可能采用相同的 QP_0 ,然而同一场景下采用不同的

收稿日期:2009-07-30;修回日期:2009-09-10。

作者简介:缪广忠(1983-),男,江苏淮安人,硕士研究生,主要研究方向:视频编码与通信;林其伟(1959-),男,福建泉州人,副教授,主要研究方向:视频编码与网络通信。

QP_0 , 峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)平均值/标准差和码率(bit-rate)相差很大。下面将通过实验数据说明。实验条件: JM 8.6 参考软件, foreman 序列(qcif), 目标码率为 64 Kbps, 帧率为 15 f/s, 编码 50 帧。实验结果如表 1。

表 1 测试结果_foreman

QP_0	PSNR/dB		码率/ Kbps	QP_0	PSNR/dB		码率/ Kbps
	平均值	标准差			平均值	标准差	
15	34.82	5.37	67.49	25	36.94	0.90	64.93
16	35.76	4.10	67.29	26	36.82	0.86	64.12
17	36.31	2.79	65.79	27	36.74	0.83	64.43
18	36.60	2.26	66.32	28	36.71	0.81	64.48
19	36.49	2.09	64.52	29	36.64	0.84	64.33
20	36.64	1.78	64.77	30	36.53	0.88	64.44
21	36.76	1.59	64.77	31	36.47	1.01	64.47
22	36.88	1.32	64.69	32	36.36	1.13	64.27
23	36.91	1.13	64.55	33	36.33	1.26	64.48
24	36.93	0.97	64.76	34	36.24	1.59	64.44

分析其原因可知, JVT-G012 没有根据不同的场景选择不同的 QP_0 , QP_0 选择不合理必将造成该 GOP 内的比特分配及 P 帧量化参数计算不准确, 而由式(3)得到的 QP_{st} 也因此计算不准确。这样就导致整个视频序列的码率和图像质量的不稳定。因此合理选取 QP_0 对提高整个视频质量有着重要意义。

2 基于 R-Q 模型的 I 帧量化参数选取

2.1 I 帧 R-Q 模型的建立

图 1 是不同复杂度视频序列第一帧在 I 帧模式下的 R-Q 曲线。可以看出, 这一组曲线接近指数分布, 因此可以用负指数函数拟合, 即:

$$R(QP) = a \times (e^{-b \times QP} + c) \quad (4)$$

其中: 参数 c 是为了提高拟合精度引入的参数。表 2 列出了不同序列对应的 a 、 b 、 c 的取值。

图 1 不同视频序列 I 帧编码时的 R-Q 曲线

表 2 不同视频第一帧 R-Q 曲线拟合系数

序列	a	b	c
mobile	371 300	0.051 9	-0.057 8
tempe	337 900	0.056 3	-0.055 1
football	361 300	0.066 5	-0.032 2
bus	327 600	0.067 5	-0.032 3
coastguard	346 300	0.081 5	-0.012 5
salesman	308 700	0.080 1	-0.011 4
foreman	292 700	0.086 9	-0.009 8
mother	264 900	0.085 8	-0.009 5
claire	134 400	0.081 3	-0.008 9

从表 2 中可以看出, 不同特征的 I 帧对应不同的 a 、 b 、 c 系数。本文用像素的平均梯度来衡量 I 帧的特征, 称为图像复杂度 Complexity^[6], 定义如下:

$$Complexity =$$

$$\frac{1}{M \times N} \left\{ \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (|I_{i,j} - I_{i+1,j}| + |I_{i,j} - I_{i,j+1}|) \right\} \quad (5)$$

其中: $I_{i,j}$ 表示坐标为 (i,j) 处的亮度像素值, M 和 N 分别是图像的宽度和高度。经大量实验证明:

$$a = A \times Complexity \quad (6)$$

则式(4)可表示为:

$$R(QP) = A \times Complexity \times (e^{-b \times QP} + c) \quad (7)$$

其中: A 、 b 、 c 与 I 帧复杂度有关, 且在某一复杂度范围内基本不变。简单起见, 在一定复杂度范围内我们将其视为常数, 对大量实验数据进行统计, 取得最优值如下:

$$\begin{cases} A = 19000, b = 0.083, c = -0.010; & 0 \leq Complexity < 20 \\ A = 13500, b = 0.067, c = -0.032; & 20 \leq Complexity < 30 \\ A = 10800, b = 0.055, c = -0.050; & Complexity \geq 30 \end{cases} \quad (8)$$

模型系数均为经验常数, 不需要更新, 因此算法复杂度较低。为了验证本文模型的准确性, 下面绘出序列的模型曲线与实际 R-Q 曲线的一组对比图, 如图 2。鉴于实际的视频编码中, QP 选取多在 15~51, 由此可见, 本文提出的模型曲线能够很好地逼近实际的 R-Q 曲线。

图 2 I 帧 R-Q 曲线拟合效果

2.2 I 帧比特分配及量化参数计算

I 帧目标比特 T_I 计算如下:

$$T_I = T_{GOP} \times \frac{W_I}{W_I + W_P \times N_P} \quad (9)$$

其中: T_{GOP} 是一个 GOP 分配的比特数, N_P 是 GOP 中 P 帧的数目, W_I 和 W_P 分别是 GOP 中 I 帧和 P 帧的权重。

通过更新 GOP 的 I 帧权重 W_I 来自适应平衡 GOP 中 I 帧和 P 帧的比特分配, 为 I 帧分配合理的比特数^[7], W_I 由下式计算得到:

$$W_I = \frac{B_I}{B_{aveP}} \times e^{\left(\frac{PSNR_{aveP} - PSNR_I + \alpha}{\beta} \right)} \quad (10)$$

其中: B_{aveP} 为 GOP 中所有 P 帧的平均实际编码比特数, B_I 为 I 帧实际编码比特数, $PSNR_{aveP}$ 为所有 P 帧的平均 PSNR 值, $PSNR_I$ 为 I 帧 PSNR 值。 α 是在最优 QP_0 下对 I 帧 PSNR 值与 P 帧平均 PSNR 值的差作统计平均得到的经验值, 设为 3, β 也为经验常数设为 8, 以 P 帧作为参考, W_P 设为 1.0。

因为 QP_0 为第一个 GOP 的 I 帧量化参数, 在此之前没有

已编码的 GOP, 所以无法根据式(10)更新 W_1 从而为 I 帧分配目标比特 T_1 , 也就无法计算 QP_0 。

本文中采用重复编码第一个 GOP 的方法来解决该问题。在预定义的 QP_0 下(本文取 28)预先编码第一个 GOP, 得到该 QP_0 下 GOP 中已编码帧的信息, 利用式(10)更新 W_1 , 通过式(9)计算目标比特 T_1 , 最后由式(7)计算出新的 QP_0' 。由于量化步长的范围是 0 ~ 51, 因此对 QP_0' 作如下限制:

$$QP_0' = \min\{51, \max\{0, QP_0'\}\} \quad (11)$$

用此 QP_0' 作为第一个 GOP 的初始量化参数对整个视频重新进行编码。 QP_{st} 仍由式(3)计算得到。

这里需要说明的是以下两点。

1) 预先设定 QP_0 对第一个 GOP 进行预编码只是为了得到该 GOP 已编码帧的信息, 从而计算出新的合理的初始量化参数 QP_0' 。因此, 最终用来对整个视频序列进行编码的初始量化参数是 QP_0' 而不是 QP_0 。

2) 重复编码一个 GOP 增加了一定的算法复杂度, 但是一个 GOP (通常为 10 ~ 30 帧) 与整个视频序列 (一个帧率为 30 帧/s 的 5 分钟长的视频就有 9 000 帧) 相比却是十分微小的。而牺牲这样一个相对较小的复杂度换来整个视频序列图像质量的提高与稳定是十分值得的。

3 实验结果及分析

为了验证本文改进的码率控制算法的有效性, 在 H. 264/AVC 的参考软件 JM8.6 上实现了本文算法, 并与提案 JVT-G012 中算法进行实验结果比较。实验条件为: 标准 qcif 测试序列, 帧率为 15 帧/s, 编码 100 帧, GOP 长度为 20, 采用 IPPP 编码模式。实验结果如图 3 ~ 5、表 3 所示。

图3 akiyo 序列不同码率下的平均 PSNR 值比较

实验证明: 对于运动平缓的序列, I 帧要作为后续 P 帧的参考, 应该分配相对较多的比特, 即采用较小的 QP_0 , 以保证后续 P 帧的质量; 对于运动剧烈的序列, P 帧预测残差较大, I 帧应分配相对较少的比特, 即采用较大的 QP_0 , 以留有足够的剩余比特给后续 P 帧保证其质量。本文正是实现了这一算法思想选取了更加合理的 QP_0 。

由图 3 和图 4 可以看出, 本文算法在不同码率下, 都能比 JVT-G012 原算法选取更优的初始量化参数 QP_0 , 使得视频图像质量有较大提高。Akiyo 序列运动比较平缓, 在高码率时, 选取不同 QP_0 对其质量影响较大; football 序列运动比较剧烈, 在低码率时, 选取不同 QP_0 对其质量影响较大。表 3 列出了其他视频序列在不同码率下的实验数据, 实验结果表明: 本文算法对不同复杂度的视频序列质量都有一定的提高, 平均 PSNR 值最大提高达到 1.23 dB, 其他也有约 0.1 ~ 0.9 dB 的提高, 而且对码率的控制更加精确。

此外, QP_0 的选取对视频质量的波动也有重要影响, 尤其是运动比较剧烈的序列。由图 5 可以看出, 由于选取了较合理

的 QP_0 , 进而 QP_{st} 的计算也更加准确, 所以各 GOP 内比特分配合理, 整个视频质量更加平稳, 主观质量得到提高。

图4 football 序列不同码率下的平均 PSNR 值比较

图5 mobile 序列每帧图像的 PSNR 值比较

表3 其他序列实验结果比较

视频序列	码率/Kbps	平均 PSNR/dB		实际码率/Kbps	
		G012	本文方法	G012	本文方法
claire	32	42.35	42.57	32.71	32.62
	48	44.09	44.46	48.43	48.33
	64	45.44	45.82	64.01	63.97
news	32	34.77	34.86	32.83	32.77
	48	37.25	37.37	49.14	48.86
	64	38.86	39.07	64.56	64.73
salesman	32	34.37	34.50	32.27	32.24
	48	36.54	36.74	48.21	48.17
	64	38.24	38.47	64.34	64.24
silent	32	33.74	33.62	32.38	32.33
	48	35.87	36.28	48.14	48.32
	64	37.45	38.05	64.40	64.41
mobile	64	26.66	27.47	64.99	64.69
	80	28.50	28.64	80.82	80.61
	96	29.53	29.50	96.62	96.52

值得指出的是: 本文算法不一定能够选取到最优的初始量化参数 QP_0 , 却能选取比预设值更优的 QP_0 。同时, 不同的视频序列也体现出不同的实际效果。

4 结语

码率控制作为视频编码中的一项重要技术, 一直以来受到广大研究学者的重视, 并提出了众多经典的码率控制算法, 值得我们借鉴和研究。本文分析了 JVT-G012 码率控制提案中 I 帧量化参数选择算法的不足, 提出一种基于 R-Q 模型的 I 帧量化参数选择算法。仿真结果证明: 该 R-Q 模型能很好地逼近实际的 R-Q 曲线, 同时重复编码的策略获得了更优的初始量化参数, 达到了整个视频质量的提高与稳定。

(下转第 340 页)

实验数据表明,本文提出的算法在解码质量和编码速度上有较大优势。但由于值域块较其他的分形压缩编码算法中的值域块小,造成在减少编码时间的同时,图像的压缩比有所降低。若提高值域块的覆盖范围同时改善值域块与定义域块的匹配方法,该方法的压缩比将有所提高。

图4 Lena 测试效果

图5 Airplane 测试效果

参考文献:

- [1] BARNESLEY M F. Fractal everywhere[M]. New York: Academic Press, 1988.
- [2] JACQUIN A E. Image coding based on a fractal theory of iterated contractive image transformation[J]. IEEE Transactions on Image Processing, 1992, 1(1): 18-30.
- [3] FISHER Y. Fractal image compression[J]. Fractals, 1994, 2(3): 321-329.
- [4] FISHER Y. Fractal image compression: Theory and application[M]. New York: Springer-Verlag, 1995.
- [5] WANG H, WANG M, HINTZ T, et al. Fractal image compression on a pseudo spiral architecture[C]// 2006 IEEE Conference on Cybernetics and Intelligent Systems. Darlinghurst: Australian Computer Society, 2006: 201-207.
- [6] THAKUR N V, KAKDE O G. Fractal color image compression on a pseudo spiral architecture[C]// IEEE Conference on Cybernetics and Intelligent Systems. Bangkok: IEEE Computer Society, 2006: 228-233.
- [7] THAKUR N V, KAKDE O G. Color image compression on spiral architecture using optimized domain blocks in fractal coding[C]// ITNG'07: International Conference on Information Technology. Washington, DC: IEEE Computer Society, 2007: 234-242.
- [8] DAVOINE F, BERTIN E, CHASSERY J M. An adaptive partition for fractal image coding[J]. Fractals, 1997, 5(S1): 243-256.
- [9] DAVOINE F, ANTONINI M, CHASSERY J M, et al. Fractal image compression based on delaunay triangulation and vector quantization[J]. IEEE Transactions on Image Processing, 1996, 5(2): 338-346.
- [10] DAVIS G M. A wavelet-based analysis of fractal image compression[J]. IEEE Transactions on Image Processing, 1998, 7(2): 141-154.
- [11] IANO Y, da SILVA F S, CRUZ A L M. A fast and efficient hybrid fractal-wavelet image coder[J]. IEEE Transactions on Image Processing, 2006, 15(1): 98-105.
- [12] ZHOU Y, ZHANG C, ZHANG Z. An efficient fractal image coding algorithm using unified feature and DCT[J]. Chaos, Solitons and Fractal, 2009, 39(4): 1823-1830.
- [13] 陈守吉, 张立明. 分形与图像压缩[M]. 上海: 上海科技教育出版社, 1998.
- [14] PYTEL A, KUSALAAS J. Engineering mechanics - Dynamics[M]. 2nd ed. 北京: 清华大学出版社, 2001.
- [15] THORNTON S T, MARION J B. Classical dynamics of particles and systems[M]. [S.l.]: Thomson Learning, 2005.
- [16] 雷蕾, 李言俊, 张科. 图像目标质心快速搜索算法[J]. 红外技术, 2007, 29(9): 548-551.
- [17] ZHOU Y, ZHANG C, ZHANG Z. Fast hybrid fractal image compression using an image feature and neural network[J]. Chaos, Solitons and Fractals, 2008, 37(2): 623-631.
- [18] TSENG C C, HSIEH J G, JENG J H. Fractal image compression using visual-based particle swarm optimization[J]. Image and Vision Computing, 2008, 26(8): 1154-1162.
- [19] YUAN W, LIN S X, ZHANG Y D. Optimum bit allocation and rate control for H. 264/AVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(6): 705-715.
- [20] KIM W J, YI J W, KIM S D. A bit allocation method based on picture activity for still image coding[J]. IEEE Transactions on Image Processing, 1999, 8(7): 974-977.
- [21] ZHOU YI-MIN, SUN YU, FENG ZHI-DAN, et al. New rate-complexity quantization modeling and efficient rate control for H. 264/AVC[C]// 2008 IEEE International Conference on Multimedia and Expo. Washington, DC: IEEE Computer Society, 2008: 717-720.
- [22] ISO-IEC/JTC1/SC29/WG11/N0400 [S]. TestModel 5, 1993.
- [23] ITU-T/SG15 [S]. Video Codec TestModel Near-Team, TMN8. ITU Study Group 16, Video Coding Experts Groups Document Q15-A-59, Portland, USA, 1997.
- [24] CHANG T, ZHANG Y. A new rate control scheme using quadratic rate distortion model[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, CSVT-7(1): 287-311.
- [25] LI Z G, PAN F, LIM K P, et al. Adaptive basic unit layer rate control for JVT [S]. Pattaya: ISO/IEC and ITUT, 2003.

(上接第332页)

参考文献:

- [1] ISO-IEC/JTC1/SC29/WG11/N0400 [S]. TestModel 5, 1993.
- [2] ITU-T/SG15 [S]. Video Codec TestModel Near-Team, TMN8. ITU Study Group 16, Video Coding Experts Groups Document Q15-A-59, Portland, USA, 1997.
- [3] CHANG T, ZHANG Y. A new rate control scheme using quadratic rate distortion model[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, CSVT-7(1): 287-311.
- [4] LI Z G, PAN F, LIM K P, et al. Adaptive basic unit layer rate control for JVT [S]. Pattaya: ISO/IEC and ITUT, 2003.
- [5] YUAN W, LIN S X, ZHANG Y D. Optimum bit allocation and rate control for H. 264/AVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(6): 705-715.
- [6] KIM W J, YI J W, KIM S D. A bit allocation method based on picture activity for still image coding[J]. IEEE Transactions on Image Processing, 1999, 8(7): 974-977.
- [7] ZHOU YI-MIN, SUN YU, FENG ZHI-DAN, et al. New rate-complexity quantization modeling and efficient rate control for H. 264/AVC[C]// 2008 IEEE International Conference on Multimedia and Expo. Washington, DC: IEEE Computer Society, 2008: 717-720.