

一种单通道 VBR 码率控制算法

陈国斌, 刘济林, 孙小叶, 姜 超

(浙江大学信息与电子工程系, 浙江杭州 310027)

摘 要: 本文提出了一种低计算复杂度的 VBR 码率控制算法, 它不需要对视频序列进行预先分析或多次编码。该算法的主要思想是根据压缩码流统计信息, 建立码率 - 失真模型, 计算目标比特开销, 然后通过一个反馈机制在帧层调节量化因子, 使得目标比特预算满足缓冲区约束条件, 并尽可能地维持量化因子的稳定。仿真结果表明, 本文提出的方法不仅能够保证总体比特预算和缓冲区约束, 获得相对稳定的视频质量, 特别是对场景切换也有很强的适应性, 而且实现简单, 运算复杂度低, 可应用于实时视频编码。

关键词: 码率控制; VBR; 目标比特分配

中图分类号: TN919.81 **文献标识码:** A **文章编号:** 0372-2112 (2004) 07-1184-04

A Single-Pass VBR Rate Control Algorithm

CHEN Guo-bin, LIU Ji-lin, SUN Xiao-ye, JIANG Chao

(Information & Electronic Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China)

Abstract: Rate control is a key module in a video encoder, and plays an important role in quality of video services. This paper provides a frame-layer variable bit-rate (VBR) rate control method without pre-analysis and pre-coding. Quantization parameters (QPs) are initialized in frame layer according to previous coding statistics firstly; then target bit is calculated according to rate-distortion (R-D) model, and a feedback mechanism regulates QPs and accommodates the target bit to buffer constraints. Simulation results show that the proposed method can not only meet the global bit budget and buffer conformance, but also result in a constant picture quality. Additionally, its low computational complexity makes it applicable to real-time video coding.

Key words: rate control; variable bit-rate; target bit allocation

1 引言

视频编码的主要目标就是提供一定压缩比, 并尽可能地提高视频质量。一般地, 视频编码算法, 如 MPEG-x 和 H.26x, 都只是定义了码流语法和解码过程, 而没有规定编码的具体算法, 这使得编码器的实现具有较大灵活性。码率控制作为视频编码算法的关键组成部分, 直接影响着编码性能。针对具体应用需求选择合适的码率控制方案, 已成为视频编码的一个重要研究方向。

在视频编码中, 变长编码的使用以及场景内容和复杂程度的变化导致了码率波动。由于信道带宽和存储空间限制, 常引入缓冲机制来平滑码率波动。在视频信号是平稳过程假设基础上, TM5 提出了基于图组 (GOP: group of picture) 分配目标比特的恒定码率 (constant bit-rate: CBR) 控制^[1]。Limin W. 在^[2]中提出了变长 GOP 码率控制方法来适应场景切换。MPEG-4 VM 15.0 分析了输入信号的空间统计特性, 提出了基于 R-D 模型的控制方案^[3]。但 CBR 控制算法有其内在缺陷, 不能提供稳定的视频服务质量。和 CBR 码率控制不同, VBR 码率控制允许码率在一定范围内波动, 因而能够获得近

似恒定的视频服务质量。基于 VBR 码率控制的视频编码往往采用双通道或多通道方法, 通过两次或多次编码, 分析视频码流的统计信息, 获取最优编码方式。这种方法计算复杂度高, 特别是对时间较长的序列^[4]。本文在分析压缩视频统计特性的基础上, 建立了码流在压缩域的码率失真模型, 提出了一种单通道的 VBR 实时码率控制方法。它无需事先获得视频特征信息, 计算复杂度低。

下文结构安排如下: 第二章讲述码率控制原理; 第三章讲述基于 R-D 模型的单通道 VBR 码率控制算法; 第四章分析仿真结果; 最后一章对全文进行总结。

2 码率控制基本原理

图 1 为一个简单的码率控制示意图。在视频编码算法中, 缓冲区被引入以平滑码率波动。码率控制器监视缓冲区状态, 根据缓冲区饱和度 B_f 改变量

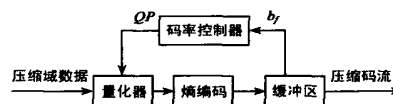


图 1 一种简单码率控制框图

化因子 QP 大小, 达到调节码率的目的。缓冲区约束和比特分

配是码率控制中需要解决的两个问题。

2.1 缓冲区约束

在码率控制中,必须防止缓冲区溢出。CBR 缓冲区模型如图 2 所示。 B_n 为图像 n 移出缓冲区之前状态。

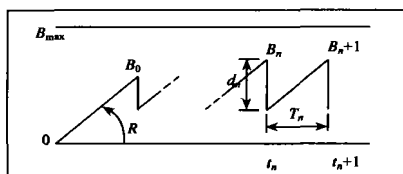


图 2 CBR 缓冲区模型

$$B_0 = B_{\max} \quad (1)$$

$$B_{n+1} = B_n - d_n + RT_n \quad (2)$$

其中, B_{\max} 为缓冲区初始状态; d_n 为第 n 帧图象大小; R 为码率; T_n 为相邻两个帧之间地间隔。为了避免缓冲区上溢和下溢,必须保证下式成立。

$$d_n \leq B_n \leq B_{n+1} \quad (3)$$

VBR 缓冲区模型如图 3 所示。在 VBR 中,码流以最大码率 R_{\max} 进入缓冲区。缓冲区充满后,停止向缓冲区输入,直到一幅图像从缓冲区中移出。我们仍以 B_n 表示图像 n 移出缓冲区之前的状态。

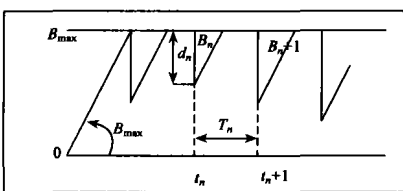


图 3 VBR 缓冲区模型

$$B_0 = B_{\max} \quad (4)$$

$$B_{n+1} = \min(B_{\max}, B_n - d_n + R_{\max} T_n) \quad (5)$$

为了防止缓冲区下溢,必须保证

$$d_n \leq B_n \quad (6)$$

若不计延时,增大缓冲区,则可以防止缓冲区溢出问题。除非缓冲区为无限大,否则必须引入码率控制。另外,对于延时要求高的应用,如何减小延时则非常重要。

2.2 比特分配

在满足缓冲区约束的基础上,提高视频质量是码率控制的另一个目标。即如何分配比特,保证图象质量的稳定。由于恒定码率的约束, CBR 不能提供恒定的视频质量。VBR 对码率约束相对宽松,在一段时间内,允许平均码率在零值和最大值之间变化。一种防止视频质量变化的方法是保持量化因子不变,但这将导致码率波动太大。TMS 提出按 GOP 分配比特的方案,但这种方法不适合 VBR,因为 VBR 并不要求每个 GOP 有相同的比特开销,另外一些控制算法则提出了按照 $R-Q$ 模型分配方案,例如对数和指数分布模型^[5~7]。所有的模型都是对实际情况某种程度的近似。合理分配比特已成为码率控制研究中的一个富有挑战性问题。

3 单通道的 VBR 码率控制算法

提出利用空间活动性度量宏块编码复杂度,自适应地调节量化因子。然而,在某些情况下,宏块的空间活动性不能有效地指示宏块编码复杂度。如宏块由亮度差别较大两部分组成,这时利用空间活动性描述编

码复杂性将夸大实际复杂度;其次,帧编码是自上而下逐宏块进行,随着可用的目标比特数减少,自适应量化可能会产生图像的上半部分被细量化,下半部分被相对粗量化,使得图象的上下两部分的质量失衡^[8]。再次,在 CBR 中,按 GOP 的比特分配方式和实际场景分布不相符,编码时尽可能地消耗所有的目标比特。在这种方式下,一旦切换到复杂场景,可用比特预算不足,造成图像质量急剧下降。

针对上述存在问题,我们采用一种 DCT 域编码复杂度的计算方法代替空间活动性。因为编码是在 DCT 域中进行的, DCT 域中的复杂度信息更能够反应实际编码的复杂度。帧、宏块和块的编码复杂度计算如下:

$$C_{\text{frm}} = C_{\text{mb}}(i), \quad i = 1, \dots, \text{mb. cnt} \quad (7)$$

$$C_{\text{mb}}(i) = C_{\text{blk}}(i, j), \quad j = 1, 1, \dots, \text{blk. cnt} \quad (8)$$

$$C_{\text{blk}}(i, j) = |\text{Err}(i, j, k, l) / Q_{\text{mat}}(k, l)| \quad (9)$$

其中 C_{frm} 为帧编码复杂度; $C_{\text{mb}}(i)$ 为第 i 宏块的编码复杂度; $C_{\text{blk}}(i, j)$ 为宏块 i 中第 j 个块的编码复杂度; mb-cnt 为一帧中宏块的总数; blk-cnt 为一个宏块中块的总数。Err(i, j, k, l) 为宏块 i 中第 j 块位于 (k, l) 动补偿后的 DCT 系数, $Q_{\text{mat}}(k, l)$ 为量化矩阵在 (k, l) 处系数。

经过大量的试验,发现帧复杂度 C_{frm} 、量化因子 QP 和 DCT 系数所用的编码比特数 Err 存在如下线性关系(如图 4 所示):

$$\text{Err} = a_1 \times \frac{C_{\text{frm}}}{QP} + a_0 \quad (10)$$

这样,一旦确定 C_{frm} 、 a_0 、 a_1 和 QP, 就能够根据式(10)预测出 DCT 系数所使用的比特数 Err。调节 QP, 直至目标比特 T 满足缓冲区约束条件。

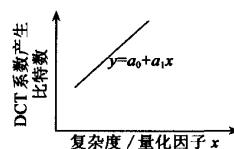


图 4 DCT 系数产生比特数-复杂度/量化因子关系

VBR 仅限制了码率的范围,允许码率在最大值和零值之间变化,为编码留下了很大的灵活性。这样,可对简单场景使用少量的比特编码,节省的比特可以为复杂场景所用。这样,出现复杂场景时,仍然可维持帧间量化因子相对稳定,也就减小了编码对场景切换的敏感性。另外,我们在宏块层上使用恒量化因子,使得每个宏块都可被平等对待,有益于抑止编码中的块效应,同时也减小了整个算法的计算复杂度。

基于上述思想,我们首先在 DCT 域计算帧编码复杂度,然后利用一个反馈机制调节量化因子,使得目标比特预算满足缓冲区约束条件。图 5 为本文提出的 VBR 码率控制框图。

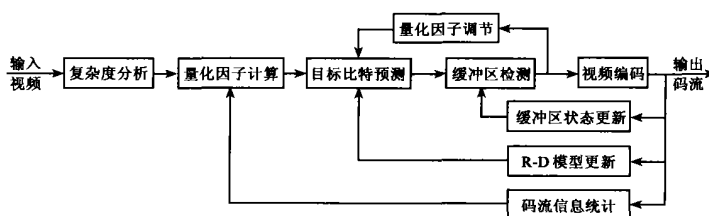


图 5 VBR 码率控制算法框图

步骤如下:

(1) 根据式(7)计算帧编码复杂度.

(2) 根据码流统计信息和编码帧类型决定量化因子 QP 的初始调节方向.

下面计算 I 帧的量化因子 QP_n^i :

$$\text{if (Bit_saving} > \text{Thres}_1^i) \quad (11)$$

$$QP_pred_n^i = QP_{n-1}^i - QP_1^i$$

$$\text{else if (Bit_saving} < \text{Thres}_2^i) \quad (12)$$

$$QP_pred_n^i = QP_{n-1}^i + QP_2^i$$

$$\text{else} \quad (13)$$

$$QP_pred_n^i = QP_{n-1}^i$$

其中, Bit_saving 为编码节省的比特数; Thres_1^i 、 Thres_2^i 为阈值. QP_1^i 、 QP_2^i 为 QP_n^i 变化增量; 如果 $n=0$, $QP_pred_n^i = 28$.

与 I 帧类似, 可以以同样的方式获得 P 帧、 B 帧的量化因子 $QP_pred_n^p$ 和 $QP_pred_n^b$; 如果 $n=0$, $QP_pred_n^p = 32$, $QP_pred_n^b = 34$.

(3) 根据 $R-D$ 模型和 QP 计算目标比特数 T .

$$T_err = a_1^i \times \frac{C_{\text{lim}}}{QP_n^i} + a_0^i \quad (14)$$

$$T_n^i = T_err + S_{n-1}^i - \text{Bit_err}_n^i \quad (15)$$

其中, T_err 为 DCT 系数的目标比特数; S_n^i 为第 n 个 I 帧编码比特总数; Bit_err_n^i 为第 n 个 I 帧 DCT 系数的实际编码比特数; T_n^i 为 I 帧的目标比特数.

与 I 帧类似, 可以以同样的方式获得 P 帧、 B 帧的目标比特 T_n^p 和 T_n^b .

(4) 检查目标比特数是否符合 VBV 缓冲区限制, 如果不满足缓冲区约束, 调整 QP_n , 转 3, 重新计算 T_n .

$$B_n = (\max(B_{\text{max}}, B_{n-1} + R_{\text{max}}/\text{Frame_rate})) \quad (16)$$

$$B_pred_n = B_n - T_n \quad (17)$$

如果 B_pred_n 小于阈值 B_{low} , 则按式(18)调节 QP_n , 转 3, 重新计算目标比特 T .

$$QP_pred_n = QP_pred_n + QP_3 \quad (18)$$

否则

$$QP_n = QP_pred_n \quad (19)$$

其中, B_n 为第 n 帧解码前的缓冲区状态; B_n 为第 n 帧解码后的缓冲区状态; B_{max} 为缓冲区大小; Frame_rate 为帧率; B_pred_n 为在目标比特数 T_n 下解码后预测的缓冲区状态; QP_3 为 QP_pred_n 的增量.

(5) 根据编码后的统计信息, 更新 $R-D$ 模型中的系数 a_0^i 、 a_1^i 、 a_0^p 、 a_1^p 和 a_0^b 、 a_1^b .

$$\text{Bit_err}_n^i = a_1^i \times \frac{C_{\text{lim}}}{QP} + a_0^i \quad (20)$$

$$\text{Bit_err}_n^p = a_1^p \times \frac{C_{\text{lim}}}{QP} + a_0^p \quad (21)$$

$$\text{Bit_err}_n^b = a_1^b \times \frac{C_{\text{lim}}}{QP} + a_0^b \quad (22)$$

其中, Bit_err_n^i 、 Bit_err_n^p 和 Bit_err_n^b 分别 I 、 P 和 B 帧的

DCT 系数的编码开销. 缓冲区状态按下式更新:

$$B_n = B_{n-1} - S \quad (23)$$

其中 S 为当前帧编码的比特开销.

4 试验结果及分析

我们在 MSSG(MPEG software simulation group) 提供的平台上实现了本文提出的单通道 VBR 码率控制算法. 测试序列为 stefan 和 concat. seq, 测试比较了 VBR 码率控制和 TM5 码率控制不同码率下的性能. 两个序列格式为 CIF, 长度为 300 帧, 色差格式为 4:2:0. 其中 concat. seq 由 foreman、container、stefan 和 coastguard 各 75 帧串接而成.

图 6 为两种算法 PSNR 比较. 其中 stefan 码率为 1.2Mbps; concat. seq 的码率为 0.5Mbps. 从图 6(a) 中可看出, 对于 stefan 序列, TM5 算法的在 200 帧后 PSNR 急剧下降, 最高和最低 PSNR 值相差 9.34dB. 而 VBR 算法能够基本保持 PSNR 的稳定. concat. seq 中则出现场景转换, 既有从简单场景切换到复杂场景, 也有从复杂场景切换到简单场景. 在 concat. seq 第 151 帧, 序列从 container 中切换到 stefan 序列, 场景由简单变为复杂. 从图 6(b) 中可以看出, TM5 算法 PSNR 下降 4.44dB, 而 VBR 控制算法仅下降 1.28dB. 一般地, 观众仅允许场景切换后开始的 2~3 帧内有一些质量下降^[2]. VBR 算法 PSNR 曲线变化平滑, 可在一定程度上减小场景切换的影响, 保证了图像质量的相对稳定.

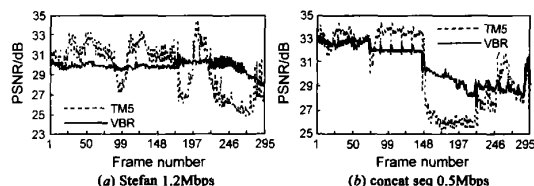


图 6 两种算法的 PSNR 曲线

从图 7(a) 中可以看出: stefan 序列的量化因子上相对稳定, 由此产生的图像质量也比较稳定, PSNR 近似于直线. 而 TM5 控制算法则对场景变化相当敏感. 量化因子的波动范围大, 造成了图像质量的不稳定. 图 7(b) 为两种算法在 concat. seq 序列上的量化因子曲线. 本文中使用的线性量化, 在很长范围内, TM5 使用了量化因子的最大值, 量化过于粗略, 图像中将出现不同程度的块效应. 而本文提出的 VBR 控制算法能够相对平等地对待不同的场景, 可以为复杂场景节省或预开销一些比特, 因而可以保持量化因子的相对稳定.

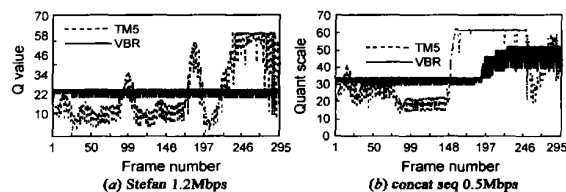


图 7 两种算法量化因子曲线

图 8 为 VBR 控制算法的缓冲区占用情况。两个序列的缓冲区大小都为半秒钟。在场景变化的地方,缓冲区状态变化趋于复杂,编码消耗的比特数相对较多。如图 8(b)所示,concat. seq 由 container 切换到 stefan 时,缓冲区变化比较剧烈。如果缓冲区充满,就停止向缓冲区中入码流,这时码率由最大码率将变为零。这样,在 VBR 码率控制中不会出现缓冲区上溢事件。无论是 stefan 在 1.2Mbps 还是 concat. seq 在较

低的 0.5Mbps 下,缓冲区都没有发生下溢事件。

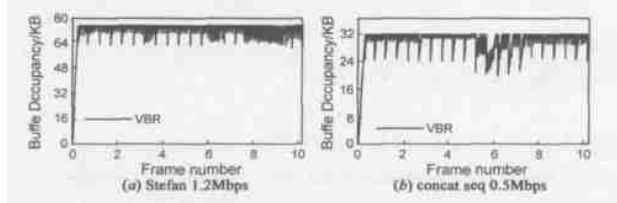


图 8 VBR 码率控制算法缓冲区状态曲线

表 1 列出了 TM5 和本文提出的 VBR 码率控制算法性能比较。我们从 PSNR 和 QP 的均值、方差、最大值和最小值几个方面衡量视频质量的稳定性。从表中可以看出:和 TM5 码率控制算法相比,在相同的平均码率下,VBR 码率控制算法能够提供更加稳定的图像质量,PSNR 的动态变化范围和方差、量化因子的变化范围和方差都比 TM5 小得多。其中 stefan 在 1.5Mbps 和 concat. seq 在 0.8Mbps 下,不但节省了比特开销,而且平均 PSNR 值比 TM5 高。

5 结论

本文实现了一个基于 R-D 模型的单通道 VBR 码率控制算法。它不需要对视频序列进行多次编码,预先获得码流的整体统计信息,因而计算复杂度低,可用于实时视频编码。它利用 R-D 模型在帧层调节量化因子,并尽量减小量化因子波动,使得算法对场景变化有很强的适应能力。试验结果表明,本文提出的算法在满足整体比特约束下可获得了一个相对稳定的视频质量。虽然该方法以 MPEG-2 算法为例,但码率控制算法有相对独立性,也很容易集成到其它基于 MC/DCT 算法的视频编码算法中。

参考文献:

[1] ISO/ IEC JTC1/ SC29/ WG11, MPEG-2 Test Model 5 (TM5) [S]. 1993.

[2] Linin, Wang. Rate control for MPEG video coding[J]. Signal Process- ing: Image Communication 2000, 15(6): 493 - 511.

表 1 VBR 算法和 TM5 性能比较

序列	算法	码率 (Mbps)	PSNR (dB)				量化因子				实际码率 (Mbps)
			均值	方差	最大值	最小值	均值	方差	最大值	最小值	
stefan	TM5	1.5	30.93	2.43	34.98	25.9	24.80	10.48	54.46	11.41	1.505
	VBR		30.54	0.57	32.64	29.73	29.38	3.08	34	18	1.366
	TM5	1.2	29.68	2.42	34.25	24.91	32.69	15.52	62	12.60	1.201
	VBR		29.83	0.51	31.22	27.92	32.70	1.86	34	28	1.199
concat. seq	TM5	0.8	32.82	2.98	36.59	26.88	23.08	10.51	56.27	9.19	0.803
	VBR		31.52	1.10	33.29	29.77	30.75	1.98	34	24	0.607
	TM5	0.5	30.31	3.04	34.14	25.51	39.34	17.83	62	14.04	0.500
	VBR		30.73	1.62	33.09	27.45	37.65	7.71	52	28	0.500

[3] ISO/ IEC JTC1/ SC29/ WG29 N3093, MPEG-4 video verification model version 15.0[S]. Dec. 1999.

[4] P H Westerink, R Rajagopalan, C A Gonzales Two-pass MPEG-2 variable-bit-rate encoding[J]. IBM J Res Develop. 1999, 43(4): 471-488.

[5] Kondo, S, Fukuda, H. A real-time variable bitrate MPEG-2 video coding method for digital storage media[J]. IEEE Trans on Consumer Electronics, 1997, 43(3): 537 - 543.

[6] Junfeng B, Chang F, Qingmin, L, and etc, Accurate estimation of R-D characteristics for rate control in real-time video coding[A]. IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing (ICASSP '01) [C]. Salt Lake City: UT USA, 2001, (3): 1381 - 1384.

[7] Michael R Izquierdo, Douglas S Reeves. A survey of statistical source models for variable-bit-rate compressed video[J]. Multimedia System, 1999, 7(3): 199-213.

[8] 李 斌, 徐维朴, 郑南宁. 一种基于片层的 MPEG-2 视频编码速率控制方法[J]. 电子学报, 2001, 29(6): 853 - 856.

作者简介:



陈国斌 男, 1972 年 6 月出生于江苏省盐城
市, 2003 年博士毕业后留校, 现为浙江大学信息
与通信工程研究所讲师, 主要研究方向为多媒体
传输和图像压缩. Email: chenguobin @zju. edu. cn



刘济林 男, 1947 年 11 月出生于湖北省孝
感市, 现为浙江大学信息与通信工程研究所教
授, 博士生导师, 主要研究方向为: 数字信号处
理、智能交通、计算机视觉。

孙小叶 女, 1979 年 8 月生于江苏太仓市, 硕士生, 主要研究方向
为图像压缩。