

一种低码率下的新型码率控制策略^{*})

朱光喜 戴声奎 李 霄 刘文予 张江山

(华中科技大学电子与信息工程系 武汉 430074)

摘 要 在大多数视频压缩标准中,码率控制方法都是针对中高码率情况而设计的,而在较低码率情况下的工作情况都不够理想。针对这个问题,本文提出了一个较低码率情况下的码率预测模型,并在该模型基础上建立起一种针对较低码率情况的码率控制策略。实验结果证明了这种码率控制策略的有效性。

关键词 视频压缩,码率控制,有效比特位,低码率

A New Rate Control Scheme on Low Bit Rate

ZHU Guang-Xi DAI Sheng-Kui LI Xiao LIU Wen-Yu ZHANG Jiang-Shan

(Department of E&I Engineering, Huazhong Univ. of Sci. & Tech. Wuhan 430074)

Abstract Most of the bit-rate control schemes in video-compression standards are designed for medium and high bit-rate conditions, while do not work well for low bit-rate. In this paper, a novel bit-rate predication model is built, and then a new rate control scheme is proposed to solve this problem. The experiments prove that the new scheme is very effective.

Keywords Video compression, Rate control, Valid bits, Low bit-rate

1 引言

码率控制虽然不是视频压缩标准的一部分,但其重要性显而易见。常见的编码标准都采用了其独特的码率控制算法,如 MPEG-2 的 TM5^[1], MPEG-4 的 VM16^[2], H. 263 中的 TMN8^[3]、TMN12^[4] 等。在最新的国际编码标准 H. 264/AVC 中,码率控制有两个比较优秀的提案: JVT-F086^[5] 和 JVT-G012^[6]。码率控制算法一般依据其码率预测模型 $QP=f(R)$, 按照输出码流的大小对量化系数进行调整,从而达到控制输出码率的目的。在中高码率情况下,以上码率控制方法都有着较好的工作效果。

随着社会的发展,人们在对移动通信环境下进行多媒体通信的需求越来越强烈。如标准 H. 264/AVC 在制定之初,就考虑到了面向移动流媒体、移动多媒体通信等新业务的需求。从 2004 年 3 月起,国内的“中国数字音视频编码标准化工作组”(AVS)开始制定面向移动多媒体通信等新业务的音视频标准 AVS-M。但在移动多媒体通信等码率较低的应用中,目前各标准中的码率控制方法的性能都有所下降。例如,对运动比较剧烈的标准序列(如 Football 序列),在低帧率、低码率时, JVT-F086 和 JVT-G012 算法在输出码率和 PSNR 值的波动性、输出码率的准确性方面都不尽理想。因此有必要专门针对低码率进行研究,设计一种新的码率预测模型和算法,从而实现码率较低时的有效控制。

针对此问题,本文首先根据一帧数据的实际编码比特与有效比特位之间的关系,建立了一种基于图像帧级的码率预测模型,然后基于此模型提出一种新的码率控制策略。本文方法在 H. 264/AVC 和 AVS-M 中进行了实验,结果表明在

低码率情况下可取得较好的控制效果,适用于移动多媒体通信等应用之中。

2 相关原理

2.1 有效比特位

文[7]中提出了一个有效比特位的定义,对于 DCT(或 ICT)系数 Y ,如果编码器的最小量化步长为 X ,其有效比特位定义为:

$$\text{ValidBit}(Y) = \begin{cases} 0 & Y=0 \text{ 时} \\ \text{int}(\log_2 \text{abs}(Y/X)+1) & Y \neq 0 \text{ 时} \end{cases} \quad (1)$$

同时文[7]验证了有效比特位与输出码率之间的相互关系:编码输出码率会随着有效比特位的增加而增加。由于 JVT 中最小量化步长为 0.625,因此本文修改(1)如下式(1')所示。

$$\text{ValidBit}(Y) = \begin{cases} \text{int}(\log_2 (\text{abs}(Y/0.625))+1) & Y > 0.625 \\ 0 & Y \leq 0.625 \end{cases} \quad (1')$$

2.2 码率预测模型

文[7,8]讨论了宏块有效比特位 V_B 和纹理编码输出码率之间的关系,提出了码率的预测模型,然后在此基础上建立了宏块级的码率控制算法。但测试发现,这种预测模型比较适合于中高码率,在低码率条件下,其预测模型不能有效地预测输出码率。为了解决这一问题,本文提出了一种低码率下的码率预测模型。类似文[7],定义 RD 优化后一个宏块内所有亮度块整数变换后系数有效比特位之和为宏块有效比特位 V_B ;一帧中所有宏块有效比特位之和定义为帧有效比特位 V_F 。本文就是使用帧有效比特位 V_F 来预测一帧图像编码的

^{*}) 本课题受 863 重大专项(2002AA119010)和华为基金资助。朱光喜 教授,博士生导师,研究方向为宽带无线系统、B3G 移动通信等。戴声奎 博士生,研究方向为图像处理、视频编码和多媒体通信等。李霄 硕士生,主要研究方向为视频编码和多媒体通信。刘文予 教授,博士生导师,研究方向为多媒体通信、数字媒体安全等。张江山 副教授,硕士生导师,主要研究方向为多媒体通信、数据编码、信号处理等。

实际输出码率。

对 QCIF 格式的 Tempete 序列,利用当前最新的国际编码标准 H. 264,按照文[9]对应的测试条件,统计量化系数分别为 25,27,29,31,33 和 35 时 P 帧编码累计的宏块有效比特位与对应的实际输出比特数之间的关系,得到如图 1 所示的关联曲线。曲线中纵坐标为一帧图像中亮度块编码后输出的实际比特数之和 *actual-bit*,横坐标为各个宏块的有效比特位 V_B 的累加值。因此,每条线段最后一个点的横坐标就是一帧的帧有效比特位 V_F ,纵坐标为该帧编码后输出的实际比特数 *Frame-actual-bit*。

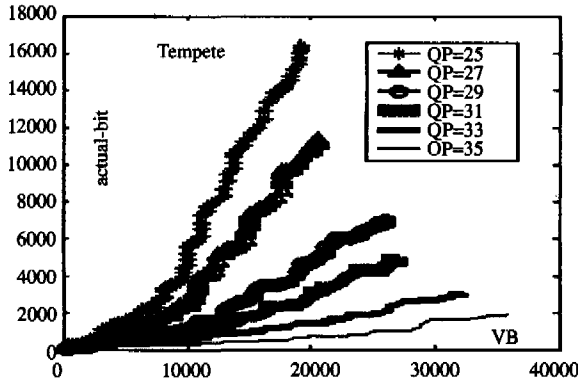


图 1 关联曲线

根据图 1 可知,每帧的帧有效比特位 V_F 与实际比特数 *Frame-actual-bit* 近似呈线性关系,因此可以假设它们之间比例参数为 K ,如式(2)所示。

$$\text{Frame-actual-bit} = K \times V_F \quad (2)$$

而在文[8]中提出了码率 R 与量化参数 QP 呈比例关系,其关系如式(3)所示。

$$R = \frac{a}{QP} + \frac{b}{QP^2} \quad a, b \text{ 均为待定系数} \quad (3)$$

由图 1 可知, K 与量化参数 QP 成反比,且通过实验测试验证近似呈二次曲线关系。这样,根据式(2)和式(3),可以建立起一个帧级的码率预测模型,如式(4)所示。

$$\text{Frame-actual-bit} = \left(\frac{a}{QP} + \frac{b}{QP^2} \right) \times V_F \quad (4)$$

a, b 均为待定系数

3 码率控制策略

在中低码率条件下,由于每个宏块纹理编码所使用的实际比特数为 0 的概率较大,输出比特中大部分是宏块头信息,宏块级码率控制性能反而不如帧级码率控制。而由图 1 可知,本文提出的码率预测模型可以较为精确地预测一帧图像编码后的输出码率。所以,本文提出的是一种帧级码率控制算法,首先确定一个 GOP 中 I 帧的量化系数,然后再根据剩余码率等参数确定 GOP 中其余 P 帧的量化系数。

3.1 I 帧码率控制

对于第 1 个 GOP 中的 I 帧,本文按下面的公式计算其量化系数 $QP-I_1$

$$QP-I_1 = \begin{cases} 40 & 0 < bpp \leq l_1 \\ 30 & l_1 < bpp \leq l_2 \\ 25 & l_2 < bpp \leq l_3 \\ 15 & l_3 < bpp \end{cases} \quad bpp = \frac{\text{bit-rate}}{\text{frame-rate} \times N_{\text{pixel}}} \quad (5)$$

其中, bit-rate 表示控制的目标码率; frame-rate 表示输出帧率; N_{pixel} 表示图像中像素的个数。实验得到:视频序列为 QCIF 格式时, $l_1=0.1, l_2=0.3, l_3=0.6$; 视频序列为 CIF 格式时, $l_1=0.2, l_2=0.6, l_3=1.2$; 视频序列分辨率大于 CIF 格式时, $l_1=0.6, l_2=1.4, l_3=2.4$ 。

3.2 P 帧码率控制

确定一个 GOP 中 P 帧的量化系数 QP 值的具体步骤如下:

(1) 如果第 i 个 GOP 中总共有 N 帧,定义目标码率为 Bit-Rate ,目标帧率为 Frame-Rate ,第 $j-1$ ($j>1$) 帧编码时实际编码使用的比特数为 $\text{Actual-Bit}(i, j-1)$,则求取第 j 帧编码时 GOP 中剩余可编码比特数的计算公式如下:

$$\text{UnAllocated-Bit}(i, j) = \begin{cases} N \times \text{Bit-Rate} / \text{Frame-Rate}, & j=1 \text{ 时} \\ \text{UnAllocated-Bit}(i, j-1) - \text{Actual-Bit}(i, j-1), & j>1 \text{ 时} \end{cases} \quad (6)$$

(2) 定义第 $i-1$ ($i>1$) 个 GOP 中第 N 帧编码时缓冲区占用率为 $V_{i-1}(N)$ 。可以采用下面公式计算第 i 个 GOP 中第 1 帧的虚拟缓冲区占用率 $V_i(1)$ 和第 j 帧的虚拟缓冲区占用率 $V_i(j)$ 。

$$V_i(1) = \begin{cases} 0 & i=1 \text{ 时} \\ V_{i-1}(N) & i>1 \text{ 时} \end{cases} \quad (7)$$

$$V_i(j) = V_i(j-1) + \text{Actual-Bit}(i, j-1) - \text{Bit-Rate} / \text{Frame-Rate} \quad j>1$$

(a) 首先对第 1 个 GOP 中的第 1 个 P 帧进行编码。在本文的算法中,认为 P 帧编码的目标比特数应该小于 I 帧编码的目标比特数。因此,对于第 1 个 GOP 中的第 1 个 P 帧,其量化系数采用 $QP-I_1+5$ 。

(b) 第 1 个 GOP 中的第 1 个 P 帧编码后,目标缓冲区占用率 $S_1(2)$ 初始化为 $V_i(2)$ 。

$$S_1(2) = V_i(2) \quad (8)$$

如果 $N_p(i)$ 代表第 i 个 GOP 中 P 帧的个数,对于第 1 个 GOP 中第 j ($j>1$) 个 P 帧,目标缓冲区占用率定义为:

$$S_1(j) = S_1(j-1) - \frac{S_1(2)}{N_p(1)-1} \quad (9)$$

(c) 第 i ($i>1$) 个 GOP 中的 I 帧编码后,目标缓冲区占用率初始化为:

$$S_i(1) = V_i(1) \quad (10)$$

对于第 i ($i>1$) 个 GOP 中第 j ($j>1$) 个 P 帧,目标缓冲区占用率定义为:

$$S_i(j) = S_i(j-1) - \frac{S_i(1)}{N_p(i)} \quad (11)$$

(3) 计算第 i 个 GOP 中的第 j ($j>1$) 帧(P 帧)编码的目标比特数 $T_i(j)$ 。

$$T_i(j) = 0.5 \times \left(\frac{\text{Bit-Rate}}{\text{frame-rate}} + \frac{\text{UnAllocated-Bit}(i, j)}{N_p(i)} + 0.5 \times (V_i(j) - S_i(j)) \right) \quad (12)$$

(4) 计算码率预测模型参数 a 和 b 。

图像序列的率失真特性随着时间变换而改变,因此需要根据已编码帧的统计数据计算编码模型式(4)中的参数。类似 JVT-G012^[6]算法,我们也采用滑动窗口技术来更新编码模型公式中的参数。滑动窗口的最大尺寸限制了模型更新的效果。通过实验测试比较,设置一个最大尺寸为 15 的滑动窗口来选择预测数据更为合理。

对于视频序列的第二个和以后的 P 帧,通过滑动窗口来选择已编码帧的统计数据。对于第二个 P 帧,滑动窗口的尺寸为 1,并把第一个 P 帧的编码数据(编码比特数、量化参数和变换参数)保存在这个窗口中;对于第三个 P 帧,滑动窗口的尺寸为 2,并把第一个 P 帧的编码数据(编码比特数、量化参数和变换参数)保存在第一个窗口中,第二个 P 帧的编码数据(编码比特数、量化参数和变换参数)保存在第二个窗口

$$b = \frac{w \sum_{i=1}^w \text{Frame_actual_bit}(i) - w \sum_{i=1}^w \text{Frame_actual_bit}(i) \times \frac{QP(i)}{V_F(i)} \times \sum_{i=1}^w \frac{QP(i)}{V_F(i)}}{w \times \sum_{i=1}^w \frac{V_F(i)^2}{QP(i)^2} - (\sum_{i=1}^w \frac{QP(i)}{V_F(i)})^2}$$

$$a = \frac{\sum_{i=1}^w [\text{Frame_actual_bit}(i) \times \frac{QP(i)}{V_F(i)} - \frac{V_F(i)}{QP(i)} \times b]}{w} \quad (13)$$

(5) 计算第 i 个 GOP 第 j 帧量化系数 $QP_i(j)$ 。

求出了 P 帧编码的目标比特数 $T_i(j)$ 后,然后根据码率预测模型参数 a 和 b ,把目标比特数 $T_i(j)$ 作为 Frame_actual_bit 代入码率预测模型式(3)中,从而求取出理想量化系数 $\text{Ideal_QP}(j)$ 。采用理想量化系数 $\text{Ideal_QP}(j)$ 与此 GOP 中其他帧量化参数的平均值 $\text{Average_QP}(j)$ 作为实际量化参数。

为了满足 H.264 的语法, $QP_i(j)$ 应在 0 和 51 之间。另一方面,为了获得稳定的图像质量,相邻帧的量化系数变化限制在 2 以内。

$$QP_i(j) = \min\{QP_i(j-1) + 2, \max\{QP_i(j-1) - 2, \text{Ideal_QP}(j)\}\} \quad (14)$$

4 实验结果与分析

为了验证本文算法(简称为 NewRC 算法)的有效性,在低码率和中码率下对其进行了测试,并与 JVT 提案算法 JVT-F086^[5]、JVT-G012^[6] 进行比较。测试条件见文[9]。实验结果如下所示:

(1) 在设定帧率为 5f/s 时,对 News(QCIF) 序列的输出码率结果如图 2 所示。

(2) 在设定帧率为 5f/s 时,对 News(QCIF) 序列的输出 PSNR 值结果如图 3 所示。

(3) 在设定帧率为 15f/s 时,对 Mobile(QCIF) 序列的输出码率结果如图 4 所示。

(4) 在设定帧率为 15f/s 时,对 Mobile(QCIF) 序列的输出 PSNR 值结果如图 5 所示。

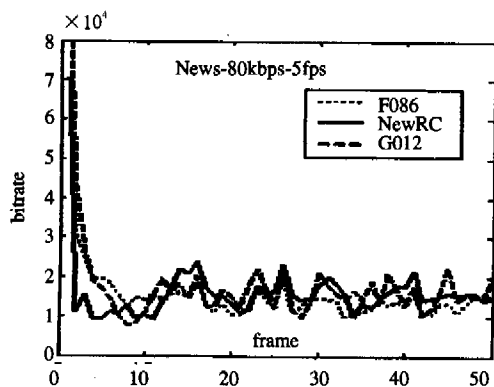


图 2 News 序列比特率变化曲线

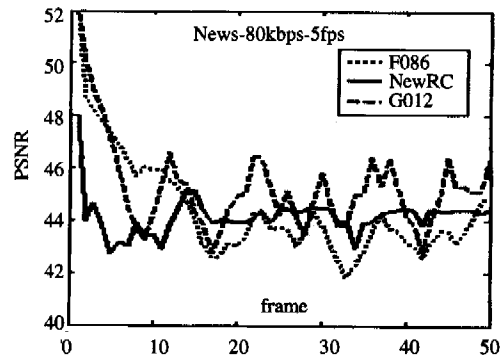


图 3 News 序列信噪比变化曲线

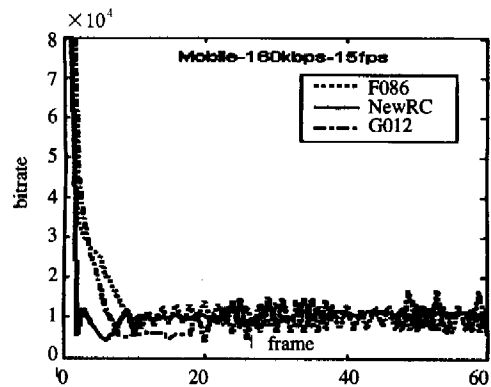


图 4 Mobile 序列比特率变化曲线

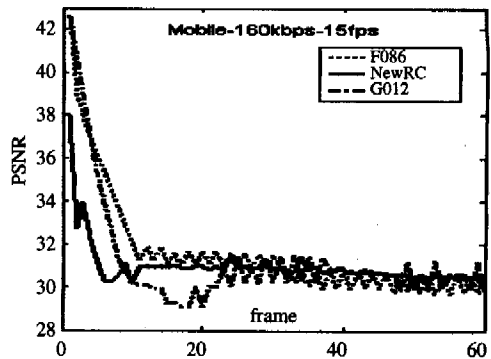


图 5 Mobile 序列信噪比变化曲线

中,依次类推,但对第十六个及其以后的 P 帧,就要把滑动窗口中保存的最先编码帧的编码数据去掉,而把当前帧的编码数据存入滑动窗口中,滑动窗口中尺寸保持为 15 不变。

假设滑动窗口的实际尺寸为 w ,第 i 个滑动窗口中存储的编码数据分别为 $VF(i)$ 、 $QP(i)$ 和 $\text{Frame_actual_bit}(i)$,可以根据下面公式来计算这一帧的码率预测模型参数 a 和 b 。

由图 2 和图 4 中可以看到,就输出比特率而言,本文提出的新算法比 JVT-F086 算法和 JVT-G012 法都要平稳;由图 3 和图 5 中可以看到,就 PSNR 值的波动性而言,本文算法比这两种算法的输出 PSNR 值平稳。

(5) 为了能够客观地评价各种算法的性能,本文定义了三种评价方式:(a) 平均码率误差,即实际输出码率平均值与目标码率的差,它反映了码率控制的精度;(b) 码率偏移 bit ...

rate-deviation,若第*i*帧的实际码率为 $actual_bit_rate_i$,目标码率为 $target_bit_rate$,按式(13)求取*N*帧图像的码率偏移值。码率偏移值反映了码率变化的剧烈程度;(c)平均质量,即实际输出图像质量的平均值。

$$bit_rate_deviation = \sum_{i=1}^N \left| \frac{actual_bit_rate_i - target_bit_rate}{target_bit_rate \times N} \right| \quad (13)$$

对 AVS 组织规定的七种测试序列,在设定帧率为 15f/s 时,采用四种典型的码率进行测试,得到以上三种评价方式的结果如表 1 所示。

表 1 三种码率控制算法的比较

序列	算法	目标码率 (kbps)	平均码率误差 (kbps)	码率偏移	平均质量 (dB)
Football QCIF	F086		0.28	0.35	27.20
	G012	64	7.3	0.58	27.13
	NewRC		2.45	0.25	27.49
Football QCIF	F086		0.53	0.20	29.09
	G012	96	7.41	0.43	29.09
	NewRC		2.1	0.31	29.12
Mobile QCIF	F086		0.75	0.36	27.52
	G012	64	11.16	0.40	27.85
	NewRC		1.31	0.24	26.67
Mobile QCIF	F086		0.96	0.32	29.57
	G012	96	11.28	0.42	29.89
	NewRC		3.14	0.23	28.53
Foreman QCIF	F086		0.24	0.28	33.5
	G012	64	2.64	0.35	33.91
	NewRC		0.29	0.31	33.49
Foreman QCIF	F086		0.35	0.26	35.38
	G012	96	2.81	0.34	35.68
	NewRC		0.48	0.30	35.43
News QCIF	F086		0.23	0.28	37.52
	G012	64	1.52	0.32	38.41
	NewRC		0.01	0.27	37.15
News QCIF	F086		0.08	0.26	40.05
	G012	96	3.39	0.30	40.55
	NewRC		0.17	0.26	40.02
Paris QCIF	F086		0.96	0.21	36.22
	G012	128	8.75	0.24	36.64
	NewRC		2.63	0.20	35.92
Paris QCIF	F086		1.09	0.17	37.7
	G012	160	8.95	0.21	38.13
	NewRC		3.67	0.19	37.35
Tempete QCIF	F086		1.07	0.34	31.07
	G012	128	9.6	0.34	31.42
	NewRC		4.74	0.20	31.42
Tempete QCIF	F086		1.55	0.27	32.26
	G012	160	9.37	0.27	32.54
	NewRC		4.59	0.19	32.44
Bus QCIF	F086		0.86	0.49	30.63
	G012	128	16.46	0.44	31.27
	NewRC		4.27	0.26	31.07
Bus QCIF	F086		2.45	0.34	32.19
	G012	160	16.25	0.33	32.62
	NewRC		1.33	0.21	32.30

由表 1 的客观评价结果可以知道本文算法的码率偏移量较小;与 JVT-G012 相比,本文算法的平均目标码率控制也要精确得多;此外,在帧率为 15fps,目标码率为 128kbps 和 160kbps 时对 football 序列使用 F086 算法会产生输出码率与目标码率相差极大的情况,其鲁棒性不好;就总体性能而言,本文算法在平均目标码率控制、码率偏移和平均质量三个方面性能都有了明显提高。

综上所述,本文算法在:(1)码率控制的精度(平均码率与目标码率的接近程度);(2)输出码率的平稳性(码率偏移较小);(3)恢复质量;(4)算法健壮性(适应应用几乎所有标准测试序列)等四个方面都要优于 F086 和 G012 算法。

本文算法作为一种新型的码率控制技术已申请相应专利,另已向 AVS 工作组提交提案,并被采纳应用于 AVS-M 测试软件之中。

结论 本文针对低码率下的码率控制问题,建立了一个基于实际比特位与有效比特位之间关系的码率预测模型,并提出了一种新的码率控制策略。本文的码率控制算法简单实用,具有可操作性;与 JVT-F086 和 JVT-G012 相比较,输出码流的 PSNR 值和码率波动曲线要相对平稳;相对于 JVT-F086 具有更强的鲁棒性,能有效地运用于七个标准测试序列;相对于 JVT-G012,本算法的实际输出码率更接近于目标码率。实验结果证明了本文的码率控制思想和算法是有效可行的。

参考文献

- 1 Test Model 5, www.mpeg.org/MPEG/MSSG/tm5
- 2 MPEG-4 Video Verification Model version 16.0, ISO/IEC JTC1/SC29/WG11 N3312
- 3 Ribas-Corbera J, Lei S. Rate Control in DCT Video Coding for Low-Delay Communications. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(1): 172~185
- 4 Test model 12/Appendix II of H. 263 Version 3, Tenth Meeting, Portland, OR, USA, Aug. 2000
- 5 Ma SiWei, et al. JVT-F086-L. doc, Proposed Draft Description of Rate Control, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 6th Meeting, Awaji, Island, JP, Dec. 2002
- 6 Li Zhengguo, et al. JVT-G012r1. doc, Adaptive Basic Unit Layer Rate Control for JVT, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 7th Meeting, Pattaya II, Thailand, March 2003
- 7 Lin Wanrong, et al. Bit number prediction for VLC coded DCT coefficients and its application in DV encoding transcoding, Multimedia Signal Processing. In: 2001 IEEE Fourth Workshop on, Oct. 2001, 135~140
- 8 Chiang T, Zhang Ya-Qin. A New Rate Control Scheme Using Quadratic Rate Distortion Model. IEEE Transactions on Circuits and Systems for Video Technology, 1997, 7(1)
- 9 AVS-M 测试通用条件, AVS-N1093. doc, 2004 年 6 月北京会议