

一种低码率下的新型宏块级码率控制算法

周全¹⁾ 刘文予²⁾ 崔春晖³⁾

(华中科技大学电子与信息工程系, 武汉 430074)

摘要 针对 H. 264 编码方案, 提出了一种实时的无需二次编码的低码率宏块级码率控制策略。为克服 H. 264 码率控制模型的缺陷, 首先提出了平均像素信息比特(APiB)的概念及其码率控制模型; 然后在宏块层, 根据各个宏块不同的编码复杂度进行准确的码字分配, 接着通过 APiB 的变化来自适应地更新码率控制模型的系数, 最后提出了完整的基于宏块的码率控制算法, 并将其在 JVT 的 JM85 平台上实现。试验结果表明, 与 H. 264 中的码率控制算法 JVT-G012 相比, 该算法不仅在控制精度上平均提高了 0. 171 kbps, 而且峰值信噪比平均提高了 0. 227 dB, 同时还较好地控制了输出码率和峰值信噪比的波动。

关键词 码率控制 H. 264 视频编码标准 平均像素信息比特 宏块

中图法分类号: TN911. 81 文献标识码: A 文章编号: 1006-8961(2006)09-1217-06

A Novel Macro-block Layer Rate Control Algorithm in Low Bit-rate

ZHOU Quan¹⁾, LIU Wen-yu²⁾, CUI Chun-hui³⁾

(Electronics & Information Engineering Department, Huazhong University of Science & Technology, Wuhan 430074)

Abstract A real-time single-pass bit rate control algorithm in low bit-rate of macro-block layer for H. 264 is presented. In order to solve the drawback of rate control model in H. 264, a new concept of average pixel information bit (APiB) and the novel rate control model based on it are represented at first; then the more precise bit allocation scheme is implemented according to the coding complexity of each macro-block, and the parameters of rate control model are updated by the change of APiB adaptively; finally, the complete rate control algorithm is proposed. This algorithm is implemented in JM85(H. 264/AVC), the experimental results show that, compared to JVT-G012 in H. 264, the proposed method gains 0. 171 kbps in control precision and 0. 227 dB in PSNR averagely, and controls the fluctuation on bit and PSNR precisely at the same time.

Keywords bit rate control, H. 264 video coding standard, average pixel information bit(APiB), macro-block

1 引言

为了在给定码率下能高效地传输编码数据, 并在接收端保证有良好和稳定的接收质量, 码率控制已成为视频编码的重要内容之一, 现有的视频压缩标准都对其进行了广泛的研究。H. 264 编码标准将量化参数同时用于码率控制和率失真优化, 但这却导致了“蛋鸡”悖论(egg-chicken absurdity): 为了有效地进行宏块的编码, 必须先给出该宏块的量化等级, 而量化等级则要根据编码之后的帧内预测或帧

间运动估计残差(mean absolutely difference, MAD)来决定。由此可见, 以往各种经典的码率控制方法在 H. 264 中并不适用^[1-3]。

由于 MAD 不能在量化前得到, 故 Li Zhengguo 提出了 JVT-G012 码率控制算法, 即先通过对当前宏块的 MAD 进行线性预测来解决这个问题, 然后使用 MPEG-4 中的二次码率预测模型(R-D 预测模型)来计算量化等级^[4]。但是, 该算法并没有突破“蛋鸡”悖论的瓶颈, 而且预测出的 MAD 并不能完全准确地反映当前宏块的编码复杂度, 另外, 以此得到的量化等级也不一定准确, 这可能导致实际编码

码字与目标码字不匹配。在目标码率和量化参数的关系上,Lei Shawmin 等人从 DCT(discrete cosine transforme)变换的角度来考虑码率控制^[5]。He Zhihai 等人又做了进一步的研究,即通过 DCT 变换量化后的零值在变换系数中的比例来确定量化参数^[6,7]。由此可知,DCT 系数的变化情况与量化参数存在一定的关系。因此,本文在 DCT 系数的基础上提出平均像素信息比特(average pixel information bit, APIB)的概念,在 MPEG-4 中的 R-D 预测模型中用 APIB 代替 MAD,以便从根本上摆脱“蛋鸡”悖论的束缚。由于 APIB 既能在量化前得到,又能比较准确地反映出当前宏块的编码复杂度,因此以此确定的量化等级更加准确。

2 平均像素信息比特及其码率预测模型

2.1 平均像素信息比特 (APIB)

为了克服 JVT-G012 算法中“蛋鸡”悖论的缺陷,在此首先提出 APIB 的概念。由于 DCT 系数是原始图像与参考图像经过运动补偿后的残差信息通过 DCT 变换来得到的,因此,DCT 系数的大小就可直接反映出图像的运动剧烈情况和内容的复杂程度。通常较大的 DCT 系数说明原始图像与参考图像间存在较大的残差冗余信息,这一般意味着图像运动剧烈或者内容变化较大;反之,较小的 DCT 系数则说明原始图像与参考图像间只有较小的残差冗余信息,这意味着图像运动平缓或者在内容上基本没有变化。

对于宏块的每个像素而言,APIB 的值不仅与 DCT 系数有关,而且还会受量化步长的影响。如果编码器选择的量化步长较大,则有较多的 DCT 系数被量化为零,这就会导致图像的一些运动信息和内

容信息被量化掉,从而降低了宏块的编码复杂度;反之,如果编码器选择较小的量化步长,则有较多的 DCT 系数被保留,这种细量化虽保留了较多的图像运动信息和内容信息,但却提高了宏块的编码复杂度。

从上面的分析看出,宏块的 APIB 值应该与 DCT 系数的绝对值大小成正比,而与编码器的量化步长成反比。由此,一个宏块的 APIB 值定义如下:

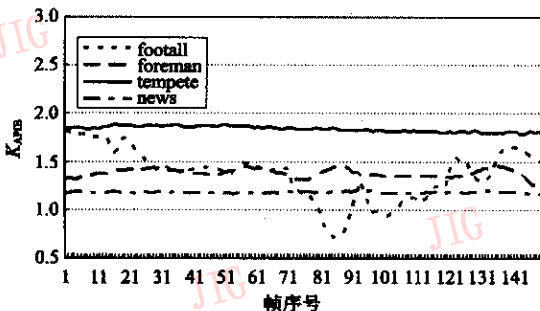
$$K_{\text{APIB}} = \sum_{i=0}^{N_{\text{pixel}}-1} \frac{\lg(|Y_i/Q_{\min}| + 1)}{N_{\text{pixel}}} \quad (1)$$

式中, Y_i 表示宏块内第 i 个像素的 DCT 系数, Q_{\min} 表示编码器的最小量化步长, N_{pixel} 表示宏块内所有像素的个数。由于 H.264 标准中定义的最小量化步长为 0.625,所以式(1)可以改写为

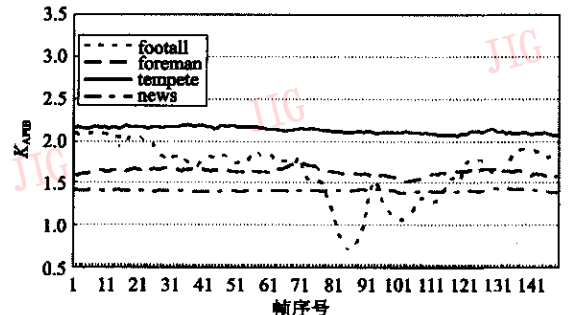
$$K_{\text{APIB}} = \begin{cases} \sum_{i=0}^{N_{\text{pixel}}-1} \frac{\lg(|Y_i/0.625| + 1)}{N_{\text{pixel}}}, & Y > 0.625 \\ 0, & Y < 0.625 \end{cases} \quad (2)$$

通过式(2)可以看出,如果一个宏块所有像素的 DCT 系数为零,就说明该宏块没有任何信息,其 APIB 的值也为零。此外,由于 APIB 是 DCT 系数 Y_i 的单调递增函数,其值随着 Y_i 的增加而增加,因此 APIB 值的大小直接反映了宏块的编码复杂度。

图 1(a)和 1(b)分别是在 JM85 平台下,量化参数(quantify parameter QP)为 28 时, qcif 和 cif 两种格式下 foreman, news, tempete, football 4 种序列在 30fps 下全 I 帧的 APIB 函数曲线。从曲线中可以看出,对于 foreman 和 news 序列而言,因为两个序列运动平缓,且图像内容相对简单,所以 APIB 的值较小,曲线变化相对平缓;而对于 tempete 和 football 序列而言,由于这两个序列运动剧烈,且图像内容复杂,所以 APIB 的值较大,曲线变化相对剧烈,这进一步说明了



(a) qcif 格式 APIB 曲线



(b) cif 格式 APIB 曲线

图 1 foreman, news, tempete, football 4 种序列在 QP = 28, 30fps 下全 I 帧的 APIB 函数曲线

Fig. 1 The APIB function curve of four test sequences(football, foreman, tempete, news) at 30fps when QP is 28 and GOP structure is full of I-frame

APIB 的值能够反映图像的编码复杂度。

2.2 基于平均像素信息比特的码率预测模型

为了更好地使编码码率与目标码率匹配,就要使用更加准确的 R-D 预测模型。因此,本文沿用了 JVT-G012 算法中的二次 R-D 预测模型结构,而把该模型中的 MAD 值用 APIB 的值来替代。新的 R-D 预测模型定义如下:

$$R_i - H_i = \left(\frac{a}{Q_i} + \frac{b}{Q_i^2} \right) \times K_i^{\text{APIB}} \quad (3)$$

式中 R_i 表示编码第 i 个宏块的目标码率, H_i 表示宏块头、运动矢量和语法元素占用的比特开销, K_i^{APIB} 表示当前宏块的 APIB 值, Q_i 表示当前宏块的量化步长,而 a, b 为模型参数,它们在编码完一个宏块后就被更新。JVT-G012 算法之所以使用 MAD,是为了使 R-D 预测模型能够自适应图像内容的变化。通过上面的分析可看出,由于 APIB 也能很好的表达图像的编码复杂度,因此用 APIB 代替 MAD 是完全可行的。

3 基于宏块的码率控制算法

为了在准确控制码率的前提下提高编码图像质量,本文利用上节提出的二次 R-D 预测模型,提出一种基于宏块的码率控制算法。此算法包括以下两点:(1)根据 DCT 系数的变化情况,首先通过式(2)来计算当前宏块的 APIB 的值,然后对每个宏块的目标码字进行更加合理的分配,以减少因码字分配不合理而造成图像质量的下降;(2)由于本文算法可使用 APIB 代替 MAD,因此本算法可根据 APIB 值的变化自适应地更新 R-D 预测模型的系数,以达到码率控制的目的。

3.1 宏块级码字分配

在给宏块进行码字分配之前,首先要对包含该宏块的帧进行码字分配。假设图像组(group of picture, GOP)结构是第 1 帧为 I 帧,后续帧为 P 帧,且给定 GOP 内第 1 个 I 帧量化参数为 24,第 1 个 P 帧量化参数与之相同,则后续 P 帧的目标码字定义为

$$B_{j+1} = \frac{U_j}{N_p} \times (1 - \rho) + T_j \times \rho \quad (4)$$

式中 U_j 表示编码当前第 j 帧时 GOP 内剩下的可分配码字, N_p 表示 GOP 内未编码的 P 帧个数, T_j 表示编码上一 P 帧使用的实际比特数, ρ 为权重常数,一般设为 0.5。从式(4)可以看出,为了使实际编码码字能更好地匹配目标码字,可将 GOP 内的可分配码字平均分配到各帧中。此外,对目标码字的估计还

应考虑到图像间的时间相关性,即应考虑到上一帧编码实际码字对本帧目标码字的影响。

算法中,每帧第 1 个宏块的量化参数由前一帧的各个宏块的平均量化参数确定。对于后续宏块而言,则先根据式(2)计算当前编码宏块 APIB 的值。由于图像内部不同部分的运动信息和内容信息各不相同,其在相同量化等级下编码产生的码字会有较大差异,因此对于包含较多运动信息和内容信息的宏块应该分配较多的码字,而对于包含较少运动信息和内容信息的宏块则应分配较少的码字。综上所述,由于每个宏块目标码字应该与 APIB 的值的平方成比例,所以当前宏块的目标码字定义为

$$V_i = (K_{\text{APIB}}^2 \times B) / \left(\sum_i^{N_{\text{MB}}} K_{\text{APIB}}^2 \right) \quad (5)$$

式中 B 表示(4)式中分配给当前帧的目标码字, K_{APIB} 表示第 i 个宏块的实际 APIB 值, N_{MB} 表示一帧图像内所有宏块的个数。与 JVT-G012 算法不同之处在于:JVT-G012 算法在确定目标码字之前要对当前宏块的 MAD 的值做线性预测,而本算法则直接使用当前宏块的实际 APIB 值,且不需要对其进行预测,因此,用 APIB 确定的目标码字比用 MAD 确定的目标码字更加准确。

3.2 码率预测模型系数更新

二次 R-D 预测模型的系数是根据已确定个数的已编码宏块的编码结果来时更新的。它具体包含以下 3 个步骤:

(1)选择滑动窗口大小。设置滑动窗口的目的,是为了自适应地平滑图像中内容复杂度不同的部分。滑动窗口的大小定义为

$$W = \begin{cases} \frac{K_i^{\text{APIB}} \times L_{\text{max}}}{K_i^{\text{APIB}}} K_{i-1}^{\text{APIB}} > K_i^{\text{APIB}} \\ \frac{K_{i-1}^{\text{APIB}} \times L_{\text{max}}}{K_i^{\text{APIB}}} K_i^{\text{APIB}} > K_{i-1}^{\text{APIB}} \end{cases} \quad (6)$$

式中 L_{max} 表示最大滑动窗口长度,一般设为 20。从式(6)中可看出,如果图像编码复杂度较小,即运动平缓,则前后宏块 APIB 的值相差不大,此时滑动窗口较大,这说明较多的宏块数据被用来估计当前宏块的模型参数,反之,如果前后宏块 APIB 的值相差较大,则滑动窗口相对缩小,这说明只有与当前宏块较近的宏块数据被采用。

(2)数据选择。确定了滑动窗口大小以后,并不是窗口内所有数据都被用来更新 R-D 预测模型系数。根据最小均方估计的理论,滑动窗口内所有

宏块的实际编码码字与目标码字的标准偏差定义为

$$e_i = \sqrt{\frac{1}{\hat{W}} \sum_{i=0}^{W-1} \left(\frac{a \times K_i^{APIB}}{Q_i} + \frac{b \times (K_i^{APIB})^2}{Q_i^2} - R_i \right)^2} \quad (7)$$

窗口内每个宏块实际编码码字与目标码字的实际偏差定义为

$$\hat{e}_i = \frac{a \times K_i^{APIB}}{Q_i} + \frac{b \times (K_i^{APIB})^2}{Q_i^2} - R_i \quad (8)$$

式(7)和式(8)中, \hat{W} 表示由式(6)中计算出来的滑动窗口大小, Q_i 表示量化步长, R_i 表示实际的编码码字。

如果实际偏差超过标准偏差, 就认为这个宏块的编码数据不可取, 其数据不能被用来对 R-D 预测模型参数做最小均方估计。为了避免窗口内所有的数据都被滤除, 应始终保留当前编码宏块的前一个宏块的数据。

(3)更新模型系数。本文沿用了 JVT-G012 算法中的线性衰减技术, 式(3)中两个模型参数的更新定义如下:

$$b = \frac{\hat{W} \times \sum_{i=1}^{\hat{W}} \frac{R_i}{K_i^{APIB}} - \left(\sum_{i=1}^{\hat{W}} Q_i^{-1} \right) \times \left(\sum_{i=1}^{\hat{W}} \frac{R_i \times Q_i}{K_i^{APIB}} \right)}{\hat{W} \times \sum_{i=1}^{\hat{W}} Q_i^{-2} - \left(\sum_{i=1}^{\hat{W}} Q_i^{-1} \right)^2} \quad (9)$$

$$a = \frac{\sum_{i=1}^{\hat{W}} \left(\frac{R_i \times Q_i}{K_i^{APIB}} - b \times Q_i^{-1} \right)}{\hat{W}} \quad (10)$$

式中, \hat{W} 表示经过数据选择后滑动窗口的大小。

4 实验结果及分析

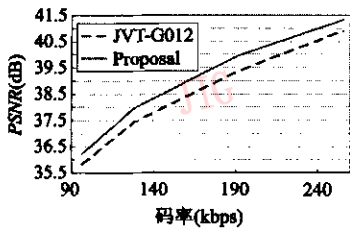
为检验本文码率控制算法的效果, 在 JVT(joint video team)的校验模型 JM85 上实现了本文提出的码率控制算法, 并与 JVT-G012 算法进行了对比。本试验采用的序列有: foreman, news, football, tempete; cif 格式序列测试帧率为 15fps, qcif 格式序列测试帧率为 5fps; GOP 结构为 IPPP; 初始量化参数设为 24, 搜索范围为 32, 参考帧为 5 帧, 熵编码方式为 CABAC (content adaptive binary coding); 执行 RDO(rate distortion optimization)。表 1 对比了本文算法和 JVT-G012 算法的控制精度和峰值信噪比。

图 2(a)和 2(b)分别是 cif 格式下的 news 序列和 qcif 格式下的 football 序列分别在 15fps 帧率和 5fps 帧率下的编码性能曲线; 图 2(c)和图 2(d)分别是两序列在码率 128kbps 和 48kbps 下的峰值信噪比波动图, 而图 2(e)和图 2(f)则分别是两序列在码率 128kbps 和 48kbps 下的峰值信噪比波动图。

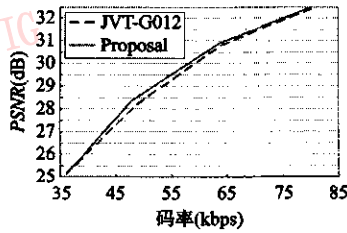
表 1 两种算法控制精度和峰值信噪比值比较

Tab. 1 Comparison of control accuracy and PSNR value with two algorithm

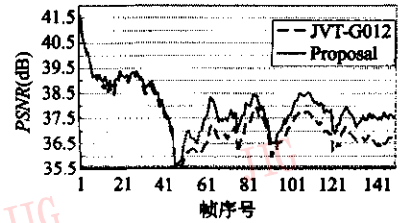
测试序列	测试码率 (kbps)	控制精度(kbps)		码率偏移(kbps)		PSNR(dB)	
		JVT-G012	本文算法	JVT-G012	本文算法	JVT-G012	本文算法
foreman (cif)	24	24.05	24.02	0.05	0.02	32.50	32.60
	48	48.03	47.94	0.03	-0.06	35.91	36.01
	64	63.99	63.95	-0.01	-0.05	37.41	37.47
	80	80.03	79.91	0.03	-0.09	38.66	38.71
	96	96.16	96.16	0.16	0.16	35.82	36.21
news (cif)	128	128.24	128.10	0.24	0.10	37.42	37.96
	192	192.26	192.10	0.26	0.10	39.36	39.94
	256	256.25	256.06	0.25	0.06	40.91	41.32
football (qcif)	24	36.03	36.03	12.03	12.03	25.10	25.10
	48	48.88	48.00	0.88	0.00	28.20	28.37
	64	64.02	63.76	0.02	-0.24	30.81	30.92
	80	80.14	79.71	0.14	-0.29	32.43	32.45
	96	96.46	96.45	0.46	0.45	24.70	24.67
tempete (cif)	128	128.64	128.6	0.64	0.60	26.92	26.93
	192	192.91	192.79	0.91	0.79	29.36	29.44
	256	257.09	256.86	1.09	0.86	30.78	30.82



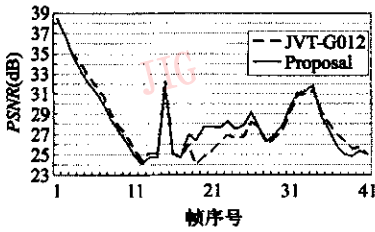
(a) cif 格式的 news 序列 帧率 15fps 时 编码性能比较曲线



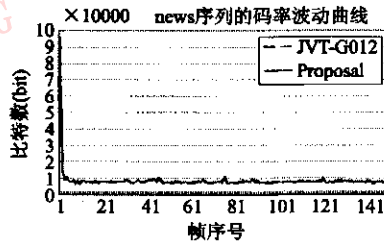
(b) qcif 格式的 football 序列 帧率 5fps 时 编码性能比较曲线



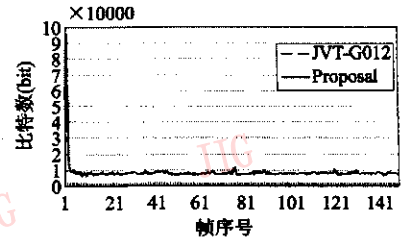
(c) cif 格式的 news 序列 帧率 15fps 时, 码率 128kbps 下的峰值信噪比波动曲线



(d) qcif 格式的 football 序列 帧率 15fps 时, 码率 48kbps 下的峰值信噪比波动曲线



(e) cif 格式的 news 序列 帧率 15fps 时, 码率 128kbps 下的比特数波动曲线



(f) qcif 格式的 football 序列 帧率 15fps 时, 码率 48kbps 下的比特数波动曲线

图 2 本文算法和 JVT-G012 算法在 H. 264/AVC 编码器下的编码性能比较曲线

Fig. 2. The comparison of coding performance curve for JVT-G012 and our proposal algorithm in H. 264 codec

从表 1 可看出,本文算法的平均码率偏移比 JVT-G012 算法减少 0.171kbps,平均控制精度优于 JVT-G012 算法,这说明本文提出的 APiB 及其 R-D

预测模型能够比较准确地估计图像编码复杂度。从表 1 和图 2 还可看出,本文算法编码图像的峰值信噪比平均提高了 0.227dB,尤其是 news 序列,平均提高

表 2 两种算法平均比特误差和平均峰值信噪比误差比较

Tab. 2 Comparison of average fluctuation error on bit and PSNR value with two algorithm

测试序列	测试码率(kbps)	平均比特误差		平均 PSNR 误差	
		JVT-G012	本文算法	JVT-G012	本文算法
foreman (cif)	24	0.194053125	0.152800666	0.062805721	0.057844611
	48	0.087808339	0.078268709	0.044963656	0.044341989
	64	0.053786906	0.063849116	0.038210308	0.037450755
	80	0.048597682	0.065355977	0.032643998	0.033718989
news (cif)	96	0.099772736	0.070877901	0.033309675	0.030263274
	128	0.070811114	0.059301013	0.024929576	0.020023237
	192	0.044527789	0.038281353	0.015538308	0.012131438
	256	0.036118220	0.046746884	0.010416431	0.009358591
football (qcif)	24	1.827078385	1.827078385	0.158092035	0.158092035
	48	0.748214286	0.685448842	0.100219832	0.092199274
	64	0.391372752	0.364661134	0.090869968	0.096475350
	80	0.265800108	0.245352909	0.091560257	0.089955110
tempete (cif)	96	0.595915362	0.593278867	0.067905028	0.067726241
	128	0.373858668	0.371056763	0.050726524	0.051688250
	192	0.178866754	0.168006338	0.030966834	0.029371363
	256	0.105574107	0.102029682	0.024164158	0.023738601

了 0.533dB。

但是,码率控制的目的是要求在固定的码率下获得最好的图像质量。这不仅要求编码每帧的实际码字与分配的目标码字相匹配,而且要求每帧图像的峰值信噪比波动平缓。因此,本文采用平均比特误差 E_{bit} 和平均峰值信噪比误差 E_{PSNR} 来分别反映图像在目标码字与实际编码码字上的匹配情况和峰值信噪比的波动情况:

$$E_{bit} = \sum_{i=1}^{N_{total}} \frac{|b_i^{actual} - b_i^{target}|}{N_{total} \times b_i^{target}} \quad (11)$$

$$E_{PSNR} = \sum_{i=1}^{N_{total}} \frac{|PSNR_i - \overline{PSNR}|}{N_{total} \times \overline{PSNR}} \quad (12)$$

式中, N_{total} 表示实际编码的帧数, b_i^{actual} 表示每帧编码的实际码字, b_i^{target} 表示每帧的目标码字, $PSNR_i$ 表示每帧的实际峰值信噪比, \overline{PSNR} 表示平均的峰值信噪比。表 2 对比了两算法的平均码率波动和平均峰值信噪比波动。

从表 2 可以看出,本文算法与 JVT-G012 算法相比,在平均比特误差上减少了 1.2%,而平均峰值信噪比误差则减少了 0.14%。这有利于低码率下的视频传输,因为它不仅降低了解码端缓冲区发生上溢和下溢的可能,同时保证了图像质量的平稳过渡,并具有较高的主观视觉质量。

5 结 论

众所周知,好的码率控制算法不仅要有准确的控制精度,而且要获得最好的图像质量。从试验数据可见,本文提出的宏块级码率控制算法能够达到这些要求。本文将 APiB 的概念引入 R-D 预测模型,这不仅从根本上避免了 H.264 中的“蛋鸡”悖论,而且只需一次编码就能达到码率控制的目的。除此之外,由于 APiB 还较好地反映了图像的编码复杂度,因此以此确定的量化参数较 MAD 更加准

确。另外,由于视频编码的计算复杂度主要集中在运动估计和帧内预测上,所以本算法不会影响编码器的实时性。

通过与 JVT 提案中的码率控制算法 JVT-G012 进行的对比试验表明,在保持编码器实时性基本不变的条件下,本文方法在控制精度上比 JVT-G012 算法平均提高了 0.171kbps,平均峰值信噪比值提高了 0.227dB。同时,本算法的平均比特误差和平均峰值信噪比误差均要优于 JVT-G012 算法,可分别减少 1.2% 和 0.14%,这说明本算法在使实际编码码字与目标码字进行更好匹配的同时,还获得了更好的主观视频质量。

参考文献 (References)

- Chiang T H, Zhang Y Q. A new rate control scheme using quadratic rate distortion mode[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, 7(1): 246 ~ 250.
- Lee H J, Chiang T H, Zhang Y Q. Scalable rate control for MPEG-4 video[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2000, 10(6): 878 ~ 894.
- Siwei Ma, Wen Gao, Peng Gao, et al. Rate control for advance video coding(AVC) Standard[A]. In: IEEE International Symposium on Circuits and Systems[C], Bangkok, Thailand, 2003, 2: 892 ~ 895.
- Li Z G, Pan F, Lim K P. Adaptive basic unit layer rate control for JVT. JVT-G012-r1[A]. In: Joint Video Team of ISO/IEC and ITU 7th Meeting[C], Pattaya, Thailand, 2003.
- Ribas-Corbera J, Lei S M. Rate control in DCT video coding for low-delay communications[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(1): 172 ~ 185.
- Kim Yong Kwan, He Zhihai, Sanjit K M. A novel linear source model and a unified rate control algorithm for H.263/MPEG-2/MPEG-4[A]. In: proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing[C], Salt Lake City, Utah, USA, 2001, 3: 1777 ~ 1780.
- Kim Yong Kwan, He Zhihai, Sanjit K M. Low-delay rate control for DCT video coding via ρ -domain source modeling[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(8): 928 ~ 940.