Frame Complexity-Based Rate-Quantization Model for H.264/AVC Intraframe Rate Control

Xuan Jing, Lap-Pui Chau, and Wan-Chi Siu

Abstract—**In this letter, we present an adaptive intraframe rate-quantization (R-Q) model for H.264/AVC video coding. The proposed method aims at selecting accurate quantization parameters (QP) for intra-coded frames according to the target bit rate. By taking gradient-based frame complexity measure into consideration, the model parameters can be adaptively updated. Experimental results show that when employing our proposed R-Q model, the intraframe target bits mismatch ratio can be reduced by up to 75% as compared to the traditional Cauchy-density-based model. Hence, this is extremely useful for H.264/AVC rate control applications.**

Index Terms—**Frame complexity, H.264/AVC, rate control, rate-quantization.**

I. Introduction

IN the latest H.264/AVC video coding standard, the QP only specifies the quantization step size which is used for scaling the coded transform coefficients. However, for the particular H.264/AVC encoder control method defined in the *joint model*, the QP also influences the rate-distortion optimized (RDO) mode decision and motion estimation process. Generally, the best coding mode for each macroblock (MB) is selected by choosing the one which can minimize the overall Lagrangian cost function $J = D + \lambda \cdot R$, where D and R are the mean square error and the number of bits for coding this MB. As QP is required when calculating the Lagrangian multiplier λ for the current MB before it is actually encoded, rate control for the H.264/AVC encoder is facing a new challenge. This phenomenon is usually referred to as a typical "chicken and egg" dilemma. Li *et al.* in JVT-G012 [1] have proposed an adaptive rate control framework for H.264/AVC. In general, for P frames, a single-pass rate control method based on the classic quadratic R-Q model is used, and a linear model for mean absolute difference (MAD) prediction is employed to solve the above dilemma. However, in the adaptive rate control scheme, no explicit R-Q model for an intraframe is discussed. Generally, the QP for an intraframe depends on the average QP of P-frames in the previous group of pictures (GOP). In addition, this QP can be adaptive to both the GOP length and the available channel bandwidth. The potential problem of this scheme is that given bit budget and buffer constraint when encoding the current intraframe, it is difficult to accurately estimate the QP since the complexity of the current frame is not considered. Thus, an efficient R-Q model for intraframes is highly desirable for H.264/AVC which must be both accurate and frame-complexity adaptive.

Although previous R-Q models [2], [3] can adaptively determine the QP for intraframes, they are not appropriate to be applied directly to H.264/AVC since actual statistics of DCT coefficients are required which are not available before RDO mode decision. Kamaci *et al.* [4] have introduced a Cauchy-density-based rate model for H.264/AVC which is more accurate for intraframes. However, with a simple scheme for updating model parameters, there is not sufficient adaptability to the changes of frame complexity. Other schemes [6], [7] aim at improving rate control for interframes without considering specific R-Q model for intraframes. In this letter, we develop an adaptive intraframe R-Q model for H.264/AVC rate control. The proposed algorithm is based on the observation that the output bit rate of each intraframe is highly correlated with its gradient-based frame complexity measure. Besides this strong correlation, there are another two advantages of using gradient as the intraframe complexity measure. One is that the computation for calculating gradient is low, and the other is that it can be obtained from the original frame without any pre-encoding process. These properties are very desirable for H.264/AVC rate control. Simulation results show that our proposed R-Q model is much more accurate than the original Cauchy-density-based method when determining QP for intraframes.

The rest of this letter is organized as follows. Section II first analyzes the gradient-based complexity measure for our R-Q model. The development of the proposed adaptive intraframe R-Q model is then discussed in Section III. Section IV demonstrates the experimental results for comparison. Finally, we draw a conclusion in Section V.

II. Gradient-Based Frame Complexity Measure

We have exhaustively encoded various test sequences only using the intra-coding scheme in the H.264/AVC using different QPs. Fig. 1(a) shows the bit rate results for the *Foreman* sequence. As it can be seen from this figure, for the same quantization value, encoding different frames will consume different numbers of bits. This difference varies with the changes of QPs. The smaller the QP, the larger the differences between frames will be observed. In addition, we can find that the shapes of these

Manuscript received July 26, 2007; revised January 22, 2008. This work was supported in part by the Centre for Multimedia Signal Processing, Department of Electronic and Information Engineering, Hong Kong Polytechnic University, and in part by the Research Grant Council of the Hong Kong SAR Government (PolyU 5268/06E). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Mila Nikolova.

X. Jing and W.-C. Siu are with the Centre for Multimedia Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong (e-mail: enxjing@polyu.edu.hk; enwcsiu@polyu.edu.hk).

L.-P. Chau is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: elpchau@ntu.edu.sg).

Digital Object Identifier 10.1109/LSP.2008.920010

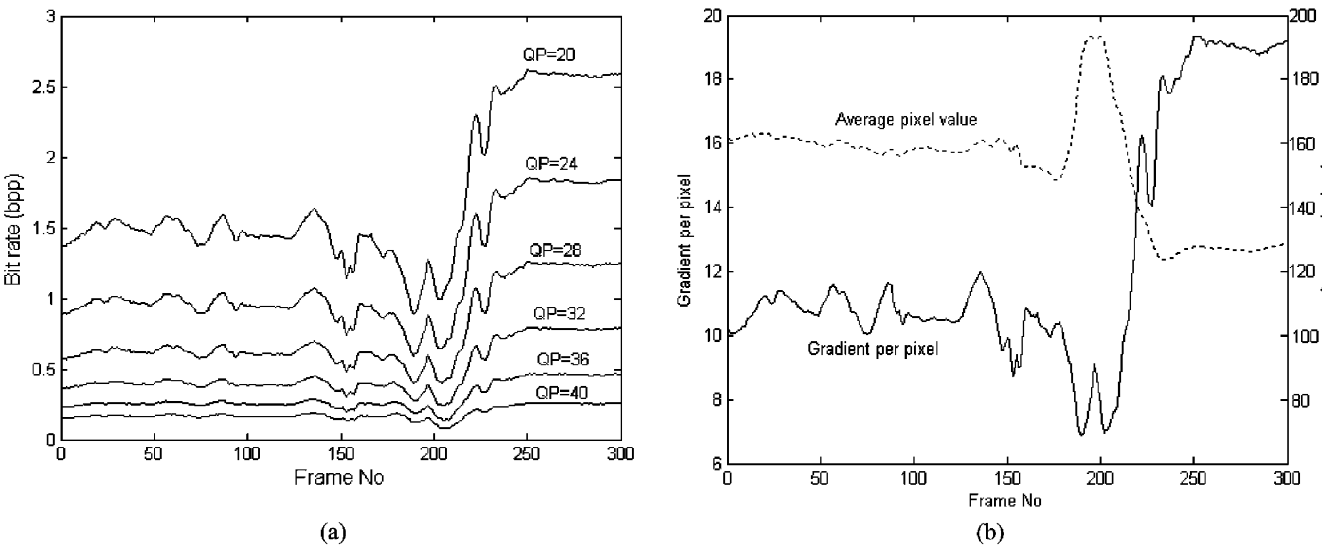

Fig. 1. (a) Intra-bit-rate of *Foreman* sequence under different QPs. (b) Gradient per pixel and average pixel value for different frames of *Foreman*.

six curves are highly correlated, which implies that the output bit rate of each intra-coded frame is inherently determined by its frame characteristics.

There are many activity measures or complexity measures for still image coding. In [5], Kim *et al.* have classified these measures into four categories and suggested that the gradient-based method is more reliable for still image coding. In our proposed intra R-Q model, we define the complexity measure of intraframe based on the average gradient per pixel of the frame. The calculation of gradient is defined by

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

where M and N are the horizontal and vertical dimensions of the frame, respectively. $I_{i,j}$ denotes the luminance value of pixel at the location of (i,j). Fig. 1(b) illustrates the average gradient value and the average pixel value for *Foreman* sequence. By comparing Fig. 1(b) with Fig. 1(a), we can find that the gradient value can better represent the complexity of an intraframe since the shapes of these two sets of curves are quite similar. For the same QP, the larger the gradient value, the larger the number of bits of this frame.

To further illustrate the relationship between the gradient and the output bit rate, we show a set of scatter plots of bit rate versus gradient value in Fig. 2. Obviously, there exists a linear relationship between these two factors. Moreover, with the QP becoming larger, the slope of each "line" is getting smaller. From this linear correlation between the number of bits and the gradient, we can assume that for a fixed QP, the output bit rate of one intra-coded frame is proportional to its average gradient.

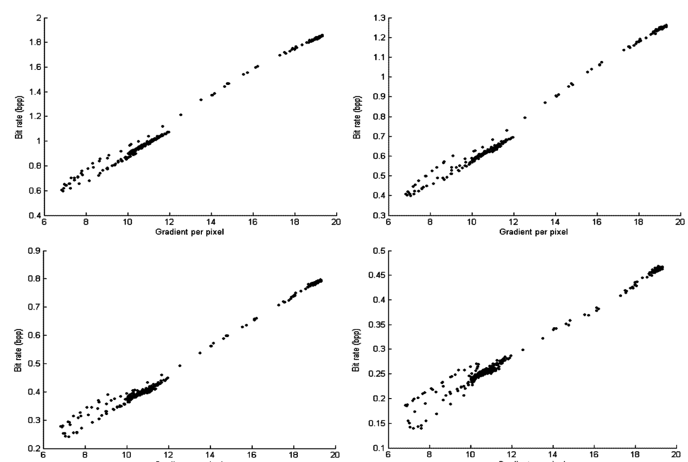

Fig. 2. Scatter plots of bit rate versus gradient value for *Foreman* sequence. Corresponding QP values (from top to bottom, left to right) are: 24, 28, 32, and 36.

III. Adaptive Intraframe R-Q Model

In the quantization scheme of previous video coding standards, the quantization parameter QP is directly used as a scaling factor which controls the coding bit rate and the picture quality. For example, in H.263, the quantization step size or scaling factor actually used has a linear relation with QP, i.e., $Q_{\text{step}} = 2QP$. This linear relation no longer exists in the H.264/AVC standard. Instead, the nonlinear relation between QP and Q_{step} in H.264/AVC is formulated as $Q_{\text{step}} = 2^{\{(QP-4)/6\}}$. Generally speaking, there are 52 values of Q_{step} which are indexed by 52 QPs ranging from 0 to 51 and the value of Q_{step} doubles for every increment of 6 in QP. Theoretically, bit rate is more directly related with Q_{step}. Although QP monotonically increases as Q_{step} increases, we model the output bit rate as a function of Q_{step} in our proposed scheme.

Experimental observations have shown that the intraframe bit rate can be formulated by

$$R(Q_{\text{step}}) = G \cdot f(Q_{\text{step}}) \quad (2)$$

$$f(Q_{\text{step}}) = a \cdot Q_{\text{step}}^{b} \quad (3)$$

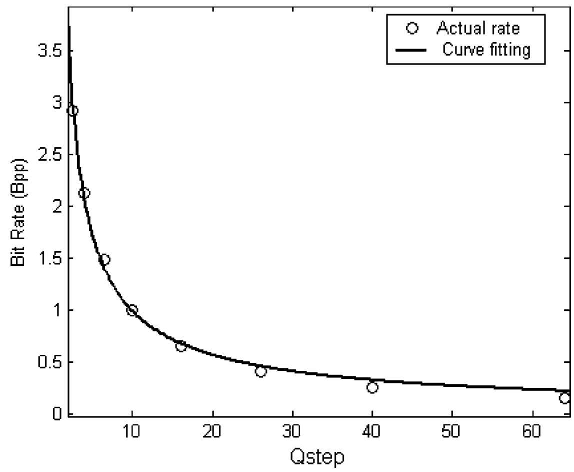

Fig. 3. Curve fitting result for the first frame of *Suzie*. The gradient value $G = 9.74$ and the model parameters in $f(Q_{\rm step}) = a \cdot Q_{\rm step}^b$ are: $a = 0.638$ and $b = -0.795$.

where G is the average gradient per pixel calculated from (1), $Q_{\rm step}$ is the quantization step size, and $a > 0$ and $b < 0$ are parameters that depend on the picture content. In [4], Kamaci *et al.* have empirically justified that the Cauchy density performs better than the traditional Laplacian density for estimating the distribution of DCT coefficients. The rate estimation accuracy is greatly improved when using Cauchy-based entropy function, especially for intra-coded frames. Furthermore, they suggested using a simplified equation as (3) to approximate the Cauchy entropy function. Based on our observations, we also adopt (3) as a part of our R-Q model (2) with the combination of the gradient value G which is an important scaling factor. In other words, our R-Q model can be explained in this way: the bit rate of one intraframe is proportional to its gradient value, and the slope of this linear relation can be represented by a power function with variable $Q_{\rm step}$. Fig. 3 gives an example of curve fitting result for the first frame of *Suzie* sequence. It shows that the proposed R-Q model achieves accurate estimation of the actual bit rate. However, we also find that it is unlikely to have one single model with fixed parameters to be used for all kinds of frames. Ideally, these two parameters (i.e., a and b) can be precisely determined by solving a set of equations based on the Cauchy density parameter [4]. Because we do not have access to the actual distribution of DCT coefficients in advance, we need to estimate and update the parameters based on previous coding results.

Based on our extensive experiments, we have found that typical values of b are in the range of -0.90 to -0.70. For simplicity, we choose a moderate value, $b = -0.80$, in our R-Q model. During our parameter updating procedure, the value of b is fixed and we only update parameter a using

$$a_{k+1} = \alpha \cdot a_k + (1-\alpha) \cdot \frac{R_k}{G_k \cdot (Q_{\rm step})_k^b} \tag{4}$$

where R_k, G_k, and $(Q_{\rm step})_k$ are the actual output bit rate (in bits per pixel), the average gradient value per pixel, and the quantization step size for the kth intraframe. Similar to [4], we adopt α as a forgetting factor with the typical value of $\alpha = 0.5$. After coding the first intraframe using the predefined $(Q_{\rm step})_0$, a_0 can be calculated as

$$a_0 = \frac{R_0}{G_0 \cdot (Q_{\rm step})_0^b}. \tag{5}$$

Comparing our proposed R-Q model with the Cauchy-based model, the major difference is that we have considered the gradient-based frame complexity measure as an important scaling factor in bit rate estimation. As a result, our model is more adaptive to the changes of frame characteristics.

IV. Experimental Results

To evaluate the performance of our proposed intraframe R-Q model, we have applied our scheme and the Cauchy method to the H.264/AVC reference encoder JM11 [8] and encoded a number of video sequences only using intra-coding mode. The experiments were conducted using the first 150 frames of six QCIF sequences (including *Foreman*, *Stefan*, *Trevor*, *Carphone*, *Suzie*, and *News*) and three CIF sequences (including *Football*, *Table Tennis*, and *Coastguard*). In addition, we synthetically generated a test sequence *Comb1* by cascading 30 frames from the six QCIF test videos (five frames from each sequence which are spaced 30 frames apart). The purpose of using *Comb1* sequence is to test the adaptability of the proposed algorithm. Since the original scheme for intraframe QP determination in [1] is not suitable for bit rate estimation, it was not included in our comparisons. In our experiments, the first frame of a sequence is coded with an initial QP, and the QP values of the following frames are determined based on the employed R-Q model. In order to measure the accuracy of each model, we define the frame bit rate mismatch ratio as

$$M = \frac{|R_{\rm target} - R_{\rm actual}|}{R_{\rm target}} \cdot 100\% \tag{6}$$

where $R_{\rm target}$ and $R_{\rm actual}$ are the bit rates for the first frame and one following frame, respectively. Table I tabulates the average value of M for different sequences with various initial QP values ($QP_0 = 26$, 32, and 38). From this table, we can see that the rate estimation accuracy has been greatly improved by the proposed R-Q model, especially for the combined sequence. On average, the mismatch of our model is about 3.5% and we have achieved up to 75% improvement over the traditional Cauchy-density-based method. Fig. 4 shows the frame-by-frame bit rate estimation result of different methods for *Comb1* with $QP_0 = 32$. As it can be seen, our model is more adaptive to frame content changes. This is due to the fact that our scheme can efficiently adjust the model parameter based on the frame complexity measure. Compared to Cauchy-based model and the one without rate control, no abrupt bit rate fluctuation is observed in our scheme. Therefore, our R-Q model is more robust and efficient for H.264/AVC intraframe rate control.

V. Conclusion

In this letter, a novel intraframe R-Q model for H.264/AVC video coding has been presented. The motivation is to adaptively enhance the QP prediction accuracy of conventional R-Q model

TABLE I
COMPARISONS OF AVERAGE INTRARATE MISMATCH RATIO (%) WITH DIFFERENT INITIAL QP VALUES

Sequence	$QP_0=26$			$QP_0=32$			$QP_0=38$		
	Cauchy	Proposed	Reduction (%)	Cauchy	Proposed	Reduction (%)	Cauchy	Proposed	Reduction (%)
Foreman	2.69	2.24	16.7	3.48	2.98	14.4	3.31	2.95	10.9
Stefan	2.15	1.92	10.7	2.22	1.99	10.4	1.65	1.51	8.5
Trevor	4.73	3.66	22.6	4.38	3.55	18.9	3.63	3.09	14.9
Carphone	2.31	2.14	7.4	2.60	2.43	6.5	3.61	3.45	4.4
Suzie	3.14	2.61	16.9	3.44	2.96	14.0	3.69	3.30	10.6
News	2.29	1.90	17.0	2.40	2.02	15.8	1.91	1.70	11.0
Football	10.81	3.69	65.9	10.77	4.80	55.4	11.37	5.29	53.5
Table Tennis	7.73	3.68	52.4	6.27	4.38	30.1	5.32	4.49	15.6
Coastguard	4.05	3.04	24.9	3.80	2.87	24.5	5.83	5.05	13.4
Comb1	25.41	6.18	75.7	27.49	7.09	74.2	28.69	9.62	66.5

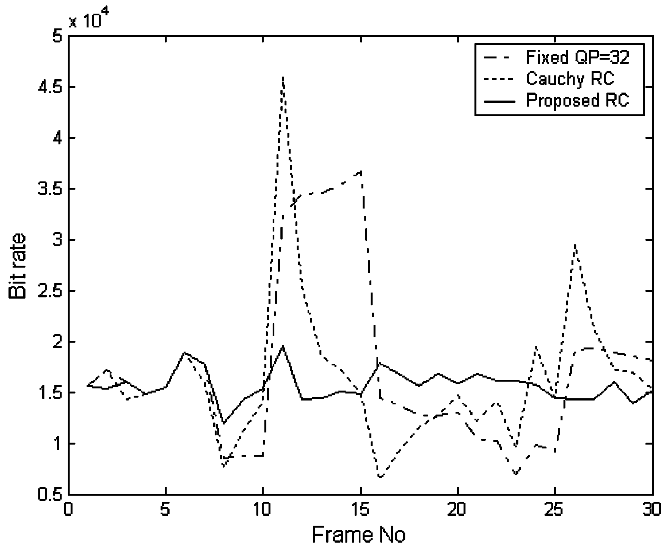

Fig. 4. Intra-frame bit rate of different rate control schemes for *Comb1* sequence with $QP_0 = 32$.

by taking gradient-based frame complexity into consideration. Experimental results show that the proposed algorithm achieves significant improvement in intraframe bit rate estimation. It is more adaptive to frame content changes. In addition, no pre-encoding is required when updating the R-Q parameters. Thus, it is suitable for H.264/AVC video rate control.

REFERENCES

[1] Z. G. Li, F. Pan, K. P. Lim, G. N. Feng, X. Lin, and S. Rahardaj, Adaptive Basic Unit Layer Rate Control for JVT, Joint Video Team, Pattaya, Thailand, 2003, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16/Q.6 Doc. JVT-G012.

[2] Z. He, Y. K. Kim, and S. K. Mitra, "Low delay rate control for DCT video coding via ρ domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 928–940, Aug. 2001.

[3] F. Pan, Z. G. Li, K. P. Lim, X. Lin, S. Rahardja, D. J. Wu, and S. Wu, "Adaptive intra-frame quantization for very low bit rate video coding," in *Proc. Int. Symp. Circuits and Systems, ISCAS'04*, May 2004, pp. 781–784.

[4] N. Kamaci, Y. Altunbasak, and R. M. Mersereau, "Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 8, pp. 994–1006, Aug. 2005.

[5] W. J. Kim, J. W. Yi, and S. D. Kim, "A bit allocation method based on picture activity for still image coding," *IEEE Trans. Image Process.*, vol. 8, no. 7, pp. 974–977, Jul. 1999.

[6] M. Jiang and N. Ling, "Low-delay rate control for real-time H.264/AVC video coding," *IEEE Trans. Multimedia*, vol. 8, no. 3, pp. 467–477, Jun. 2006.

[7] S. Ma, W. Gao, and Y. Lu, "Rate-distortion analysis for H.264/AVC video coding and its application to rate control," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 12, pp. 1533–1544, Dec. 2005.

[8] JM11 Reference Software. [Online]. Available: http://iphome.hhi.de/suehring/tml/download/.