A Study on the Quantization Scheme in H.264/AVC and Its Application to Rate Control*

Siwei Ma[1], Wen Gao[1], Debin Zhao[1], and Yan Lu[2]

[1] Institute of Computing Technology,
Chinese Academy of Science, Beijing, China
{swma,wgao,dbzhao}@jdl.ac.cn
[2] Microsoft Research Asia, Beijing, China
t-yanlu@microsoft.com

Abstract. Compared with previous video coding standards, H.264/AVC employs a division-free quantization scheme. The relation between quantization parameter and quantization step changes from general linear to exponential as well. In this paper, we first analyze the relation between rate and quantization step-size and derive a new rate-distortion (R-D) model. An efficient rate control scheme is then developed based on the new R-D model. The proposed rate control scheme is implemented into the H.264/AVC reference software, with which the better coding performance can be achieved. Experimental results show that the PSNR of the proposed rate control scheme is averagely 0.23dB over the current rate control scheme in H.264/AVC test model and 0.41dB over the coding scheme with fixed quantization parameter, and meanwhile, the complexity of the proposed rate control scheme is much lower than the original one.

1 Introduction

Rate control plays an important part in any standard-compliant video codec. Without rate control, any video coding standard would be practically useless. As a conse-quence, a proper rate control scheme was usually recommended for a standard during the development, e.g. TM5 [1] for MPEG-2, TMN8 [2] for H.263 and VM8 [3] for MPEG-4, etc. H.264/AVC is the newest international video coding standard, and some work about rate control has been done for H.264/AVC too. In [5], a rate control scheme based on VM8 has been proposed to and adopted by H.264/AVC test model. In our previous work [6], rate distortion optimization and hypothetical reference decoder (HRD) have been jointly considered in rate control implementation process, part of which has also been adopted by H.264/AVC test model.

Generally, the rate control can be implemented in two steps. The first step is to allocate appropriate bits for each picture. The bit allocation process is

---

* This paper has been partly supported by NSFC(60333020) and National Hi-Tech Research Program(2002AA119010).

K. Aizawa, Y. Nakamura, and S. Satoh (Eds.): PCM 2004, LNCS 3333, pp. 192–199, 2004.

constrained by a HRD model defined in the standard specification. The second step is to adjust quantization parameter (*QP*) for each coding unit (e.g. the macroblock) so as to fulfill the target bitrate constraint. In other words, the key point is to find the relation between the rate and *QP*. Since the source distortion is closely related with the quantization errors as well as the *QP*, the relation between rate and *QP* is usually developed based on a rate-distortion (R-D) model. For example, in TM5 a simple linear rate-distortion model is introduced. In TMN8 and VM8, the more accurate R-D models are used, which can reduce rate control error and provide better performance but have relatively higher computational complexity. In [4,7], the relation between rate and *QP* is indirectly represented with the relation between rate and , where is the percent of zero coefficients after quantization. In [8] and [9], a modified linear R-D model with an offset indication overhead bits is used for rate control on H.26x.

In conclusion, these rate control schemes are mostly associated with the true relation between rate and quantization parameters in the video codec. Since many coding tools in H.264/AVC, in particular the quantization scheme, differ from the previous video coding standards, it is desirable to derive the new relation between the rate and distortion as well as *QP* for rate control on H.264/AVC. This paper is an extension and refinement of our previous researches. The quantization scheme in H.264/AVC is first fully studied to reveal the true relation between rate and quantization parameter and derive the new R-D model. Based on this new R-D model, a macroblock-layer rate control is proposed and implemented on the H.264/AVC reference software while with lower complexity compared with current rate control scheme [5] in H.264/AVC. Because the complicated MAD prediction and R-D model in [5] makes it difficult to be used in real time encoder.

The rest of the paper is organized as follows. Section 2 presents the detailed analysis on the quantization scheme in H.264/AVC, and then describes the proposed R-D model. In Section 3, the strategy of the proposed rate control algorithm is described. The experimental results are presented in Section 4. And finally, Section 5 concludes this paper.

2 Quantization Scheme in H.264/AVC

Before we present the proposed rate control algorithm, we first make a study on the quantization scheme and the rate-distortion relation in H.264/AVC. Above all we should distinguish *QP* and quantization step-size (referred to as Q_{step} hereafter). *QP* denotes the quantization scale indirectly, whereas Q_{step} is the true value used in quantization. In the previous video coding standards, the relation between *QP* and Q_{step} is usually linear. For example, in H.263 quantization scheme, in terms of the quantization parameter *QP*, a coefficient *COF* is quantized to:

$$LEVEL = \{ \begin{matrix} |COF|/(2 \times QP) \\ |COF - QP/2|/(2 \times QP) \end{matrix} \tag{1}$$

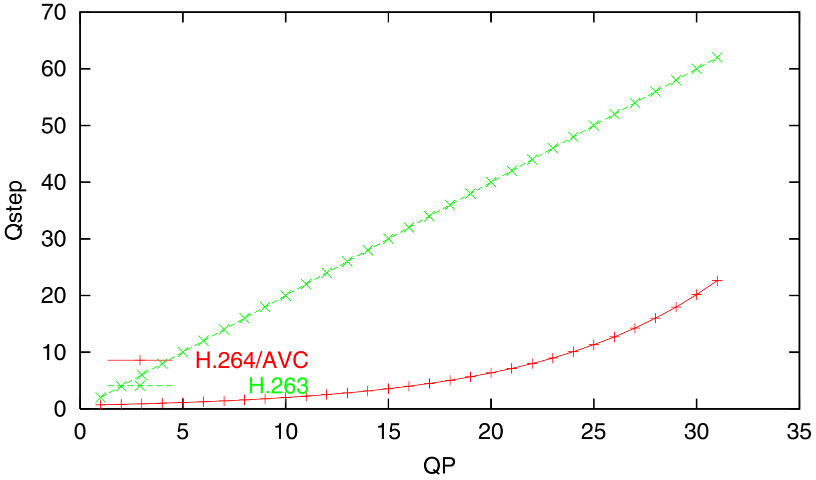

Fig. 1. The relation between QP and Qstep in H.264/AVC

where quantization step-size $Q_{step} = 2QP$. However, in H.264/AVC, the relation between QP and Q_{step} is that $Q_{step} = 2^{(QP/6)}$, as shown in Fig. 1. The underlying reason for this change is that an integer transform and division free quantization scheme is adopted by H.264/AVC. In H.264/AVC, the following integer transform is used to do transformation [10]:

$$Y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \begin{bmatrix} \\ X \\ \\ \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \tag{2}$$

Since the integer transform is not a unitary matrix, Y must be normalized as follows:

$$W = \begin{bmatrix} \\ Y \\ \\ \end{bmatrix} \otimes \begin{bmatrix} \\ E \\ \\ \end{bmatrix} = \begin{bmatrix} \\ Y \\ \\ \end{bmatrix} \begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \tag{3}$$

where a=1/2, b=$1/\sqrt{10}$. In H.264/AVC, in terms of quantization parameter QP, the transform normalization combining with the quantization is implemented for division free as:

$$LEVEL_{i,j} = round(W_{i,j}/Q_{step}) = round(Y_{i,j}S_{i,j}/2^{qbits}) \tag{4}$$

where $S_{i,j}/2^{qbits} = E_{i,j}/Q_{step}$, $Q_{step} = 2^{((QP-4)/6)}$ and $qbits = 15 + floor(QP/6)$.

To model the relation between the rate R and the quantization step-size Q_{step} as well as quantization parameter QP, we make the statistics as follows. Figure 2 shows the relations of $R - (1/QP)$ and $R - (1/Q_{step})$ for the News

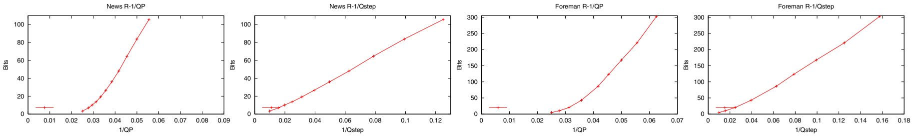

Fig. 2. The relation between R and 1/QP(1/Qstep) on News and Foreman.

sequence, respectively, wherein R denotes the bits for coefficients of luma and chroma components. Table 1 shows the statistics on the Foreman sequence. The correlation factor $\rho_{R-(1/QP)}$ for R and $1/QP$ is 0.991, and the correlation factor $\rho_{R-(1/Qstep)}$ for R and Q_{step} is 0.999. Though R and $1/QP$ is also highly linear-correlated, the sum of squared error (R_i' is the linear aproximation of R_i according to the linear model resolved by least square error multiplier) for the linear approximation of $R-(1/QP)$ is much larger than that of $R-(1/Q_{step})$. Test results on other sequences also show the similar statistical result. More experiments on some other sequences also prove the similar results.

Table 1. Experimental results on News

QP	18	20	22	24	26	28	30	32	34	36	40
R	105.7	83.88	64.65	48.11	36.07	26.53	19.17	13.69	10.05	6.944	3.324
$\rho_{R-1/QP}$	0.991										
$\rho_{R-1/Q_{step}}$	0.999										
$\sum(R-R')^2_{1/QP}$	236.003										
$\sum(R-R')^2_{1/Q_{step}}$	15.1208										

Therefore, we can draw a conclusion that the relation between R and $1/Q_{step}$ can be thought as linear in H.264/AVC. The $R-Q_{step}$ model is then derived as:

$$R_i^t = K^t SAD_i / Q_{stepi} + C^t, t = I, P, B \tag{5}$$

where R_i^t is the estimated number of bits of a macroblock, SAD_j is the sum of absolute difference of a motion compensated macroblock. The first item reflects the bits used to code transform coefficients. The second item is the bits used to code header information of a macroblock. Compared with the linear R-D model, e.g. R-QP model in TM5, the $R-Q_{step}$ model is more accurate for H.264/AVC due to the new quantization scheme.

3 Rate Control Algorithm

In the previous section, a linear $R-Q_{step}$ model has been proposed to reveal the relation of rate and quantization step-size in H.264/AVC. In this section, an efficient rate control scheme for H.264/AVC is presented. Concretely, the proposed rate control algorithm is performed as follows:

Step 1. Bit allocation.
In this step, a target bit is allocated to each picture in a group of picture (GOP) as TM5. For the first I/P/B picture, bit allocation is not performed, and a fixed QP_0^t is then used. After coding the picture, parameters X_0^t, K^t and C^t are initialized, respectively.

$$X_0^t = S_0^t {Q_{step}}_0^t = S_0^t 2^{((QP_0^t - 4)/6))}, C^t = B_{head} \; and$$
$$K^t = B_{coeff} {Q_{step}}_0^t / SAD_t = B_{coeff} 2^{((QP_0^t - 4)/6))} / SAD_t \qquad (6)$$

S_0^t is the coded bits of the frame; SAD_t is the average SAD of all macroblocks in the frame; B_{head} is the average header bits for a macroblock, including motion and mode information; and B_{coeff} is the bits used to code luma and chroma coefficients. Set $j^t = 0$ for the first picture with type t, which is used to update K^t in the R-D model. Set $K_1 = KK_0 = K^t, C_1 = CC_0 = C^t$. N^t is the number of t type pictures in the current GOP.

Step 2. Initialization of the current macroblock.
In this step, we initialize some parameters. Let $i = 1$ for the first macroblock. Assume B_1 is the number of available bits for coding this frame, and L_1 is the number of remained not coded macroblocks in the current frame.

Step 3. Rate distortion optimization mode selection for the current macroblock.
If the current macroblock is the first one in a frame, QP is set to be the average quan-tization parameter QP_{prev} of previous frame; otherwise, if $B_i/L_i < C^t, QP = QP_{prev}$, else a new Q_{step} for the current macroblock is calculated as:

$$Q_{step} = K^t SAD_i - 1/(B_i/L_i - C^t) \qquad (7)$$

Therefore, $QP = round(6 log_2 Q_{step}) + 4$. QP is then clipped with:

$$QP = min(max(QP_{prev} - 3, QP), QP_{prev} + 3) \qquad (8)$$

QP must be clipped to be in the range from 0 to 51. The new QP is used in the coding mode selection of the current macroblock.

Step 4. Counter updating.
In this step, the remaining bits and the number of not coded macroblocks of the frame are updated as follows:

$$B_{i+1} = B_i - R_i, and L_{i+1} = L_i - 1$$

Assume MB_CNT is the total number of macroblocks in the frame. If $i = MB\_CNT$, all macroblocks in the frame are coded; otherwise, let $i = i + 1$, and go to Step 2.

Step 5. R-D model parameter updating.
The R-D mode parameters K and C are updated similar to [2], but at frame level. First calculate:

$$K' = \frac{R_{C,n} Q_{step}}{SAD_n^t MB\_CNT}, C' = \frac{R_n - R_{C,n}}{MB\_CNT} \qquad (9)$$

where $R_{C,n}$ is the number of bits spent for the luminance and chrominance of the nth picture with picture type t. If $(K' >= 0andK' <= threshold)$, set $j^t = j^t + 1$ and:

$$KK_{j^t} = KK_{j^t-1}(j^t - 1)/j^t + K'/j^t, CC_n = CC_{n-1}/n + C'/n \qquad (10)$$

K^t and C^t are updated as a weighted average of the initial estimates and the current average:

$$K^t = KK_{j^t} n/N^t + K_1(N^t - n)/N, C^t = CC_n n/N^t + C_1(N^t - n)/N^t \qquad (11)$$

4 Experiments and Results

In order to evaluate the performance of the proposed algorithm, some experiments have been done on the typical test sequences. Three rate control schemes, i.e. the proposed rate control scheme, the modified TM5 rate control scheme with some parameters adjustment for H.264/AVC and the rate control scheme currently adopted in H.264/AVC test model are implemented on the same H.264/AVC reference software, respectively. The comparisons are then performed among these rate control schemes and fixed QP coding.

The experimental results are listed in Table 2. According to the table, we can see that the proposed algorithm can effectively control the bit-rate at different resolution, frame rate, and meanwhile it can also achieves better coding efficiency than the rate control schemes and fixed QP coding. The maximum improvement compared to the current rate control in H.264/AVC test model [5]

Table 2. Experimental results on test sequence

Sequence	Squence Type	Target Bit-rate (kbps)	Rate Control Scheme	Coded Bit-rate (kbps)	PSNRY (dB)	Improvement over (dB)		
						[5]	TM5	Fixed QP
News	QCIF 10f/s, IPP	10.19	Proposed	10.19	28.78	0.48	0.98	0.78
			[5]	10.23	28.30			
			TM5	10.20	27.80			
			None(Fixed-QP)	10.19	28.00			
Foreman	QCIF 15f/s, IBBP	98.00	Proposed	97.76	36.35	0.06	0.50	0.25
			[5]	97.76	36.29			
			TM5	98.00	35.85			
			None(Fixed-QP)	98.00	36.10			
Container	QCIF 10f/s, IPP	12.84	Proposed	12.92	33.56	0.15	0.71	0.46
			[5]	12.99	33.41			
			TM5	13.15	32.85			
			None(Fixed-QP)	12.84	33.10			
Bus	QCIF 30f/s, IPP	93.84	Proposed	93.87	30.78	0.24	0.53	0.13
			[5]	93.88	30.54			
			TM5	94.36	30.23			
			None(Fixed-QP)	93.84	30.65			

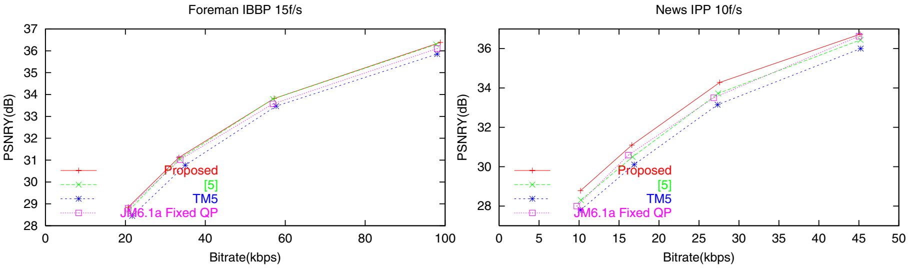

Fig. 3. PSNR curve of News at 10f/s and Foreman at 15fps, QCIF

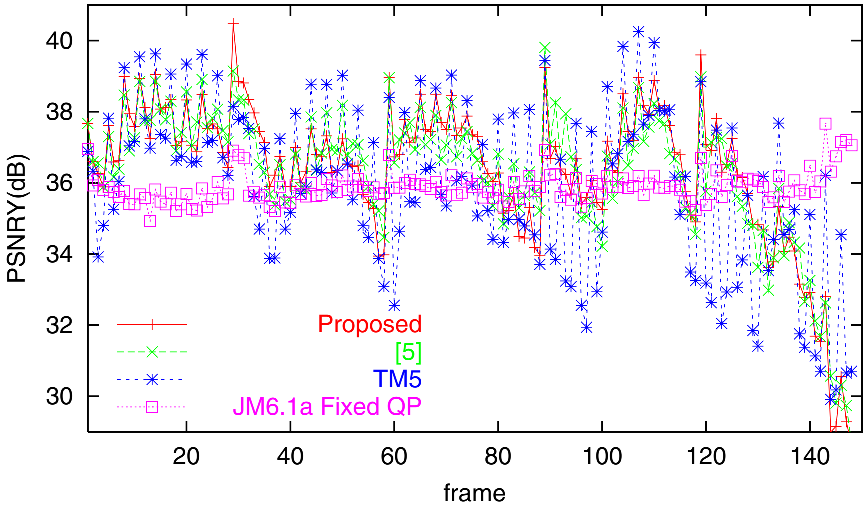

Fig. 4. PSNR per frame for Foreman 98.0kbps, 15f/s, QCIF

is 0.48dB and the average improvement reaches 0.16dB. Since the proposed algorithm only employs a simple linear model without performing the complicated MAD prediction, the complexity is also much lower than the current rate control in H.264/AVC test model. Compared with fixed *QP* coding, the proposed rate control can further improve coding efficiency with the average PSNR improvement reaching at 0.45dB. Compared with the modified TM5 implementation in H.264/AVC, the improvement is up to 0.98dB.

In Fig. 3, the rate-distortion curves of the proposed rate control, TM5, current rate control in H.264/AVC test model and fixed *QP* coding are shown, respectively. From the curves we can see that the proposed rate control can achieve better performance than any other schemes, which demonstrates that the R-D model in the proposed rate control scheme is more accurate for H.264/AVC video coding. Since the bits of coding the motion vectors play a part in the overall rate, the motion in the sequence may also influence the rate control scheme.

For the News sequence with small motion, the proposed rate control can achieve much better performance than the other schemes. The performance is almost the optimum. For the Foreman sequence with high motion, the performance of the proposed scheme is also a little better than the other schemes, which demonstrates that the proposed R-D model is still more accurate than the others even for the sequence with high motion.

Figure 4 shows the PSNR per frame in terms of the Foreman sequence. The figure further indicates that the proposed rate control scheme shows better performance and has improved the coding efficiency of the original H.264/AVC test model.

5 Conclusion

This paper has presented a detailed study on the quantization scheme in H.264/AVC, from which a more accurate rate distortion model has been derived. Based on the proposed rate distortion model, an efficient rate control scheme has been proposed for H.264/AVC video coding. The experimental results have shown that the proposed scheme can further improve the coding efficiency of the original H.264/AVC test model. In addition, the proposed rate control scheme can also outperform the current rate control scheme in H.264/AVC test model, and meanwhile its computational complexity is also lower.

References

1. ISO-IEC/JTC1/SC29/WG11, "Test model 5," JTC1/SC29/WG11 Coding of Moving Pictures and Associated Audio, MPEG 1994
2. J. Corbera and S. Lei, "Rate control for low-delay video communications," ITU Study Group 16, Video Coding Experts Group Documents Q15-A-20, Portland, June 1997
3. T. Chiang, Y. Zhang, "A new rate control scheme using quadratic rate distortion model," IEEE Trans. Circuits Syst. Video Technol. Vol. 7, pp. 287–311, Apr. 1997
4. S. Milani, L. Celetto, G. A. Mian, "A rate control algorithm for the H.264 encoder"
5. ISO IEC TC JTC 1/SC29 N5821, Draft ISO/IEC 14496-5:2002/PDAM6: 2003
6. S. Ma, W. Gao, F. Wu, Y. Lu, "Rate control for AVC video coding scheme with HRD considerations," IEEE International Conference on Image Processing, ICIP 2003, Barcelona, Spain, 14–17 Sept. 2003
7. Z. He, S. K. Mitra, "A unified rate-distortion analysis frame work for transform coding," IEEE Trans. on Circuits Syst. Video Technology. Vol. 11. No. 12, Dec. 2001
8. J. Wang, Z. Chen, Y. He, Y. Chen, "A MAD-based rate control strategy," Document JVT-D070, Klagenfurt, Austria, July 2002
9. C. Wong, O. C. Au, B. Meng, H. Lam, "Novel H.26X optimal rate control for low-delay communicaitons," ICICS-PCM 2003, 15–18 Dec. 2003, Sigapore
10. H.264/MPEG-4 Part 10: Transform & Quantization, http://www.vcodex.com