Efficient Bit Allocation and CTU level Rate Control for High Efficiency Video Coding

Junjun Si, Siwei Ma, Wen Gao
Institute of Digital Media, Peking University
Beijing, China, 100871
{jjsi, swma, wgao}@pku.edu.cn

Abstract—**In this work, a CTU level rate control algorithm is proposed for High Efficiency Video Coding (HEVC). On top of the CTU level rate control algorithm, an efficient bit allocation method considering the HEVC hierarchical coding structure is specifically designed. Instead of directly applying the rate distortion model at the CTU level, it's proposed to derive the CTU level quantization parameter (QP) based on the frame level QP with feedback of the coding status of CTUs. To further improve the coding performance of the CTU level rate control, a QP adjustment strategy is proposed and incorporated into the rate control algorithm. Experimental results show that the proposed CTU level rate control algorithm can achieve better coding performance than the state-of-the-art rate control scheme for HEVC (1.9% for RA, 3.2% for LD). Moreover, the coding quality become much smoother with the proposed algorithm.**

Index Terms—**HEVC, rate control, bit allocation.**

I. Introduction

The ultimate aim of video coding is to provide the best video quality at the receiver side, subject to the available channel bandwidth. Basically, best video quality means that the compressed video must provide high quality of the video content and no sharp quality fluctuation shall be noticed. Given a bit budget, the best coding quality can be achieved by optimal bit allocation and accurate rate control schemes. Specifically, the bit allocation scheme distributes the available bit budget to the overall video data according to a principle that the distortion should be minimized. And the rate control scheme is to meet the bit allocation budget by setting an appropriate quantization parameter (QP) for each coding unit of the video. Therefore, bit allocation and rate control for various video coding standards (H.264/AVC, etc.) has been extensively studied in the literature.

High Efficiency Video Coding (HEVC) is the up-to-date video coding standard, which achieves a compression ratio of 200 times. It's developed by the Joint Collaborative Team of Video Coding (JCT-VC), based on conventional block-based motion-compensated hybrid video coding framework. However, many new coding tools are adopted in HEVC, like the extended coding unit scheme, Sample Adaptive Offset (SAO), merge mode, etc. When the new coding tools are well used together, HEVC can achieve approximately 50% bit rate reduction with equivalent perceptual quality relative to the prior coding standards [1].

For HEVC, two rate control proposals have been adopted, which are JCTVC-H0213 [2] and JCTVC-K0103 [3]. JCTVC-H0213 proposes a rate control scheme similar to the H.264/AVC rate control but has significant performance loss. Consequently, JCTVC-K0103 is then adopted into HEVC. It proposes a hierarchical bit allocation scheme based on the hierarchical coding structure of HEVC, then employs a R-λ (R: rate, λ: Lagrange multiplier) model based rate control scheme. Compared to JCTVC-H0213 rate control, it can improve the coding performance as well as the rate control accuracy. However, it has been noticed that the coding quality fluctuates greatly, especially in the Low Delay (LD) configuration, which may not be appropriate for practical applications of video coding.

In view of this, a Coding Tree Unit (CTU) level rate control algorithm along with an efficient bit allocation scheme is proposed in this paper to generate higher and smoother coding quality for HEVC encoder. As for the bit allocation scheme, it takes into account the hierarchical structure of HEVC. The bit allocation is performed at GOP level, frame level and CTU level, respectively. For the CTU level rate control, QPs for CTUs are derived based on the frame level QP with feedback of coding status of CTUs. It's widely known that with limited budget bits, the posterior coding units in a frame usually get very little budget, therefore, the quantization parameter for those coding units are usually very large. As a result, the coding quality of those coding units is poor. The poor quality can propagate to the subsequent frames, which may create significant coding performance loss. Constructively, we designed a simple method to overcome this problem by adjusting the QP of these coding units according to their complexity. Experimental results show that the proposed CTU level rate control algorithm can achieve 1.9% and 3.2% coding gain against JCTVC-K0103 rate control under RA and LD configuration, respectively. The control accuracy is with 1%. Besides, the coding quality is smoother.

The rest of the paper is organized as follows. Section II analyzes the hierarchical structure of HEVC and details the proposed bit allocation scheme. Section III presents the CTU level rate control scheme. Experimental results are demonstrated in Section IV. Finally, Section V concludes this paper.

II. PROPOSED BIT ALLOCATION SCHEME

A. Brief Review of Bit Allocation

First of all, we briefly review some typical bit allocation schemes developed in the literature. In transform coding, bit allocation is performed by allocating the bits budget to different groups of transformed coefficients to minimize the overall quantization distortion. In [4], Huang first presents the optimal bit allocation problem. Optimal bit allocation is incorporated with rate control scheme in video coding to improve the coding quality. It is based on the rate distortion functions of the encoder. The analytic formulas of the rate and distortion functions are employed to derive the closed-form expression of the optimal bit allocation scheme. However, the analytic models of rate and distortion usually exists great estimation error [5]. Therefore, it cannot be truly optimal. For this reason, operational rate distortion curves are proposed to be applied in practical image/video coding. But a critical problem is the computational complexity to generate the operational rate distortion curves is relatively high. Consequently, it cannot work well for practical video applications, especially real-time video coding. Therefore, we propose a simple but efficient bit allocation scheme for HEVC. It operates at three different granularities, which are GOP, frame and CTU levels.

B. GOP Level Bit Allocation

In our previous work [6, 7], a frame level rate control scheme was proposed for HEVC. Average bit allocation is performed for each frame. That's because in practical video applications, it's always expected that the fluctuation of coding quality should be small. Though the average bit allocation scheme is simple, it didn't take the hierarchical structure characteristic into account. In this work, we first perform average bit allocation at GOP level with the assumption that the bits per GOP consumed shall be equally balanced.

$$T_{GOP} = (R/F)\cdot N \tag{1}$$

where R is the target bitrate, F is the frame rate of the sequence and N is the number of frames in a GOP.

C. Frame Level Bit Allocation

In HEVC, hierarchical coding structure is employed. Frames are divided into different layers according to their Picture Order Count (POC) and the size of *gop*. It should be noticed that "*gop*" is different from "GOP". It is a small group of pictures which has certain reference relationship. Its size is set to 4 in the low delay configuration and 8 in the random access configuration. And a syntax Reference Picture Set (RPS) is used in HEVC to specify the reference relationship. Take the low delay configuration as an example, the reference relationship in a *"gop"* is shown in Table 1. It selects one nearest frame and three frames from depth 0 as references.

As can be seen from Table 1, frames in depth 0 are critical in the reference relationship. Therefore, it's reasonable to allocate more bits to these frames so that they can be coded as high quality. I frame's residual is much bigger than inter frames, and its coding quality affects the following inter frames greatly.

Table 1: The reference picture set in low delay configuration,

POC	Depth	POC of Reference pictures			
4n+1	2	-1	-5	-9	-13
4n+2	1	-1	-2	-6	-10
4n+3	2	-1	-3	-7	-11
4n+4	0	-1	-4	-8	-12

Consequently, the proposed frame level bit allocation scheme firstly distribute bits to I frames as:

$$T_I = T_{GOP}\cdot\alpha \tag{2}$$

where α is a constant, set to 0.15 in the experiment. Secondly, for frames of depth 0, their target bit budget is calculated as:

$$T_0 = T_{GOP}\cdot(1-\alpha)\cdot\beta/N_0 \tag{3}$$

where β is a constant set to 0.3 (RA) and 0.45 (LD) in the experiment and N_0 is the number of frames of depth 0 in the GOP. Finally, for the remaining frames in the GOP, the target bit budget for each frame is set as:

$$T_r = (T_{GOP} - T_I - T_0\cdot N_0)/N_r \tag{4}$$

where N_r is the number of remaining frames in the current GOP, and $N = 1+N_0+N_r$.

D. CTU Level Bit Allocation

Assume that the target bit budget for the current frame is T, excluding the bits used to encode the slice head etc. Then, the bits allocated for each CTU is:

$$\begin{aligned} T' &= T - head \\ T_{ctu} &= T'/(N_{ctu} - ctu\_done) \end{aligned} \tag{5}$$

head is the head bits of the frame including slice header etc. N_{ctu} is the total number of CTUs in the current frame. *ctu_done* is the number of CTUs coded in the frame. *T'* is updated with the actual coding bits *m* of the CTU:

$$T' = T' - m \tag{6}$$

E. Effectiveness of the Proposed Bit Allocation Method

To verify the effectiveness of the proposed bit allocation method, experiments are implemented on HM10.0. The coding

Table 2: Coding gain of the proposed bit allocation method

Sequence	BD-RATE		
	Y	U	V
BasketballPass	-0.2%	2.7%	0.9%
BQSquare	-3.3%	-5.5%	-6.7%
BlowingBubbles	-0.8%	0.5%	-0.6%
RaceHorses	-1.2%	0.1%	1.0%
Average	**-1.4%**	**-0.5%**	**-1.4%**

gain against average bit allocation in [6] is shown in Table 2. From the table, we can obtain the conclusion that the proposed bit allocation scheme is efficient for HEVC encoder.

III. Proposed CTU Level Rate Control Scheme

For CTU level rate control, the traditional way is implementing the rate distortion model at CTU level. It's favorable to achieve accurate rate control, but the computational complexity of updating model parameters is relatively high and the model sometimes fails because of a shortage of bit budget for the posterior CTUs of a frame. Therefore, a novel CTU level rate control scheme is proposed in this paper.

In our previous work [6, 7], we proposed a frame level rate control scheme. In this paper, the frame level rate control scheme is incorporated with the proposed bit allocation method detailed in Section II. Then, the improved frame level rate control is employed to generate a frame level QP. At CTU level, the quantization parameter for each CTU is decided with a reaction factor based on the frame level QP. Besides, a CTU level QP adjustment strategy is designed. During the pre-analysis process, the SATD value for each CTU in the frame is calculated and stored for CTU level QP adjustment. The following parts of this section give the details of the proposed CTU level rate control.

A. CTU Level Rate Control

Instead of applying the rate distortion model at CTU level, we propose to use a feedback strategy to generate the CTU level quantization parameter based on the coding status of CTUs. First of all, a reaction factor γ is defined as:

$$\gamma = \sum_i T_i / s \tag{7}$$

T_i is the target bit budget for the i-th frame, s is a variable calculated as:

$$s = \sum_i (R_i \cdot QP_i) / \sum_i (w_i \cdot SATD_i) \tag{8}$$

where R_i is the actual coding bits of the i-th frame, QP_i is the average QP value, w_i is a weighting factor defined in [7].

At frame level, the reaction factor γ is calculated with the target bit budget for the frame and actual coding status of previously encoded frames. On the other hand, at CTU level, it can be updated with the CTU level bit allocation and actual coding status of CTUs. With the frame level and CTU level reaction factor, the quantization parameter for the i-th CTU can be derived as:

$$q_{ctu\_i} = q_{frame} \cdot \gamma_{frame} / \gamma_{ctu\_i} \tag{9}$$

q is the quantization scale and can be mapped to QP:

$$QP = 12 + 6 \times \log_2(q / 0.85) \tag{10}$$

For the first CTU in each frame, its quantization parameter is set to the frame level QP value directly. Then the CTU level reaction factor can be derived.

B. QP Adjustment Strategy

In CTU level rate control, it's very likely that few bits can be remained for CTUs in the posterior part of the frame, especially under the low bitrate conditions. As a result, QP for these CTUs increase dramatically to a very large value and the CTUs are consequently encoded in poor quality. Nonetheless, the distortion will propagate to the following frames through reference and decrease the overall coding quality. To solve this problem, we propose a CTU level QP adjustment strategy.

If the target bit budget T' for the uncoded CTUs is less than a certain value t, then the following CTUs are divided into two different categories according to their complexity. In the experiment, t is set to 0 and Sum of Absolute Transformed Difference (SATD) is used to specify the complexity. For CTUs which has larger SATD value, the QP value should be smaller to enable enough details to be remained. For CTUs which has smaller SATD value, QP is allowed to be larger. Specifically, the QP adjustment strategy can be expressed as:

$$QP = \begin{cases} QP_{frame}, & satd > avg\_satd \\ QP_{frame} + 2, & satd \le avg\_satd \end{cases} \tag{11}$$

where avg_satd is the average SATD value of a CTU in the current frame.

$$avg\_satd = \sum_{all\ ctu} satd / N_{ctu} \tag{12}$$

Table 3 shows the coding gain of the proposed QP adjustment strategy over some test sequences. The coding gain can be much bigger under low bitrate or low resolution sequences, therefore, it's beneficial for practical applications such as video chat.

C. QP clipping

Finally, the QP value for each CTU should be clipped into an expected range. Since the QP values for neighboring CTUs should not fluctuate sharply. Besides, the CTU level QP shall be within a certain range around the frame level QP value. Based on the analysis, the CTU level QP is clipped as:

$$\begin{aligned} QP_{ctu\_i-1} - 3 \le QP_{ctu\_i} \le QP_{ctu\_i} + 3 \\ QP_{frame} - 3 \le QP_{ctu\_i} \le QP_{frame} + 3 \end{aligned} \tag{13}$$

In the end, the QP shall be clipped into the allowed QP value (0~51).

Table 3: The coding gain of the proposed CTU level QP adjustment strategy

Sequence	BD-RATE		
	Y	U	V
BasketballPass	-0.2%	0.6%	0.8%
BQSquare	-1.0%	1.4%	0.2%
BlowingBubbles	-0.6%	-1.7%	-0.9%
RaceHorses	-0.7%	1.3%	0.7%
Average	**-0.6%**	**0.4%**	**0.2%**

IV. Experimental Results

To verify the effectiveness and coding performance of the proposed CTU level rate control scheme, experiments are designed and implemented on HEVC reference software HM10.0. The rate distortion performance of the proposed rate control scheme is compared with the state-of-the-art rate control algorithm in JCTVC-K0103. The target bitrates are set as the default hierarchical QP coding results. The test conditions of JCTVC-K0103 rate control are shown in Table 4.

Table 4: Test condition of JCTVC-K0103 rate control

RateControl	1
KeepHierarchicalBit	1
LCULevelRateControl	1
RCLCUSeparateModel	1
InitialQP	0
RCForceIntraQP	0

The detailed experimental results of the proposed CTU level rate control algorithm are demonstrated as follows:

A. *Rate Distortion Performance*

In this experiment, we employ the JCTVC-K0103 rate control as anchor, and the rate distortion performance of the proposed CTU level rate control scheme in terms of BD-RATE is show in Table 5. It can be seen that the average coding gain is 1.9% and 3.2% over test sequences of various resolution and frame rate for RA and LD configuration, respectively. The average bitrate error is 0.96% for RA configuration and 0.78% for LD configuration.

Table 5: Rate distortion performance against JCTVC-K0103 rate control algorithm

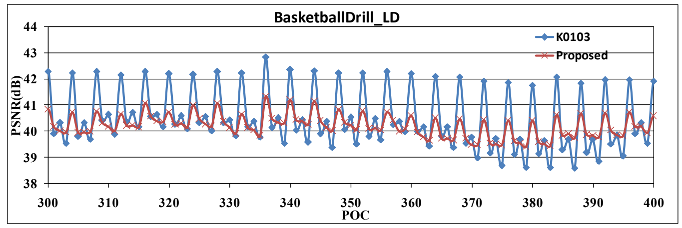
Sequence	**Resolution**	**Frame rate**	**BD-RATE**	
			RA	**LD**
PeopleOnStreet	2560x1600	30	-4.1%	
Kimono	1920x1080	24	0.7%	-2.2%
BasketballDrive	1920x1080	50	-2.2%	-5.9%
BQMall	832x480	60	-1.0%	0.3%
RaceHorsesC	832x480	30	-3.9%	-5.0%
BlowingBubbles	416x240	50	-2.5%	-1.1%
BasketballPass	416x240	50	-0.1%	-0.4%
KristenAndSara	1280x720	60		-7.9%
Average			-1.9%	-3.2%

B. *Smoothness of the Coding Quality*

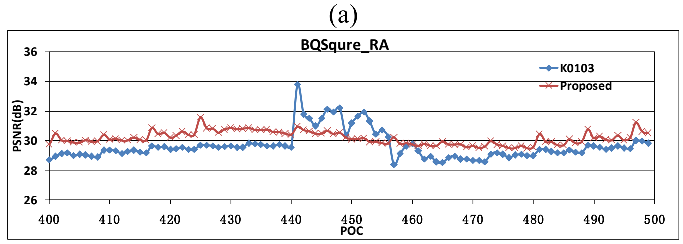
Another advantage of the proposed rate control scheme is it helps to smooth the coding quality and reduce the quality fluctuation. In Fig. 1, some PSNR curves of the proposed rate control algorithm and JCTVC-K0103 rate control are presented. For comparison convenience, only part of the PSNR curve is demonstrated. It can be seen that the proposed rate control can efficiently reduce the coding quality fluctuation.

V. Conclusion

This paper proposes an accurate CTU level rate control scheme incorporated with an efficient bit allocation method for HEVC. The proposed bit allocation method takes the hierarchical coding structure of HEVC into consideration. Experimental results show that the proposed bit allocation scheme is efficient for HEVC encoder. For CTU level rate control, QP determination at CTU level is based on the frame level QP with feedback of the coding status of CTUs. Besides, a CTU level QP adjustment strategy is proposed to improve the coding performance. It's verified that the proposed CTU level rate control scheme can achieve better coding performance than the state-of-the-art rate control for HEVC. The bitrate error is within 1% and the coding quality is smoother.



(a)



(b)

Figure 1: Comparison of PSNR curves of the proposed rate control algorithm and JCTVC-K0103 rate control. (a) LD configuration; (b) RA configuration.

VI. Acknowledgement

This work was supported in part by National High-tech R&D Program of China (863 Program, 2012AA010805) and National Science Foundation (61210005 and 61103088).

References

[1] G. J. Sullivan, J.-R. Ohm, "Recent developments in standardization of high efficiency video coding (HEVC)," SPIE Applications of Digital Image Processing XXXIII, Proc. SPIE, Vol. 7798, Aug. 2010.

[2] H. Choi, J. Nam, J. Yoo, D. Sim and I.V. Bajić, "Rate control based on unified RQ model for HEVC," ITU-T SG16 Contribution, JCTVC-H0213, San José, Feb. 2012.

[3] B. Li, H. Li, , L. Li, and J. Zhang, "Rate control by R-lambda model for HEVC," ITU-T SG16 Contribution, JCTVC-K0103, Shanghai, Oct.2012

[4] J. J. Huang and P. M. Schultheiss, "Block quantization of correlated Gaussian random variables," IEEE Trans. Commun. Syst., vol. CS-11, pp. 289–296, Sept. 1963.

[5] Z. He and S. K. Mitra, "Optimum bit allocation and accurate rate control for video coding via rho-domain source modeling", IEEE Trans. Circuits Syst. Video Technol., vol. 12, no. 10, pp.840 -849, 2002.

[6] J. Si, S. Ma, X. Zhang, W. Gao, "Adaptive Rate control for High Efficiency video Coding," in Proceedings of SPIE Conference on Visual Communications and Image Processing, VCIP, San Diego, USA, Nov.2012

[7] J. Si, S. Ma, W. Gao, And M. Yang. "Adaptive Rate Control For HEVC," ITU-T Sg16 Contribution, Jctvc-I0433, Geneva, Ch, Apr.-May 2012