Low Delay Rate Control for HEVC

Zhongzhu Yang, Li Song, Zhengyi Luo and Xiangwen Wang

Abstract—**This paper presents a rate control scheme for low delay video communication of the High Efficiency Video Coding (HEVC) standard. To prevent the buffer overflow and underflow under small buffer size constraint in low delay communication, the state-of-the-art *R-λ* algorithm is improved for more accurate bit allocation. A new bit allocation method based on the buffer status is proposed to control the buffer better, which is very important in low delay applications. Experimental results show that compared with the original *R-λ* algorithm, the proposed scheme has a smaller bit fluctuation and can avoid buffer overflow and underflow with a smaller buffer size, which often means a lower delay, and the PSNR degradation is negligible.**

Index Terms—**Video coding, Rate Control, HEVC, R-λ model, low delay.**

I. INTRODUCTION

HIGH Efficiency Video Coding (HEVC) is the latest video coding standard developed by JCT-VC (Joint Collaborative Team on Video Coding) [1], which significantly improves the coding efficiency than the previous video coding standards.

Due to the limited storage capacity or communication bandwidth, quantization is introduced to reduce the bit rate of the compressed video signal such that the capacity or bandwidth limitation can be met properly. To achieve a target bitrate, rate control, which dynamically adjusts the quantization parameters (QPs), is usually adopted in the video coding framework. Many rate control algorithms and rate quantization (*R-Q*) models have been developed for previous video coding standards. As far as HEVC is concerned, the rate control problems have been discussed in [2] [3] [4].

It's flexible for designers to develop suitable schemes for specific applications of HEVC. Choi et al. [2] proposed a pixel-wise *R-Q* model with a quadratic form and a temporal linear model for predicting the mean absolute difference (MAD) value. Li et al. [3] developed a rate control algorithm based on the *R-λ* model for HEVC temporal scalability, considering that λ is the key factor to determine bitrate. Wang et al. [4] modified the ρ-domain rate control according to the new feature of quadtree coding structure in HEVC. It should be noted that all the above algorithms mainly take into account the coding efficiency and video quality, but are not designed for low delay applications of HEVC.

For low delay video communication systems such as video conferencing, the buffer size has to be limited to meet a low delay constraint. Besides, to avoid buffer overflow and underflow, more accurate bit allocation and encoding parameter adjustment is desired in low delay applications. The state-of-the-art *R-λ* algorithm [3] in HM10.0 cannot be used for low delay cases directly because the number of frames to be coded is not available on the fly. There are two problems with the rate control algorithm in HM10.0.

- *Buffer overflow and underflow*: If the number of frames to be coded is set to a very large number, the buffer occupancy will increase or decrease at a constant speed, so the buffer may overflow or underflow. And if it is used in the low delay applications, the buffer size should be set a large size, which is not suitable for low delay requirement.
- *Inaccurate CTU bit estimation*: In HM10.0, the bit allocation for CTU is according to the MAD ratio of the same position of the previous frame. But inaccurate MAD estimation which causes an inaccurate CTU bit estimation will bring about an unstable buffer occupancy. So the simple CTU bit allocation will impair the final rate control result, especially in the low delay applications.

In this paper, a low delay rate control scheme for HEVC is proposed. We improve the algorithm in HM10.0 for more accurate bit allocation. A new bit allocation method based on the buffer status is proposed to control the buffer better, and a weight ratio using a numerical method is adopted to get over the inaccurate MAD estimation. Experimental results show that the algorithm has a lower buffer occupancy and a smaller bit fluctuation, which enables a lower delay and adapts better to low delay applications of HEVC.

The remainder of this paper is organized as follows. In Section II, we briefly describe three rate control models that have been used for HEVC. In Section III, the rate control scheme for low delay video communication of HEVC is proposed. The experimental results and discussions are presented in Section IV. The conclusion is drawn in Section V.

This work was supported by National 863 project (2012AA011703), National Key Technology R&D Program of China (2013BAH53F04), NSFC (61221001, 61271221), the 111 Project (B07022) and the Shanghai Key Laboratory of Digital Media Processing and Transmissions.

Z. Yang is with the Department of Electronic Engineering, Shanghai Jiao Tong University, 200240, China (e-mail: yangzhongzhu@sjtu.edu.cn).

L. Song is with the Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, 200240, China (86 21 34204468; e-mail: song_li@sjtu.edu.cn).

Z. Luo and X. Wang are with the Shanghai University of Electric Power, 200240, China (e-mail: lzy@shiep.edu.cn, wxw21st@gmail.com).

II. STATE-OF-THE-ART R-Q MODELS

Generally rate control can be divided into two steps and the first is bit allocation. The hierarchical level, complexity, the buffer status and sensitivity of each frame are usually considered during bit allocation. The bit allocation of HEVC can be conducted at three levels—GOP level, frame level and CTU level. The other step of rate control is to produce the bits close to the allocated bits, which is realized by building the models between the target bitrate (*R*) and quantization parameter (*Q*) directly or indirectly. Three *R-Q* models that have been used for rate control are introduced as follows.

A. Quadratic R-Q model

One *R-Q* model builds the relationship between bitrate R and quantization parameter *Q* in a quadratic form [5]

$$R = a \times Q^{-1} + b \times Q^{-2} \tag{1}$$

where *a* and *b* are model parameters related to the video content. Choi et al. applied the above model to rate control of HEVC in [2], where the relationship between *R* and *Q* of HEVC is improved as

$$bpp_i(j) = a \times \frac{MAD_{pred,i}(j)}{QP_i(j)} + b \times \frac{MAD_{pred,i}(j)}{QP_i^2(j)} \tag{2}$$

where $bpp_i(j)$ is the target bits per pixel of the *j*-th picture in the *i*-th GOP, $MAD_{pred,i}(j)$ denotes the predicted MAD, *a* and *b* are the model parameters. And $MAD_{pred,i}(j)$ is predicted from the MAD of the previous frame using the linear model

$$MAD_{pred,i}(j) = \alpha_1 \times MAD_{actual,i}(j-1) + \alpha_2 \tag{3}$$

where $MAD_{actual,i}(j\text{-}1)$ is the actual MAD of the previous frame, α_1 and α_2 are two parameters in the prediction model.

B. R-ρ model

Another model examines the relationship between *R* and the percentage of zeros among quantized transform coefficients-*ρ* [6], which is called *R-ρ* model and can be expressed in a linear form as

$$R = \theta(1-\rho) \tag{4}$$

where *θ* is a model parameter directly related to the image content. Considering the transform coefficients follow the laplacian distribution, Wang et al. [4] proposed a one-to-one correspondence between *ρ* and *Q* in rate control of HEVC.

C. R-λ model

The latest model builds the relationship between target bitrate *R* and lagrangian multiplier *λ* for HEVC [3], which is called *R- λ* model and is expressed as

$$R = \alpha \times \lambda^{\beta} \tag{5}$$

where *α* and *β* are the model parameters related to the video source. When *λ* is determined by the target bit, all the coding parameters including QP can be selected by the RDO process. But considering the encoding complexity, QP is determined by the following equation for simplification [7]

$$QP = a \cdot \ln(\lambda) + b \tag{6}$$

where *a* and *b* are constant parameters, and are set to 4.2005 and 13.7122 empirically.

Rate control algorithms based on different models have different features. In the quadratic model based rate control algorithms the MAD has to be predicted accurately, but the linear prediction model in (3) is not accurate in high motion scenes. Because the transform coefficients are necessary to calculate the QP beforehand, the *R-ρ* model based rate control algorithms are usually used in two-pass rate control and are not suitable for low delay cases. The state-of-the-art *R-λ* model has a high accuracy in producing target bits, which is crucial in low delay applications. And it has been used in HM10.0, so in this paper, the *R-λ* model is adopted to compute the QP and *λ*.

III. LOW DELAY RATE CONTROL FOR HEVC

Our proposed rate control algorithm is composed of three levels—GOP level rate control, frame level rate control and CTU level rate control. In the GOP level, the total number of remaining bits for the left non-coded frames should be computed. The frame level rate control allocates a number of bits for the current frame to be coded, where our buffer status-based bit allocation scheme is applied. CTU level rate control gets the target bit number according to a new CTU allocation scheme. After the target bits are computed, the *λ* and QP are calculated by (5), (6). Finally, the model parameters in (5) are updated after the frame or the CTU is coded.

A. Buffer status based bit allocation

To solve the problem of overflow and underflow of the buffer, we propose the buffer status-based bit allocation for a frame. At first, the GOP-level controller manages the target bit number of a GOP. The total bit budget $B_i(j)$ for a GOP is calculated by

$$B_i(j) = \begin{cases} \left(\frac{R}{f} - \theta \times \frac{V_i(j)}{f}\right) \times N_{GOP} & if\ j = 1 \\ B_i(j-1) - \mathrm{b}_i(j-1) & if\ j = 2,3,\ldots,N_{GOP} \end{cases} \tag{7}$$

where *f* is the frame rate, *R* is the target bit rate and N_{GOP} is the number of frames in a GOP, $V_i(j)$ is the buffer occupancy and *θ* is a parameter controlling the convergence speed.

After finishing encoding a frame in current GOP, the budget and the buffer occupancy is updated using the actual generated bit number $b_i(j\text{-}1)$.

$$V_i(j)=\begin{cases} V_{i-1}(1) & if\ j=1 \\ V_i(j-1)+\mathrm{b}_i(j-1)-\dfrac{R}{f} & if\ j=2,3,\ldots,N_{GOP} \end{cases} \tag{8}$$

For the frame-level rate control, the target bit number of current frame depends on two target bit budgets: one is the target bit budget based on the buffer occupancy and target buffer level, and the other is the target bit budget based on the total bit budget in a GOP. In the calculation of the target bit budget $\widetilde{T}_i(j)$ based on buffer status, a feedback from the buffer Δ_T is proposed, which is defined as

$$\Delta_T=\gamma\times(L-V_i(\mathrm{j})) \tag{9}$$

where L is the target buffer level, which has a constant value of $0.5B_s$ with B_s being the buffer size. And γ is a constant, which is set to 0.5 for low delay cases. So the target bit budget based on buffer status is calculated by

$$\widetilde{T}_i(j)=\frac{R}{f}+\Delta_T \tag{10}$$

The other target bit budget $\widehat{T}_i(j)$ based on the left bit budget in current GOP is calculated by

$$\widehat{T}_i(j)=\frac{B_i(j)}{N_{left}} \tag{11}$$

where N_{left} represents the number of not-yet-coded frames in current GOP. After the two part bit budget is calculated, the final target bit budget for a frame is a weighted average of $\widetilde{T}_i(j)$ and $\widehat{T}_i(j)$

$$T_i(j)=\beta\times\widehat{T}_i(j)+(1-\beta)\times\widetilde{T}_i(j) \tag{12}$$

where β is a weighting factor, and it is set to 0.9 for low delay case.

B. Accurate CTU-level bit allocation

To get over the inaccurate CTU-level bit estimation due to the poor MAD prediction in high motion scenes, a numerical method is adopted to calculate the weight for CTU-level bit allocation. After the λ of the frame level rate control is calculated, the weight w_m of the m-th CTU is computed by

$$w_m=\frac{\alpha_m\times\lambda^{\beta_m}}{\sum\limits_{k=0}^{N_{CTU}-1}\alpha_k\times\lambda^{\beta_k}},m=0,1,2,\ldots,N_{CTU}-1 \tag{13}$$

where α_m and β_m are the model parameters for the m-th CTU, and N_{CTU} is the number of CTU in a frame.

At the CTU level, the total bit budget for the non-coded CTUs is calculated by

$$B_i(j,m)=\begin{cases} T_i(j) & m=0 \\ B_i(j,m-1)-b_i(j,m-1) & otherwise \end{cases} \tag{14}$$

where $b_i(j,\mathrm{m}-1)$ is the actual generated bits of (m-1)-th CTU in the j-th frame of the i-th GOP. We first calculate the target bit number $\widehat{T}_i(j,m)$ according to the left bit budget $B_i(j,m)$ in the frame and the weight ratio of the CTU calculated in (13) by

$$\widehat{T}_i(j,m)=B_i(j,m)\times\frac{w_m}{\sum\limits_{k=m}^{N_{CTU}-1}w_k},\ \mathrm{m}=0,1,2,\ldots,N_{CTU}-1 \tag{15}$$

where N_{CTU} is the number of CTU in a frame. Then the frame level buffer occupancy $V_i(j,m)$ is defined as

$$V_i(j,m)=\begin{cases} 0 & m=0 \\ V_i(j,m-1)+b_i(j,m-1)-\dfrac{T_i(j)}{N_{CTU}} & otherwise \end{cases} \tag{16}$$

And the target bit number $\widetilde{T}_i(j,m)$ is calculated by

$$\widetilde{T}_i(j,m)=B_i(j)\times\frac{w_m}{\sum\limits_{k=1}^{N_{CTU}}w_k}-\frac{V_i(j,m)}{N_{CTU\_left}},\ m=0,1,2\ldots,N_{CTU}-1 \tag{17}$$

where N_{CTU_left} denotes the number of remaining CTUs in the frame. Now the target bit number for the m-th CTU is the weighted average of the two part target bit budgets

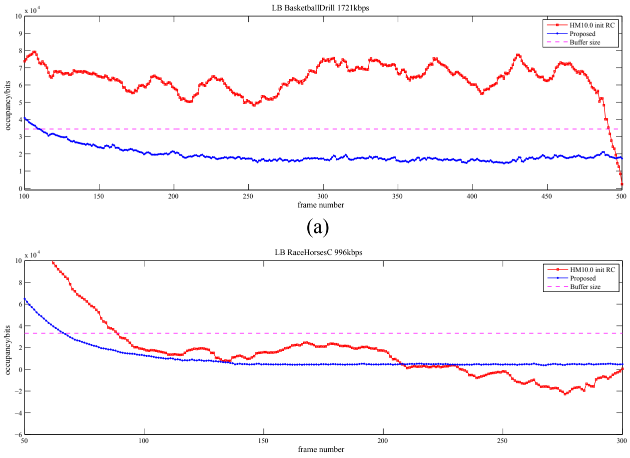
$$T_i(j,m)=\beta\times\widehat{T}_i(j,m)+(1-\beta)\times\widetilde{T}_i(j,m) \tag{18}$$

where β is the same weighting factor as in (12).

After the target bit number is estimated, the R-λ model and the λ-QP model in (5) and (6) are used to get the λ and QP for encoding the m-th CTU. And after that, the parameters a and b are updated.

IV. EXPERIMENTAL RESULTS

Experiments are conducted to test the performance of the proposed rate control algorithm. We aim to prove that the proposed rate control algorithm has better buffer status, which can better adapt to low delay applications. And the bitrate error and PSNR performance are compared with the HM10.0 anchor rate control algorithm. The proposed algorithm is implemented in HM10.0 with low delay B coding structure. The output bitrate of the HM10.0 encoder with QPs (22, 27, 32 and 37) is used both as the target bitrate for the anchor rate control algorithm in HM10.0 and our proposed algorithm. And the standard test sequences in different classes provided by HEVC are adopted.



(a)

(b)

Fig. 1. Fluctuation of buffer occupancy of the rate control in HM10.0 and the proposed rate control. (a)BasketballDrill(832x480) @1721kbps in LB mode. (b)RaseHorsesC(416x240) @996kbps in LB mode.

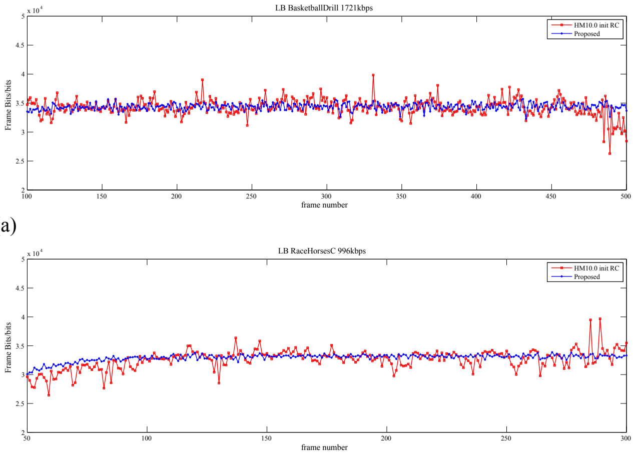
In our proposed rate control algorithm, the buffer size B_s is set to *R/f* in the experiments for low delay applications. To avoid the buffer overflow and underflow are major objectives.

If the buffer is full, the encoder skips frames until there is available space in the buffer. If the buffer is empty and the input bit number is less than the output bandwidth, buffer underflow will happen and cause low channel utilization.

Fig. 1 shows occupancy curves of the buffer over frames of *RaceHorses* and *BaskeballDrill*, for the anchor and proposed algorithm. Except for the high occupancy at startup due to model adjustment, the proposed rate control algorithm maintains a lower buffer occupancy than the preset buffer size and achieves much steadier buffer occupancy. In comparison, the anchor algorithm cannot keep the buffer status stable and the occupancy of buffer fluctuates heavily, which causes overflow and underflow easily at a small buffer size. In *BasketballDrill* case, the anchor algorithm has a much higher buffer occupancy at the startup and cannot keep the buffer occupancy under the buffer size, that is, it encounters buffer overflow and in *RaceHorsesC* case, the buffer occupancy falls below zero and it yields a buffer underflow problem. Since lower and steadier buffer occupancy makes a smaller buffer size possible, which means lower delay in practice, our algorithm is more suitable for low delay applications.

Fig. 2 shows the bit fluctuation curves over frames for various sequences for the anchor and the proposed algorithm. It can be seen that compared with the anchor algorithm, the proposed algorithm has less bit fluctuation. Because of the accurate CTU bit estimation, the proposed algorithm controls the bits more accurately and provides smoother frame bit curve. And the frame bits is close to the buffer output bandwidth R/f, which makes a steadier buffer occupancy and a lower delay.

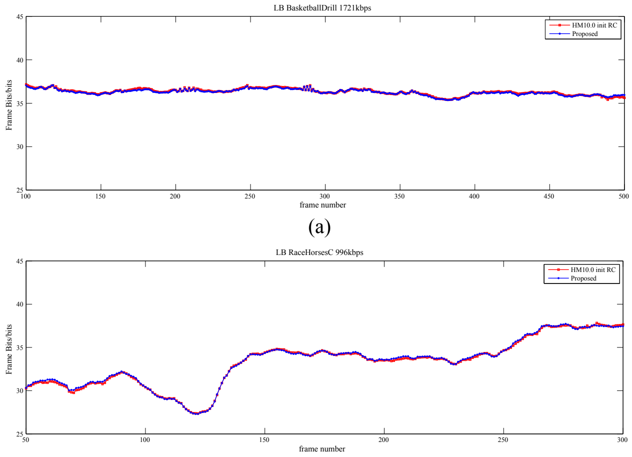
Although the buffer status and the frame bit are controlled more steadily compared with the anchor algorithm, the quality of the proposed algorithm doesn't degrade. Fig. 3 shows the fluctuation curves of quality of *BasketballDrill* and *RaceHorsesC*. It can be seen that the proposed algorithm achieves almost the same quality consistency as the anchor algorithm.



(a)

(b)

Fig. 2. Fluctuation of frame bits of the rate control in HM10.0 and the proposed rate control. (a)BasketballDrill(832x480) @1721kbps in LB mode. (b)RaseHorsesC(416x240) @996kbps in LB mode.



(a)

(b)

Fig. 3. Fluctuation of frame PSNR of the rate control in HM10.0 and the proposed rate control. (a)BasketballDrill(832x480) @1721kbps in LB mode. (b)RaseHorsesC(416x240) @996kbps in LB mode.

To compare the rate estimation accuracy of our proposed algorithm, a mismatch ratio is defined by

$$M\% = \frac{|R_{actual} - R_{target}|}{R_{target}} \times 100\% \tag{19}$$

where R_{target} and R_{actual} denote the target bit rate and the actual bit rate of the test video sequences respectively. As stated before, a sequence adopts the same R_{target} for the anchor algorithm and the proposed algorithm.

Table I states the performance comparisons and the bit rate mismatch comparisons for the HM anchor rate control algorithm and the proposed rate control algorithm. It can be observed that the proposed algorithm has a small mismatch between the actual bit rate and the target bit rate. But the coding efficiency of our proposed algorithm has little

TABLE I
PERFORMANCE COMPARISON OF THE HM10.0 ANCHOR ALGORITHM AND THE PROPOSED ALGORITHM FOR LB-MAIN CONFIGURATION

Sequences	vs HM10.0 without rate control(LB-main)			
	Anchor algorithm		**Proposed algorithm**	
	BD-Rate	***Mismatch***	***BD-Rate***	***Mismatch***
ClassB	8.1%	0.17%	10.2%	0.13%
ClassC	11.2%	0.17%	12.2%	0.07%
ClassD	15.7%	0.44%	17.0%	0.10%
ClassE	23.0%	0.19%	27.2%	0.02%
Avg	14.5%	0.24%	16.7%	0.08%

degradation compared with the anchor algorithm, spending only 2.2% more bit rate for the same objective quality.

V. CONCLUSION

In low delay applications, small buffer size poses a great challenge to rate control algorithms. In this paper, we presented an efficient rate control algorithm of HEVC for low delay video communication. The proposed algorithm adopted a bit allocation scheme based on buffer status and a more accurate bit estimation for CTU using a numerical method. The performance of the proposed algorithm is verified by lots of experimental results. The results show that the algorithm has less bit fluctuation and lower and steadier buffer occupancy, which enables a lower delay in applications.

REFERENCES

[1] B. Bross, W.-j. Han, G. J. sullivan, J.-R. Ohm, T. Wiegand (Editors), “High Efficiency Video Coding (HEVC) test specification draft 9”, Document: JCTVC-K1003, Joint Collaborative Team on Video Coding.

[2] H. Choi, J. Nan, J. Yoo, D. Sim, and I. Bajic, “Rate control based on unified RQ model for HEVC,” Document JCTVC-H0213, San Jos, CA, USA, Feb 2012.

[3] Houqiang Li, Li Li and Bin Li, “Rate control for HEVC Temporal Scalability using R-λ Model,” IEEE Journal of Selected Topics in Signal Processing, Jan.2013.

[4] Shanshe Wang, Siwei Ma, Shiqi Wang, Debin Zhao, Wen Gao, “Quadratic ρ-domain based rate control algorithm for HEVC,” IEEE ICASSP2013, Apr.2013

[5] K.-p. Lim, G. J. Sulivan, and T.Wiegand, “Text Description of Joint Model Reference Encoding Method and Decoding Concealment,” Document JVT-N046, Hong Kong, CN, Jan 2005.

[6] Z. He, Y. K. Kim, and S. Mitra, “Low-delay rate control for dct video coding via rho-domain source modeling,” Circuits and Systems for Video Technology, IEEE Transactions on, vol. 11, no. 8, pp. 928-940,aug 2001.

[7] B. Li, D. Zhang, H. Li, and J. Xu, “QP determination by lambda value”, Document JCTVC-I0426, Geneva, CH, 27 April - 7 May 2012.