A NOVEL METHOD ON OPTIMAL BIT ALLOCATION AT LCU LEVEL FOR RATE CONTROL IN HEVC

Shengxi Li, Mai Xu* †, and Zulin Wang**

∗ School of Electronic and Information Engineering, Beihang University, Beijing, 100191, China
† EDA Lab, Research Institute of Tsinghua University in Shenzhen, Shenzhen, China
ShengxiLi2014@gmail.com, MaiXu@buaa.edu.cn, and wzulin@buaa.edu.cn

ABSTRACT

In this paper, we propose a new method, namely recursive Taylor expansion (RTE) method, for optimally allocating bits to each LCU in the R-λ rate control scheme for HEVC. Specifically, we first set up an optimization formulation on optimal bit allocation. Unfortunately, it is intractable to achieve a closed-form solution for this formulation. We therefore propose a RTE solution to iteratively solve the formulation with a fast convergence speed. Then, an approximate closed-form solution can be obtained. This way, the optimal bit allocation can be achieved at little encoding complexity cost. Finally, the experimental results validate the effectiveness of our method in three aspects: compressed distortion, bit-rate control error, and bit fluctuation.

Index Terms— HEVC, rate control, Taylor expansion

1. INTRODUCTION

With the explosive increasing of multimedia data, especially the video data, the limited bandwidth issue becomes more and more serious. Therefore, video coding is required to save bits of video stream at the cost of video quality loss. For video coding, rate control aims at minimizing distortion of video quality and meanwhile satisfying constraints on bit-rate. There are many rate control schemes for different video coding standards (TM5 for MPEG-2 [6], VM8 for MPEG-4 [17], and JVT-N046 [14] for H.264). For example, Liu *et al.* [10] proposed a novel optimization rate control scheme to improve rate distortion performance of H.264. It predicts mean absolute difference (MAD) first and then solves an optimization formulation to minimize overall distortion at a specific target bits. However, it relies on the rate quantization (RQ) model, which is inaccurate in the state-of-the-art rate control schemes.

Recently, high efficiency video coding (HEVC) standard has been formally established, with more eminent compression performance than the preceding H.264/AVC standard [19]. Since HEVC utilizes the flexible picture partition, parallel coding, and some other cutting-edge technologies, the rate control schemes need to be redeveloped corresponding to these new features. In [2], a pixel-wise unified rate control (URQ) scheme has been proposed to allocate bits via a term bpp, and to assign quantization parameter (QP) using a quadratic equation of QP and target bpp. However, according to [9], the RQ relationship is hard to be precisely estimated. It thus makes the URQ scheme lack of efficiency. Fortunately, the relationship between parameter λ[1] [15] and bit-rate R can be better characterized than previous ways modeling RQ relationship. Therefore, a new scheme, namely R-λ scheme, was proposed in [9], applying the predicted MAD to bit allocation.

Most recently, there have been many methods proposed to improve the R-λ rate control scheme. Rather than the predicted MAD, the sum of absolute transformed difference (SATD) is utilized in [13, 16] to allocate target bits. Moreover, in [4], a pre-encoding of multiple-QPs is proceeded to estimate the SATD-rate-distortion relationship for the bit allocation in rate control. As the cost, it largely increases encoding complexity due to the pre-encoding process. In addition, Yang *et al.* [20] proposed a bit allocation method, adopting bit budget and buffer occupancy to smooth the bit fluctuation. Besides, Si *et al.* [12] used a feedback strategy to avoid sudden increasing of QP in poor quality regions that can lead to propagation for the following frames. However, all existing bit allocation methods above for the R-λ scheme are not on optimizing distortion, which is the ultimate aim of rate control.

To this end, we propose a new method, called recursive Taylor expansion (RTE) method, to achieve the optimal bit allocation at largest coding unit (LCU) level for HEVC. Specifically, we first establish a formulation on optimizing bit allocation to each LCU, achieving minimal distortion over the video frame. However, it is impractical to obtain the closed-form solution to this formulation. Thus, we propose the RTE solution to obtain the approximate closed-form solution to the bit allocation formulation, via iterating the Taylor expansion. In our RTE method, the convergence speed is indeed fast, as the approximation error decreases to 10^{-10} with no more than three iterations. This way, our RTE method can achieve the optimal bit allocation with little encoding complexity cost. In

This work was supported by NSFC under grant number 61202139 and China 973 program under grant number 2013CB329006.

[1] Lagrange multiplier λ stands for the slope of rate distortion curve.

addition, the whole distortion is able to be minimized without extra consumed bits, ensuring the stability of bit-rate control.

2. THE EXISTING R-λ RATE CONTROL SCHEME

The main idea of rate control is optimally allocating bits to each coding part via rate-distortion optimization (RDO). As such, distortion of the compressed video can be minimized at a given bit-rate. Since our scheme mainly works at LCU level, we only focus on reviewing LCU level rate control. The main steps of the R-λ rate control scheme [9] are illustrated in Figure 1, to be discussed in the following.

At LCU level, λ for the i-th LCU (denoted as λ_i), indicates the slope of rate distortion (RD) curve for this LCU. It can be formulated [7] by

$$\lambda_i = -\frac{\partial d_i}{\partial r_i}, \tag{1}$$

where d_i and r_i represent the distortion and target bits for the i-th LCU.

There are many models [3, 11, 15, 18] to work out the relationship between d_i and r_i by RD curve fitting. However, as HEVC has many new features on video coding, Li. *et al.* [9] found out that the Hyperbolic model [3, 11] performs better than other models. Such a model can be expressed by

$$d_i = c_i {r_i}^{-k_i}, \tag{2}$$

where c_i and k_i are fitting parameters of $d_i(r_i)$, related to the content of the i-th LCU [9]. Then, based on (1) and (2), R-λ relationship can be established:

$$\lambda_i = -\frac{\partial d_i}{\partial r_i} = c_i k_i \cdot (\mathrm{bpp_i} \cdot w_i \cdot h_i)^{-k_i - 1} = \alpha_i \cdot {\mathrm{bpp_i}}^{\beta_i}, \tag{3}$$

where w_i, h_i, and $\mathrm{bpp_i}$ are the width, height, and target bits per pixel for the i-th LCU. In addition, $\alpha_i = c_i k_i \cdot (w_i h_i)^{-k_i-1}$ and $\beta_i = -k_i - 1$ are the parameters for fitting λ_i and $\mathrm{bpp_i}$. Therefore, with (3), we can acquire the "best" λ_i to minimize distortion of the i-th LCU at a given bpp_i.

However, in practice, α_i and β_i are unknown, and may be different from one LCU to another. So, they need to be updated for the subsequent frames after encoding each LCU, according to its real consumed bpp_i and actual λ_i. Details about the updating method can refer to [9].

According to (3), λ_i for each LCU can be acquired once $\mathrm{bpp_i}$ is determined. The remaining task for λ_i is thus the bit allocation for each LCU (i.e., computing $\mathrm{bpp_i}$). In the R-λ rate control scheme [9], the way to allocate target bits is

$$\mathrm{bpp_i} = \frac{T_i}{N_i}, \text{ and } T_i = (\widehat{T} - B) \cdot h_i / (\sum_{j=i}^{M} h_j), \tag{4}$$

where N_i and T_i denote the number of pixels and target bits for the i-th LCU, respectively. In (4), M denotes the number of LCUs in the currently compressed frame, $\widehat{T}$ is the target bits left for encoding this frame, and B is the estimated header bits. Moreover, h_i stands for the weight of the i-th LCU, which may be represented by the predicted MAD of the collocated picture in previous coded frames [9]. Finally, with estimated $\mathrm{bpp_i}$ in (4), λ_i of the i-th LCU can be output upon (3) for determining its QP value.

In summary, as illustrated in Figure 1, once bpp_i is allocated for each LCU by (4), rate control can be achieved via estimating its corresponding λ_i and "best" QP_i. Bit allocation to each LCU is thus crucial in the R-λ scheme. However, the state-of-the-art bit allocation in (4) is not an optimal one, as it does not target at minimizing the overall distortion within a video frame. To avoid such a disadvantage, we propose the RTE method on optimally assigning bits to each LCU, achieving minimal distortion over the whole video frame.

3. RTE BASED OPTIMAL BIT ALLOCATION

This section proposes a new method, namely RTE method, towards optimal bit allocation to LCUs, for rate control in HEVC. Specifically, we first establish a formulation on optimal LCU bit allocation. However, it is intractable to obtain the closed-form solution to the proposed formulation. We thus utilize Taylor expansion to solve this formulation. Nevertheless, the Taylor expansion solution has large approximation error, resulting in non-optimal bit allocation. Thereby, we propose the RTE method to obtain approximate closed-form solution to optimal LCU bit allocation.

3.1. Formulation on Optimal LCU Bit Allocation

The rate control within a video frame is minimizing overall distortion D at a given amount of target bits R, formulated by

$$\min_{\{r_i\}_{i=1}^{M}} D = \Sigma_{i=1}^{M} d_i \quad \text{s.t. } \Sigma_{i=1}^{M} r_i \le R, \tag{5}$$

where D and R denote the distortion and target bits for the currently compressed video frame. Recall that M is the total number of LCUs in the frame. Given Lagrange multiplier λ, (5) can be converted to an unconstrained problem [15]:

$$\min_{\{r_i\}_{i=1}^{M}} \Sigma_{i=1}^{M} \left(d_i(r_i) + \lambda r_i\right). \tag{6}$$

Note that λ_i in (3) is the Lagrange multiplier for finding optimal QP_i to minimize the distortion in an LCU. Different from λ_i, λ in (6) is the Lagrange multiplier for optimally allocating bits with a minimal overall distortion in a frame.

Next, setting the derivative of $\Sigma_{i=1}^{M}(d_i(r_i) + \lambda r_i)$ to zero, (6) can be uniquely solved as follows,

$$\lambda = -\frac{\partial d_i}{\partial r_i} \quad \text{and} \quad \Sigma_{i=1}^{M} r_i = R, \qquad i = 0, 1, 2 \ldots M. \tag{7}$$

As we have mentioned in Section 2, the Hyperbolic model $d_i = c_i r_i^{-k_i}$ is able to better characterize the relationship between distortion and target bits. Thereby, for (7), we have the following formulation on optimal LCU bit allocation:

$$\sum_{i=1}^{M} r_i = \sum_{i=1}^{M} \left(\frac{\lambda}{c_i k_i}\right)^{-\frac{1}{k_i+1}} = \sum_{i=1}^{M} \left(\frac{a_i}{\lambda}\right)^{b_i} = R, \tag{8}$$

where $a_i = c_i k_i$ and $b_i = \frac{1}{k_i+1}$. Note that c_i and k_i are known parameters introduced in Section 2. Therefore, once the "best" λ is achieved in currently compressed video frame, the bits can be optimally allocated to each LCU via (8). Then, the remaining task is to calculate the "best" λ.

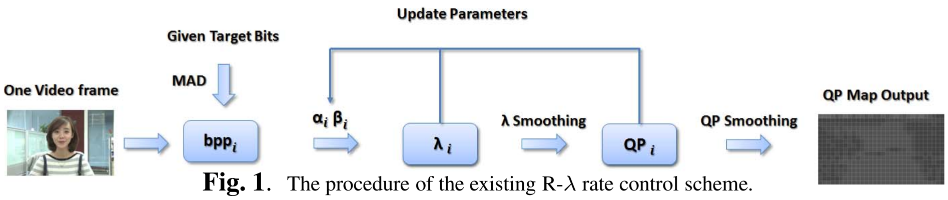

Fig. 1. The procedure of the existing R-λ rate control scheme.

However, since b_i varies among LCUs, it is hard to derive a closed-from solution to λ in (8). The existing greedy search algorithms can solve this equation, but they usually cannot ensure the global optimum. Besides, the computational complexity of greedy search algorithms is heavy, which may hugely increase the encoding complexity of HEVC. Thus, the RTE method is proposed to solve (8), for obtaining the optimal bit allocation with little complexity cost.

3.2. Taylor Expansion Solution

To deal with different exponent b_i in (8), we rewrite $(\frac{a_i}{\lambda})^{b_i}$ utilizing Taylor expansion:

$$(\frac{a_i}{\lambda})^{b_i} = 1 + \frac{\ln(\frac{a_i}{\lambda})}{1!}b_i + \cdots + \frac{(\ln\frac{a_i}{\lambda})^n}{n!}b_i^n + \cdots. \quad (9)$$

Then, we discard the biquadratic and higher-order terms in the above Taylor expansion. As a result, the following approximation holds:

$$\begin{aligned}
&(\frac{a_i}{\lambda})^{b_i}\\
&\approx 1 + \frac{\ln(\frac{a_i}{\lambda})}{1!}b_i + \frac{(\ln\frac{a_i}{\lambda})^2}{2!}b_i^2 + \frac{(\ln\frac{a_i}{\lambda})^3}{3!}b_i^3\\
&= -\frac{b_i^3}{6}\ln^3\lambda + (\frac{b_i^2}{2} + \frac{b_i^3}{2}\ln a_i)\ln^2\lambda \qquad (10)\\
&-(b_i^2\ln a_i + b_i + \frac{b_i^3}{2}\ln^2 a_i)\ln\lambda + (1 + b_i\ln a_i + \frac{b_i^2}{2}\ln^2 a_i + \frac{b_i^3}{6}\ln^3 a_i).
\end{aligned}$$

Accordingly, (8) can be approximated by

$$\begin{aligned}
R = \sum_{i=1}^{M}(\frac{a_i}{\lambda})^{b_i} \approx &-\underbrace{\sum_{i=1}^{M}(\frac{b_i^3}{6})}_{A}\ln^3\lambda\\
&+\underbrace{\sum_{i=1}^{M}(\frac{b_i^2}{2} + \frac{b_i^3}{2}\ln a_i)}_{B}\ln^2\lambda - \underbrace{\sum_{i=1}^{M}(b_i^2\ln a_i + b_i + \frac{b_i^3}{2}\ln^2 a_i)}_{C}\ln\lambda\\
&+\underbrace{\sum_{i=1}^{M}(1 + b_i\ln a_i + \frac{b_i^2}{2}\ln^2 a_i + \frac{b_i^3}{6}\ln^3 a_i)}_{D}. \qquad (11)
\end{aligned}$$

Applying Shengjin formula [5], the cubic equation in (11) can be worked out to obtain the estimated λ (denoted by $\widehat{\lambda}$) as[2]:

$$\widehat{\lambda} = e^{\frac{-B-(\sqrt[3]{Y_1}+\sqrt[3]{Y_2})}{3A}}, Y_{1,2} = BE + 3A(\frac{-F \pm \sqrt{F^2 - 4EG}}{2}), \quad (12)$$

where $E = B^2 - 3AC$, $F = BC - 9A(D - R)$, and $G = C^2 - 3B(D - R)$. Finally, given $\widehat{\lambda}$, the bit allocation can be achieved using (8).

However, since the value of $\ln\frac{a_i}{\lambda}$ is normally very large in practical encoding, the truncation of higher-order terms in Taylor expansion may lead to great approximation error in (11). As a result, $\widehat{\lambda}$ estimated by (12) is not the "best", such that the bit allocation is not sufficiently optimal. Therefore, the RTE method is proposed to recursively implement Taylor expansion for significantly improving the approximation accuracy.

3.3. Recursive Taylor Expansion Solution

As discussed above, the truncation of higher-order terms in (9) may result in large approximation error of Taylor expansion solution in Section 3.2. In fact, the approximation error depends on the decay rate of Taylor expansion, and large decay rate leads to small approximation error. It is defined as

$$\left|\frac{\frac{(\ln\frac{a_i}{\lambda})^n}{n!}{b_i}^n}{\frac{(\ln\frac{a_i}{\lambda})^{n+1}}{(n+1)!}{b_i}^{n+1}}\right| = \left|\frac{n+1}{(\ln\frac{a_i}{\lambda})\cdot b_i}\right|. \quad (13)$$

As seen from (13), it is possible to increase the decay rate of Taylor expansion by reducing the value of $|\ln\frac{a_i}{\lambda}|$ (i.e., $\frac{a_i}{\lambda} \to 1$). To this end, we replace $(\frac{a_i}{\lambda})^{b_i}$ by $\widetilde{r}_i(\frac{\widetilde{\lambda}}{\lambda})^{b_i}$ via separating a_i, where $\widetilde{\lambda}$ ($\frac{\widetilde{\lambda}}{\lambda} \to 1$) is the pre-estimated λ of the currently compressed frame, and $\widetilde{r}_i = (\frac{a_i}{\widetilde{\lambda}})^{b_i}$ is the pre-estimated target bits for the i-th LCU. Then, one can obtain

$$\sum_{i=1}^{M}(\frac{a_i}{\lambda})^{b_i} = \sum_{i=1}^{M}\widetilde{r}_i(\frac{\widetilde{\lambda}}{\lambda})^{b_i} = R. \quad (14)$$

Similar to (10), the following approximation on $(\frac{a_i}{\lambda})^{b_i}$ holds

$$\begin{aligned}
&(\frac{a_i}{\lambda})^{b_i} = \widetilde{r}_i(\frac{\widetilde{\lambda}}{\lambda})^{b_i}\\
&= \widetilde{r}_i + \widetilde{r}_i\frac{\ln(\frac{\widetilde{\lambda}}{\lambda})}{1!}b_i + \widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^2}{2!}b_i^2 + \cdots + \widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^n}{n!}b_i^n + \cdots\\
&\approx \widetilde{r}_i(1 + \frac{\ln(\frac{\widetilde{\lambda}}{\lambda})}{1!}b_i + \frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^2}{2!}b_i^2 + \frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^3}{3!}b_i^3). \qquad (15)
\end{aligned}$$

The decay rate of Taylor expansion in (15) is $|\frac{n+1}{(\ln\frac{\widetilde{\lambda}}{\lambda})\cdot b_i}|$. It is much larger than the decay rate in (9), when $|\ln\frac{\widetilde{\lambda}}{\lambda}| \ll |\ln\frac{a_i}{\lambda}|$ and $b_i > 0$. Consequently, the approximation error of Taylor expansion can be reduced. Consider a special case that $\widetilde{\lambda} = \lambda$. In this case, there exists no approximation error of (15) as $\ln\frac{\widetilde{\lambda}}{\lambda} = 0$ makes the decay rate approaching to infinity. Accordingly, $\widetilde{r}_i$ is the optimal bit allocation for each LCU.

Therefore, based on the 3rd order Taylor approximation of (15), (14) can be rewritten as:

[2] As $\Delta = F^2 - 4EG > 0$ in practical encoding, there only exists one real root for (11). Therefore, $\widehat{\lambda}$ value is unique.

$$R=\sum_{i=1}^{M}\widetilde{r}_i(\frac{\widetilde{\lambda}}{\lambda})^{b_i}\approx\underbrace{-\sum_{i=1}^{M}(\widetilde{r}_i\frac{b_i^3}{6})\ln^3\lambda}_{A'}$$

$$+\underbrace{\sum_{i=1}^{M}\widetilde{r}_i(\frac{b_i^2}{2}+\frac{b_i^3}{2}\ln\widetilde{\lambda})\ln^2\lambda}_{B'}\underbrace{-\sum_{i=1}^{M}\widetilde{r}_i(b_i^2\ln\widetilde{\lambda}+b_i+\frac{b_i^3}{2}\ln^2\widetilde{\lambda})\ln\lambda}_{C'}$$

$$+\underbrace{\sum_{i=1}^{M}\widetilde{r}_i(1+b_i\ln\widetilde{\lambda}+\frac{b_i^2}{2}\ln^2\widetilde{\lambda}+\frac{b_i^3}{6}\ln^3\widetilde{\lambda})}_{D'}. \quad (16)$$

Finally, $\widehat{\lambda}$ can be obtained according to (12), with A, B, C, and D replaced by A', B', C', and D'. Note that $\widehat{\lambda}$ acquired from (16) is with less approximation error than that from (11).

Next, it is possible to further reduce the approximation error of (16) by making $\widetilde{\lambda}$ close to the "best" λ. However, it is impossible to obtain the "best" λ in practical encoding, as the "best" λ is the variable to be solved. So, there is a chicken-and-egg dilemma between $\widetilde{\lambda}$ and the "best" λ. In fact, we can iterate the Taylor expansion via utilizing the estimated $\widehat{\lambda}$ as the input $\widetilde{\lambda}$ for the next iteration. Then, the approximation error can be reduced alongside the iterations, once Lemma 2 exists. Note that if $\widetilde{\lambda}>\lambda>0$ at the first iteration, its output $\widehat{\lambda}$ is smaller than λ, as pointed out by Proposition 1. Then, for the subsequent iterations that replace the value of $\widetilde{\lambda}$ by $\widehat{\lambda}$, $0<\widetilde{\lambda}<\lambda$ can be achieved.

Proposition 1 *Consider $\widetilde{\lambda}>0$, $\lambda>0$, $b_i>0$, $\lambda\neq\widetilde{\lambda}$, and $R>0$ for* (16)*. If $\widehat{\lambda}$ is the estimated solution of λ to* (16)*, then the following holds*

$$\widehat{\lambda}<\lambda. \quad (17)$$

Proof: *Towards the Taylor expansion of $\sum_{i=1}^{M}\widetilde{r}_i(\frac{\widetilde{\lambda}}{\lambda})^{b_i}$ in* (16)*, we can obtain the following equations:*

$$R=\sum_{i=1}^{M}\widetilde{r}_i(\frac{\widetilde{\lambda}}{\lambda})^{b_i}$$
$$=\sum_{i=1}^{M}\widetilde{r}_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{\ln(\frac{\widetilde{\lambda}}{\lambda})}{1!}b_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^2}{2!}b_i^2+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^3}{3!}b_i^3$$
$$=\sum_{i=1}^{M}\widetilde{r}_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{\ln(\frac{\widehat{\lambda}}{\lambda})}{1!}b_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^2}{2!}b_i^2+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^3}{3!}b_i^3$$
$$+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^4}{4!}b_i^4+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^5}{5!}b_i^5+\cdots. \quad (18)$$

For $\widetilde{\lambda}>\lambda>0$ and $b_i>0$, we can obtain $(\ln\frac{\widetilde{\lambda}}{\lambda})\cdot b_i>0$. It is known that (18) *holds with $R>\sum_{i=1}^{M}\widetilde{r}_i>0$ and $\ln\frac{\widetilde{\lambda}}{\widehat{\lambda}}>0$, because of $\frac{\widetilde{\lambda}}{\widehat{\lambda}}>1$. Thus, $\ln\frac{\widetilde{\lambda}}{\widehat{\lambda}}>\ln\frac{\widetilde{\lambda}}{\lambda}>0$ exists, such that $\widehat{\lambda}<\lambda$ can be achieved.*

For $\lambda>\widetilde{\lambda}>0$ and $b_i>0$, we have

$$\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^4}{4!}b_i^4+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^5}{5!}b_i^5+\cdots>0. \quad (19)$$

Then, with (18)*, the following inequality exists:*

$$\sum_{i=1}^{M}\widetilde{r}_i\frac{\ln(\frac{\widetilde{\lambda}}{\lambda})}{1!}b_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^2}{2!}b_i^2+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\lambda})^3}{3!}b_i^3$$
$$>\sum_{i=1}^{M}\widetilde{r}_i\frac{\ln(\frac{\widetilde{\lambda}}{\widehat{\lambda}})}{1!}b_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\widehat{\lambda}})^2}{2!}b_i^2+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\widehat{\lambda}})^3}{3!}b_i^3. \quad (20)$$

Moreover, the following function monotonously decreases to 0 *along with the increasing of variable x (up to $\widetilde{\lambda}$):*

$$\sum_{i=1}^{M}\widetilde{r}_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{\ln(\frac{\widetilde{\lambda}}{x})}{1!}b_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{x})^2}{2!}b_i^2+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{x})^3}{3!}b_i^3. \quad (21)$$

Therefore, we can obtain that $\widehat{\lambda}<\lambda$ by combining (20) *and* (21)*. This completes the proof.*

Lemma 2 *Consider $\lambda>\widetilde{\lambda}>0$, $b_i>0$, and $R>0$ for* (16)*. When the estimated solution of λ to* (16) *is $\widehat{\lambda}$, the following inequality holds for $\widehat{\lambda}$,*

$$|\widehat{\lambda}-\lambda|<|\widetilde{\lambda}-\lambda|. \quad (22)$$

Proof: *Since $\widehat{\lambda}$ is the solution of λ to the 3rd order Taylor expansion on $\sum_{i=1}^{M}\widetilde{r}_i(\frac{\widetilde{\lambda}}{\lambda})^{b_i}$, the following equation exists:*

$$R=\sum_{i=1}^{M}\widetilde{r}_i(\frac{\widetilde{\lambda}}{\lambda})^{b_i} \quad (23)$$
$$=\sum_{i=1}^{M}\widetilde{r}_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{\ln(\frac{\widetilde{\lambda}}{\widehat{\lambda}})}{1!}b_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\widehat{\lambda}})^2}{2!}b_i^2+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\widehat{\lambda}})^3}{3!}b_i^3.$$

In fact, $0<(\frac{\widetilde{\lambda}}{\lambda})^{b_i}<1$ holds for $0<\widetilde{\lambda}<\lambda$ and $b_i>0$. Besides, there exists $R=\sum_{i=1}^{M}\widetilde{r}_i(\frac{\widetilde{\lambda}}{\lambda})^{b_i}$ in (23)*. Therefore, $\sum_{i=1}^{M}\widetilde{r}_i>R$ can be worked out.*

Next, assuming that $\widehat{\lambda}\leq\widetilde{\lambda}$, we have $\ln\frac{\widetilde{\lambda}}{\widehat{\lambda}}\geq 0$. Due to $\sum_{i=1}^{M}\widetilde{r}_i>R$, $\ln\frac{\widetilde{\lambda}}{\widehat{\lambda}}\geq 0$, and $b_i>0$, the inequality below can be achieved,

$$\sum_{i=1}^{M}\widetilde{r}_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{\ln(\frac{\widetilde{\lambda}}{\widehat{\lambda}})}{1!}b_i+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\widehat{\lambda}})^2}{2!}b_i^2+\sum_{i=1}^{M}\widetilde{r}_i\frac{(\ln\frac{\widetilde{\lambda}}{\widehat{\lambda}})^3}{3!}b_i^3>R, \quad (24)$$

which is contradictory with the equation in (23)*. Therefore, it can be proved that $\widetilde{\lambda}<\widehat{\lambda}$. Then, combining with Proposition 1, $\widetilde{\lambda}<\widehat{\lambda}<\lambda$ can be achieved. As a result, $|\widehat{\lambda}-\lambda|<|\widetilde{\lambda}-\lambda|$ exists. This completes the proof of Lemma 2.*

Utilizing the iteration process, the RTE method can be achieved via the following steps:

- **Step 1**: **Initialize** pre-estimated $\widetilde{\lambda}$ to be the picture λ.
- **Step 2**: **Calculate** A', B', C', and D' of (16) with $\widetilde{\lambda}$.
- **Step 3**: **Obtain** $\widehat{\lambda}$ estimated by (12).

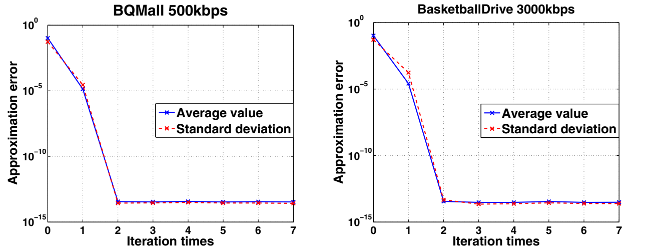

Fig. 2. The average value and standard deviation of approximation error for each iteration. The approximation error is defined as $|\Sigma_{i=1}^{M}(\frac{a_i}{\widehat{\lambda}})^{b_i} - R|/R$.

- **Step 4**: **Update** $\widetilde{\lambda}$ to be $\widehat{\lambda}$ obtained from step 3, for the next iteration.
- **Step 5**: **Judge** whether $\widehat{\lambda}$ meets the convergence criterion. Go to step 6 once satisfying the convergence criterion. Otherwise, go to step 2.
- **Step 6**: **Apply** (14) with $\widehat{\lambda}$ to optimally allocating bits to each LCU.

Note that for the first iteration of our RTE method, we set the initial $\widetilde{\lambda}$ as the picture λ (denoted by λ_P) [9], which is calculated for the frame level rate control. In practice, it is closer to the "best" λ than a_i. Moreover, the convergence criterion is set as $|\Sigma_{i=1}^{M}(\frac{a_i}{\widehat{\lambda}})^{b_i} - R|/R < 10^{-10}$.

After several iterations, RTE can reduce the difference between $\widehat{\lambda}$ and the "best" λ to a small range, meeting the convergence criterion. Thus, $\widehat{\lambda}$ can be seen as the approximate closed-form solution to the "best" λ of (14). As such, the bits can be almost optimally allocated to each LCU. The approximation error of each iteration is illustrated in Figure 2. Similar results can be found in other cases. So, we can conclude that our RTE method converges fast, and three iterations can guarantee the approximation error of $\widehat{\lambda}$ to be less than 10^{-10}.

4. EXPERIMENTAL RESULTS

4.1. Parameter setting

In this section, experimental results are presented to validate our RTE method. We chose four video sequences (*BasketballDrive*, *KristenAndSara*, *Vidyo3*, and *BQMall*), which belong to B, C, and E classes from the standard video database [1].

In addition, we set HM 14.0 as an anchor [8] in our experiment. Furthermore, the low delay IPPP structure was chosen for comparison, using the configuration file *encoder_lowdelay_P_main.cfg*. Since we mainly focus on the bit allocation at LCU level, non-hierarchical bit allocation at frame level was applied in our experiments to reduce the influence by other frames. Other parameters were set by default.

We utilized the state-of-the-art bit allocation method by Li *et al.* [9] for comparison. Note that such a method is not optimal. Moreover, both Li *et al.* and our RTE methods were implemented on HM 14.0 platform, and the difference merely exists in the bit allocation for R-λ rate control scheme.

4.2. Evaluation on Rate and Distortion

This section verifies the performance of the bit allocation methods in terms of distortion and bit-rate control error. Here, the distortion is assessed by PSNR. The bit-rate control error (denoted as e) is calculated [16] as follows,

$$e = |\frac{\widehat{E} - \overline{E}}{\overline{E}}|, \tag{25}$$

where $\overline{E}$ and $\widehat{E}$ are target and actual bit-rates.

Table 1 shows the comparison of PSNR and bit-rate control error between two bit allocation methods. As can be seen from this table, for all video sequences at various bit-rates, our RTE method has the larger PSNR improvement. Besides, smaller bit-rate control error is achieved by our RTE method in nearly all cases. To be more specific, for PSNR, our method gains average 0.211 dB increment with up to 0.378 dB improvement. Meanwhile, the accuracy of our RTE method can be doubled over Li *et al.* method. Therefore, we can conclude from Table 1 that our RTE method outperforms Li *et al.* method in terms of both distortion and rate control accuracy.

Table 2. Comparison on standard deviation of bit fluctuation between Li *et al.* [9] and our RTE methods.

Sequences	Bit-rates (kbps)	Li *et al.* [9] (bits)	RTE (bits)
BasketballDrive	1000	1947.85	**1938.95**
	3000	4159.69	**3489.16**
KristenAndSara	1000	1821.39	**1115.99**
	1500	2738.22	**1473.66**
Vidyo3	1000	1537.91	**1104.94**
	1500	2329.93	**1472.13**
BQMall	500	767.14	**759.69**
	1500	1449.40	**1428.88**
Average Standard Deviation		2093.94	**1597.93**

4.3. Evaluation on Bit Fluctuation

Bit fluctuation is also a crucial factor when evaluating rate control, especially for low-delay cases. If the actual bits are equal to the target bits assigned for each frame, there is no bit fluctuation for non-hierarchical bit allocation. On the other hand, the large bit fluctuation may cause the overflow or underflow of buffers. Therefore, the comparison of bit fluctuation for Li *et al.* [9] and our RTE methods is conducted to further validate the effectiveness of our RTE method.

Figure 3 shows the comparison of bit fluctuation for all frames between two methods. Note that the first frame is not considered as it is an I-frame. The figure indicates that our RTE method is more steady than Li *et al.* method, with less abrupt fluctuation and smaller deviation from the average value.

To quantify the bit fluctuation, Table 2 presents the standard deviations for each method. From this table, we can conclude that for all videos at various bit-rates, the standard deviations of our RTE method are smaller with average 23.69% reduction compared with Li *et al.* method. It means that our RTE method has smaller bit fluctuation. This may be due to the fact that the optimal bit allocation can improve the PSNR for each frame without extra bits allocated to this frame. The actual bits for each frame are thus with less bit fluctuation.

5. CONCLUSION

In this paper, we have proposed a new method, called RTE method, to optimally allocate bits for LCU level rate control

Table 1. Comparison on PSNR and bit-rate control error between Li *et al.* [9] and our RTE methods of HM 14.0.

Sequences	Target Bit-rates	Li *et al.* Method [9]			Our RTE Method			
	kbps	PSNR (dB)	Bit-rates (kbps)	e (‰)	PSNR (dB)	Increment (dB)	Bit-rates (kbps)	e (‰)
BasketballDrive	1000	30.913	1000.014	0.014	**31.291**	0.378	999.996	**0.004**
	3000	34.797	3000.092	**0.031**	**35.014**	0.217	3000.119	0.040
KristenAndSara	1000	41.507	1000.160	0.160	**41.670**	0.163	1000.027	**0.027**
	1500	42.357	1500.189	0.126	**42.517**	0.160	1500.082	**0.054**
Vidyo3	1000	40.289	1000.010	0.010	**40.519**	0.231	999.990	**0.010**
	1500	41.202	1500.046	0.031	**41.441**	0.239	1500.034	**0.023**
BQMall	500	31.276	500.025	0.050	**31.455**	0.179	500.010	**0.020**
	1500	35.742	1500.038	0.025	**35.865**	0.123	1500.032	**0.021**
Average PSNR Increment		− −			**0.211** dB			
Average e		0.056 ‰			**0.025** ‰			

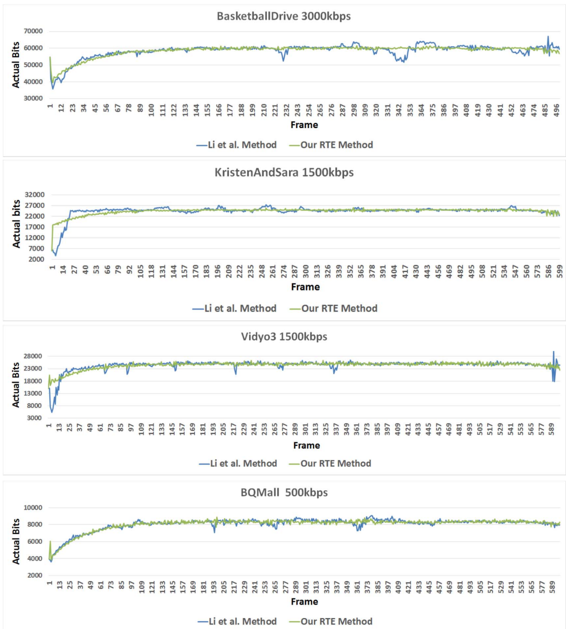

Fig. 3. Bit fluctuation comparison for Li *et al.* [9] and our RTE methods.

in HEVC. First, we established a formulation for optimal bit allocation, using Lagrange multiplier on distortion minimization and bit-rate constraint. Then, since it is impossible to acquire the closed-form solution, RTE was developed to obtain the approximate closed-form solution for the optimal bit allocation formulation. Due to the quick convergence property of our RTE method, the approximate closed-form solution can be reached during several iterations (no more than three iterations). Thus, the optimal bit allocation can be achieved in our RTE method with less encoding complexity cost. Finally, experimental results show that our RTE method not only performs well in quality, but also in accuracy and stationary of bit-rate control.

6. REFERENCES

[1] F. Bossen, "Common test conditions and software reference configurations," *Document: JCTVC-G1200, Joint Collaborative Team on Video Coding*, Nov. 2011.

[2] H. Choi, J. Yoo, J. Nam, D. Sim, and I. Bajic, "Pixel-wise unified rate-quantization model for multi-level rate control," *Journal of Selected Topics in Signal Processing*, 2013.

[3] M. Dai, D. Loguinov, and H. Radha, "Rate-distortion modeling of scalable video coders," in *ICIP*, vol. 2, 2004, pp. 1093–1096.

[4] L. Deng, F. Pu, S. Hu, and C.-C. J. Kuo, "HEVC encoder optimization based on a new RD model and pre-enoding," in *30th PCS*, 2014.

[5] S. Fan, "A new extracting formula and a new distinguishing means on the one variable cubic equation," *Nat. Sci. J. Hainan Teach. Coll*, vol. 2, pp. 91–98, 1989.

[6] B. G. Haskell, *Digital Video: An Introduction to MPEG-2: An Introduction to MPEG-2*. Springer, 1997.

[7] Y.-H. Huang, T.-S. Ou, P.-Y. Su, and H. H. Chen, "Perceptual rate-distortion optimization using structural similarity index as quality metric," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 20, no. 11, pp. 1614–1624, 2010.

[8] JCT-VC, "HM 14.0," http://hevc.hhi.fraunhofer.de/.

[9] B. Li, H. Li, L. Li, and J. Zhang, "λ domain based rate control for high efficiency video coding," *Image Processing, IEEE Transactions on*, vol. 23, no. 9, Sep. 2014.

[10] Y. Liu, Z. G. Li, and Y. C. Soh, "A novel rate control scheme for low delay video communication of H.264/AVC standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 1, pp. 68–78, 2007.

[11] S. Mallat and F. Falzon, "Analysis of low bit rate image transform coding," *Signal Processing, IEEE Transactions on*, vol. 46, no. 4, pp. 1027–1042, 1998.

[12] J. Si, S. Ma, and W. Gao, "Efficient bit allocation and CTU level rate control for high efficiency video coding." in *PCS*, 2013, pp. 89–92.

[13] J. Si, S. Ma, S. Wang, and W. Gao, "Laplace distribution based CTU level rate control for HEVC," in *VCIP 2013*. IEEE, 2013, pp. 1–6.

[14] G. J. Sullivan, T. Wiegand, and K.-P. Lim, "Text description of joint model reference encoding methods and decoding concealment methods," *Document: JVT-N046, Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG*, Jan. 2005.

[15] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *Signal Processing Magazine, IEEE*, vol. 15, no. 6, pp. 74–90, 1998.

[16] H. Sun, S. Gao, and C. Zhang, "Adaptive bit allocation scheme for rate control in high efficiency video coding with initial quantization parameter determination," *Signal Processing: Image Communication*, vol. 29, no. 10, pp. 1029–1045, 2014.

[17] A. Vetro, H. Sun, and Y. Wang, "MPEG-4 rate control for multiple video objects," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 9, no. 1, pp. 186–199, 1999.

[18] T. Wiegand and B. Girod, "Lagrange multiplier selection in hybrid video coder control," in *ICIP*, vol. 3, 2001, pp. 542–545.

[19] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 560–576, 2003.

[20] Z. Yang, L. Song, Z. Luo, and X. Wang, "Low delay rate control for HEVC," in *BMSB 2014*. IEEE, 2014, pp. 1–5.