# Rate-GOP Based Rate Control for High Efficiency Video Coding

Shanshe Wang, Siwei Ma, Shiqi Wang, Debin Zhao, and Wen Gao, *Fellow, IEEE*

***Abstract*—In this paper, a Rate-GOP based frame level rate control scheme is proposed for High Efficiency Video Coding (HEVC). The proposed scheme is developed with the consideration of the new coding tools adopted into HEVC, including the quad-tree coding structure and the new reference frame selection mechanism, called reference picture set (RPS). The contributions of this paper mainly include the following three aspects. Firstly, a RPS based hierarchical rate control structure is designed to maintain the high video quality of the key frames. Secondly, the inter-frame dependency based distortion model and bit rate model are proposed, considering the dependency between a coding frame and its reference frame. Thus the distortion and bit rate of the coding frame can be represented by the distortion and bit rate of its reference frame. Accordingly, the Rate-GOP based distortion model and rate model can be achieved via the inter-frame dependency based distortion model and bit rate model. Thirdly, based on these models and a mixed Laplacian distribution of residual information, a new $\rho$-domain Rate-GOP based rate control is proposed. Experimental results demonstrate the proposed Rate-GOP based rate control has much better R-D performance. Compared with the two state-of-the-art rate control schemes for HEVC, the coding gain with BD-PSNR can be up to 0.87 dB and 0.13 dB on average respectively for all testing configurations. Especially for random access low complexity testing configuration, the BD-PSNR gain can be up to 1.30 dB and 0.23 dB respectively.**

***Index Terms*—Bit allocation, HEVC, inter-frame dependency, rate control, rate-GOP, video coding.**

## I. Introduction

HIGH efficiency video coding (HEVC) [1] is the latest video coding standard which significantly improves the coding efficiency over the preceding standards, such as H.264/AVC [2]. In HEVC, many new coding tools are adopted to improve the coding performance, and they bring great challenges to establish an accurate rate quantization (R-Q) model and bit allocation scheme in rate control for HEVC. Although HEVC still adopts the traditional hybrid prediction/transform coding framework, in which many rate control algorithms have been proposed before, such as TM5 for MPEG-2 [3], VM8 for MPEG-4 [4] and TMN8 for H.263 [5], yet these rate control schemes cannot work well if they are applied directly to HEVC without considering the new coding tools.

To establish a more accurate R-Q model and a more efficient bit allocation scheme, the rate control algorithm must consider the rate-distortion (R-D) characteristics. How to use the inter-frame dependency to improve the rate control performance is also a difficult problem which may not be addressed well in the past.

Many R-Q models have been proposed in the previous rate control methods. These R-Q models always consider the distribution of the residual information [6]. Under the assumption that the residual information follows a Laplacian distribution, a quadratic R-Q model is proposed in [4] which employs the mean of absolute difference (MAD) to estimate the complexity of basic coding units. Many improved algorithms based on this quadratic model are proposed to improve the coding performance [7]–[10] for H.264/AVC. In [11], it is stated that the Cauchy distribution is more accurate than Laplacian distribution when the distribution of residual information has a long tail. The related rate and distortion models are proposed to provide better rate control performance. In [12], a more accurate linear model in $\rho$-domain is proposed, which is proved to be very effective in most distributions including in generalized Gaussian distribution (GGD).

In predictive coding, inter-frame dependency plays a crucial role. It can improve the coding performance since a better reference frame can lead to better prediction. The dynamic programming algorithm [13] can obtain an optimal solution for inter-frame dependency, however with the great computational complexity, which is unaffordable for real-time video coding. For rate control, Ramchandran *et al.* [14] analyzed the inter-frame dependency of I, P and B frames in MPEG-2. A trellis-based optimal scheme was proposed to solve the dependent bit allocation problem. Lin *et al.* [15] used a piecewise linear approximation model for the rate and distortion dependency among successive frames for rate control. Both [14] and [15] can obtain good coding performance for MPEG-2, but the computational complexity of the proposed algorithm grows exponentially with the increasing of number of dependent frames. In H.264/SVC, some rate control schemes were proposed based on inter-frame dependency. In [16], Liu *et al.* proposed a GOP-based dependent distortion model of different temporal layers and incorporated it into rate control, which results in better coding performance. But when the length of the GOP is too long, the long coding delay cannot be avoided because it must utilize multi-pass coding process, which is not suitable for real time video coding. Hu *et al.* [17] proposed an enhanced rate control

Manuscript received January 30, 2013; revised May 10, 2013; accepted June 20, 2013. Date of publication July 03, 2013; date of current version November 18, 2013. This work was supported in part by the Major State Basic Research Development Program of China (973 Program, 2009CB320900), National High-tech R&D Program of China (61272386, 61121002 and 61103088), National High-tech R&D Program of China (863 Program, SS2012AA010805), and National Sci-Tech Support Plan (2011BAH08B01). The guest editor coordinating the review of this manuscript and approving it for publication was Prof. Nam Ling.

S. Wang and D. Zhao are with the School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150080, China.

S. Ma, S. Wang, and W. Gao are with the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China (e-mail: swma@pku.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTSP.2013.2272240

for H.264/SVC considering inter-frame dependency. In [18], an efficient bit allocation algorithm using two-dimensional interpolation of rate distortion applications of one-way video was proposed and a pruning algorithm was provided to avoid the complexity increasing exponentially.

The rate control for HEVC has not been thoroughly studied yet. Based on the quadratic R-Q model in [6], a unified R-Q model was proposed in [19] called quadratic pixel-based unified rate-quantization (URQ) model, which is the recommended rate control algorithm in HM6.0. This model considered the new feature that the size of prediction unit varies so the bit allocation must be accordance with the number of pixels. But the R-D performance is not as good as HM anchor without rate control. In [20], a linear $R-\lambda$ model based rate control algorithm was proposed and shows lower bit rate mismatch than the URQ model. The linear $R-\lambda$ model is the recommended rate control algorithm in present HEVC. In our previous work [21], a multi-layer bit allocation based rate control was proposed.

Owing to the new coding tools adopted in HEVC, many new features related to rate control, such as R-Q model, bit allocation and header information prediction, are all different with the previous used features. In this paper, based on Rate-GOP provided in HEVC, a frame level rate control algorithm is proposed by taking into account the inter-frame dependency. The main contributions of this paper can be summarized as the following three aspects.

Firstly, a hierarchical rate control structure is proposed based on Rate-GOP to ensure the high video quality of the key frames. Therefore, the following frames can have a better quality reference frame.

Secondly, considering the inter-frame dependency between a coding frame and its reference frame, an inter-frame dependency based distortion model and an inter-frame dependency based bit rate model are proposed respectively.

Thirdly, based on these models and a mixed Laplacian distribution of residual information, a new $\rho$-domain Rate-GOP based rate control is proposed.

The rest of the paper is organized as follows. Section II introduces the quad-tree coding structure, the specialized Rate-GOP based QP determination and RPS determination in HEVC. The proposed hierarchical rate control structure and problem formulation are also given in this section. Section III analyses the inter-frame dependency relation between a coding frame and its reference frame. Based on this analysis, the inter-frame dependency based distortion model and the inter-frame dependency based bit rate model are proposed. The bit rate model and distortion model for Rate-GOP are presented in Section IV. The proposed hierarchical bit allocation scheme for HEVC is described in Section V. Section VI provides the experimental results. Finally, we conclude the paper in the last section.

## II. Hierarchical Rate Control Structure of Rate-GOP and Problem Formulation

### A. *Quad-Tree Coding Structure and Reference Picture Set*

HEVC is based on traditional hybrid prediction/transform coding framework as described in Fig. 1. Many new coding tools are adopted to improve the coding performance, such as

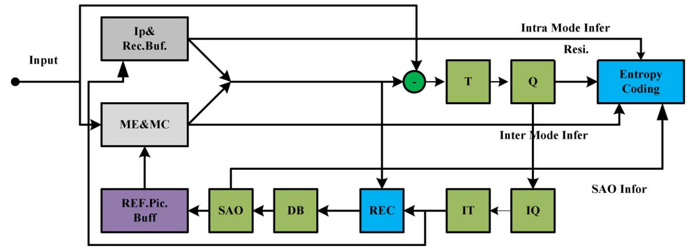


Fig. 1. HEVC coding framework, IP: Intra prediction, T: Transform, Q: Quantization, IT: Inverse Transform, IQ: Inverse Quantization, REC: Reconstruction, DB: Deblock-filter, ME: Motion Estimation, MC: Motion Compensation.

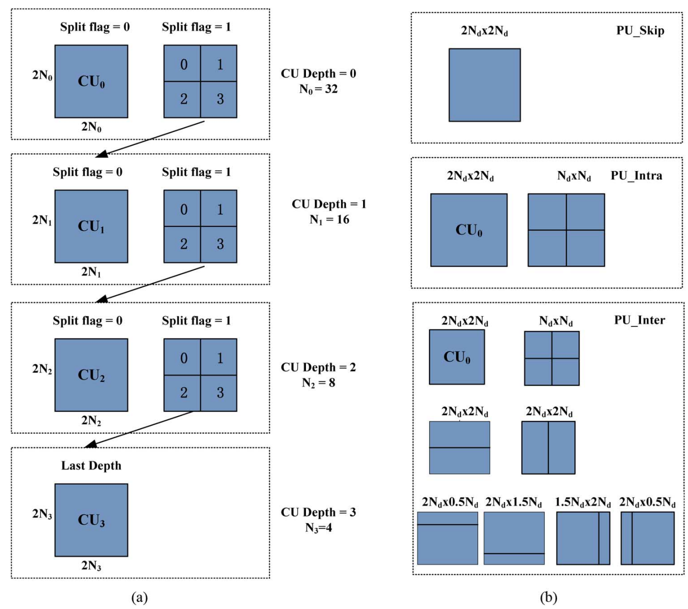


Fig. 2. (a) Recursive CU structure in HM. (LCU size $=64$, maximum hierarchical depth $=4$), (b) PU splitting for skip, intra and inter in HM.

the coding tree unit (CTU) based adaptive quad-tree coding structure.

CTU is the basic coding unit in HEVC which is similar to macro-block in H.264/AVC. In CTU, three new concepts named coding unit (CU), prediction unit (PU) and transform unit (TU) [22] are introduced to specify the basic processing unit of coding, prediction and transform respectively. CU can have various sizes and allows recursive quad-tree splitting. Given the size of CTU and the maximum hierarchical depth, CU can be expressed in a recursive quad-tree representation adapted to the picture content as illustrated in Fig. 2(a). Once the splitting of CU hierarchical tree is finished, the leaf node CUs can be further split into PUs. PU is the basic unit for prediction and it allows multiple different shapes to encode irregular image patterns as shown in Fig. 2(b). The size of PU is limited to that of CU with square or rectangular shape. However, for intra CU and PU splitting, $2\mathrm{N}\times 2\mathrm{N}$ and $\mathrm{N}\times\mathrm{N}$ partion mode are used, and NxN partition mode is allowed only when the corresponding CU size is equal to the minimum CU size. TU is defined to represent the basic unit for transform. For inter mode, the size of TU is independent with the size of PU; while for intra mode, the size of TU cannot exceed the size of PU. The size of TU cannot exceed the size of CU for both intra and inter mode. This highly flexible coding structure

TABLE I
REFERENCE PICTURE SET (RPS) IN HEVC FOR LD CONFIGURATION

| GOPId | Delta POC | | | |
|---|---|---|---|---|
| 0 | -1 | -5 | -9 | -13 |
| 1 | -1 | -2 | -6 | -10 |
| 2 | -1 | -3 | -7 | -11 |
| 3 | -1 | -4 | -8 | -12 |

TABLE II
REFERENCE PICTURE SET (RPS) IN HEVC FOR RA CONFIGURATION

| GOPId | Delta POC | | | |
|---|---|---|---|---|
| 0 | -8 | -10 | -12 | -16 |
| 1 | -4 | -6 | 4 | |
| 2 | -2 | -4 | 2 | 6 |
| 3 | -1 | -1 | 3 | 7 |
| 4 | -1 | -3 | 1 | 5 |
| 5 | -2 | -4 | -6 | -2 |
| 6 | -1 | -5 | 1 | 3 |
| 7 | -1 | -3 | -7 | 1 |

provides great flexibility to improve the coding efficiency. But it also makes the distribution of residual information more complex. The distribution functions used in the previous coding standards may not be suitable for HEVC.

Besides the quad-tree coding structure, Rate-GOP [23] is also employed to improve the R-D performance. Rate-GOP is a coding group including some successive frames with the default number as four in Low Delay (LD) configuration and eight in random access (RA) configuration.

In Rate-GOP, the determinations of QP and RPS are both different from that of in the previous coding standard.

For QP determination, QP of each frame has fixed difference with QP of I frame: $\mathrm{QP_I}$. Eqs. (1) and (2) illustrate the QP determination for LD and RA configurations respectively.

$$QP = \begin{cases} QP_I + 1 & if(POC\%4 == 0) \\ QP_I + 2 & if(POC\%4 == 2) \\ QP_I + 3 & else \end{cases} \tag{1}$$

$$QP = \begin{cases} QP_I + 1 & if(POC\%8 == 0) \\ QP_I + 2 & if(POC\%8 == 4) \\ QP_I + 3 & if(POC\%8 == 2) \\ QP_I + 4 & else \end{cases}. \tag{2}$$

$POC$ denotes Picture Order Count and represents an output order of the pictures in the video stream.

For RPS determination, it is not composed of some successive decoded frames but of one frame with the nearest temporal distance and three frames with lowest QP in Decoded Picture Buffer (DPB). For the coding frame, the determination of RPS can be presented by the following two steps. First, the $GOPId$ is calculated by

$$GOPId = (POC - 1)\%4. \tag{3}$$

Then RPS for the current coding frame is determined by the corresponding $Delta\ POC$ set depending on the $GOPId$. $Delta\ POC$ is the difference between $POC$ of the current frame and $POC$ of previous coded frame. Tables I and II present the $Delta\ POC$ set of the reference frame for different $GOPId$ in LD and RA configurations respectively.

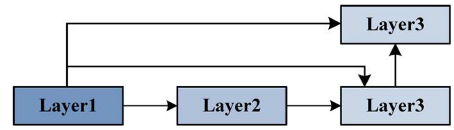


Fig. 3. Hierarchical rate control structure of Rate-GOP for LD.

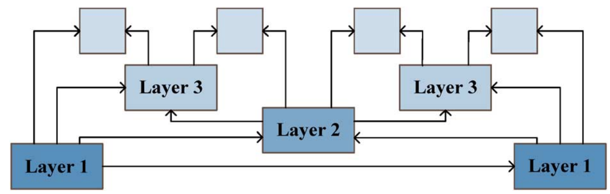


Fig. 4. Hierarchical rate control structure of Rate-GOP for RA.

Such QP determination and RPS determination in Rate-GOP take more consideration of inter-frame dependency and thus improves encoding efficiency significantly.

## B. Proposed Hierarchical Rate Control Structure

From Tables I and II and (1) and (2), it can be concluded that different frame in a Rate-GOP has different influence on the following frames. Generally, the frame with $\mathrm{QP_I}+1$ with relatively high reconstructed video quality will be referred more than once and thus plays a crucial role, which is called the key frame in our paper. Other frames in the same Rate-GOP are referenced less with less importance than the key frame. Based on this observation, taking Rate-GOP as a rate control unit, the hierarchical rate control structures of Rate-GOP for LD and RA are proposed respectively, which are shown in Fig. 3 and Fig. 4.

For LD configuration, the composition of Rate-GOP is made up of four frames including one key frame and its following three frames. The frames are divided into three layers according to their importance. The first layer consists of the first frame, as it plays a crucial role and will be referred by the following frame more times. The second layer is made up of the second frame, as it will mainly refer to the first frame. The remaining two frames belong to the third layer.

For RA configuration, the proposed hierarchical rate control structure is similar to the hierarchical depth in HEVC. The frame with hierarchical depth $= 0$ composes of the first layer. The second layer is made up of the frame with hierarchical depth $= 1$. The frames with hierarchical depth $= 2$ compose of the third layer. The other four frames with the maximum hierarchical depth will not be used as a reference frame, so these frames are not included in the proposed hierarchical structure.

## C. Problem Formulation

Based on the proposed hierarchical rate control structure, the constrained optimization problem for the bit allocation of a Rate-GOP can be stated as follows:

$$q^* = (q_1, \ldots, q_k) = \arg\min_{q_i \in Q} D_{RG} \quad s.t.\ R_{RG} \le R_T, \tag{4}$$

where $q^* = (q_1, \ldots, q_k)$ denotes the optimal QP vector in the set of all possible QP candidates: $Q$ and $k$ is the number of frame in a Rate-GOP. $D_{RG}$ and $R_{RG}$ denote the total distortion and the bit rate of the Rate-GOP respectively. $R_T$ is the constrained channel bandwidth. The constrained optimization problem can be solved with Lagrangian multiplier optimization method by incorporating the constraint function into the objective function with a Lagrange multiplier $\lambda$ [24].

$$q^* = \arg\min_{q_i \in Q} J(q^*, \lambda) \tag{5}$$

$$J(q^*, \lambda) = \sum_{k=1}^{k} D_k(q_1, \ldots, q_k) + \lambda \cdot \left( \sum_{k=1}^{k} R_k(q_1, \ldots, q_k) - R_T \right). \tag{6}$$

One optimal solution to this optimization problem is to execute a full search over all possible combinations of all the quantization steps. However, the computational complexity grows exponentially with the increasing of number of frames in a Rate-GOP. This optimization problem is addressed in this paper by analyzing the inter-frame dependency of distortion and inter-frame dependency of bit rate in a Rate-GOP.

## III. Inter-Frame Dependency Based Distortion Model and Rate Model

Inter-frame prediction results in the distortion dependency and bit rate dependency between a coding frame and its reference frames. Generally, a better reference frame leads to better coding performance in predictive coding. However the dependency makes the rate distortion optimization (RDO) based optimization problems more complicated. The influence of the dependency for the final coding performance cannot be addressed well since the problem is difficult to formulate and solve. Thus, the improvement in coding efficiency is not so significant. In this section, we will analyze the distortion dependency and bit rate dependency in a Rate-GOP under the assumption that each Rate-GOP is independent with others.

### A. Inter-Frame Dependency Based Distortion Model

The distortion of the coding frame has great dependency with its reference frames. This kind of distortion dependency is one of the main factors affecting the bit allocation scheme for rate control.

Experiments are conducted to investigate the distortion dependency between the coding frame and its reference frame. Fig. 5 shows the distortion dependency of *Rasehorses* and *BascketballPass*. Only the nearest frame in DPB, which has the minimum POC difference with the POC of the current coding frame, is selected as the reference frame in our experiment. This is reasonable for actual video coding even with more reference frames since the distortion of the coding frame mainly depends on its nearest reference frame. The QP of coding frame is fixed, while the QP of reference frame varies. The $QP$ of coding frame is set to 20, 22, 25 and 27 respectively while the $QP$ of reference frame changes from 10 to 20, 10 to 22, 10 to 25 and 10 to 27 respectively. The distortion is measured by mean square error (MSE).

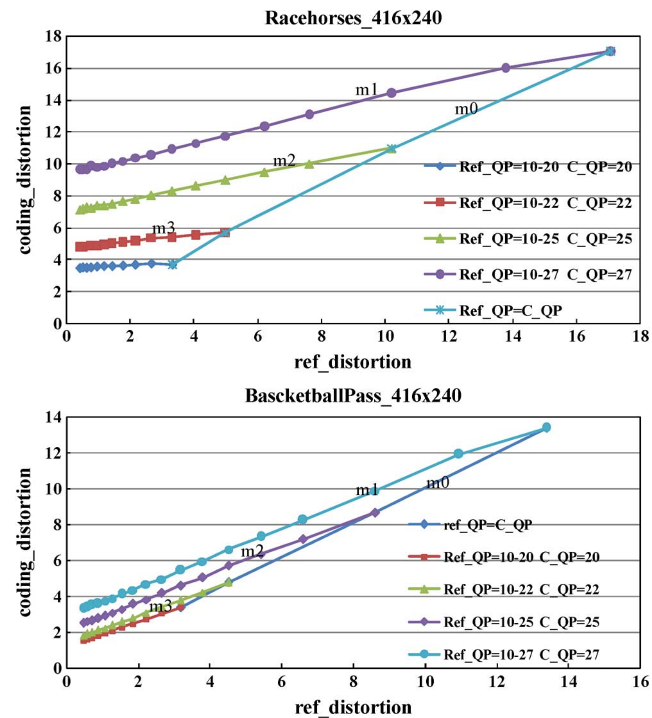


Fig. 5. Distortion dependency curve between reference frame and coding frame with fixed QP of coding frame and varied QP of reference frame.

It can be observed that the distortion of the coding frame has an approximate linear correlation with the distortion of the reference frame and then the relationship can be described as follows:

$$\begin{aligned} D(q_0, q_1) &= m_0 \times D_{ref}(q_0) - m_1 \times (D_{ref}(q_0) - D_{ref}(q_1)) \\ &= m_1 \times D_{ref}(q_1) + (m_0 - m_1) \times D_{ref}(q_0) \\ &= \zeta_1 \cdot D_{ref}(q_1) + \zeta_2 \cdot D_{ref}(q_0), \end{aligned} \tag{7}$$

where $q_0$ and $q_1$ are the QPs of the reference frame and coding frame. $D_{ref}(q_0)$ denotes the distortion of the reference frame. $D(q_0, q_1)$ denotes the distortion of the coding frame. $m_0$ and $m_1$ are the slopes of the distortion dependency curves reflecting the dependency under different QP. Based on the distortion model, the distortion of the coding frame can be simply represented by its reference frame.

### B. Inter-Frame Dependency Based Rate Model

The bit rate dependency between the coding frame and its reference frame is also investigated aiming at a deeper exploration for R-D relations. Fig. 6 shows the rate dependency between reference frame and coding frame. The reference frame selection is the same as the previous experiment. The $QP$ value of the coding frame is fixed and the amount of output bits is recorded by changing $QP$ of the reference frame. The $QP$ of coding frame is set as 20, 22, 25 and 27 respectively while the $QP$ of reference frame changes from 10 to 20, 10 to 22, 10 to 25 and 10 to 27 respectively.

Unlike the distortion dependency, it can be observed that the number of coded bits of reference frame has limited influence on the number of bits of the coding frame. The bits of reference frame vary in a large extent while the bits of coding frame vary in a small domain. The observation indicates that the bit rate

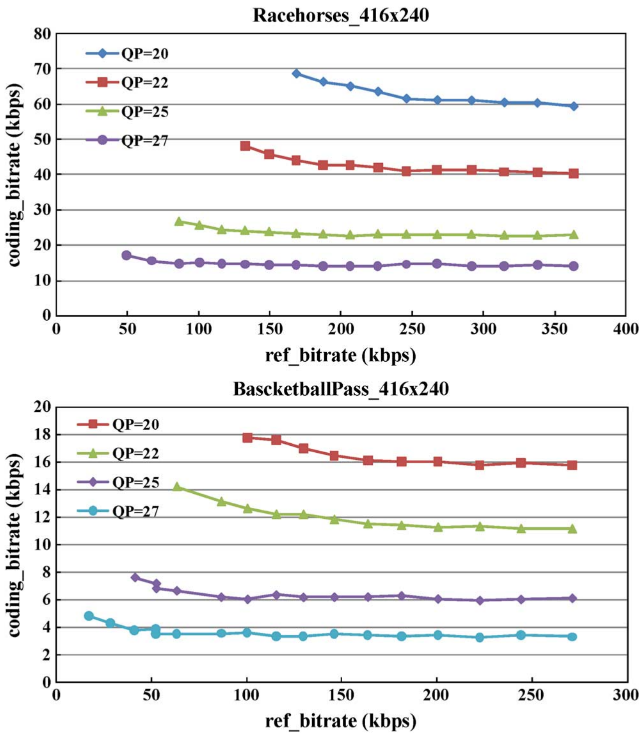


Fig. 6. Rate dependency curve between reference frame and coding frame with fixed QP of reference frame and changed QP of coding frame.

of coding frame is mainly determined by its own QP. In [16], the bit rate of coding frame is represented as a piecewise linear model for H.264/SVC.

$$R(q_0, q_1) = \begin{cases} r \cdot R(q_0) + (s - r) \cdot R\left(\frac{q_1}{2}\right) & q_1 \le 2q_0 \\ s \cdot R\left(\frac{q_1}{2}\right) & q_1 > 2q_0 \end{cases}. \quad (8)$$

However, for rate control, it always limits the QP variance between adjacent frames in a small domain, e.g. 4 or even smaller, to maintain the smoothness of video quality. So the bit rate can be represented as a function of its own QP.

$$R(q_0, q_1) \approx R(q_1). \quad (9)$$

## IV. R-Q Model and D-Q Model for HEVC

In this section, the Rate-GOP based R-Q model and D-Q model for HEVC are presented. Since the derivation of the two models for RA is similar with that of LD. Only the derivation of for LD is described.

### *A. Quadratic ρ-Domain Based R-Q Model for HEVC*

The most crucial step for rate control is to compute a suitable QP which heavily depends on the R-Q model to meet the target bits. Generally speaking, R-Q model has great correlation with the distribution of the residual information. It was stated in [25] that a single Laplacian distribution is not accurate enough to capture the distribution due to the quad-tree coding structure. Thus a mixed Laplacian distribution is used to represent the distribution as follows.

$$f(x) = \sum_{j=0}^{k} f_j(x) = \sum_{j=0}^{k} \frac{1}{2}\lambda_j e^{-\lambda_j |x|}, \quad (10)$$

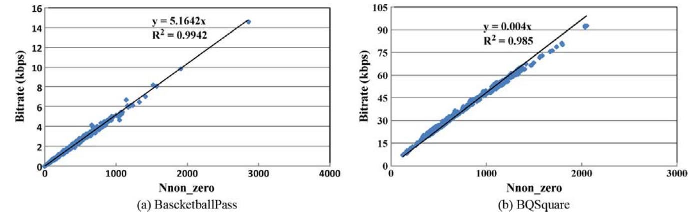


Fig. 7. The linear relation between the number of non-zero transformed coefficients and the bits in $\rho$-domain. (a) BascketballPass; (b) BQSquare.

where $f_j(x)$ denotes the distribution function of transformed coefficients of CUs in $j$-*th* depth and $k$ indicates the maximum value of the depth which is set to 3 in default configuration. It has been proved in [12] that an accurate linear relation in $\rho$-domain between the bits and the number of non-zero transformed coefficients holds firmly via the theoretical derivation for a Laplacian distribution as follows:

$$R_i = \theta_i' \cdot (1 - \rho_i) = \theta_i \cdot N_{i,non\_zero}, \quad (11)$$

where $\rho_i$ and $R_i$ denote the percentage of zero transformed coefficients and the bit rate of coding frame $i$. $N_{i,non\_zero}$ indicates the total number of non-zero transformed coefficients of the frame. $\theta_i'$ and $\theta_i$ are two constants and their relation can be represented as $\theta_i' = \theta_i \cdot N$, where $N$ is the number of pixels in a frame. Fig. 7 shows the linear relation in $\rho$-domain for the sequence *BascketballPass* and *BQSquare* in HEVC. It can be seen that an accurate linear relation holds firmly with different slopes and zero intercept.

According to the quantization in HEVC, for a certain $j$-*th* depth, the distortion $D_j(q)$ can be calculated by (12) and the percentage of zero coefficients $\rho_j(q)$ can be computed by (13).

$$D_j(q) = 2\int_0^{\delta q} f_j(x)dx + 2\sum_{l=1}^{\infty} \int_{(l-\delta)q}^{(l+\delta)q} f_j(x)|x - lq|dx \quad (12)$$

$$\rho_j(q) = \int_{-\delta q}^{\delta q} f_j(x)dx = \int_{-\delta q}^{\delta q} \frac{\lambda_j}{2} e^{-\lambda_j |x|}dx = 1 - e^{-\delta\lambda_j q}. \quad (13)$$

As illustrated in [12], for the $j$-*th* depth in frame $i$ with a Laplacian distribution, we can obtain the linear relation as follows:

$$R_{i,j} = \theta_j' \cdot (1 - \rho_j), \quad (14)$$

where $\rho_j$ and $R_{i,j}$ denote the percentage of zero transformed coefficients and the bit rate of the $j$-*th* depth. Combined with (13), the bit rate of frame $i$ can be represented as:

$$R_i = \sum_{j=0}^{k} \theta_j' \cdot (1 - \rho_j) = \sum_{j=0}^{k} \theta_j' \cdot e^{-\delta\lambda_j q_i}. \quad (15)$$

For a single Laplacian distribution, a mapping scheme between $\rho$ and QP is provided in [12] to determine a suitable QP to meet the target bits. It may be not suitable to utilize the scheme to HEVC due to the complicated distribution as indicated in (10). Consequently, it is reasonable to find a suitable mapping scheme between the number of non-zero transformed coefficients and the QP.

Assume $N_{i,j,non\_zero}$ and $N_{i,j}$ denote the number of non-zero transformed coefficients and the number of pixels of $j$-$th$ depth in frame $i$ respectively. Combine with (13), we can get

$$N_{i,j,non\_zero} = N_{i,j} \cdot (1-\rho_j) = N_{i,j} \cdot e^{-\delta\lambda_j q_i}. \quad (16)$$

Thus the total number of non-zero transformed coefficients of frame $i$ can be obtained by

$$N_{i,non\_zero} = \sum_{j=0}^{k} N_{i,j,non\_zero} = \sum_{j=0}^{k} N_{i,j} \cdot e^{-\delta\lambda_j q_i}, \quad (17)$$

where $k$ indicates the maximum value of the depth.

By a Taylor expansion, the non-zero number of transformed coefficients can be represented by the quadratic function of quantization step.

$$N_{i,non\_zero} = N + a_i \cdot q_i^2 + b_i \cdot q_i, \quad (18)$$

where $N$ denotes the total number of pixel of the frame and the two parameters can be represented as:

$$a_i = \sum_{j=0}^{3} (\delta_i\lambda_j)^2, b_i = -\sum_{j=0}^{3} \delta_i\lambda_j. \quad (19)$$

Combining with (11), we can get the R-Q model for frame $i$:

$$R_i = \theta_i \cdot (N + a_i \cdot q_i^2 + b_i \cdot q_i). \quad (20)$$

Finally, the R-Q model for Rate-GOP can be obtained by

$$R_{RG} = \sum_{i=0}^{3} \theta_i \cdot N_{i,non\_zero} = \sum_{i=0}^{3} \theta_i \cdot \left(N + a_i \cdot q_i^2 + b_i \cdot q_i\right). \quad (21)$$

In actual coding process, $a_i$ and $b_i$ are obtained via linear regression scheme to avoid the complicated calculation in (19).

### B. D-Q Model of Rate-GOP

As illustrated in the above subsection, the linear relation between the bit rate and non-zero transformed coefficients still holds firmly in HEVC under the mixed Laplacian distribution. Consequently, for the first layer, the distortion can be represented by the following as stated in [12].

$$D_0(\rho_0) = \sigma_0^2 \cdot e^{-\alpha_0(1-\rho_0)} = \sigma_0^2 \cdot e^{-\alpha_0(1-f(q_0))}, \quad (22)$$

where $\alpha_0$ is a constant parameter that normally ranges from 10 to 20, $\sigma_0^2$ is the variance of the picture related to video content. $f(q_0)$ is a function of the zero transformed coefficients.

For the second layer and third layer, we take into account the inter-frame dependency in (7) to model the distortion as follows.

$$\begin{aligned} D_1(q_0,q_1) &= \alpha_1 \times D_0(q_1) + \alpha_2 \times D_0(q_0) \quad (23)\\ D_2'(q_1,q_2) &= \beta_1 \times D_1(q_2) + \beta_2 \times D_1(q_1)\\ &= \beta_1 \times D_0(q_2) + \beta_2 \times D_0(q_1) \quad (24)\\ D_2''(q_2,q_3) &= \gamma_1 \times D_2'(q_3) + \gamma_2 \times D_2'(q_2)\\ &= \gamma_1 \times D_0(q_3) + \gamma_2 \times D_0(q_2) \quad (25)\end{aligned}$$

Subsequently, the distortion of a Rate-GOP can be presented as follows:

$$\begin{aligned} D_{RG} &= D_0(q_0) + D_1(q_0,q_1) + D_2'(q_1,q_2) + D_2''(q_2,q_3)\\ &= D_0(q_0) + [\alpha_1 D_0(q_1) + \alpha_2 D_0(q_0)]\\ &\quad + [\beta_1 D_0(q_2) + \beta_2 D_0(q_1)]\\ &\quad + [\gamma_1 D_0(q_3) + \gamma_2 D_0(q_2)]\\ &= \xi_0 \cdot D_0(q_0) + \xi_1 \cdot D_0(q_1) + \xi_2 \cdot D_0(q_2)\\ &\quad + \xi_3 \cdot D_0(q_3)\\ &= \sum_{i=0}^{3} \xi_i \cdot D_0(q_i) = \sum_{i=0}^{3} \xi_i \cdot \sigma_0^2 \cdot e^{-\alpha_0(1-f(q_i))}\\ &= \sigma_0^2 \cdot \sum_{i=0}^{3} \xi_i \cdot e^{-\alpha_0(1-f(q_i))}. \quad (26)\end{aligned}$$

## V. Proposed Rate Control Algorithm Based on R-D Optimization

In this section, we will provide the proposed rate control algorithm from the following parts: Rate-GOP level bit allocation, optimal hierarchical bit allocation for frames in a Rate-GOP and header bits prediction. The ultimate task is to obtain accurate QP to achieve better R-D performance and lower bit rate mismatch.

### A. Rate-GOP Level Bit Allocation

For the bit allocation at the Rate-GOP level, the target bits for a Rate-GOP are allocated according to the bandwidth after coding the previous Rate-GOPs, which can be calculated as

$$T_{RateGop} = N_{RateGop} \times R_{remaining}/F_{num}, \quad (27)$$

where $T_{RateGop}$ denotes the target bits for a Rate-GOP, $R_{remaining}$ is the remaining bits after encoding the previous Rate-GOPs, $N_{RateGop}$ indicates the number of the frames in a Rate-GOP and $F_{num}$ denotes the remaining frames to be coded.

Considering the temporal correlation between the adjacent Rate-GOPs and in order to get a smoother bit rate, the allocated bits for a Rate-GOP is further modified as

$$T'_{RateGop} = \alpha \times T_{RateGop} + (1-\alpha) \times T_{RateGop\_pre}, \quad (28)$$

where $T_{RateGop\_pre}$ denotes the actual consumed bits of the previous Rate-GOP and $\alpha$ is an empirical parameter, which is set to 0.5 in our experiments.

After completing the Rate-GOP level bit allocation, the next crucial step is to allocate target bits for each frame in one Rate-GOP. And the target bits will be utilized in the calculation of QP to meet the target bits.

### B. Optimal Hierarchical Bit Allocation

Based on the proposed R-Q model, inter-frame dependency model of distortion and inter-frame dependency model of bit rate, in this subsection we will provide an optimal bit allocation scheme for the proposed hierarchical structure of a Rate-GOP.

For a Rate-GOP, given the bit budget $R_T$ allocated as in the previous subsection, the optimization problem is to choose appropriate QPs to minimize the total distortion of the Rate-GOP as the following.

$$\min D_{RG}(q^*) \quad s.t.\ R_{RG}(q^*) \leq R_T. \tag{29}$$

By Lagrangian multiplier $\lambda$, the optimization problem can be expressed as a problem to minimize the R-D cost $J$ as follows.

$$q^* = \arg\min_{q_i \in Q} J(q^*, \lambda)\,. \tag{30}$$

Combining the rate model in (21) and the distortion model in (26), (30) can be rewritten as

$$\begin{aligned} q^* &= \arg\min_{q_i \in Q} J(q^*, \lambda) \\ &= \arg\min \left(D_{RG}(q^*) + \lambda\left(R_{RG}(q^*) - R_T\right)\right) \\ &= \arg\min \left\{ \sigma_0^2 \cdot \sum_{i=0}^{3} \xi_i \cdot e^{-\alpha_0(1-f(q_i))} \right. \\ &\qquad \left. + \lambda\left(\sum_{i=0}^{3} \theta_i \cdot (N + a_i \cdot q_i^2 + b_i \cdot q_i) - R_T\right)\right\}. \end{aligned} \tag{31}$$

By computing partial derivation with respect to $q_i$,

$$\frac{\partial D_{RG}(q^*)}{\partial q_i} = -\lambda \frac{\partial R_{RG}(q^*)}{\partial q_i} \tag{32}$$

$$\sigma_0^2 \cdot \xi_i \cdot e^{-\alpha_0(1-f(q_i))} \cdot \alpha_0 f'(q_i) = -\lambda\theta_i \cdot (2a_i q_i + b_i). \tag{33}$$

From (18), we can get

$$\rho_i = f(q_i) = \frac{(N - N_{i,non\_zero})}{N} = -\left(a_i \cdot q_i^2 + b_i \cdot q_i\right)/N. \tag{34}$$

Combining (20) and (34), we can further obtain

$$1 - f(q_i) = \frac{R_i}{(N\theta_i)}. \tag{35}$$

By a Taylor expansion,

$$e^{-\alpha_0(1-f(q_i))} \approx 1 - \alpha_0\left(1 - f(q_i)\right) = 1 - \alpha_0\left(\frac{R_i}{(N\theta_i)}\right). \tag{36}$$

Taking (36) into (33), the bit rate for $i$-th frame can be obtained by

$$R_i = f\left(\alpha_0, \sigma_0^2, \xi_i, \theta_i\right) = \frac{\theta_i N}{\alpha_0} - \frac{\lambda N^2 \theta_i^2}{\sigma_0^2 \alpha_0^2 \xi_i}. \tag{37}$$

Once $\theta_i$ is determined, $R_i$ can be obtained. Due to the correlation among adjacent frames, the average value of $\theta_i$ of previous coded frames is used to predict $\theta_i$ of the coding frame by

$$\theta_i = \frac{1}{n}\sum_{k=1}^{n} \theta_k = \frac{1}{n}\sum_{k=1}^{n} \frac{R_k}{N_{k,non\_zero}}. \tag{38}$$

By using the predicted $\theta_i$ to calculate $R_i$ may not be accurate enough, thus a weighting factor $\omega_i$ can be derived as

$$\omega_i = \frac{R_i}{R_0} = \frac{f\left(\alpha_0, \sigma_0^2, \xi_i, \theta_i\right)}{f\left(\alpha_0, \sigma_0^2, \xi_0, \theta_0\right)}, \tag{39}$$

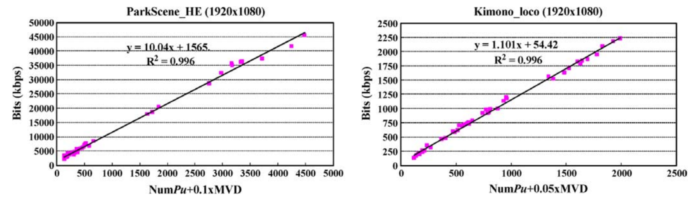


Fig. 8. The relationship between $R_h$ and PU number $Num_{pu}$ and Motion Vector Difference $(MVD)$.

where $R_i$ is the number of bits consumed by the frames in the corresponding layer and $R_0$ is the number of bits consumed by the key frame.

By the given target bits of a Rate-GOP and $\omega_i$, $R_i$ can be obtained. Using $R_i$ and $\theta_i$, the final QP can be calculated by proposed R-Q model in (20).

To avoid the complicated calculation of $\omega_i$ in (39), $\omega_i$ is estimated from that of the previous GOPs. Given a set of initial values for the weighting factors, the bits for $j$-th layer in $i$-th Rate-GOP is allocated as

$$T_{i,j} = \alpha_{i,j} \times T'_{RateGop} \quad j = 1, 2, 3, \tag{40}$$

where $T_{i,j}$ and $\alpha_{i,j}$ denote the target bits and weighting factors of $j$-th layer of $i$-th Rate-GOP respectively. $T'_{RateGop}$ is the given bit rate for the Rate-GOP. The initial values of the parameters $\alpha_{i,j}$ are set to 0.5, 0.3 and 0.2. The bit rates for the two frames in the third layer are allocated averagely. After encoding one Rate-GOP, if $|R_{actual} - T'_{RateGop}| \leq bitrate/framerate$, then the weighting factors are calculated and utilized for the next Rate-GOP. Otherwise, QP is adjusted based on the QP and bit rate of the previous Rate-GOP.

For $j$-th layer in $k$-th Rate-GOP,

$$\alpha_{k,j} = (\alpha_{r,j} + \alpha_{s,j} + \alpha_{t,j})/3, \quad j = 1, 2, 3, \tag{41}$$

where $\alpha_{r,j}, \alpha_{s,j}, \alpha_{t,j}$ are the weighting factors of $j$-th layer in the three nearest Rate-GOPs to $k$-th Rate-GOP, satisfying $|R_{actual} - T'_{RateGop}| \leq bitrate/framerate.$

After the determination of the target bits, $\alpha_{i,j}$ and $\theta_i$, the final QP can be calculated by proposed R-Q model in (20) to meet the target bits.

It should be noted that since the frames with the maximum hierarchical depth in RA configuration will not be a reference frame, the QP of these frames is obtained by

$$QP = \max(QP_1, QP_2), \tag{42}$$

where $QP_1$ and $QP_2$ are the QPs of the adjacent frames.

The variation of QP in a Rate-GOP is limited in 4 for LD and 8 for RA.

### C. Header Bits Prediction

Header bits prediction is also a component of bit allocation. In [10], the header bits for inter coding macro-block are modeled with the number of nonzero motion vector elements and motion vectors. Inspired by [10], in the experiments it was found that the header bits has strong linear relationship with the number of PU and the value of MVD, as shown in Fig. 8. From the curves, it can be seen that the header bits $R_h$ can be modeled as

$$R_h = \alpha \times (Num_{pu} + \beta \times MVD) + \gamma, \tag{43}$$

where $R_h$ indicates the header bits and $Num_{pu}$ is the number of PU in one frame. The parameter $\gamma$ is usually small which can be set to zero and other parameters can be updated adaptively by linear regression method.

### D. Proposed Rate-GOP Based Rate Control Scheme

The proposed Rate-GOP based rate control scheme can be summarized in Algorithm 1.

**Algorithm 1. Summary of the proposed rate control scheme**

**begin**

Allocate target bits for $i$-th RateGOP by (28);

**begin**

For each frame in the $i$-th RateGOP

1) Calculate $\alpha_{i,j}$ by (41);

2) Allocate target bits for $j$-th frame by (40);

3) Calculate $\theta_i$ by (38);

4) Calculate QP by the R-Q model in (20);

5) Encoding with the QP;

**end**

Updating the parameters;

**end**

## VI. Experimental Results

To evaluate the performance of the proposed algorithm, the experiments are conducted on HEVC test model HM8.0. The test sequences and 6 testing configurations are detailed in [26].

### A. Control Accuracy

The accuracy of bit mismatch is investigated in terms of mismatch error as follows.

$$M = \frac{|R_{target} - R_{actual}|}{R_{target}} \times 100\%, \tag{44}$$

where $R_{target}$ and $R_{actual}$ are the number of target bits and the actual output bits for a coding frame respectively.

Table III shows the bit rate mismatch ratio comparisons. It can be seen that the proposed algorithm generates smaller mismatch between target bits and actual output bits.

Usually, the larger the size of the basic unit, the better video quality the rate control algorithm can achieve, but at the cost of degradation of bit rate accuracy. The proposed frame level rate control can also obtain better accuracy of bit rate.

### B. R-D Performance

Firstly, experiments are conducted with the proposed bit allocation scheme but without using the proposed R-Q model. QP is obtained by the quadratic model as in H.264/AVC. The R-D performances in terms of BD-PSNR and BD-Rate [27] results are presented for comparison with URQ algorithm as illustrated in the Table IV. The positive BD-PSNR or the negative BD-Rate denotes the corresponding algorithm achieves better R-D performance. It can be seen that the proposed bit allocation can significantly improve the rate control performance. It is because the proposed bit allocation ensures the high quality of the frames in low layer. Therefore the frame in high layer can have better reference frame. Thus better R-D performance can be achieved.

Secondly, experiments are implemented with both the bit allocation scheme and the proposed R-Q model. As illustrated in Tables V–VII, the R-D performances in terms of BD-PSNR and BD-Rate results are presented for comparison with URQ algorithm and $\mathrm{R}-\lambda$ algorithm. From the tables, it can be observed the proposed algorithm shows much better R-D performance than the other two algorithms in a significant

TABLE III
Bit Rate Mismatch Comparison for the Proposed Algorithm With $\mathrm{R}-\lambda$ Model [20]

| Sequences | LB-main | | LB-HE | |
|---|---|---|---|---|
| | R-λ[20] | Proposed | R-λ[20] | Proposed |
| ClassB | 0.28% | 0.51% | 0.94% | 0.74% |
| ClassC | 0.21% | 0.27% | 0.26% | 0.28% |
| ClassD | 0.16% | 0.95% | 0.29% | 0.81% |
| ClassE | 0.41% | 0.31% | 1.56% | 0.42% |
| Average | 0.27% | 0.51% | 0.94% | 0.74% |
| Sequences | LP-main | | LP-HE | |
| | R-λ[20] | Proposed | R-λ[20] | Proposed |
| ClassB | 0.21% | 0.59% | 1.01% | 0.64% |
| ClassC | 0.46% | 0.29% | 0.30% | 0.34% |
| ClassD | 0.16% | 0.92% | 0.36% | 0.96% |
| ClassE | 0.32% | 0.28% | 1.65% | 0.37% |
| Average | 0.29% | 0.52% | 0.83% | 0.58% |
| Sequences | RA-main | | RA-HE | |
| | R-λ[20] | Proposed | R-λ[20] | Proposed |
| ClassB | 1.08% | 1.73% | 1.39% | 1.34% |
| ClassC | 0.65% | 1.58% | 0.44% | 1.49% |
| ClassD | 0.72% | 2.92% | 1.86% | 0.63% |
| Average | 0.82% | 2.08% | 1.23% | 1.15% |

TABLE IV
Performance Comparison for the Proposed Bit Allocation Scheme With URQ Model [19]

| Sequences | Gain over URQ[19] | | | |
|---|---|---|---|---|
| | LP-HE | | LB-HE | |
| | BD-Rate | BD-PSNR | BD-Rate | BD-PSNR |
| ClassB | -16.19% | 0.43dB | -36.70% | 1.05 dB |
| ClassC | -9.80% | 0.40 dB | -15.80% | 0.66 dB |
| ClassD | -11.25% | 0.48 dB | -16.26% | 0.68 dB |
| ClassE | -10.87% | 0.34 dB | -18.87% | 0.68 dB |
| Average | -12.03% | 0.41 dB | -21.91% | 0.77 dB |
| Sequences | LP-main | | LB-main | |
| | BD-Rate | BD-PSNR | BD-Rate | BD-PSNR |
| ClassB | -18.93% | 0.52dB | -31.28% | 0.95 dB |
| ClassC | -5.76% | 0.22 dB | -15.12% | 0.63 dB |
| ClassD | -11.94% | 0.47 dB | -15.27% | 0.63 dB |
| ClassE | -8.67% | 0.28 dB | -16.08% | 0.62 dB |
| Average | -11.33% | 0.37 dB | -19.44% | 0.71 dB |
| Sequences | RA-main | | RA-HE | |
| | BD-Rate | BD-PSNR | BD-Rate | BD-PSNR |
| ClassB | -15.9% | 0.56dB | -10.2% | 0.35dB |
| ClassC | -13.2% | 0.48dB | -7.60% | 0.28dB |
| ClassD | -10.9% | 0.36dB | -1.50% | 0.07dB |
| Average | -13.3% | 0.47dB | -6.43% | 0.23dB |

TABLE V
PERFORMANCE COMPARISON FOR THE PROPOSED ALGORITHM WITH URQ MODEL [19] AND $\mathrm{R}-\lambda$ MODEL [20]

| Sequences | LB-HE | | | |
|---|---|---|---|---|
| | Proposed vs.URQ[19] | | Proposed vs. R-λ[20] | |
| | BD-Rate | BD-PSNR | BD-Rate | BD-PSNR |
| ClassB | -52.6% | 1.89dB | -1.4% | 0.03dB |
| ClassC | -26.1% | 1.20dB | -3.2% | 0.15dB |
| ClassD | -12.4% | 0.54dB | -7.6% | 0.25dB |
| ClassE | -32.3% | 1.16dB | -1.3% | 0.05dB |
| Average | -30.85% | 1.20dB | -3.38% | 0.12dB |
| Sequences | LP-HE | | | |
| | Proposed vs.URQ[19] | | Proposed vs. R-λ[20] | |
| | BD-Rate | BD-PSNR | BD-Rate | BD-PSNR |
| ClassB | -52.5% | 1.95dB | -3.7% | 0.11dB |
| ClassC | -25.9% | 1.15dB | -2.3% | 0.12dB |
| ClassD | -12.7% | 0.53dB | -6.3% | 0.21dB |
| ClassE | -31.4% | 1.06dB | -1.0% | 0.02dB |
| Average | -30.63% | 1.17dB | -3.33% | 0.12dB |

TABLE VI
PERFORMANCE COMPARISON FOR THE PROPOSED ALGORITHM WITH URQ MODEL [19] AND $\mathrm{R}-\lambda$ MODEL [20]

| Sequences | LB-main | | | |
|---|---|---|---|---|
| | Proposed vs.URQ[19] | | Proposed vs. R-λ[20] | |
| | BD-Rate | BD-PSNR | BD-Rate | BD-PSNR |
| ClassB | -23.64% | 0.77dB | -1.6% | 0.05dB |
| ClassC | -13.90% | 0.57dB | -3.7% | 0.11dB |
| ClassD | -12.09% | 0.50dB | -5.9% | 0.23dB |
| ClassE | -23.21% | 0.78dB | -1.3% | 0.05dB |
| Average | -18.21% | 0.66dB | -3.13% | 0.11dB |
| Sequences | LP-main | | | |
| | Proposed vs.URQ[19] | | Proposed vs. R-λ[20] | |
| | BD-Rate | BD-PSNR | BD-Rate | BD-PSNR |
| ClassB | -24.5% | 0.76dB | -4.60% | 0.11dB |
| ClassC | -14.2% | 0.58dB | -3.60% | 0.12dB |
| ClassD | -13.10% | 0.52dB | -6.20% | 0.21dB |
| ClassE | -24.30% | 0.75dB | -1.60% | 0.05dB |
| Average | -19.03% | 0.65dB | -4.00% | 0.13dB |

TABLE VII
PERFORMANCE COMPARISON FOR THE PROPOSED ALGORITHM WITH URQ MODEL [19] AND $\mathrm{R}-\lambda$ MODEL [20]

| Sequences | RA-main | | | |
|---|---|---|---|---|
| | Proposed vs.URQ[19] | | Proposed vs. R-λ[20] | |
| | BD-Rate | BD-PSNR | BD-Rate | BD-PSNR |
| ClassB | -30.9% | 1.12dB | -6.1% | 0.18dB |
| ClassC | -29.9% | 1.52dB | -4.9% | 0.19dB |
| ClassD | -25.9% | 1.25dB | -6.8% | 0.31dB |
| Average | -28.9% | 1.30dB | -6.0% | 0.23dB |
| Sequences | RA-HE | | | |
| | Proposed vs.URQ[19] | | Proposed vs. R-λ[20] | |
| | BD-Rate | BD-PSNR | BD-Rate | BD-PSNR |
| ClassB | -18.2% | 0.55dB | 2.9% | -0.01dB |
| ClassC | -9.4% | 0.41dB | -2.3% | 0.08dB |
| ClassD | -2.5% | 0.13dB | -1.3% | 0.12dB |
| Average | -10.03% | 0.26dB | -0.23% | 0.06dB |

margin for both high resolution and low resolution sequences. The maximum BD-PSNR gain can be over 1.8 dB and 0.3 dB respectively.

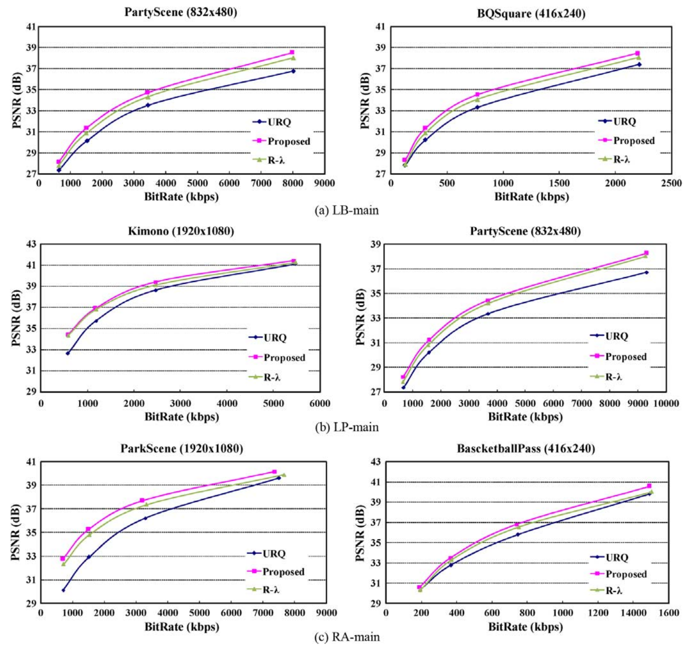


Fig. 9. RD curves of the performance comparison for the proposed algorithm under different testing configurations. (a) LB-main; (b) LP-main; (c) RA-main.

TABLE VIII
PERFORMANCE COMPARISON FOR THE PROPOSED ALGORITHM WITH HM ANCHOR

| Sequences | Proposed vs. HM8.0 without rate control | | | |
|---|---|---|---|---|
| | LB-main | | LP-main | |
| | BD-Rate | BD-PSNR | BD-Rate | BD-PSNR |
| ClassB | 3.5% | -0.07 | 2.0% | -0.05 |
| ClassC | -1.4% | 0.09 | -2.4% | 0.09 |
| ClassD | -2.0% | 0.08 | -1.6% | 0.07 |
| ClassE | -3.5% | 0.11 | -3.1% | 0.11 |
| Average | -0.9% | 0.05 | -2.3% | 0.06 |

The proposed Rate-GOP level bit allocation scheme and proposed header bit prediction model make small contribution for the coding performance. The average coding gain of these two tools for all testing configurations is about 1.86%.

Furthermore, six R-D curves are shown in Fig. 9 for different coding configurations. It can be observed the proposed algorithm has much better R-D performance than the other two algorithms for both high bit rate and low bit rate.

Compared with HM anchor without rate control, the URQ model and $\mathrm{R}-\lambda$ model generates performance loss [19], [20], especially for URQ model. Table VIII presents the comparison results between the proposed algorithm and HM anchor without rate control in LD low complexity configuration. It can be observed the proposed algorithm achieves a little better R-D performance than HM anchor without rate control.

## C. Complexity Analysis

In the proposed Rate-GOP based rate control, the parameters $a_i$ and $b_i$ in (20) are obtained by linear regression and the parameter $\theta_i$ in (20) are predicted from the previous coded frames as in (38) to avoid the complicated calculations in (19) and (37), respectively. The parameters $\alpha$ and $\beta$ in (43) are also obtained

TABLE IX
ENCODING COMPUTATION COMPARISON OF THE PROPOSED SCHEME WITH HM ANCHOR

| Configuration | $\Delta T$ | Configuration | $\Delta T$ |
|---|---|---|---|
| LP-main | 0.34% | LD-HE | 0.72% |
| LB-main | -0.87% | LB-HE | -0.10% |
| RA-main | -2.58% | RA-HE | -1.96% |
| Average | -1.04% | average | -0.45% |

by linear regression. The parameter $\omega_i$ in (39) is predicted from the previous Rate-GOPs.

Table IX shows the computation comparison of the proposed rate control scheme with HM anchor without rate control, where $\Delta T$ is calculated by

$$\Delta T = \frac{T_{pro} - T_{org}}{T_{org}} \times 100\%, \tag{45}$$

where $T_{org}$ and $T_{pro}$ indicate the total coding time of HM anchor without rate control and the proposed rate control scheme respectively.

From Table IX, it can be concluded that the computation overhead of the proposed method is no more than 0.72%, in some testing configurations, even faster than HM anchor without rate control. This is because HM anchor without rate control is implemented under fixed QP as in default configuration, while the proposed rate control scheme adjusted QP according to the bandwidth. If the calculated QP is bigger than QP of HM anchor without rate control, the encoding computation time may be reduced.

## VII. CONCLUSIONS

In this paper, a frame level rate control scheme based on Rate-GOP is proposed. The proposed rate control algorithm considered the Rate-GOP as a rate control unit and provided a hierarchical structure for Rate-GOP. The distortion dependency and the bit rate dependency are investigated between the coding frame and its reference frame. Then the inter-frame dependency based distortion model and bit rate model for Rate-GOP are derived. Based on these models and a mixed Laplacian distribution of residual information, a new $\rho$-domain Rate-GOP based rate control is proposed. The proposed algorithm was incorporated into the HM software HM8.0. The experimental results demonstrated that the proposed algorithm can achieve much better R-D performance than the two state of the art rate control algorithms for HEVC, even faster than HM software HM8.0 without rate control averagely.

## REFERENCES


[1] B. Bross, W.-J. Han, G. J. Sullivan, J.-R. Ohm, and T. Wiegand, Eds., "High efficiency video Coding (HEVC) text specification draft 6," in *Document: JCTVC-H1003, Joint Collaborative Team on Video Coding (JCT-VC)of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG118th Meeting*, San José, CA, USA, Feb. 1–10, 2012.

[2] *Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification*, ITU-T Rec. H.264 and ISO/IEC 14 496-10 AVC, Joint Video Team, Mar. 2003.

[3] Test Model 5 [Online]. Available: http://www.mpeg.org/MPEG/ MSSG/ tm5

[4] T. Chiang and Y. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 2, pp. 287–311, Apr. 1997.

[5] J. Corbera and S. Lei, Rate control for low-delay video communications ITU Study Group 16, Video Coding Experts Group. Portland, OR, USA, Documents Q15-A-20, 1997.

[6] F. Müller, "Distribution shape of two-dimensional DCT coefficients of natural images," *Electron. Lett.*, vol. 29, no. 22, pp. 1935–1936, Oct. 1993.

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

[8] Z. G. LI, F. Pan, K. P. Lim, and G. N. Feng, "Adaptive basic unit layer rate control for JVT," in *Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-G012, 7th Meeting*, Pattaya, Thailand, 2003, pp. 7–14, JVT-G012.

[9] M. Jiang and N. Ling, "On enhancing H.264/AVC video rate control by PSNR-based frame complexity estimation," *IEEE Trans. Consum. Electron.*, vol. 51, no. 1, pp. 281–286, Feb. 2005.

[10] D. Kwon, M. Shen, and C. C. Jay Kuo, "Rate control for H.264 video with enhanced rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 5, pp. 517–529, May 2007.

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

[12] Z. He and S. 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, Oct. 2002.

[13] J. Wen, M. Luttrell, and J. Villasenor, "Trellis-based r-d optimal quantization in H.263," *IEEE Trans. Image Process.*, vol. 9, no. 8, pp. 1431–1434, Aug. 2000.

[14] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit allocation for dependent quantization with applications to multi resolution and MPEG video coders," *IEEE Trans. Image Process.*, vol. 3, no. 5, pp. 533–545, Sep. 1994.

[15] L. J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 8, no. 4, pp. 446–459, Aug. 1998.

[16] J. Liu, Y. Cho, Z. Guo, and C. C. Jay Kuo, "Bit Allocation for Spatial Scalability Coding of H.264/SVC With Dependent Rate-Distortion Analysis," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 20, no. 7, pp. 967–981, Jul. 2010.

[17] S. Hu, S. Kwong, T. Zhao, and C. C. Jay Kuo, "Rate control optimization for temporal-layer scalable video coding," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 21, no. 8, pp. 1152–1162, Aug. 2011.

[18] S. H. Jang and N. Jayant, "An efficient bit allocation algorithm in dependent coding framework an one-way video applications," in *Proc. ICME*, 2004, pp. 383–386.

[19] H. Choi, J. Nam, J. Yoo, D. Sim, and I. V. Bajić, "Rate control based on unified RQ model for HEVC," in *Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG118th Meeting*, San José, CA, USA, Feb. 1–10, 2012.

[20] B. Li, "Rate control by R-lambda model for HEVC," in *Proc. 11th Meeting Joint Collaborative Team Video Coding (JCT-VC)of ITU-T 1/SC 29/WG 11*, Shanghai, China, Oct. 10–19, 2012.

[21] S. Wang, S. Ma, S. Wang, D. Zhao, and W. Gao, "Multi layer rate control algorithm for HEVC," in *Proc. ISCAS 2013*, Beijing, China.

[22] I. K. Kim, J. Min, T. Lee, W. J. Han, and J. Park, "Block partitioning structure in the HEVC standard," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 22, no. 12, pp. 1697–1706, Dec. 2012.

[23] R. Sjöberg and J. Samuelsson, "Absolute signaling of reference pictures," in *Proc. Joint Collaborative Team on Video Coding (JCT-VC)of ITU-T of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG116th Meeting*, Torino, Italy, 2011.

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

[25] B. M. Kim, "Modeling rates and distortion based on a mixture of laplacian distribution for inter-predicted residuals in quadtree coding of HEVC," *IEEE Signal Process. Lett.*, vol. 18, no. 10, pp. 571–574, Oct. 2011.

[26] F. Bossen, "Common test conditions and software reference configurations," in *Document: JCTVC-E700 Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/ WG114th Meeting*, Daegu, Korea, Jan. 20–28, 2011.

[27] G. Bjontegaard, "Calculation of average PSNR difference between RD curves," in *Proc. ITU-T Q.6/SG16 VCEG 13th Meeting*, Austin, TX, USA, Apr. 2001, document VCEG-M33.

**Shanshe Wang** received the B.S. degree in the Department of Mathematics from Heilongjiang University, Harbin, China, in 2004 and received M.S. degree in computer software and theory from Northeast Petroleum University, Daqing, China, in 2010. He is currently working toward the Ph.D. degree in computer science from the Harbin Institute of Technology, Harbin, China. His current research interests include video compression, image and video quality assessment.

**Siwei Ma** received the B.S. degree from Shandong Normal University, Jinan, China, in 1999, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005. From 2005 to 2007, he held a post-doctorate position with the University of Southern California, Los Angeles. Then, he joined the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing, where he is currently an Associate Professor. He has published over 100 technical articles in refereed journals and proceedings in the areas of image and video coding, video processing, video streaming, and transmission.

**Shiqi Wang** received the B.S. degree in computer science from the Harbin Institute of Technology, Harbin, China, in 2008. He is currently working toward the Ph.D. degree in computer science at Peking University, Beijing, China. He was a Visiting Student with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada from 2010 to 2011. From Apr. 2011 to Aug. 2011, he was with Microsoft Research Asia, Beijing, as an Intern. His current research interests include video compression, image and video quality assessment and multi-view video coding.

**Debin Zhao** received the B.S., M.S., and Ph.D. degrees in computer science, all from the Harbin Institute of Technology, Harbin, China, in 1985, in 1988, and in 1998, respectively. He was an Associate Professor in the Department of Computer Science, Harbin Institute of Technology, and a Research Fellow in the Department of Computer Science, City University of Hong Kong, Hong Kong, China, from 1989 to 1993. He is currently a Professor with the Department of Computer Science, Harbin Institute of Technology. His research interests include data compression, image processing, and human-machine interface.

**Wen Gao** (M'92–SM'05–F'09) received the Ph.D. degree in electronics engineering from the University of Tokyo, Japan, in 1991. He is a professor of computer science at Peking University, China. Before joining Peking University, he was a professor of computer science at Harbin Institute of Technology from 1991 to 1995, and a professor at the Institute of Computing Technology of Chinese Academy of Sciences. He has published extensively including five books and over 600 technical articles in refereed journals and conference proceedings in the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics.

Dr. Gao served or serves on the editorial board for several journals, such as IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Multimedia, IEEE Transactions on Autonomous Mental Development, EURASIP Journal of Image Communications, Journal of Visual Communication and Image Representation. He chaired a number of prestigious international conferences on multimedia and video signal processing, such as IEEE ICME and ACM Multimedia, and also served on the advisory and technical committees of numerous professional organizations.