λ-domain Perceptual Rate Control for 360-degree Video Compression

Li Li, *Member, IEEE*, Ning Yan, Zhu Li, *Senior member, IEEE*, Shan Liu, Houqiang Li, *Senior member, IEEE*

***Abstract*—The 360-degree video is currently projected to 2-D formats using various projection methods for efficient compression. As a necessary part for video compression, rate control is also indispensable for the projected 360-degree video compression. However, the current rate control algorithm has not been optimized for the 360-degree video compression yet. The current rate control algorithms especially the Coding Tree Unit (CTU) level bit allocation algorithms usually have not taken into consideration the characteristic that various pixels in 2-D formats may have different influences on the visual experiences. In this paper, we first propose an optimal CTU level weight taking this characteristic into consideration. The CTU level weight is an approximation to the pixel level weight since the smallest granularity of a rate control algorithm is usually CTU. Second, based on the CTU level weight, a weighted coding tree unit (CTU) level bit allocation algorithm is proposed to achieve better coding performance. The bits of each CTU are assigned that the λ of a CTU is inverse proportional to its CTU level weight. This CTU level bit allocation scheme can be applied to all the 360-degree video projection formats. Third, we propose a CTU row (CR) level rate control algorithm for the Equi-Rectangle Projection (ERP) format. Different CTUs in the same row in the ERP format are combined into a CR to better control the target bitrate and improve the performance. The proposed algorithms are implemented in the newest video coding standard High Efficiency Video Coding (HEVC) reference software. The experimental results show that the proposed algorithm is able to achieve much better subjective and objective qualities as well as smaller bitrate errors compared with the state-of-the-art rate control algorithms.**

***Index Terms*—High Efficiency Video Coding, Rate control, Rate distortion optimization, λ-domain rate control, 360-degree video compression**

I. Introduction

Innovative applications of the 360-degree video [1] are expected to become widespread in the near future due to its capability to bring immersive experiences to the users. Especially, thanks to the emergence of the various head-mounted displays [2] such as HTC Vive, Samsung Gear VR, and Oculus Rift, there are more and more 360-degree videos distributed online. However, since the 360-degree video is a bounding sphere containing the whole surroundings, the resolution of the 360-degree video should be $8K$ or even higher to provide satisfiable visual experience. The high resolution would absolutely bring high bitrate, which is now preventing the 360-degree video from its wide use. Therefore, there is an urgent need to compress the 360-degree video efficiently.

L. Li and Z. Li are with the University of Missouri-Kansas City, 5100 Rockhill Road, Kansas City, MO 64110, USA. L. Li is also with the CAS Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, University of Science and Technology of China, Hefei 230027, China. The email addresses are (lil1, lizhu)@umkc.edu.

N. Yan and H. Li are with the CAS Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, University of Science and Technology of China, No. 443 Huangshan Road, Hefei 230027, China. The email addresses are nyan@mail.ustc.edu.cn and lihq@ustc.edu.cn.

S. Liu is with Tencent America, 661 Bryant St, Palo Alto, CA 94301, USA. The email address is shanl@tencent.com.

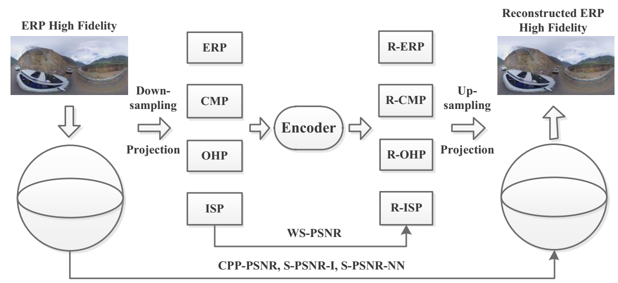


Fig. 1. Pipeline of the 360-degree video compression defined by JVET.

Currently, the Joint Video Exploration Team (JVET) of ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Pictures Experts Group (MPEG) is studying the potential need to include 360-degree video coding technologies in a future video coding standard. Fig. 1 shows the pipeline of the 360-degree video compression defined by JVET [3]. The high fidelity $8K$ Equi-Rectangle Projection (ERP) video source is first down-sampled and projected to a 2-D video such as ERP, Cube Map Projection (CMP), OctaHedron Projection (OHP), and IcoSahedron Projection (ISP) [4]. The 2-D video is then compressed using the video coding standard such as High Efficiency Video Coding (HEVC) [5] or the coming Versatile Video Coding (VVC) [6]. In the final, the reconstructed ERP, CMP, OHP, or ISP is up-sampled and projected back to the sphere for the users. Since various pixels in a projected 2-D video correspond to different areas in 3-D space, some specific quality measurements [4] other than Peak Signal to Noise Ratio (PSNR), such as Weighted Sphere-PSNR (WS-PSNR) [7], Crasters Parabolic Projection-PSNR (CPP-PSNR), Sphere PSNR-Nearest Neighbor (SPSNR-NN) [8], and Sphere PSNR-Interpolation (SPSNR-I) are provided to better approximate the subjective quality of the 360-degree video. In addition, some works try to use the 360-degree padding [9] [10] [11] [12] or the advanced motion model [13] [14] to improve the performance of the 360-degree video compression. However, the basic framework of using HEVC or VVC keeps unchanged.

In the 2-D video compression, a suitable rate control scheme is necessary to make full use of the bandwidth to improve the reconstructed video quality as much as possible. According to the key parameters used to control the bitrate, the rate control algorithms can be roughly divided into three groups: the Q-domain rate control algorithm, the ρ-domain rate control algorithm, and the λ-domain rate control algorithm. As its name implies, the Q-domain rate control algorithm considers the quantization step (Q) as the key factor to determine the bitrate and models the relationship between the target bitrate R and Q as linear [15] or second-order [16]. The ρ-domain rate control algorithm considers the percentage of non-zero transform coefficients after quantization ρ as a more robust factor to characterize the bitrate than Q and proposes a linear relationship between R and ρ [17]. However, as pointed out in [18], both the Q and ρ can only determine the residue bitrate and the Lagrangian Multiplier λ is the key factor to determine the overall bitrate. A hyperbolic model between R and λ is further proposed to accurately characterize their relationship. The λ-domain rate control algorithm has been integrated into the HEVC and VVC reference software. However, these rate control algorithms designed for general video compression, especially the CTU level bit allocation algorithms, have not taken the characteristics "unequal characteristics" of the 360-degree video that different pixels correspond various areas of the sphere into consideration. Therefore, the optimal performance for 360-degree video rate control has not been achieved yet.

Some researchers propose using spherical rate distortion optimization (RDO) [19] [20] to take the "unequal characteristics" in the 360-degree video compression into consideration. The mean square error (MSE) is replaced by the weighted MSE to better reflect the features of the 360-degree video. The spherical RDO is further combined with a general λ-domain rate control algorithm to improve the rate control performance. However, without explicit changes to bit allocation, the performance improvement provided by the rate control algorithm is limited. In addition, the repetitive calculations of the complex weighted MSE will increase the complexity of the encoder. Furthermore, the float weights per pixels used to calculate the weighted MSE make the RDO process unfriendly to the hardware design.

Therefore, in this paper, we design specified rate control algorithms for the 360-degree video compression by taking the "unequal characteristics" into consideration. The proposed rate control algorithms mainly have the following contributions.

- We propose a Coding Tree Unit (CTU) level weight as an approximation of the pixel level weight to consider the "unequal characteristics". We prove that our proposed CTU level weight is optimal if we consider the whole CTU as a unit.
- We propose a CTU level bit allocation algorithm Utilizing the CTU level weight to achieve better compression performance. The bits of each CTU are assigned that the λ of a CTU is inverse proportional to its CTU level weight. Such a CTU level bit allocation algorithm can be applied to various projection formats including ERP, CMP, and OHP.
- We further provide a CTU Row (CR) level rate control algorithm for the ERP format since the CTUs in the same row have the same CTU level weight. The CR level rate control algorithm is between the picture level bit allocation algorithm and the CTU level bit allocation algorithm. The CR level rate control algorithm can have more stable model parameters for the bit allocation than the CTU level and can make each CTU be assigned a more reasonable number of bits.

The proposed algorithms are implemented in the HEVC reference software. Very obvious bitrate savings, as well as smaller bitrate errors, are achieved compared with the state-of-the-art rate control algorithms.

The rest of this paper is organized as follows. We will review the related works on general rate control algorithms and spherical RDO in Section II. In Section III, we will introduce the derivation of the optimal CTU level weight from the pixel level weight. In Section IV and Section V, we will introduce the proposed CTU and CR level bit allocation algorithms in detail. The detailed experimental results are shown in Section VI. Section VII concludes the whole paper.

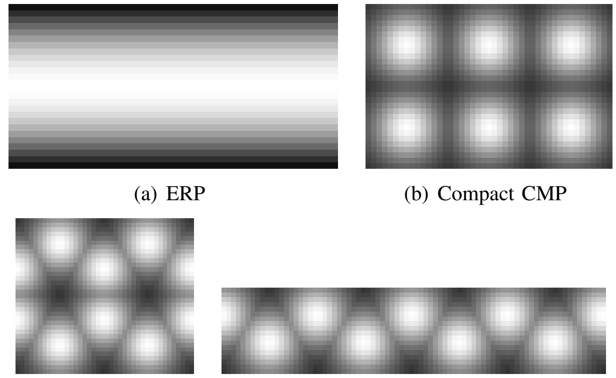
II. Related work

A. General rate control algorithm

The current rate control algorithms can be roughly divided into three groups: Q-domain rate control algorithm, ρ-domain rate control algorithm, and λ-domain rate control algorithm. In the Q-domain rate control algorithm, Q is considered as the key factor to determine the bitrate and a relationship between R and Q is built to control the bitrate. Chiang and Zhang [16] proposed a quadratic relationship between R and Q and provided a corresponding rate control algorithm for MPEG-4. Ma *et al.* [15] proposed a linear R-Q relationship and an efficient rate control algorithm for H.264/Advanced Video Coding. Kwon *et al.* [21] introduced a relationship between R and Q, which varies depending on the frame types. As we can see from the previous works that there are multiple relationships between R and Q, which shows that the relationship between R and Q is not robust enough which limits the performance of the Q-domain rate control algorithm. In addition, the header bits estimation [22] and initial quantization parameter (QP) [23] are also widely studied for the Q-domain rate control algorithms.

He *et al.* [24] first proposed that there is a more robust relationship between R and ρ compared with R and Q. They proposed a linear R-ρ relationship and provide a ρ-domain rate control algorithm to control the bitrate [24]. A bit allocation algorithm is further provided to improve the performance [25]. Pitrey *et al.* [26] [27] extended the ρ-domain rate control algorithm to H.264/Scalable Video Coding [28]. Wang *et al.* [29] proposed using a quadratic R-ρ model and a corresponding rate control algorithm for HEVC. However, as we have mentioned in Section I, both the Q and ρ can only determine the residue bitrate but have no influences on the non-residue bits. Especially, along with the increase of the header bits of HEVC and VVC, the Q and ρ-domain rate control algorithms become unsuitable for them.

To address this problem, Li *et al.* [18] first revealed that the λ is essentially the key factor to determine the overall bitrate. A hyperbolic R-λ model and a corresponding rate control algorithm are proposed to precisely control the bitrate for HEVC. They [30] and Li *et al.* [31] further provided picture level and CTU level bit allocation algorithms to further improve the performance. The λ-domain rate control algorithm has been implemented in the reference software of HEVC and VVC and used as the recommended rate control algorithm. There are also some algorithms focusing on the λ-domain intra frame rate control [32] [33] and initial encoding parameters determination [34]. However, most current CTU level bit allocation algorithms always treat all the CTUs with equal importance and are unsuitable for the projected 360-degree video. Li *et al.* [35] proposed assigning the bits inverse proportional to the CTU level weights. However, as far as we can see, those weights are not used in a proper way.



(a) ERP (b) Compact CMP (c) Compact OHP1 (d) Compact OHP2

Fig. 2. CTU level weights illustration of different projection formats of the 360-degree video.

B. Spherical rate distortion optimization

Some works have considered the "unequal characteristics" of the CTUs of the 360-degree video in the RDO process. Tang *et al.* [36] proposed using an adaptive quantization scheme to apply different QPs for various pixels according to their importance for the ERP format. Li *et al.* [19] [20] introduced a spherical RDO by applying a pixel level weight according to the area in the sphere the pixels correspond to improve the 360-degree video compression performance. Such a method can be applied to various projection formats of the 360-degree video. In addition, Liu *et al.* [37] proposed using the spherical MSE as the distortion metric and proposed a rate control algorithm by combining it with the general λ-domain rate control algorithm. However, the repetitive calculations of the spherical MSE will increase the complexity of the encoder.

III. OPTIMAL CTU LEVEL WEIGHT

All projections from sphere to the 2-D projection formats are performed pixel by pixel. In these projections, different small units containing various pixels in the projected plane may correspond to different surface areas on the sphere. For example, under the ERP format, the pixels near the equatorial are corresponding to a larger area compared with those near the north and south poles. Since the quality of the 360-degree video is essentially measured on the sphere, the pixels corresponding to a larger area can be considered as more "important". Therefore, the weights to indicate the "importance" of various pixels can be different. However, when we design a rate control algorithm, the smallest granularity is usually CTU. Therefore, we need to first derive an optimal CTU level weight to replace the pixel level weight for the following bit allocation and rate control processes.

We first define the weighted distortion $D_{i,j}^{\omega}$ of position (i,j) as

$$D_{i,j}^{\omega} = \omega_{i,j} D_{i,j}, \tag{1}$$

where $\omega_{i,j}$ is the pixel level weight of position (i,j). $D_{i,j}$ is the distortion of that pixel. The distortion is usually the square error between the original pixel and the reconstructed pixel. If we unify the weights of all the pixels in the kth CTU as ω_k, the difference E_k between the pixel level weighted distortion and the block level weighted distortion can be calculated as

$$E_k = \sum_{(i,j)\in CTU_k} (D_{i,j}^{\omega} - \omega_k D_{i,j})^2, \tag{2}$$

where CTU_k is the pixel set of the kth CTU.

We try to minimize the difference of the distortion calculated by the CTU level weight and the pixel level weight to optimize the CTU level weight,

$$\min_{\omega_k} E_k = \min_{\omega_k} \sum_{(i,j)\in CTU_k} (D_{i,j}^{\omega} - \omega_k D_{i,j})^2. \tag{3}$$

Such an unconstrained problem can be solved by setting the derivative of E_k with respect to ω_k to 0,

$$\frac{\partial E_k}{\partial \omega_k} = \frac{\partial \sum\limits_{(i,j)\in CTU_k} (D_{i,j}^{\omega} - \omega_k D_{i,j})^2}{\partial \omega_k} = 0. \tag{4}$$

If we substitute (1) into (4), (4) can be converted as

$$\omega_k = \frac{\sum\limits_{(i,j)\in CTU_k} \omega_{i,j} D_{i,j}^2}{\sum\limits_{(i,j)\in CTU_k} D_{i,j}^2}. \tag{5}$$

In theory, ω_k is not only related to $\omega_{i,j}$ but also related to the distortion of each pixel $D_{i,j}$. However, $D_{i,j}$ is only available after the encoding of the current CTU. Therefore, we have to find a way to estimate ω_k.

Within a local region such as a CTU, we assume that the $D_{i,j}$ of various pixels is the same and denoted as D_k,

$$D_{i,j} = D_k, \quad for\ any\ (i,j) \in CTU_k. \tag{6}$$

In this way, ω_k can be estimated as

$$\omega_k = \frac{\sum\limits_{(i,j)\in CTU_k} \omega_{i,j} D_k^2}{\sum\limits_{(i,j)\in CTU_k} D_k^2} = \frac{\sum\limits_{(i,j)\in CTU_k} \omega_{i,j}}{N_k}, \tag{7}$$

where N_k is number of pixels in CTU_k. The typical value of N_k is 64×64 except for the CTUs in right or bottom borders. Eq. (7) indicates that the CTU level optimal weight ω_k is equal to the average weight of all the pixels in the current CTU. Therefore, to derive the optimal weight ω_k, the only problem is to derive the $\omega_{i,j}$, which is related to each specified projection.

According to [7] and [20], it is assumed that each pixel (i, j) lies in the center of a very small unit, whose area is $P(i, j)$. The area of the corresponding unit in the sphere is $S(i, j)$. Then the $\omega_{i,j}$ can be calculated as

$$\omega_{i,j} = P(i, j)/S(i, j). \tag{8}$$

For the ERP format, the $\omega_{i,j}$ can be calculated as

$$\omega_{i,j} = cos\frac{(j + 0.5 - H/2)\pi}{H}, \tag{9}$$

where H is the height of the ERP format. We can see from (9) that the $\omega_{i,j}$ is only related to the longitude of the pixel, which is in accordance with our common sense. For the CMP and OHP formats, the $\omega_{i,j}$ can be calculated face by face. Using CMP as an example, the $\omega_{i,j}$ of the top left face is calculated by

$$\omega_{i,j} = (1 + \frac{d^2(i, j)}{r^2})^{(-3/2)}, \tag{10}$$

where r is the radius of the sphere, which is equal to half of the edge length of the cube. $d^2(i, j) = (i+0.5-r)^2+(j+0.5-r)^2$, which is the squared distance between position (i, j) and the center point of the face. The $\omega_{i,j}$ of the other faces can be calculated accordingly.

After the calculation of the $\omega_{i,j}$, the ω_k of the kth CTU can be calculated using (7). Fig. 2 shows the CTU level weights for various projection formats of the 360-degree video. They are in accordance with the pixel level weights as shown in [20] but with a coarser granularity. In Fig. 2, the extent of the brightness indicates the importance of each CTU. The brighter a CTU is, the more important a CTU will be and vice versa. The more important CTUs will be assigned more bits in the following bit allocation process. For example, for the ERP format, the CTUs near the equatorial will be assigned the most number of bits while those near the poles will be assigned the least number of bits.

IV. THE PROPOSED CTU LEVEL RATE CONTROL

As indicated in [18], λ is a more robust factor to determine the bitrate compared with Q and ρ. Therefore, the λ-domain rate control algorithm has been integrated into the HEVC reference software. In this paper, we also follow the λ-domain rate control algorithm to determine a proper λ for each picture or CTU to achieve accurate bitrate control. Note that the CTU level weights mainly have influences on the bit allocation of various CTUs in a CTU level rate control algorithm. For the GOP and picture level bit allocation and rate control algorithms, we still follow the original rate control algorithm for general videos as in [30].

For the CTU level rate control algorithm, under the constraint of the picture level target bits R_t, the optimization target of the CTU level rate control is to determine the λ_k for each CTU to minimize the sum of the weighted distortion of all the CTUs,

$$\min_{\lambda_k} \sum_{i=1}^{N} \omega_i D_i \quad s.t. \sum_{i=1}^{N} R_i \leq R_t, \quad k = 1, 2, ..., N, \tag{11}$$

where D_i and R_i are the distortion and bits of the ith CTU, respectively. N is the number of CTUs. λ_k is the λ of the kth CTU. ω_i is the CTU level weight derived in the last section. The constrained problem can be converted to the following unconstrained problem by introducing the Lagrangian Multiplier λ,

$$\min_{\lambda_k} \sum_{i=1}^{N} \omega_i D_i + \lambda \sum_{i=1}^{N} R_i, \quad k = 1, 2, ..., N. \tag{12}$$

The unconstrained problem can then be solved using the Lagrangian method,

$$\frac{\partial \sum_{i=1}^{N} \omega_i D_i}{\partial \lambda_k} + \lambda \frac{\partial \sum_{i=1}^{N} R_i}{\partial \lambda_k} = 0, \quad k = 1, 2, ..., N. \tag{13}$$

In an inter frame, most CTUs will use inter prediction and obtain the prediction from the reference frames instead of the neighboring CTUs. Therefore, these CTUs can be considered independent of each other,

$$\frac{\partial D_i}{\partial \lambda_k} = 0, \quad \frac{\partial R_i}{\partial \lambda_k} = 0, \quad i \neq k. \tag{14}$$

If we substitute (14) into (13), (13) can be converted to

$$\omega_k \frac{\partial D_k}{\partial \lambda_k} + \lambda \frac{\partial R_k}{\partial \lambda_k} = 0, \quad k = 1, 2, ..., N. \tag{15}$$

Since the λ_k is slope of the Rate-Distortion (R-D) curve of the kth CTU, λ_k can be expressed as

$$\lambda_k = -\frac{\partial D_k}{\partial R_k} = -\frac{\partial D_k}{\partial \lambda_k} / \frac{\partial R_k}{\partial \lambda_k}, \quad k = 1, 2, ..., N. \tag{16}$$

Then (15) can be solved as

$$\lambda_k = \frac{\lambda}{\omega_k} \quad k = 1, 2, ..., N. \tag{17}$$

Since ω_k can better reflect the perceptual quality of the 360-degree video, the proposed algorithm is able to improve both the objective and subjective qualities significantly.

Except for the constraints shown in (17), the sum of the target bits of all the CTUs should be equal to the picture level target bits,

$$\sum_{i=1}^{N} R_i = R_t. \tag{18}$$

In addition, in the λ-domain rate control algorithm, the λ_i follows a hyperbolic relationship with R_i,

$$\lambda_i = \alpha_i R_i^{\beta_i}, \quad i = 1, 2, ..., N. \tag{19}$$

Through combining (17), (18), and (19), we can derive the following equation as

$$\sum_{i=1}^{N} (\frac{\lambda}{\alpha_i \omega_i})^{\frac{1}{\beta_i}} = R_t. \tag{20}$$

In (20), there is only one unknown parameter λ. However, it is still difficult for us to obtain the analytic solution of λ since the β_i is a negative decimal number. In fact, Eq. (20) can be solved using numerical methods such as the bisection method. However, since N is usually very large especially for the high-resolution 360-degree videos, such a method will still increase the encoder complexity.

In this work, we consider the average λ of all the CTUs as the picture level λ_p. Note that the average of λs is the geometric average instead of arithmetic average,

$$\sqrt[N]{\prod_{i=1}^{N}\frac{\lambda}{\omega_i}} = \lambda_p. \tag{21}$$

In this way, the λ can be easily approximated using the following equation without complex calculations,

$$\lambda = \lambda_p \sqrt[N]{\prod_{i=1}^{N}\omega_i}. \tag{22}$$

Therefore, the target bits of each λ can be calculated as

$$R_i = \left(\frac{\lambda_p \sqrt[N]{\prod_{i=1}^{N}\omega_i}}{\alpha_i\omega_i}\right)^{\frac{1}{\beta_i}}, \quad i = 1, 2, ..., N. \tag{23}$$

As we can see from (23), the target bits are related to the content-related parameters α_i, β_i, and ω_i. The content-related characteristics are the key why the proposed CTU-level bit allocation algorithm can achieve a very good performance. Note that the actual bits cost of each CTU will not match exactly with the CTU level target bits. Therefore, the CTU level target bits TR_i should be also related to the remained picture level target bits LR_P and calculated as

$$TR_i = R_i - (\sum_{j=i}^{N} R_j - LR_P)/SW, \tag{24}$$

where SW is the slide window size, which we set as 4 in our experiments. Eq. (24) indicates that we will assign a less number of bits to the current CTU when the actual bits of the previous coded CTUs are more than the target bits and vice versa. We try to balance the budget within SW.

After determining the CTU level target bits, the λ will be then calculated using (19) for the following encoding process. In the general rate control algorithm, the λ of each CTU is clipped within an offset of 2 using the λ_P to prevent the quality of the current frame from fluctuating seriously. In this work, we intentionally change the λs of all the CTUs. Therefore, the clipping operation is also changed accordingly as follows.

$$\lambda_{clip} = \frac{\lambda_P \sqrt[N]{\prod_{i=1}^{N}\omega_i}}{\omega_k}. \tag{25}$$

$$\lambda_k = clip(\lambda_{clip} \times 2^{(-\frac{2}{3})}, \lambda_{clip} \times 2^{\frac{2}{3}}, \lambda_k). \tag{26}$$

The λ_k is clipped in the desired range within an offset of 2 to avoid the quality fluctuation in the sphere. The QP_k of the kth CTU is finally calculated according to λ_k using the following equation [38],

$$QP_k = 4.2003 \times ln(\lambda_k) + 13.7122. \tag{27}$$

After the determination of the λ_k and QP_k, we can finish the whole encoding process using RDO.

V. The proposed CTU row level rate control

For the ERP format, we can observe from Fig. 2 (a) that all the CTUs in the same row are with the same weights. From (17), we can see that the CTUs with the same weights should be encoded using the same λ to achieve the best performance. However, for each co-located CTU pair in the same hierarchical level especially in the lower levels of the random access coding structure, the content of each CTU may be unstable. The change of the content may lead to the change of the α_i and β_i of the ith CTU. Even if the λ_i is clipped in a suitable range, it may still lead to an inaccurate bit allocation of the ith CTU and bad influences on the bitrate accuracy and compression performance. Therefore, we try to organize all the CTUs with the same weights together to be a CTU row (CR) and propose a CR level rate control algorithm in this section. Compared with the CTU level model parameters, the model parameters of the CR level can be much more accurate since the video content of a CTU row is relatively stable.

The proposed CR level rate control algorithm is between the picture level and CTU level. The CR level rate control algorithm can be roughly divided into two processes: bit allocation and bitrate achievement. It will also have influences on the CTU level bit allocation algorithm. As the basic process is similar to the CTU level rate control, we will just introduce these concepts briefly in the following paragraphs.

The CR level bit allocation algorithm tries to assign the picture level target bits to each CR to minimize the picture level distortion. Such a formulation is similar to the CTU level bit allocation algorithms introduced in the last section. By solving a similar optimization problem, we can obtain a λ constraint similar to (17).

$$\lambda_k = \frac{\lambda}{\Omega_k}. \tag{28}$$

The main difference is that ω_k which indicates the weight of the kth CTU becomes Ω_k, which is the weight of the kth CR. The weight of the kth CR is equal to the sum of all the CTUs in the CR,

$$\Omega_k = \sum_{i=1}^{N_k}\omega_i, \tag{29}$$

where N_k is the number of CTUs in the kth CR. Then under the constraint of the picture level target bits, we can estimate the target bits of each CR using a similar equation as (24). The λ_k of the kth CR can then be calculated using (19) and clipped using (25) and (26). The QP_k of the kth CR is computed using (27). Since the model parameters are much more stable in CR level compared with the CTU level, the CR level λ_k can be a very good constraint for the CTU level rate control.

After we obtain the CR level target bits of the kth CR R_k, the optimization target of the CTU level rate control becomes minimizing the distortion of the kth CR under the R_k,

$$\min_{\lambda_k^i}\sum_{i=1}^{N_k} D_k^i, \quad \sum_{i=1}^{N_k} R_k^i \le R_k, \quad i = 1, 2, ..., N_k. \tag{30}$$

Different from (11), the CTU level rate control here is without any weight in front of the distortion since all the CTUs in a

CR has the same importance. Using the Lagrangian method, Eq. (30) can be solved as

$$\lambda_k^i = \lambda_k, \quad i = 1, 2, ..., N_k. \tag{31}$$

Then the target bits R_k^i of the ith CTU in the kth CR can be calculated as

$$R_k^i = (\frac{\lambda_k}{\alpha_k^i})^{\frac{1}{\beta_k^i}}, \quad i = 1, 2, ..., N_k, \tag{32}$$

where the α_k^i and β_k^i are the model parameters of the ith CTU of the kth CR. Similar to the CTU level rate control, the real target bits of each CTU should be also related to the remained bits in each CR. The λ_k^i can then be calculated using (19). Since the CTUs are in the same row, the λ_k^i is clipped within a limited range of λ_k to prevent the quality fluctuation for the CTUs with the same weights,

$$\lambda_k^i = clip(\lambda_k \times 2^{(-\frac{1}{3})}, \lambda_k \times 2^{\frac{1}{3}}, \lambda_k^i). \tag{33}$$

The QP_k^i is then computed using (27). The QP_k^i is clipped using

$$QP_k^i = clip(QP_k - 1, QP_k + 1, QP_k^i). \tag{34}$$

The RDO process of each CTU is performed based on the λ_k^i and QP_k^i.

VI. Experimental results

A. Simulation setup

The proposed algorithms are implemented in the HEVC reference software HM16.9 [39] to compare with HM anchor with and without rate control. The performance of a rate control algorithm can mainly be determined in two ways: the bitrate accuracy and video quality improvement. To better compare the proposed rate control algorithm with the HM anchor, we propose to generate the target bitrate using the following steps. We first run the anchor under the common test condition (CTC) for the 360-degree video compression [3]. Then the bitrates generated by the anchor are rounded and used as the target bitrate for the rate control algorithms. In this way, the comparison becomes easier since the bitrate generated by the rate control algorithms will be similar to that of the HM anchor. However, the bitrates generated by all the algorithms are still not the same, the Bjontegaard Delta rate (BD-rate) [40] is used to measure the performance. The quality metrics WS-PSNR, S-PSNR-NN, S-PSNR-I, CPP-PSNR which are more appropriate for the 360-degree video are used as the quality measurements. We test all the bitrates generated by the HM without rate control using QPs 22, 27, 32, and 37. Since the proposed algorithms are mainly proposed for the inter frames, we test the random access main10 configuration defined in the CTC.

We test all the 360-degree videos defined in the CTC to validate the performance of the proposed algorithms [3]. As shown in Fig. 1, the 8K or 4K video source is first projected to the 2D video formats for compression. We test the ERP, compact CMP (CCMP), and two compact OHP cases (COHP1 and COHP2) to explain the benefits of the proposed algorithms. The detailed resolutions of the projected 2D videos are shown in Table I. In the following, we will first show the bitrate accuracy and the R-D performance of the proposed CTU and CR level rate control algorithms. We will then show some R-D curves and examples of the subjective improvements to further explain the benefits.

TABLE I
Frame sizes for different projections

Format	8K		4K		Bit Depth
	width	height	width	height	
ERP	4096	2048	3328	1664	10
CCMP	3552	2368	2880	1920	10
COHP1	3096	2672	2536	2192	10
COHP2	6192	1336	5072	1096	10

B. Performance of the proposed CTU level rate control

1) R-D performance of the ERP format: Table II shows the R-D performance of the proposed CTU level rate control algorithm compared with the HM with and without rate control respectively. From Table II, we can see that the proposed algorithm can achieve an average of 4.1%, 4.1%, 4.0%, and 4.0% bitrate savings compared with the HM with the rate control algorithm for the WS-PSNR, S-PSNR-NN, S-PSNR-I, and CPP-PSNR, respectively. Comparing with the HM anchor without rate control, the proposed algorithm just suffers about 2.8% bitrate saving accordingly. The experimental results obviously demonstrate that the proposed CTU level rate control algorithm can lead to very significant bitrate savings compared with the state-of-the-art algorithm. In addition, the experimental results show that the performance improvements brought by the proposed algorithm are very consistent under different quality metrics. In the following, we will only show the performance of the proposed algorithms under the WS-PSNR metric.

In Table II, we can also see that the proposed algorithm can achieve obvious performance improvements for 90% test sequences compared with the HM with rate control. However, the proposed algorithm suffers 1.5% performance loss for the sequence Train. Compared with the HM without rate control, the performance loss can be as high as 27.0%. This loss is mainly caused by the following two reasons. First, the characteristics of the sequence Train change seriously. The first part of the sequence is stationary while the second part has many motions. Therefore, fewer bits are assigned to the first part while more bits are assigned to the second part in the HM without rate control. However, in a rate control algorithm, we need to keep the GOP level target bits stable, which inevitably leads to some performance losses. Second, the very stationary characteristic of the first part leads to many skip coding blocks, which leads to the CTU level model to be quite inaccurate. The inaccurate model has very significant influences on the following bit allocation and rate control algorithms. Such a problem can be partially alleviated by the proposed CR level rate control algorithm which will be illustrated in the next subsection.

In terms of the complexity, the proposed algorithm leads to slightly less encoding and decoding time compared with both the HM with and without rate control. The encoding and

TABLE II
THE PERFORMANCE OF THE PROPOSED CTU LEVEL RATE CONTROL PROBLEM FOR THE ERP FORMAT

Class	Sequence	HM with rate control as anchor				HM without rate control as anchor			
		WS-PSNR	S-PSNR-NN	S-PSNR-I	CPP-PSNR	WS-PSNR	S-PSNR-NN	S-PSNR-I	CPP-PSNR
8K	Train	1.5%	1.5%	1.6%	1.6%	27.0%	26.9%	27.1%	27.0%
	SkateboardingTrick	–0.9%	–0.9%	–0.9%	–0.9%	–2.7%	–2.8%	–2.8%	–2.8%
	SkateboardingInLot	–8.2%	–8.3%	–7.9%	–7.8%	–3.8%	–3.8%	–3.2%	–3.2%
	ChairLift	–6.8%	–6.9%	–6.7%	–6.6%	–4.3%	–4.3%	–4.2%	–4.2%
	KiteFlite	–3.1%	–3.0%	–3.0%	–3.1%	4.5%	4.6%	4.1%	4.0%
	Harbor	–1.5%	–1.5%	–1.5%	–1.5%	13.0%	13.0%	12.5%	12.5%
4K	PoleVault	–6.4%	–6.4%	–6.3%	–6.3%	–4.3%	–4.3%	–4.4%	–4.4%
	AerialCity	–6.2%	–6.3%	–6.2%	–6.2%	4.7%	4.7%	4.7%	4.7%
	DrivingInCity	–1.2%	–1.2%	–1.2%	–1.2%	2.3%	2.3%	2.3%	2.3%
	DrivingInCountry	–8.4%	–8.3%	–8.1%	–8.1%	–8.5%	–8.6%	–8.3%	–8.4%
Average		–4.1%	–4.1%	–4.0%	–4.0%	2.8%	2.8%	2.8%	2.8%
Enc. Time		98%				92%			
Dec. Time		96%				92%			

decoding time change mainly comes from the following two aspects. First, the CTU level weights proposed in Section III are only calculated once for one sequence. Therefore, the method itself will not increase any complexity to the encoder. Second, under the proposed algorithms, more blocks encoded with larger λs will choose the large blocks and skip mode. So some encoding complexities can be saved due to the early determination of the RDO process. Some decoding complexities can be saved because of the motion compensation with a large block size.

2) Performance of the other projection formats: Table III shows the WS-PSNR performance of the proposed CTU level rate control under the CCMP, COHP1, and COHP2 formats compared with HM with and without rate control. From Table III, we can see that, compared with the HM with rate control, the proposed algorithm achieves an average of 3.5%, 2.8%, and 3.0% bitrate savings for the CCMP, COHP1, and COHP2, respectively. The proposed algorithm achieves very consistent bitrate savings for all the test sequences under all the projected formats. Compared with the HM without rate control, the proposed algorithm only suffers 2.7%, 2.7%, and 2.4% performance losses accordingly. The proposed algorithm obviously demonstrate the effectiveness of the proposed CTU level rate control algorithm. In terms of the complexity, the proposed algorithm leads to a similar encoding/decoding time deduction compared with the ERP case.

3) Bitrate accuracy of the proposed CTU level rate control: Table IV shows the bitrate accuracy comparison of the proposed CTU level rate control method compared with the HM16.9 rate control. In terms of the average bitrate accuracy, the proposed rate control algorithm achieves an average of 0.22%, 0.09%, 0.05%, and 0.21% bitrate error reduction compared with the original rate control method in HM in ERP, CCMP, COHP1, and COHP2 cases, respectively. The experimental results show consistently better average bitrate accuracy in all these projection formats, which obviously demonstrate that the proposed algorithm is able to achieve slightly smaller bitrate errors. The benefits mainly come from a more reasonable number of bits assigned to each CTU, which can effectively prevent the cases where the bits are unable to be achieved. In addition, the proposed CTU level rate control algorithm leads to smaller maximum bitrate errors for the ERP and COHP2 cases, while larger maximum bitrate errors for the CCMP and COHP1 cases. This is caused by the fact that the last intra frame in the sequence accidentally spends much more bits under the CTU level rate control method compared with HM rate control in the lowest bitrate case. A very small number of bits left for the remaining frames lead to the larger bitrate errors in some extreme cases.

C. Performance of the CR level rate control on ERP format

Table V shows the R-D performance of the proposed CR level rate control algorithm in terms of WS-PSNR. Compared with the HM with rate control, the proposed CR level rate control algorithm is able to achieve an average of 5.3% performance improvements. Except for the sequence SkateboardingTrick which achieves almost the same performance, the proposed algorithm can bring very obvious performance improvements for all the other test sequences. The experimental results obviously demonstrate the effectiveness of the CR level rate control algorithm.

Compared with the proposed CTU level rate control, the proposed CR level rate control algorithm can lead to averagely 1.1% performance improvement. The performance improvements mainly come from the sequence Train while the performance of the other sequences keeps almost unchanged. For the sequence Train, the proposed algorithm can lead to much more accurate bit allocation and rate control by providing an accurate CR level model. The CR level λ calculated from the accurate CR level model can be used as a good constraint for the CTU level λ. Note that comparing with the HM without rate control, the proposed CR level rate control algorithm only suffers about 1.4% performance loss. In terms of complexity, the proposed algorithm brings slightly higher complexity compared with the CTU level rate control algorithm since we added a new level rate control algorithm and some corresponding bit allocation and rate control operations. Table VI shows the bitrate accuracy of the proposed CR level rate control algorithm. The proposed CR level rate control algorithm brings smaller bitrate error on average but a little bit higher bitrate error in the worse case compared with the CTU level rate control algorithm.

TABLE III
THE WS-PSNR PERFORMANCE OF THE PROPOSED CTU LEVEL RATE CONTROL PROBLEM FOR THE CMP, COHP1, AND COHP2 FORMATS

Class	Sequence	HM with rate control as anchor			HM without rate control as anchor		
		CCMP	COHP1	COHP2	CCMP	COHP1	COHP2
8K	Train	–3.4%	–6.3%	–6.5%	14.9%	13.0%	10.7%
	SkateboardingTrick	–4.6%	–2.0%	–1.3%	–7.6%	–4.2%	–4.0%
	SkateboardingInLot	–4.1%	–3.0%	–3.3%	2.0%	3.0%	2.9%
	ChairLift	–3.2%	–2.4%	–3.8%	1.8%	1.4%	0.7%
	KiteFlite	–2.1%	–2.7%	–2.8%	3.6%	1.9%	1.8%
	Harbor	–3.4%	–0.9%	–1.5%	8.8%	8.8%	8.5%
4K	PoleVault	–4.0%	–2.3%	–2.7%	–1.0%	0.8%	0.8%
	AerialCity	–3.5%	–2.2%	–2.4%	4.1%	2.7%	2.8%
	DrivingInCity	–2.8%	–3.0%	–2.7%	1.5%	0.1%	0.0%
	DrivingInCountry	–4.0%	–2.9%	–2.9%	–0.9%	–0.3%	–0.3%
Average		–3.5%	–2.8%	–3.0%	2.7%	2.7%	2.4%
Enc. Time		95%	96%	100%	88%	96%	99%
Dec. Time		96%	95%	98%	94%	96%	94%

TABLE IV
THE BITRATE ACCURACY COMPARISON OF THE CTU LEVEL RATE CONTROL AND HM WITH RATE CONTROL

Projection	HM with rate control		CTU level rate control	
	Average	Maximum	Average	Maximum
ERP	0.97%	4.64%	0.75%	3.91%
CCMP	0.90%	4.85%	0.81%	6.62%
COHP1	0.82%	4.89%	0.77%	5.35%
COHP2	0.96%	6.49%	0.75%	4.30%

TABLE V
THE WS-PSNR PERFORMANCE OF THE PROPOSED CR LEVEL RATE CONTROL PROBLEM FOR ERP FORMAT

Class	Sequence	HM w RC	CTU level	HM w/o RC
8K	Train	–9.0%	–10.2%	14.8%
	SkateboardingTrick	0.2%	1.2%	–1.6%
	SkateboardingInLot	–9.4%	–1.3%	–5.0%
	ChairLift	–7.4%	–0.6%	–4.9%
	KiteFlite	–2.8%	0.3%	4.8%
	Harbor	–2.9%	–1.4%	11.4%
4K	PoleVault	–5.6%	0.8%	–3.5%
	AerialCity	–6.9%	–0.8%	3.9%
	DrivingInCity	–0.8%	0.4%	2.7%
	DrivingInCountry	–7.9%	0.5%	–8.1%
Average		–5.3%	–1.1%	1.4%
Enc. Time		102%	106%	96%
Dec. Time		103%	105%	95%

D. R-D curves

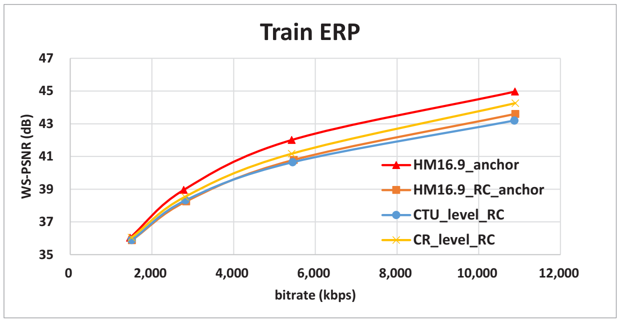
Some examples of R-D curves for different projection formats are shown in Fig. 3. From Fig. 3, we can see that the proposed CTU level rate control algorithm can lead to significant bitrate savings compared with the original rate control algorithm in HM for various projection formats. We can also see that the benefits mainly come from the high bitrate case. In high bitrate case, more blocks with small weights and large λ will choose skip mode and thus can provide more bits for the blocks with large weights and improve the quality. In low bitrate case, the situation is not as obvious as the high bitrate case. In addition, we can see that for the sequence train, the CR level rate control algorithm can bring very significant gain compared with the CTU level rate control. However, it still suffers obvious losses compared with the HM without rate control. This is mainly due to the change of the characteristics of the test sequence. The motion of the first half of the sequence Train is very slow while that of the second half becomes relatively fast. For the HM16.9 without rate control, the bits can be assigned according to the content. However, the rate control algorithm needs to make the GOP level target bits relatively steady and thus leads to some performance losses. For the other test sequences, the proposed rate control will not suffer serious losses compared with the HM without rate control.

TABLE VI
THE BITRATE ACCURACY RESULTS OF THE CR LEVEL RATE CONTROL

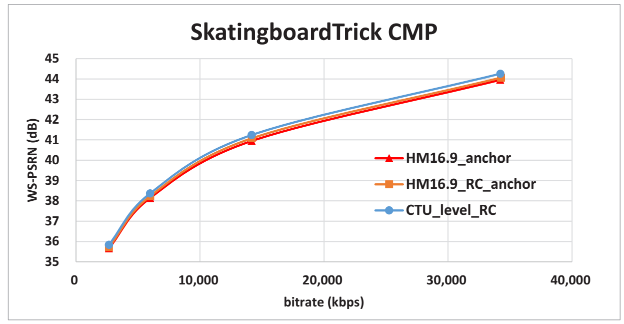
Test Case	Average	Maximum
HM w RC	0.97%	4.64%
CTU level	0.75%	3.91%
CR level	0.72%	4.93%

E. Subjective quality

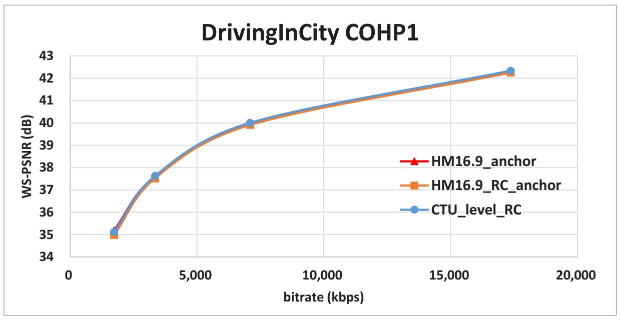
Some examples of the subjective quality improvements of the proposed rate control algorithms for ERP and CMP formats are shown in Fig. 4 and Fig. 5, respectively. Fig. 4 is a 200×200 region cropped from the sequence DrivingInCountry, picture order count (POC) 2, view port $(75, 75, 0, 0)$ with size 856×856. Fig. 4 (a), (b), (c), and (d) are the HM without rate control, HM with rate control, the proposed CTU level rate control, and the proposed CR level rate control algorithms, respectively. From Fig. 4, we can see that both the taillights and its left mountain are encoded with much better quality under the proposed CTU and CR level rate control algorithms compared with the HM16.9 original rate control algorithm. Fig. 5 is a 400×400 region cropped from the sequence SkateboardInLot, POC 16, view port $(75, 75, 0, 0)$ with size 1816×1816. Fig. 4 (a), (b), and (c) are the HM without rate control, HM with rate control, and the proposed CTU level rate control, respectively. From Fig. 5, we can see that the rearview mirror is lost under the HM16.9 original rate control algorithm while it is kept completely under the proposed CTU level rate control algorithm. All these results are obtained under the same target bitrate. The experimental



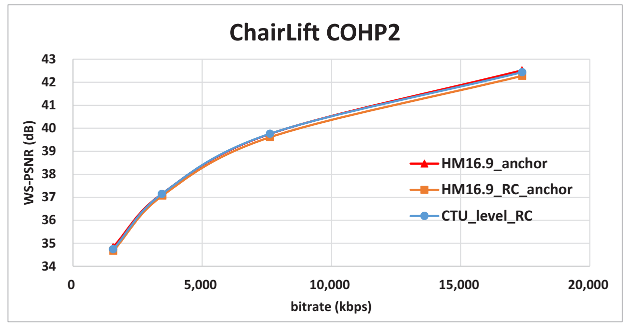
(a)



(b)



(c)



(d)

Fig. 3. Some examples of the R-D curves for different projection formats.

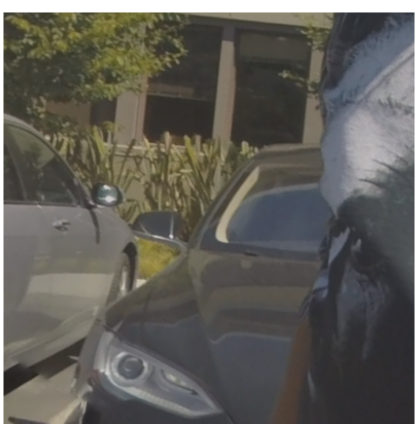
(a) ERP Original

(b) ERP HM16.9 RC

(c) ERP CTU level

(d) ERP CR level

Fig. 4. View port subjective improvement of the ERP format. Test sequence: DrivingInCountry. Picture order count: 2. The region shown is a cropped 200×200 zone staring at $(300, 500)$ from the view port $(75, 75, 0, 0)$ with size 856×856. (a) original sequence; (b) HM16.9 RC (bitrate: $2055.9kbps$); (c) HM16.9 CTU level (bitrate: $2040.2kbps$); (d) HM16.9 CR level (bitrate: $2037.2kbps$).



(a) CMP Original

(b) CMP HM16.9 RC

(c) CMP CTU level

Fig. 5. View port subjective improvement of the CMP format. Test sequence: SkateboardInLot. Picture order count: 16. The region shown is a cropped 400×400 zone staring at $(500, 800)$ from the view port $(75, 75, 0, 0)$ with size 1816×1816. (a) original sequence; (b) HM16.9 RC (bitrate: $2021.3kbps$); (c) HM16.9 CTU level (bitrate: $2020.9kbps$).

results obviously demonstrate that the proposed algorithms are able to significantly improve the subjective quality of the 360-degree video compression.

VII. Conclusion

In this paper, we propose a λ-domain perceptual rate control algorithms for the 360-degree video compression. We first propose a Coding Tree Unit (CTU) level weight for each CTU and prove that such a CTU level weight is optimal to measure the importance of each CTU. Then based on the CTU level weight, we propose a CTU level rate control algorithm which can be applied to all the projection formats of the 360-degree video. Furthermore, for the ERP format, a CTU Row (CR) level rate control algorithm is proposed to make each CR achieve the target bitrate more precisely. The proposed algorithms are implemented in the High Efficiency Video Coding reference software. The experimental results show that the proposed algorithm is able to provide significant bitrate savings for multiple projection formats compared with the original rate control algorithm in the HEVC reference software. The experimental results obviously demonstrate the effectiveness of the proposed algorithm.

References

[1] 360-degree video. [Online]. Available: https://en.wikipedia.org/wiki/360-degree_video

[2] Head-mounted display. [Online]. Available: https://en.wikipedia.org/wiki/Head-mounted_display

[3] J. Boyce, E. Alshina, A. Abbas, and Y. Ye, "JVET common test conditions and evaluation procedures for 360 video," Document JVET-D1030, Oct. 2016.

[4] Y. Ye, E. Alshina, and J. Boyce, "Algorithm descriptions of projection format conversion and video quality metrics in 360Lib Version 5," Document JVET-H1004, Geneva, CH, Oct. 2017.

[5] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[6] B. Bross, J. Chen, and S. Liu, "Versatile Video Coding (Draft 5)," Document JVET-N1001, Mar. 2019.

[7] Y. Sun, A. Lu, and L. Yu, "Weighted-to-spherically-uniform quality evaluation for omnidirectional video," *IEEE Signal Processing Letters*, vol. 24, no. 9, pp. 1408–1412, Sep. 2017.

[8] M. Yu, H. Lakshman, and B. Girod, "A framework to evaluate omnidirectional video coding schemes," in *2015 IEEE International Symposium on Mixed and Augmented Reality*, Sep. 2015, pp. 31–36.

[9] L. Li, Z. Li, X. Ma, H. Yang, and H. Li, "Advanced spherical motion model and local padding for 360 video compression," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2342–2356, May 2019.

[10] Y. He, Y. Ye, P. Hanhart, and X. Xiu, "AHG8: Geometry padding for 360 video coding," Document JVET-D0075, Oct. 2016.

[11] J. Sauer and M. Wien, "AHG8: Results for geometry correction for motion compensation of planar-projected 360VR video with JEM4.1 and 360Lib," Document JVET-E0026, Jan. 2017.

[12] M. Coban, G. V. der Auwera, and M. Karczewicz, "AHG8: Reference picture extension of ACP format 360-degree video," Document JVET-G0058, Jul. 2017.

[13] L. Li, Z. Li, M. Budagavi, and H. Li, "Projection based advanced motion model for cubic mapping for 360-degree video," in *2017 IEEE International Conference on Image Processing (ICIP)*, Sep. 2017, pp. 1427–1431.

[14] B. Vishwanath, T. Nanjundaswamy, and K. Rose, "Rotational motion model for temporal prediction in 360 video coding," in *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*, Oct 2017, pp. 1–6.

[15] S. Ma, W. Gao, and Y. Lu, "Rate-distortion analysis for h.264/avc video coding and its application to rate control," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 12, pp. 1533–1544, Dec 2005.

[16] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 1, pp. 246–250, Feb 1997.

[17] Z. He and S. K. Mitra, "A linear source model and a unified rate control algorithm for dct video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 11, pp. 970–982, Nov 2002.

[18] B. Li, H. Li, L. Li, and J. Zhang, "λ domain rate control algorithm for high efficiency video coding," *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp. 3841–3854, Sep. 2014.

[19] Y. Li, J. Xu, and Z. Chen, "Spherical domain rate-distortion optimization for 360-degree video coding," in *2017 IEEE International Conference on Multimedia and Expo (ICME)*, July 2017, pp. 709–714.

[20] ——, "Spherical domain rate-distortion optimization for omnidirectional video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2018.

[21] D. Kwon, M. Shen, and C. . J. Kuo, "Rate control for h.264 video with enhanced rate and distortion models," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 5, pp. 517–529, May 2007.

[22] S. Q. He, Z. J. Deng, and C. Shi, "A novel header bits estimation scheme for h.264/avc standard," in *Measurement Technology and Engineering Researches in Industry*, ser. Applied Mechanics and Materials, vol. 333. Trans Tech Publications, 9 2013, pp. 787–790.

[23] H. Wang and S. Kwong, "Rate-distortion optimization of rate control for h.264 with adaptive initial quantization parameter determination," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 1, pp. 140–144, Jan 2008.

[24] Z. He, Y. K. Kim, and S. K. Mitra, "Low-delay rate control for dct video coding via /spl rho/-domain source modeling," *IEEE Transactions on Circuits and Systems*

for Video Technology, vol. 11, no. 8, pp. 928–940, Aug 2001.

[25] Z. He and S. K. Mitra, "Optimum bit allocation and accurate rate control for video coding via /spl rho/-domain source modeling," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 10, pp. 840–849, Oct 2002.

[26] Y. Pitrey, Y. Serrand, M. Babel, and O. Dforges, "Rho-domain for low-complexity rate control on mpeg-4 scalable video coding," in *2008 Tenth IEEE International Symposium on Multimedia*, Dec 2008, pp. 89–96.

[27] Y. Pitrey, M. Babel, and O. Deforges, "One-pass bitrate control for mpeg-4 scalable video coding using -domain," in *2009 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, May 2009, pp. 1–5.

[28] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the h. 264/avc standard," *To appear in IEEE Transactions on Circuits and Systems for Video Technology*, p. 1, 2007.

[29] S. Wang, S. Ma, S. Wang, D. Zhao, and W. Gao, "Rate-gop based rate control for high efficiency video coding," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 1101–1111, Dec 2013.

[30] L. Li, B. Li, H. Li, and C. W. Chen, "λ -domain optimal bit allocation algorithm for high efficiency video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 1, pp. 130–142, Jan 2018.

[31] S. Li, M. Xu, Z. Wang, and X. Sun, "Optimal bit allocation for ctu level rate control in hevc," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 11, pp. 2409–2424, Nov 2017.

[32] M. Wang, K. N. Ngan, and H. Li, "An efficient frame-content based intra frame rate control for high efficiency video coding," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 896–900, July 2015.

[33] Y. Li, B. Li, D. Liu, and Z. Chen, "A convolutional neural network-based approach to rate control in hevc intra coding," in *2017 IEEE Visual Communications and Image Processing (VCIP)*, Dec 2017, pp. 1–4.

[34] L. Li, B. Li, D. Liu, and H. Li, "λ-domain rate control algorithm for hevc scalable extension," *IEEE Transactions on Multimedia*, vol. 18, no. 10, pp. 2023–2039, Oct 2016.

[35] B. Li, L. Song, R. Xie, and W. Zhang, "Weight-based bit allocation scheme for vr videos in hevc," in *2017 IEEE Visual Communications and Image Processing (VCIP)*, Dec 2017, pp. 1–4.

[36] M. Tang, Y. Zhang, J. Wen, and S. Yang, "Optimized video coding for omnidirectional videos," in *2017 IEEE International Conference on Multimedia and Expo (ICME)*, July 2017, pp. 799–804.

[37] Y. Liu, L. Yang, M. Xu, and Z. Wang, "Rate control schemes for panoramic video coding," *Journal of Visual Communication and Image Representation*, vol. 53, pp. 76–85, 2018.

[38] B. Li, J. Xu, D. Zhang, and H. Li, "Qp refinement according to lagrange multiplier for high efficiency video coding," in *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*, May 2013, pp. 477–480.

[39] "High Efficiency Video Coding test model, HM-16.9," https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/, accessed: 2018.

[40] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," Document VCEG-M33, Austin, Texas, USA, Apr. 2001.