

λ -Domain Perceptual Rate Control for 360-Degree Video Compression

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

I. INTRODUCTION

Abstract—The 360-degree video is projected to 2-D formats using various projection methods for efficient compression. As a necessary part of general-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 Coding Tree Unit (CTU) level bit allocation in the rate control algorithm has not taken into consideration the characteristic that various pixels in 2-D formats have different influences on the visual experiences. In this article, 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 CTU level bit allocation algorithm is proposed to achieve better coding performance. The bits of each CTU are assigned that the Lagrange multiplier λ of a CTU is inversely proportional to its CTU level weight. This CTU level bit allocation scheme is 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 provide more stable model parameters. 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 algorithm.

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

Manuscript received April 16, 2019; revised September 18, 2019 and December 13, 2019; accepted December 22, 2019. Date of publication December 31, 2019; date of current version February 5, 2020. This work was partially supported by the NSF I/UCRC on Big Learning and Tencent Media Lab. This work was partially supported by the National Key Research and Development Plan under Grant 2017YFB1002401. The guest editor coordinating the review of this manuscript and approving it for publication was Dr. Ce Zhu. (*Corresponding author: Zhu Li*.)

L. Li is with the University of Missouri-Kansas City, Kansas City, MO 64110 USA, and 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 (e-mail: lil1@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, Hefei 230027, China (e-mail: nyan@mail.ustc.edu.cn; lihq@ustc.edu.cn).

Z. Li is with the University of Missouri-Kansas City, Kansas City, MO 64110 USA (e-mail: lizhu@umkc.edu).

S. Liu is with the Tencent America, Palo Alto, CA 94301 USA (e-mail: shanl@tencent.com).

Digital Object Identifier 10.1109/JSTSP.2019.2963154

INNOVATIVE applications of 360-degree videos [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 various head-mounted displays [2] such as High Tech Computer (HTC) Vive, Samsung Gear Virtual Reality (VR), and Oculus Rift, there are more and more 360-degree videos distributed online. However, since a 360-degree video is a bounding sphere containing the whole surroundings, the resolution of a 360-degree video should be 8 K or even higher to provide satisfactory 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.

Currently, the Joint Video Exploration Team (JVET) of Video Coding Experts Group (VCEG) and 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 8 K 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 researchers try to use the 360-degree padding [9]–[12] or the advanced motion model [13], [14] to improve the performance of the 360-degree video compression. However, the 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

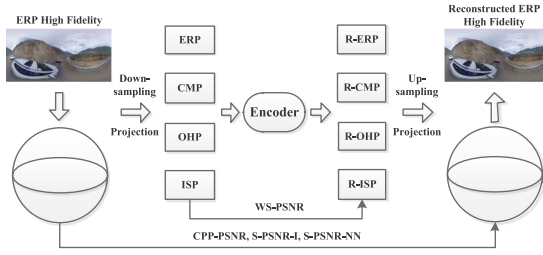


Fig. 1. Pipeline of the 360-degree video compression defined by JVET.

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 Lagrange multiplier λ is essentially the key factor to determine the overall bitrate. A hyperbolic model is further proposed in [18] to accurately characterize the relationship between R and λ . The λ -domain rate control algorithm is integrated into the HEVC and VVC reference software. However, the rate control algorithm designed for general videos has not been optimized for the 360-degree video compression yet. The Coding Tree Unit (CTU) level bit allocation in the rate control algorithm has not taken into consideration the “unequal characteristics” that various pixels in 2-D formats have different influences on the visual experiences. 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 [21]. However, without explicit changes to the 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 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 CTU as a basic 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 inversely proportional to its CTU level weight. This CTU level bit allocation algorithm is 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 and the CTU level bit allocation. It can provide more stable model parameters than the CTU level and make each CTU be assigned a more reasonable number of bits.

The proposed algorithms are implemented in the newest video coding standard High Efficiency Video Coding (HEVC) reference software. The source codes of the proposed algorithms are released at https://github.com/xiaoliqiu/360_Rate_Control. The experimental results show that the proposed algorithms can achieve obvious bitrate savings and smaller bitrate errors 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 V-B 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 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 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.* [22] introduced a relationship between R and Q that 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. Therefore, the performance of the Q -domain rate control algorithm is limited. In addition, the header bits estimation [23], initial quantization parameter (QP) determination [24], and frame skip avoidance [25] are also widely studied for the Q -domain rate control algorithms.

He *et al.* [26] 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 provided a ρ -domain rate control algorithm for video compression [26]. A bit allocation algorithm is further provided to improve the performance [27]. Pitrey *et al.* [28], [29] extended the ρ -domain rate control

algorithm to H.264/Scalable Video Coding [30]. Wang *et al.* [31] proposed a quadratic R - ρ model and a corresponding rate control algorithm for HEVC. Gao *et al.* [32] introduced a ρ -domain frame-level bit allocation method based on a synthesized Laplacian distribution model. 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 bitrate. 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 [33] and Li *et al.* [34] further provided picture level and CTU level bit allocation algorithms to improve the performance. The λ -domain rate control algorithm is implemented in the reference software of HEVC and VVC and used as the recommended rate control algorithm. In addition, Guo *et al.* [35] proposed a λ -domain frame level bit allocation considering the content-related quality dependencies. They further extended the frame level bit allocation to CTU level in [36]. Cai *et al.* [37] introduced a real-time constant quality rate control strategy for HEVC. Chen *et al.* [38] took the distribution of the R - λ model parameters into consideration and proposed a CTU level bit allocation to improve the performance. Gao *et al.* [39] proposed a structure similarity-based game theory approach for CTU level bit allocation. There are also some algorithms focusing on the λ -domain intra frame rate control [40]–[42] and initial encoding parameters determination [43]. In addition to the rate control for general videos, some researchers extended the λ -domain rate control schemes to scalable video coding [44], [45], high dynamic range video coding [46], multi-view video coding [47]–[49], and 3D video coding [50], [51]. However, most current CTU level bit allocation algorithms treat all the CTUs with equal importance and are unsuitable for the projected 360-degree video compression. For the rate control for 360-degree video compression, Li *et al.* [52] proposed assigning the bits inversely proportional to the CTU level weights. However, as far as we can see, those weights are not used in a proper way.

B. Spherical Rate Distortion Optimization

Some works considered the “unequal characteristics” of the CTUs of the 360-degree video in the RDO process. Tang *et al.* [53] 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 to improve the 360-degree video compression performance. This method was applied to various projection formats of the 360-degree video. In addition, Liu *et al.* [21] proposed using the spherical MSE as the distortion metric and provided 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 correspond to larger areas 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 larger areas should be considered as more “important”. Therefore, the weights to indicate the “importance” of various pixels are 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

$$D_{i,j}^\omega = \omega_{i,j} D_{i,j}, \quad (1)$$

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 k th CTU as ω_k , the difference E_k between the pixel level weighted distortion and the block level weighted distortion is calculated as

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

where CTU_k is the pixel set of the k th CTU.

We try to minimize E_k to optimize the CTU level weight,

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

This unconstrained problem is solved by setting the derivative of E_k with respect to ω_k to 0,

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

If we substitute (1) into (4), (4) is converted as

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

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, $D_{i,j}$ will only be available after the encoding of the current CTU. Therefore, we have to find another way to estimate ω_k .

Eq. (5) can be converted to the following form,

$$\omega_k = \frac{\sum_{(i,j) \in CTU_k} (\omega_{i,j} - C_k + C_k) D_{i,j}^2}{\sum_{(i,j) \in CTU_k} D_{i,j}^2}. \quad (6)$$

Here, we set C_k as

$$C_k = \frac{\sum_{(m,n) \in CTU_k} \omega_{m,n}}{N_k}, \quad (7)$$

TABLE I
THE AVERAGE RATIOS BETWEEN THE SECOND TERM AND THE FIRST TERM
IN (8) FOR ALL THE CTUs FOR ALL THE TEST SEQUENCES UNDER
DIFFERENT BITRATES

Class	Sequence	Q22	Q27	Q32	Q37
8K	Train	2.7%	2.6%	2.7%	2.8%
	SkateboardingTrick	3.0%	2.9%	3.2%	3.1%
	SkateboardingInLot	1.1%	1.2%	1.3%	1.4%
	ChairLift	1.2%	1.4%	1.7%	2.3%
	KiteFlite	1.5%	1.7%	1.8%	1.9%
	Harbor	0.9%	0.9%	1.1%	1.2%
	Trolley	1.0%	1.1%	1.3%	1.5%
	GasLamp	0.8%	1.0%	1.1%	1.6%
4K	Balboa	1.4%	1.5%	1.8%	2.0%
	Broadway	1.7%	2.1%	2.2%	2.4%
	Landing2	1.8%	1.9%	2.0%	2.1%
	BranCastle2	1.9%	2.4%	2.4%	2.2%
4K	PoleVault	1.8%	1.9%	2.1%	2.1%
	AerialCity	1.6%	1.6%	1.8%	2.1%
	DrivingInCity	1.4%	1.5%	1.6%	2.1%
	DrivingInCountry	2.0%	1.9%	1.9%	2.0%
Average		1.61%	1.73%	1.88%	2.05%

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. (6) can be further composed into two terms,

$$\omega_k = C_k + \frac{\sum_{(i,j) \in CTU_k} (\omega_{i,j} - C_k) D_{i,j}^2}{\sum_{(i,j) \in CTU_k} D_{i,j}^2}. \quad (8)$$

From (8), we can see that if the second term is much smaller than the first term, we can approximate the value of ω_k as the first term. Therefore, we count the average ratio between the second term and the first term for all the CTUs for many 360-degree sequences under different bitrates as shown in Table I. We can see that the average ratio is small under both low and high bitrates. This demonstrates that the second term can be ignored compared with the first term. Therefore, the ω_k is approximated C_k ,

$$\omega_k \approx C_k = \frac{\sum_{(m,n) \in CTU_k} \omega_{m,n}}{N_k}. \quad (9)$$

Eq. (9) 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 $\omega_{i,j}$ that 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 small unit, whose area is $P(i, j)$. The area of the corresponding unit on the sphere is $S(i, j)$. Then the $\omega_{i,j}$ is calculated as

$$\omega_{i,j} = P(i, j) / S(i, j). \quad (10)$$

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

$$\omega_{i,j} = \cos \frac{(j + 0.5 - H/2)\pi}{H}, \quad (11)$$

where H is the height of the ERP format. We can see that the $\omega_{i,j}$ is only related to the longitude of the pixel. For the CMP and OHP formats, the $\omega_{i,j}$ is calculated face by face. Using CMP as an example, the $\omega_{i,j}$ of the top left face is calculated by

$$\omega_{i,j} = \left(1 + \frac{d^2(i, j)}{r^2}\right)^{(-3/2)}, \quad (12)$$

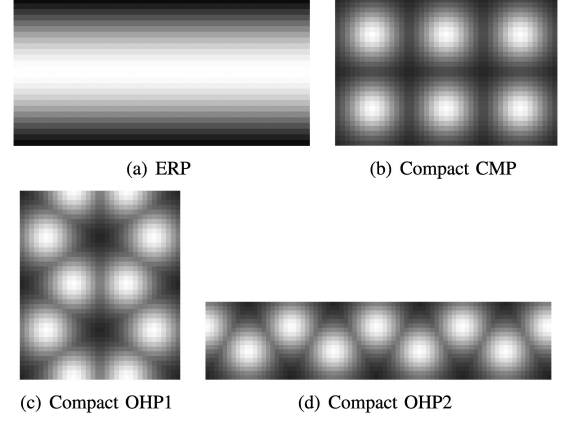


Fig. 2. CTU level weights illustration of different projection formats of the 360-degree video.

where r is the radius of the sphere that is equal to half of the edge length of the cube. $d^2(i, j) = (i + 0.5 - r)^2 + (j + 0.5 - r)^2$ is the squared distance between position (i, j) and the center position 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 k th CTU is calculated using (9). 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 largest number of bits while those near the poles will be assigned the smallest 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 and VVC reference software. In this paper, we 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. Therefore, for the group of pictures (GOP) and picture level bit allocation and rate control algorithms, we follow the original rate control algorithm for general videos as in [33].

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 distortions of all the CTUs under the constraint of the picture level target bits R_t ,

$$\min_{\lambda_k} \sum_{i=1}^N \omega_i D_i \quad \text{s.t.} \quad \sum_{i=1}^N R_i \leq R_t, \quad k = 1, 2, \dots, N, \quad (13)$$

where D_i and R_i are the distortion and bits of the i th CTU, respectively. N is the number of CTUs in the frame. λ_k is the λ

of the k th CTU. ω_i is the CTU level weight derived in (9). The constrained problem is converted to the following unconstrained problem by introducing the Lagrange multiplier λ ,

$$\min_{\lambda_k} \sum_{i=1}^N \omega_i D_i + \lambda \sum_{i=1}^N R_i, \quad k = 1, 2, \dots, N. \quad (14)$$

The unconstrained problem can then be solved using the method of Lagrange multipliers,

$$\frac{\partial \sum_{i=1}^N \omega_i D_i}{\partial \lambda_k} + \lambda \frac{\partial \sum_{i=1}^N R_i}{\partial \lambda_k} = 0, \quad k = 1, 2, \dots, N. \quad (15)$$

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

$$\frac{\partial D_i}{\partial \lambda_k} = 0, \quad \frac{\partial R_i}{\partial \lambda_k} = 0, \quad i \neq k. \quad (16)$$

If we substitute (16) into (15), (15) is converted to

$$\omega_k \frac{\partial D_k}{\partial \lambda_k} + \lambda \frac{\partial R_k}{\partial \lambda_k} = 0, \quad k = 1, 2, \dots, N. \quad (17)$$

Since the λ_k is slope of the rate distortion (R-D) curve of the k th CTU, λ_k can be expressed as

$$\lambda_k = -\frac{\partial D_k}{\partial R_k} = -\frac{\partial D_k}{\partial \lambda_k} \bigg/ \frac{\partial R_k}{\partial \lambda_k}, \quad k = 1, 2, \dots, N. \quad (18)$$

Therefore, (17) is solved as

$$\lambda_k = \frac{\lambda}{\omega_k}, \quad k = 1, 2, \dots, N. \quad (19)$$

Eq. (19) indicates that the λ_k of various CTUs should be inversely proportional to its ω_k to optimize the R-D performance. In addition, since ω_k can reflect the perceptual quality of the 360-degree video, the proposed algorithm is able to improve both the objective and subjective qualities significantly.

In addition to the constraint shown in (19), the sum of the target bits of all the CTUs should be equal to the picture level target bits,

$$\sum_{k=1}^N R_k = R_t. \quad (20)$$

Moreover, in the λ -domain rate control algorithm, the λ_k follows a hyperbolic relationship with R_k ,

$$\lambda_k = \alpha_k R_k^{\beta_k}, \quad k = 1, 2, \dots, N. \quad (21)$$

Through combining (19), (20), and (21), we can derive the following equation as

$$\sum_{k=1}^N \left(\frac{\lambda}{\alpha_k \omega_k} \right)^{\frac{1}{\beta_k}} = R_t. \quad (22)$$

In (22), there is only one unknown parameter λ . However, it is difficult for us to obtain the analytic solution of λ since the β_k is a negative decimal number. In fact, (22) can be solved using numerical methods such as the bisection method. However,

since N is large for the high-resolution 360-degree videos, the bisection method will increase the encoder complexity.

In this work, we consider that the average λ of all the CTUs is approximated as the picture level λ_p . The average λ of all the CTUs is the picture level actual λ . The picture level λ_p is the calculated λ according to the picture level R - λ model,

$$\lambda_p = \alpha_p R_t^{\beta_p}, \quad (23)$$

where α_p and β_p are the picture level R - λ model parameters. As the picture level model parameters are stable, we approximate the picture level actual λ as the picture level λ_p . Note that the average of λ s is the geometric average instead of arithmetic average as we try to keep the exponential relationships between QP and λ in both picture and CTU levels,

$$\sqrt[N]{\prod_{i=1}^N \frac{\lambda}{\omega_i}} = \lambda_p. \quad (24)$$

In this way, the λ is calculated using the following equation without complex calculations,

$$\lambda = \lambda_p \sqrt[N]{\prod_{i=1}^N \omega_i}. \quad (25)$$

Therefore, the target bits of each CTU are calculated as

$$R_k = \left(\frac{\lambda_p \sqrt[N]{\prod_{i=1}^N \omega_i}}{\alpha_k \omega_k} \right)^{\frac{1}{\beta_k}}, \quad k = 1, 2, \dots, N. \quad (26)$$

As we can see from (26), the target bits are related to the content-related parameters α_k , β_k , and ω_k . The content-related characteristics are the key why the proposed CTU-level bit allocation algorithm can achieve good performance. Note that the actual bit cost of each CTU will not match exactly with the CTU level target bits. Therefore, the CTU level target bits TR_k should be related to the remained picture level target bits LR_P and calculated as

$$TR_k = R_k - \left(\sum_{j=k}^N R_j - LR_P \right) / SW, \quad k = 1, 2, \dots, N, \quad (27)$$

where SW is the slide window size that is set as 4 in our experiments. Eq. (27) indicates that we will assign a less number of bits to the current CTU when the actual bit cost of the previous coded CTUs is 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 is then calculated using (21) 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 λ s of all the CTUs. Therefore, the clipping operation is changed accordingly as follows,

$$\lambda_{clip,k} = \frac{\lambda_P \sqrt[N]{\prod_{i=1}^N \omega_i}}{\omega_k}, \quad k = 1, 2, \dots, N, \quad (28)$$

$$\lambda_k = clip(\lambda_{clip,k} \times 2^{(-\frac{2}{3})}, \lambda_{clip,k} \times 2^{\frac{2}{3}}, \lambda_k) \quad k = 1, 2, \dots, N. \quad (29)$$

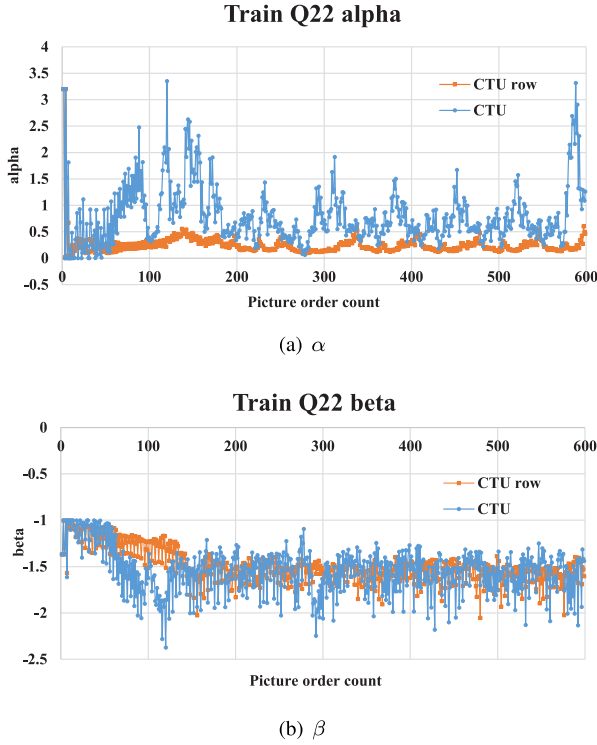


Fig. 3. The change of the model parameters α and β in the CTU level and CTU row level.

The λ_k is clipped in the desired range within an offset of 2 to avoid the quality fluctuation on the sphere. The QP_k of the k th CTU is calculated according to λ_k using the following equation [54],

$$QP_k = 4.2003 \times \ln(\lambda_k) + 13.7122, \quad k = 1, 2, \dots, N. \quad (30)$$

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

V. THE PROPOSED CTU ROW LEVEL RATE CONTROL

For the ERP format, we can see from Fig. 2(a) that all the CTUs in the same row have the same weights. From (19), 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 i th CTU. Even if the λ_i is clipped in a suitable range, it may still lead to an inaccurate bit allocation of the i th 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 are much more accurate since the video content of a CR is relatively stable. We give a typical example of the change of the model parameters α and β in the CTU level and CR level in Fig. 3. We can see that the model parameters of the CTU level change

much more seriously than that of the CR level. Fig. 3 obviously demonstrates that the model parameters in the CR level are more stable than that in the CTU level.

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: CR level λ and QP determination, CTU level λ and QP determination. As the basic process is similar to the CTU level rate control, we will just introduce these concepts briefly in the following.

A. CTU Row Level λ and QP Determination

The CR level bit allocation algorithm tries to assign the picture level target bits to each CR to minimize the picture level distortion. The problem formulation is similar to the CTU level bit allocation. Through solving a similar optimization problem, we can obtain a CR level λ constraint similar to (19).

$$\lambda_k = \frac{\lambda}{\Omega_k}. \quad (31)$$

The main difference is that ω_k which indicates the weight of the k th CTU is replaced by Ω_k which is the weight of the k th CR. The weight of the k th CR is equal to the sum of the weights of all the CTUs in the CR,

$$\Omega_k = \sum_{i=1}^{N_k} \omega_i, \quad (32)$$

where N_k is the number of CTUs in the k th 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 (27). The λ_k of the k th CR is then calculated using (21) and clipped using (28) and (29). The QP_k of the k th CR is computed using (30). Since the model parameters are much more stable in CR level compared with that in the CTU level, the CR level λ_k will be a good constraint for the CTU level rate control.

B. CTU Level λ and QP Determination

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

$$\min_{\lambda_k} \sum_{i=1}^{N_k} D_k^i, \quad \sum_{i=1}^{N_k} R_k^i \leq R_k, \quad i = 1, 2, \dots, N_k. \quad (33)$$

Different from (13), the CTU level bit allocation under the CR level rate control does not have any weight in front of the distortion since all the CTUs in a CR have the same importance. Using the method of Lagrange multipliers, (33) is solved as

$$\lambda_k^i = \lambda_k, \quad i = 1, 2, \dots, N_k. \quad (34)$$

Then the target bits R_k^i of the i th CTU in the k th CR are calculated as

$$R_k^i = \left(\frac{\lambda_k}{\alpha_k^i} \right)^{\frac{1}{\beta_k^i}}, \quad i = 1, 2, \dots, N_k, \quad (35)$$

Algorithm 1: CTU Row Level Rate Control.

- 1: **Input:** Picture level target bits R_t , Picture level λ_p ,
CTU row level α_k and β_k , CTU level α_k^i and β_k^i
- 2: **Output:** CTU level λ_k^i , CTU level QP_k^i
- 3: **CTU row level λ_k and QP_k determination**
- 4: $\Omega_k \leftarrow \omega_i$ ▷Eq. (32)
- 5: $R_k \leftarrow \Omega_k, \alpha_k, \beta_k, \lambda_p$ ▷Eq. (26)
- 6: $TR_k \leftarrow R_k, LR_p, SW$ ▷Eq. (27)
- 7: $\lambda_k \leftarrow TR_k, \alpha_k, \beta_k$ ▷Eq. (21)
- 8: λ_k clip operation ▷Eq. (28)(29)
- 9: $QP_k \leftarrow \lambda_k$ ▷Eq. (30)
- 10: **CTU level λ_k^i and QP_k^i determination**
- 11: $R_k^i \leftarrow \lambda_k, \alpha_k^i, \beta_k^i$ ▷Eq. (35)
- 12: $\lambda_k^i \leftarrow R_k^i, \alpha_k^i, \beta_k^i$ ▷Eq. (21)
- 13: λ_k^i clip operation ▷Eq. (36)
- 14: $QP_k^i \leftarrow \lambda_k^i$ ▷Eq. (30)
- 15: QP_k^i clip operation ▷Eq. (37)

where the α_k^i and β_k^i are the model parameters of the i th CTU of the k th CR. Similar to the CTU level rate control, the real target bits of each CTU are related to the remained bits in each CR and calculated using a similar equation as (27). The λ_k^i is then calculated using (21). 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 fluctuations of the CTUs with the same weights,

$$\lambda_k^i = \text{clip}(\lambda_k \times 2^{(-\frac{1}{3})}, \lambda_k \times 2^{\frac{1}{3}}, \lambda_k^i). \quad (36)$$

The QP_k^i is then computed using (30). The QP_k^i is clipped using

$$QP_k^i = \text{clip}(QP_k - 1, QP_k + 1, QP_k^i). \quad (37)$$

The RDO process of each CTU is performed based on the λ_k^i and QP_k^i . A detailed description of the proposed CR level rate control method is shown in Algorithm 1 for a better illustration.

VI. EXPERIMENTAL RESULTS

A. Simulation Setup

The proposed algorithms are implemented in the HEVC reference software HM16.20 [55] to compare with HEVC test Model (HM) anchor with and without rate control. We also compare the proposed algorithms with the state-of-the-art 360-degree video rate control methods in [21] and [52]. To better compare the proposed rate control algorithms with the HM anchor, we generate the target bitrates 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 bitrates 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 exactly the same. Therefore, the Bjontegaard Delta rate (BD-rate) [56] is used to measure the performance. The quality metrics WS-PSNR, S-PSNR-NN, S-PSNR-I, CPP-PSNR that are more appropriate for the 360-degree video are used as the quality measurements.

TABLE II
RESOLUTIONS FOR DIFFERENT PROJECTIONS

Format	8K/6K		4K		Bit Depth
	width	height	width	height	
ERP	4096	2048	3328	1664	10
CCMP	3840	2560	2880	1920	10
COHP1	2688	3112	2176	2520	10
COHP2	6224	1344	5040	1088	10

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 and low delay main10 configurations to demonstrate the effectiveness of the proposed algorithms.

We test all the 360-degree videos defined in [3] and [57] to validate the performance of the proposed algorithms. As shown in Fig. 1, the 8 K, 6 K, or 4 K video source is first projected to the 2D video formats for compression. We test the ERP, compact CMP (CCMP), and two compact OHP (COHP1 and COHP2) projection formats to demonstrate the benefits of the proposed algorithms. The detailed resolutions of the projected 2D videos are shown in Table II. 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 examples of bit cost per picture, R-D curves, and the subjective improvements to further demonstrate the benefits of the proposed algorithms.

B. Performance of the Proposed CTU Level Rate Control

1) *R-D Performance of the ERP Format:* Table III and Table IV show the R-D performance of the proposed CTU level bit allocation algorithm compared with the HM with and without rate control in random access main10 and low delay main10, respectively. From Table III, we can see that the proposed algorithm can achieve an average of 3.1%, 3.1%, 3.1%, and 3.1% bitrate savings in random access main10 case compared with the HM with rate control for the WS-PSNR, S-PSNR-NN, S-PSNR-I, and CPP-PSNR, respectively. Compared with the HM anchor without rate control, the proposed algorithm suffers about 8.5% performance losses accordingly. From Table IV, we can see that the proposed algorithm can achieve 5.5%, 5.5%, 5.5%, and 5.5% R-D performance improvements on average in low delay main10 case compared with the HM with rate control for the WS-PSNR, S-PSNR-NN, S-PSNR-I, and CPP-PSNR, respectively. Compared with the HM anchor without rate control, the proposed algorithm can lead to about 5.0% performance improvements accordingly. The experimental results obviously demonstrate that the proposed CTU level rate control algorithm can lead to significant bitrate savings compared with the HM with rate control. In addition, the experimental results show that the performance improvements brought by the proposed algorithm are consistent under different quality metrics. In the following, we will only show the performance of the proposed algorithms under the WS-PSNR metric. Furthermore, we can see that the encoder and decoder complexities of the proposed algorithm are similar to the HM with and without rate control.

TABLE III
THE PERFORMANCE OF THE PROPOSED CTU LEVEL RATE CONTROL PROBLEM FOR THE ERP FORMAT IN RANDOM ACCESS MAIN10 CASE

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	0.0%	0.1%	0.1%	0.0%	40.6%	40.7%	40.8%	40.7%
	SkateboardingTrick	-0.7%	-0.7%	-0.7%	-0.7%	1.8%	1.7%	1.7%	1.7%
	SkateboardingInLot	-6.1%	-6.1%	-6.1%	-6.1%	-1.0%	-1.0%	-1.1%	-1.1%
	ChairLift	-5.6%	-5.7%	-5.7%	-5.7%	-0.9%	-0.9%	-1.0%	-1.0%
	KiteFlite	-2.7%	-2.6%	-2.7%	-2.7%	9.1%	9.1%	9.2%	9.2%
	Harbor	-1.0%	-1.1%	-1.1%	-1.1%	18.2%	18.1%	18.2%	18.2%
	Trolley	-6.9%	-6.8%	-6.9%	-6.9%	20.8%	20.9%	20.9%	20.9%
6K	GasLamp	-4.5%	-4.5%	-4.5%	-4.5%	30.2%	30.0%	30.0%	30.3%
	Balboa	-1.1%	-1.0%	-1.0%	-1.0%	0.7%	0.8%	0.7%	0.6%
	Broadway	1.6%	1.7%	1.7%	1.7%	4.8%	4.9%	4.8%	4.7%
	Landing2	-0.8%	-0.7%	-0.8%	-0.8%	-1.9%	-1.9%	-2.1%	-2.1%
4K	BranCastle2	-3.8%	-3.8%	-3.8%	-3.8%	1.7%	1.7%	1.6%	1.6%
	PoleVault	-5.4%	-5.4%	-5.4%	-5.4%	0.5%	0.5%	0.3%	0.3%
	AerialCity	-5.5%	-5.6%	-5.6%	-5.6%	9.2%	9.2%	9.2%	9.2%
	DrivingInCity	-0.3%	-0.3%	-0.2%	-0.2%	5.4%	5.4%	5.3%	5.3%
	DrivingInCountry	-7.3%	-7.2%	-7.1%	-7.2%	-4.0%	-3.9%	-3.9%	-4.0%
Average		-3.1%	-3.1%	-3.1%	-3.1%	8.5%	8.5%	8.4%	8.4%
Enc. Time		102%				102%			
Dec. Time		101%				102%			

TABLE IV
THE PERFORMANCE OF THE PROPOSED CTU LEVEL RATE CONTROL PROBLEM FOR THE ERP FORMAT IN LOW DELAY MAIN10 CASE

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	-2.1%	-2.1%	-2.2%	-2.2%	-29.1%	-29.1%	-29.2%	-29.2%
	SkateboardingTrick	-2.3%	-2.3%	-2.3%	-2.3%	-7.1%	-7.2%	-7.2%	-7.1%
	SkateboardingInLot	-8.8%	-8.9%	-8.9%	-8.9%	-6.8%	-6.9%	-6.9%	-6.9%
	ChairLift	-9.6%	-9.7%	-9.7%	-9.6%	-9.1%	-9.1%	-9.1%	-9.1%
	KiteFlite	-7.2%	-7.1%	-7.2%	-7.3%	-3.6%	-3.6%	-3.6%	-3.7%
	Harbor	-6.4%	-6.4%	-6.5%	-6.3%	1.2%	1.2%	1.2%	1.4%
	Trolley	-2.7%	-2.6%	-2.6%	-2.7%	-0.3%	-0.1%	-0.2%	-0.3%
6K	GasLamp	-6.0%	-6.1%	-6.2%	-6.0%	13.8%	13.8%	13.8%	13.8%
	Balboa	-1.2%	-1.1%	-1.2%	-1.2%	-2.6%	-2.5%	-2.5%	-2.6%
	Broadway	0.5%	0.7%	0.7%	0.5%	1.7%	1.8%	1.8%	1.6%
	Landing2	-1.2%	-1.2%	-1.3%	-1.3%	-6.9%	-6.9%	-6.9%	-6.9%
4K	BranCastle2	-5.7%	-5.7%	-5.7%	-5.7%	-3.2%	-3.3%	-3.3%	-3.3%
	PoleVault	-10.6%	-10.5%	-10.5%	-10.6%	-8.9%	-8.8%	-8.8%	-9.1%
	AerialCity	-9.7%	-9.7%	-9.8%	-9.7%	-2.9%	-3.0%	-3.1%	-3.1%
	DrivingInCity	-1.0%	-1.0%	-1.0%	-1.0%	-0.9%	-1.0%	-1.0%	-1.0%
4K	DrivingInCountry	-14.0%	-14.0%	-13.9%	-14.0%	-14.8%	-14.8%	-14.8%	-14.8%
Average		-5.5%	-5.5%	-5.5%	-5.5%	-5.0%	-4.9%	-5.0%	-5.0%
Enc. Time		102%				101%			
Dec. Time		100%				102%			

In Table III and Table IV, we can see that the proposed CTU level bit allocation algorithm can achieve obvious performance improvements for 90% test sequences compared with the HM with rate control. However, the proposed algorithm suffers 1.6% and 0.5% performance losses for the sequence Broadway in the random access main10 and low delay main10 cases, respectively. Compared with the HM without rate control, the performance loss of the sequence Train can be as high as 40.6%. 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 stationary characteristic of the first part results in many skip coding blocks that lead

to inaccurate CTU level models. The inaccurate models have significant influences on the following bit allocation and rate control algorithms. This problem can be partially alleviated by the proposed CR level rate control algorithm that will be illustrated in the next subsection.

2) *Performance of the Other Projection Formats:* Table V and Table VI show 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 in random access main10 and low delay main10 cases, respectively. From Table V, we can see that, compared with the HM with rate control, the proposed algorithm achieves an average of 3.6%, 2.8%, and 2.7% bitrate savings in random access main10 case for the CCMP, COHP1, and COHP2 formats, respectively. From Table VI, we can see that the proposed algorithm can lead to 3.6%, 2.8%, and 3.7% R-D performance improvements on average in low delay main10 case for the CCMP, COHP1,

TABLE V
THE WS-PSNR PERFORMANCE OF THE PROPOSED CTU LEVEL RATE CONTROL PROBLEM FOR THE CMP, COHP1, AND COHP2 FORMATS IN RANDOM ACCESS MAIN10 CASE

Class	Sequence	HM with rate control as anchor			HM without rate control as anchor		
		CCMP	COHP1	COHP2	CCMP	COHP1	COHP2
8K	Train	-4.7%	-0.3%	-5.6%	28.6%	27.5%	26.8%
	SkateboardingTrick	-4.0%	-2.6%	-1.1%	-1.9%	0.8%	0.6%
	SkateboardingInLot	-3.7%	-2.9%	-3.1%	3.8%	4.6%	4.1%
	ChairLift	-2.5%	-2.4%	-3.0%	6.3%	5.2%	4.8%
	KiteFlite	-2.0%	-1.1%	-2.4%	7.9%	9.3%	7.4%
	Harbor	-2.5%	-2.4%	-1.1%	14.2%	15.1%	14.3%
	Trolley	-4.6%	-4.8%	-4.1%	19.2%	21.6%	20.4%
6K	GasLamp	-3.6%	-2.4%	-0.3%	29.0%	32.2%	28.6%
	Balboa	-3.7%	-3.5%	-3.8%	1.1%	0.2%	1.2%
	Broadway	-4.4%	-3.5%	-2.7%	1.9%	1.9%	3.0%
	Landing2	-3.4%	-2.2%	-3.4%	-3.4%	-1.7%	-3.3%
4K	BranCastle2	-4.8%	-3.8%	-3.0%	3.6%	3.9%	3.2%
	PoleVault	-3.8%	-3.2%	-2.5%	3.1%	3.8%	4.4%
	AerialCity	-3.7%	-2.8%	-1.8%	8.1%	7.6%	6.9%
	DrivingInCity	-2.7%	-3.3%	-2.3%	4.0%	3.5%	3.7%
	DrivingInCountry	-3.9%	-3.0%	-2.6%	1.5%	1.2%	1.6%
	Average	-3.6%	-2.8%	-2.7%	7.9%	8.5%	8.0%
	Enc. Time	101%	98%	100%	102%	103%	101%
	Dec. Time	102%	99%	98%	104%	104%	99%

TABLE VI
THE WS-PSNR PERFORMANCE OF THE PROPOSED CTU LEVEL RATE CONTROL PROBLEM FOR THE CMP, COHP1, AND COHP2 FORMATS IN LOW DELAY MAIN10 CASE

Class	Sequence	HM with rate control as anchor			HM without rate control as anchor		
		CCMP	COHP1	COHP2	CCMP	COHP1	COHP2
8K	Train	4.0%	-0.7%	-6.4%	-27.5%	-28.7%	-38.0%
	SkateboardingTrick	-6.0%	-3.4%	-0.7%	-10.8%	-7.1%	-5.6%
	SkateboardingInLot	-3.9%	-1.4%	-4.1%	-0.2%	3.2%	0.1%
	ChairLift	-4.8%	-4.0%	-6.0%	-4.2%	-3.5%	-5.4%
	KiteFlite	-2.6%	-0.7%	-5.2%	-1.4%	0.4%	-4.7%
	Harbor	-2.3%	-3.4%	1.4%	-4.3%	-3.7%	1.1%
	Trolley	0.6%	4.0%	-2.9%	-2.3%	-2.1%	-6.0%
6K	GasLamp	-3.2%	-1.8%	-4.3%	12.5%	11.5%	7.1%
	Balboa	-4.5%	-3.9%	-4.0%	-3.1%	-3.4%	-1.8%
	Broadway	-4.5%	-2.8%	-3.1%	-2.2%	-0.9%	0.2%
	Landing2	-4.1%	-1.8%	-3.5%	-7.4%	-5.8%	-7.0%
4K	BranCastle2	-4.7%	-3.5%	-2.9%	-0.7%	0.2%	0.1%
	PoleVault	-5.9%	-5.4%	-4.3%	-4.9%	-4.1%	-3.0%
	AerialCity	-6.6%	-5.3%	-4.6%	-1.8%	-1.1%	-1.3%
	DrivingInCity	-3.0%	-4.6%	-6.1%	-2.7%	-4.2%	-6.1%
	DrivingInCountry	-6.6%	-5.4%	-3.0%	-6.7%	-6.0%	-3.6%
	Average	-3.6%	-2.8%	-3.7%	-4.2%	-3.5%	-4.6%
	Enc. Time	101%	101%	99%	102%	101%	101%
	Dec. Time	100%	103%	93%	103%	103%	101%

and COHP2, respectively. The proposed algorithm achieves consistent bitrate savings for almost all the test sequences under all projection formats. The proposed algorithm obviously demonstrate the effectiveness of the proposed CTU level rate control algorithm. Compared with the HM without rate control, the proposed algorithm suffers about 8% performance losses in random access main10 case while it achieves about 4% performance improvement in low delay main10 case. In terms of the complexity, the proposed algorithm leads to a similar encoding/decoding time compared with HM with and without rate control.

3) *Bitrate Accuracy of the Proposed CTU Level Rate Control:* Table VII and Table VIII show the bitrate accuracy of the proposed CTU level rate control method compared with the HM16.20 rate control in random access main10 and low delay

TABLE VII
THE BITRATE ACCURACY COMPARISON OF THE CTU LEVEL RATE CONTROL AND HM WITH RATE CONTROL IN RANDOM ACCESS MAIN10 CASE

Projection	HM with rate control		CTU level bit allocation	
	Average	Maximum	Average	Maximum
ERP	0.87%	7.89%	0.78%	5.47%
CCMP	0.94%	7.09%	0.79%	6.45%
COHP1	0.84%	7.10%	0.80%	5.85%
COHP2	0.90%	6.95%	0.71%	5.25%

main10 cases, respectively. We can see that the proposed CTU level bit allocation algorithm achieves both smaller average bitrate error and maximum bitrate error compared with the original rate control method in HM for ERP, CCMP, COHP1, and COHP2 formats, respectively. The experimental results

TABLE VIII

THE BITRATE ACCURACY COMPARISON OF THE CTU LEVEL RATE CONTROL AND HM WITH RATE CONTROL IN LOW DELAY MAIN10 CASE

Projection	HM with rate control		CTU level bit allocation	
	Average	Maximum	Average	Maximum
ERP	0.17%	2.25%	0.07%	0.49%
CCMP	0.05%	0.37%	0.04%	0.37%
COHP1	0.08%	0.76%	0.05%	0.40%
COHP2	0.05%	0.34%	0.04%	0.34%

TABLE IX

THE WS-PSNR PERFORMANCE OF THE PROPOSED CR LEVEL RATE CONTROL ALGORITHM FOR THE ERP FORMAT IN RANDOM ACCESS MAIN10 CASE

Class	Sequence	HM RC	CTU level	HM w/o RC
8K	Train	-8.9%	-8.9%	28.1%
	SkateboardingTrick	0.7%	1.4%	3.2%
	SkateboardingInLot	-9.4%	-1.4%	-2.4%
	ChairLift	-7.4%	-0.7%	-1.5%
	KiteFlite	-6.1%	0.2%	9.2%
	Harbor	-2.5%	-1.4%	16.5%
	Trolley	-10.5%	-3.8%	16.2%
	GasLamp	-9.3%	-5.0%	24.6%
4K	Balboa	-2.8%	-1.8%	-1.0%
	Broadway	-1.9%	-3.4%	1.2%
	Landing2	-1.6%	-0.8%	-2.7%
	BranCastle2	-6.4%	-2.7%	-1.1%
4K	PoleVault	-5.8%	-0.4%	0.0%
	AerialCity	-6.5%	-1.0%	8.1%
	DrivingInCity	-0.8%	-0.5%	4.9%
	DrivingInCountry	-7.9%	-0.7%	-4.6%
Average		-5.0%	-1.9%	6.2%
Enc. Time		100%	98%	100%
Dec. Time		100%	99%	101%

show consistently better average bitrate accuracy in all these projection formats. The benefits mainly come from a more reasonable number of bits assigned to each CTU that can effectively prevent the cases where the bits are unable to be achieved.

C. Performance of the CR Level Rate Control on ERP Format

Table IX shows the R-D performance of the proposed CR level rate control algorithm in terms of WS-PSNR in random access main10 case. Compared with the HM with rate control, the proposed CR level rate control algorithm is able to achieve an average of 5.0% performance improvements. Except for the sequence SkateboardingTrick that suffers a few performance losses, the proposed algorithm can bring significant performance improvements for all the other test sequences. The experimental results obviously demonstrate the effectiveness of the proposed CR level rate control algorithm.

Compared with the proposed CTU level bit allocation, the proposed CR level rate control algorithm can lead to 1.9% bitrate savings on average. For the sequence Train, the proposed CR level rate control algorithm brings 8.9% performance improvements compared with the CTU level bit allocation. The proposed algorithm can lead to 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 is used as a good constraint for the CTU level λ , and thus the proposed CR level rate control can further improve the R-D performance. In terms of complexity, the proposed algorithm has approximately the

TABLE X

THE BITRATE ACCURACY OF THE CR LEVEL RATE CONTROL IN RANDOM ACCESS MAIN10 CASE

Test Case	Average	Maximum
HM w RC	0.87%	7.89%
CTU level	0.78%	5.47%
CR level	0.65%	4.36%

TABLE XI

THE S-PSNR PERFORMANCE OF THE PROPOSED CTU LEVEL BIT ALLOCATION AND THE METHOD IN [21] COMPARED WITH THE HM ANCHOR WITHOUT RATE CONTROL

Sequence	CTU level	Liu <i>et al.</i> [21]
Fengjing_1	-4.8%	-8.3%
Tiyu_1	-3.3%	-7.3%
Yanchanghui_2	-2.8%	-6.6%
Hangpai_2	-15.7%	-4.2%
AerialCity	-6.1%	-3.6%
DrivingInCountry	-16.4%	-4.8%
Average	-8.2%	-5.8%

same encoding and decoding time compared with the CTU level bit allocation algorithm. Table X shows the bitrate accuracy of the proposed CR level rate control algorithm. The proposed CR level rate control algorithm brings both smaller average bitrate error and maximum bitrate error compared with the CTU level bit allocation algorithm.

D. Comparison With the State-of-the-Art Rate Control for 360-Degree Video

To compare with the method in [21], we test our proposed algorithm under low delay P main configuration for low bitrates ranging from QPs 32 to 47. We test the sequences defined in the IEEE 1857 working group [58]. Table XI shows the experimental results comparison of the proposed CTU level bit allocation with the method proposed by Liu *et al.*. We can see that the proposed algorithm achieves an average of 2.4% R-D performance improvement compared with Liu *et al.*'s method. The method proposed by Liu *et al.* uses the pixel level weight that has higher precision than the CTU level weight proposed in this paper. Therefore, it leads to some performance improvements compared with the proposed method for some sequences. However, without an explicit change to the bit allocation algorithm, it may lead to some performance losses for the other sequences. On average, the proposed algorithm leads to some performance improvements compared with [21]. In addition, although no complexity is shown in Liu *et al.*'s paper, we consider that the algorithm will bring some complexity increase as it needs to calculate the sphere MSE many times during the RDO.

We show the comparison between the proposed CTU level bit allocation method and the method proposed by Li *et al.* [52] in random access main10 and low delay main10 cases in Table XII. We can see that the proposed algorithm can bring 4.2% and 4.4% R-D performance improvements on average in random access main10 and low delay main10 cases, respectively. The performance improvement is consistent for almost all the sequences since the CTU level weights are used in a more proper way under the proposed bit allocation method.

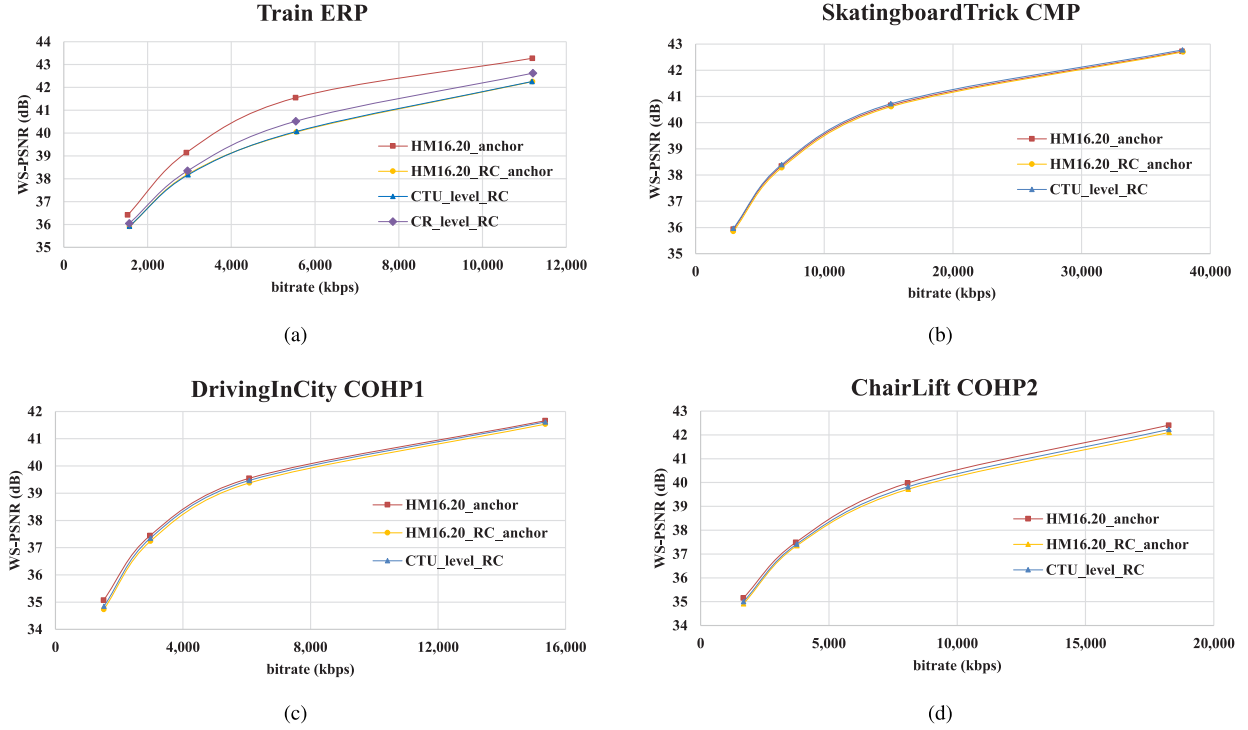


Fig. 4. Some examples of the R-D curves for various projection formats.

TABLE XII
THE WS-PSNR PERFORMANCE COMPARISON BETWEEN THE PROPOSED CTU LEVEL BIT ALLOCATION AND THE METHOD PROPOSED BY LI *et al.* [52]

Class	Sequence	Random Access	Low Delay
8K	Train	-0.3%	-3.5%
	SkateboardingTrick	-3.0%	-1.2%
	SkateboardingInLot	-6.8%	-9.2%
	ChairLift	-4.7%	-3.9%
	KiteFlite	-2.9%	-6.9%
	Harbor	-1.9%	1.9%
	Trolley	-4.9%	-4.4%
	GasLamp	-8.6%	-4.9%
4K	Balboa	-4.8%	-6.5%
	Broadway	-4.3%	-7.1%
	Landing2	-3.4%	-2.9%
	BranCastle2	-6.3%	-6.8%
4K	PoleVault	-2.5%	-5.3%
	AerialCity	-4.8%	-4.0%
	DrivingInCity	-3.2%	-2.4%
	DrivingInCountry	-4.0%	-3.6%
Average		-4.2%	-4.4%

E. R-D Curves

Some examples of R-D curves for different projection formats in random access main10 case are shown in Fig. 4. 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. We can also see that the benefits are more obvious in high bitrate case compared with that in low bitrate case. In high bitrate case, more blocks with small weights and large λ s 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 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 slow while that of the second half becomes fast. For the HM anchor 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 steady, and therefore leads to some performance losses.

F. Bit Cost Per Frame

Fig. 5 shows some examples of the bit cost per frame for various projection formats in random access main10 and low delay main10 cases. We can see from Fig. 5 that the bit cost of various pictures in a GOP follows the hierarchical coding structure in both cases. The intra frame consumes more bits compared with the inter frames. Especially, in the low delay main10 case, the difference between the picture level bit cost under the original and proposed rate control algorithms is small. The benefit of the proposed algorithm mainly comes from the proposed CTU level bit allocation algorithm.

G. Subjective Quality

Some examples of the subjective quality improvements of the proposed rate control algorithms for ERP and CMP formats are shown in Figs. 6, 7, 8, and 9. Fig. 6 and Fig. 7 are 200×200 regions cropped from the sequence DrivingInCountry, dynamic view port 0 with size 856×856 . The sub-figures (a), (b), (c),

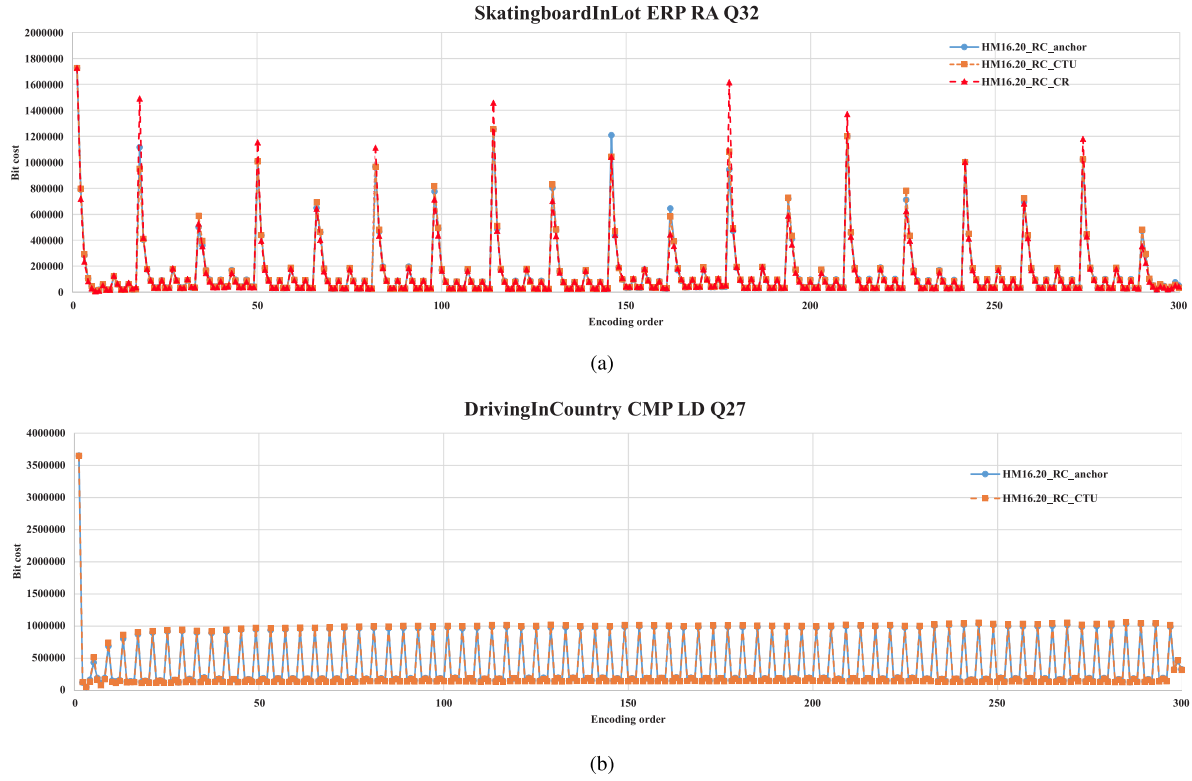


Fig. 5. Bit cost per frame for various projection formats.



Fig. 6. View port subjective improvement of the ERP format. Test sequence: DrivingInCountry. Picture order count: 187. Configuration: random access. Target bitrate: 2130 kbps. The region shown is a cropped 200×200 zone starting at (300, 500) from the dynamic view port 0 with size 856×856 . (a) original sequence; (b) HM16.20 RC (frame bit cost: 9376); (c) HM16.20 CTU level (frame bit cost: 7544); (d) HM16.20 CR level (frame bit cost: 7800).

and (d) are the original sequence, HM with rate control, the proposed CTU level rate control, and the proposed CR level rate control, respectively. From Fig. 6, we can see that the taillights are encoded with better qualities under the proposed CTU and CR level rate control algorithms compared with the HM16.20 original rate control algorithm. From Fig. 7, we can see that the edges between the car and the mountain are encoded with better qualities accordingly.

Fig. 8 and Fig. 9 are 500×500 regions cropped from the sequence SkateboardInLot, dynamic view port 0 with size 1816×1816 . The sub-figures (a), (b), and (c) are the original sequence, HM with rate control, and the proposed CTU level rate control, respectively. From Fig. 8 and Fig. 9, we can see that

the car is encoded with a better quality under the proposed CTU level rate control algorithm compared with the HM16.20 original rate control algorithm. All these results are obtained under the same target bitrate. The experimental results obviously demonstrate that the proposed algorithms are able to significantly improve the subjective qualities of the 360-degree video compression.

In addition, we show the BD-rate performance of various view ports and dynamic view ports in Table XIII as the PSNR of the view ports can partially reflect the subjective quality. Note that the view ports and dynamic view ports are not selected by ourselves but defined in the CTC. We can see that the proposed algorithms show consistently better R-D performance

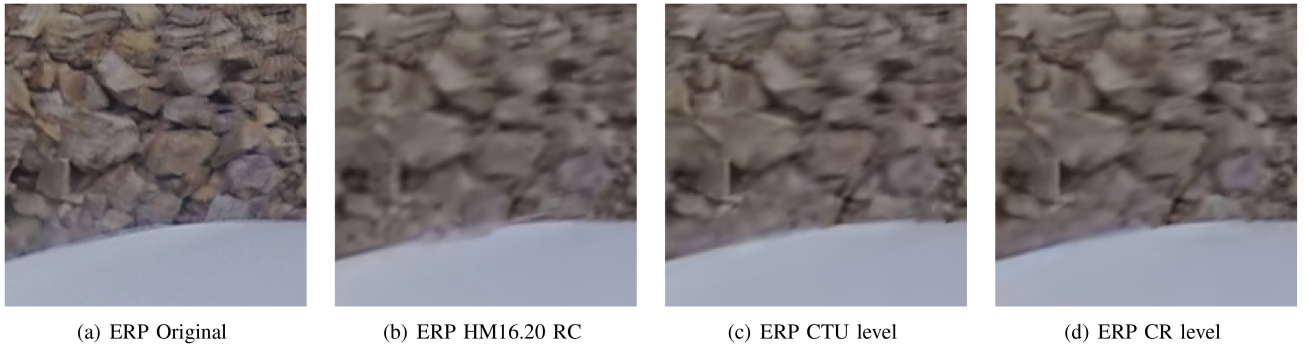


Fig. 7. View port subjective improvement of the ERP format. Test sequence: DrivingInCountry. Picture order count: 115. Configuration: random access. Target bitrate: 2130 kbps. The region shown is a cropped 200×200 zone staring at (0,600) from the dynamic view port 0 with size 856×856 . (a) original sequence; (b) HM16.20 RC (frame bit cost: 7760); (c) HM16.20 CTU level (frame bit cost: 6264); (d) HM16.20 CR level (frame bit cost: 6336).



Fig. 8. View port subjective improvement of the CMP format. Test sequence: SkateboardInLot. Picture order count: 9. Configuration: random access. Target bitrate: 2357 kbps. The region shown is a cropped 500×500 zone staring at (1200,1300) from the dynamic view port 0 with size 1816×1816 . (a) original sequence; (b) HM16.20 RC (frame bit cost: 9304); (c) HM16.20 CTU level (frame bit cost: 8696).



Fig. 9. View port subjective improvement of the CMP format. Test sequence: SkateboardInLot. Picture order count: 165. Configuration: random access. Target bitrate: 2357 kbps. The region shown is a cropped 500×500 zone staring at (1200,1200) from the dynamic view port 0 with size 1816×1816 . (a) original sequence; (b) HM16.20 RC (frame bit cost: 20424); (c) HM16.20 CTU level (frame bit cost: 20304).

TABLE XIII

THE PERFORMANCE COMPARISON BETWEEN THE PROPOSED RATE CONTROL ALGORITHMS COMPARED WITH THE HM WITH RATE CONTROL FOR THE VIEW PORTS DEFINED IN THE CTC IN RANDOM ACCESS MAIN10 CASE

Projection	VPort0	VPort1	DVPort0	DVPort1
ERP CTU	-7.2%	-3.4%	-6.9%	-3.8%
ERP CR	-8.6%	-5.2%	-8.4%	-6.2%
CCMP	-6.4%	-5.7%	-6.5%	-4.8%
COHP1	-4.2%	-1.0%	-4.4%	-3.1%
COHP2	-3.1%	-2.5%	-1.4%	-1.9%

for various view ports compared with the HM with rate control. We can also see that the proposed CR level rate control can bring better R-D performance compared with the CTU level bit allocation for the ERP format.

VII. CONCLUSION

In this paper, we propose a λ -domain perceptual rate control algorithm for the 360-degree video compression. We first propose a Coding Tree Unit (CTU) level weight for each CTU and prove that this CTU level weight is optimal to measure the importance of each CTU. Then based on the CTU level weight, we propose a CTU level bit allocation algorithm that 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 algorithms are 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 demonstrate the effectiveness of the proposed algorithm. In our future work, we will consider providing a more accurate CTU level weight to optimize the performance of the proposed algorithms. In addition, we will adapt the proposed algorithms to more projection formats especially those with some padding operations.

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, Chengdu, China, 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 Syst. Video Technol.*, *IEEE Trans.*, 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, Geneva, CH, Mar. 2019.
- [7] Y. Sun, A. Lu, and L. Yu, "Weighted-to-Spherically-Uniform Quality Evaluation for Omnidirectional Video," *IEEE Signal Process. Lett.*, 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 *Proc. IEEE Int. Symp. Mixed 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 Trans. Image Process.*, 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, Chengdu, China, 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, Geneva, CH, 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, Torino, Italy, 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 *Proc. IEEE Int. Conf. Image Process.*, Sep. 2017, pp. 1427–1431.
- [14] B. Vishwanath, T. Nanjundaswamy, and K. Rose, "Rotational motion model for temporal prediction in 360 video coding," in *Proc. IEEE 19th Int. Workshop Multimedia Signal Process. (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 Trans. Circuits Syst. Video Technol.*, 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 Trans. Circuits Syst. Video Technol.*, 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 Trans. Circuits Syst. Video Technol.*, 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 Trans. Image Process.*, 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 *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2017, pp. 709–714.
- [20] Y. Li, J. Xu, and Z. Chen, "Spherical domain rate-distortion optimization for omnidirectional video coding," *IEEE Trans. Circuits Syst. Video Technol.*, pp. 1–1, 2018.
- [21] Y. Liu, L. Yang, M. Xu, and Z. Wang, "Rate Control Schemes for Panoramic Video Coding," *J. Visual Commun. Image Represent.*, vol. 53, pp. 76–85, 2018.
- [22] D. Kwon, M. Shen, and C. J. 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.
- [23] S. Q. He, Z. J. Deng, and C. Shi, "A novel header bits estimation scheme for H.264/AVC standard," in *Proc. Meas. Technol. Eng. Researches Ind.*, ser. Applied Mechanics and Materials, vol. 333, Jul. 2013, pp. 787–790.
- [24] H. Wang and S. Kwong, "Rate-distortion optimization of rate control for H.264 with adaptive initial quantization parameter determination," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 1, pp. 140–144, Jan. 2008.
- [25] H. Yuan, J. Liu, J. Ren, and Y. Li, "Coding modes-based frame skip avoidance scheme for low bit rate video coding," *J. Real-Time Image Process.*, vol. 9, no. 4, pp. 609–619, Dec. 2014.
- [26] Z. He, Y. K. Kim, and S. K. Mitra, "Low-delay rate control for DCT video coding via p -domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 928–940, Aug. 2001.
- [27] Z. He and S. K. Mitra, "Optimum bit allocation and accurate rate control for video coding via p -domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 10, pp. 840–849, Oct. 2002.
- [28] Y. Pitrey, Y. Serrand, M. Babel, and O. Dforges, "Rho-domain for low-complexity rate control on MPEG-4 scalable video coding," in *Proc. Tenth IEEE Int. Symp. Multimedia*, Dec. 2008, pp. 89–96.
- [29] Y. Pitrey, M. Babel, and O. Deforges, "One-pass bitrate control for MPEG-4 scalable video coding using p -Domain," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast.*, May 2009, pp. 1–5.
- [30] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.
- [31] S. Wang, S. Ma, S. Wang, D. Zhao, and W. Gao, "Rate-GOP based rate control for high efficiency video coding," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1101–1111, Dec. 2013.
- [32] W. Gao, S. Kwong, H. Yuan, and X. Wang, "DCT coefficient distribution modeling and quality dependency analysis based frame-level bit allocation for HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 139–153, Jan. 2016.
- [33] L. Li, B. Li, H. Li, and C. W. Chen, " λ -Domain optimal bit allocation algorithm for high efficiency video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 1, pp. 130–142, Jan. 2018.

- [34] S. Li, M. Xu, Z. Wang, and X. Sun, "Optimal bit allocation for CTU level rate control in HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 11, pp. 2409–2424, Nov. 2017.
- [35] H. Guo, C. Zhu, S. Li, and Y. Gao, "Optimal bit allocation at frame level for rate control in HEVC," *IEEE Trans. Broadcast.*, vol. 65, no. 2, pp. 270–281, Jun. 2019.
- [36] H. Guo, C. Zhu, M. Xu, and S. Li, "Inter-block dependency-based CTU level rate control for HEVC," *IEEE Trans. Broadcast.*, pp. 1–14, 2019.
- [37] Q. Cai, Z. Chen, D. Wu, and B. Huang, "Real-time constant objective quality video coding strategy in high efficiency video coding," *IEEE Trans. Circuits Syst. Video Technol.*, pp. 1–1, 2019.
- [38] Z. Chen and X. Pan, "An Optimized Rate Control for Low-Delay H.265/HEVC," *IEEE Trans. Image Process.*, vol. 28, no. 9, pp. 4541–4552, Sep. 2019.
- [39] W. Gao, S. Kwong, Y. Zhou, and H. Yuan, "SSIM-based game theory approach for rate-distortion optimized intra frame CTU-Level bit allocation," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 988–999, Jun. 2016.
- [40] M. Wang, K. N. Ngan, and H. Li, "An efficient frame-content based intra frame rate control for high efficiency video coding," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 896–900, Jul. 2015.
- [41] Y. Li, B. Li, D. Liu, and Z. Chen, "A convolutional neural network-based approach to rate control in HEVC intra coding," in *Proc. IEEE Visual Commun. Image Process.*, Dec. 2017, pp. 1–4.
- [42] P. Wang, C. Ni, G. Zhang, and K. Li, "R-lambda model based CTU-level rate control for intra frames in HEVC," *Multimedia Tools Appl.*, vol. 78, no. 1, pp. 125–139, Jan. 2019.
- [43] W. Gao, S. Kwong, Q. Jiang, C. Fong, P. H. W. Wong, and W. Y. F. Yuen, "Data-driven rate control for rate-distortion optimization in HEVC based on simplified effective initial QP learning," *IEEE Trans. Broadcast.*, vol. 65, no. 1, pp. 94–108, Mar. 2019.
- [44] L. Li, B. Li, D. Liu, and H. Li, " λ -Domain rate control algorithm for HEVC scalable extension," *IEEE Trans. Multimedia*, vol. 18, no. 10, pp. 2023–2039, Oct. 2016.
- [45] Y. Yang, L. Shen, H. Yang, and P. An, "A content-based rate control algorithm for screen content video coding," *J. Visual Commun. Image Representation*, vol. 60, pp. 328–338, 2019.
- [46] J. Mir, D. S. Talagala, A. Fernando, and S. S. Husain, "Improved HEVC λ -Domain rate control algorithm for HDR video," *Signal, Image Video Process.*, vol. 13, no. 3, pp. 439–445, Apr. 2019.
- [47] T. Li, L. Yu, S. Yu, and Y. Chen, "The bit allocation method based on inter-view dependency for multi-view texture video coding," in *Proc. Data Compression Conf.*, Mar. 2019, pp. 428–437.
- [48] J. Lei, X. He, H. Yuan, F. Wu, N. Ling, and C. Hou, "Region adaptive R- λ model-based rate control for depth maps coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 6, pp. 1390–1405, Jun. 2018.
- [49] H. Yuan, S. Kwong, X. Wang, W. Gao, and Y. Zhang, "Rate distortion optimized inter-view frame level bit allocation method for MV-HEVC," *IEEE Trans. Multimedia*, vol. 17, no. 12, pp. 2134–2146, Dec. 2015.
- [50] C.-C. Wang and C.-W. Tang, "Region-based rate control for 3D-HEVC based texture video coding," *J. Visual Commun. Image Represent.*, vol. 54, pp. 108–122, 2018.
- [51] H. Yuan, S. Kwong, C. Ge, X. Wang, and Y. Zhang, "Interview rate distortion analysis-based coarse to fine bit allocation algorithm for 3-D video coding," *IEEE Trans. Broadcast.*, vol. 60, no. 4, pp. 614–625, Dec. 2014.
- [52] B. Li, L. Song, R. Xie, and W. Zhang, "Weight-based bit allocation scheme for VR videos in HEVC," in *Proc. IEEE Visual Commun. Image Process.*, Dec. 2017, pp. 1–4.
- [53] M. Tang, Y. Zhang, J. Wen, and S. Yang, "Optimized video coding for omnidirectional videos," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2017, pp. 799–804.
- [54] B. Li, J. Xu, D. Zhang, and H. Li, "QP refinement according to lagrange multiplier for high efficiency video coding," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS2013)*, May 2013, pp. 477–480.
- [55] "High Efficiency Video Coding test model, HM-16.20," https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/, accessed: 2019.
- [56] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-Curves," Document VCEG-M33, Austin, Texas, USA, Apr. 2001.
- [57] P. Hanhart, J. Boyce, K. Choi, and J.-L. Lin, "JVET Common Test Conditions and Evaluation Procedures for 360 Video," Document JVET-L1012, Macao, China, Oct. 2018.
- [58] "IEEE 1857 working group," <http://www.ieee1857.org>, accessed: 2019.



Li Li (M'17) received the B.S. and Ph.D. degrees in electronic engineering from the University of Science and Technology of China, Hefei, China, in 2011 and 2016, respectively. He is currently a Visiting Assistant Professor with the University of Missouri-Kansas City, Kansas City, MO, USA.

His research interests include image/video coding and processing. He received the Best 10% Paper Award at the 2016 IEEE Visual Communications and Image Processing (VCIP) and the 2019 IEEE International Conference on Image Processing (ICIP).



Ning Yan received the B.S. degree in information engineering from the China University of China University of Mining and Technology, Xuzhou, China, in 2015. He is currently working toward the Ph.D. degree in electronic engineering and information science with the University of Science and Technology of China, Hefei, China.

His research interests include video coding/processing and machine learning.



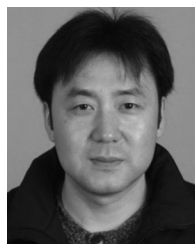
Zhu Li (M'02–SM'07) received the Ph.D. degree in electrical and computer engineering from Northwestern University, Evanston, IL, USA, in 2004. He is an Associated Professor with the Department of CSEE, University of Missouri, Kansas City, MO, USA, directs the NSF I/UCRC Center for Big Learning at UMKC. He was with the AFRL Summer Faculty, US Air Force Academy, UAV Research Center, 2016, 2017, and 2018, respectively. He was a Senior Staff Researcher/Senior Manager with Samsung Research America's Multimedia Core Standards Research Lab in Dallas, from 2012 to 2015, a Senior Staff Researcher with FutureWei, from 2010 to 2012, an Assistant Professor with the Department of Computing, The Hong Kong Polytechnic University from 2008 to 2010, and a Principal Staff Research Engineer with the Multimedia Research Lab, Motorola Labs, Schaumburg, Illinois, from 2000 to 2008. His research interests include image/video analysis, compression, and communication and associated optimization and machine learning problems.

He has 46 issued or pending patents, more than 100 publications in book chapters, journals, conference proceedings and standards contributions in these areas. He is an Associate Editor-in-Chief (AEiC) for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, since 2020, and served and serving as an Associated Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING (since 2019), IEEE TRANSACTIONS ON MULTIMEDIA (2015–2019), and IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (2016–2019). He received a Best Paper Award from IEEE International Conference on Multimedia & Expo (ICME) at Toronto, 2006, and a Best Paper Award from IEEE International Conference on Image Processing (ICIP) at San Antonio, 2007.



Shan Liu received the B.Eng. degree in electronics engineering from Tsinghua University, Beijing, China, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, CA, USA. She is a Tencent Distinguished Scientist and a General Manager of Tencent Media Lab. Prior to joining Tencent, she was the Chief Scientist and the Head of America Media Lab, Futurewei Technologies. She was formerly Director of Multimedia Technology Division with MediaTek USA. She was also formerly with MERL, Sony, and

IBM. She has been actively contributing to international standards since the last decade and has numerous proposed technologies adopted into various standards. She holds more than 100 granted US and global patents, some of which have been productized to serve millions of users daily. She served the Industrial Relationship Committee of IEEE Signal Processing Society 2014–2015 and the VP of Industrial Relations and Development of Asia-Pacific Signal and Information Processing Association (APSIPA) 2016–2017. She was named APSIPA Industrial Distinguished Leader in 2018.



Houqiang Li (M'10–SM'12) received the B.S., M.Eng., and Ph.D. degrees in electronic engineering in 1992, 1997, and 2000, respectively, from the University of Science and Technology of China, Hefei, China, where he is currently a Professor with the Department of Electronic Engineering and Information Science.

His research interests include video coding and communication, multimedia search, image/video analysis. He has authored or coauthored more than 100 papers in journals and conferences. He was an

Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY from 2010 to 2013, and has been with the Editorial Board of the *Journal of Multimedia* since 2009. He was the recipient of the Best Paper Award for Visual Communications and Image Processing (VCIP) in 2012, the recipient of the Best Paper Award for International Conference on Internet Multimedia Computing and Service (ICIMCS) in 2012, the recipient of the Best Paper Award for the International Conference on Mobile and Ubiquitous Multimedia from ACM (ACM MUM) in 2011, and a senior author of the Best Student Paper of the 5th International Mobile Multimedia Communications Conference (MobiMedia) in 2009.