Rate Control for Video-Based Point Cloud Compression

Li Li, *Member, IEEE*, Zhu Li, *Senior Member, IEEE*, Shan Liu, *Senior Member, IEEE*, Houqiang Li, *Senior Member, IEEE*

***Abstract*— Rate control is a necessary tool for video-based point cloud compression (V-PCC). However, there is no solution specified on this topic yet. In this paper, we propose the first rate control algorithm for V-PCC. Generally, a rate control algorithm is divided into two processes: bit allocation and bitrate control. In V-PCC, the total bits are composed of three parts: the header information including the auxiliary information and occupancy map, the geometry video, and the attribute video. The bit allocation aims to assign the total bits to these three parts. Since the auxiliary information and occupancy map are encoded losslessly, the bit cost of the header information is fixed. Therefore, we only need to assign bits between the geometry and attribute videos. Our first key contribution is the proposed video-level bit allocation algorithm between the geometry and attribute videos to optimize the overall reconstructed point cloud quality. Then we assign geometry and attribute video bits to each group of pictures (GOP), each frame, and each basic unit (BU). Our second key contribution is that we assign zero bits to the BUs with only unoccupied pixels. The unoccupied pixels are useless for the reconstructed quality of the point cloud and therefore should be assigned zero bits. In the bitrate control process, the encoding parameters are determined, and the model parameters are updated for each frame and BU to achieve the target bits. Our third key contribution is that we propose a BU-level model updating scheme to handle the case where various patches may be placed in different positions in neighboring frames. We use the auxiliary information to find the corresponding BU in the previous frame and apply its model parameters to the current BU. The proposed algorithms are implemented in the V-PCC and High Efficiency Video Coding (HEVC) reference software. The experimental results show that the proposed rate control algorithm can achieve significant bitrate savings compared with the state-of-the-art method.**

***Index Terms*— Bit allocation, high efficiency video coding, point cloud compression, rate control, video-based point cloud compression.**

Manuscript received July 28, 2019; revised February 16, 2020 and March 24, 2020; accepted April 14, 2020. Date of publication April 28, 2020; date of current version May 8, 2020. This work was partially supported by NSF grant 1747751 and a gift grant from Tencent America. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yui-Lam Chan. *(Corresponding author: Zhu Li.)*

Li Li and Zhu Li are with the Department of Computer Science and Electrical Engineering, University of Missouri–Kansas City, Kansas City, MO 64110 USA (e-mail: lil1@umkc.edu; lizhu@umkc.edu).

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

Houqiang Li is with the Chinese Academy of Sciences Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, University of Science and Technology of China, Hefei 230027, China (e-mail: lihq@ustc.edu.cn).

Digital Object Identifier 10.1109/TIP.2020.2989576

I. Introduction

A POINT cloud is a set of points in 3D space that can be used to represent a 3D surface. Not only the geometry information but also each point contains some specific attributes, such as colors and material reflection. The capability of the point cloud for recovering 3D objects makes it promising for extended virtual-reality applications such as 3D immersive telepresence [1] and virtual-reality viewing with interactive parallax [2]. However, the high data rate of the point cloud prevents the adoption of this media format. For example, for a typical dynamic point cloud (DPC) captured by 8i with 30 frames per second [3], each frame usually has approximately one million points. If 30 and 24 bits are used to represent the geometry and attribute of the point cloud, the bitrate of the DPC can be as high as $180 Mbytes$ per second without compression. The Moving Pictures Experts Group Immersive media working group (MPEG-I) is currently working on a video-based point cloud compression (V-PCC) standard [4] utilizing the existing video coding technologies to solve this problem.

Briefly, V-PCC projects the DPC to the geometry and attribute videos and encodes these videos using video compression standards such as High Efficiency Video Coding (HEVC) [5]. Some header bits indicating the occupancy map and auxiliary information are also transmitted in addition to the geometry and attribute videos. The occupancy map indicates whether a pixel in the geometry or attribute video corresponds to a valid 3D point. The auxiliary information together with the geometry video constructs the complete 3D geometry. V-PCC is already in the committee draft stage [6]. However, no rate control scheme has yet been designed for V-PCC. Similar to video compression standards such as HEVC, V-PCC can be applied to scenarios that have a limited storage size or communication bandwidth [7]. In these scenarios, rate control is needed to control the bitrate as well as optimizing the quality of the compressed video or point cloud such that the size or bandwidth limitation can be met properly [8], [9]. Therefore, rate control is a necessary tool for V-PCC for point cloud transmission and storage. In this paper, we propose the first rate control scheme for V-PCC.

Generally, a rate control scheme can be divided into two processes: bit allocation and bitrate control. The bit allocation is responsible for assigning the total bits to each subunit such as video, frame, and basic unit (BU). The bitrate control aims to achieve the assigned bits for each subunit. In V-PCC, the total bits are composed of three parts: the header information

including the auxiliary information and occupancy map, the geometry video, and the attribute video. The auxiliary information is encoded losslessly. The occupancy map is first downsampled and then encoded losslessly. Because of this, the bit cost of the header information is fixed. Therefore, the first step of the bit allocation focuses on the bit allocation among the geometry video and attribute video. This bit allocation step is called as video-level bit allocation in the following sections. The current bit allocation algorithms on video compression mainly focus on picture-level and BU-level bit allocation [10], [11]. The picture-level and BU-level bit allocation algorithms take the inter-frame or inter-BU dependency into consideration to optimize the overall video or frame quality under the target bitrate. The influence of the geometry and attribute videos on the quality of the reconstructed point cloud, which is the key to video-level bit allocation, is not considered.

After video-level bit allocation, we then adapt the existing rate control mechanisms designed for general videos to geometry and attribute videos projected from the point cloud. The state-of-the-art rate control algorithm for general videos is the λ-domain rate control algorithm [12]. The λ-domain rate control algorithm considers the Lagrange Multiplier λ as the key factor for determining the bitrate, and proposes a hyperbolic R-λ model to characterize the relationship between bitrate R and λ. It has been adopted by both HEVC [13] and Versatile Video Coding (VVC) [14] and integrated into the reference software. However, two new characteristics of the projected geometry and attribute videos make the existing rate control algorithms difficult to apply. The first is that some unoccupied BUs existing in both geometry and attribute videos are useless for reconstructing the point cloud. The second is that corresponding BUs in neighboring frames may be placed in different positions, which affects the R-λ model accuracy.

One typical example of the projected neighboring attribute frames of the point cloud "RedAndBlack" is shown in Fig. 1. We can see that some BUs indicated by the blue squares are unoccupied. These unoccupied BUs are useless for the reconstructed point cloud quality and should be assigned zero bits. However, all the existing BU-level bit allocation schemes are designed on videos instead of DPCs. They treat all the BUs with equal importance and thus cannot solve the problem of the unoccupied BUs in V-PCC. The problem can be solved only if the effects of the unoccupied BUs are considered. Additionally, we can see from Fig. 1 that corresponding BUs in neighboring frames may be placed in different positions as indicated by the red squares. This has a significant influence on the R-λ model accuracy as we obtain the model parameters of the current BU from the co-located BU in the previous frame at the same hierarchical level. Inaccurate model parameters lead to serious rate-distortion (RD) performance losses.

To address the above problems, we propose the first rate control framework designed specified for V-PCC in this paper. The proposed framework mainly has the following key contributions.

- We propose a video-level bit allocation algorithm between the geometry and attribute videos. Specifically, we set the λ ratio between the geometry video and the attribute video inversely proportional to its influence on the reconstructed quality of the point cloud. However, an accurate estimation of the influences is difficult due to the lack of a simple scheme for quantifying the importance of geometry and attribute in various bitrate scenarios. We currently set the λ ratio with a fixed value according to some subjective observations.
- We propose assigning zero bits to the unoccupied BUs that are useless for the reconstructed quality of the point cloud. These BUs are encoded using large λs and quantization parameters (QPs) to minimize their bit cost. In addition, the RD characteristics of these BUs are not included in RD characteristics of the picture level.
- We propose estimating the R-λ model parameters of the BU with occupied pixels from the corresponding BU in the previous frame instead of the co-located BU. To be more specific, we use the auxiliary information to find the corresponding BU and obtain its model parameters. For the BU with no occupied pixels, we propose obtaining the model parameters from the BU with no occupied pixels in the previous frame.

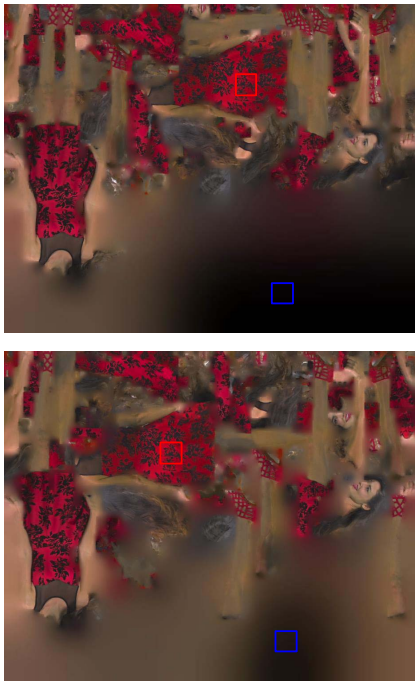


Fig. 1. Typical example of the projected neighboring attribute frames of the point cloud "RedAndBlack". The picture order counts of the top and bottom frames are 1450 and 1451, respectively.

The proposed algorithms are implemented in the V-PCC reference software [15] and the corresponding HEVC reference software (HM) [16]. The experimental results show that the proposed algorithm can achieve significant bitrate savings compared with the state-of-the-art method. Note that assigning zero bits to the unoccupied pixels can be applied whether or not rate control is desired. The experimental results of assigning zero bits to the unoccupied pixels can be found in [17].

The rest of this paper is organized as follows. In Section II, we introduce the related works on rate control and some

basic concepts of V-PCC. We introduce the proposed rate control algorithm for V-PCC in Section III. In Section IV, the experimental results are presented in detail. Section V concludes the paper.

II. Related Works

In this section, we introduce some related works on bit allocation and rate control. We also briefly introduce the V-PCC framework.

A. Rate Control

The 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. The Q-domain rate control algorithm [9], [18] considers the quantization step Q as the key factor for determining the bitrate. The ρ-main rate control algorithm [19], [20] considers the percentage of zeros among the quantized coefficients ρ as the key factor to determine the bitrate. However, Q or ρ can only determine the residue bitrate. Li *et al.* [12] noted that, along with the header bit increase in HEVC and VVC, Q or ρ can no longer determine the overall bitrate. They proposed that the Lagrange Multiplier λ is essentially the key factor in determining the overall bitrate. A corresponding hyperbolic R-λ model-based rate control algorithm is proposed and integrated into the HEVC and VVC reference software. Therefore, in this paper, we use the λ-domain rate control algorithm as the basis of the proposed rate control algorithm for V-PCC.

In addition to the rate control algorithms focusing on bitrate control, there are also many works on bit allocation. The bit allocation algorithms can be divided into two levels: picture level and BU level. The picture-level bit allocation is closely related to the encoding structure. The early bit allocation algorithms mainly deal with the simple IPPP structure. In the IPPP structure, each P frame is referenced by the immediate subsequent frame. Therefore, all the P frames are considered to have equal importance. These bit allocation algorithms consider the frame complexity to perform picture-level bit allocation. Jiang *et al.* [21] first introduced the mean absolute difference (MAD) ratio to characterize the frame complexity. They further proposed using the weighted sum of the MAD ratio and Peak-Signal-to-Noise-Ratio (PSNR) decrease to enhance the frame complexity [22]. To handle scene change, Zhou *et al.* [23] proposed using the histogram of difference frame (HOD) to describe the frame complexity.

Along with the introduction of the hierarchical-B coding structure to the video coding framework [24], the reference relationships of various frames became much more complex. An increasing number of works have focused on characterizing the inter frame dependency among various frames. Hu *et al.* [25] proposed a linear model to characterize the quality dependency and deduced a Q-domain bit allocation algorithm for H.264/Advanced Video Coding [26]. Wang *et al.* [27] extended the linear model to group of pictures (GOP)-level distortion and rate models and derived a ρ-domain bit allocation algorithm for HEVC. Gao *et al.* [28] introduced a synthesized Laplacian model to describe the Discrete Cosine Transform (DCT) coefficient distribution and introduced a picture-level bit allocation algorithm for HEVC. Under the λ-domain rate control algorithm, Li *et al.* [10] proved that the λ ratios of various pictures should be inversely proportional to their influences on the sequence to achieve the optimal performance. Gao *et al.* [29], [30] further carefully analyzed the distortion propagation in the hierarchical-B coding structure and provided a content-related bit allocation scheme. These picture-level bit allocation algorithms carefully investigated the distortion-propagation influences of various frames on the sequence. However, the picture-level distortion-propagation characteristics are different from those at the video level.

In addition to picture-level bit allocation algorithms, there are many works focusing on BU-level bit allocation. Some studies on BU-level bit allocation consider various BUs as independent units since most BUs in inter frames obtain the prediction from the previously coded frames. The most straightforward method uses the complexity of a BU as the measurement to determine its target bits. Seo *et al.* [31] introduced a combination of the variance of difference (VOD) and MAD as the complexity measurement and performed BU-level bit allocation based on this combination. However, the BU-level bit allocation should be essentially determined by the RD characteristics instead of only the distortion. Yuan *et al.* [32] proposed a linear model between the distortion and Q and derived the BU-level bit allocation based on the model. He and Mitra [33] introduced a second-order model between the distortion and ρ and deducted the BU-level bit allocation algorithm. In addition, Li *et al.* [11] proposed the λ-domain BU-level bit allocation algorithm under the constraint of picture-level target bits. Guo *et al.* [34] provided BU-level rate control by considering the inter dependency directly in the block level.

There are also some works that focused on the dependent BU-level bit allocation that considers the dependency among various BUs within one frame. The dependent BU-level bit allocation is usually considered for the intra frame instead of the inter frame. Ferguson and Allinson [35] introduced the dependent quantization to video coding and proposed the modified steepest-descent algorithm to solve the BU-level bit allocation problem. This algorithm showed significant gains on the I frames with strong dependency but only very small gains on the P frames with weak dependency. Lee and Song [36] proposed using the gradient as the complexity measurement and introduced an intra BU-level bit allocation algorithm based on the complexity. Wang *et al.* [37] extended this idea for the λ-domain rate control algorithm and applied it to HEVC intra-frame rate control. In addition, Gao *et al.* [38] attempted to solve the bit allocation using game theory and proposed optimizing the structural similarity index (SSIM) [39] instead of PSNR. However, all these BU-level bit allocation schemes are unable to solve the problem of unoccupied BUs in the V-PCC framework.

Except for the bitrate control and bit allocation algorithms, some other factors may have significant influences on the rate control such as the initial encoding parameter determination [40], [41] and the accuracy of the picture-level

TABLE I
BRIEF INTRODUCTIONS OF THE HEADER INFORMATION

Auxiliary information	Notation	Explanation
Projection plane	n	Indication of the projected plane; value range: 0, 1, 2.
2D bounding box	$(u0, v0)$ $(u1, v1)$	Top-left 2D position of the bounding box; bpr basis. Size of the 2D bounding box; bpr basis.
3D location	$\delta 0$ $s0$ $r0$	Minimum depth of the current patch. Tangential shift of the top-left position of the current patch. Bi-tangential shift of the top-left position of the current patch.
Block-to-patch	bpr BP	Block resolution; typical value: 16. Block-to-patch array; bpr basis.
Occupancy map	opr OM	Occupancy map precision; typical value: 4 or 2. Occupancy map array; or basis.

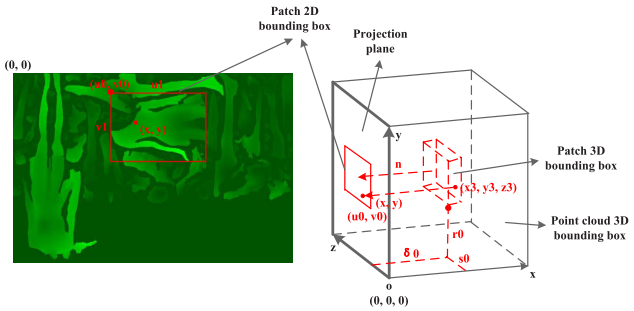


Fig. 2. Illustration of the header information.

and BU-level model parameters. For example, Li *et al.* [11] proposed calculating the model parameters from the RD characteristics from the co-located BU. Instead of estimating the model parameters from the previous BU, Chen and Pan [42] taken the distributions of the model parameters into consideration. In addition, Li *et al.* [43] proposed using the convolutional neural network (CNN) [44] to estimate the model parameters more accurately. However, these methods are unable to solve the problem that the current BU and its corresponding BU are placed in different positions.

B. Video-Based Point Cloud Compression

The conversion from a point cloud to geometry and attribute videos can be roughly divided into three stages: clustering, packing, and padding. Initial clustering of the point cloud is obtained by associating each point with the face having the most similar normal. The cluster is then refined by smoothing the generated patches. After patch generation, the patch-based projection method uses a simple packing strategy to organize the patches into videos. The patch location is determined through an exhaustive search in the raster scan order. The padding process then fills the empty space between the patches to make the generated frames more suitable for video coding. Note that the videos include geometry and attribute videos to record the geometry and attribute information.

To recover the point cloud from the compressed videos, some header information is transmitted to the decoder. An illustration of the header information is shown in Fig. 2. In Fig. 2, we give an example of the patch projected to the yoz plane. For each patch, we need to signal the index of the projected plane n, the 2D bounding box $(u0, v0, u1, v1)$, and the 3D location $(\delta 0, s0, r0)$. Additionally, for each block with size $bpr \times bpr$, the array BP indicates which patch it belongs to. For each block with size $opr \times opr$, the array OM indicates whether it is occupied. A detailed explanation of the header information is shown in Table I. Note that both the 2D bounding box and the block-to-patch information are transmitted based on bpr to save bits.

Based on the header information, from Fig. 2, we can derive the 3D-to-2D correspondence from $(x3, y3, z3)$ to (x, y) as follows,

$$\begin{cases} x = (z3 - s0) + u0 \cdot bpr \\ y = (y3 - r0) + v0 \cdot bpr \\ h(x, y) = x3 - \delta 0, \end{cases} \tag{1}$$

where $h(x, y)$ is the pixel value of position (x, y) in the geometry frame. Additionally, we can derive the 2D-to-3D correspondence as follows,

$$\begin{cases} x3 = \delta 0 + h(x, y) \\ y3 = y - v0 \cdot bpr + r0 \\ z3 = x - u0 \cdot bpr + s0. \end{cases} \tag{2}$$

These 3D-to-2D relationships are used to find the corresponding BU in the previous frame.

III. PROPOSED ALGORITHMS

As we have mentioned in Section I, a rate control algorithm can be divided into bit allocation and bitrate control. We introduce the proposed bit allocation and bitrate control algorithms in Section III-A and Section III-B, respectively.

A. Bit Allocation

Our proposed bit allocation algorithm includes four levels: video level, GOP level, picture level, and BU level.

1) Video-Level Bit Allocation: The video-level bit allocation aims to assign the total bits to the header information including auxiliary information and occupancy map, the geometry video, and the attribute video. The auxiliary information is always encoded losslessly. The occupancy map is encoded losslessly after down-sampling. Only after we determine the occupancy map resolution opr that indicates the down-sampling ratio can we obtain the bit cost of the occupancy map.

In the V-PCC common test condition (CTC), opr is set according to the bitrate. It is set to 2 and 4 in the highest bitrate

TABLE II

PERFORMANCE OF DOWN-SAMPLING THE OCCUPANCY MAP BY 4 TIMES FOR ALL THE BITRATES COMPARED WITH THE V-PCC REFERENCE SOFTWARE

Test point cloud	Geom.BD-GeomRate D1	Geom.BD-GeomRate D2	Attr.BD-AttrRate Luma	Attr.BD-AttrRate Cb	Attr.BD-AttrRate Cr
Loot	0.5%	0.8%	0.3%	0.0%	–0.3%
RedAndBlack	0.3%	0.9%	0.1%	0.4%	0.2%
Soldier	–0.2%	0.6%	0.8%	0.8%	0.9%
Queen	0.2%	0.8%	–0.5%	–1.1%	–1.1%
LongDress	0.1%	0.6%	0.4%	0.3%	0.2%
Avg.	**0.1%**	**0.7%**	**0.2%**	**0.1%**	**0.0%**

scenario r5 and all the other bitrate scenarios from r1 to r4, respectively. The setting in the CTC is in accordance with our common sense that we should use coarse and fine occupancy maps for the low and high bitrates, respectively. However, how to determine *opr* according to the target bitrate and the content of the DPC in a rate control algorithm remains a problem. As *opr* is selected between 4 and 2 in the CTC, which covers a large bitrate range, we attempt to choose an *opr* between 4 and 2 according to the target bits in the proposed rate control algorithm.

To choose a suitable *opr* for various target bits, we test 32 frames in the random access case by setting *opr* to 4 instead of 2 in the highest bitrate scenario r5 and compare the performance with the CTC in Table II. In Table II, D1 and D2 are point-to-point and point-to-plane PSNRs for the geometry, respectively. The experimental results show that setting *opr* to 4 suffers only an average of 0.1% and 0.7% losses for the geometry under D1 and D2 measurements, respectively. In terms of the attribute, it leads to 0.2%, 0.1%, and 0.0% RD performance losses for the Luma, Cb, and Cr components, respectively. The experimental results show that setting *opr* to 4 leads to almost no performance loss in the highest bitrate compared with setting *opr* to 2 for all the tested point clouds. The experimental results indicate that we can set *opr* to 4 in a large bitrate range. Only when the target bitrate is extremely high do we need to set *opr* to 2. Therefore, *opr* is determined using the following equation according to the target bits per point $TBpp_T$,

$$\begin{cases} opr = 2, & TBpp_T > 1.0 \\ opr = 4, & TBpp_T \leq 1.0. \end{cases} \tag{3}$$

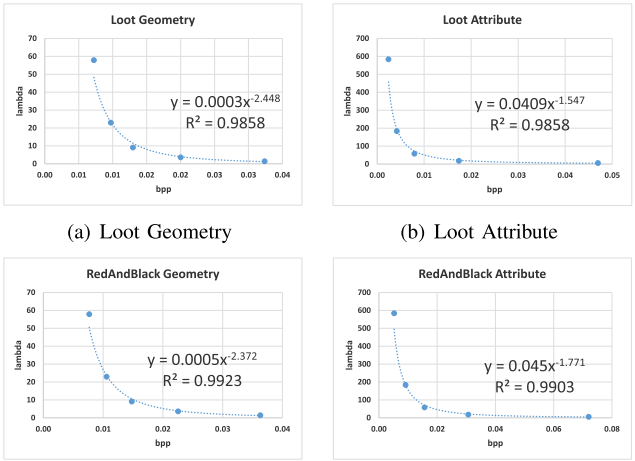
After *opr* is determined, the actual bits of the header information AR_H are fixed. The sum of the target bits of the geometry and attribute videos TR_V are calculated by subtracting AR_H from TR_T,

$$TR_V = TR_T - AR_H. \tag{4}$$

Before introducing the video-level bit allocation to assign TR_V to the geometry and attribute videos, we first validate that the hyperbolic $R-\lambda$ model proposed in [12] is still valid for the geometry and attribute videos,

$$\lambda = \alpha \cdot TR^{\beta}, \tag{5}$$

where λ and TR are the Lagrange multiplier and bitrate, respectively. α and β are the model parameters related to the video content. We test 32 frames in the random access case under the V-PCC CTC and count the bit cost TR and λ for the geometry and attribute videos. We fit the data using the hyperbolic model as shown in Fig. 3. From Fig 3, we can see that the determination coefficients R^2 of all the fitted curves are larger than 0.98. The experimental results demonstrate that the hyperbolic model is still valid for the geometry and attribute videos. Additionally, note that αs of the geometry and attribute videos show a large difference. The α of the attribute is approximately 100 times larger than that of the geometry.



(a) Loot Geometry (b) Loot Attribute

(c) RedAndBlack Geometry (d) RedAndBlack Attribute

Fig. 3. Validation of the hyperbolic R-λ relationship.

The target of video-level bit allocation is to minimize the distortion of the reconstructed point cloud under bit constraint TR_V. The distortion of the reconstructed point cloud is modeled as a weighted combination of the geometry and attribute distortions,

$$\min_{TR_G, TR_A} \omega TD_G + TD_A, \quad TR_G + TR_A \leq TR_V, \tag{6}$$

where TD_G and TD_A are the distortions for the geometry and attribute videos, respectively. TR_G and TR_A are the target bits for the geometry and attribute videos, respectively. ω indicates the relative importance of the geometry and attribute videos. Generally, geometry distortion is more important than attribute distortion. This can be partially explained by the V-PCC CTC. Under the V-PCC CTC, the geometry video is encoded with much smaller QPs compared with that of the attribute video. We give some examples of the PSNRs of the geometry and attribute videos, under the V-PCC CTC as shown in Table III. In the table, r1 to r5 indicate the bitrate scenario from the low bitrate case to the high bitrate case defined in the V-PCC CTC [45]. We can see that the PSNRs of the geometry videos are 30 to 40 dB higher than those of the attribute videos. Additionally, we give some examples of the subjective qualities by setting different ωs, as shown in Fig. 4. We can see that when ω decreases, the shape of the man's head becomes more irregular, and some extra points appear near the man's hand. These examples demonstrate that subjective quality becomes worse when the geometry and attribute videos are compressed with equal importance. Therefore, ω should be larger than 1 to obtain good subjective quality.

TABLE III

SOME EXAMPLES OF THE PSNRS OF THE GEOMETRY AND ATTRIBUTE VIDEOS UNDER THE V-PCC CTC

	PSNR(dB)	r1	r2	r3	r4	r5
Loot	Geometry (D1)	66.8	68.5	69.8	70.8	72.1
	Attribute (Luma)	31.7	33.8	36.3	38.7	41.2
RedAndBlack	Geometry (D1)	66.0	67.7	69.1	70.0	71.2
	Attribute (Luma)	31.1	33.0	35.0	37.2	39.7
Soldier	Geometry (D1)	66.4	68.1	69.4	70.5	71.6
	Attribute (Luma)	29.3	31.7	34.1	36.5	38.8
Queen	Geometry (D1)	67.1	68.8	70.0	70.9	71.9
	Attribute (Luma)	30.0	31.6	33.6	35.5	37.0
LongDress	Geometry (D1)	66.3	68.0	69.4	70.4	71.5
	Attribute (Luma)	26.0	28.1	30.4	32.8	35.5

(a) Original (b) ω = 1

(c) ω = 3 (d) ω = 8

Fig. 4. Examples of the subjective qualities by setting different ωs.

The constrained problem in (6) is converted to an unconstrained problem by introducing a Lagrangian Multiplier μ,

$$\min_{TR_G, TR_A} \omega TD_G + TD_A + \mu(TR_G + TR_A). \tag{7}$$

The unconstrained problem is solved using the Lagrangian method by setting the derivatives to TR_G and TR_A as 0. Since the geometry and attribute videos are encoded independently, the unconstrained problem is solved using the following two equations,

$$\omega \cdot \frac{\partial TD_G}{\partial TR_G} + \mu = 0, \tag{8}$$

$$\frac{\partial TD_A}{\partial TR_A} + \mu = 0. \tag{9}$$

According to the RD curves of the geometry and attribute videos, we can obtain the following equations,

$$\lambda_G = -\frac{\partial TD_G}{\partial TR_G}, \quad \lambda_A = -\frac{\partial TD_A}{\partial TR_A}, \tag{10}$$

where λ_G and λ_A are the slopes of the tangent line of the RD curves of the geometry and attribute videos, respectively. By substituting (10) into (8) and (9), we can derive the solution as follows,

$$\lambda_G = \frac{\lambda_A}{\omega}. \tag{11}$$

As we can see from (11), the larger ω is, the smaller λ_G is. This leads to more bits assigned to the geometry video and fewer bits assigned to the attribute video. Therefore, the larger the ω is, the better the RD performance of the geometry is. However, this leads to worse RD performance for the attribute.

In addition to the constraint shown in (11), we have TR_V as the other constraint for the video-level bit allocation,

$$TR_G + TR_A = TR_V. \tag{12}$$

By substituting (12) and the hyperbolic model in (5) into (11), we can derive the following equation,

$$\omega \cdot \alpha_G TR_G^{\beta_G} - \alpha_A (TR_V - TR_G)^{\beta_A} = 0. \tag{13}$$

In (13), to derive TR_G, the only unknown parameters are α_G, β_G, α_A, β_A, and ω. α_G, β_G, α_A, and β_A are sequence-related parameters, which are obtained by encoding the first 32 frames of the point cloud twice using different bitrates. We choose the bitrate scenarios $r2$ and $r4$ defined in the V-PCC CTC [45] to calculate the parameters. ω is set to 8 in the overall experiment according to our experience. We analyze the influence of ωs on the performance of video-level bit allocation in the experimental results.

As shown in Fig. 3, α_G and α_A are positive values, and β_G and β_A are negative values. Therefore, the left term of (13) is a monotonic function of TR_G. We can solve this problem using the Bisection Method. After TR_G is determined, we encode the geometry video and obtain its actual bits AR_G. Then, we calculate the target bits of the attribute video TR_A as follows,

$$TR_A = TR_V - AR_G. \tag{14}$$

The target bits guide the encoding process of the attribute video.

2) GOP-Level Bit Allocation: The GOP-level bit allocation has a simple aim in making the bitstream more adaptable to the bandwidth. We follow the current method in the HM software to determine the GOP-level target bits TR_{GOP} [12],

$$TR_{GOP} = \frac{AVGR_{Pic} \cdot (N_{Coded} + SW) - R_{Coded}}{SW} \cdot N_{GOP}, \tag{15}$$

where $AVGR_{Pic}$ is the average bits per picture. N_{Coded} and R_{Coded} are the number of coded pictures and spent bits, respectively. N_{GOP} is the number of pictures in a GOP. SW is the size of the sliding window that aims to make the bit adjustment smooth. The SW used in our experiments is set to 40.

3) Picture-Level Bit Allocation: The picture-level bit allocation can be divided into intra-frame bit allocation and inter-frame bit allocation. We follow the current method in the HM software to determine the bits of the intra frame. The intra-frame target bits are proportional to the frame complexity that is measured using the sum of absolute transformed difference

TABLE IV
Ω_{Pic_i} SETTING FOR THE GEOMETRY AND ATTRIBUTE VIDEOS

Frame level	Geometry	Attribute
0	1.0	1.0
1	1.0	1.5874
2	1.0	3.1748
3	1.0	5.0397
4	1.0	6.3496

(SATD). The inter-frame bit allocation is related to both the reference structures and video content. As proved in [10], the λ ratio of various pictures should be inversely proportional to their influences on the subsequent pictures,

$$\lambda_{Pic_i} \cdot \Omega_{Pic_i} = \lambda_{Pic_j} \cdot \Omega_{Pic_j} \triangleq \lambda, \tag{16}$$

where λ_i and λ_j are the Lagrange multipliers for pictures i and j, respectively. Ω_{Pic_i} and Ω_{Pic_j} are the distortion-propagation influences on the subsequent pictures accordingly. λ is the video-level optimization parameter. Additionally, the GOP-level target bits TR_{GOP} are used as a constraint of the sum of the target bits of all the pictures in the GOP,

$$\sum_{i=1}^{N_{GOP}} TR_{Pic_i} = TR_{GOP}, \tag{17}$$

where TR_{Pic_i} is the picture-level target bits of the ith picture in the GOP.

Combining (5), (16), and (17), we can derive the following equation,

$$\sum_{i=1}^{N_{GOP}} \left(\frac{\lambda}{\Omega_{Pic_i} \cdot \alpha_{Pic_i}}\right)^{\frac{1}{\beta_{Pic_i}}} = TR_{GOP}, \tag{18}$$

where α_{Pic_i} and β_{Pic_i} are the R-λ model parameters of picture i. To solve this equation, we need to derive the unknown parameters including α_{Pic_i}, β_{Pic_i}, and Ω_{Pic_i}. We follow the method in HM software to determine α_{Pic_i} and β_{Pic_i} according to the model parameters of the previous picture in the same hierarchical level.

Then, the only unknown parameter in (18) is Ω_{Pic_i}. It is not easy to find the optimal Ω_{Pic_i} for different point clouds under various target bitrates. In this paper, we follow the V-PCC CTC to set Ω_{Pic_i} for simplification, as we consider the V-PCC CTC to be able to provide good RD performance. Under the V-PCC CTC, the pictures at the same hierarchical level have the same Ω_{Pic_i}. Under the random access coding structure with GOP size 16, there are five hierarchical levels. The detailed settings of Ω_{Pic_i} for geometry and attribute videos in different hierarchical levels are shown in Table IV. Note that the settings are different from the rate control designed for general videos for the following two reasons. First, corresponding patches that are placed in different positions of neighboring frames decrease the inter correlation for both geometry and attribute videos. Second, the geometry video has less correlation compared with the attribute video, as its pixel values of each patch are subtracted by the minimum value in the patch.

4) Basic-Unit-Level Bit Allocation: The optimization target of the BU-level bit allocation is to minimize the reconstructed quality of the point cloud frame under the constraint of TR_{Pic}. When performing the BU-level bit allocation, it is difficult to accurately measure the influence of each BU on the reconstructed quality of the point cloud frame. For simplification, we consider the BUs containing occupied pixels, which are called occupied BUs, to have the same importance. The BUs containing only unoccupied pixels, which are called unoccupied BUs, have no influence on the reconstructed quality. Therefore, we formulate the BU-level bit allocation as

$$\min_{\lambda_{BU_j}} \sum_{i=1}^{N_{Pic}} \sum_{BU_i \in BU_O} TD_{BU_i}, \quad s.t. \sum_{i=1}^{N_{Pic}} TR_{BU_i} = TR_{Pic}, \tag{19}$$

where TD_{BU_i} and TR_{BU_i} are the target distortion and bits for the ith BU, respectively. N_{Pic} is the number of BUs in the picture. BU_O is the set of occupied BUs in the picture. As shown in (19), we care about the distortions of occupied BUs and ignore the distortions of unoccupied BUs.

The constrained problem is converted to the following unconstrained problem by introducing the Lagrangian Multiplier μ,

$$\min_{\lambda_{BU_j}} \sum_{i=1}^{N_{Pic}} \sum_{BU_i \in BU_O} TD_{BU_i} + \mu \sum_{i=1}^{N_{Pic}} TR_{BU_i}. \tag{20}$$

This unconstrained problem is solved by setting its derivative to λ_{BU_j} as 0,

$$\frac{\partial \sum_{i=1}^{N_{Pic}} \sum_{BU_i \in BU_O} TD_{BU_i}}{\partial \lambda_{BU_j}} + \mu \frac{\partial \sum_{i=1}^{N_{Pic}} TR_{BU_i}}{\partial \lambda_{BU_j}} = 0. \tag{21}$$

If $BU_j \in BU_O$, (21) is converted to the following equation,

$$\frac{\partial TD_{BU_j}}{\partial \lambda_{BU_j}} + \mu \frac{\partial TR_{BU_j}}{\partial \lambda_{BU_j}} = 0, \quad if\ BU_j \in BU_O. \tag{22}$$

According to the RD curves of BUs, we have

$$\lambda_{BU_j} = \mu, \quad if\ BU_j \in BU_O, \tag{23}$$

where λ_{BU_j} is the slope of the tangent line of the RD curves of the jth BU. Eq. (23) is in accordance with the conclusion in [10] and [11]. λ_{BU_j} of each BU should be set as equal as possible to optimize the RD performance of the current frame.

If $BU_j \notin BU_O$, (21) is converted to the following equation,

$$\mu \frac{\partial TR_{BU_j}}{\partial \lambda_{BU_j}} = 0, \quad if\ BU_j \notin BU_O. \tag{24}$$

From (5), we have

$$\frac{\partial TR_{BU_j}}{\partial \lambda_{BU_j}} = \frac{1}{\beta_{BU_j}} \left(\frac{1}{\alpha_{BU_j}}\right)^{\frac{1}{\beta_{BU_j}}} \lambda_{BU_j}^{\left(\frac{1}{\beta_{BU_j}} - 1\right)} \triangleq \alpha_1 \lambda_{BU_j}^{\beta_1}. \tag{25}$$

As β_1 is negative, by substituting (25) into (24), we obtain

$$\lambda_{BU_j} = +\infty, \quad if\ BU_j \notin BU_O, \tag{26}$$

where α_{BU_j} and β_{BU_j} are the R-λ model parameters of the jth BU. According to (5), we can see that (26) indicates that the

unoccupied BUs should be assigned zero bits. This conclusion is in accordance with our common sense that the unoccupied BUs that have no influence on the reconstructed quality should be assigned zero bits. Theoretically, zero bits are only obtained if λ_{BU_j} equals infinity. However, infinity λ_{BU_j} is not used because it would result in a very large RD cost that may not be correctly represented by the double-precision floating-point format. In addition, the encoding choices are a discrete combination of modes, motions, and QPs. The best we can do is to find the combination with the least number of bits cost. Therefore, we set λ_{BU_j} to 160000 in our implementation. It is a practical choice not only to make the RD cost correctly represented by the double-precision floating-point format but also to find the combination of the encoding choices with the least number of bits.

In addition to the constraint shown in (23), the sum of the target bits of all the occupied BUs should follow the picture-level target bits TR_{Pic},

$$\sum_{i=1}^{N_{Pic}} \sum_{BU_i \in BU_O} TD_{BU_i} = TR_{Pic}. \tag{27}$$

Combining (23), (5), and (27), we can derive the target bits TD_{BU_i} for each BU. Additionally, we modify two other aspects of the original BU-level rate control algorithm in HM software to make it correspond with the proposed BU-level bit allocation. First, only λ_{BU_j}s of the occupied BUs are clipped in a limited range of the picture-level λ_{Pic}. λ_{BU_j}s of unoccupied BUs are not restricted within a limited range of λ_{Pic} in the proposed algorithm. Second, only the distortions of the occupied BUs are counted in the overall distortion of the picture.

B. Bitrate Control

The above bit allocation processes assign the bits to each picture and BU. In this section, we determine the λ and QP for each picture and BU to finish the encoding process. We use (5) to determine the λ. Then, QP is determined as follows [46],

$$QP = 4.3281 \times ln\lambda + 14.4329. \tag{28}$$

The only problem left in both bit allocation and rate control processes is how to determine the model parameters α and β. In [11], the R-λ model parameters are calculated using

$$\beta = -\frac{R\lambda}{D} - 1, \tag{29}$$

$$\alpha = \lambda R^{(1+\frac{R\lambda}{D})}. \tag{30}$$

During the rate control process, the current picture or BU has not yet been encoded. λ, R, and D of the previously encoded picture or the co-located BU at the same hierarchical level is used to estimate the model parameters of the current picture or BU.

The above method works well for the update of the picture-level model parameters under the V-PCC framework. However, for BU-level bitrate control, the corresponding BUs may be placed in different positions. An occupied BU in the current frame may correspond to an unoccupied BU in the co-located position. Therefore, using the co-located BU to estimate the model parameters does not work at all. In this paper, we use the auxiliary information to find the corresponding BU in the previous frame to obtain more accurate model parameters for the current BU.

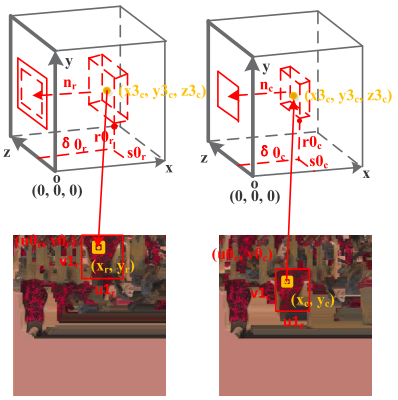


Fig. 5. Illustration of finding the corresponding BU.

We first handle the case where some pixels are occupied in the current BU according to the occupancy map of the current picture OM_c. If the center pixel of the current BU is occupied, we use it as the representative of the current BU. If the center pixel of the current BU is unoccupied, we go through the current BU using the raster scan and use the first occupied pixel as the representative of the current BU. The representative pixel is denoted as (x_c, y_c). The idea of finding the corresponding BU (x_r, y_r) is shown in Fig. 5. Based on the block-to-patch information BP_c, we find the current patch to which the current BU (x_c, y_c) belongs. The auxiliary information of the current patch provides us with the patch projected plane n_c, the patch 2D bounding box $(u0_c, v0_c)$, and the patch 3D location $(\delta 0_c, s0_c, r0_c)$. With all this information, we can calculate the corresponding 3D coordinate $(y3_c, z3_c)$ using (2).

We then search all the patches in the previously coded frame at the same hierarchical level to find the corresponding patch. The corresponding patch should at least follow two constraints. First, the projected plane of the reference patch n_r should be the same as n_c. If the projected planes are different, the possibility of finding the corresponding patch is very small. Even if the current BU finds the corresponding BU in the previous frame, different projection planes lead to various shape changes of the corresponding BUs. This makes the estimation of the model parameters inaccurate. Therefore, we bypass the reference patches with different projected planes from the current patch. Second, $(y3_c, z3_c)$ should be within the range of the reference patch,

$$\begin{cases} s0_r \le z3_c \le s0_r + u1_r \times bpr - 1 \\ r0_r \le y3_c \le r0_r + v1_r \times bpr - 1, \end{cases} \tag{31}$$

where $(u1_r, v1_r)$ is the bounding box size of the reference patch and $(s0_r, r0_r)$ is the 3D coordinate of the top-left position of the reference patch. If the reference patch does not contain $(y3_c, z3_c)$, it is impossible for the reference patch to have the current BU.

TABLE V
SOME EXAMPLES OF αS OF THE GEOMETRY AND ATTRIBUTE VIDEOS

Test point cloud	Geometry	Attribute
Loot	0.64	20.84
RedAndBlack	1.07	17.60
Soldier	0.90	38.16
Queen	0.75	26.42
LongDress	0.83	97.84

There may be multiple patches following the above two constraints. For example, if one reference patch satisfies the constraint, the patch in its opposite position in 3D space also has a good chance of satisfying the constraint. The differences between these patches are the minimum depth of the patch $\delta 0_r$. We choose the patch with the smallest difference between $\delta 0_r$ and the minimum depth of the current patch $\delta 0_c$ since the motion is usually not large between the current frame and the previously coded frame at the same hierarchical level.

After the reference patch is found, based on the assumption that the 3D coordinates of the current pixel are the same as that of the reference pixel,

$$\begin{cases} s0_r + (x_r - u0_r \times bpr) = s0_c + (x_c - u0_c \times bpr) \\ r0_r + (y_r - v0_r \times bpr) = r0_c + (y_c - v0_c \times bpr). \end{cases} \tag{32}$$

The coordinate of the reference BU (x_r, y_r) can be calculated as

$$\begin{cases} x_r = x_c + (s0_c - s0_r) + (u0_r - u0_c) \times bpr \\ y_r = y_c + (r0_c - r0_r) + (v0_r - v0_c) \times bpr. \end{cases} \tag{33}$$

As we can see from (33), the right term is a combination of the 3D and 2D patch offsets. Therefore, this coordinate offset is essentially a global patch offset. We use λ, R, and D of the BU containing (x_r, y_r) to update the model parameters of the current BU using (29) and (30).

All the unoccupied BUs that are padded based on the occupied pixels can share the model parameters as these BUs are usually smooth. We obtain the model parameters from the unoccupied BUs in the previous frame at the same hierarchical level.

In addition, as we mentioned in Fig. 3, the αs of the geometry and attribute videos differ significantly. We also give some examples of αs of the geometry and attribute frames of typical point clouds under bitrate scenario r2, as shown in Table V. All the numbers in Table V are from the 48th frame in the geometry and attribute videos. We can see that αs of the geometry frames are much smaller than that of the attribute frames. In addition, we know that the initial α is set to 3.2003 in HM software for the natural videos. Comparing α of the natural videos with that of the geometry and attribute videos, we set the initial αs of the attribute and geometry videos as the default one in HM software multiplying 10 and dividing 10, respectively. Note that the initial αs of the geometry and attribute videos do not have to be accurate. They can be updated during the encoding process to approximate the actual αs. They only need to be within a range of the actual αs to avoid irregular bit cost at the beginning stage of rate control.

TABLE VI
CHARACTERISTICS OF THE TEST DYNAMIC POINT CLOUDS

Test point cloud	Frame rate	Points per frame	Geometry precision	Attributes
Loot	30	$\sim$ 780000	10 bit	RGB
RedAndBlack	30	$\sim$ 700000	10 bit	RGB
Soldier	30	$\sim$ 1500000	10 bit	RGB
Queen	50	$\sim$ 1000000	10 bit	RGB
LongDress	30	$\sim$ 800000	10 bit	RGB

IV. EXPERIMENTAL RESULTS

The proposed algorithms are implemented in the V-PCC reference software TMC2-5.0 [15] and the corresponding HM software [16] to compare with the V-PCC anchor without rate control. We also use the λ-domain rate control algorithm designed for the general videos as the anchor for comparison. Note that the differences between the proposed algorithm and the anchor are only the proposed BU-level bit allocation and model updating algorithms. We set all the other parameters the same, including the initial model parameters, for a fair comparison. We use the same video-level bit allocation algorithm for the anchor and the proposed rate control algorithm since there is no video-level bit allocation algorithm yet designed for V-PCC. For both the anchor and the proposed rate control algorithm, the target bits are generated as follows. We run the TMC2-5.0 anchor without rate control and count the number of bits ranging from the low bitrate scenario r1 to the high bitrate scenario r5 following the V-PCC CTC. The counted bits are then used as the target bits for both the anchor and the proposed rate control algorithm.

We test the lossy geometry, lossy attribute, random access (RA) case to demonstrate the effectiveness of the proposed algorithms. We perform experiments on the five DPCs defined in the V-PCC CTC [45]. The detailed characteristics of the test point clouds are shown in Table VI. In the current V-PCC reference software, every 32 frame is encoded as independent sequences. Therefore, we test 32 frames as a good representative for the whole point cloud to verify the performance of the proposed algorithms. Since the bits generated by the anchor and the proposed algorithms are not the same, the Bjontegaard-Delta-rate (BD-rate) [47] is used to compare the respective RD performance.

We show the benefits of the proposed algorithm in two aspects: the RD performance and the bit error between the target bits and actual bits. For the RD performance of the geometry, we report the BD-rates for both point-to-point PSNR (D1) and point-to-plane PSNR (D2) [45]. For the RD performance of the attribute, the BD-rates for the Luma, Cb, and Cr components are reported. The bit error E between the target bits and actual bits is defined as follows,

$$E = \frac{abs(R_t - R_a)}{R_t}, \tag{34}$$

where R_t and R_a are the target bits and actual bits, respectively. In the following subsections, we first introduce the overall performance of the proposed rate control algorithm. Then, we report the performance and analysis of the proposed algorithms individually.

TABLE VII

OVERALL PERFORMANCE OF THE PROPOSED RATE CONTROL ALGORITHM COMPARED WITH THE ORIGINAL RATE CONTROL ALGORITHM

Test point cloud	Geom.BD-TotalRate D1	D2	Attr.BD-TotalRate Luma	Cb	Cr
Loot	–5.8%	–6.8%	–1.9%	–1.6%	–1.6%
RedAndBlack	–1.3%	–2.6%	–3.3%	–3.8%	–4.4%
Soldier	–3.3%	–3.9%	–0.3%	0.2%	0.8%
Queen	–6.7%	–7.2%	–2.2%	–0.3%	0.2%
LongDress	–0.4%	–0.7%	–3.2%	–2.9%	–3.7%
Avg.	**–3.5%**	**–4.2%**	**–2.2%**	**–1.7%**	**–1.7%**

TABLE VIII

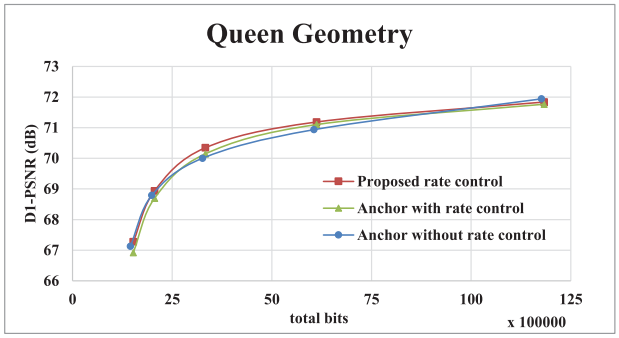
OVERALL PERFORMANCE OF THE PROPOSED RATE CONTROL ALGORITHM COMPARED WITH THE TMC2-5.0 ANCHOR WITHOUT RATE CONTROL

Test point cloud	Geom.BD-TotalRate D1	D2	Attr.BD-TotalRate Luma	Cb	Cr
Loot	2.7%	1.5%	4.9%	2.7%	3.5%
RedAndBlack	12.4%	11.8%	6.0%	6.1%	6.7%
Soldier	–7.7%	–7.6%	1.9%	–1.2%	0.6%
Queen	–6.2%	–6.2%	5.2%	0.0%	2.1%
LongDress	17.2%	12.5%	10.5%	3.1%	6.4%
Avg.	**3.7%**	**2.4%**	**5.7%**	**2.1%**	**3.9%**

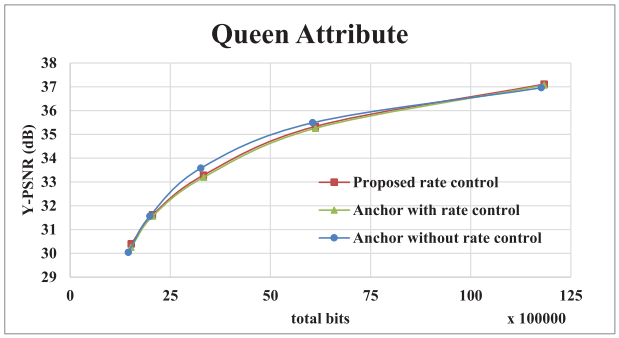
A. Overall Performance of the Proposed Rate Control Algorithm

Table VII shows the overall RD performance of the proposed rate control algorithm compared to the anchor with rate control. Note that ω is set to 8 in the overall experimental result. We show the influences of different ωs in the next subsection. In addition, we only enable the proposed BU-level bit allocation algorithm for the attribute instead of the geometry. This is also explained in the following subsections. From Table VII, we can see that the proposed algorithm can lead to 3.5% and 4.2% RD performance improvements on average for the geometry under the D1 and D2 measurements, respectively. In terms of the attribute, the proposed algorithm can save an average of 2.2%, 1.7%, and 1.7% bits for the Luma, Cb, and Cr components, respectively. The proposed algorithms bring up to 6.7% and 7.2% RD performance improvements for the geometry for the point cloud "Queen". In addition, it saves 2.2% bits saving for the Luma component. The experimental results demonstrate that the proposed rate control algorithm can lead to obvious bit savings compared to the original rate control algorithm.

Table VIII shows the overall RD performance of the proposed rate control algorithm compared to the TMC2-5.0 anchor without rate control. We can see that the proposed algorithm leads to an average of 3.7% and 5.7% RD performance losses compared to the anchor without rate control for the geometry (D1) and attribute (Luma), respectively. As the intra frame costs much more bits than the inter frames, the immediate inter frames after the intra frame are assigned much fewer bits compared with the anchor to maintain the balance of the buffer. Those frames suffer quality degradation that propagates to the subsequent frames. That is why we suffer some performance losses on average.



(a) Queen Geometry



(b) Queen Attribute

Fig. 6. Typical examples of the RD curves.

TABLE IX

BIT ERRORS COMPARISON BETWEEN THE ANCHOR AND THE PROPOSED RATE CONTROL ALGORITHM

Test case	Point cloud	r1	r2	r3	r4	r5
Anchor	Loot	2.9%	2.6%	2.0%	0.8%	0.4%
	RedAndBlack	5.8%	4.9%	3.9%	2.2%	0.6%
	Soldier	4.7%	3.3%	2.3%	1.3%	0.3%
	Queen	4.9%	3.5%	2.1%	1.1%	0.6%
	LongDress	4.2%	3.0%	1.5%	0.4%	0.5%
	Avg.			**2.4%**		
	Max			**5.8%**		
Proposed	Loot	2.9%	2.5%	1.7%	0.9%	0.4%
	RedAndBlack	6.4%	5.1%	3.2%	1.8%	0.8%
	Soldier	4.8%	3.2%	2.3%	1.3%	0.6%
	Queen	4.9%	3.5%	2.1%	1.1%	0.6%
	LongDress	4.4%	2.7%	2.2%	0.9%	0.1%
	Avg.			**2.4%**		
	Max			**6.4%**		

In addition, we show the geometry and attribute RD curves for the DPC "Queen" for validation of the RD performance in Fig. 6. The proposed rate control algorithm shows better RD performance compared with the anchor with rate control. However, it leads to some improvements for some bitrate scenarios and some losses for the other bitrate scenarios compared to the anchor without rate control.

Table IX shows the comparison of the bit error between the anchor and the proposed rate control algorithm. We can see that the proposed rate control algorithm achieves the same average bit error and a slightly higher maximum bit error compared with the anchor. The bit error is mainly determined by picture-level bit allocation and rate control. As the proposed algorithm is mainly designed on the BU level, it leads to the

TABLE X

PERFORMANCE OF THE PROPOSED RATE CONTROL ALGORITHM WITH ω SET TO 6 COMPARED TO THAT WITH ω SET TO 8

Test point cloud	Geom.BD-TotalRate D1	D2	Attr.BD-TotalRate Luma	Cb	Cr
Loot	9.8%	10.3%	–6.9%	–5.1%	–6.7%
RedAndBlack	10.3%	10.7%	–4.6%	–5.4%	–4.7%
Soldier	13.7%	15.5%	–3.8%	–1.8%	–3.1%
Queen	19.1%	18.6%	–0.4%	–2.9%	–3.6%
LongDress	23.2%	21.6%	–2.2%	–5.3%	–4.6%
Avg.	**15.2%**	**15.3%**	**–3.6%**	**–4.1%**	**–4.6%**

TABLE XI

PERFORMANCE OF THE PROPOSED RATE CONTROL ALGORITHM WITH ω SET TO 10 COMPARED TO THAT WITH ω SET TO 8

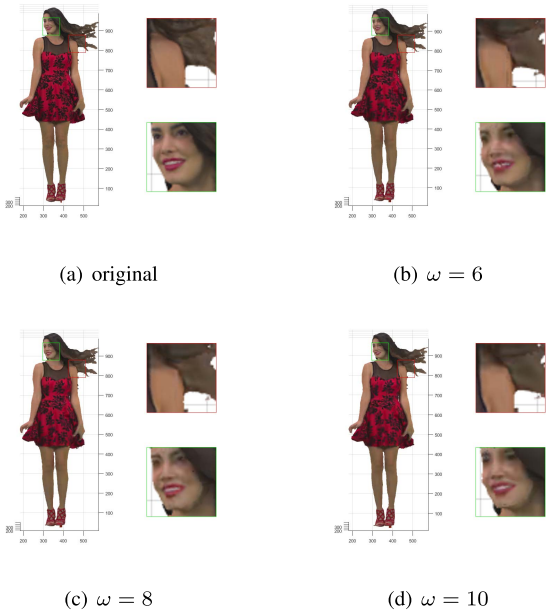
Test point cloud	Geom.BD-TotalRate D1	D2	Attr.BD-TotalRate Luma	Cb	Cr
Loot	–7.0%	–7.5%	7.7%	9.8%	8.9%
RedAndBlack	–9.1%	–8.7%	6.2%	7.3%	7.0%
Soldier	–11.3%	–10.4%	5.4%	9.2%	7.3%
Queen	–14.7%	–14.8%	2.3%	3.6%	3.4%
LongDress	–14.8%	–14.4%	1.7%	1.8%	2.0%
Avg.	**–11.4%**	**–11.2%**	**4.7%**	**6.4%**	**5.7%**

same average bit error compared with the original rate control algorithm. Additionally, we can see that both the anchor and the proposed algorithm lead to higher bit errors in the low bitrate case compared with those in the high bitrate case. The portion of the intra bits in the low bitrate case is higher than that in the high bitrate case. The larger portion of intra bits leads to fewer bits available for the subsequent frames and creates larger errors.

B. Performance of the Video-Level Bit Allocation

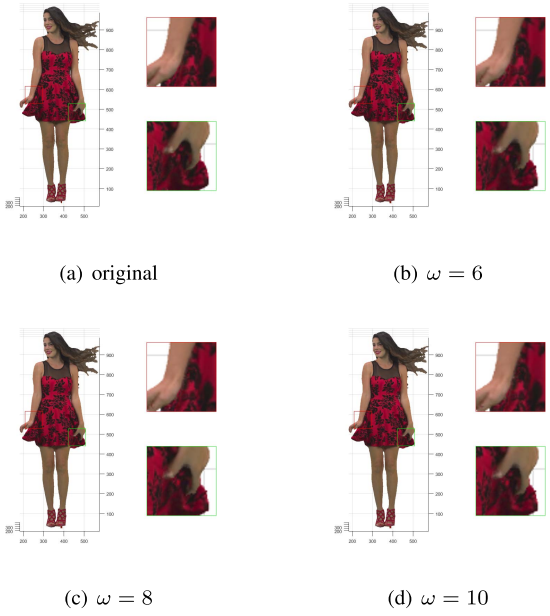
In this section, we analyze the performance of the video-level bit allocation algorithm under different ωs. Table X and Table XI show the performance of the proposed algorithm with ω set to 6 and 10 compared to ω set to 8, respectively. We can see that the larger the ω is, the better RD performance we can achieve for the geometry, however, the worse the RD performance we obtain for the attribute. This is in accordance with our analysis in Section III-A that a larger λ corresponds to better RD performance for the geometry and worse RD performance for the attribute. However, how to select a better ω for the overall reconstructed point cloud quality remains a problem.

Essentially, to obtain the optimal ω, we need a full experimental result on the subjective quality of the reconstructed point cloud under different ωs. However, this is not the main focus of this work. We only show a few examples of subjective qualities in high and low bitrates to find a suitable ω, as shown in Fig. 7 and Fig. 8. We can see obvious geometry distortions in the low bitrate case if we set ω to 6. The woman's hand and teeth have some extra points that significantly influence the subjective quality. In addition, we can observe that the woman's clothes have extra colors in the high bitrate case if we set ω to 10. The artifacts in the high bitrate case are not as obvious as those in the low bitrate case since the qualities of the reconstructed point cloud in the high bitrate case are much higher than those in the low bitrate case. The optimal ω may vary under different target bits. In this paper, we choose ω as 8 to obtain a better performance balance in the low and high bitrate cases.



(a) original (b) $\omega = 6$ (c) $\omega = 8$ (d) $\omega = 10$

Fig. 7. Subjective qualities of setting different ωs in the low bitrate case. The example is from "Redandblack" with picture order count 1450. The target bits are 2234336.



(a) original (b) $\omega = 6$ (c) $\omega = 8$ (d) $\omega = 10$

Fig. 8. Subjective qualities of setting different ωs in the high bitrate case. The example is from "Redandblack" with picture order count 1450. The target bits are 14059352.

C. Performance of the Proposed Model Parameter Estimation

Table XII shows the performance of the proposed model updating algorithm compared with the original model updating

TABLE XII

PERFORMANCE OF THE PROPOSED MODEL UPDATING ALGORITHM COMPARED TO THE ORIGINAL MODEL UPDATING ALGORITHM

Test point cloud	Geom.BD-TotalRate D1	Geom.BD-TotalRate D2	Attr.BD-TotalRate Luma	Attr.BD-TotalRate Cb	Attr.BD-TotalRate Cr
Loot	–5.7%	–6.7%	1.2%	1.2%	1.2%
RedAndBlack	–1.6%	–2.8%	–0.1%	0.5%	0.2%
Soldier	–3.3%	–3.9%	0.8%	1.8%	2.3%
Queen	–6.7%	–7.2%	–1.4%	1.0%	1.1%
LongDress	–0.4%	–0.7%	–0.5%	–0.4%	–0.9%
Avg.	**–3.6%**	**–4.3%**	**0.0%**	**0.8%**	**0.8%**

TABLE XIII

PERFORMANCE OF THE BU-LEVEL BIT ALLOCATION ALGORITHM FOR GEOMETRY COMPARED WITH THE ORIGINAL BU-LEVEL BIT ALLOCATION ALGORITHM

Test point cloud	Geom.BD-TotalRate D1	Geom.BD-TotalRate D2	Attr.BD-TotalRate Luma	Attr.BD-TotalRate Cb	Attr.BD-TotalRate Cr
Loot	0.2%	0.9%	–0.6%	–0.6%	–2.6%
RedAndBlack	–0.7%	0.2%	–0.3%	–0.5%	0.1%
Soldier	–0.1%	0.4%	0.3%	1.1%	–0.8%
Queen	3.6%	4.4%	0.6%	–0.1%	–1.3%
LongDress	–2.0%	–1.0%	–0.7%	–1.0%	–0.6%
Avg.	**0.2%**	**1.0%**	**–0.1%**	**–0.2%**	**–1.0%**

TABLE XIV

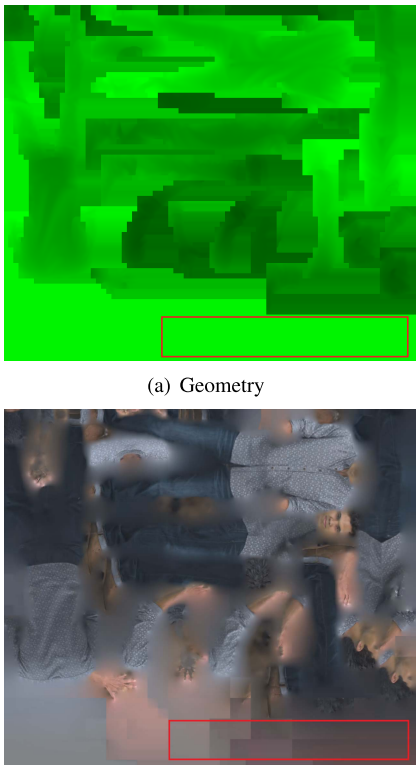
PERFORMANCE OF THE BU-LEVEL BIT ALLOCATION ALGORITHM FOR ATTRIBUTE COMPARED WITH THE ORIGINAL BU-LEVEL BIT ALLOCATION ALGORITHM

Test point cloud	Geom.BD-TotalRate D1	Geom.BD-TotalRate D2	Attr.BD-TotalRate Luma	Attr.BD-TotalRate Cb	Attr.BD-TotalRate Cr
Loot	–0.1%	–0.1%	–3.1%	–2.8%	–2.8%
RedAndBlack	0.3%	0.3%	–3.2%	–4.3%	–4.5%
Soldier	0.0%	0.0%	–1.1%	–1.6%	–1.4%
Queen	0.0%	0.0%	–0.7%	–1.3%	–1.0%
LongDress	0.0%	0.0%	–2.6%	–2.5%	–2.8%
Avg.	**0.0%**	**0.0%**	**–2.2%**	**–2.5%**	**–2.5%**

algorithm. We can see that the proposed model updating algorithm leads to 3.6% and 4.3% bit savings on average under D1 and D2 quality measurements, respectively. It leads to the same or slightly worse performance improvements for the Luma, Cb, and Cr components, respectively. The experimental results show that the proposed model updating algorithm can bring better RD performance for the geometry compared with the attribute. As we have shown in Section III-B, we use the center pixel to represent the current BU to find the corresponding pixel. However, the corresponding pixel may not be the center pixel of the corresponding BU. Therefore, the corresponding BUs usually do not have exactly the same content. This situation is more serious for attribute videos whose texture is less smooth compared with that of the geometry video.

D. Performance of the Proposed BU-level Bit Allocation

Table XIII and Table XIV show the performance of the BU-level bit allocation algorithm for the geometry and attribute compared with the original BU-level bit allocation, respectively. We can see that the proposed BU-level bit allocation algorithm leads to obvious RD performance improvements for the attribute. However, it leads to some RD performance losses for the geometry. Therefore, we disabled the BU-level bit allocation algorithm for the geometry in the overall experimental results.



(a) Geometry

(b) Attribute

Fig. 9. Typical example of the reconstructed frame of the geometry and attribute under the proposed BU-level bit allocation algorithm. The example is from the 5th frame of the geometry and attribute videos from "Loot" with target bits 8097008.

The different padding algorithms for the geometry and attribute are the key reasons why the proposed BU-level bit allocation leads to different performances. An unoccupied geometry BU is padded using the horizontal extension from the left border pixels or the vertical extension from the top border pixels if the left border pixels are unavailable. The unoccupied geometry BU does not require many bits through intra horizontal prediction or vertical prediction even if the proposed BU-level bit allocation is not used. However, an unoccupied attribute BU is padded using a pull-push algorithm to minimize the overall bitrate for both the completely unoccupied BU and partially unoccupied BU [48]. Therefore, the proposed BU-level bit allocation has a larger chance to save some bits and improve the overall performance for the attribute. To better illustrate this problem, we show an example of the reconstructed frames of the geometry and attribute videos in Fig. 9. We can see very obvious blocking artifacts for the unoccupied BUs, as indicated by the red triangles in the attribute frame. This is due to the bit savings for the residue of the unoccupied BUs. However, we cannot see similar artifacts in the geometry frame, as the unoccupied BUs in the geometry frame are very smooth.

V. CONCLUSION

In this paper, we propose the first rate control algorithm for the video-based point cloud compression (V-PCC) framework.

We mainly introduce the following key algorithms to address the new features brought by the V-PCC framework. First, we propose a video-level bit allocation algorithm to assign the bits between the geometry and the attribute videos. Second, we introduce a basic-unit-level (BU-level) bit allocation algorithm assigning zero bits to the unoccupied BUs by ignoring their distortions. Third, we propose a more accurate model updating scheme by finding the corresponding BUs with similar video content in the previous frame at the same hierarchical level. The proposed rate control algorithms are implemented in the V-PCC and the corresponding High Efficiency Video Coding (HEVC) reference software. The experimental results show that the proposed algorithms can lead to obvious rate-distortion performance improvements. We will investigate the optimal video-level, picture-level bit allocation algorithms in our future work to achieve even better performance.

REFERENCES

[1] H. Fuchs, A. State, and J.-C. Bazin, "Immersive 3D telepresence," *Computer*, vol. 47, no. 7, pp. 46–52, Jul. 2014.

[2] M.-L. Champel, R. Doré, and N. Mollet, "Key factors for a high-quality VR experience," *Proc. SPIE*, vol. 10396, Sep. 2017, Art. no. 103960Z.

[3] E. d'Eon, B. Harrison, T. Myers, and P. A. Chou, *Input to Ad Hoc Groups on MPEG Point Cloud Compression and JPEG PLENO*, document ISO/IEC JTC1/SC29/WG11 m40059, Geneva, Switzerland, Jan. 2017.

[4] S. Schwarz *et al.*, "Emerging MPEG standards for point cloud compression," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 9, no. 1, pp. 133–148, Mar. 2019.

[5] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[6] *Continuous Improvement of Study Text of ISO/IEC CD 23090-5 Video-Based Point Cloud Compression*, document ISO/IEC JTC1/SC29/WG11 w18479, 3DG, Geneva, Switzerland, Mar. 2019.

[7] C. Tulvan, R. Mekuria, Z. Li, and S. Laserre, *Use Cases for Point Cloud Compression (PCC)*, document ISO/IEC JTC1/SC29/WG11 MPEG2015/N16331, Geneva, Switzerland, Jun. 2016.

[8] Z. Chen and K. N. Ngan, "Recent advances in rate control for video coding," *Signal Process., Image Commun.*, vol. 22, no. 1, pp. 19–38, Jan. 2007.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] B. Li, H. Li, L. Li, and J. Zhang, *Rate Control by R-lambda Model for HEVC*, document ISO/IEC JTC1/SC29/WG11 JCTVC-K0103, Shanghai, China, Oct. 2012.

[14] Y. Li, Z. Chen, X. Li, and S. Liu, *Rate Control for VVC*, document ISO/IEC JTC1/SC29/WG11 JVET-K0390, Ljubljana, Slovenia, Jul. 2018.

[15] *Point Cloud Compression Category 2 Reference Software, TMC2-5.0*. Accessed: 2019. [Online]. Available: http://mpegx.int-evry.fr/software/MPEG/PCC/TM/mpeg-pcc-tmc2.git

[16] *High Efficiency Video Coding Test Model, HM-16.18+SCM8.7*. Accessed: 2019. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/

[17] L. Li, Z. Li, S. Liu, and H. Li, *[VPCC] [New Contribution] Occupancy-Map-Based Rate Distortion Optimization*, document m52765, Brussels, Belgium, Jan. 2020.

[18] 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.

[19] Z. He, Y. Kwan Kim, and S. K. Mitra, "Low-delay rate control for DCT video coding via ρ-domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 928–940, Aug. 2001.

[20] 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.

[21] M. Jiang, X. Yi, and N. Ling, "Frame layer bit allocation scheme for constant quality video," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jun. 2004, pp. 1055–1058.

[22] M. Jiang and N. Ling, "Low-delay rate control for real-time H.264/AVC video coding," *IEEE Trans. Multimedia*, vol. 8, no. 3, pp. 467–477, Jun. 2006.

[23] S. Min Zhou, J. Li, J. Fei, and Y. Zhang, "Improvement on rate-distortion performance of H.264 rate control in low bit rate," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 8, pp. 996–1006, Aug. 2007.

[24] H. Schwarz, D. Marpe, and T. Wiegand, "Analysis of hierarchical b pictures and MCTF," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2006, pp. 1929–1932.

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

[26] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[27] 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.

[28] 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.

[29] Y. Gao, C. Zhu, S. Li, and T. Yang, "Source distortion temporal propagation analysis for random-access hierarchical video coding optimization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 2, pp. 546–559, Feb. 2019.

[30] T. Yang, C. Zhu, X. Fan, and Q. Peng, "Source distortion temporal propagation model for motion compensated video coding optimization," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2012, pp. 85–90.

[31] C.-W. Seo, J. W. Kang, J.-K. Han, and T. Q. Nguyen, "Efficient bit allocation and rate control algorithms for hierarchical video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 9, pp. 1210–1223, Sep. 2010.

[32] W. Yuan, S. Lin, Y. Zhang, W. Yuan, and H. Luo, "Optimum bit allocation and rate control for H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 6, pp. 705–715, Jun. 2006.

[33] Z. He and S. K. Mitra, "Optimum bit allocation and accurate rate control for video coding via ρ-domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 10, pp. 840–849, Oct. 2002.

[34] H. Guo, C. Zhu, M. Xu, and S. Li, "Inter-block dependency-based CTU level rate control for HEVC," *IEEE Trans. Broadcast.*, vol. 66, no. 1, pp. 113–126, Mar. 2020.

[35] K. L. Ferguson and N. M. Allinson, "Modified steepest-descent for bit allocation in strongly dependent video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 7, pp. 1057–1062, Jul. 2009.

[36] Y. Gu Lee and B. Cheol Song, "An intra-frame rate control algorithm for ultralow delay H.264/Advanced video coding (AVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 5, pp. 747–752, May 2009.

[37] 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.

[38] 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.

[39] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[40] 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.

[41] 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.

[42] 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.

[43] 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 Vis. Commun. Image Process. (VCIP)*, Dec. 2017, pp. 1–4.
[44] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," in *Adv. neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[45] S. Schwarz, G. Martin-Cocher, D. Flynn, and M. Budagavi, *Common Test Conditions for Point Cloud Compression*, document ISO/IEC JTC1/SC29/WG11 w17766, Ljubljana, Slovenia, Jul. 2018.
[46] K. Andersson *et al.*, *Non-normative HM Encoder Improvements*, document JCTVC-W0062, San Diego, CA, USA, Feb. 2016.
[47] G. Bjontegaard, *Calculation of Average PSNR Differences Between RD-Curves*, document VCEG-M33, Austin, TX, USA, Apr. 2001.
[48] D. Graziosi, *[V-PCC] TMC2 Optimal Texture Packing*, document ISO/IEC JTC1/SC29/WG11 m43681, Ljubljana, Slovenia, Jul. 2018.

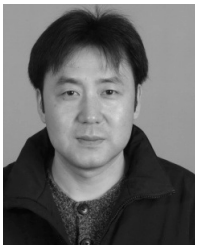
Li Li (Member, IEEE) received the B.S. and Ph.D. degrees in electronic engineering from the University of Science and Technology of China (USTC), Hefei, Anhui, China, in 2011 and 2016, respectively. He is currently a Visiting Assistant Professor with the University of Missouri–Kansas City. His research 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).

Zhu Li (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from Northwestern University in 2004.

He was the AFRL Summer Faculty with the U.S. Air Force Academy, UAV Research Center, in 2016, 2017, and 2018. He was Principal Staff Research Engineer with the Multimedia Research Lab (MRL), Motorola Labs, Schaumburg, IL, USA, from 2000 to 2008, an Assistant Professor with the Department of Computing, The Hong Kong Polytechnic University, from 2008 to 2010, a Senior Staff Researcher with FutureWei, from 2010 to 2012, and a Senior Staff Researcher/Senior Manager with the Samsung Research America's Multimedia Core Standards Research Lab, Dallas, TX, USA, from 2012 to 2015. He is currently an Associated Professor with the Department of CSEE, University of Missouri–Kansas City, Kansas City, MO, USA, and also directs the NSF I/UCRC Center for Big Learning, UMKC. 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, and conference proceedings and standards contributions in these areas. He received the Best Paper Award from the IEEE Int'l Conf on Multimedia and Expo (ICME) at Toronto, in 2006, and the Best Paper Award from the IEEE Int'l Conf on Image Processing (ICIP) at San Antonio, in 2007. He has been the Associate Editor-in-Chief (AEiC) of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, since 2020, and has served and serving as an Associated Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING since 2019, the IEEE TRANSACTIONS ON MULTIMEDIA from 2015 to 2019, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY from 2016 to 2019.

Shan Liu (Senior Member, IEEE) received the B.Eng. degree in electronics engineering from Tsinghua University, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California.

She was formerly the Director of the Media Technology Division, 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, such as HEVC, VVC, OMAF, DASH, and PCC. She is currently a Tencent Distinguished Scientist and the General Manager of Tencent Media Lab. She served as a Co-Editor of HEVC SCC and the emerging VVC. At the same time, technologies and products developed under her leadership have reached more than 10 million DAU. She holds more than 150 granted U.S. and global patents and has authored or coauthored more than 80 peer-reviewed technical articles. She was in the committee of Industrial Relationship of the IEEE Signal Processing Society from 2014 to 2015. She also served the VP of Industrial Relations and Development of Asia-Pacific Signal and Information Processing Association from 2016 to 2017. She was named APSIPA Industrial Distinguished Leader in 2018. She is currently on the Editorial Board of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY from 2018 to 2021. She has been serving as the Vice Chair for the IEEE Data Compression Standards Committee since 2019.



Houqiang Li (Senior Member, IEEE) received the B.S., M.Eng., and Ph.D. degrees in electronic engineering from the University of Science and Technology of China, Hefei, China, in 1992, 1997, and 2000, respectively.

He is currently a Professor with the Department of Electronic Engineering and Information Science, University of Science and Technology of China. He has authored or coauthored over 200 articles in journals and conferences. His research interests include video coding and communication, multimedia search, image/video analysis. He was a recipient of the Best Paper Awards for VCIP 2012, ICIMCS 2012, and ACM MUM in 2011, and the National Technological Invention Award of China (second class) in 2019 and the National Natural Science Award of China (second class) in 2015. He was the winner of the National Science Funds (NSFC) for Distinguished Young Scientists, the Distinguished Professor of Changjiang Scholars Program of China, and the Leading Scientist of Ten Thousand Talent Program of China. He served as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY from 2010 to 2013 and the TPC Co-Chair of VCIP 2010. He will serve as the General Co-Chair of ICME 2021.