IEEK Transactions on Smart Processing and Computing

Largest Coding Unit Level Rate Control Algorithm for Hierarchical Video Coding in HEVC

Yeo-Jin Yoon, Hoon Kim, Seung-Jin Baek, and Sung-Jea Ko

Department of Electronics Engineering, Korea University / Seoul, Rep. of Korea sjko@korea.ac.kr

* Corresponding Author: Sung-Jea Ko

Received November 2, 2012; Revised December 11, 2012; Accepted December 20, 2012; Published December 31, 2012

Abstract**:** In the new video coding standard, called high efficiency video coding (HEVC), the coding unit (CU) is adopted as a basic unit of a coded block structure. Therefore, the rate control (RC) methods of H.264/AVC, whose basic unit is a macroblock, cannot be applied directly to HEVC. This paper proposes the largest CU (LCU) level RC method for hierarchical video coding in a HEVC. In the proposed method, the effective bit allocation is performed first based on the hierarchical structure, and the quantization parameters (QP) are then determined using the Cauchy density based rate-quantization (RQ) model. A novel method based on the linear rate model is introduced to estimate the parameters of the Cauchy density based RQ model precisely. The experimental results show that the proposed RC method not only controls the bitrate accurately, but also generates a constant number of bits per second with less degradation of the decoded picture quality than with the fixed QP coding and latest RC method for HEVC.

Keywords**:** Bit allocation, Hierarchical video coding, High efficiency video coding, Rate control

1. Introduction

Currently, high efficiency video coding (HEVC), a next generation standard video codec, is being developed by JCT-VC [1]. The goal of this new codec is to maximize the coding performance with the least increment of computational complexity compared to H.264/AVC [2]. HEVC is expected to achieve more efficient video compression performance, particularly for high definition (HD) and ultra-HD (UHD) resolution video contents. To this end, HEVC adopts a highly flexible coding block structure based on the quad-tree scheme. Therefore, the coding unit (CU), which is an expended concept of macroblock (MB) of H.264/AVC, is used as a basic unit for region splitting. The CUs can have a range of block sizes from 8×8 to 64×64 pixels, and recursive splitting into smaller CUs is allowed for the contents in the image. As a result, advanced rate control (RC) based on the new coding block structure of HEVC is needed because the conventional MB level RC methods cannot be applied to HEVC directly. In [3], the first RC method for HEVC was introduced, in which the MB level RC scheme of H.264/AVC [4, 5] was applied according to the coding structure of HEVC with minor modifications to control the bitrate based on the largest CU (LCU).

High coding performance can be achieved using the hierarchical-B structure, as shown in Fig. 1 [6]. Hierarchical video coding supports both the temporal scalability and network condition adaptability, which are useful for broadcast and communication applications. Therefore, a range of RC methods, which employ the bit allocation and quantization parameter (QP) decision methods for B-frames in the hierarchical structure, have been proposed. In [7], the RC_MODE_2 and RC_MODE_3 methods were proposed for hierarchical video coding and adopted as RC methods in H.264/AVC reference software [8]. Li et al. proposed an improved frame level RC method of H.264/AVC by integrating the temporal level RC phase [9]. The RC algorithm in [10] improves the bit allocation performance of H.264/AVC RC using the hierarchical level, complexity, and sensitivity of each frame.

This paper proposes an LCU level RC method for hierarchical video coding in HEVC. The proposed RC algorithm aims to achieve constant bitrate (CBR) control while ensuring minimal fluctuations in the decoded picture quality. First the number of target bits is assigned to each

---

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No. 2012R1A2A4A01008384).

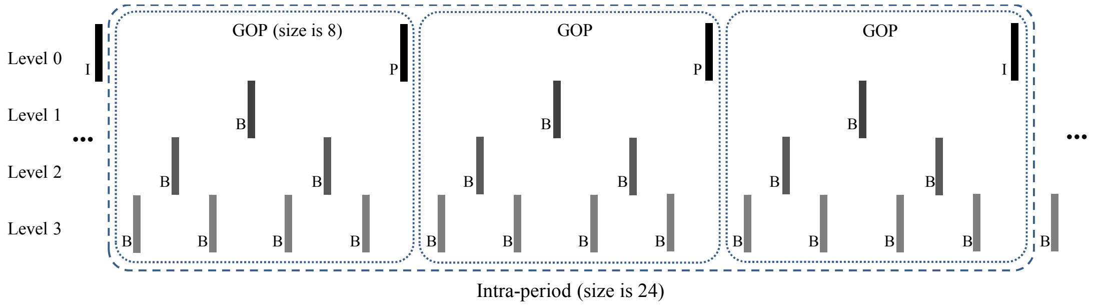

Fig. 1. Hierarchical coding structure with the intra-period of 24 frames and GOP of 8 frames. Each rectangle represents a frame.

intra-period, group of picture (GOP), frame, and LCU by considering the hierarchical structure. The QPs for the frame and LCU are then determined to generate the target bits using the Cauchy density based rate-quantization (RQ) model [11]. In addition, a novel parameter estimation method for the RQ model based on the linear rate model is presented.

This paper is organized as follows. In Section 2, the proposed bit allocation method is explained. Section 3 describes the RQ model used in the algorithm and presents the new parameter estimation and QP decision methods. Performance analysis is presented in Section 4, and the conclusions and future work are discussed in Section 5.

2. Proposed Bit Allocation Method

When a video is encoded in the hierarchical-B structure without a RC, the number of output bits for each frame is different depending on the frame type and temporal level. For temporal level 0, the number of output bits for the P-frame is smaller than that for the I-frame. For higher temporal levels, the frames are encoded in B-picture which allows a bi-prediction, and the closer neighboring frames are referred for the prediction in the higher temporal level. Therefore, the number of output bits for each frame decrease with increasing temporal level. As a result, it is necessary to allocate the target bits by considering the hierarchical structure for RC. In the proposed bit allocation method, the number of bits for an intra-period is first assigned to generate a constant number of bits per second. Then, the target bits for a GOP are allocated by considering the generated bits of the previous GOP as well as the remaining bits of the current intra-period and distributed to each frame using the frame weights that differ according to the temporal level and frame type. The number of target bits for each LCU in a frame is allocated from the target bits of the frame using the number of sub-CUs in the LCU. The frequently used notations for describing the bit allocation scheme in detail are summarized in Table 1.

2.1 Intra-period Level Bit Allocation

The size of the intra-period is determined by the random access interval of approximately 1 second in [12]. Therefore, in the proposed method, the target bits are first assigned to the intra-period for CBR control. The number of target bits for an intra-period is allocated by using the scheme in [10] as follows:

$$T_{ip} = \frac{TB}{FR} \cdot N_{ip} + t_{r,ip}. \tag{1}$$

The number of generated bits per second can be controlled uniformly by constraining the number of target bits for an intra-period.

2.2 GOP Level Bit Allocation

In the proposed method, the target bits for the current GOP are allocated within the remaining bits of an intra-period by taking the generated bits of the previous GOP into consideration. Using the hierarchical structure, the target bit T_{gop} is assigned to the current GOP as follows:

$$T_{gop} = \beta \cdot \frac{\varpi_1}{\varpi_2} \cdot G_{p,gop} + (1-\beta) \cdot \frac{\varpi_1}{\varpi_3} \cdot R_{ip}, \tag{2}$$

where β is a parameter which is selected experimentally for the best performance, and ϖ_1, ϖ_2, and ϖ_3 are given by the following:

$$\begin{aligned} \varpi_1 &= \sum_{l=0}^{l_{\max}} \sum_{f \in \{I,P,B\}} w_c^{l,f} \cdot N_{c,gop}^{l,f}, \\ \varpi_2 &= \sum_{l=0}^{l_{\max}} \sum_{f \in \{I,P,B\}} w_p^{l,f} \cdot N_{p,gop}^{l,f}, \\ \varpi_3 &= \sum_{l=0}^{l_{\max}} \sum_{f \in \{I,P,B\}} w_c^{l,f} \cdot n_{ip}^{l,f}. \end{aligned} \tag{3}$$

ϖ_1 and ϖ_2 represent the sum of the weighting factors for the frames in the current GOP and previous GOP, respectively, and ϖ_3 is the sum of $w_c^{l,f}$'s for the frames that have not been encoded yet in the current intra-period. T_{gop} is equal to the weighted average of the generated bits

Table1. Notations used in the rate control algorithm.

Variable	Definition		Variable	Definition	
f	frame type, f={I,P,B}		l	Temporal level	
l_{max}	Maximum value of temporal level		$w_p^{l,f}$	Weighting factor of l-level and f-type frames	in the previous GOP
TB	Target bitrate		$w_c^{l,f}$		in the current GOP
FR	Frame rate of the sequence		N_{ip}	The number of frames	in the current intra-period
$t_{r,ip}$	The number of bits that remains after the frames in the previous intra-period are encoded.		$N_{p,gop}^{l,f}$		with level l and type f in the previous GOP
T_{ip}	The number of target bits	for the current intra-period	$N_{c,gop}^{l,f}$		with level l and type f in the current GOP
T_{gop}		for the current GOP	$n_{ip}^{l,f}$	The number of remaining frames with level l and type f	in the current intra-period
$T_{frm}^{l,f}$		for the current frame in the temporal level l and type f	$n_{gop}^{l,f}$		in the current GOP
T_{lcu}		for the current LCU	$N_{cu,lcu}$	The number of 8×8 CUs	in the current LCU
R_{ip}	The remaining number of target bits	in the current intra-period	$N_{cu,frm}$		in the remaining region of the current frame
R_{gop}		in the current GOP	$Q_c^{l,f}$	Quantization step	for the current frame with level l and type f
R_{frm}		in the current frame	$Q_{c,lcu}$		for the current LCU
$G_{p,gop}$	The number of generated bits	in the previous GOP	$QP_p^{l,f}$	Quantization parameter	for the previous frame with level l and type f
$G_{c,frm}^{l,f}$		in the current frame with level l and type f	$QP_c^{l,f}$		for the current frame with level l and type f
$G_{c,lcu}$		in the current LCU	$QP_{c,lcu}$		for the current LCU

of the previous GOP and the remaining bits in the current intra-period.

After the first two GOPs are encoded with the initial QP (IQP), the initial value of $w_c^{l,f}$ is set to the ratio of the average output bits of the frames in level l and type f to the output bits of the P-frame. From the third GOP, the weighting factor $w_c^{l,f}$ is updated using the generated bits of the frames in the current GOP for the frames of the same frame type and temporal level in the next GOP.

2.3 Frame Level Bit Allocation

For the frame level bit allocation, the scheme in [7] is used, where the number of target bits, $T_{frm}^{l,f}$, for a f-type frame in the temporal level l is assigned using the following equation:

$$T_{frm}^{l,f} = \frac{w_c^{l,f}}{w_c^{l,f} \cdot n_{gop}^{l,f} + \sum_{\substack{i=l+1,\\ f=B}}^{l_{\max}} w_c^{i,f} \cdot N_{c,gop}^{i,f}} \cdot R_{gop}. \tag{4}$$

Note that the remaining number of bits in the current GOP, R_{gop}, is distributed to each frame using the frame weights to assign the target bits adaptively according to the frame type and temporal level.

2.4 LCU Level Bit Allocation

In HEVC, a picture is composed of a sequence of LCUs with a size of M×M pixels, and each LCU can be split into smaller CUs based on a quad-tree structure. The target bits T_{lcu} are assigned to the LCU according to the size of the LCU as follows:

$$T_{lcu} = \frac{N_{cu,lcu}}{N_{cu,frm}} \cdot R_{frm}. \tag{5}$$

Note that T_{lcu} is allocated from R_{frm} using the ratio of $N_{cu,lcu}$ to $N_{cu,frm}$. The LCUs in the right-end and bottom boundaries of the frame can have different sizes, as shown in Fig. 2. Therefore, the smallest CU with 8×8 pixels, which forms any size of LCU, is exploited to assign the target bits to the LCU according to its block size.

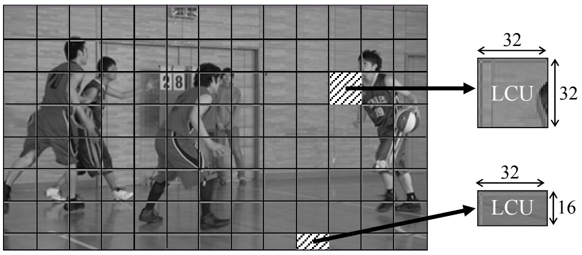

Fig. 2. Example of region partitioning with the 32×32 LCU.

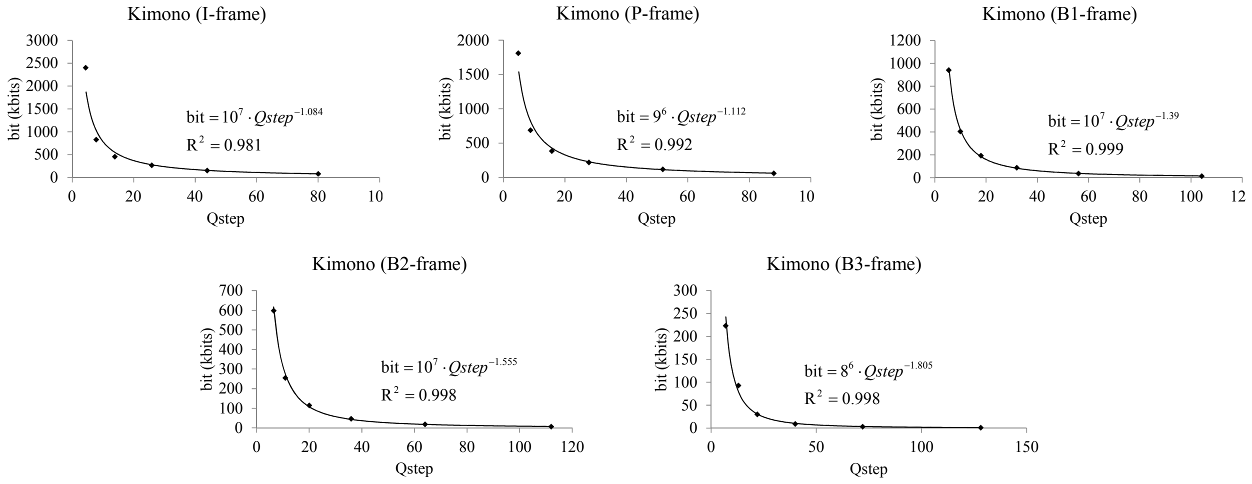

Fig. 3. Bit-*Qstep* relationship for different frame types and temporal levels for the Kimono sequence encoded by HEVC.

3. Parameter Estimation of Rate Control Model and QP Determination Methods

To generate the target bits after encoding, proper QPs should be determined by exploiting an RQ model that defines the relationship between the output bitrate and quantization. In this section, the RQ model is presented for the proposed RC algorithm and a new parameter estimation method is introduced for the RQ model. A QP decision method, in which the QP obtained from the RQ model is adjusted by considering the hierarchical structure to determine the final QP, is also proposed.

3.1 Rate Control Method for HEVC

A range of RQ models have been introduced to improve the RC efficiency of H.264/AVC. The quadratic RQ model [4, 13] adopted in the H.264/AVC reference software [8] was designed by modeling the discrete cosine transform (DCT) coefficients as the Laplacian probability density function (pdf). In [14], a linear rate model was applied to H.264/AVC by representing the rate with a proportion of non-zero quantized coefficients. Kamaci et al. [11] proposed a new RQ model based on the Cauchy pdf to estimate the distribution of DCT coefficients accurately and simply. We use the Cauchy density based RQ model defined as

$$BR \approx aQ^{-\alpha}, \tag{6}$$

where BR is the bitrate, Q is the quantization step (*Qstep*), and a and α are the parameters of the Cauchy distribution function [11]. Because the relationship between the number of bits and *Qstep* is dependent on the frame type and the temporal level in the hierarchical structure, as shown in Fig. 3, we apply the model separately for different frame types and levels by adjusting the a and α values.

3.2 Parameter Estimation Method for Rate Control Method

To utilize the RQ model for RC, a and α in Eq. (6) must first be determined. In [11], the value of α for the I-frame is determined from a predefined set of values by calculating the Cauchy parameter after the first I-frame is encoded using IQP. In addition, for P-frames and B-frames, the α values are selected from the given values by calculating the bits per pixel after encoding the first P-frame and B-frame, respectively. Then, the values of a for each frame type are calculated using Eq. (6). On the other hand, they cannot be used directly for sequences with other resolutions because the predetermined sets of α values for each frame type in [11] are defined only for low resolution sequences, such as CIF (352×288) and QCIF (176×144) sequences. In addition, the adaptability of the conventional parameter decision method is low because the value of α is selected from only three values. To this end, an efficient method is introduced to calculate the parameter values of the RQ model depending on the frame size. Furthermore, the hierarchical structure is also considered to determine the parameter values in the proposed method.

We first calculate α using the linear rate model [14] defined as

$$BR = \theta \cdot (1 - \rho), \tag{7}$$

where ρ is the fraction of zeros among the DCT coefficients after quantization, and θ is a frame-dependent constant. The linear rate model was used for some standard video codecs, such as H.263, MPEG-4, and H.264 [14, 15]. Our experiments indicate that the model also fits well to the HEVC in Fig. 4. Therefore, the rate model in Eq. (7) is used to calculate α. From Eqs. (6) and (7), the ρ-*Qstep* relationship can be expressed as

$$(1 - \rho) = cQ^{-\alpha}, \tag{8}$$

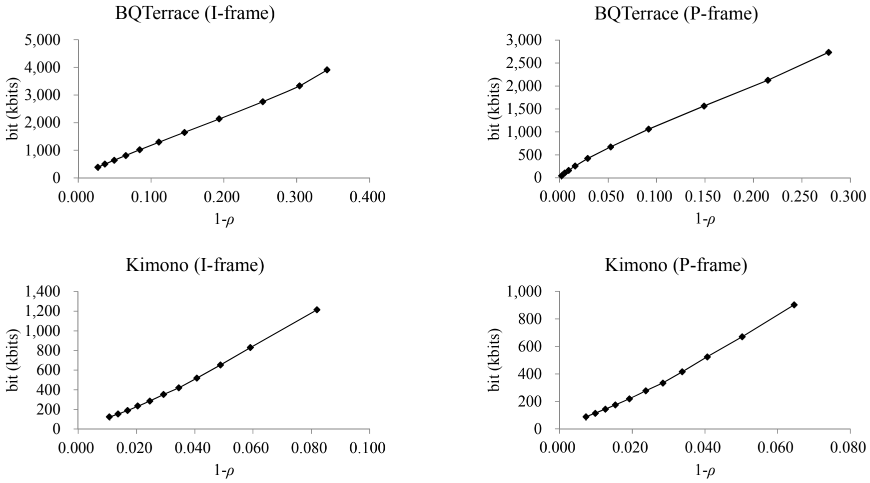

Fig. 4. Linear relationship between the proportion of nonzero quantized coefficients (1-*ρ*) and the number of output bits of I-frame and P-frame in video coding of HEVC for the 1920×1080 sequences BQTerrace and Kimono.

where c is a parameter used only to calculate α. Then, α for frame type f, α_f, which is defined in Eq. (9), is calculated using Eq. (8).

$$\alpha_f = \begin{cases} \alpha_{\mathrm{I}}, & \text{if } f = \mathrm{I} \\ \alpha_{\mathrm{P/B}}, & \text{if } f = \mathrm{P} \text{ or } \mathrm{B}, \end{cases} \tag{9}$$

After the first f-type frame (I or P) in the sequence is encoded using the IQP, the quantized DCT coefficients are collected, and the proportion of the number of zero-quantized coefficients to the total number of quantized coefficients for the IQP, ρ_0, is calculated. In addition, the DCT coefficients before quantization are quantized using IQP-2 and IQP+2 to estimate the values of ρ, i.e. ρ_1 and ρ_2 for IQP-2 and IQP+2, respectively. The estimated values of ρ for IQP-2 and IQP+2 without the re-encoding process are used because calculating the ρ values for the other QPs requires an additional encoding process with huge computational complexity. The error between the real and estimated ρ values is relatively small and acceptable when a QP in the range of IQP-2 to IQP+2 is used. Table 2 lists the average accuracy of the estimated ρ values for IQP-2 and IQP+2. Consequently, by using the obtained ρ's and corresponding *Qstep* values Q_i's, α_f^* is calculated as follows:

Table 2. Mean accuracy (%) of the estimated *ρ* values for IQP-2 and IQP+2, respectively, for each frame type (totally 10 IQPs in the range of 20 to 40 are tested).

Target QP of ρ estimation	Frame type	
	I-frame	P/B-frame
IQP-2	98.21	98.68
IQP+2	99.31	97.99

$$\alpha_f^* = \arg\min \left[\sum_{i=0}^{2} \left(1 - \rho_i - c Q_i^{-\alpha_f} \right)^2 \right]. \tag{10}$$

It is verified from our experiments that the value of α calculated using Eq. (10) is similar to that obtained from the actual encoding, as shown in Table 3.

Table 3. Comparison of *α* values, which are obtained from the actual encoding and using Equation (10) when the Kimono sequence is encoded using IQP=22 and 32.

	IQP=22		IQP=32	
Frame type	actual encoding	(10)	actual encoding	(10)
I-frame	1.20	1.26	0.90	0.76
P/B-frame	1.13	0.96	0.94	0.96

The a and α values are different depending on the frame type and temporal level, as shown in the equations in Fig. 3. However, in the proposed method, $\alpha_{\mathrm{P/B}}$ is commonly used for P-frames and B-frames regardless of the temporal level, whereas an independent value α_{I} is used for I-frames. $\alpha_{\mathrm{P/B}}$ and α_{I} are fixed for all LCUs and frames. On the other hand, the value of a in Eq. (6) is allowed to vary according to the frame type and temporal level and to be updated for each frame and LCU to accommodate the change in the original α value [16].

The a for the temporal level l and frame type f, $a_{l,f}$, is initially calculated using Eq. (6), α_f, and the generated bits obtained after the first two GOPs are encoded using the IQP. In addition, the initial value of $a_{l,f}$ is used as the initial value of a for LCU, a_{lcu}. For the next frame in the sequence with the same frame type and temporal level, the parameter $a_{l,f}$ is updated at each frame as follows [11]:

$$\tilde{a}_{l,f} = \gamma \times a_{l,f} + (1-\gamma) \times \frac{G_{c,frm}^{l,f}}{(Q_c^{l,f})^{-\alpha_f}}, \tag{11}$$

where $\tilde{a}_{l,f}$ indicates the updated value, and γ denotes the weight factor. Similarly, the updated value of a_{lcu}, $\tilde{a}_{lcu}$, is calculated at each LCU as follows:

$$\tilde{a}_{lcu} = \gamma \times a_{lcu} + (1-\gamma) \times \frac{G_{c,lcu}}{(Q_{c,lcu})^{-\alpha_f}}. \tag{12}$$

3.3 QP Determination Method

The QP is calculated using the RQ model in Eq. (6) to generate the target bits. However, because the values of the parameters a and α need to be determined before the RQ model can be used, the first two GOPs of the sequence are encoded using the IQP. The IQP for each frame type f and temporal level l is defined as

$$\mathrm{IQP} = \begin{cases} \mathrm{IQP_I}, & \text{if } l = 0, f = \mathrm{I} \\ \mathrm{IQP_P}, & \text{if } l = 0, f = \mathrm{P} \\ \mathrm{IQP}_{l,\mathrm{B}}, & \text{if } l > 0, f = \mathrm{B}. \end{cases} \tag{13}$$

For 4:2:0 format sequences, the IQP is set to

$$\mathrm{IQP_I} = \begin{cases} 20, & s_1 < bpp \\ 25, & s_2 < bpp \le s_1 \\ 30, & s_3 < bpp \le s_2 \\ 35, & s_4 < bpp \le s_3 \\ 40, & bpp \le s_4, \end{cases} \tag{14}$$

$$\mathrm{IQP_P} = \mathrm{IQP_I} + 1, \tag{15}$$

$$\mathrm{IQP}_{l,\mathrm{B}} = \mathrm{IQP_P} + l, \tag{16}$$

where the values of s_i's, $1 \le i \le 4$, are predetermined experimentally by encoding all sequences using the QPs from 15 to 45 with an increment of 5 and averaging the output bits per pixel for each QP, and

$$bpp = \frac{TB}{FR \times H_f \times W_f \times 1.5}, \tag{17}$$

where H_f and W_f represent the height and width of the frame, respectively. The constant value 1.5 is used to calculate the number of pixels in a frame with YUV 4:2:0 format.

From the third GOP, the QP of the current frame with level l and type f, $QP_c^{l,f}$, is calculated from the RQ model in Eq. (6). Initially, the preliminary QP of frame, QP_{frm}^*, is obtained by calculating the *Qstep* value, Q_{frm}^*, for the given target bits $T_{frm}^{l,f}$. The Q_{frm}^* is calculated using the following equation:

$$Q_{frm}^* = \left(\frac{T_{frm}^{l,f}}{a_{l,f}} \right)^{-\frac{1}{\alpha_f}}. \tag{18}$$

The obtained QP_{frm}^* is adjusted to determine the final QP of frame $QP_c^{l,f}$, i.e.

$$QP_c^{l,f} = \min\{QP_p^{l,f} + offset_l, \max\{QP_p^{l,f} - offset_l, QP_{frm}^*\}\}. \tag{19}$$

To minimize the fluctuation in the decoded picture quality between consecutive frames, the variation in QP is limited using the offset in Eq. (19). In the GOP, because the number of frames in temporal levels with $l > 1$ is greater than two, the frame QPs can be changed more frequently than in the temporal levels with l=0 or 1. Therefore, the value of $offset_l$ in levels $l > 1$ is set to be smaller than that in level 0 and 1 as follows:

$$offset_l = \begin{cases} 2, & \text{if } l = 0,1 \\ 1, & \text{otherwise.} \end{cases} \tag{20}$$

On the other hand, if $T_{frm}^{l,f} < 0$, $QP_c^{l,f} = QP_p^{l,f} + 2$ without a QP decision process in Eqs. (18) and (19).

A QP for the current LCU, $QP_{c,lcu}$, is determined on a case by case basis.

Case 1) When the LCU is the first one in the current frame, the QP is given by

$$QP_{c,lcu} = QP_c^{l,f}. \tag{21}$$

Case 2) When $T_{lcu} < 0$, QP is determined to be greater than that of the current frame as follows:

$$QP_{c,lcu} = QP_c^{l,f} + 2. \tag{22}$$

Case 3) Otherwise, the preliminary QP of LCU, QP_{lcu}^*, is first determined using Q_{lcu}^* in Eq. (23) for the given target bits T_{lcu}.

$$Q_{lcu}^* = \left(\frac{T_{lcu}}{a_{lcu}} \right)^{-\frac{1}{\alpha_f}}. \tag{23}$$

Then, it is further bounded to obtain the final QP of the LCU as follows:

$$QP_{c,lcu} = \min\{QP_c^{l,f} + 5, \max\{QP_c^{l,f} - 5, QP_{lcu}^*\}\}. \tag{24}$$

The $QP_{c,lcu}$, which is restricted to be within the range of 0 to 51, is also used for every CUs of the current LCU. In HEVC, the encoded CU structure in an LCU is determined by recursive encoding and rate-distortion optimization process for all possible CUs. Therefore, determining the proper QPs for all CUs in the LCU is a difficult and complicated task in RC. Furthermore, it is cautious to find the QP for each CU because it is unclear if the CU is used as a real encoded CU. Therefore, the $QP_{c,lcu}$ is simply used for all CUs in the LCU regardless of the CU size and encoded CU structure in LCU.

4. Experimental Results

The proposed RC algorithm is implemented using the HEVC reference software, HM4.0 [17]. Various test sequences with 2560×1600 and 1920×1080 resolutions in [12] are encoded in the hierarchical coding structure using the random access-high efficiency (RA-HE) test conditions in [12]. In the configurations of RA-HE, because the B-frame is encoded in the temporal level 0 instead of the P-frame, the value of $w_c^{l,f}$ is set to the ratio of the average output bits of the frames in level l and type f to the output bits of the B-frame in level 0. The proposed techniques for the P-frame are performed on the B-frame in level 0. The performance of the proposed RC algorithm is compared with that of [3] and that of the fixed QP coding without RC of HEVC. The target bitrate is set to the value obtained by encoding the sequence using HM with QP=22 and 32. In a fixed QP coding, the QP of the I-frame is given, and the QPs for the B-frames are varied in increments of 1 according to the hierarchy of the frame as indicated in Eqs. (15) and (16). Table 4 lists the s_i values in Eq. (14) for different resolutions. To obtain the best performance, the values of β in Eq. (2) and γ in Eqs. (11) and (12) are set to 0.3 and 0.5, respectively. To find the best values of β and γ, simulations were performed with various possible values of β and γ.

Tables 5 and 6 show the RC results associated with the output bitrate and Y-PSNR. In Table 5, the target bitrate is given by the result of fixed QP coding with QP=22. The output bitrate of fixed QP coding with QP=32 is used as the target bitrate in Table 6. Table 5 shows that the proposed algorithm controls the bitrate accurately for all test sequences. The proposed method achieves approximately 99% RC accuracy for most sequences, and an approximately 99.99% bitrate accuracy is obtained for the BasketballDrive sequence with the Y-PSNR loss of 0.28dB. Moreover, although Y-PSNR obtained after RC using [3] is slightly higher than that using the proposed method for most sequences, the accuracy of the result bitrate of the proposed method is higher than that of [3] for all sequences. In Table 6, although the RC performance is relatively low when the target bitrate is low, the proposed RC method controls the bitrate precisely and achieves high PSNR performance without significant Y-PSNR degradation. For the Nebuta sequence, the Y-PSNR of the proposed RC algorithm is the same as that of fixed QP coding, with a bitrate accuracy of 99.31%. Compared to the RC results of [3], the proposed method shows better performance for most sequences.

Table 4. s_i values in Eq. (14) for different frame resolutions.

Size	s_1	s_2	s_3	s_4
2560×1600	0.8	0.5	0.3	0.2
1920×1080	0.7	0.3	0.2	0.1

Fig. 5 compares the number of bits per second generated by the proposed RC method and the target bits per second. The proposed method can achieve the CBR

Table 5. Experimental results on the test sequences when the target bitrate is determined by fixed QP coding with QP=22.

size	sequence	Frame number	Frame rate (fps)	Rate control method	Bitrate (kbps)	Y-PSNR (dB)	Inaccuracy compared to target	
							Bitrate (%)	Y-PSNR (Δ dB)
2560×1600	Nebuta	300	60	Target	218548.20	39.69	-	-
				[3]	217289.08	39.34	0.58	-0.35
				Proposed	217769.10	39.02	0.36	-0.67
	SteamLocomotive	300	60	Target	23392.29	41.67	-	-
				[3]	24823.39	41.12	6.12	-0.55
				Proposed	23416.02	40.89	0.10	-0.78
	PeopleOnStreet	150	30	Target	33106.84	40.50	-	-
				[3]	33620.84	39.74	1.55	-0.76
				Proposed	33163.65	38.61	0.17	-1.89
1920×1080	BQTerrace	600	60	Target	39752.68	37.59	-	-
				[3]	39860.57	37.30	0.27	-0.29
				Proposed	39705.23	37.40	0.12	-0.19
	BasketballDrive	500	50	Target	17386.66	39.33	-	-
				[3]	17297.49	38.71	0.51	-0.62
				Proposed	17384.22	39.05	0.01	-0.28
	Cactus	500	50	Target	18127.03	38.63	-	-
				[3]	18063.47	37.96	0.35	-0.67
				Proposed	18099.77	37.70	0.15	-0.93

Table 6. Experimental results of test sequences when the target bitrate is determined by fixed QP coding with QP=32.

size	sequence	Frame number	Frame rate (fps)	Rate control method	Bitrate (kbps)	Y-PSNR (dB)	Inaccuracy	
							Bitrate (%)	Y-PSNR (Δ dB)
2560×1600	Nebuta	300	60	Target	28513.34	29.71	-	-
				[3]	28370.56	29.40	0.50	-0.31
				Proposed	28317.59	29.71	0.69	0.00
	SteamLocomotive	300	60	Target	2497.76	38.94	-	-
				[3]	2457.44	36.88	1.61	-2.06
				Proposed	2503.59	37.80	0.23	-1.14
	PeopleOnStreet	150	30	Target	8326.86	34.51	-	-
				[3]	8200.15	33.35	1.52	-1.16
				Proposed	8260.73	33.72	0.79	-0.79
1920×1080	BQTerrace	600	60	Target	2274.15	33.99	-	-
				[3]	2329.59	33.14	2.44	-0.85
				Proposed	2250.12	32.65	1.06	-1.34
	BasketballDrive	500	50	Target	2848.75	35.81	-	-
				[3]	2841.71	35.17	0.25	-0.64
				Proposed	2840.29	35.27	0.30	-0.54
	Cactus	500	50	Target	2723.51	35.09	-	-
				[3]	2715.39	34.37	0.30	-0.72
				Proposed	2717.11	33.76	0.23	-1.33

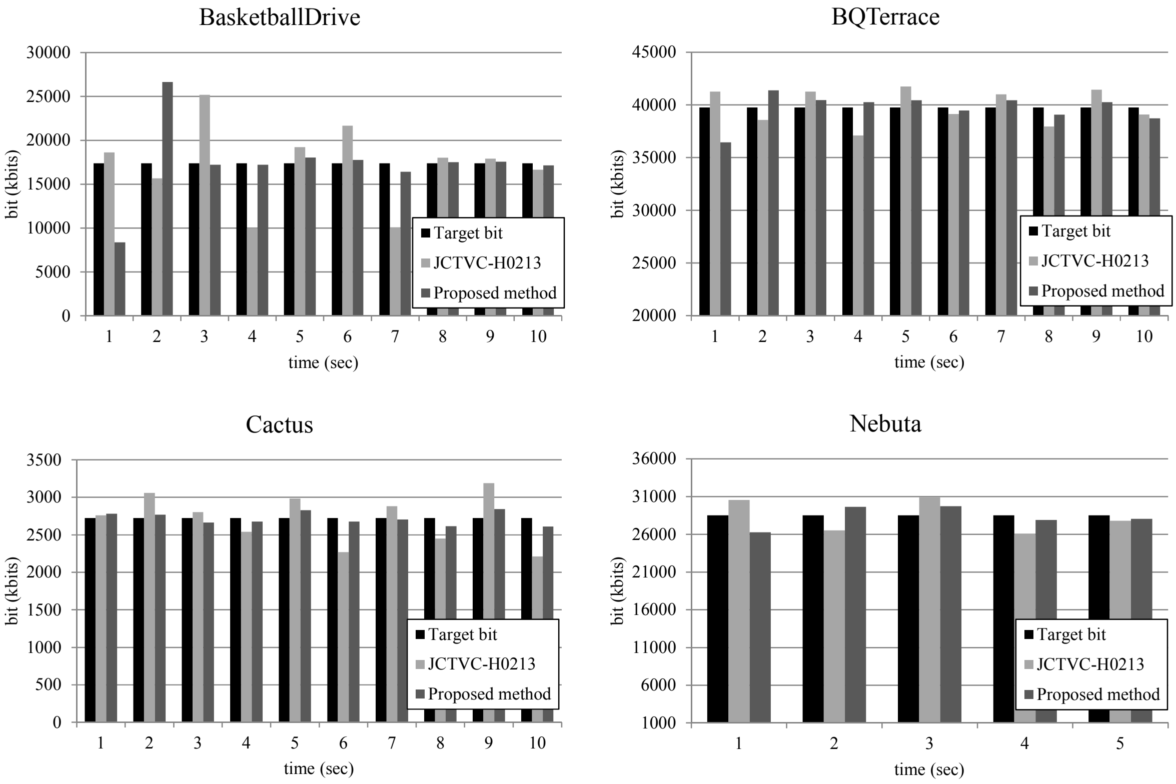

Fig. 5. Bits per second generated by the proposed rate control method and method in [3] and the target bits per second. Target bit is set by fixed QP coding with QP=22 for the BasketballDrive and BQTerrace sequences and with QP=32 for the Cactus and Nebuta sequences.

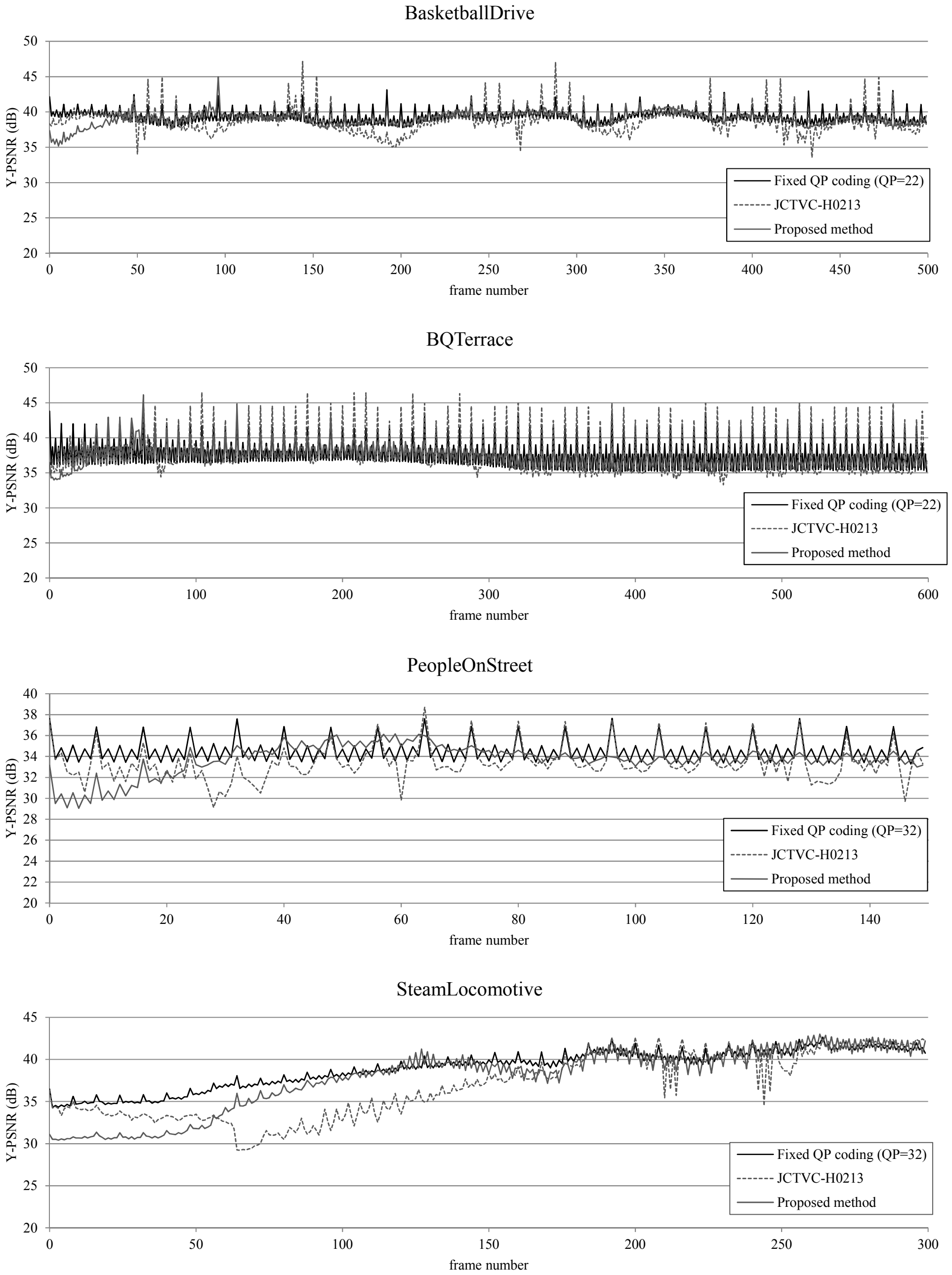

Fig. 6. Y-PSNR performance of the fixed QP coding, the method in [3], and the proposed rate control method.

RC because the number of output bits generated per second by the proposed RC method is almost constant and close to the target bits. On the other hand, the number of bits generated using the method in [3] is less stable and not maintained uniformly compared to the proposed method. Because the intra-period level bit allocation in the proposed method controls the number of bits generated per second, the CBR RC can be accomplished using the proposed method. With the method of [3], however, it is difficult to obtain the constant number of bits generated per second because the number of output bits for an intra-period is not constrained.

Fig. 6 shows the Y-PSNR curve versus frame number for some test sequences. Generally, the fixed QP coding method is more effective in terms of the PSNR performance than the RC scheme. This is because fixed QP coding can effectively utilize the characteristics of the hierarchical structure to improve the coding performance without any constraints on the generated coding bits, whereas the proper QP should be calculated by considering the target bit budgets in the RC scheme [9]. Nevertheless, the Y-PSNR of the proposed method is similar to that of fixed QP coding for most frames except at the start of the sequence. Moreover, the proposed method produces smaller differences in Y-PSNR between consecutive frames than with fixed QP coding. Compared to the method of [3], Y-PSNR of the proposed method has similar or higher values except at the start of the sequence.

The proposed method does not create significant fluctuations in the decoded picture quality and also provides high PSNR performance. On the other hand, the fluctuations of Y-PSNR obtained from [3] are very large for many frames, and the method of [3] sometimes produces steep degradation between consecutive frames.

5. Conclusion

An LCU level RC algorithm that employs the framework of hierarchical video coding for HEVC is presented. An efficient bit allocation method, which considers the different coding properties of each frame type and the temporal level in hierarchical coding, was proposed. An LCU level bit allocation scheme was introduced based on the new coding block structure of a HEVC. A new parameter estimation method for the Cauchy density based RQ model, where the linear rate model is used to simply and accurately calculate the parameter values, was also proposed. The experimental results suggest that the proposed algorithm achieves high performance of CBR control with negligible degradation of the decoded video quality for most sequences tested. Overall, the proposed RC algorithm produces successful results when the hierarchical structure is used for video coding of HEVC. In future work, the proposed LCU level RC will be extended to a smaller CU level using the contents complexity of the CU and develop the initial QP setting method to improve the current RC performance.

References

[1] B. Bross, W.-J. Han, J.-R. Ohm, G. J. Sullivan, and T. Wiegand, “Working draft 4 of high-efficiency video coding,” JCTVC-F803, JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Jul. 2011.

[2] T. Wiegand, J.-R. Ohm, G. J. Sullivan, W.-J. Han, R. Joshi, T. K. Tan, and K. Ugur, “Special section on the joint call for proposals on high efficiency video coding (HEVC) standardization,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 12, pp. 1661-1666, Dec. 2011. Article (CrossRefLink)

[3] H. Choi, J. Nam, J. Yoo, D. Sim, and I. V. Bajić, “Rate control based on unified RQ model for HEVC,” JCTVC-H0213, JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Feb. 2012.

[4] Z. Li, F. Pan, K. P, Lim, G. Feng, X. Lin, and S. Rahardja, “Adaptive basic unit layer rate control for JVT,” JVT-G012-r1, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Mar. 2003.

[5] Z. G. Li, W. Gao, F. Pan, S. W. Ma, K. P. Lim, G. N. Feng, X. Lin, S. Rahardja, H. Q. Lu, and Y. Lu, “Adaptive rate control for H.264,” *J. Vis. Commun. Image Represent.*, vol. 17, no. 2, pp. 376-406, Apr. 2006. Article (CrossRef Link)

[6] H. Schwarz, D. Marpe, and T. Wiegand, “Hierarchical B pictures,” JVT-P014, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Jul. 2005.

[7] A. Leontaris and A. M. Tourapis, “Rate control reorganization in the Joint Model (JM) reference software,” JVT-W042, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Apr. 2007.

[8] JM reference software, Article (CrossRef Link).

[9] M. Li, Y. Chang, F. Yang, S. Wan, S. Lin, and L. Xiong, “Frame layer rate control for H.264/AVC with hierarchical B-frames,” *Signal Process: Image Commun.*, vol. 24, no. 3, pp. 177-199, Mar. 2009. Article (CrossRef Link)

[10] 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. Article (CrossRef Link)

[11] N. Kamaci, Y. Altunbasak, and R. Mersereau, “Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 8, pp. 994-1006, Aug. 2005. Article (CrossRef Link)

[12] F. Bossen, “Common test conditions and software reference configurations,” JCTVC-F900, JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Jul. 2011.

[13] T. Chiang and Y. 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. Article (CrossRef Link)

[14] Z. He and D. O. Wu, “Linear rate control and optimum statistical multiplexing for H.264 video broadcast,” *IEEE Trans. Multimedia*, vol. 10, no. 7, pp. 1237-1249, Nov. 2008. Article (CrossRef Link)

[15] 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. Article (CrossRef Link)

[16] J.-J. Chen and H.-M Hang, “Source model for transform video coder and its application-Part II: Variable frame rate coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 2, pp. 299-311, Apr. 1997. Article (CrossRef Link)

[17] HM reference software, https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware.

Yeo-Jin Yoon received the B.S. degree in Electronics Engineering from Korea University, Seoul, in 2008. She is a Ph.D. candidate in the Department of Electronics Engineering at Korea University. Her research interests are image processing and video coding.

Hoon Kim received a B.S. degree from Korea University in electronic engineering, in 2010. He is currently pursuing a Ph.D. degree at the School of Electrical Engineering at Korea University. His research interests are in the areas of image processing and computer vision.

Seung-Jin Baek received the B.S. degree in Electrical Engineering from Korea University, Seoul, in Feb. 2007, where he is currently pursuing the Ph.D. degree in Electrical Engineering. His research interests include video signal processing, computer vision, and video compression codec.

Sung-Jea Ko received the Ph.D. degree in 1988 and the M.S. degree in 1986, both in Electrical and Computer Engineering, from State University of New York at Buffalo, and the B.S. degree in Electronic Engineering at Korea University in 1980. In 1992, he joined the Department of Electronic Engineering at Korea University where he is currently a Professor. From 1988 to 1992, he was an Assistant Professor of the Department of Electrical and Computer Engineering at the University of Michigan-Dearborn. He has published over 160 international journal articles. He also holds over 60 patents registered on video signal processing, computer vision, and multimedia communications. He is the 1999 Recipient of the LG Research Award given to the Outstanding Information Technology and Communication Researcher. He received the Hae-Dong best paper award from the Institute of Electronics Engineers of Korea (IEEK) (1997), the best paper award from the IEEE Asia Pacific Conference on Circuits and Systems (1996), the research excellence award from Korea University (2004), and the technical achievement award from the IEEE Consumer Electronics (CE) Society (2012). He has served as a TPC member for the IEEE Conference on Consumer Electronics (ICCE) since 1997 and received a 10-year service award from the TPC of ICCE in 2008. He is a member of the editorial board of the IEEE Transactions on Consumer Electronics. He is the President of the IEEK, the Vice-President of the IEEE CE Society, and an IEEE Fellow.