HEVC Rate Control Optimization Algorithm Based on Video Characteristics

Qiang Li and Jun Nie(✉)

School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China
liqiang@cqupt.edu.cn, 724413152@qq.com

Abstract. Based on the relationship between the frame-level and the large coding unit (LCU)-level in the high-efficiency video coding (HEVC) rate control scheme, this paper proposes novel optimization algorithms for bit allocation at the frame-level and improves algorithms for code rate control in LCU level for the space-time domain. For the frame-level, the statistical characteristics of the whole source are considered, and the information entropy of the coded source is added to the R-lambda model. An optimization model of bit allocation considering the current coded frame is used to guide the bit allocation in the frame layer. This algorithm improves the bit accuracy in the frame level. For the LCU-level, the Hadamard transform algorithm detects the energy distribution region and motion region. It constructs a new complexity from the energy value and the predicted residual value of the image so that the target bits of the LCU layer can be reasonably adjusted and accurately assigned. The coding structure in LCU-level achieves a more accurate update of the model parameters. The experimental results show that the algorithm in this paper reduces the relative error of bit rate by 0.017% and 0.016% on average and improves the rate-distortion performance by 2.7% and 2.6% on average compared with the adaptive ratio bit allocation algorithm under low-delay B configuration and P configuration.

Keywords: HEVC · Rate control · Space-time domain · Information entropy

1 Introduction

With the rapid development of computers, the Internet, intelligent terminals in today's era, people's video demands are increasing. It also causes the increasing amount of data. The International Telecommunication Union ITU-T and the International Organization for Standardization ISO/IEC proposed the high-efficiency video coding (HEVC) standard [1] in 2013 to improve the coding efficiency of video compression and meet the needs of high-definition and ultra-high-definition video compression in practical applications. In practical applications such as surveillance and video conferencing, ensuring the quality of encoded video with limited bandwidth resources is the problem to solve by bit rate control technology. Therefore, bit rate control technology is an essential process

H. Gao et al. (Eds.): ChinaCom 2021, LNICST 433, pp. 235–249, 2022.
https://doi.org/10.1007/978-3-030-99200-2_19

in video compression coding. Code rate control is a decision problem in video compression. How to allocate the code rate efficiently and how to convert the target allocated code rate into the actual encoded bit rate are the two most important problems need to be solved by code rate control techniques. These two problems also affect the final coding efficiency and observation effect. In the evolving video compression technology, the bit rate control part is also ceaselessly updated. The bit rate control algorithm has been a hot research topic in video compression technology. Therefore, the study of bit rate control algorithm in video compression technology has significant practical value.

The rate control algorithm allocates the target number of bits to the group of pictures (GOP) level, image level, a large coding unit (LCU)-level through a particular strategy and then calculates the Lagrange multiplier (λ) by the target number of bits, Update the model parameters. In the process of formulating the high-efficiency video coding standard HEVC/H.265, a series of continuously improving bit rate allocation schemes have also been successfully formed. The main influential ones are JCTVC-H0213 [1] and JCTVC-K0103 [2]. The JCTVC-H0213 proposal is the first-rate control algorithm proposed by HEVC. This algorithm has certain defects, mainly due to the allocation model's unreasonable setting, which leads to a large gap between the code rate generated after the final encoding and the expected rate, and the video. The quality after compression is not very good either. Because of the JCTVC-H0213 rate control algorithm's shortcomings, the rate implementation model in the JCTVC-K0103 control algorithm is modified, and a rate implementation model based on the R-λ-QP model is proposed, which significantly improves the insufficiency of previous rate control algorithm. K0103 is also not the optimal algorithm. Because of its shortcomings, the follow-up meeting offered JCTVC-M0257 [3], JCTVC-M0036 [4], and other proposals. In recent years, many rate control algorithms based on the R-λ model have emerged. The work [5] found that there is a more robust correspondence between R and the λ, and proposed a novel λ-domain rate control algorithm, which has been implemented in the latest video coding standard. The authors [6] proposed image-level and basic unit-level bit allocation algorithms based on basic RD optimization theory, which makes full use of video content to guide bit allocation. The work [7] took into account the characteristics of inter-frame correlation. A precise H.265/HEVC frame-level bit allocation algorithm is developed, which improves coding efficiency. In [8], considering the recursive rate-distortion model with dependence between frames caused by motion compensation, the algorithm treats frame-level rate allocation as a convex optimization problem.[9] used the recursive Taylor expansion equation to solve the constraint equation, and the optimized R-D model performs target bit allocation on the CTU layer. [10] proposed an improved R-λ to establish a rate control model based on joint time-space information and HVS characteristics. In this model, mutual time-space information based on gradient information is used to guide bit allocation at the frame and CTU level, where the time coefficient is adaptively corrected. The work [11] proposed a gradient-based R-lambda model for intra-frame rate control, which can effectively measure the frame content complexity and enhance traditional R-lambda methods' performance. And the work developed LCU-level bit allocation method.

Although there are many work designs on rate control algorithms, most studies have not considered the impact of the frame levels target bit on the bit allocation of the LCU-level and the coding characteristics of HEVC itself. This paper proposes an optimized Frame-level and LCU-level rate control algorithm. The code rate control algorithm further improves the coding performance. Unlike the traditional way of frame-level bit weight allocation, it also considers the amount of information in the aggregation feature of the gray distribution in the current frame. The purpose is to make the bit allocation of the current frame fully consider the current source's overall information measurement. Besides, based on the space-time theory of the LCU-level, a new linear bit allocation weight calculation method for the LCU-level is proposed. Experimental results show that the proposed rate control algorithm further improves the coding performance. In summary, the main contributions of this article are as follows:

In this article, a rate control algorithm incorporating the current coded frame's information entropy into the R-D model is proposed. Therefore, compared with the traditional frame-level bit allocation algorithm, the proposed rate control algorithm can make the actual rate closer to the target rate and improve the encoding quality to a certain extent.

The new LCU-level bitrate control algorithm takes account of the Hadamard transform algorithm, and takes more fully into account each frame's texture characteristics, time-domain prediction information, and the video content characteristics. Besides, the algorithm designs two parts to control them.

The rest of this article is organized as follows. The second section introduces the relevant technical principles, and the third section discusses the rate control scheme at the frame and LCU-level. Next, the experimental results and conclusions are presented in Sect. 4 and Sect. 5, respectively.

2 Related Technical Principles

Like other rate control schemes, the R-λ rate control model is mainly divided into two parts, the one is bit allocation, the other is the calculation of the λ and the quantization parameter QP. Literature [5] established the exponential relationship between the code rate and the λ, and its model is:

$$\lambda = -\frac{\partial D}{\partial R} = CK \cdot R^{-K-1} \triangleq \alpha R^{\beta} \tag{1}$$

Among them, D is the encoding distortion. C and K are model parameters related to the video sequence. α and β are model parameters related to the characteristics of the video content. R represents the coding bit, the unit is *bpp* (bit per pixel). If the target bit of a particular frame or a certain LCU is T and the number of pixels is N, the calculation equation is:

$$\text{bpp} = \frac{T}{N} \tag{2}$$

Among them α and β are model parameters, which will be updated as each LCU or frame is encoded. Then QP can be determined by the empirical equation [5].

$$QP = 4.2005 \ln \lambda + 13.7122 \tag{3}$$

Bit allocation will be implemented in the GOP level, frame-level, and basic coding unit level. First, calculate the target number of bits for each picture.

$$R_{PicAvg} = \frac{R_{\text{tar}}}{\text{f}} \tag{4}$$

Suppose the number of encoded pictures is $\text{N}_{\text{coded}}\text{N}_{\text{coded}}\text{N}_{\text{coded}}$ N_{coded}, the number of bits used by these pictures is R_{coded}, the number of frames in the current GOP is N_{GOP}, SW is the size of the sliding window for smooth bit allocation, which is used to make bit consumption changes, and the quality of encoded pictures more smooth. The bit allocation of the GOP level is:

$$T_{AvgPic} = \frac{R_{PicAvg} \times (N_{\text{coded}} + SW) - R_{\text{coded}}}{SW} \tag{5}$$

$$T_{GOP} = T_{AvgPic} \cdot N_{GOP} \tag{6}$$

We hope to reach the target bit rate after the SW frame. If the SW frame can consume bits in each frame T_{AvgPic}, the above equation can be rewritten as:

$$T_{AvgPic} = R_{PicAvg} + \frac{R_{PicAvg} \cdot N_{\text{coded}} - R_{coded}}{SW} \tag{7}$$

The first part of equation represents the target bit rate, and the second part describes the buffer state.

Then there is the bit allocation at the picture level. Suppose the number of bits used in the current GOP is $Coded_{GOP}$, ω is the bit allocation weight of each picture, so the target bit rate of the recent frame is:

$$T_{CurrPic} = \frac{T_{GOP} - Coded_{GOP}}{\sum\limits_{NotCodedPictures} \omega_{\text{i}}} \cdot \omega_{CurrPic} \tag{8}$$

The bit allocation of the LCU-level is considered in the proposal that a fundamental unit contains each LCU, and the following equation determines the target number of bits:

$$T_{CurrLCU} = \frac{T_{CurrPic} - Bit_{\text{header}} - Coded_{Pic}}{\sum\limits_{NotCodedLCUs} \omega_{\text{i}}} \cdot \omega_{CurrLCU} \tag{9}$$

Among them Bit_{header} is the estimated value of all header information bits, which is estimated by the actual number of header information bits of the coded picture in the same layer.

3 Proposed Rate Control Model

3.1 Frame-Level Bit Allocation Model

In this section, this paper first proposes a frame-level bit allocation model based on information entropy. The amount of information is used to measure an event's uncertainty, and image entropy refers to the average amount of information in the image. The greater the probability of an event, and the smaller the uncertainty, the smaller the amount of information it carries; conversely, the greater the entropy, the richer the information. To describe that the amount of information in different messages is different, we use the mathematical expectation of self-information to the amount of information, called information entropy, $\chi_i(i = 1, 2, \cdots, \mathrm{n})$ means that an information source sends out n symbolic messages, $p(\chi_i)$ means the probability of occurrence of different symbolic messages. The calculation equation is shown in (10):

$$H(X) = E[I(\chi_\mathrm{i})] = E\left(\log_2 \frac{1}{\mathrm{p}(\chi_\mathrm{i})}\right) = -\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{p}(\chi_\mathrm{i}) \times \log_2 \mathrm{p}(\chi_\mathrm{i}) \tag{10}$$

Among them $I(\chi_i)$ is the self-information of the symbolic message χ_i, $p(\chi_i)$ represents the probability of the symbolic message χ_i appearing, and n represents the total number of symbolic messages.

Because the image information entropy can characterize the overall characteristics of the source and describe the local details of the image, it can well reflect the amount of information contained in the aggregation features of the grayscale distribution in the image. Image information entropy can well reflect the complexity of the image. The calculation equation of image information entropy is as follows:

$$EI = -\sum_{\chi=0}^{N-1} \mathrm{p}(\chi) \times \log_2\left[\mathrm{p}(\chi)\right] \tag{11}$$

EI is the image's information entropy, $p(\chi)$ is the proportion of pixels with gray value χ in the image, and N is the total number of image gray levels. Each pixel gray level is quantized with 8 bits, and the total number of gray image levels is 256 gray levels.

This algorithm considers the frame-level fixed weight allocation model's information entropy in the original R-lambda model rate control algorithm. Unlike the traditional linear allocation model, the higher the frame level's information entropy, the more allocated bits. The current encoded frame has different bit demands on the frame level, and the information entropy can represent the value of the information. Equation (12) accumulates all the weights and information entropy within a GOP and finds the multiplier between them, ensuring that the weights and information entropy are guaranteed to be of the same order of magnitude.

$$A\mathrm{f} = \sum_{AllPictures} \omega_\mathrm{i} / \sum_{AllPictures} EI_\mathrm{i} \tag{12}$$

The bit allocation algorithm takes into account the information entropy of the currently encoded frame. The calculation equations for the bit allocation weight ω'_{pic} and total bit weight ω_{total} of the new frame-level can be obtained as follows:

$$\omega'_\mathrm{pic} = \mathrm{Af} \times EI_\mathrm{i} + \omega_{\mathrm{original}_\mathrm{i}} \tag{13}$$

$$\omega_\mathrm{total} = \sum_{NotCodedPictures} \omega_\mathrm{i} + \sum_{NotCodedPictures} (EI_\mathrm{i} \times A\mathrm{f}) \tag{14}$$

Among them EI_i is the information entropy of the current frame image, $\omega_{original_i}$ representing the weight of the current frame image in the rate control algorithm based on the R-λ model, and $\sum_{NotCodedPictures} \omega_i$ assigns the sum of all uncoded images in the current image group.

Because the frame-level bit allocation weight does not fully consider each frame's content's characteristics and allocates the target bits of the frame-level more reasonably, this section proposes a novel optimized frame-level target bit allocation algorithm. The layer bit allocation weight ω'_{pic} and the new total bit weight ω_{total} perform high-precision bit allocation to the frame. Therefore, equation for the allocation of target bits at the frame-level in this algorithm is shown in Eq. (15):

$$T_{\mathrm{Curr}P\mathrm{ic}} = \frac{T_{GOP} - C\mathrm{oded}_{GOP}}{\omega_{\mathrm{total}}} \cdot \omega'_{\mathrm{pic}} \tag{15}$$

3.2 LCU-Level Bit Allocation Model

Hadamard transform is often used in image and video processing to calculate the residual signal's SATD (sum of absolute transformed difference). SATD calculates the sum of each element's absolute value after Hadamard transforms the residual signal. Suppose the square matrix of residual signal is X, then SATD is:

$$SATD = \sum_{M} \sum_{M} |HXH| \tag{16}$$

Where M is the size of the square matrix, and H is the $M \times M$ normalized Hadamard matrix.

Since the SATD value reflects some extent, the energy magnitude of the residual signal in the frequency domain and the size of the coded output stream, the energy of the LCU can be statistically derived by calculating the SATD value of each sub-block of the LCU. The energy magnitude of each LCU in a frame is allocated accordingly to improve the quality of video coding.

First, split each LCU into 8 × 8 sub-blocks and using the Hadamard matrix $H_{8\times 8}$. Use Eq. (16) to calculate each sub-block's SATD value and then using Eq. (17) to calculate the SATD value D'_1 of each LCU.

$$D'_1 = \sum_{\mathrm{i}=1}^{\mathrm{m}} SATD_{\mathrm{i}} \tag{17}$$

Among them, m is the number of LCU sub-blocks; $SATD_i$ is the SATD value of the i-th LCU sub-block.

Then, using Eq. (18) to calculate the average SATD value D'_2 of a frame.

$$D'_2 = \frac{1}{\mathrm{n}} \sum_{\mathrm{j}=1}^{\mathrm{n}} SATD_{\mathrm{j}} \tag{18}$$

Among them, n is the number of LCUs in a frame; $SATD_j$ is the SATD value of the j-th LCU in the current frame.

Finally, according to D_1' of each LCU and D_2' of the current frame, Eq. (19) calculates the energy proportion factor θ_1 that can reflect each LCU in the frame. The larger the value of θ_1, the larger the energy ratio of the LCU, and the more bits should be allocated.

$$\theta_1 = \frac{D_1'}{D_2'} \tag{19}$$

Inter-frame prediction refers to the method which uses the video time domain's correlation to predict the pixels of the current image based on adjacent coded image pixels in order to remove redundant information in the video time domain. Since video sequences usually include solid temporal correlation, the prediction residuals are generally flat. That is, many residual values are close to "0". The residual signal is used as the input of the subsequent module for transformation, quantization, and entropy coding, realizing the video signal's high-efficiency compression. The residual value of the current LCU can be obtained by estimating the actual reconstruction value of the LCU corresponding to the previous coded position, which can well reflect the relevant information between the video frames. The SATD algorithm is used to count the predicted residual value R_1' of each LCU in a frame, as shown in Eq. (20):

$$R_1' = \sum_{\mathrm{i}=1}^{\mathrm{m}} SATD_{\mathrm{i}}' \tag{20}$$

Where m is the number of sub-pictures in each LCU. $SATD_i'$ is the residual prediction value of one sub-picture in each LCU. The average prediction residual R_2' in the current frame is counted using LCU as the basic unit, as shown in Eq. (21):

$$\mathrm{R}_2' = \frac{1}{\mathrm{k}} \sum_{\mathrm{j}=1}^{\mathrm{k}} SATD_{\mathrm{j}}' \tag{21}$$

Where k is the number of LCUs in a frame and $SATD_j'$ is the predicted residual value of the j-th LCU in the current frame.

According to the residual prediction value R_1' of the current LCU and the average prediction residual value R_2' of the current frame, Eq. (22) calculates the residual prediction factor θ_2 of the LCU block.

$$\theta_2 = \frac{R_1'}{R_2'} \tag{22}$$

3.3 New Complexity Based on Time and Space

Because of the shortcomings of the rate control algorithm based on the R-λ model in HEVC, this section weights the energy proportion factor and the residual prediction factor of the LCU-level to obtain a new type of complexity *NC*. It can effectively distinguish the motion area, and perform a good fitting based on the actual complexity of the

LCU, that is a linear relationship. After adopting this weighted combination method, this layer's target bits can be allocated more reasonably. More target bits can be distributed to areas in HEVC with inaccurate predictions, high complexity, or intense motion, and vice versa. To better detect the texture area and the residual prediction area, combining Eq. (15) and Eq. (19), the calculation equation *NC* is shown in Eq. (23):

$$NC = \mathrm{a} \times \eta_1 + (1 - \mathrm{a}) \times \eta_2 \tag{23}$$

a is a weighting coefficient, and the coefficient value is greater than 0 and less than 1. In this paper, by performing feature quantization on all LCU energy values and predicted residual values in a frame, the coefficients normalized by the linear function are used as the weights of η_1 and η_2 convenient for the indicators of energy ratio and residual factor. Perform comparison and weighting, and improve the accuracy of the algorithm to a certain extent. Its feature quantification equation is as follows:

$$\mathrm{a}_{\eta_1} = \frac{D'_1 - D'_{MIN}}{D'_{MAX} - D'_{MIN}} \tag{24}$$

$$\mathrm{a}_{\eta_2} = \frac{R'_1 - R'_{MIN}}{R'_{MAX} - R'_{MIN}} \tag{25}$$

Among them, D'_{MIN}, D'_{MAX}, R'_{MIN}, and R'_{MAX} are the minimum and maximum values of all LCU energy values and the minimum and maximum values of residual prediction values in a frame, respectively. The feature quantization weighting coefficient a is calculated by Eqs. (24) and (25), as shown in Eq. (26):

$$\mathrm{a} = \frac{\mathrm{a}_{\eta_1}}{\mathrm{a}_{\eta_1} + \mathrm{a}_{\eta_2}} \tag{26}$$

3.4 LCU-Level Bit Allocation Equation

For the B frame or P frame, the target bit allocation algorithm that takes the time and space domain of the LCU-level into consideration is adopted, that is, after the new weight NC is obtained by (23), and use it as the allocation weight to allocate target bits to the LCU-level.

According to the weight of the current LCU, through Eq. (27), the initial target bit allocation for the LCU is performed using the target remaining bit number of the current frame.

$$T_{LCU} = (1 - \mathrm{b}) \times \frac{T_{\mathrm{Curr}Pic} - Bit_{\mathrm{header}} - Coded_{Pic}}{\sum\limits_{NotCodedLCUs} \omega_{\mathrm{i}}} \cdot \omega_{\mathrm{Curr}LCU} + \mathrm{b} \times \frac{T_{\mathrm{Curr}Pic} - Bit_{\mathrm{header}}}{\sum\limits_{AllLCUs} NC_{\mathrm{i}}} \cdot NC_{\mathrm{Curr}LCU} \tag{27}$$

Among them, $T_{CurrPic}$ represents the current frame's target bit; Bit_{header} is the estimated number of bits of the header information; $\omega_{CurrLCU}$ represents the adaptive bit allocation weight of the original platform LCU; $NC_{CurrLCU}$ represents the bit allocation weight of each LCU of the proposed algorithm. *b* is a weighting coefficient, and its value is greater than 0 and less than 1. In this paper, several test sequences were tested for coding performance to determine the value of b. The effect of different values of b on RD performance and code rate control accuracy was calculated. Finally, the value of b was set to 0.4.

According to the buffer status, use Eq. (28) to dynamically adjust the target bit T_{LCU} initially allocated by the current LCU to obtain the final target bit number $T_{CurrLCU}$.

$$T_{\mathrm{CurrLCU}} = T_{LCU} - (\mathrm{totalWeight} - B_{left})/\mathrm{realInfluenceLUC} + 0.5 \tag{28}$$

Equation B_{left} represents the actual remaining bits in the current frame, *totalWeight* represents the sum of bits required by the remaining LCU including the current LCU, *realInfluenceLCU* represents the actual smoothing window size.

3.5 Model Parameter Update

The video sequence's RD characteristic is significant for bit allocation and rate control in the video encoding process. With the changes in the encoded frame's complexity, the motion area, and the reference frame structure, the R-λ of each encoded frame, The λ model parameters (α and β) are not the same. Its rate distortion relationship is not available until the next frame is encoded. The frame's R-λ model parameters to be encoded in [5, 6] are estimated using the closest encoded frame information that belongs to the same level and is updated. The method only uses the actual coded bits and Lagrange multipliers of the coded frame. The convergence speed and accuracy of its parameter update are not ideal, which affects the rate control's RD performance. Using the frame-level and LCU-level bit allocation algorithm proposed in this article can make the target bit allocation more accurate and reasonable. As the encoding quality of the previous encoded CTU frame becomes higher, more valid reference information can be obtained for the encoded frame, including the actual coded bit R_{real}, the Lagrange multiplier λ_{real}, and the coding distortion D_{real}. The updated R-D model parameters can be directly derived from the R-λ function and R-D function of Eq. (1) and Eq. (29):

$$D(R) = CR^{-K} \tag{29}$$

$$C_{\mathrm{new}} = \frac{D_{\mathrm{real}}}{R_{\mathrm{real}}^{-\lambda_{\mathrm{real}} \times R_{\mathrm{real}}/D_{\mathrm{real}}}} \tag{30}$$

$$K_{\mathrm{new}} = \frac{\lambda_{\mathrm{real}} \times R_{\mathrm{real}}}{D_{\mathrm{real}}} \tag{31}$$

Then, the R-λ model parameters are expressed as:

$$\alpha_{new} = C_{new} \cdot K_{new} \tag{32}$$

$$\beta_{new} = -K_{new} - 1 \tag{33}$$

The α_{new} and β_{new} calculated by Eqs. (32) and (33) will be more accurate because the next GOP after encoding can obtain the R-lambda model parameters so that the parameter update can make full use of the information of the relevant coding unit, which will get better RD performance and higher bit rate accuracy.

4 Experimental Result

The HM-16.7 default algorithm without rate control is used as the benchmark scheme to prove the proposed algorithm's effectiveness. The algorithms in [5] and [6] have been applied to the software HM as a comparison solution. All algorithms are implemented on the HEVC reference platform HM-16.7, and the target bit rate of each sequence is under the condition of a fixed QP (QP = 22, 27, 32, 37). After following the general test under the platform of the HM-16.7 reference software, get. Sixteen video sequences with different content and characteristics were tested using the default low-latency B and P configurations. The encoding results include actual bit rate, PSNR, BD-Rate, and RCE calculations thoroughly compare the encoding performance. Table 1 shows the test sequences of Type B, C, D, and E attributes, including resolution, frame rate, and the number of coded frames.

Table 1. Basic information of the test sequence

Sequences	Number of sequences	Resolution	Frame count	Frame rate
Class B	5	1920 × 1080	240&500&600	24&50&60
Class C	4	832 × 480	300&500&600	30&50&60
Class D	4	416 × 240	300&500&600	30&50&60
Class E	3	1280 × 720	600	60

4.1 R-D Performance

RD performance is an important indicator to measure the video coding system. BDBR represents the bit rate saving percentage of the comparison scheme relative to the original method under the same video coding quality. A positive value indicates that the compression performance becomes worse, and a negative value indicates that the compression performance becomes better. HM16.7 without code rate control is the baseline solution, and HM16.7 with code rate control algorithm is the test solution. This article tested RC [5], RC [6], and the proposed algorithm under LDP and LDB configuration. BD-Rate, the test results are shown in Table 2.

Under normal circumstances, compared with the encoder without rate control, RD performance of the encoder with rate control will be reduced due to the encoding bit restriction on GOP, frame, and LCU. The average BDBR of RC [5] and RC [6] under LDP and LDB configurations are 1.7% and 2.7%, 0.1%, and 0.7%, respectively. Although the coding performance of some test sequences has been improved, the overall performance of these two algorithms the rate-distortion performance is not as good as the benchmark scheme; the average BDBR of the algorithm in the LDP and LDB configuration is −2.5% and −2.0%, respectively, and the RD performance of all test sequences is better than that of [5] and [6]. Table 2 allows to derive the average BDBR values for each type of test sequence. This paper's algorithm can improve the C-type sequence's performance the most. The texture of this type of sequence is relatively complex, and the movement is fierce. Because the algorithm in this paper can detect the texture complexity and motion area of the image area well, the energy proportion factor and residual prediction factor proposed can better reflect the sequence's texture characteristics and motion characteristics. More bits are allocated to areas with more intense motion, which improves the coding quality of the video, and provides a better reference frame for the next GOP, thereby improving the RD performance of the entire coding sequence.

Table 2. Performance comparison of BD-Rate (%) in LDP and LDB configuration

Class	Sequence	RC [7]		RC [8]		Proposed	
		LDP	LDB	LDP	LDB	LDP	LDB
B	Cactus	−2.7	−1.1	−2.2	−1.2	−5.5	−5.0
	BasketballDrive	6.8	8.4	3.2	3.5	1.1	1.3
	BQTerrace	3.9	7.1	4.5	7.1	−1.0	1.5
	Kimono	11.3	11.5	6.6	6.5	4.5	4.4
	ParkScene	2.4	2.7	1.7	2.1	−2.5	-2.2
	Average	4.3	5.7	2.8	3.6	−0.7	0.0
C	BasketballDrill	−5.9	−5.9	−5.9	−5.5	−8.4	−8.4
	BQMall	−1.6	−1.4	−1.8	−1.7	−3.9	−4.1
	PartyScene	−1.3	−1.2	−2.3	−2.4	−5.0	−5.5
	RaceHorsesC	1.5	2.7	−1.4	−0.2	−2.8	−1.8
	Average	−1.8	−1.5	−2.8	−2.5	−5.0	−5.0
D	BasketballPass	−2.7	−2.5	−4.8	−4.9	−6.9	−7.1
	BlowingBubbles	0.9	0.6	−1.7	−2.1	−4.3	−4.8
	BQSquare	5.4	7.9	1.1	2.3	0.6	1.4
	RaceHorses	1.2	1.9	0.6	0.8	−0.4	−0.2
	Average	1.2	2.0	−1.2	−1.0	−2.8	−2.7
E	FourPeople	2.3	2.8	1.2	1.7	−3.2	−2.5
	Johnny	5.9	8.1	3.0	4.9	1.5	3.7
	KristenAndSara	0.4	0.9	−0.8	0.0	−4.0	−3.4
	Average	2.9	4.0	1.1	2.2	−1.9	−0.7
Total average		1.7	2.7	0.1	0.7	−2.5	−2.0

Figure 1 and Fig. 2 are the comparison diagrams of Bits cost and Y-PSNR of each frame obtained by encoding the class B sequence BQTerrace with the target code rate of 841.254 kbps the algorithm of this paper and RC [6] under the LDP configuration. The actual output bit rate and Y-PSNR are 841.278 kbps and 31.3942 dB. The actual output bit rate and Y-PSNR of RC [6] are 841.343 kbps and 31.1997 dB. Figure 3 and Fig. 4 are the RD curves of the Basketballpass sequence and the Cactus sequence. It can be seen from Figs. 1 and 2 that for the non-homogeneous BQTerrace sequence, the algorithm in this paper has better rate control performance. Most of the coded frames can get a higher Y with a relatively small number of bits. -PSNR, which improves the average PSNR of the sequence and increases the smoothness. Among them, there is a very obvious shot switch from frame 150 to frame 250. This paper's algorithm can detect the image texture complexity and motion areas better than RC [6] to obtain a very high coding quality. Figures 3 and 4 reflect the non-homogeneous and violent motion sequence with complex background and object rotation. This paper's algorithm is used to adjust and accurately allocate the number of bits in the frame and the LCU-level. It has better RD performance under the condition of high bit rate and high bit rate.

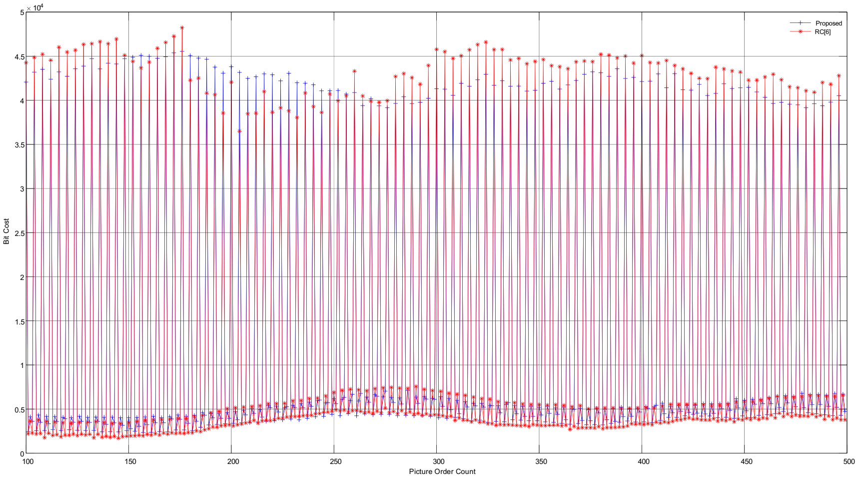

Fig. 1. Bits cost comparison chart of BQTerrace sequence

4.2 Rate Control Performance

The purpose of code rate control is to make the encoder's actual output code rate equal to the target code rate as much as possible while minimizing the coding distortion of the sequence. Therefore, the bit rate accuracy is another crucial performance index measured by the bit rate's relative error (R_{Err}). The inaccuracy of the code rate is defined as:

$$\mathrm{R}_{Err} = \frac{\left|R_{\text{actual}} - R_{\text{target}}\right|}{R_{\text{target}}} \times 100\% \tag{34}$$

R_{actual} and R_{target} represent the actual code rate and target code rate. The trimmer R_{Err} is, the higher the rate control accuracy is. In the experiment, each test sequence is set with four target bit rates from low to high. Due to space limitations, the average value of the four QPs is used as the result. As shown in Table 3, it can be observed that this

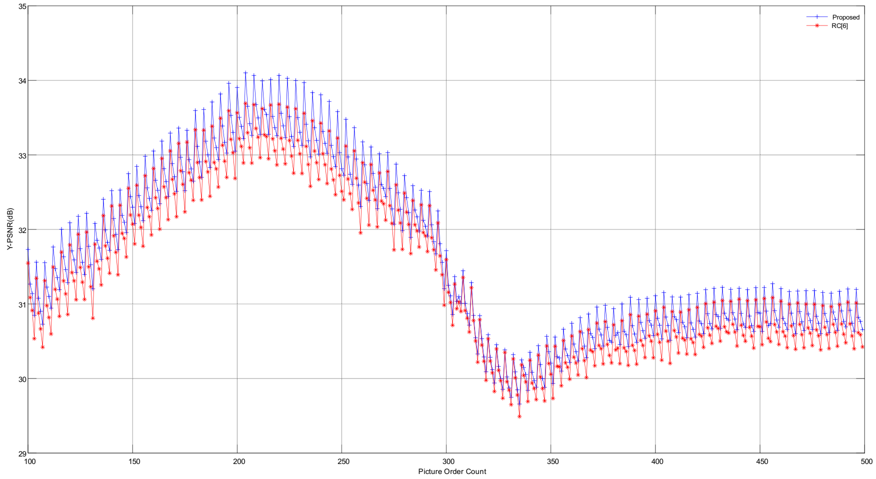

Fig. 2. Y-PSNR comparison chart of BQTerrace sequence

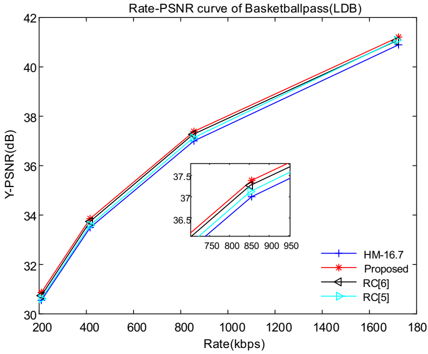

Fig. 3. RD curve of Basketballpass sequence

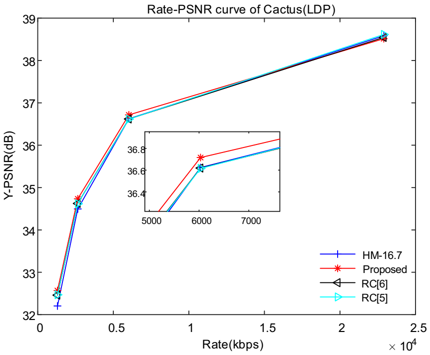

Fig. 4. RD curve of Cactus sequence

paper can achieve more accurate rate control accuracy than RC [5] and RC [6]. Under the configuration of LDP and LDB, the algorithm's average in this paper is 0.007% and 0.006%, respectively. Among them, the relative error of the average code rate of Class B, Class C, and Class D is better than that of the comparison method, and the accuracy reaches 0.004%, 0.003%, and 0.010%. The rate control accuracy of Class D is 0.063% higher than that of RC [6]. It is worth noting that the algorithm in this paper achieves a considerable improvement in RD performance and improves the rate control accuracy.

Table 3. Comparison of R_{Err} (%) Bit Rate Accuracy under LDP and LDB Configurations

Class	Sequence	RC [5]		RC [6]		Proposed	
		LDP	LDB	LDP	LDB	LDP	LDB
B	Cactus	0.002	0.005	0.002	0.006	0.001	0.001
	BasketballDrive	0.003	0.001	0.003	0.001	0.002	0.002
	BQTerrace	0.002	0.002	0.003	0.003	0.003	0.001
	Kimono	0.011	0.009	0.013	0.003	0.011	0.011
	ParkScene	0.022	0.028	0.006	0.009	0.003	0.001
	Average	0.008	0.009	0.005	0.005	0.004	0.003
C	BasketballDrill	0.021	0.029	0.012	0.011	0.006	0.005
	BQMall	0.002	0.002	0.002	0.002	0.003	0.004
	PartyScene	0.004	0.003	0.004	0.003	0.001	0.003
	RaceHorsesC	0.003	0.035	0.002	0.015	0.002	0.001
	Average	0.007	0.017	0.005	0.008	0.003	0.003
D	BasketballPass	0.224	0.218	0.260	0.261	0.011	0.005
	BlowingBubbles	0.009	0.008	0.013	0.011	0.007	0.006
	BQSquare	0.007	0.005	0.012	0.015	0.015	0.019
	RaceHorses	0.005	0.009	0.008	0.006	0.008	0.006
	Average	0.061	0.060	0.073	0.073	0.010	0.009
E	FourPeople	0.007	0.003	0.006	0.006	0.005	0.007
	Johnny	0.004	0.001	0.001	0.002	0.005	0.002
	KristenAndSara	0.027	0.031	0.025	0.019	0.024	0.023
	Average	0.013	0.012	0.011	0.009	0.011	0.011
Total average		0.022	0.024	0.023	0.023	0.007	0.006

5 Conclusion

This article aims to improve the visual quality and coding performance, so a novel frame-level precision bit allocation optimization algorithm and LCU-level rate control improvement algorithm for image complexity are proposed. The frame level bit allocation is performed based on the average amount of information in the image by the new type of weight bit allocation algorithm, and then the Hadamard transform algorithm is performed to detect the texture complexity area and the residual prediction area, besides, take the image's complexity and the unique coding features of HEVC into account to allocate the bits of the LCU-level reasonably. This paper's proposed algorithm has excellent benefits compared with other advanced bit allocation algorithms and can achieve more stable rate control accuracy, more minor bit fluctuations, and better RD performance.

References

1. Choi, H., Nam, J., Yoo, J., et al.: Rate control based on unified RQ model for HEVC. ITU-T SG16 Contribution, JCTVC-H0213, 1–13 (2012)
2. Li, B., Li, H., Li, L., et al.: Rate control by R-lambda model for HEVC. ITU-T SG16 Contribution, JCTVC-K0103, 1–5 (2012)
3. Karczewicz, M., Wang, X.: Intra frame rate control based on SATD. Document JCTVC-M0257, Incheon, Korea, April 2013
4. Li, B., Li, H., Li, L.: Adaptive bit allocation for R-lambda model rate control in HM. Document JCTVC-M0036, Incheon, Korea, April 2013
5. Li, B., Li, H., Li, L., Zhang, J.: λ domain rate control algorithm for high efficiency video coding. IEEE Trans. Image Process. **23**(9), 3841–3854 (2014)
6. Li, L., Li, B., Li, H., Chen, C.W.: λ-domain optimal bit allocation algorithm for high efficiency video coding. IEEE Trans. Circuits Syst. Video Technol. **28**(1), 130–142 (2018)
7. Liu, M., Ren, P., Xiang, Z.: Frame-level bit allocation for hierarchical coding of H.265/HEVC considering dependent rate-distortion characteristics. Sig. Image Video Process. **10**(8), 1457–1463 (2016)
8. Fiengo, A., Chierchia, G., Cagnazzo, M., Pesquet-Popescu, B.: Rate allocation in predictive video coding using a convex optimization framework. IEEE Trans. Image Process. **26**(1), 479–489 (2017)
9. Lu, X., et al.: Improved Bit Allocation Using Distortion for the CTU-Level Rate Control in HEVC (2018)
10. Zhao, Z., Xiong, S., Sun, W., He, X., Zhang, F.: An improved R-λ rate control model based on joint spatial-temporal domain information and HVS characteristics. Multimed. Tools Appl. **80**(1), 345–366 (2020). https://doi.org/10.1007/s11042-020-09721-9
11. Wang, M., Ngan, K.N., Li, H.: An efficient frame-content based intra frame rate control for high efficiency video coding. IEEE Sig. Process. Lett. **22**(7), 896–900 (2015)