Journal of Research in Science and Engineering (JRSE)
ISSN: 1656-1996 Volume-4, Issue-2, February 2022
DOI: 10.53469/jrse.2022.04(02).06

Rate Control Optimization Algorithm Based on R-λ Model in H.266/VVC

Yue Sha[1], Jing Yang[2]

[1,2]School of Information Engineering, Shanghai Maritime University, Shanghai, China
[1]1490003862@qq.com, [2]jingyang@shmtu.edu.cn

Abstract: ***In this paper, we propose a new frame-level rate control scheme based on quality dependent factor by studying the effect of skipping blocks on rate-distortion parameters and quality dependence. Experimental results show that compared with the previously commonly used R-λ rate control scheme, the proposed method has an average BD-rate value of -1.98% and -1.75% under the configuration of low delay and random access respectively. And the bit error rate has been reduced by 0.03% and 0.09% respectively. The improved rate control algorithm improves the rate-distortion performance while reducing the bit error rate.***

Keywords: Quality dependent factor, Skipping blocks, Rate-distortion parameters, Rate control**.**

1. Introduction

Video coding originated in 1950, thus opening a new era of digital video coding. From the initial 360p, it has developed to SD480p, HD720p, Blu-ray, etc. Nowadays, the video resolution has reached 4k or even 8k. This is becoming more and more challenging for video compression. After the High Efficiency Video Coding (HEVC)/H.265 standard, VCEG and ISO/IEC jointly initiated the establishment of the Joint Video Expert Group (JVET) and began to explore the next generation of video coding technology, namely H.266/VVC. As the latest generation of video coding technology, H.266/VVC started in 2016, and the first version was finalized in July 2020. On the whole, H.266 still follows the block-based hybrid coding framework in H.265. But for different coding parts, the H.266 standard has more or less added some new technologies. These new technologies increase the coding efficiency of H.266 by approximately 50% compared to H.265/HEVC.

In video communication, video compression efficiency and coding video quality, as well as the adaptability of transmission bandwidth and coding rate are always the most concerned issues of researchers. In order to solve this problem, the current mainstream video coding standard H.265/HEVC and the latest video coding standard H.266/VVC have adopted key rate control technologies. Rate control technology has always been a research hotspot. Researchers at home and abroad have successively proposed a variety of rate control algorithms. In the past video coding standards, the representative rate control algorithms mainly include algorithms based on the R-Q model, ρ domain model algorithms and algorithms based on the R-λ model. Among them, the R-λ model algorithm is widely used because of its better rate control effect and lower bit fluctuation. The current R-λ standard model rate algorithm first obtains the Lagrangian factor λ through the bit rate distortion model, According to the linear relationship between λ and quantization parameter (Quantization Parameter, QP), the QP required for encoding is obtained, so that the actual output code rate is closer to the set target code rate. But there are still shortcomings: In VVC, with the further improvement of coding efficiency, especially between frames, the reference frame can better predict the block, and most of the blocks can be skipped without residual information. Since skipped blocks and non-skip blocks have different characteristics, this is not considered in the previous rate control algorithm based on the R-λ model. The skipped blocks and non-skip blocks are mixed for parameter calculation, this will lead to inaccurate updates of model parameters. In response to this problem, this paper proposes a frame-level bit allocation scheme based on quality-dependent factors by analyzing the statistical information of the skipped block ratio. The proposed scheme can make the actual output bit rate closer to the allocated target bit rate, thereby improving the coding performance to a certain extent.

2. R-λ Rate Control Model

For the rate control algorithm, the main function is to assign the target bit rate to each image group as accurately as possible under the given target bit rate, then redistribute it to each frame and finally to each coding unit. After the rate control model has gone through the R-Q model and the R-ρ model, the R-λ model is now usually used.

The idea of rate control for each group of pictures (Group of Pictures, GOP) layer is mainly to use a series of parameters to obtain the target number of bits in the picture layer, the LCU layer target bit allocation depends on the weight of the complexity of the current largest coding unit in the complexity of the largest uncoded coding unit in the frame. Since this article is a proposed algorithm for bit allocation at the frame layer, the previous scheme is used for bit allocation at the GOP layer and LCU layer. The number of target bits in the frame layer is determined by the weight of the current frame in the GOP, etc. For non-I frames, the ω of each frame is different, so a different target number of bits is allocated to each frame. The target bit allocation model of the frame layer is as follows:

$$T_{pic} = \frac{T_{Gop} - R_{G,c}}{\sum_{i=1}^{N_G} \omega_i} \times \omega_i \tag{1}$$

In the formula: T_{Gop} represents the target bit of the GOP layer, the current GOP coded bit is represented by $R_{G,c}$, ω_i is the bit allocation weight, and the total number of frames contained in the image group is N_G. In particular, for a complete full I frame, the size of the GOP is 1, the number of target bits in the frame layer is the target number of bits in the GOP layer.

The existing rate control scheme uses the R-λ rate-distortion model to select coding parameters—Lagrangian factors and quantization parameters. Then use the quantization parameter and the assigned target bit rate to optimize the rate distortion of the unit layer (RDO), select the optimal coding mode for the coding unit. Finally, the image is compressed by using quantization parameters. It is formulated by a constrained optimization problem and can be transformed into an unconstrained form:

$$\min J_{\text{mode1...mode}N} \quad where J = D + \lambda R \tag{2}$$

The R-D relationship is:

$$D(R) = CR^{-K} \tag{3}$$

In the formula: C and K are model parameters related to the characteristics of the video content, which are calculated by least squares algorithm and updated gradually. The result is an R-D curve with λ as the slope, and the formula is as follows:

$$\lambda = -\frac{\partial D}{\partial R} = CKR^{-K-1} = \alpha R^{\beta} \tag{4}$$

In the formula: the two parameters α and β are related to the characteristics of the video content. In addition, through a large number of experimental fittings, there is a linear relationship between the quantitative parameters QP and ln λ:

$$QP = m \ln \lambda + n \tag{5}$$

In the formula: the value of m is 4. 2005, the value of n is 13. 7122.

Since the content structure transformation characteristics of a coded video sequence are closely related to the parameters α and β, and different frames or coding units have different video content characteristics. Therefore, in the actual encoding process, after encoding one frame of image or one coding unit, the model parameters α and β need to be updated before executing the next frame of image or the next coding unit. After comparing the average target bit number bpp obtained by bit allocation with the actual target bit number obtained after encoding, it is found that there is still a certain error between the two. In the R-λ model, the actual target bit number bpp_{act} is used to update the parameters:

$$\lambda_{comp} = \alpha_{old}(bpp_{act})^{\beta_{old}} \tag{6}$$

$$\alpha_{new} = \alpha_{old} + \delta_{\alpha}(\ln \lambda_{old} - \ln \lambda_{comp})\alpha_{old} \tag{7}$$

$$\beta_{new} = \beta_{old} + \delta_{\beta}(\ln \lambda_{old} - \ln \lambda_{comp}) \ln bpp_{act} \tag{8}$$

In the formula, α_{old} and α_{new}, β_{old} and β_{new} are the rate-distortion model parameters. α_{old} and β_{old} are the rate-distortion model parameters of the current frame or the current coding unit, which are used to calculate the λ_{comp} parameter; λ_{comp} is the Lagrangian factor calculated from the actual bit rate in formula (6); The λ_{old} parameter is the Lagrangian factor calculated under the target bit rate; α_{new} and β_{new} are the model parameters of the next frame or coding unit; bpp_{act} is the bit rate value of each pixel under the actual bit rate. δ_{α} and δ_{β} are the iterative steps when updating the rate-distortion model parameters, which are determined by the bit (bpp) of each pixel under the target bit. The relationship between them is shown in the following table:

Table 1: The relationship between δ_{α} and δ_{β} parameters and bpp

Update parameters	bpp >=0.08	0.08> bpp >=0.03	bpp <0.03
δ_{α}	0.1	0.5	0.1
δ_{β}	0.05	0.025	0.005

3. Rate Control Optimization Algorithm based on Quality Dependent Factor

3.1 Quality Dependent Factor

A GOP in RA configuration contains 16 pictures belonging to 5 coding levels, where the pictures in higher level will choose the pictures in the lower as their reference pictures. Thus, there exists quality dependency among different pictures in different coding level, which can be expressed as:

$$\delta_{i,j} = \frac{\partial D_j}{\partial D_i} \tag{9}$$

where i, j mean the coding levels and $j > i$, D means the distortion. The quality dependent factor (QDF) can be used to describe the impact on layer j when the quality of layer i changes.

For the relationship between the quality dependent factor and the ratio of skipped pixels, Gao[11] first demonstrated in the paper and applied it to ρ domain rate control algorithm, but this is not applicable to the R-λ model in VVC. Later, Liu et al. demonstrated the QDF in VVC through experiments and explained it in the literature[12]. According to the experimental results, it is found that the quality dependent factor has a strong relationship with the skip rate of the coded picture. QDF can be obtained through a linear function, and the skip ratio is shown in[10]:

$$\begin{aligned} \delta_{1,2} &= A_{1,2}S_2 + B_{1,2} \\ \delta_{1,3} &= A_{1,3}S_3 + B_{1,3} \\ \delta_{1,4} &= A_{1,4}S_4 + B_{1,4} \\ \delta_{1,5} &= A_{1,5}S_5 + B_{1,5} \end{aligned} \tag{10}$$

Where S_j represents the skip ratio of j th coding levels which can be acquired from last encoded GOP, $A_{1,j}$ and $B_{1,j}$ are the parameters obtained by some analysis of VTM software when rate control is not applied. The parameters $A_{1,j}$ and $B_{1,j}$ can be seen in Table 2.

Table 2: Parameters of relationships between QDF and skip ratio

Level i	$A_{1,j}$	$B_{1,j}$
2	0.5847	-0.0782
3	0.5468	-0.1364
4	0.6539	-0.203
5	0.8623	-0.4676

The QDF of other coding levels can be deduced from the known relationship:

$$\delta_{i,j} = \delta_{1,j} \times (1-\delta_{1,i}) \; (i > 1, j > i.) \quad (11)$$

3.2 QDF Based Frame-level Bit Allocation

Supposing the target number of bits for this GOP is T_{Gop} and N is the GOP size, the picture level bit allocation problem is to solve the constrained optimization problem as:

$$\min \sum_{i=1}^{N} D_i \text{ s.t. } \sum_{i=1}^{N} R_i \le T_{GOP} \quad (12)$$

This problem can be converted to an unconstrained optimization problemby introducing a Lagrange multiplier:

$$J = \min_{\lambda_1,...,\lambda_N} \sum_{i=1}^{N} D_i + \lambda(\sum_{i=1}^{N} R_i - T_{GOP}) \quad (13)$$

This can be solved by making the derivative zero:

$$\frac{\partial \sum_{i=1}^{N} D_i}{\partial \lambda_j} + \lambda \frac{\partial \sum_{i=1}^{N} R_i}{\partial \lambda_j} = 0, \; j = 1,2,...N \quad (14)$$

By introducing the R-D relational expression shown by equation (3), in addition to the analysis quality and bit rate dependence, the problem in equation (14) can finally be transformed into the following equation:

$$\lambda_j = \frac{\lambda}{\frac{\partial \sum_{i=j}^{Np} D_i}{\partial D_j}} = \omega_j \lambda, \; j = 1,2,...N \quad (15)$$

where ω_j is the quality dependency on j th frame. Because high-level frames may choose low-level frames as their reference frames, the quality dependence comes from the reference relationship. ω_j is determined by QDFs from the pictures that will take j th picture as their reference picture:

$$\omega_j = \begin{cases} \frac{1}{1+2\times(\delta_{1,2}+2\times\delta_{1,3}+4\times\delta_{1,4}+8\times\delta_{1,5})}, & j \in level\,1 \\ \frac{1}{1+(3\times\delta_{2,3}+5\times\delta_{2,4}+8\times\delta_{2,5})}, & j \in level\,2 \\ \frac{1}{1+(2\times\delta_{3,4}+4\times\delta_{3,5})}, & j \in level\,3 \\ \frac{1}{1+(2\times\delta_{4,5})}, & j \in level\,4 \\ 1.0, & j \in level\,5 \end{cases} \quad (16)$$

The quality dependent factor $\delta_{i,j}$ is calculated by formulas (10) and (11). Combining (15) and (16), and using the dichotomy in the literature[14], the solution of λ can be calculated. Finally, the final bit allocation weight of each frame can be calculated by the following formula, where C_i and K_i are model parameters related to the characteristics of the video content.

$$\Omega_i = (\frac{\lambda_i}{C_i K_i})^{-\frac{1}{K_i+1}} = (\frac{\omega_i \lambda}{C_i K_i})^{-\frac{1}{K_i+1}} \quad (17)$$

4. Experiment and Result Analysis

The simulation experiment is based on the H.266 default test software VTM 10.0. The computer used in the experiment is Intel Core i5-6500 cpu @ 3.2GHz, the memory is 16G, the operating system is Windows 10 X64, and the running environment is Microsoft Visual Studio 2017. The test sequence is H.266 default test sequence A1, A2, B, C, D, E, a total of 6 sets of sequences. The test configuration environment is LD (low delay) and RA (random access), and QP (quantization parameter) is 22, 27, 32, 37 in turn. The BD-rate in the experimental data represents the average change of the bit rate when the restored quality of the decoded image is consistent. If the value is negative, it means that the bit rate is saved and the coding performance is improved; otherwise, it means that the bit rate is lost. The bit rate error is defined as ΔR, which reflects the degree of deviation between the target bit rate and the actual output bit rate:

$$\Delta R = \frac{|R_{act} - R_{tar}|}{R_{tar}} \times 100\% \quad (18)$$

Where: R_{tar} is the target bit rate, and R_{act} is the actual bit rate. The smaller the ΔR, the better the accuracy of rate control and the better the video coding performance. The test results are shown in Table 3 and Table 4 below:

Table 3: Comparision of BD-rate(Y) and Bit-error in LD configuration

	Sequence	Proposed vs Default (%)	ΔR (%)	
		BD-rate (Y)	Proposed	Default
A1	Tango2	-5.66	0.02	0.03
	FoodMarket4	1.23	0.01	0.02
	Campfire	-3.63	0.03	0.01
A2	CatRobot1	-1.77	0.04	0.05
	DaylightRoad2	0.82	0.02	0.01
	ParkRunning3	-1.33	0.01	0.03
B	MarketPlace	-8.26	0.08	1.25
	RitualDance	-1.69	0.09	0.04
	Cactus	-1.23	0.05	0.03
	BasketballDrive	-1.12	0.07	0.02
	BQTerrace	-3.24	0.13	0.08
C	RaceHorseC	-0.82	0.12	0.09
	BQMall	-0.59	0.25	0.18
	PartyScene	-1.43	0.11	0.08
	BasketballDrill	-0.47	0.19	0.15
D	RaceHorses	-0.69	0.29	0.26
	BQSquare	0.87	0.65	0.68
	BlowingBubbles	-0.92	0.49	0.43
	BasketballPass	-1.52	0.36	0.35
E	FourPeople	-2.88	0.41	0.39
	Johnny	-4.36	0.82	0.86
	KristenAndSara	-4.77	0.53	0.55
Average		- 1.98	0.22	0.25

It can be seen from the experimental results that for the luminance component Y, the BD-rate value of almost all test sequences is negative, which means that it saves the bit rate. In the low delay and random access configuration, the average bit rate is saved by 1.98% and 1.75% respectively.

It is worth noting that especially for some video sequences with the characteristics of homogeneity, flatness and no strenuous motion, the improvement of rate-distortion performance is more obvious. The reason is that test sequences with more skip rates and stable content texture can benefit more from the proposed method. In addition, the bit rate control algorithm proposed in this paper reduces the bit error rate by 0.03% and 0.09% on average compared with the default algorithm under the configuration of low delay and random access. This algorithm makes the actual output bit rate closer to the target bit rate, with better consistency.

Table 4: Comparision of BD-rate(Y) and Bit-error in RA configuration

	Sequence	Proposed vs Default (%)	Δ*R* (%)	
		BD-rate (Y)	Proposed	Default
A1	Tango2	-4.36	0.92	0.83
	FoodMarket4	2.34	0.37	0.63
	Campfire	-1.69	0.05	0.06
A2	CatRobot1	-1.25	0.40	0.14
	DaylightRoad2	-2.62	0.35	0.41
	ParkRunning3	-0.93	0.17	0.20
B	MarketPlace	-2.65	0.76	1.46
	RitualDance	-1.54	0.21	0.37
	Cactus	-2.04	0.25	0.23
	BasketballDrive	-1.53	0.16	0.19
	BQTerrace	-2.66	0.34	0.55
C	RaceHorseC	-1.03	0.23	0.18
	BQMall	-1.28	0.74	1.02
	PartyScene	-0.76	0.19	0.32
	BasketballDrill	-1.15	0.68	0.71
D	RaceHorses	-0.76	0.76	0.69
	BQSquare	-2.56	1.24	1.36
	BlowingBubbles	-1.87	1.57	1.84
	BasketballPass	-1.08	0.62	0.88
E	FourPeople	-1.93	0.83	0.73
	Johnny	-2.59	0.77	0.82
	KristenAndSara	-4.64	0.94	0.79
Average		- 1.75	0.57	0.66

In order to further illustrate the effectiveness of the algorithm in this paper, the algorithm in this paper is compared with the rate control algorithm used in literature[19] and literature[20]. Since the test sequences used in different papers are different, here we only compare the results of the same test sequence. In Table 5 below, the BD-Rate values of this algorithm for the reference sequence are all negative numbers, and the average values are -1.00% and -1.52% under low delay and random access configurations respectively, which are lower than the reference algorithm. This shows to some extent that the rate-distortion performance of the rate control algorithm in this paper is better than that of the reference algorithm.

Table 5: Comparison of BD-rate values of three rate control algorithms

	Sequence	BD-rate/%			
		[19]	[20]	Proposed (LD)	Proposed (RA)
B	Cactus	0.4	-0,46	-1.23	-2.04
	BQTerrace	2.9	-0.54	-3.24	-2.66
C	RaceHorseC	-0.3	-0.56	-0.82	-1.03
	BQMall	-1.5	-0.70	-0.59	-1.28
	PartyScene	-3.0	-0.72	-1.43	-0.76
	BasketballDrill	-2.3	-0.58	-0.47	-1.15
D	RaceHorses	1.8	-0.66	-0.69	-0.76
	BQSquare	3.7	-0.74	-0.87	-2.56
	BlowingBubbles	-3.8	-0.72	-0.92	-1.87
	BasketballPass	-4.6	-0.76	-1.52	-1.08
Average		-0.67	-0.64	-1.00	-1.52

5. Conclusion

Based on the analysis of skip block information, this paper proposes a frame-level bit allocation scheme based on quality dependent factors, which effectively improves the video coding quality. Experimental results show that the proposed algorithm reduces the bit error rate while improving the rate-distortion performance. For video images with stable content texture, the optimization effect is particularly obvious.

References

[1] B. Bross, J. Chen, S. Liu, Versatile Video Coding (Draft 5), JVET N1001 (2019).

[2] 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.

[3] 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.

[4] L. Xu, S. Li, K. N. Ngan, and L. Ma, “Consistent visual quality control in video coding,” IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 6, pp. 975-989, Jun. 2013.

[5] M. Wang, K. N. Ngan, and H. Li, “Low-delay rate control for consistent quality using distortion-based Lagrange multiplier,” IEEE Trans. Image Process., vol. 25, no. 7, pp. 2943-2955, Jul. 2016.

[6] 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.

[7] 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.

[8] B. Li, H. Li, L. Li, and J. Zhang, “Rate control by R-lambda model for HEVC, JCTVC-K0103,” in Proc. 11th JCTVC Meeting, Shanghai, China, 2012, pp. 1-11.

[9] 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.

[10] Y. Li, Z. Chen, X. Li, and S. Liu, “Rate Control for VVC, JVET-K0390,” in Proc. 11th JVET Meeting, Ljubljana, Slovenia, 2018, pp. 1-5.

[11] 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 Techn., vol. 26, no. 1, pp. 139-153, 2015.

[12] Z. Liu, Z. Chen, Y. Li, Y. Wu, and S. Liu, “AHG10: Quality dependency factor based rate control for VVC,” in Joint Video Exploration Team (JVET) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-M0600, Marrakech, 2019.

[13] 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 Techn., vol. 21, no. 8, pp. 1152-1162, 2011.

[14] L. Li, B. Li, H. Li, and C. Chen, “λ-domain optimal bit allocation algorithm for high efficiency video coding,” IEEE Trans. Circuits Syst. Video Techn., vol. 28, no. 1, pp. 130-142, 2018.

[15] H. Yang, L. Shen, X. Dong, Q. Ding, P. An, and G. Jiang, "Low-complexity ctu partition structure decision and fast intra mode decision for versatile video coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 6, pp. 1668-1682, 2020.
[16] V. Sanchez, "Rate control for predictive transform screen content video coding based on ransac," IEEE Transactions on Circuits and Systems for Video Technology, pp. 1-1, 2020.
[17] B. Bross, J. Chen, and S. Liu, Versatile Video Coding (Draft 5), document JVET-N1001-v10, 14th meeting, Geneva, China, 2019.
[18] J. Chen, Y. Ye, and S. H. Kim, Algorithm Description for Versatile Video Coding and Test Model 5, document JVET-N1002, 14th meeting, Geneva, China, 2019.
[19] F. Liu, Z.Z. Chen, "Multi-Objective Optimization of Quality in VVC Rate Control for Low-Delay Video Coding," IEEE Transactions on Image Processing, vol. 22, no.4, pp. 4706-4718,April.2021.
[20] Y. Chen, S. Kwong, M. Zhou, S. Wang, G. Zhu and Y. Wang, "Intra Frame Rate Control for Versatile Video Coding with Quadratic Rate-Distortion Modelling," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4422-4426, 2020.