

## Improved Rate Control Scheme for Consistent Video Quality

Yuan Li<sup>+</sup>, Huizhu Jia, Xiaodong Xie and Tiejun Huang

National Engineering Laboratory for Video Technology, Peking University, Beijing, China

**Abstract.** Consistent video quality is an important optimization purpose for rate control except the bit rate accuracy. In this paper, we propose a rate control scheme for the consistent video quality. Firstly, the frame bit allocation is improved by the MAD-based method, which can follow the characteristic of the video content. Then, to minimize the fluctuation of the frame distortion, D-Q model is used to regulate the QP after QP decision by the R-Q model. Experimental results show that the proposed rate control scheme can achieve better R-D performance than benchmark and the fluctuation of the frame distortion is decreased up to 40.68%.

**Keywords:** HEVC, Rate control, Video quality, Bit allocation, D-Q model

### 1. Introduction

Rate control is a hot topic with the improvement of video codecs though it is not specified by the video coding standards. In various video coding applications, rate control is employed to regulate the bit stream to adapt the limited transmission channel. Except achieving the target bit rate accurately, some other optimizations also need to be considered in the rate control process, such as minimizing the total distortion, reducing the system latency, maintain the consistent video quality etc.

For the purpose of video quality optimization, various technologies have been proposed in the literature. In [1] and [2], the relationship between QP and rate-distortion optimization (RDO) is studied to improve the R-D performance. Li et al. [1] proposed a linear model to predict the coding complexity of the current basic unit, which was used to obtain the QP before the RDO process. While in [2], Ma et al. proposed a partial two-pass scheme for RDO process at macroblock (MB) level to make the mode selection more accurate. Besides, R-D optimized bit allocation (BA) algorithms are also implemented widely [3-4]. Both [3] and [4] explored the BA of hierarchical structure in HEVC, where the basic concept was to ensure the high video quality of the key frames. In Xie's work [5], a frame-level BA framework with a rate-complexity model was proposed to track the nonstationary characteristics, which achieved the smooth video quality.

In this paper, we propose a rate control scheme for the consistent video quality. Firstly, the frame bit allocation is improved by the MAD-based method, which can follow the characteristic of the video content. Then, to minimize the fluctuation of the frame distortion, D-Q model is used to regulate the QP after QP decision by the R-Q model. Experimental results show that the proposed rate control scheme can achieve better R-D performance than benchmark and the fluctuation of the frame distortion is decreased up to 40.68%.

The rest of this paper is organized as follows. Section 2 describes the rate control scheme in HEVC. In Section 3, the rate control scheme is proposed to reduce the fluctuation of the video quality. Section 4 shows the experimental results and the conclusion is given in Section 5.

---

<sup>+</sup> Corresponding author. Tel.: +86-13810334558  
E-mail address: yuanli@jdl.ac.cn

## 2. Rate Control in HEVC

Rate control algorithm in video coding mainly contains two important parts, the first one is the target bit allocation for different level of coding units and the second one is the QP decision according to the R-Q model. In HEVC the target bit allocation has three levels which are GOP level, frame level and LCU level. The R-Q model for QP decision is R- $\lambda$  model [6].

### 2.1. Target bit allocation

For the GOP level bit allocation, the average bit rate for each frame ( $R_{PicAvg}$ ) is obtained first as

$$R_{PicAvg} = R_{tar} / f \quad (1)$$

where  $f$  is the frame rate,  $R_{tar}$  is the target bit rate. Then the target bit for a GOP ( $T_{GOP}$ ) is

$$T_{AvgPic} = \frac{R_{PicAvg} \cdot (N_{coded} + SW) - R_{coded}}{SW} \quad (2)$$

$$T_{GOP} = T_{AvgPic} \cdot N_{GOP} \quad (3)$$

where  $N_{GOP}$  is the frame number in one GOP,  $N_{coded}$  is the number of coded frames,  $R_{coded}$  is the bit cost for the these frames,  $SW$  is a smooth window which is set as 40.

Frame bit allocation is according to the weight for each frame in current GOP, which is

$$T_{CurrPic} = \frac{T_{GOP} - Coded_{GOP}}{\sum_{NotCodedPictures} \omega_i} \cdot \omega_{CurrPic} \quad (4)$$

where  $Coded_{GOP}$  is the coded bits of the current GOP.  $\omega$  is the bit allocation weight of every frame in the current GOP.

Then LCU level bit allocation is determined by

$$T_{CurrLCU} = \frac{T_{CurrPic} - Bit_{header} - Coded_{Pic}}{\sum_{NotCodedLCUs} \omega_i} \cdot \omega_{CurrLCU} \quad (5)$$

where The  $Bit_{header}$  is the estimated bits of all headers, including slice header.  $Bit_{header}$  is estimated according to the actual header bits of previous coded frames belonging to the same level,  $\omega$  is the weight of each LCU, estimated by the prediction error (in form of MAD) of collocated frame in the previous coded frame belonging to the same level.

$$MAD_{LCU} = \frac{1}{N_{pixels}} \sum_i |pred_i - org_i| \quad (6)$$

$$\omega_{LCU} = MAD_{LCU}^2 \quad (7)$$

### 2.2. QP decision

After the bit allocation for different level of coding units, QP of each level need to be decided according to the R-Q model. In HEVC, the Hyperbolic R-D model can better represent the relationship between the bit rate and the distortion after the R-D performance analysis,

$$D(R) = CR^{-K} \quad (8)$$

where  $C$  and  $K$  are model parameters according to the video content. As is well known,  $\lambda$  is the slope of R-D curve. Then  $\lambda$  can be derived from (8) which is

$$\lambda = -\frac{\partial D}{\partial R} = CKR^{-K-1} = \alpha R^\beta \quad (9)$$

where  $\alpha$  and  $\beta$  are model parameters. Then  $bpp$  (bit per pixel, which is  $T_{cur}/N_{pixel}$ ) is used to represent the different level of target bit rate and (9) can be rewritten as

$$\lambda = \alpha \cdot bpp^\beta \quad (10)$$

Then QP can be calculated by

$$QP = 4.2005 \ln \lambda + 13.7122 \quad (11)$$

To better follow the characteristic of the video content, the model parameter  $\alpha$  and  $\beta$  need to be updated after encoding one LCU or one frame.

### 3. Proposed algorithm

In this section, a rate control scheme for both bit rate accuracy and video quality consistency is proposed. The GOP level bit allocation, denoted as  $T_{GOP}$ , is obtained using the method mentioned in Section II, which is also the recommended method in HEVC reference software. Our work mainly focuses on the frame level bit allocation and the video quality optimization.

#### 3.1. Frame bit allocation

Frame bit allocation is used to obtain the target bit for each frame in a GOP after the  $T_{GOP}$  has been done. The target bit of current frame  $T_{Cur}$  is usually calculated as follows,

$$T_{Cur} = \frac{T_{GOP} - T_{coded}}{\sum_{NotCoded} \omega_i} \cdot \omega_{Cur} \quad (12)$$

where  $T_{coded}$  denotes the bits of encoded frame in current GOP,  $\omega_i$  is bit allocation weighting factor of each frame in current GOP. Traditionally,  $\omega_i$  is set up equally or at fixed ratio, which will cause the fluctuation of video quality for ignoring the coding complexity and scene change. In this work, we propose a frame bit allocation according to the coding complexity, which is,

$$T_{Cur} = \frac{T_{GOP} - T_{coded}}{(N_{GOP} - N_{coded} - 1)MAD_{R_{avg}} + MAD_{O_{Cur}}} MAD_{O_{Cur}} \quad (13)$$

where  $N_{GOP}$  and  $N_{coded}$  denotes the number of frames for each GOP and number of coded frames in current GOP respectively,  $MAD_{R_{avg}}$  denotes the average real MAD of previous encoded M frames (M is set as 5 in this work),  $MAD_{O_{Cur}}$  is the MAD of current frame which is obtained from the pre-analysis process. Considering both computational complexity and data accuracy, pre-analysis is carefully designed to get the MAD between current frame and its reference frame. For inter frames, only 8x8 motion search with search window 16 is used, while for intra frames, only few directions of prediction (such as horizontal, vertical and diagonal) are used. The computational complexity of the pre-analysis is much lower compared to the complete encoding process.

The proposed method of frame bit allocation is based on the concept of the coding complexity, which could catch the trend of the video content better than the traditional method especially when scene change happens or high motion occurs. After the frame bit allocation, we use R- $\lambda$  model to derive the basic QP (denoted as  $QP_R$ ) of current frame.

#### 3.2. Video quality regulation

The rate control in HM is mainly designed on targeting the bit rate accuracy. However, the consistent video quality is also important in many video applications. The traditional method in rate control to achieve the smooth video quality is to regulate the QP of current frame by the QP of previous coded frames. This method is based on the assumption that the video quality, represents as distortion here, has a linear relationship with the QP. While this assumption is not always suitable for video coding. In this work, we propose a model based method to regulate the video quality.

The distortion is represented as MSE between the original frame and the reconstructed frame. Since the distortion is mainly caused by the quantization process after the transformation in encoding process, the transformed coefficients need to be analysed first. In H.264/AVC, it is well known that the distribution of transformed coefficients is modelled with Laplacian, Cauchy, and generalized Gaussian probability density functions (pdfs). However in HEVC, the Laplace distribution is also well suited for the transformed coefficients [7]. Based on this pdf, D-Q model can be obtained as follows [8],

$$D_L = \frac{Q^2}{12} \quad (14)$$

$$D_H = \frac{2}{\lambda^2} - \frac{10^4}{\lambda e^{7.94\lambda}} Q^{-2} \quad (15)$$

where  $D_L$  and  $D_H$  are the distortions in low bit rate and high bit rate respectively,  $\lambda$  is the model parameter which can be calculated by  $\lambda = \sqrt{2}/\sigma$ , where  $\sigma$  is the standard deviation of transformed coefficients,  $Q$  is the quantization step size. The threshold  $D_{TH}$  between the low bit rate and high bit rate is

$$D_{TH} = 0.79\lambda^{-2} \quad (16)$$

Then we use this D-Q model to regulate the QP of current frame accurately. From our previous work, to minimize the fluctuation of the video quality (minimize the variance of D), the distortion of the current frame should equal to the average distortion of the previous coded frames, which is

$$D_{Cur} = \frac{1}{N} \sum_N D_{coded} \quad (17)$$

where  $D_{Cur}$  is the expected distortion of current coding frame,  $D_{coded}$  are distortions of the previous coded frames,  $N$  is the number of frames that used to calculate the current distortion which is set as 30 here. According to (17), we can get the expected distortion of current frame. Then the expected QP aimed for the smooth video quality (denoted as  $QP_D$ ) is obtained by the D-Q model (14) (15). Finally, the exact QP of current frame is obtained from  $QP_R$  by the regulating of  $[QP_D-2, QP_D+2]$ .

#### 4. Experimental result

In order to evaluate the proposed rate control scheme, we implement the algorithm in HM 16.0. The testing configuration is set as low delay (LD) with 2 reference frames. The target of rate control is CBR and the target bit rate is generated with fixed QP sets {22, 27, 32, 37}. The rate control in HM 16.0 is used as benchmark. The experimental results of bit rate error and BD-Rate are shown in Table I, where the bit rate error is measured as

$$BR \text{ error} = \frac{|Target \ BR - Actual \ BR|}{Target \ BR} \times 100\% \quad (18)$$

And BD-rate is calculated by four actual BR and the corresponding PSNR.

From Table I we can see that the proposed algorithm has the same bit rate accuracy as the benchmark, which is because that both algorithm adopt the R- $\lambda$  model to derive the QP. The BD-rate shows that the proposed algorithm has a 0.33% bit rate saving than benchmark, which means a better coding performance. This is because that the frame bit allocation of the proposed algorithm is based on the coding complexity, which can distribute more bits to the scene with high motion or complicate texture and less bits to the scene with slow motion or simple texture to achieve the better coding performance. The sequences with high motion such as “BasketballDrill” and “BasketballPass” have a bit rate saving 0.43% and 0.45% respectively.

Table 1: Bit rate error and performance (BD-rate)

Class	Sequence	Target BR	Benchmark		Proposed		BD-rate
			Actual BR	BR error	Actual BR	BR error	
B	BasketballDrive	28638	28597	0.14%	28582	0.20%	-0.21%
	BQTerrace	6773	6775	0.03%	6785	0.18%	-0.19%
	Cactus	2131	2132	0.05%	2138	0.33%	-0.23%
	Kimono	4684	4685	0.02%	4702	0.38%	-0.15%
	ParkScene	2007	2004	0.15%	2011	0.20%	-0.38%
C	BasketballDrill	1178	1182	0.34%	1183	0.42%	-0.43%
	BQMall	3770	3765	0.13%	3773	0.08%	-0.41%
	PartyScene	1281	1279	0.16%	1282	0.08%	-0.55%
	RaceHorses	775	773	0.26%	773	0.26%	-0.38%
D	BasketballPass	1552	1548	0.26%	1554	0.13%	-0.45%
	BlowingBubbles	501	496	1.00%	497	0.80%	-0.33%
	BQSquare	1824	1823	0.05%	1825	0.05%	-0.18%
	RaceHorses	245	245	0.00%	245	0.00%	-0.31%
E	FourPeople	425	423	0.47%	427	0.24%	-0.25%
	Johnny	531	532	0.19%	533	0.19%	-0.41%
	KristenAndSara	310	309	0.32%	311	0.32%	-0.39%
<b>Average</b>		-	-	<b>0.22%</b>	-	<b>0.24%</b>	<b>-0.33%</b>

The experimental results of the video quality fluctuation, which is measured by the standard deviation of PSNR, are shown in Table II. The average of the standard deviation of the PSNR are 1.18 and 1.00 for benchmark and proposed algorithm, where the proposed algorithm has a gain of 17.37% on average. The improvement of the video quality fluctuation is achieved by the regulation of QP using the D-Q model. In the sequence “Kimono” which has a scene changing, the proposed algorithm with D-Q model can make the distortion smoother when the scene change occurs. While the algorithm in benchmark, which does not take the distortion into consideration, has a fluctuation on the video quality at the time of scene change. The gain in this sequence reaches 40.68%.

Table 2: Video quality fluctuation

Class	Sequence	Target BR	Benchmark		Proposed		ΔSD
			PSNR-Y	Standard D	PSNR-Y	Standard D	
B	BasketballDrive	28638	39.66	1.12	39.72	0.93	16.95%
	BQTerrace	6773	33.73	0.85	33.68	0.58	31.43%
	Cactus	2131	32.98	0.69	32.93	0.48	30.28%
	Kimono	4684	40.71	0.80	40.85	0.47	40.68%
	ParkScene	2007	34.38	0.99	34.28	0.90	8.42%
C	BasketballDrill	1178	35.17	1.08	35.22	1.01	6.82%
	BQMall	3770	37.21	1.18	37.12	0.98	17.48%
	PartyScene	1281	27.49	1.46	27.55	1.41	3.68%
	RaceHorses	775	29.58	1.31	29.61	1.12	15.05%
D	BasketballPass	1552	42.38	2.50	42.41	2.45	1.77%
	BlowingBubbles	501	31.78	1.37	31.9	1.21	11.68%
	BQSquare	1824	34.77	1.31	34.83	0.91	30.08%
	RaceHorses	245	29.79	1.32	29.96	1.06	19.83%
E	FourPeople	425	36.71	0.83	36.78	0.66	19.88%
	Johnny	531	39.6	0.91	39.68	0.77	16.05%
	KristenAndSara	310	37.34	1.13	37.37	1.04	7.91%
<b>Average</b>		-	-	<b>1.18</b>	-	<b>1.00</b>	<b>17.37%</b>

## 5. Conclusion

In this paper, we propose a rate control scheme for the consistent video quality. Firstly, the frame bit allocation is improved by the MAD-based method, which can follow the characteristic of the video content. Then, to minimize the fluctuation of the frame distortion, D-Q model is used to regulate the QP after QP decision by the R-Q model. Experimental results show that the proposed rate control scheme can achieve better R-D performance than benchmark and the fluctuation of the frame distortion is decreased up to 40.68%.

## 6. Acknowledgements

This work is partially supported by the National Science Foundation of China under contract No. 61502013, China Postdoctoral Science Foundation funded project and National Key Research and Development Program of China under contract No. 2016YFB0401904.

## 7. Reference

- [1] Z. G. Li, F. Pan, K. P. Lim, and G. N. Feng, "Adaptive basic unit layer rate control for JVT," in *Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-G012, 7th Meeting*, Thailand, 2003, pp. 7–14, JVT-G012.
- [2] S. Ma, W. Gao, and Y. Lu, "Rate-distortion analysis for H.264/AVC video coding and its application to rate control," *IEEE TCSVT.*, vol. 15, no. 12, pp. 1533–1544, Dec. 2005.
- [3] S. Wang, S. Ma, S. Wang, D. Zhao, and W. Gao, "Rate-GOP based rate control for high efficiency video coding," *IEEE Journal of Selected Topics in Signal Process.*, vol. 7, no. 6, pp. 1101–1111, Dec. 2013.
- [4] S. Sanz-Rodriguez, and T. Schierl, "A rate control algorithm for HEVC with hierarchical GOP structures," in *Proc. IEEE ICASSP*, Vancouver, May 2013, pp. 1719–1723.
- [5] B. Xie and W. Zeng, "A sequence-based rate control framework for consistent quality real-time video," *IEEE TCSVT*, vol. 16, no. 1, pp. 56–71, Jan. 2006.
- [6] B. Li, H. Li, L. Li, and J. Zhang, "Rate control by R-lambda model for HEVC," in *Proc. 11<sup>th</sup> Meeting Joint Collaborative Team Video Coding (JCT-VC) of ITU-T/SC 29/WG 11*, Shanghai, China, Oct. 10–19, 2012.
- [7] B. Lee and M. Kim, "Modeling rates and distortions based on a mixture of Laplacian distributions for inter-predicted residues in quadtree coding of HEVC," *IEEE Signal Process. Lett.*, vol. 18, pp. 571–574, Oct. 2011.
- [8] Chan-Won Seo, Joo-Hee Moon, and Jong-Ki Han, "Rate Control for Consistent Objective Quality in High Efficiency Video Coding," *IEEE TCSVT.*, vol. 22, no. 6, pp. 2442–2454, Jun. 2013.