An Efficient Frame-Level Rate Control Algorithm for High Efficiency Video Coding

Yubei Lin[1], Xingming Zhang[1], Jianen Xiao[2] and Shengkai Su[2]
[1] School of Computer Science and Engineering, South China University of Technology
Guangzhou, Guangdong – P.R. China
[e-mail: yupilin@scut.edu.cn]
[2] School of Software Engineering, South China University of Technology
Guangzhou, Guangdong – P.R. China
[e-mail: 350877702@qq.com]
*Corresponding author: Yubei Lin

Received October 18, 2015; revised January 25, 2016; accepted February 12, 2016; published April 30, 2016

Abstract

In video coding, the goal of rate control (RC) is not only to avoid the undesirable fluctuation in bit allocation, but also to provide a good visual perception. In this paper, a novel frame-level rate control algorithm for High Efficiency Video Coding (HEVC) is proposed. Firstly a model that reveals the relationship between bit per pixel (*bpp*), the bitrate of the intra frame and the bitrate of the subsequent inter frames in a group of pictures (GOP) is established, based on which the target bitrate of the first intra frame is well estimated. Then a novel frame-level bit allocation algorithm is developed, which provides a robust bit balancing scheme between the intra frame and the inter frames in a GOP to achieve the visual quality smoothness throughout the whole sequence. Our experimental results show that when compared to the RC scheme for HEVC encoder HM-16.0, the proposed algorithm can produce reconstructed frames with more consistent objective video quality. In addition, the objective visual quality of the reconstructed frames can be improved with less bitrate.

Keywords: High Efficiency Video Coding (HEVC), video coding, rate control, structural similarity (SSIM) index, consistent video quality

This work is financially supported by the Fundamental Research Funds for the Central Universities of China (Grant No. 2014ZM0049).

http://dx.doi.org/10.3837/tiis.2016.04.023 **ISSN : 1976-7277**

1. Introduction

High Efficiency Video Coding (HEVC), the newest video coding standard, has been proved superior in coding efficiency over its precedents [1]. A few new techniques, such as larger coding tree unit (CTU) sizes and improved parallel processing methods, have been adopted in HEVC to improve two aspects: increasing video resolution and increasing use of parallel processing architectures, which cannot be settled very well by the prior standard H.264/AVC [2].

In general, the quantitation parameter (QP) should be specified when using the HEVC codec to encode video stream. Also, a target bitrate is often given as a parameter for video encoding according to the application environment. In order to regulate the encoded bit stream such that the best video quality can be achieved without violating the constraints imposed by the encoder/decoder buffer size and the available channel bandwidth, rate control (RC) is widely adopted in the video coding standard-based encoders, such as MPEG-4 [3, 4], H.264/AVC [2] and HEVC [1]. The new characteristics of HEVC pose a new challenge for designing accurate and robust RC and remain RC a hot research issue.

In Section 2, we review the previous work in RC. Section 3 establishes a model that reveals the relationship between bit per pixel (*bpp*), the bitrate of the intra frame and the bitrate of the subsequent inter frames in a group of pictures (GOP), and details how to determine the target bitrate of the first intra frame. A robust and adaptive frame-level RC scheme is given in Section 4. Simulation results are presented in Section 5, followed by the conclusion in Section 6.

2. Related Work

There have been a number of investigations for RC. The existing RC schemes can be roughly categorized into three classes: Q-domain model, ρ-domain model, and λ-domain model. The Q-domain model builds direct relationship between bitrate and QP. In [5], Cauchy-distribution-based R-Q model was used to determine QP. In [6], a Laplace-distribution-based CTU level RC algorithm for HEVC was proposed. To overcome the high computational complexity, Choi *et al*. presented a pixel-wise unified R-Q model for multi-level RC [7]. In the ρ-domain-based RC algorithms [8, 9], bitrate was modeled as a linear function of ρ, which was the percentage of zeroes in discrete cosine transform (DCT) coefficients. Since it is assumed that there is a one-to-one relationship between ρ and QP, the suitable QP can be determined to meet the target bitrate through ρ. In fact, both Q-domain and ρ-domain RC models utilize a close relationship between bitrate and QP. However, it becomes more difficult to accurately characterize the relationship between bitrate and QP when the video coding scheme is becoming more flexible. In [10], a λ-domain RC algorithm for HEVC was proposed for inter-frame coding. Since it achieves high coding performance, it has been already adopted by Joint Collaborative Team on Video Coding (JCT-VC) and integrated into the state-of-the-art RC scheme for HEVC encoder HM-16.0 [11], together with the λ-domain-based RC scheme for intra-frame coding which is proposed in [12].

In the state-of-the-art RC scheme for HM-16.0, sum of absolute transformed difference (SATD) is employed to measure the complexity of intra frames [12], and mean absolute difference (MAD) of the CTU at the same position in the previous decoded frame is used to predict the complexity of the current CTU [10]. These frame complexity measures are very simple. However, they perform poorly in allocating the target bits. It can be observed in our experiments that a large number of bits are over-spent in the intra frame and there are not enough target bits left to the subsequent inter frames in the same GOP. This may unavoidably lead to degradation of visual quality for these inter frames and undesirable fluctuation in the actual frame-level bitrate allocation. In order to solve this problem, some efforts have been done for seeking a more accurate content complexity measure. Several edge-based [13] and gradient-based [14] content complexity measures have been developed for H.264/AVC. However, since coding characteristics of HEVC are quite different from H.264/AVC, these methods for H.264/AVC are no longer applicable for HEVC. In [15, 16], variance-based methods were proposed to measure the complexity for HEVC intra prediction. In [17, 18], gradient was used to denote the picture content complexity for HEVC intra frame RC. Sun *et al*. proposed an edge-based frame complexity measure using the Gaussian gradient operator [19]. In [20], a model considering the spatial-temporal correlations was developed to measure the texture complexity, in which spatial complexity and temporal complexity referred to the texture similarities inside a single video frame and the stillness between consecutive frames in the temporal dimension, respectively.

The goal of RC is to avoid the undesirable fluctuation in bit allocation. For providing a good visual perception, it is also very important to avoid the video quality fluctuation. In [21], a RC algorithm was proposed to keep the consistent objective quality for HEVC, where distortion-quantization and rate-quantization models were derived using the Laplacian function. Unfortunately, this algorithm was developed in Q-domain, which is not appropriate for the state-of-the-art RC scheme for HM-16.0.

To better address the issues mentioned above, we attempt to achieve more accurate bit allocation and keep consistent object video quality with a different approach for HEVC in this paper. Unlike the complexity-based methods, the bit proportion of a GOP allocated to the intra frame is investigated first. Based on the research results, a novel frame-level bit allocation algorithm is developed, which provides a robust bit balancing scheme between intra frame and inter frame in a GOP to achieve the visual quality smoothness throughout the whole video sequence. Note that structural similarity (SSIM) index [22] is employed to measure the image quality in this paper since it has been shown to be effective and well matched to the perceived quality [23].

3. Initial Target Bitrate of the First Intra Frame

An accurate estimation of the initial target bitrate of the first intra frame is vital to improve the overall performance of RC. The more accurate the estimation is, the less time it will take to adjust the bit cost to a steady state. However, the initialization scheme in HM-16.0 only takes *bpp* into consideration, which is certainly not accurate [24]. Those complexity-based methods mentioned in Section 2 are not appropriate for HM-16.0 since new characteristics have been adopted in the state-of-the-art HEVC. In this section, we develop a novel but simple model to estimate the initial target bitrate of the first intra frame with the following new characteristics: (1) it assumes that the GOP structure is IB…B (an I frame followed by n B frames) and (2) the relationship between *bpp*, the bitrate of the I frame and the bitrate of the B frames in a GOP is investigated.

Denote R_{GOP} as the target bitrate of a GOP, then

$$R_{GOP} = R_I + n \cdot \bar{R}_B \tag{1}$$

where R_I is the target bitrate of the I frame and $\bar{R}_B$ is the average target bitrate of the B frames in the GOP. Let

$$y = \frac{R_I}{\bar{R}_B} \tag{2}$$

Then from Eqs. (1) and (2), we can obtain

$$R_{GOP} = R_I + n \cdot \frac{R_I}{y} \tag{3}$$

i.e.

$$R_I = \frac{R_{GOP}}{1 + \frac{n}{y}} \tag{4}$$

In order to discover the relationship between *y* and *bpp*, some experiments have been performed in HM-16.0, in which *n* = {3, 7, 11, 15} and flat QP (QPs = {17, 22, 27, 32, 37, 42}) are used in encoding. By performing curve fitting on extensive data, we find that this relationship can be accurately modeled by a Hyperbolic function as follows, which is represented by colorful curves in **Fig. 1**:

$$y = a \cdot bpp^b \tag{5}$$

where *a* and *b* are the model parameters, and *bpp* can be calculated by:

$$bpp = \frac{R}{f \cdot w \cdot h} \tag{6}$$

where *R* is the bitrate of the sequence, *f* is the frame rate, *w* and *h* are the width and height of the picture respectively.

Note that R^2 in **Fig. 1** is the correlation coefficient, which is between 0 and 1. The bigger R^2 is, the closer the approximated curve is to the actual data. From the curve fitting results in **Fig. 1**, we can conclude that the model can fit the actual data points very well. According to Eqs. (4) and (5), R_I can be obtained by:

$$R_I = \frac{R_{GOP}}{1 + \frac{n}{a \cdot bpp^b}} \tag{7}$$

Once the model parameters *a* and *b* in Eq. (5) are determined, R_I can be well estimated. It can be observed from **Fig. 1** that *a* and *b* are quite different for different video sequence, but for the same sequence, *a* and *b* for different GOP sizes are quite similar. Therefore, a simple solution is developed to obtain the parameters *a* and *b* for a specified sequence as follows:

(1) Pre-encode the first GOP using the original HM-16.0 with RC off.

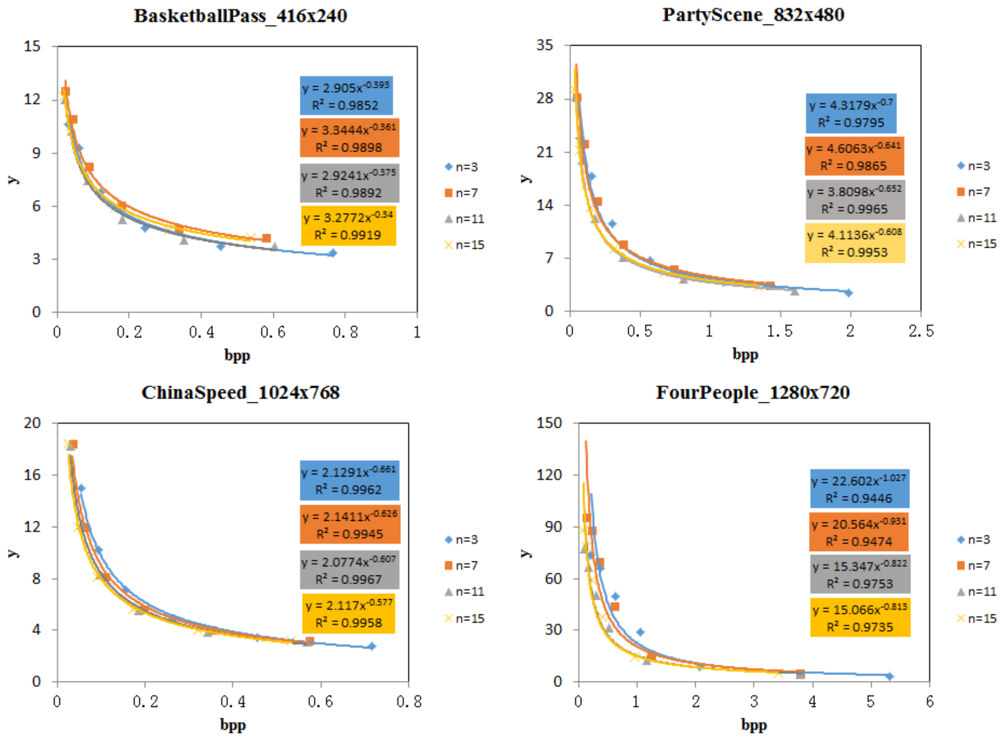

Fig. 1. *bpp-y* curves fitting according to Eq. (5)

(2) Compute the average QP of the B frames in the first GOP, and round it to the nearest integer, which is denoted as QP_0.

(3) Use QP_0 to encode the first two GOPs, from which two groups of actual *bpp* and *y* are obtained and denoted as (bpp_1, y_1) and (bpp_2, y_2), respectively.

(4) According to Eq. (5), the two sets of *bpp* and *y* from step (3) yield two equations:

$$\begin{cases} y_1 = a \cdot bpp_1^b \\ y_2 = a \cdot bpp_2^b \end{cases} \tag{8}$$

Hence, the parameters *a* and *b* can be determined by solving Eq. (8) and represented as follows:

$$\begin{cases} b = (\ln y_1 - \ln y_2)/(\ln bpp_1 - \ln bpp_2) \\ a = y_1/bpp_1^b \end{cases} \tag{9}$$

Denote R_{I1} as the target bitrate of the first intra frame, then according to Eq. (7), there is

$$R_{I1} = \frac{R_{GOP1}}{1 + \dfrac{n}{a \cdot bpp^b}} \tag{10}$$

where R_{GOP1} is the target bitrate of the first GOP.

4. Proposed Frame-Level RC Algorithm

In Section 3, an algorithm to estimate the target bitrate of the first intra frame is proposed, in which some frames need to be pre-encoded to get the model parameters. If all the intra frames in the video sequence are encoded in this way, extra complexity will be introduced and it will bring negative effects for real-time application. Furthermore, the RC scheme should be adaptive to the video content and achieve the visual quality smoothness throughout the whole video sequence. To intelligently balance bit allocation between intra and inter frames in a GOP, *y* should be dynamically updated. If the visual quality of the coded intra frame is higher than the average one of the inter frames in the same GOP, it means relatively more bits have been allocated to the intra frame and *y* of the next GOP should be decreased. Otherwise, more bits should be allocated to the next intra frame to improve its quality. After encoding the *i*th GOP, *y* of the next GOP, denoted as y_{i+1}, can be updated as follows:

$$y_{i+1} = \frac{n}{\frac{\tilde{R}_{GOPi}}{\tilde{R}_{Ii}} - 1} \cdot e^{k \cdot (\overline{SSIM}_{Bi} - SSIM_{Ii})} \tag{11}$$

where $\tilde{R}_{GOPi}$ is the actual bitrate of the *i*th GOP, $\tilde{R}_{Ii}$ is the actual bitrate of the intra frame in the *i*th GOP, *k* is an adjustment factor, $\overline{SSIM}_{Bi}$ is the average SSIM value of the inter frames in the *i*th GOP, and $SSIM_{Ii}$ is the SSIM value of the intra frame in the *i*th GOP. Note that *k* is empirically set as follows:

$$k = \begin{cases} 10 & if\ \overline{SSIM}_{Bi} - SSIM_{Ii} \leq 0.001 \\ 20 & else\ if\ 0.001 < \overline{SSIM}_{Bi} - SSIM_{Ii} \leq 0.005 \\ 40 & otherwise \end{cases} \tag{12}$$

By combining the *y*-updating strategy and the initial target bitrate estimation for the first intra frame described in Section 3, our proposed frame-level RC algorithm can be summarized as follows:

(1) Obtain *a* , *b*, and R_{I1}, as described in Section 3, and set the initial y_1 as follows:

$$y_1 = \frac{n}{\frac{R_{GOP1}}{R_{I1}} - 1} \tag{13}$$

(2) For the *i*th GOP
  a. Estimate the target bitrate of the intra frame according to Eq. (4).
  b. Encode each frame in the current GOP and obtain $\tilde{R}_{Ii}$, $\tilde{R}_{GOPi}$, $\overline{SSIM}_{Bi}$ and $SSIM_{Ii}$.
  c. Update y_{i+1} according to Eq. (11).

(3) Go to step (2) until the end of a sequence.

5. Experimental Results and Analysis

To evaluate the performance of the proposed algorithm, numerous experiments have been conducted. Twenty sequences from Class A, B, C, D, and F as specified in [25] are used for simulation. The detail information of the tested sequences is summarized in **Table 1**. In the experiments, Random Access (RA) Main Profile configuration is used. To conduct a fair comparison between our algorithm and the original RC scheme [10, 12] in HM-16.0, we assign a target bitrate for each sequence which is obtained by performing the original HM-16.0 with RC disabled according to the HEVC common test conditions, then perform these two RC methods with the same configuration. It should be noted that the standard deviation of SSIM is implemented to measure the variation of video quality, and Bitrate error in Eq. (14) is calculated to measure the RC accuracy:

$$Bitrate\ error = \frac{R_{act} - R_{tar}}{R_{tar}} \times 100\% \tag{14}$$

where R_{act} is the actual bitrate and R_{tar} is the target bitrate of the sequence.

Table 1. Information of test sequences used for simulation

Class	Resolution	Sequence	Abbreviation	Tested frames	Frame rate(fps)
A	2560×1600	Traffic	TRF	First 100	30
		PeopleOnStreet	POS	First 150	30
		Nebuta	NBT	First 100	60
		SteamLocomotive	SLM	First 100	60
B	1920×1080	Kimono	KMN	First 200	24
		ParkScene	PKS	First 200	24
		Cactus	CAC	First 200	50
		BQTerrace	BQT	First 200	60
		BasketballDrive	BBD	First 200	50
C	832×480	RaceHorses	RHC	First 200	30
		BQMall	BQM	First 200	60
		PartyScene	PTS	First 200	50
		BasketballDrill	BBR	First 200	50
D	416×240	RaceHorses	RHD	First 200	30
		BQSquare	BQS	First 200	60
		BlowingBubbles	BWB	First 200	50
		BasketballPass	BBP	First 200	50
F	1024×768	ChinaSpeed	CNS	First 200	30
		SlideEditing	SDE	First 200	30
		SlideShow	SDS	First 200	20

First we perform experiments with coding structure IBBBIBBB… and GOP size equal to 4. All the sequences in **Table 1** are tested. We use ten sequences as example (two sequences from each class) to show the efficiency of the proposed method in **Tables 2** and **3**. The overall results are illustrated in **Table 4**. From **Tables 2**, **3** and **4**, it can be observed that when compared with the HM-16.0 RC method, our method significantly reduces the bitrate error. The average bitrate error of the HM-16.0 RC method is 7.16% while that of our proposed method is only 0.46%. Meanwhile, our algorithm saves more than 428 kbps bitrate on average when compared with the HM-16.0 RC method. In addition, we can find that when compared with the HM-16.0 RC method, our proposed algorithm provides better visual quality of sequences, which obtains up to 0.027440 SSIM value improvement. Moreover, the SSIM variation values of our proposed method are much smaller than that of HM-16.0 RC method, which shows the proposed method has the ability to keep more consistent objective video quality.

Table 2. Simulation results on RC (coding structure: IBBBIBBB…, GOP size: 4)

Seq.	Target bitrate (kbps)	RC in HM-16.0		Proposed		Bitrate reduction (kbps)
		Actual bitrate(kbps)	Bitrate error(%)	Actual bitrate(kbps)	Bitrate error(%)	
POS	63788.067	64596.707	1.27	63786.1424	0.00	810.5646
	30411.923	30924.771	1.69	30478.4144	0.22	446.3566
	15349.37	16370.066	6.65	15386.7632	0.24	983.3028
	8346.659	9364.485	12.19	8363.1408	0.20	1001.3442
NBT	127576.134	130422.749	2.23	126815.7264	-0.60	3607.0226
	60823.846	62981.155	3.55	60541.2144	-0.46	2439.9406
	30698.74	32199.038	4.89	30646.3344	-0.17	1552.7036
	16693.318	18131.63	8.62	16680.6144	-0.08	1451.0156
PKS	10574.594	10809.126	2.22	10560.42624	-0.13	248.69976
	5192.2	5514.745	6.21	5187.17472	-0.10	327.57028
	2625.252	3033.22	15.54	2622.88992	-0.09	410.33008
	1283.081	1616.676	26.00	1282.09824	-0.08	334.57776
BQT	26436.485	26979.288	2.05	26385.0672	-0.19	594.2208
	12980.5	14717.083	13.38	12975.5424	-0.04	1741.5406
	6563.13	7572.13	15.37	6560.7456	-0.04	1011.3844
	3207.702	3476.633	8.38	3207.0624	-0.02	269.5706
RHC	12712.984	12789.715	0.60	12687.5832	-0.20	102.1318
	6585.082	6620.632	0.54	6570.6648	-0.22	49.9672
	3152.902	3205.718	1.68	3147.3612	-0.18	58.3568
	1283.384	1384.958	7.91	1281.2052	-0.17	103.7528
PTS	21188.308	21319.66	0.62	21130.606	-0.27	189.054
	10975.138	11238.702	2.40	10953.44	-0.20	285.262
	5254.838	5524.758	5.14	5246.41	-0.16	278.348
	2138.974	2456.06	14.82	2136.386	-0.12	319.674
BWB	2539.17	2563.16	0.94	2532.894	-0.25	30.266
	1339.414	1361.234	1.63	1336.4	-0.23	24.834
	647.34	696.998	7.67	646.112	-0.19	50.886
	313.14	357.334	14.11	312.636	-0.16	44.698
BBP	2539.17	2563.16	0.94	2532.894	-0.25	30.266
	1339.414	1361.234	1.63	1336.4	-0.23	24.834
	647.34	696.998	7.67	646.112	-0.19	50.886
	313.14	357.334	14.11	312.636	-0.16	44.698
CNS	10135.331	10355.599	2.17	10111.4352	-0.24	244.1638
	5782.771	6129.001	5.99	5773.7892	-0.16	355.2118
	3149.878	3436.21	9.09	3145.806	-0.13	290.404
	1724.377	1989.652	15.38	1723.3548	-0.06	266.2972

SDS	1927.172	2023.562	5.00	2081.292	8.00	-57.73
	1248.699	1297.197	3.88	1400.9792	12.20	-103.7822
	771.733	845.818	9.60	787.0352	1.98	58.7828
	470.82	560.086	18.96	481.164	2.20	78.922

Table 3. Simulation results on objective quality (coding structure: IBBBIBBB…, GOP size: 4)

Seq.	Target bitrate (kbps)	RC in HM-16.0		Proposed		SSIM improvement
		SSIM	SSIM variation	SSIM	SSIM variation	
POS	63788.067	0.996820	0.007436	0.999436	0.000122	0.002616
	30411.923	0.991607	0.015208	0.998588	0.000449	0.006981
	15349.37	0.991688	0.010954	0.996926	0.001017	0.005238
	8346.659	0.987074	0.011631	0.993555	0.002098	0.006481
NBT	127576.134	0.989814	0.011985	0.996703	0.000829	0.006889
	60823.846	0.983944	0.014015	0.991265	0.003340	0.007321
	30698.74	0.977563	0.015271	0.986742	0.003939	0.009179
	16693.318	0.969456	0.017159	0.983370	0.002865	0.013915
PKS	10574.594	0.991060	0.012815	0.996310	0.000603	0.005250
	5192.2	0.987183	0.009731	0.992798	0.001411	0.005616
	2625.252	0.981580	0.009469	0.987163	0.002434	0.005583
	1283.081	0.972914	0.010171	0.977117	0.004681	0.004203
BQT	26436.485	0.996962	0.003719	0.998347	0.000241	0.001385
	12980.5	0.996517	0.001588	0.997353	0.000612	0.000836
	6563.13	0.994645	0.002050	0.995973	0.001438	0.001327
	3207.702	0.986024	0.008425	0.992516	0.004338	0.006492
RHC	12712.984	0.997349	0.002990	0.998258	0.001253	0.000909
	6585.082	0.990643	0.012467	0.995897	0.003066	0.005254
	3152.902	0.980506	0.019306	0.991046	0.006077	0.010540
	1283.384	0.965938	0.022473	0.979368	0.012017	0.013430
PTS	21188.308	0.993739	0.016603	0.998062	0.001080	0.004323
	10975.138	0.984941	0.026131	0.995043	0.001704	0.010102
	5254.838	0.973780	0.023303	0.986672	0.004879	0.012892
	2138.974	0.951955	0.019723	0.964439	0.007579	0.012484
BWB	2539.17	0.988298	0.023477	0.996438	0.001664	0.008139
	1339.414	0.978016	0.030329	0.992339	0.003288	0.014324
	647.34	0.973754	0.018650	0.984210	0.004879	0.010456
	313.14	0.963975	0.018765	0.969723	0.007926	0.005748
BBP	2539.17	0.991350	0.021248	0.997541	0.001720	0.006191
	1339.414	0.978928	0.036912	0.994218	0.002570	0.015290
	647.34	0.973231	0.025718	0.986561	0.004405	0.013330
	313.14	0.954612	0.029621	0.971076	0.007856	0.016464
CNS	10135.331	0.997133	0.007749	0.999543	0.000112	0.002410
	5782.771	0.996162	0.004408	0.998705	0.000376	0.002543

	3149.878	0.992353	0.006113	0.996242	0.001181	0.003889
	1724.377	0.985711	0.007854	0.990751	0.002363	0.005040
SDS	1927.172	0.999783	0.000453	0.999925	0.000320	0.000142
	1248.699	0.996822	0.005008	0.999695	0.001466	0.002873
	771.733	0.997711	0.004095	0.999294	0.001292	0.001583
	470.82	0.998523	0.002512687	0.999209	0.001297	0.002616

Table 4. Overall results of simulation (coding structure: IBBBIBBB…, GOP size: 4)

		RC in HM-16.0		**Proposed**		**Bitrate reduction (kbps)**	**SSIM improvement**
		Bitrate error(%)	**SSIM variation**	**Bitrate error(%)**	**SSIM variation**		
Class A	**Avg.**	3.81	0.010627	-0.12	0.001369	1167.04	0.005898
	Max.	12.19	0.017159	0.24	0.003939	3607.02	0.013915
Class B	**Avg.**	9.95	0.006336	-0.15	0.001633	532.04	0.003183
	Max.	26.00	0.013379	-0.02	0.004681	1741.54	0.006492
Class C	**Avg.**	3.63	0.014451	-0.20	0.002720	172.42	0.006706
	Max.	14.82	0.026131	-0.12	0.012017	319.67	0.013430
Class D	**Avg.**	8.43	0.023034	-0.19	0.004531	55.60	0.011412
	Max.	26.46	0.037720	-0.07	0.011076	133.99	0.027440
Class F	**Avg.**	9.99	0.018470	3.99	0.006620	106.58	0.005506
	Max.	18.96	0.048496	17.23	0.033129	355.21	0.016689
Overall	**Avg.**	7.16	0.013977	0.46	0.003125	428.01	0.006425
	Max.	26.46	0.048496	17.23	0.033129	3607.02	0.027440

Then we perform experiments with coding structure IBBBBBBBIBBBBBBB… and GOP size equal to 8 to show the performance of the proposed method with bigger GOP size. All the sequences in Class C and D are tested. The overall results are illustrated in Table 5. We can find that in such configuration the proposed algorithm also works better than the HM-16.0 RC method. The bitrate error and SSIM variation of the proposed algorithm are far less than those of the HM-16.0 RC method. Meanwhile, the proposed RC scheme saves up to 793.50 kbps bitrate when compared with the HM-16.0 RC method. Besides bitrate reduction, the proposed method also has 0.001087 SSIM value on average, up to 0.009122 SSIM value gain over the HM-16.0 RC scheme.

Table 5. Overall results of simulation (coding structure: IBBBBBBBIBBBBBBB…, GOP size: 8)

		RC in HM-16.0		**Proposed**		**Bitrate reduction (kbps)**	**SSIM improvement**
		Bitrate error(%)	**SSIM variation**	**Bitrate error(%)**	**SSIM variation**		
Class C	**Avg.**	1.16	0.008125	-1.44	0.004122	63.10	0.000403
	Max.	4.54	0.024266	-0.88	0.017339	135.77	0.005312
Class D	**Avg.**	1.84	0.011243	-11.06	0.003946	115.90	0.001770
	Max.	4.89	0.029432	-0.73	0.010870	793.50	0.009122
Overall	**Avg.**	1.50	0.009684	-6.25	0.004034	89.50	0.001087
	Max.	4.89	0.029432	-0.73	0.017339	793.50	0.009122

Fig. 2 presents the frame-level bit allocation comparison of these two RC methods. It is obvious that when compared with the HM-16.0 RC method, the proposed RC method can keep the bit cost of different pictures in a video sequence within a narrower range.

Fig. 3 demonstrates the frame-level objective visual quality comparison of these two RC methods. It can be observed that the SSIM curves of our proposed algorithm are consistently higher than those of the HM-16.0 RC method throughout the whole sequences. The HM-16.0 RC method unavoidably leads to obvious degradation of visual quality in the later part of the sequences. But our proposed algorithm can perform very well without much fluctuation in objective visual quality of the reconstructed frames under the same conditions.

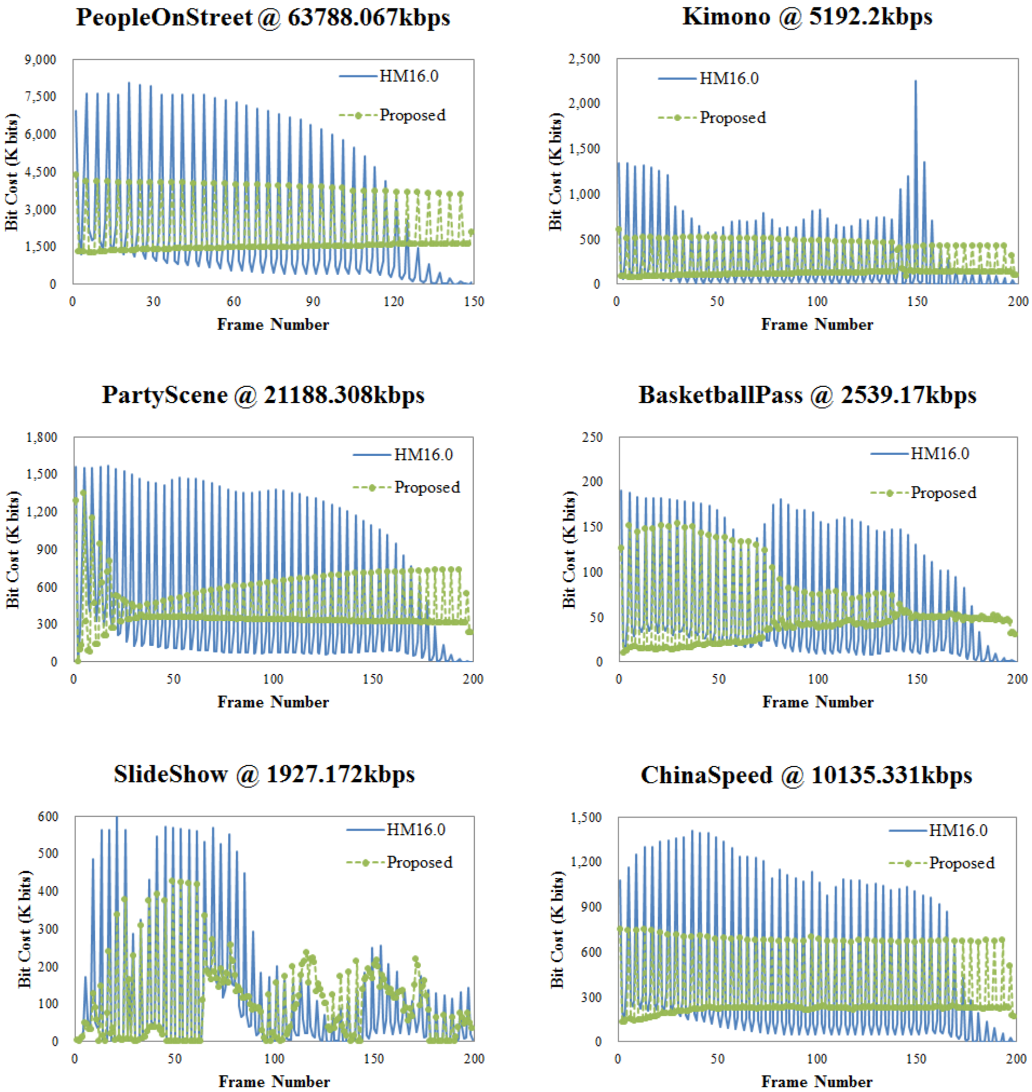

Fig. 2. Bit cost comparison between the HM-16.0 RC method and our proposed algorithm (coding structure: IBBBIBBB…, GOP size: 4)

6. Conclusion

This paper presents a novel and efficient frame-level rate control algorithm for HEVC. An accurate estimation of the initial target bitrate of the first intra frame is proposed. Then a balanced frame-level bit allocation strategy is designed to improve the overall performance of RC scheme for HM-16.0. The simulation results show that the proposed algorithm is able to achieve more accurate RC and obtain better and smoother visual quality of reconstructed pictures with less bitrate when compared to the HM-16.0 RC method.

Regarding future work directions, with the objective to further enhance the overall performance of RC scheme, we will continue our research on exploring perceptual approaches for basic-unit-level bit allocation.

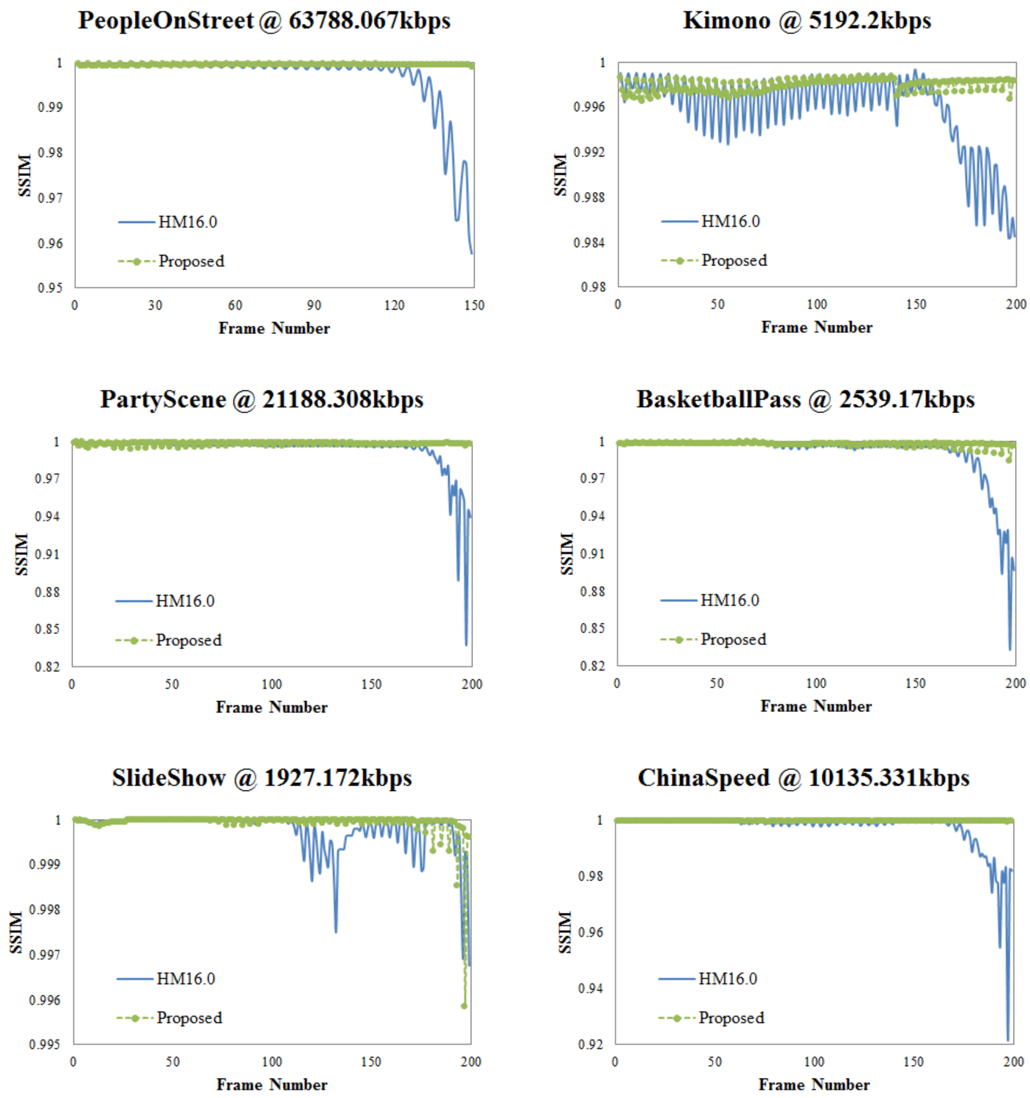

Fig. 3. Objective visual quality comparison between the HM-16.0 RC method and our proposed algorithm (coding structure: IBBBIBBB…, GOP size: 4)

References

[1] J. R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan and T. Wiegand, "Comparison of the coding efficiency of video coding standards-including High Efficiency Video Coding (HEVC)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1669-1684, December, 2012. Article (CrossRef Link).

[2] T. Wiegand, G. J. Sullivan, G. Bjontegaard and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-576, July, 2003. Article (CrossRef Link).

[3] A. Vetro, H. Sun and Y. Wang, "MPEG-4 rate control for multiple video objects," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 1, pp. 186-199, February, 1999. Article (CrossRef Link).

[4] H. J. Lee, T. H. Chiang and Y. Q. Zhang, "Scalable rate control for MPEG-4 video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 6, pp. 878-894, September, 2000. Article (CrossRef Link).

[5] Y. J. Yoon, H. Kim, S. H. Jung, D. S. Jun, Y. H. Kim, J. S. Choi and S. J. Ko, "A new rate control method for hierarchical video coding in HEVC," in *Proc. of IEEE International Symp. on Broadband Multimedia Systems and Broadcasting (BMSB)*, pp. 1-4, June 27-29, 2012. Article (CrossRef Link).

[6] J. J. Si, S. W. Ma, S. Q. Wang and W. Gao, "Laplace distribution based CTU level rate control for HEVC," in *Proc. of Visual Communications and Image Processing (VCIP)*, pp. 1-67, November 17-20, 2013. Article (CrossRef Link).

[7] H. M. Choi, J. H. Yoo, J. H. Nam, D. Y. Sim and I. V. Baji, "Pixel-wise unified rate-quantization model for multi-level rate control," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 1112-1123, December, 2013. Article (CrossRef Link).

[8] T. Biatek, M. Raulety, J. F. Traversz and O. Deforges, "Efficient quantization parameter estimation in HEVC based on ρ-domain," in *Proc. of the 22nd European Signal Processing Conference (EUSIPCO)*, pp. 296-300, September 1-5, 2014. Article (CrossRef Link).

[9] S. S. Wang, S. W. Ma, S. Q. Wang, D. B. Zhao and W. Gao, "Quadratic ρ-domain based rate control algorithm for HEVC," in *Proc. of IEEE International Conf. of Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1695-1699, May 26-31, 2013. Article (CrossRef Link).

[10] B. Li, H. Li, L. Li and J. Zhang, "Rate control by R-lambda model for HEVC," *ITU-T SG16 Contribution, JCTVC-K0103*, Shanghai, October, 2012. Article (CrossRef Link).

[11] JCT-VC of ISO/IEC MPEG and ITU-T VCEG, "HM Reference Software 16.0 [Online]," Article (CrossRef Link).

[12] M. Karczewicz and X. L. Wang, "Intra frame rate control based on SATD," in *Proc. of 13th Meeting of JCTVC-M0257, ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11*, Incheon, KR, April 18-26, 2013. Article (CrossRef Link).

[13] Z. G. Cui and X. C. Zhu, "Image complexity adaptive intra-frame rate control algorithm for H.264," *Journal of Electronics & Information Technology*, vol. 32, no. 11, pp. 2547-2553, November, 2010. Article (CrossRef Link).

[14] Y. M. Zhou, Y. Sun, Z. D. Feng and S. X. Sun, "New rate-distortion modeling and efficient rate control for H.264/AVC video coding," *Signal Processing: Image Communication*, vol. 24, no. 5, pp. 345-356, May, 2009. Article (CrossRef Link).

[15] W. Q. Zhao, L. Q. Shen, Z. M. Cao and Z. Y. Zhang, "Texture and correlation based fast intra prediction algorithm for HEVC," in *Proc. of 9th International Forum on Digital TV and Wireless Multimedia Communication*, pp. 284-291, November 9-10, 2012. Article (CrossRef Link).

[16] G. F. Tian and S. Goto, "Content adaptive prediction unit size decision algorithm for HEVC intra coding," in *Proc. of 2012 Picture Coding Symposium*, pp. 405-408, May 7-9, 2012. Article (CrossRef Link).

[17] L. Tian, Y. M. Zhou and X. J. Cao, "A new rate-complexity-QP algorithm (RCQA) for HEVC intra-picture rate control," in *Proc. of 2014 International Conf. on Computing, Networking and Communications*, pp. 375-380, February 3-6, 2014. Article (CrossRef Link).

[18] M. H. Wang, K. N. Ngan and H. L. Li, "An efficient frame-content based intra frame rate control for High Efficiency Video Coding," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 896-900, July, 2015. Article (CrossRef Link).

[19] L. Sun, O. C. Au, W. Dai, Y. F. Guo and R. B. Zou, "An adaptive frame complexity based rate quantization model for intra-frame rate control of High Efficiency Video Coding (HEVC)," in *Proc. of 2012 Asia-Pacific Signal & Information Processing Association Annual Summit and Conference*, pp. 1-6, December 3-6, 2012. Article (CrossRef Link).

[20] H. Sun, S. S. Gao and C. Zhang, "Adaptive bit allocation scheme for rate control in High Efficiency Video Coding with initial quantization parameter determination," *Signal Processing: Image Communication*, vol. 29, no. 10, pp. 1029-1045, November, 2014. Article (CrossRef Link).

[21] C. W. Seo, J. H. Moon and J. K. Han, "Rate control for consistent objective quality in High Efficiency Video Coding," *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2442-2454, June, 2013. Article (CrossRef Link).

[22] W. Zhou, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April, 2004. Article (CrossRef Link).

[23] W. Zhou and A. C. Bovik, "Mean squared error: love it or leave it? A new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98-117, January, 2009. Article (CrossRef Link).

[24] H. Choi, J. Nam, J. Yoo, D. Sim and I. Baji ć, "Improvement of the rate control based on pixel-based URQ model for HEVC," in *Proc. of 9th Meeting of JCTVCI0094, ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11*, 27 April -7 May, 2012. Article (CrossRef Link).

[25] F. Bossen, "HM 8 common test conditions and software reference configurations," *ITU-T SG16 Contribution, JCTVC-J1100*, Stockholm, July, 2012. Article (CrossRef Link).

Yubei Lin received the B. S. degree in computational mathematics and M. Eng. degree in computer software and theory from the Sun Yat-sen University, Guangzhou, China, in 2002 and 2005, respectively. She joined the Hong Kong Polytechnic University as a Research Assistant in 2004. She has been an Engineer with the Computing Center of South China University of Technology, Guangzhou, China, since 2005, where she is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering. Her research interests include video coding, image processing and information processing.

Xingming Zhang received the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 1996. He is currently a Professor, Doctoral Supervisor, and Vice-dean with the School of Computer Science and Engineering, South China University of Technology, Guangzhou, China. He is a member of the Standing Committee of the Technical Committee of Education, China Computer Federation, and an Executive Director of the Computer Federation of Guangdong Province. His research focuses on image processing, video coding and surveillance.

Jianen Xiao is currently pursuing the B. S. degree with the School of Software Engineering, South China University of Technology, Guangzhou, China. His research interests include video coding and communication.

Shengkai Su is currently pursuing the B. S. degree with the School of Software Engineering, South China University of Technology, Guangzhou, China. His research interests include video coding and communication.