A Deep Learning-Based Rate Control for HEVC Intra Coding

Ismail Marzuki, *Donggyu Sim
Kwangwoon University
ismail@kw.ac.kr, *dgsim@kw.ac.kr

Abstract

This paper proposes a rate control algorithm for intra coding frame in HEVC encoder using a deep learning approach. The proposed algorithm is designed for CTU level bit allocation in intra frame by considering visual features spatially and temporally. Our features are generated using visual geometry group (VGG-16) with deep convolutional layers, then it is used for bit allocation per each CTU within an intra frame. According to our experiments, the proposed algorithm can achieve -2.04% Luma component BD-rate gain with minimal bit accuracy loss against the HM-16.20 rate control model.

Keywords: Deep neural network, High efficiency video coding (HEVC), Rate control, VGG-16, Video coding

1. Introduction

High Efficiency Video Coding (HEVC) is the latest video coding standard. It has been developed with many advanced coding tools to improve its predecessor, H.264/Advanced Video Coding (AVC) standard [1]. Thanks to its remarkable performances, HEVC has now been used as a basic technology for next video coding standard, Versatile Video Coding (VVC) [2]. Among those coding tools adopted in HEVC, rate control has always been one of the most important tools for HEVC encoder in maintaining its entire performances: coding efficiency and visual quality. Adopting a λ-domain method rate control [3], which outperforms a pixel based unified rate-quantization (URQ) rate control model [4], HEVC achieves significantly improvements in coding efficiency [5]. However, λ-domain or rate (R)-λ has been organized without considering any visual aspect. Many algorithms have been developed R-λ rate control model by taking visual characteristics. Few studies have provoked deep neural network to get better performance over the traditional model in HEVC encoders [6].

In this paper, we propose a CTU level bit allocation for intra coding frame in HM-16.20 by taking visual geometry group (VGG-16) network [7], one of famous deep neural network (DNN) architectures in image classification. We use the predefined VGG-16 model for extracting our feature descriptors. Two features, spatial and temporal features, are introduced for the algorithm. We estimate bit budgets, lambda, and quantization parameter (QP) values for each coding tree unit (CTU) in intra frame with the extracted features based on spatially block characteristics and VGG-16 model. Our experiments yield that the proposed algorithm achieves better coding gain over the rate control in HM.

This paper is organized as follows. In Section 2, the proposed algorithm will be presented. In Section 3 and 4, we evaluate and conclude the proposed algorithm, respectively.

2. The proposed rate control algorithm

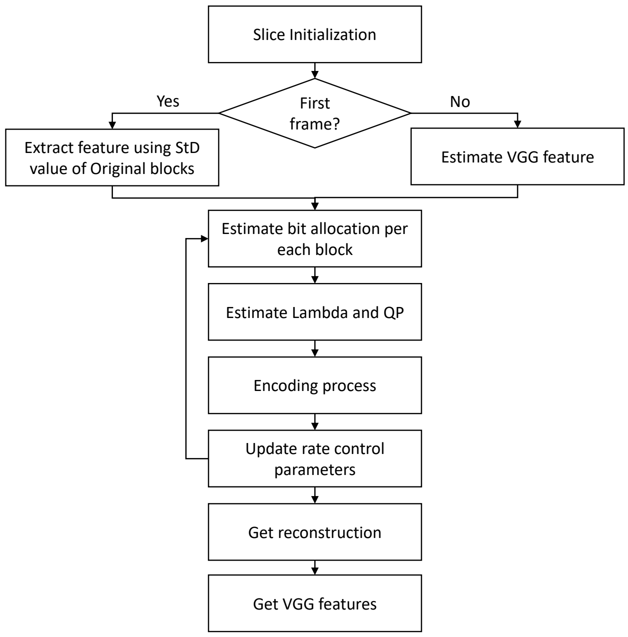

Figure 1: Flowchart of the proposed algorithm

The proposed algorithm is applied for handling bit budgets in intra frames. Controlling bit for intra frames is important for video coding system due to two main reasons. First, intra frames usually spends more bits than those in inter frames. Second, the quality of intra frames will significantly affect both coding efficiency and visual quality of the proceeding inter frames since inter prediction process. Accordingly, it is necessary to always consider the quality of intra pictures by controlling their bit budgeting.

We first extracted our spatial feature by taking original CTU block characteristics. Standard deviation value of each original CTU block is considered as spatial feature for only the first intra picture in a sequence. For the rest of intra frames, we estimate feature of the VGG-16 predefined model of previously coded frame. We consider this VGG-16 feature as temporal feature. Then, we estimate bit allocation in each CTU, followed by determining lambda and QP values. All these operations are processed with the extracted features involved. Note that our VGG feature is derived by calculating the Euclidean distance of two different feature maps from original and reconstruction blocks. Hence, the VGG feature is extracted after a frame is reconstructed. Flow of our proposed algorithm is illustrated in Figure 1.

3. Experimental results

In comparison against rate control model in HM-16.20, the proposed algorithm was tested using sequences in Class D based on the common test conditions (CTC) for HEVC. We coded 100 frames of each test sequences with all QP values 22, 27, 32, and 37 under All Main Intra configuration. The algorithm was evaluated on Windows 10 (64-bits) OS over 3.00GHz Intel (R) Core (TM) i7-5960X CPU with 32GB RAM. Table 1 shows RD performance and Table 2 shows bit accuracy between the proposed algorithm over the HM-16.20. In our experiment, our algorithm yields coding gain about -2.04%, -3.96%, and -4.75% in average for Luma and both chroma components, respectively, with some bitrate accuracy loss about 0.41% in average over the anchor.

Table 1: Overall BD-rate performance

No	Sequence	BD-BR		
		Y	U	V
1	BasketballPass	-1.50%	-1.14%	-2.55%
2	BQSquare	-2.44%	-1.98%	-3.33%
3	BlowingBubbles	-2.36%	-3.50%	-2.65%
4	RaceHorses	-1.87%	-9.20%	-10.48%
Average		**-2.04%**	**-3.96%**	**-4.75%**

Table 2: Overall bit accuracy performance

No	QP	Target bit (kbps)	HM-16.20 (kbps)	Prop. (kbps)	Accuracy		Diff. Acc
					HM-16.20	Prop.	
1	22	1504.98	1510.82	1511.62	0.39%	0.44%	0.05%
	27	753.14	758.93	759.65	0.79%	0.88%	0.10%
	32	371.68	377.87	375.54	1.58%	0.95%	-0.63%
	37	193.82	200.04	209.55	3.12%	8.02%	4.90%
2	22	1622.51	1629.79	1629.55	0.42%	0.40%	-0.01%
	27	631.19	638.32	638.72	1.16%	1.22%	0.06%
	32	283.17	290.37	295.08	2.60%	4.27%	1.67%
	37	139.29	280.63	281.86	101.89%	102.78%	0.89%
3	22	1639.64	1645.70	1645.23	0.35%	0.32%	-0.03%
	27	752.66	758.44	759.14	0.72%	0.82%	0.09%
	32	351.46	357.66	356.89	1.90%	1.68%	-0.22%
	37	163.58	169.61	169.16	3.42%	3.15%	-0.27%
4	22	1196.26	1200.00	1199.74	0.33%	0.31%	-0.02%
	27	587.09	590.46	590.13	0.59%	0.53%	-0.06%
	32	285.25	288.74	288.76	1.31%	1.32%	0.01%
	37	141.01	144.57	144.70	2.53%	2.62%	0.09%
Average					**7.69%**	**8.11%**	**0.41%**

Acknowledgement

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program(IITP-2019-2016-0-00288) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2018R1A2B2008238)

References

[1] Sullivan, G.J., Ohm, J.R.: Recent developments in standardization of high efficiency video coding (HEVC). In: Proc. 33rd SPIE Appl. Digital Image Process. 7798 (2010)

[2] I. Marzuki and D. Sim, "Overview of the Potential Technologies for Future Video Coding Standards in JEM: Status and Review," IEIE Transactions on Smart Processing and Computing, 2018

[3] B. Li, H. Li, L. Li, J. Zhang (Lambda) domain rate control algorithm for high efficiency video coding," IEEE Trans. Image Process., 23 (9) (2014), pp. 3841- 3854.

[4] H. Choi, J. Yoo, J. Nam, D. Sim, and I. V. Bajic, "Pixel-wise unified rate-quantization model for multilevel rate control," IEEE J. Sel. Topics Signal Process., vol. 7, no. 6, pp. 1112-1123, 2013.

[5] I. Marzuki, Y. Ahn, and D. Sim, "Tile-level Rate Control for HEVC Video Encoders on Multi-core Platforms," Journal of Real-Time Image Processing, DOI 10.1007/s11554-017-0720-5 (2017)

[6] Y. Li, B. Li, D. Liu, and Z. Chen, "A Convolutional Neural Network-Based Approach to Rate Control in HEVC Intra Coding," in IEEE Visual Communications and Image Processing (VCIP) (2017)

[7] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," [Online]. Available: https://arxiv.org/abs/1409.1556 (2014)