Fast CU Size Decision for HEVC Intra Coding by Using Local Characteristics and RD Costs

Dokyung Lee and Jechang Jeong
Department of Electronics and Computer Engineering, Hanyang University, Seoul, Republic of Korea
Email: dky1006@gmail.com, jjeong@hanyang.ac.kr

***Abstract*—The newest video coding compression standard is High–Efficiency Video Coding (HEVC) which supports high-resolution video sequences. Also, HEVC has about 50% bitrate saving compared with the former standard H.264/AVC. The quad-tree-based Coding Unit (CU) partitioning process is one of the most efficient technologies in HEVC. Its computational complexity, however, becomes main issue when attempting implementation with an encoder. A fast CU size decision algorithm for the intra-predicted picture of HEVC is proposed in this paper. We utilize the image complexity and RD cost values for early split decision. Also, the Bayesian decision rule are used for early termination of the CU partitioning process. Our experimental results show that the proposed algorithm significantly reduces the encoding time by about 55.7% with small BD-BR loss (1.56%) compared to the HEVC reference software HM 16.0.**

***Index Terms*—High-Efficiency Video Coding (HEVC), fast coding unit size decision, Bayesian decision rule, video compression**

I. Introduction

The latest video coding standard, called High Efficiency Video Coding (HEVC) [1], is developed by the Joint Collaborative Team on Video Coding (JCT-VC) established by ISO/IEC MPEG and ITU-T VCEG. HEVC achieves great coding improvements compared with the former video coding standard, H.264/AVC [2]. Especially, HEVC can efficiently compress video sequences has various resolutions (4K, 2K, 1080p, 720p, and etc.) by adopting novel techniques: quad-tree based block partitioning, 33 intra prediction directions, Discrete Cosine Transform (DCT) based interpolation filter for Motion Compensation (MC), sampled adaptive offset (SAO), and Discrete Sine Transform (DST).

The main difference between HEVC and H.264.AVC is a basic block size. The macro block (MB) is a basic unit of encoding and decoding process in H.264/AVC and it has fixed block size (16×16). Since the target video sequences of H.264/AVC is smaller than HEVC, it can be enough to compress video contents. However, as the resolution of video contents is much bigger than before, a basic unit of HEVC become larger and also even smaller than H.264/AVC. The Coding Unit (CU) which can have flexible block sizes from (64×64 to 8×8) is proposed for HEVC. The optimal block size is determined by quad-tree based block partitioning process based on Rate-Distortion (RD) cost, as shown in Fig. 1, and this is the most effective algorithm of HEVC in point of coding gain. The CU partitioning result is described in Fig. 2. We can observe that the CU size in flag region is bigger than complex region.

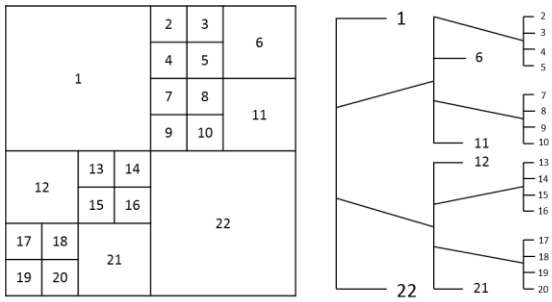

Figure 1. CU partitioning process

Figure 2. CU partitioning result

The intra prediction of HEVC have 35 prediction modes (planar, DC and 33 angular modes) and also includes mode-dependent smoothing filter, reference sample padding algorithm, and Rough Mode Decision (RMD). The whole procedure of intra prediction is explained as follows:

1) The reference sample preparation: For intra prediction, reference samples nearby current block are needed. If they are not reconstructed yet, the reference sample padding should be applied. Also, since quantization errors in reference samples are not removed, the mode-dependent smoothing filter is applied for better coding performance.

Manuscript received October 19, 2016; revised March 24, 2017

doi: 10.18178/ijsps.5.1.7-11

2) The RMD process: In RMD, the candidates for full RD calculation is determined (3, 3, 8, and 8 candidates for 64×64, 32×32, 16×16, 8×8, and 4×4 PUs, respectively) based on the following equation,

$$J_{SATD} = SATD + \lambda_{pred} \cdot B_{mode} \quad (1)$$

where *SATD* is the sum of the absolute Hadamard-transformed differences, λ_{pred} denotes the Lagrange multiplier for RMD and B_{mode} specifies the bit cost of prediction mode. The candidate modes which has the smallest J_{SATD} are chosen for the final best mode.

3) Full RD cost calculation: The final best mode of intra prediction is chosen among candidates determined through RMD. The full RD cost is calculated by

$$J_{Full} = SSE + \lambda \cdot B \quad (2)$$

here, SSE denotes a sum of squared error between reconstructed and original signal. λ represents the Lagrange multiplier for full RD cost and *B* is encoded bit cost of prediction mode, residual signal and other syntax elements.

Even the coding efficiency of the intra prediction is improved a lot compared with H.264/AVC, the encoding time also increase dramatically since the full RD costs should be calculated about all prediction modes and block sizes. Therefore, in recent years, many CU size decision algorithms for HEVC intra coding are proposed to reduce computational complexity [3]-[10]. L. Shen *et al.* [3] proposed CU size decision method using motion homogeneity and RD cost. Also, they predict depth level of current CU from neighboring CU and co-located CU for depth range determination. The statistical fast cu size decision based on Bayesian decision is proposed by S. Cho and M. Kim [4]. They utilize statistical characteristics of RD costs (J_{SATD} and J_{Full}) obtained by online update phase. However, J_{SATD} is not accurate value to determine early split and the image complexity is critical factor of CU partitioning process. In [8], C. Tseng and Y. Lai introduce fast coding unit decision algorithm using the standard deviation of CU and RD costs. The standard deviation in high resolution video sequences can have difference statistics according to local complexity.

Therefore, in this paper, the fast CU size decision algorithm is proposed by using the local complexity and RD cost. First, in online update phase, the statistical data is obtained and the proposed algorithm calculate thresholds for fast CU size decision. We utilize the local edge magnitude difference to measure complexity based on the Sobel operator for early split decision. Then, to determine early termination, the combined cost of the edge magnitude difference and the full RD cost (J_{Full}) is calculated. The Bayesian decision rule is employed for threshold.

The remainder of this paper is organized as follows. In Section II, we propose early spit decision and early termination decision for CUs. The experimental results of proposed algorithm is shown in Section III. Finally, we conclude our works in Section IV.

II. Proposed Algorithm

A. Early Splitting Decision (ESD)

In proposed Early Splitting Decision (ESD), we employ edge magnitude difference obtained by the Sobel operator as the measure of the local complexity. First, the magnitude of the gradient vector $\mathbf{D}(i, j)_k$ at depth level *k* is calculated for a pixel $P_{i,j}$, which can be represented by

$$M_k = \left|\mathbf{D}(i, j)_k\right| = \sqrt{(G_x(i, j)^2 + G_y(i, j)^2)} \quad (3)$$

where,

$$\begin{aligned} G_x(i, j) = &P_{i+1,j+1} + 2 \cdot P_{i+1,j} + P_{i+1,j-1} \\ &- P_{i-1,j+1} - 2 \cdot P_{i-1,j} - P_{i-1,j-1}, \textit{ and} \end{aligned} \quad (4)$$

$$\begin{aligned} G_y(i, j) = &P_{i-1,j-1} + 2 \cdot P_{i,j-1} + P_{i+1,j-1} \\ &- P_{i-1,j+1} - 2 \cdot P_{i,j+1} - P_{i+1,j+1} \end{aligned} \quad (5)$$

As mentioned before, M_k itself may not be efficient value to ESD since its average can be relatively high in complex area. Therefore, we define the local complexity defined as, into early split CU class and immediately move to the encoding process of the next depth.

$$\varepsilon = \frac{1}{4}\sum_{i=0}^{3}\left|M_k - M_{k+1}^i\right| \quad (6)$$

where M_{k+1}^i denotes the gradient magnitude of *i*-th sub-CU, as described in Fig. 3. It is critical to choose appropriate feature to solve problems for online learning based algorithm. Using the relative complexity between CU and 4 sub-CUs, we can more accurately classify early-split CUs.

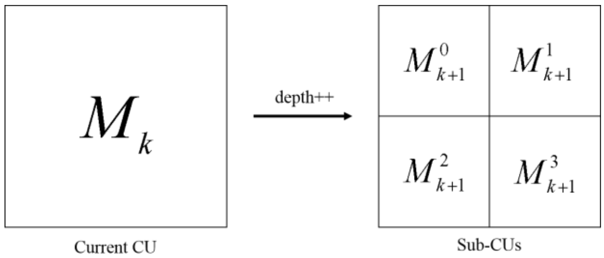

Figure 3. CU partitioning structure

In online update phase, ε values of split CU and un-split CU is stored. The encoder calculates T_S using the statistics of stored ε values and determine early split CU by compared with its ε value. First, the accumulated number of CUs, H_k^C is defined as follows,

$$H_k^i(\varepsilon) = \sum_{C}^{\infty} h_k^i(\varepsilon),\ i \in \{S, U\} \quad (7)$$

here, $h_k^i(\varepsilon)$ represents the number of *i*-class CUs of depth *k* with the local complexity, ε . *S* and *U* denote split CU class and un-split class, respectively. The proposed algorithm calculate the probability $P_S(\varepsilon)$ as

$$P_S(\varepsilon) = \frac{H_k^S(\varepsilon)}{H_k^S(\varepsilon) + H_k^U(\varepsilon)} \quad (8)$$

The threshold T_S is set to the largest ε when $P_S(\varepsilon)$ is lager or equal to 0.9. If the local complexity (ε) of the current CU is larger than T_S, the current CU is classified

B. Early Termination Decision (ETD)

When a current CU is not determined to early-split, the Early Termination Decision (ETD) with Bayesian decision rule is performed. In conventional algorithm ([4], [7], and [9]), they apply the Bayesian decision for fast CU size decision using full RD cost. The local complexity of image, however, should be considered because image can have locally high RD costs. In case Kimono sequence, RD cost values in back ground is much higher than object (woman) in the image. The ETD based on only RD costs cannot classify accurately in this situation. Therefore, we propose a combined cost for the Bayesian decision, as follows

$$J_C = J_{Full} + \frac{\mu_J}{\mu_J + \mu_\varepsilon} \cdot \varepsilon \tag{9}$$

where μ_J and μ_ε are mean of stored J_{Full} and ε in online update phase, respectively. In online update phase, the encoder collect J_C of split class and un-split class and calculate the threshold for ETD, which is defined based on the Bayesian decision rule.

The classification problem of proposed ETD is defined as binary classification (split class: S and un-split class: U). The posteriori $P(i\,|\,x)$ for the Bayesian decision rule when x is J_C is given as

$$P(i\,|\,x) = \frac{p(x\,|\,i)\cdot P(i)}{p(x)},\ i \in \{S, U\} \tag{10}$$

here, $p(x\,|\,i)$ is called likelihood, $P(i)$ denotes the class prior and $p(x)$ is the evidence. Also, the probability density function (PDF) of J_C can be assumed to be the Gaussian distribution based on our experiment. Therefore, the likelihood, $p(x\,|\,i)$ can be represented by

$$p(J_C\,|\,i) = \frac{1}{(2\pi)^{1/2}\delta_i}\exp\left\{-\frac{(J_C-\mu_i)^2}{2\delta_i^{\,2}}\right\}, i \in \{S, U\} \tag{11}$$

where δ_i denote a standard deviation of J_C, μ_i is mean value of J_C. As shown in Fig. 4, the threshold (T_U) of ETD is determined based on the minimum risk error. The error rate (α) is defined by

$$\alpha = \int_{-\infty}^{T_U} p(x\,|\,U)\cdot P(U)dx \tag{12}$$

The threshold T_U can be obtained by a cumulative standard normal distribution table with $P(U) = N_S / N_T$, where N_S and N_T are the numbers of split CUs and total CUs, respectively. If the combined cost J_C of the current block is smaller than T_U, the encoding process of next depth is terminated.

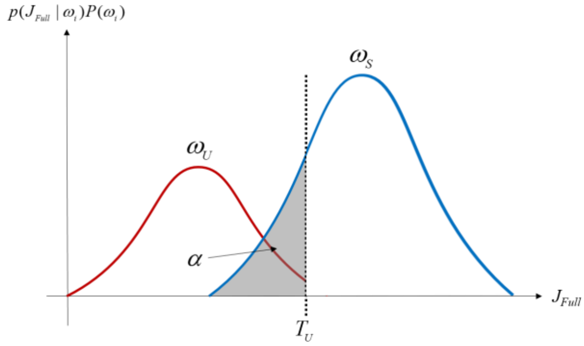

Figure 4. The Bayesian decision example for early termination decision

III. Experimental Results

In order to evaluate the performance of the proposed algorithm, we implement them on the reference software HM-16.0 [11]. The test sequences and experimental configuration of the reference software are specified by JCT-VC [12]. For our algorithm, the all intra (AI) configuration is used. The CTU size is 64×64 (depth level is 0) and the maximum depth level for CU partitioning is 3 (8×8). The error rate α is set to 0.1 and the online update phase is a first frame of each second. There is tradeoff between time saving and coding performance by the error rate. The proposed algorithm is compared with conventional algorithms in terms of the Bjontegaard Delta bit rate (BD-BR) and PSNR (BD-PSNR) of luminance component [13] when quantization parameter (QP) is set to 22, 27, 32, and 37. The encoding time saving is calculated by

$$\Delta Time = \frac{T_{HM16.0} - T_{Proposed}}{T_{HM16.0}} \tag{13}$$

where $T_{HM16.0}$ and $T_{Proposed}$ denote the encoding time of HM-16.0 and the proposed algorithm, respectively.

The experimental results is shown in Table I in terms of BD-BR, BD-PSNR, and $\triangle Time$. The proposed algorithm is averagely 55.7% faster than HM-16.0 with small coding loss (about 1.56% BD-BR loss). The local complexity of *NebutaFestival* sequence has spatial consistency, so the proposed algorithm get great results of this sequence.

In Table II, the comparison results are listed with other relative algorithms. The proposed algorithm has similar coding performance with [3], but the encoding time is much faster. These experimental results imply that the proposed algorithm is efficient and accurate method to fast CU size decision.

TABLE I. Experimental Results in Terms of BD-BR (%), BD-PSNR (dB), and Δ Time (%) for the Proposed Algorithm When A Is Set to 0.1

Class	Test sequences	Proposed algorithm		
		BD-rate (%)	BD-PSNR (dB)	ΔTime (%)
Class A 2560x1600	NebutaFestival	0.433	-0.031	68.8
	PeopleOnStreet	1.704	-0.097	56.2

	SteamLocomotiveTrain	1.178	-0.034	59.4
	Traffic	1.673	-0.090	51.4
Class B 1920x1080	BasketballDrive	1.313	-0.039	59.6
	BQTerrace	1.178	-0.059	60.8
	Cactus	1.511	-0.055	54.2
	Kimono	1.259	-0.044	57.7
	ParkScene	1.121	-0.050	48.7
Class C WVGA	BasketballDrill	1.288	-0.063	47.4
	BQMall	1.829	-0.102	50.7
	PartyScene	1.110	-0.081	44.7
	RaceHorses	1.586	-0.090	54.1
Class D WQVGA	BasketballPass	1.550	-0.094	59.9
	BlowingBubbles	1.172	-0.080	43.3
	BQ square	1.732	-0.139	48.4
	RaceHorses	1.646	-0.101	50.6
Class E 720p	FourPeople	2.086	-0.120	57.1
	Johnny	2.408	-0.101	66.5
	KristenAndSara	2.385	-0.122	67.0
Average		**1.559**	**-0.083**	**55.7**

TABLE II. Comparison with Other Fast CU Decision Algorithms

Algorithms	BD-BR (%)	∆Time (%)
Proposed algorithm	**1.56**	**55.7**
[3]	1.49	42.0
[4] ($\alpha = 0.2$)	1.20	51.4
[6]	0.82	50.1
[10]	1.08	47.0

IV. Conclusion

In this paper, we proposed a fast CU decision algorithm for intra coding in HEVC using statistical data related to the local complexity and RD costs. The proposed algorithm includes two main ideas (ESD and ETD). The edge magnitude of the CU are employed for the ESM, and early-terminating CUs are chosen by analyzing the combined cost based on the Bayesian decision rule. The experimental results show that the proposed algorithm successfully reduces the encoding time with negligible coding efficiency degradation compared with HM software version 16.0 in the AI configuration.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and future Planning (NRF-2015R1A2A2A01006004).

References

[1] 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.

[2] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[3] L. Shen, Z. Liu, X. Zhang, W. Zhao, and Z. Zhang, "An effective cu size decision method for HEVC encoders," *IEEE Trans. Multimedia*, vol. 15, no. 2, pp. 465–470, Feb. 2013.

[4] S. Cho and M. Kim, "Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 9, pp. 1555-1564, Sep. 2013.

[5] L. Shen, Z. Zhang, and P. An, "Fast CU size decision and mode decision algorithm for HEVC intra coding," *IEEE Trans. Consum. Electron.*, vol. 59, no. 1, pp. 207–213, Feb. 2013.

[6] B. Min and R. C. C. Cheung, "A fast CU size decision algorithm for the HEVC intra encoder," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 5, pp. 892-896, May 2015.

[7] K. Lim, J. Lee, S. Kim, and S. Lee, "Fast PU skip and split termination algorithm for HEVC intra prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 8, pp. 1335–1346, Aug. 2015.

[8] C. Tseng and Y. Lai, "Fast coding unit decision and mode selection for intra-frame coding in high-efficiency video coding," *IET Image Processing*, vol. 10, no. 3, pp. 216-221, Mar. 2016.

[9] H. S. Kim and R. H. Park, "Fast CU partitioning algorithm for HEVC using online learning based Bayesian decision rule," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, Jan. 2016.

[10] L. Shen, Z. Zhang, and Z. Liu, "Effective CU size decision for HEVC intra coding," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4232-4241, Oct. 2014

[11] HM 16.0 Reference Software. [Online]. Available: http://hevc.kw.bbc.co.uk/trac/browser/tags/HM-16.0

[12] F. Bossen, "Common test conditions and software reference configurations," Document JCTVC-K1100, Shanghai, China, Oct. 2012.

[13] G. Bjøntegaard, "Calculation of Average PSNR Differences between RD curves," Document VCEG-M33, Austin, TX, USA, Apr. 2001.

Dokyung Lee received a B.S. from the Department of Electrical and Electronics Engineering, Chung-ang University, Seoul, Korea, in 2010. He is currently working toward a Ph.D. degree in Electronic and Computer Engineering at Hanyang University, Seoul, Korea.
His research interests include image processing, H.264/AVC, high-efficiency video coding standards, pattern recognition, and machine learning.

Jechang Jeong received a B.S. degree in Electronic Engineering from Seoul National University, Korea, in 1980; an M.S. degree in Electrical Engineering from the Korea Advanced Institute of Science and Technology in 1982; and a Ph.D. degree in Electrical Engineering from the University of Michigan, Ann Arbor, in 1990. From 1982 to 1986, he was with the Korean Broadcasting System, where he helped develop teletext systems. From 1990 to 1991, he worked as a postdoctoral research associate at the University of Michigan, Ann Arbor, where he helped to develop various signal-processing algorithms. From 1991 through 1995, Dr. Jeong was with the Samsung Electronics Company, Korea, where he was involved in the development of HDTV, digital broadcasting receivers, and other multimedia systems. Since 1995, he has conducted research at Hanyang University, Seoul, Korea. His research interests include digital signal processing, digital communication, and image and audio compression for HDTV and multimedia applications. He has published numerous technical papers. Dr. Jeong received the Scientist of the Month Award in 1998 from the Ministry of Science and Technology of Korea, and he was the recipient of the 2007 IEEE Chester Sall Award and the 2008 ETRI Journal Paper Award. He was also honored with a government commendation in 1998 from the Ministry of Information and Communication of Korea.