LETTER

Fast Coding Unit Size Decision in HEVC Intra Coding

Tao FAN[†a)], ***Member***, **Guozhong WANG**[†], ***and*** **Xiwu SHANG**[†], ***Nonmembers***

SUMMARY The current high efficiency video coding (HEVC) standard is developed to achieve greatly improved compression performance compared with the previous coding standard H.264/AVC. It adopts a quadtree based picture partition structure to flexibility signal various texture characteristics of images. However, this results in a dramatic increase in computational complexity, which obstructs HEVC in real-time application. To alleviate this problem, we propose a fast coding unit (CU) size decision algorithm in HEVC intra coding based on consideration of the depth level of neighboring CUs, distribution of rate distortion (RD) value and distribution of residual data. Experimental results demonstrate that the proposed algorithm can achieve up to 60% time reduction with negligible RD performance loss.
key words: *HEVC, CU, size decision, intra coding*

1. Introduction

The new high efficiency video coding (HEVC) standard developed by the joint collaborative team on video coding (JCT-VC) provides a bit rate reduction of 50% with the same subjective quality compared with the former coding standard H.264. It adopts a flexible quadtree structure in which the largest coding unit (LCU) is 64×64 and the smallest coding unit (SCU) is 8×8. Each CU is recursively split into four equal-sized sub-CUs until SCU is reached. These various partition sizes can flexibly predict the objects with different shapes, which is totally different from the term marcroblock utilized in HEVC [1]. Prediction Unit (PU) is the basic unit utilized for prediction existing in each CU. It has two types of partition size $2N\times2N$ and $N\times N$ (only under 8×8 CU) [2]. Therefore, there are five depth levels of PU in HEVC intra coding ranging from 64×64 to 4×4. In PU, the number of prediction modes is 35. During the process of intra prediction, all CU sizes and prediction modes are exhausted to find the one with the minimum rate distortion (RD) cost. As a result, the computational complexity is much higher than H.264.

Currently, a number of fast algorithms are presented to speed up intra prediction process in HEVC encoder. Some methods mainly focus on reducing the CU depth levels [3]–[7]. Tian et al. [3] analyze the texture complexity of downsampled LCU and its four sub-CUs measured by variance to skip unnecessary CU levels. Shen et al. [4], [5] utilize the coding information from neighboring coded PUs to make an early CU size decision. Lee et al. [6] reduce the depth range according to the depth level of temporally co-located CU. In this algorithm, the encoding time reduction is limited, because at most one depth is eliminated during the coding process. Mu et al. [7] model the process of CU splitting as a binary decision problem and resolve it by support vector machine (SVM) which utilizes mean square error (MSE) and number of coding bits (NEB) as the selected features to early terminate the CU splitting process. Other algorithms devote to reducing the number of prediction modes [8]–[12]. Kim et al. [8] propose a hierarchical mode decision algorithm, which consists of two stages: candidate update stage (CUS) and decision stage (DS). In CUS, it selects modes which can cover the probable area to lower the computational complexity. In DS, only the RD cost of two best candidate modes is calculated to find the one with the minimum cost as the optimal mode. Gradient directions with Sobel operator are calculated to generate gradient-mode histogram for each CU [9], [10]. Then, according to the distribution of the histogram, only a small candidate mode set is selected for the rough mode decision and rate distortion optimization process. Ma et al. [11] select eleven candidate modes for rough mode decision and obtain the best six modes based on sum of absolute transformed differences (SATD) costs. Then two-spaced adjacent modes of the six best candidates are tried to select the best three modes. Finally, the one-spaced modes plus the most probable mode (MPM) and DC mode are compared with RD costs to choose the optimal one.

In this paper, we propose a fast CU size decision algorithm to speed up the coding process while maintains the coding performance. First, we analyze the CU depth correlation among neighboring CUs to skip unnecessary depth levels of the current CU. Then, according to the distribution of RD cost from the training frame, the prediction process is bypassed early when the depth level of CU with the lowest RD cost meets certain conditions. Finally, we analyze the distributions of residual data under horizontal and vertical partitions. If there is no significant difference between the two distributions, the current CU is no longer split.

The rest of paper is organized as follows. Section 2 presents the fast intra coding algorithm. Section 3 shows the simulation results. Conclusion is given in Sect. 4.

Manuscript received November 10, 2015.
Manuscript revised March 1, 2016.
Manuscript publicized April 15, 2016.
[†]The authors are with School of Communication and Information Engineering, Shanghai University, No.333 Nanchen Road, Baoshan District, Shanghai, 200444 China.
a) E-mail: fantao19830627@163.com
DOI: 10.1587/transinf.2015EDL8231

2. Proposed Scheme

2.1 Content Based CU Size Decision

In HEVC encoder, four types of CU are supported from 64×64 to 8×8. For SCU, it can also be split into four equal size blocks. Thus, five coding blocks from 64×64 to 4×4 exist in intra prediction process. In this paper, 4×4 block is regarded as a CU whose depth level is four for simplicity. In video coding, neighboring CUs usually have similar texture features and the depth level of current CU possesses high correlation with its neighboring CUs. If the depth level of CU (edge CU excluded) differs by two or more from that of neighboring CUs including left CU, above CU, and left above CU simultaneously, it is defined as discontinuous unit (DU). Figure 1 illustrates the final partition result under the original encoder. As we can see that the depth level cannot vary dramatically. Only one CU marked by green circle is DU as shown in Fig. 1. In fact, the coding content in CUs is closely related with nearby CUs and usually cannot isolate with neighboring CUs at the same time. Thus, the ratio of DU's occurrence is low.

To verify this assumption, extensive experiments have been done on six typical sequences with different resolutions and texture information. The test complexity in "BQmall" and "PartyScene" is high, while in "Basketballdrive" and "SlideEditting" is medium, and in "Vidyo1" and "Flowervase" is low. Simulation conditions are as follows: all intra (AI) mode is configured with quantization parameters (QPs) 20, 30, 40. Through the experiments performed on the original encoder of HEVC, we analyze DU's distribution in different test sequences. Table 1 illustrates the percentage of DU in five depth levels respectively. As we can see from Table 1, the percentage of DU is small with the maximum percentage 3.80% DU in depth 0 and the minimum percentage 0.02% DU in depth 2 on average. During the process of video coding, when the current CU is detected as DU, the intra prediction can be skipped early. The criteria is

$$\begin{cases} depth_i - depth_{cur} \geq 2, \text{ skip coding the current CU} \\ depth_{cur} - depth_i \geq 2, \text{ stop splitting the current CU} \end{cases} \quad (1)$$

where $depth_i$ is the depth level of neighboring CU (i = $left, above \text{ or } left - above$), and $depth_{cur}$ is the depth of the current CU. If the difference of the depth of neighboring CU minus that of the current CU is equal to or more than two, skip the prediction on the current depth and split the CU into four sub-CUs directly. Otherwise, if the difference of the depth of current CU minus that of the neighboring CU is equal to or more than two, stop splitting the current CU into four sub-CUs. In Fig. 1, for the CU marked by green rectangle, the depths of its neighboring CUs are equal to 3. According to Eq. (1), the intra prediction on its up layer CU whose depth is equal to 1 is bypassed. Similarity, for the CU marked by parallelogram, the depths of its neighboring CUs are equal to 1. Its down layer CU whose depth is 3 is performed to stop splitting early.

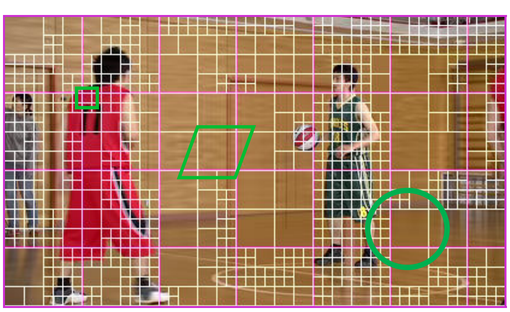


Fig. 1 CU partition result in "BasketballPass".

2.2 RD Cost Based CU Size Decision

During the process of CU size selection, the RD cost of the current CU is compared with that of its four sub-CUs. When the RD cost is smaller enough, the splitting process can be terminated early. In our algorithm, we obtain the threshold on line. In the training frame, if the RD cost of the current CU in depth D is less than the sum of the RD cost of its sub-CUs, we calculate the sum of the RD cost value of the current CU $RD\cos t_D$ and count the number Num_D. Then the average of the RD cost is computed as follows

$$Avg_D = \frac{RD\cos t_D}{Num_D} \quad (2)$$

By threshold Th_D, unnecessary CU size is bypassed in advance. Th_D is defined as

$$Th_D = \beta \cdot Avg_D \quad (3)$$

where β is an adjust parameter and is set to 0.9 in our experiment for having good tradeoff between coding performance and time saving. Since temporally adjacent frames have similar contents, which results in a similar RD cost distribution. Therefore in the non-training frames, when the RD cost of the current CU in depth D is less than Th_D, which means the current RD cost is possibly less than the sum of the sub-CUs' RD cost, the current CU is no longer to be partitioned into four sub-CUs. To cope with different features of test sequences, threshold Th_D is updated in every

Table 1 Distribution of DUs

Sequence	64x64 depth 0	32x32 depth 1	16x16 depth 2	8x8 depth 3	4x4 depth 4
BQMall	4.94%	4.42%	0.02%	0.00%	0.76%
Flowervase	1.02%	0.48%	0.02%	0.15%	3.59%
SlideEditing	0.30%	2.81%	0.00%	0.01%	0.23%
vidyo1	7.59%	1.64%	0.01%	0.09%	2.16%
BaskeballDrive	4.68%	3.55%	0.03%	0.18%	3.08%
ParkScene	4.29%	3.74%	0.01%	0.09%	3.14%
Average	3.80%	2.77%	0.02%	0.09%	2.16%

one second.

2.3 Distribution of Residual Data Based CU Size Decision

By analyzing the distribution of residual coefficients, we aim to decide whether the current CU should be split or not. Figure 2 demonstrates the probability of residual coefficients in different CU depth levels for “BasketballDrive”, and other sequences have similar results. It can be seen in Fig. 2 that residual coefficients satisfy Gaussian distribution. To decide whether the current CU should be split, we first split the residual coefficients of the current CU into two parts in horizontal and vertical direction respectively, as shown in Fig. 3. Then check whether the means of the two partition parts (A and B or C and D) are similar with each other. When there is no significant difference in mean between block A and B, and block C and D, it shows that the current CU size can well represent the prediction content and the early termination of CU size decision is enabled.

Since residual data meets Gaussian distribution, we exploit t examination to detect whether the two partitions have significant difference. The hypothesis is established as following:

$$H_0 : \mu_1 = \mu_2; H_1 : \mu_1 \neq \mu_2 \tag{4}$$

where μ_1 and μ_2 are the true mean of residual data in block A and block B or in block C and block D. Hypothesis H_0 denotes the residual coefficients in two partitions are from the same distribution, while H_1 denotes they are from different distribution. When H_0 is accepted, the current CU size is decided as the optimal size early.

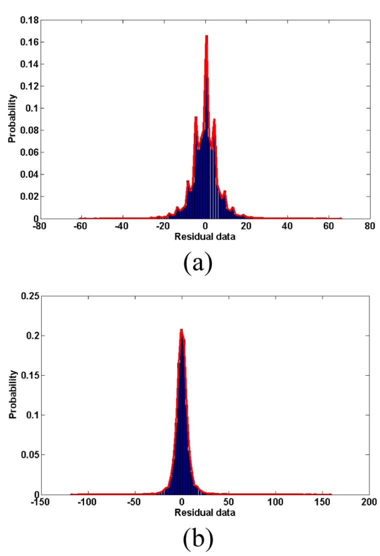


Fig. 2 Probability of residual data for “BasketballDrive” in different depth levels, (a) depth = 2 (16 × 16 CU), (b) depth = 1 (32 × 32 CU).

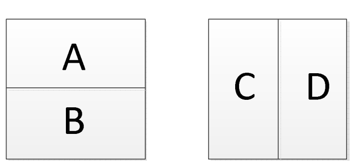


Fig. 3 Illustration of residual coefficients split in horizontal and vertical direction respectively.

The detection statistics satisfied t distribution, which is defined as:

$$t = \frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2) \tag{5}$$

where $\overline{X}$ and $\overline{Y}$ are the sample means of block A and B or block C and D. n_1 and n_2 are the number of pixels in two partitions respectively. S_w is calculated as:

$$S_w = \sqrt{\frac{\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_1}(Y_i - \overline{Y})^2}{n_1 + n_2 - 2}} \tag{6}$$

X_i and Y_i is the pixel value in block A and B or block C and D.

In this paper, we define the confident interval α as 0.05. If $|t| < t_{\alpha/2}$ for both block A, B and block C, D, hypothesis H_0 is accepted.

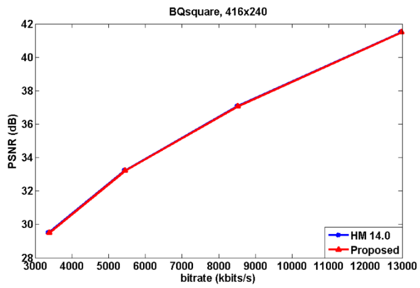
3. Experimental Results

To verify the performance of the proposed algorithm, we implement the algorithm on HEVC reference encoder HM 14.0 under all-intra mode configuration. The test platform is that: Windows 7 operation system, Intel core i5 of 3.10 GHz with 4 GB memory. Simulations are implemented under QP = 22, 27, 32, 37 with six different resolution sequences. Coding performance is measured by BDBR (%) and BDPSNR (dB) which denote the average change of bit rate and PSNR [12]. Coding time saving is calculated as

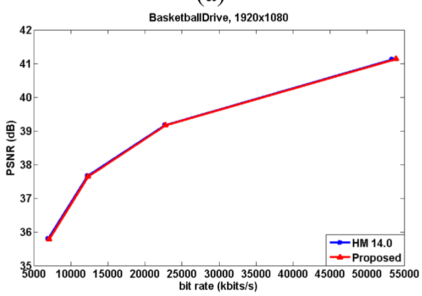
$$TS(\%) = \frac{T_{HM} - T_{proposed}}{T_{HM}} \times 100\% \tag{7}$$

Table 2 Performance comparison between Shen et al. [5] and the proposed algorithm.

		Shen [5]			Proposed		
Resolu-tion	Sequence	BDBR (%)	BDPSNR (dB)	TS (%)	BDBR (%)	BDPSNR (dB)	TS (%)
Class A	Traffic	0.97	-0.06	37	1.75	-0.09	52
	PeopleOn street	1.28	-0.07	41	1.54	-0.09	49
Class B	BasketballDrive	2.50	-0.06	61	1.05	-0.03	52
	BQterrace	0.62	-0.04	39	0.69	-0.04	46
	Kimono1	0.81	-0.03	38	0.97	-0.03	59
	Tennis	2.86	-0.09	57	1.61	-0.05	58
Class C	BasketballDrill	0.79	-0.04	29	1.38	-0.07	49
	Bqmall	1.16	-0.07	36	1.27	-0.08	46
	RaceHorses	0.55	-0.04	32	1.93	-0.12	52
	PartyScene	0.14	-0.01	24	0.50	-0.04	44
Class D	BasketballPass	1.41	-0.08	38	1.21	-0.07	49
	BQsquare	0.32	-0.03	23	0.33	-0.03	40
	BlowingBubbles	0.05	0.00	17	0.55	-0.03	40
	Keiba	0.73	-0.05	29	1.39	-0.09	47
Class E	FourPeople	2.19	-0.13	49	2.08	-0.12	54
	Johnny	3.88	-0.16	59	2.21	-0.09	60
	KristenAndSara	3.31	-0.17	59	1.99	-0.10	60
Class F	SlideEditting	4.32	-0.67	43	1.64	-0.26	48
	BasketballdrillText	1.39	-0.07	34	1.51	-0.08	47
	Chinaspeed	2.55	-0.23	47	1.73	-0.16	55
Average		1.59	-0.11	40	1.37	-0.08	50



(a)



(b)

Fig. 4 RD curve of HM 14.0 and the proposed algorithm, (a) "BQsquare", 416 × 240, (b) "BasketballDrive", 1920 × 1080.

where T_{HM} and $T_{proposed}$ is the encoding time under HM 14.0 encoder and the presented encoder, respectively.

Table 2 demonstrates the experimental results of the presented algorithm. From the table, we observe that the proposed algorithm achieves 50% time reduction with 0.08 BDPSNR loss and 1.37% BDBR increase. Generally, time saving in high resolution sequences is greater than that of low resolution sequences. The reason is that the spatial correlation of sequence with high resolution is higher than that of low resolution sequences. Figure 4 shows the RD curves of "BQsquare", "BasketballDrive", where the blue curves indicate the RD performance of original encoder HM 14.0, and the red curves indicate the RD performance of the proposed algorithm. It demonstrates that the RD curves of the proposed algorithm is almost similar with that of HM 14.0, which indicates our algorithm has nearly no RD performance loss. Besides, compared with Shen's method [5], the proposed algorithm achieves higher time reduction, as Shen's method provides 40% time reduction with 0.11 dB BDPSNR loss and 1.59% BDBR increase.

4. Conclusion

In this paper, we present a fast CU size decision algorithm for HEVC intra coding. It consists of three methods: (1) content based CU size decision, (2) RD cost based CU size decision, and (3) distribution of residual data based CU size decision. By utilizing the combination of spatial, temporal and statistical correlation, the proposed algorithm reduces the coding time by 50% with negligible RD performance loss. In addition, compared with Shen's algorithm [5], our algorithm obtains about an additional 10% time reduction while keeping a better RD performance.

Acknowledgments

This work was supported by National Science Foundation of China under Grant No. 61271212, and National High-tech R&D program of China (863 Program, 2015AA015903).

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, 2012.
[2] J. Lainema, F. Bossen, W.-J. Han, J. Min, and K. Ugur, "Intra Coding of the HEVC Standard," IEEE Trans. Circuits Syst. Video Technol., vol.22, no.12, pp.1792–1801, 2012.
[3] T.S. Kim, M.H. Sunwoo, and J.-G. Chung, "Hierarchical fast mode decision algorithm for intra prediction in HEVC," IEEE International Symposium on Circuits and System, pp.2792–2795, 2015.
[4] 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, 2013.
[5] L. Shen, Z. Zhang, and Z. Liu, "Effective CU size decision for HEVC intra coding," IEEE Trans. Image Processing, vol.23, no.10, pp.4232–4241, 2014.
[6] H. Lee, K. Kim, T. Kim, and G. Park, "Fast encoding algorithm based on depth of coding-unit for high efficiency video coding," Jounal of Optical Engineering, vol.51, no.6, 1–11, 2012.
[7] F. Mu, L. Song, X. Yang, and Z. Luo, "Fast coding unit depth decision for HEVC," IEEE International Conference on Multimedia and Expo Workshops (ICMEW), pp.1–6, 2014.
[8] T.S. Kim, M.H. Sunwoo, and J.-G. Chung, "Hierarchical fast mode decision algorithm for intra prediction in HEVC," IEEE International Symposium on Circuits and Systems (ISCAS), pp.2792–2795, 2015.
[9] W. Jiang, H. Ma, and Y. Chen, "Gradient based fast mode decision algorithm for intra prediction in HEVC," International Conference on Consumer Electronics, Communications and Networks (CECNet), pp.1836–1840, 2012.
[10] M. Jamali, S. Coulombe, and F. Caron, "Fast HEVC Intra Mode Decision Based on Edge Detection and SATD Costs Classification," Data Compression Conference (DCC), pp.43–52, 2015.
[11] H. Zhang and Z. Ma, "Fast intra mode decision for high efficiency video coding (HEVC)," IEEE Trans. Circuits Syst. Video Technol., vol.24, no.4, pp.660–668, 2014.
[12] G. Bjontegaard, "calculation of average PSNR difference between RD curves," 13th VCEG-M33 Meeting, Austin, TX, April 2001.