Fast Coding Unit Decision Algorithm for HEVC Intra Coding

Zhilong Zhu, Gang Xu, and Fengsui Wang
Anhui Key Laboratory of Detection Technology and Energy Saving Devices, Anhui Polytechnic University, Wuhu, Anhui Province, 241000, China
College of Electrical Engineering, Anhui Polytechnic University, Wuhu, Anhui Province, 241000, China
Email: {zhuzhilong919, xugang, wangfengsui}@ahpu.edu.cn

Abstract —High Efficiency Video Coding (HEVC) exploits quad-tree structured Coding Unit (CU) to improve compression efficiency. It saves about 50% coding bits as compared with the former standard H.264/AVC high profile. However, the computational complexity is dramatically increased for more partition blocks and coding modes supported in HEVC. In this paper, a fast CU decision algorithm is proposed to reduce the number of candidate CUs for HEVC intra coding, which is consisted of two algorithms: a spatial correlation based early CU decision algorithm (SECU) and a Rate Distortion (RD) cost based early CU decision algorithm (RDCU). The depths of spatially neighboring CUs are exploited to skip unnecessary CU size tests first. Then the distribution of RD cost is utilized on the selection of CU sizes. Experimental results demonstrate that the proposed algorithm can achieve the time saving by 46% on average as compared with original encoder with 0.92% BDBR increase and 0.05 dB BDPSNR decrease.

Index Terms—HEVC, coding unit, intra coding, computational complexity

I. Introduction

With the popularity of High Definition (HD) or Ultra-High Definition (UHD) televisions, the huge data generated by high resolution videos challenge the current limited bandwidth. Traditional coding standards cannot satisfy the demand of compression efficiency. High Efficiency Video Coding (HEVC) is a new generation of video coding standard exploited to compress high resolution videos. By adopting quad-tree based recursive splitting coding structure, the coding efficiency is improved by 50% as compared with the former coding standard H.264 [1]. Analogous to the macroblock in H.264, the Coding Unit (CU) in HEVC is ranged from 64×64 (depth 0) to 8×8 (depth 3) [2]. Flexible combinations of different CU sizes can well represent coding contents with different texture complexities. Fig. 1 shows an example of CU splitting procedure from 64×64 to 16×16 CU size. The number in the node denotes the coding order of CU. The CU in depth X is coded first, and then it is split into four equal-sized CUs named CUs in depth X+1. To achieve the optimal CU size, each depth of CUs will be tested using Rate Distortion (RD) cost, which is calculated as

$$RDcost = D + \lambda \cdot R \tag{1}$$

where D is the distortion between the original pixels and the reconstructed pixels. R and λ is the coding bits and the lagrange multiplier. The way of exhausting all type of CUs to find the optimal one definitely consumes much time [3], which obstructs HEVC from real-time applications.

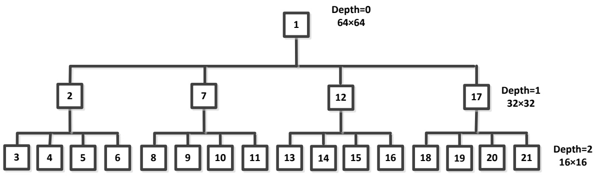


Fig. 1. Illustration of quad-tree based recursive splitting coding structure.

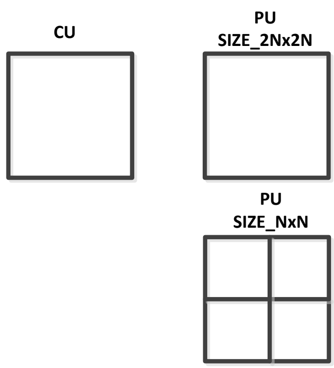


Fig. 2. PU types for intra coding in a CU.

As shown in Fig. 2, under intra coding, each depth of CU includes SIZE_2N×2N prediction unit (PU). SIZE_N×N PU is only available for 8×8 CU. PU is the basic unit carrying the prediction information (eg. prediction mode, residual). There are up to 35 prediction modes in HEVC including a planar mode, a DC mode and 33 angular prediction modes [4]. In order to reduce the coding time, a fast intra coding algorithm is adopted in HEVC test mode HM, which reduces the candidate modes by comparing the value of Hadmard cost [5]. The

Manuscript received May 11, 2016; revised October 18, 2016.
This work was supported by the Open Research Fund of AnHui Key Laboratory of Detection Technology and Energy Saving Devices, AnHui Polytechnic University, China (1506c085002), the Major Program of Scientific Research of Anhui Province, China (KJ2015A071), and the Scientific Research Foundation for Talents, Anhui Polytechnic University of China (2014YQQ006).
Corresponding author email: xugang@ahpu.edu.cn
doi:10.12720/jcm.11.10.950-956

©2016 Journal of Communications

number of candidate mode selected by Hadamard cost is equal to {8, 8, 3, 3, 3} for 4×4, 8×8, 16×16, 32×32 and 64×64 PU, respectively. Then the RD cost of each candidate mode is calculated and the mode with the minimum cost is selected as the final prediction mode. However, various combinations of CU size require great computational burden.

Currently, many algorithms [6]-[16] have been proposed to reduce the coding complexity for HEVC video coding. Reference [6] and [7] reduced the prediction directions utilized to perform RD optimization (RDO) by calculating the predominant angles. Reference [8] presents a content based complexity reduction scheme, which predicts the partitioning size prior to RDO calculations in LCU based on the texture property of the coding frame. In reference [9], Shen et al. divided CUs into two parts: texture homogeneity CUs and texture non-homogeneity CUs. For texture homogeneity CUs, CUs are directly no longer to be split. For texture non-homogeneity CUs, whether to split or not depends on the depths of neighboring CUs. In this algorithm, the parameters utilized to denote the correlations among neighboring CUs are fixed, while the correlations will vary for different sequences with different texture characteristics. Reference [10] and [11] exploited Hadamard cost to select intra prediction mode to avoid RDO process whose computational complexity is high. In reference [12], Lee et al. shrinked the depth range by the depth information of the collocated LCU in neighboring frames. When the depth of the collocated CU is 0, the current CU will skip intra prediction for CUs in depth 3. Otherwise, when the depth of the collocated CU is 3, the current CU will skip intra prediction for CUs in depth 0. Time reduction is limited for only one depth is excluded in this method. In reference [13] and [14], Shang et al. speeded up the coding process based on the coding information from neighboring coded CUs. Reference [15] and [16] modeled CU splitting as a binary classification problem and solved it by support vector machine (SVM). The performance of this method depends on the selection of feature and the accuracy of prediction. When the accuracy is low, it may result in a large RD performance loss. Reference [17] used Bayesian decision rule to tackle the binary problem: splitting or non-splitting. It chooses the variance of the residual coefficients as the feature for transform unit (TU) splitting, which may not be accurate to reflect the different distributions under different coding configurations (e.g. different quantization parameters (QPs)).

In this paper, we propose a spatial correlation based early CU decision algorithm (SECU) and a RD cost based early CU decision algorithm (RDCU) for HEVC intra coding. For SECU, we predict the depth of the current CU by the depth distributions of neighboring CUs. Instead of regarding 8×8 CU with mode SIZE_N×N as depth 3, coding block size of 4×4 is regarded as depth 4, which can make an accurate prediction on the CU's size. For RDCU, the distribution of RD cost obtained from the former coded frame is exploited to make an early decision on whether the CU should be split or not. To adapt to different coding contents, parameters utilized to make an early CU decision algorithm are periodically updated to prevent error propagation.

The rest of the paper is organized as follows. In Section 2, SECU and RDCU are presented respectively. Simulation results and analyses are demonstrated in Section 3. Section 4 concludes the paper.

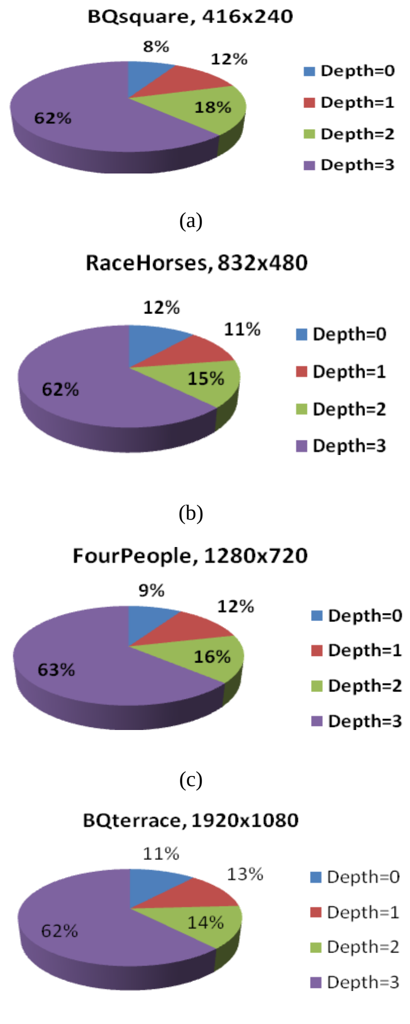


Fig. 3. Time-consumption distribution of different CU depths.

II. The Proposed Algorithm

A. *Spatial Correlation Based Early CU Decision Algorithm (SECU)*

Four test video sequences with different texture and motion properties are encoded to analyse the time-consumption distribution of different CU depths. RaceHorses (832×480) and BQterrace (1920×1080) are sequences with fast motion activities. The difference between them is that RaceHorses is a complex texture sequence, while BQterrace is a simple texture sequence. BQsquare (416×240) and FourPeople (1280×720) are slow motion sequences. The texture complexity in

BQsquare and FourPeople is complex and simple respectively. Fig. 3 shows the time-consumption distribution of different CU depths for all intra configuration in HM 14.0, where the QP is 27. We can observe that the probability of time-consumption in depth 3 is about 62% for sequences with different motion and texture features, while the probability of time-consumption of other depths is about 10%. The reason is that when CUs in depth 3 are tested, they are forcedly split into four 4×4 blocks indicated by SIZE_N×N. Besides, the number of prediction mode for CUs in depth 3 is 8, while for other depths is 3. Thus, making a decision on whether CUs in depth 3 should perform SIZE_N×N mode or not is significant for reducing the intra coding time in HEVC. In the proposed algorithm, we define depth 4 to denote 4×4 block for simplicity to determine CUs in depth 3 to split or not.

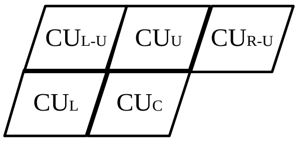


Fig. 4. Current CU and its neighboring CUs. CUL: left CU; CUL-U: left-up CU; CUU: up CU; CUR-U: right-up CU

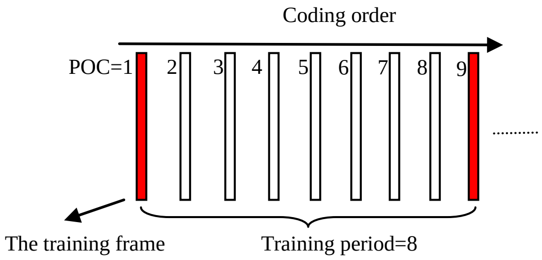


Fig. 5. Illustration of online training frames period

In natural video, there are strong spatial correlations especially for high resolution sequences. Through the depth information of neighboring CUs which is shown in Fig. 4, the depth level of the current CU is predicted early to bypass prediction process on some CU size. The optimal depth level of the current CU is predicted as follows

$$d_{pred} = \sum_{i=0}^{3} w_i \cdot d_i \tag{2}$$

where d_i is the depth level of the neighboring CU. w_i is the weighting factor derived from the correlation between the current CU and its neighboring CU. i is ranged from 0 to 3 to contain all of the four nearby CUs shown in Fig. 4. Since the correlation varies for different sequences under different coding environment, we update w_i in every eight frames shown in Fig. 5 to prevent error propagation. The error e^2 between the predicted value and actual value can be calculated by

$$e^2 = (d_{actual} - d_{pred})^2 \tag{3}$$

where d_{actual} is the actual value of depth. The optimal weighted vector $W\{w_0, w_1, w_2, w_3\}$ correspond to the one that minimizes the error between the actual depth value and the predicted depth value, which is denoted as

$$W = \min(e^2) \tag{4}$$

According to the least square approach, the optimal weighted vector W is computed as

$$W = (\mathrm{D}_N{}^T \mathrm{D}_N)^{-1} (\mathrm{D}_N{}^T \mathrm{D}_C) \tag{5}$$

where D_N is the depth level of neighboring CUs in the training frame which is a K×4 matrix (K is the number of CUs in the training frame), and D_C is the depth of the current CU which is a column vector with length of K.

After obtaining the optimal weighted factors in training frames, we predict the depth of the current CU by equation (1) in non-training frames. The criterion of performing splitting or non-splitting is

$$\begin{cases} non-splitting, d_{cur} - d_{pred} \geq 1.5 \\ splitting, d_{cur} - d_{pred} < -1.5 \end{cases} \tag{6}$$

If the predicted depth is less than or equal to the current CU's depth minus 1.5, the current CU is determined to not split early. Otherwise, if the predicted depth is greater than the current CU's depth plus 1.5, the current CU is split into four equal-size blocks directly instead of coding the current depth CU. Here the threshold set to 1.5 is because it can obtain good trade off between coding performance and computational complexity.

B. RD Cost Based Early CU Decision Algorithm (RDCU)

The CU's size depends on the value of RD cost. So it's necessary for us to investigate the correlation between RD cost distribution and CUs' depth. Fig. 6 shows an example of RD cost distribution for CUs with non-splitting mode and splitting mode in BQMall. Other test sequences have a similar result. From Fig. 6, we can see that CUs with lager RD cost probably select splitting mode, and CUs with smaller RD cost probably select non-splitting mode. The reason is that CUs with small RD cost indicate that the current partition style can well predict the coding area.

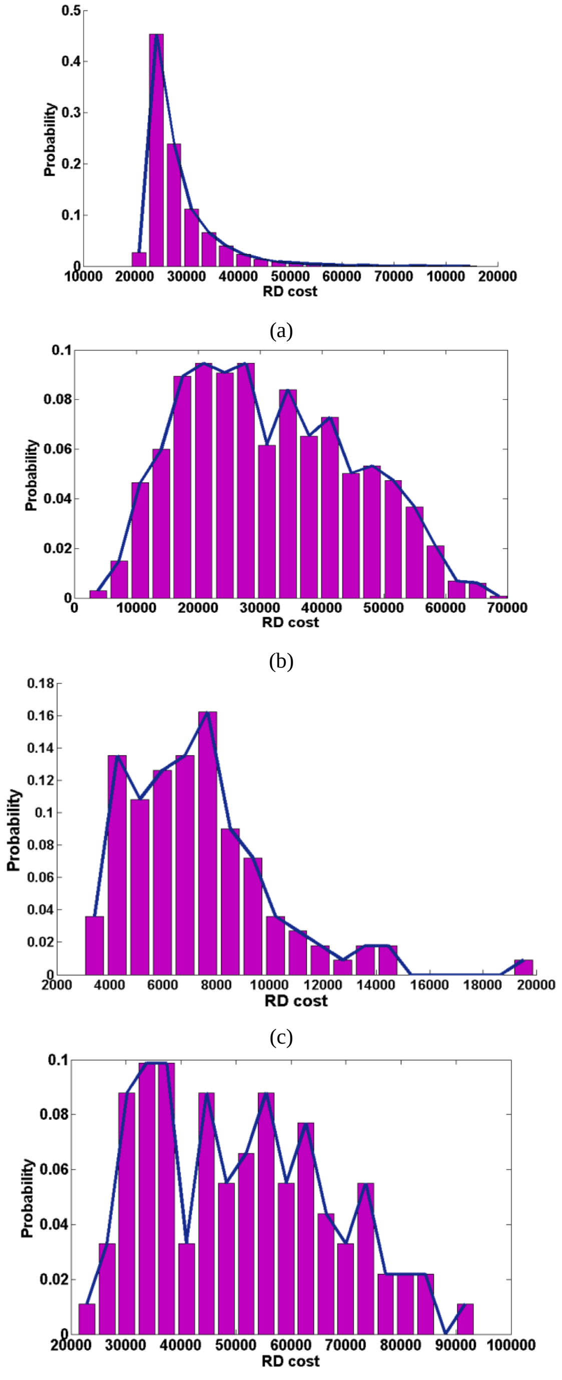
In training frames, when the sum of the RD cost of its sub-CUs is larger than the RD cost of the CU in depth D, the sum of the RD cost value of the CU Rd_D is computed and the number of CUs satisfying this condition is N_D. We then calculate the average of the RD cost by

$$Avg_D = \frac{Rd_D}{N_D} \tag{7}$$

By threshold T_D, some CU size is bypassed in advance. In our algorithm, T_D is obtained as follows

$$T_D = \alpha \cdot Avg_D \tag{8}$$

where α is the adjust parameter set to 0.8 in our simulation for reducing the coding time with nearly no coding performance loss. Generally neighboring frames in time domain possess similar texture information. Thus the RD cost distribution is almost similar. In this scheme, if the RD cost of the current CU in depth D in the non-training frames is less than T_D, the current CU is determined to not split into four equal-sized CUs.



(a) (b) (c) (d)

Fig. 6. Example of RD cost distribution for CU with non-splitting and splitting mode in BQMall. (a) CUs in depth 2 with non-splitting mode, (b) CUs in depth 2 with splitting mode, (c) CUs in depth 0 with non-splitting mode, (d) CUs in depth 0 with splitting mode.

C. The Proposed Algorithm

Based on the analysis above, the proposed fast CU decision algorithm for HEVC intra coding is presented as:

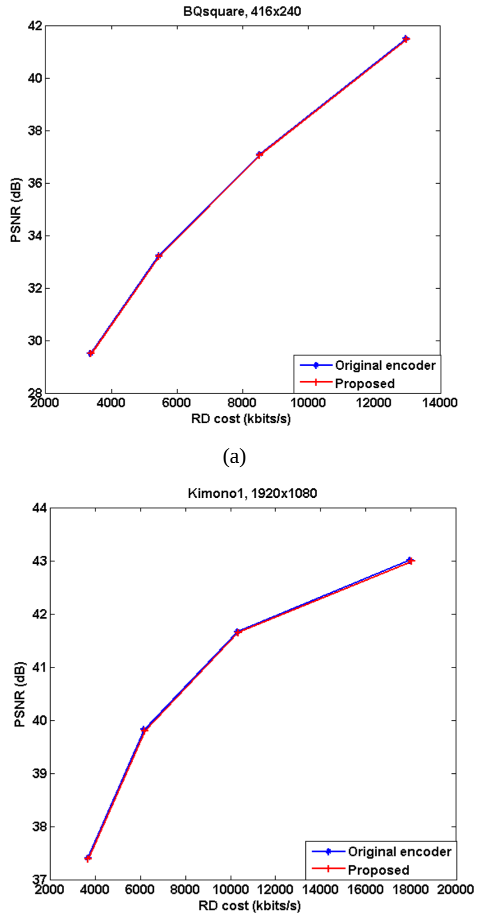
Step 1: Perform intra prediction for a CU. When the current frame is the training frame, code the current frame with the original encoder to update the statistical parameters and go to step 5, else go to step 2.

Step 2: Calculate the predicted depth and test whether the criterion of performing splitting or non-splitting is satisfied. Then go to step 3.

Step 3: Perform the RDO process on the current CU to derive the RD cost and compare it with the threshold T_D. If the RD cost is less than T_D, the current CU is not split. Otherwise, the current CU goes into the normal coding process. Then go to step 4.

Step 4: Decide the optimal CU depth.

Step 5: Code the next frame.



(a) (b)

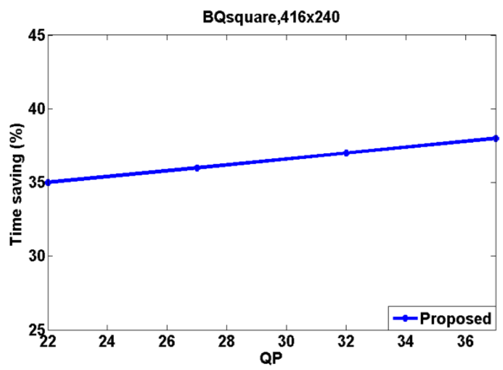
Fig. 7. RD curves of BQsquare and Kimono1 under QP=22, 27, 32, 37.

III. Experimental Results

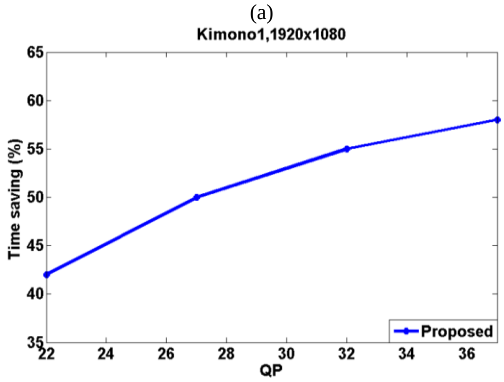
To justify the effectiveness of the proposed early termination algorithm, the proposed algorithm is performed on HEVC test model (HM14.0) under the common condition defined in [19]. All-Intra-configuration is used for the simulation with QP of 22, 27, 32 and 37. Five resolution sequences Class A (4K×2K), Class B (1080p), Class C (WVGA), Class D (OWVGA) and Class E (720p) are all tested for performance verification. Time saving (TS), BDBR (%) and BDPSNR (dB) [20] are utilized to measure the coding efficiency, where BDBR and BDPSNR represent the average bitrate and PSNR differences as compared with the original encoder.

TABLE I: RESULTS OF THE PROPOSED ALGORITHM COMPARED WITH THE EXISTING ALGORITHM

		Lee [12]			Shen [9]			Proposed		
Resolution	Sequence	BDBR (%)	BDPSNR (dB)	TS (%)	BDBR (%)	BDPSNR (dB)	TS (%)	BDBR (%)	BDPSNR (dB)	TS (%)
Class A	Traffic	0.15	-0.01	21	0.97	-0.06	37	1.10	-0.06	45
	PeopleOn street	0.24	-0.01	17	1.28	-0.07	41	1.29	-0.07	48
Class B	BasketballDrive	4.17	-0.10	51	2.50	-0.06	61	0.63	-0.02	49
	BQterrace	0.06	0.00	22	0.62	-0.04	39	0.62	-0.04	44
	Kimono1	0.46	-0.02	57	0.81	-0.03	38	0.73	-0.03	52
	Tennis	3.27	-0.10	52	2.86	-0.09	57	0.87	-0.03	53
Class C	BasketballDrill	0.32	-0.02	16	0.79	-0.04	29	1.19	-0.06	46
	Bqmall	0.28	-0.02	15	1.16	-0.07	36	1.05	-0.06	47
	RaceHorses	1.25	-0.08	24	0.55	-0.04	32	1.46	-0.09	48
	PartyScene	0.19	-0.01	12	0.14	-0.01	24	0.52	-0.04	42
Class D	BasketballPass	0.16	-0.01	17	1.41	-0.08	38	0.93	-0.05	44
	BQsquare	0.00	0.00	10	0.32	-0.03	23	0.29	-0.03	37
	BlowingBubbles	0.05	0.00	10	0.05	0.00	17	0.52	-0.03	36
	Keiba	1.56	-0.10	20	0.73	-0.05	29	1.23	-0.08	42
Class E	FourPeople	0.06	0.00	21	2.19	-0.13	49	1.10	-0.06	48
	Johnny	0.22	-0.01	38	3.88	-0.16	59	1.11	-0.05	54
	KristenAndSara	0.31	-0.02	35	3.31	-0.17	59	1.03	-0.05	53
Average		0.75	-0.03	26	1.39	-0.07	39	0.92	-0.05	46



(a)



(b)

Fig. 8. Time saving of BQsquare and Kimono1 under QP=22, 27, 32, 37.

Table I shows that, in comparison with the original encoder, the proposed algorithm achieves 46% encoding time saving for all intra coding with 0.92% BDBR increase and 0.05 dB BDPSNR decrease on average. Fig. 7 and Fig. 8 demonstrate the simulation results of the proposed method under four QPs (22, 27, 32 and 37) for sequence BQsquare and Kimono1. We can observe that the proposed method can significantly reduce the coding time by 36%~54%, while the RD curve of the proposed algorithm is almost the same as that of the original encoder indicating no RD performance loss. Meanwhile, we also compare the proposed algorithm with Lee's [12] and Shen's algorithm [9]. As shown in Table 1, Lee's algorithm reduces the coding time by 26% with 0.75% BDBR increase. Although it has a better RD performance, while the time saving is limited for only skipping one depth in the coding process of LCU. As compared with Shen's algorithm, the proposed method saves more than 7% coding time, while saves 0.47% BDBR and keeps a better quality. In addition, the rangeability of time saving and BDBR is 17% ~61% and 0.05%~3.38% in Shen's algorithm, with 36%~54% and 0.29%~1.46% in the proposed algorithm. The instability of Shen's algorithm may cause by the way of deciding the texture homogeneity region. By computing the mean absolute deviation of pixels in coding areas, the CU is performed to make different size decisions. When the coding configuration is changed (eg. under different QPs), using the same threshold to determine the CU's size can result in the instability of RD performance. Besides, the optimal depth may be predicted insufficiently for analysing the depth correlation among coding blocks with the size from

64×64 to 8×8 instead of from 64×64 to 4×4. In the proposed algorithm, SECU predicts the CU's size by updating adaptively the depth correlation between the current CU and its neighboring CU in all block sizes. RDCU exploits RD cost to predict the optimal CU depth and update the prior information constantly to improve the accuracy of prediction.

IV. CONCLUSION

In this paper, we propose a fast CU decision algorithm to reduce the computational complexity for HEVC intra coding. The CU's size is determined by the depth of neighbouring CU and the RD cost distribution to terminate the process of intra prediction early. Experimental results demonstrate that the proposed algorithm can achieve 46% time saving with 0.92% BDBR increase and 0.05dB BDPSNR loss as compared with the original encoder. In addition, the proposed algorithm outperforms the state-of-the-art fast CU size decision algorithm, with about additional 7%~20% time saving. Further work will focus on combining the proposed algorithm with other fast mode decision algorithm for inter frames.

ACKNOWLEDGMENT

This work is supported by the Open Research Fund of AnHui Key Laboratory of Detection Technology and Energy Saving Devices, AnHui Polytechnic University, China (1506c085002), the Major Program of Scientific Research of Anhui Province, China (KJ2015A071), and the Scientific Research Foundation for Talents, Anhui Polytechnic University of China (2014YQQ006).

REFERENCES

[1] ITU-T Rec. H.264 and ISO/IEC 14496-10 (MPEG-4 AVC), Version 8, document ITU-T and ISO/IEC, Advanced Video Coding for Generic Audiovisual Services, 2007.

[2] 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.

[3] F. Bossen, B. Bross, K. Suhring, and D. Flynn, "HEVC complexity and implementation analysis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1685–1696, 2012.

[4] J. Lainema, F. Bossen, W. 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.

[5] JCTVC-D283, JCT-VC Meeting, Further Encoder Improvement for Intra Mode Decision, 2011.

[6] J. Mohammadreza, C. Stephane, and C. Francois, "Fast HEVC intra mode decision based on edge detection and SATD costs classification," in *Proc. Data Compression Conference (DCC)*, Snowbird, 2015, pp. 43-52.

[7] Y. Ting and T. Chang, "Gradient-based PU size selection for HEVC intra prediction," in *Proc. International Symposium on Circuits and Systems*, Melbourne, Australia, 2014, pp. 1929-1932.

[8] M. U. K. Khan, M. Shafique, and J. Henkel, "An adaptive complexity reduction scheme with fast prediction unit decision for HEVC intra encoding," in *Proc. Int. Conf. Image Processing*, Melbourne, Australia, 2013, pp. 1578-1582.

[9] L. Shen, Z. Zhang, and Z. Liu, "Effective CU size decision for HEVC intracoding," *IEEE Trans. Image Processing*, vol. 23, no. 10, pp. 4232-4241, 2014.

[10] D. Kim, S. Jun, J. Jung, J. Choi, and J. Kim, "A fast intra-prediction method in HEVC using rate-distortion estimation based on Hadamard transform," *ETRI Journal*, vol. 35, no. 2, pp. 270-280, 2013.

[11] H. Zhang and Z. Ma, "Early termination schemes for fast intra mode decision in high efficiency video coding," in *Proc. International Symposium on Circuits and Systems*, China, pp.45-48, 2013.

[12] H. S. Lee, K. Y. Kim, T. R. Kim, and G. Park, "Fast encoding algorithm based on depth of coding-unit for high efficiency video coding," *Opt. Eng.*, vol. 51, no. 6, 2012.

[13] X. Shang, G. Wang, T. Fan, and Y. Li, "Statistical and spatiotemporal correlation based low-complexity video coding for high-efficiency video coding," *J. Electron. Imaging*, vol. 24, no. 2, 2015.

[14] X. Shang, G. Wang, T. Fan, and Y. Li, "Fast CU size decision and PU mode decision algorithm in HEVC intra coding," in *Proc. IEEE International Conference on Image Processing*, 2015, pp. 1593-1597.

[15] X. Shen and L. Yu, "CU splitting early termination based on weighted SVM," *EURASIP Journal on Image and Video Processing*, pp. 1-11, 2013.

[16] Y. Liu, Z. Chen, J. Fang, and P. Chang, "SVM-Based fast intra CU depth decision for HEVC," in *Proc. Data Compression Conference*, 2015, pp. 458-458.

[17] L. Shen, Z. Zhang, X. Zhang, P. An, and Z. Liu, "Fast TU size decision algorithm for HEVC encoders using Bayesian theorem detection," *Signal Processing: Image Communication*, pp. 121-128, 2015.

[18] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Wiley, 2001.

[19] Joint Collaborative Team on Video Coding of ISO/IEC and ITU-T, *JCTVC-K1100*, Common HM test condition and Software Reference Configurations, 2012.

[20] ITU-T SG16 Q.6 Document, *VCEG-M33*, Calculation of Average PSNR Differences between RD-Curves, 2001.

Zhilong Zhu received the MD degree from Anhui Agricultural University, Hefei, China, in 2004, in Agricultural Mechanization Engineering. Since 2004, he has been a Faculty Member with Anhui Polytechnic University, Wuhu, China. He has attempted many scientific research projects in Artificial Intelligence. His main research areas are in digital electronic technology, computer vision and pattern recognition.

Gang Xu received the MD degree from Guilin University of Electronic Technology, Guilin, China, in 2004, in Electronic Technology and now is PhD student in Jiangnan University, Wuxi, China, in Control Science and Engineering. Since 1995, he has been a Faculty Member in Anhui Polytechnic University, Wuhu, China, and is currently an associate Professor of Electronic Information Science and Technology. His main research areas are in computer vision, feature extraction and pattern recognition.

Fengsui Wang received the Ph.D. degree from Nanjing University, Nanjing, China, in 2013, in electronic engineering. Since 2013, he has been a Faculty Member with Anhui Polytechnic University, Wuhu, China, and is currently an associate Professor of electronic and communication engineering. His current research interests include the areas of image processing, video signal processing, and computer vision.