Low Complexity I-frame Coding for HEVC Standard

Mohammadreza Ramezanpour
PHD candidate in Department of Computer Engineering
Science and Research branch, Islamic Azad University
Tehran, Iran
Email: mr.ramezanpoor@srbiau.ac.ir

Farzad Zargari
Information Technology Research Institute
Iran Telecom Research Center (ITRC)
Tehran, Iran
Email: zargari@itrc.ac.ir

Abstract— **High Efficiency Video Coding (HEVC), the latest international video coding, greatly outperforms previous standards such as H.264/AVC in terms of coding bitrate and video quality. The coding efficiency improvement in HEVC is achieved by accepting several new techniques such as recursive quad-tree structure and increasing the number of intra prediction modes. However, computational load is also increased due to these new techniques. In this paper, we propose fast decision algorithm according to homogeneity of Coding Unit (CU). In the proposed method, we calculate CU smoothness based on edge strength in four different directions and predict CU size by this means. Experimental results indicate that the proposed algorithm can provide on average 34.8% savings on coding time with only 1.03% BD-rate loss, whereas it maintains the same coding video quality compared with HEVC test model, HM15.0, in all intra-main configuration.**

Keywords—***I-frame coding; CU smoothness; CU size decision; HEVC***

I. INTRODUCTION

Joint Collaborative Team on Video Coding (JCT-VC) which consisted of ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) started working on the development of High Efficiency Video Coding (HEVC) standard [1] in April 2010. The main purpose of this group was to prepare a standard to be able to provide a significant improvement in coding efficiency compared with the H.264/AVC, while maintaining the same video quality mainly in high resolution videos [1], [2]. The standard approved by both ISO/IEC and ITU/T in January 2013. The approved HEVC standard could achieve 50% lower bitrate than H.264/AVC in the similar coding qualities. The high coding efficiency in HEVC is achieved at the expense of increasing the computational complexity. One of the techniques that is used in HEVC is the recursive quad-tree structure which first partitions a frame into 64×64 Coding Units (CUs) and each CU may be recursively split into four sub-CUs. For each CU, a prediction is produced based on the previously coded data, either from the current frame using intra prediction or from other frame using inter prediction. Coding efficiency improvement in HEVC I-frame coding is achieved partly by increasing the number of intra Prediction Modes (PMs) from 9 in H.264/AVC to 35. These features lead to much higher coding efficiency while increasing the computational complexity compared with intra prediction in H.264/AVC, which substantially is computational intensive. In this paper, we propose a fast CU size decision algorithm which is derived from the analysis of intra prediction complexity of the HM test model. Experimental results indicate that our proposed method has on average 34.8% lower encoding time for I-frames compared with the HEVC test model and for the same PSNR quality imposes 1.03% increase in the bit-rate. Comparison of the proposed method with other methods, indicate that our method achieves high reduction in coding time with negligible impact in coding efficiency.

The rest of this paper is organized as follows. Section II reviewed the related works on fast intra PMs in HEVC. Section III gives explanation on intra prediction in HEVC standard along with discussions about the complexity of intra prediction mode decision. In section IV description of the proposed method is provided in more detail. Experimental results and concluding remarks are given in section V and VI, respectively.

II. A REVIEW ON FAST HEVC I-FRAME CODING

High number of prediction modes and various CU sizes employed in HEVC I-frame coding process provide high coding efficiency, but they result in extremely large encoding time. Therefore, it is necessary to develop a method that can reduce complexity of I-frame coding with minimal loss of video quality. Recently, a number of efforts have been proposed to reduce the computational complexity for HEVC I-frame coding. A fast CU size decision algorithm is proposed in [3] to reduce computational complexity of I-frame coding based on the previous decisions in spatially nearby CUs. They reported about 21% average reduction of coding time and about 1.74% average increase in the bit rate for similar coding qualities. Shen and Yu [4] utilized machine learning to accelerate decision on CU size. In [5], a fast CU size decision algorithm was proposed based on CU texture complexity. They skip the 64×64 and 32×32 CU sizes if CU complexity is larger than a predefined threshold and skip the 8×8 and 4×4 partitions if the complexity is below a second predefined threshold. The authors reported 29% average time reduction with average 0.47% penalty in bit rate. Kim et al. [6] applied early termination pruning to the tree to avoid the evaluation of smaller CU sizes, when the Rate-Distortion (R-D) cost of the best mode in the Rate Distortion Optimization (RDO) stage is lower than a predefined threshold.

In order to reduce the computational complexity of I-frame coding, Yan et al. [7] merged the adjacent candidate mode obtained through the Rough Mode Decision (RMD) stage into same groups to reduce the number of PMs in the RDO stage. Moreover, they applied a pixel based edge detection algorithm to select the best prediction mode. The algorithm achieved average time saving about 23.5% whereas bit rate increases 1.3% on average. Wang and Siu [8] utilized variance of reference samples to measure the smoothness of the entire reference samples. If the variance is smaller than a selected threshold, only planar mode is used for predicting the current CU and there is no need to evaluate the other modes. Zhao et al. [9] employed RMD stage to reduce the number of intra modes in the RDO stage. In their proposed method, the first N modes with least R-D cost are selected by using Hadamard transform instead of Discrete Cosine Transform (DCT) and then the MPMs that derived from the intra modes of the left and top neighboring PUs, are added to the N modes that take part in the RDO stage. By using this method about 6% reduction in coding time is reported.

In previous works [10-11] we proposed two stage fast mode decision method for HEVC that significantly reduced the number of evaluated modes in intra prediction process. In the first stage, only 19 modes used in the RMD stage instead of 35 modes and the number of best candidate modes reduced based on the correlations among directional modes. In the second stage, the number of selected modes through the RMD stage is reduced as well. Thus, RDO is applied to lower number of modes. In this paper, we improve and extend the fast mode decision algorithm for intra prediction of HEVC in the Yao's proposed method [12] to reduce the number of CU sizes. The proposed method differs with the Yao's proposed method [12] in two aspects. On one hand, we use edge strength to reduce the evaluated CU-sizes in the RMD and RDO stages. On the other hand, edge strength is calculated only for 4×4 blocks when CU size is 64×64 which it causes the computational complexity to reduce significantly. Before presenting the proposed method an overview about I-frame coding in the HEVC standard is provided in the next section to the extent which is required for understanding the rest of paper.

III. INTRA FRAME CODING IN HEVC

In HEVC I-frame coding, each picture is divided to equal size blocks named Coding Tree Unit (CTU). CTUs are the root of quad-tree partitioning which employed in HEVC quad-tree partitioning divides each CTU to CUs. Each CU consists of one luma and two chroma blocks and its size is $2^n \times 2^n$ where n is an integer number ranging from three to six. Fig. 1 indicates partitioning of a 64×64 CTU into 32×32 to 8×8 CUs and the arrows in Fig. 1 indicates the coding order of CUs in a CTU. Each CU includes one or more Prediction Units (PUs). CU is the unit that is coded either by intra frame or inter frame and for each PU the same prediction mode is employed. PU sizes are in the form of $2^m \times 2^m$ where m is an integer number ranging from two to six. There are various Discrete Cosine Transform (DCT) sizes in HEVC and Transform Unit (TU) indicates transform size. As a result, TU size correspond to DCT sizes in HEVC which is in the form of $2^k \times 2^k$ where k is an integer number ranging from two to five. Hence the residues for each PU can be transformed by various sizes of DCT which means each PU consists of one or more TUs. Partitioning a PU to one or more TUs is indicated by residual quad-tree structure. Since each CU can be divided to one or more PUs, the intra prediction process in HEVC includes two stages. These two stages are deciding on the appropriate size of PU and selecting the proper prediction mode.

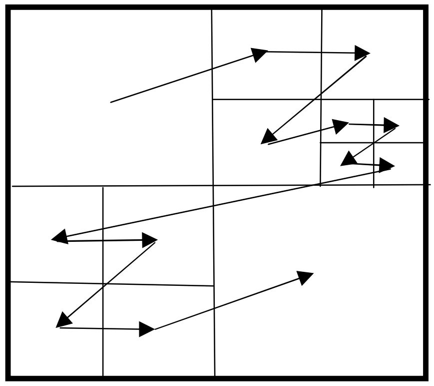

Fig. 1 An example of the partitioning of a 64×64 CTU into CUs.

I-frame coding consists of 35 intra PMs including 33 directional modes and two non-directional (DC and planar) modes. The non-directional PMs are mainly employed for smooth areas. HEVC uses RDO process to select the best prediction mode. RDO is defined as.

$$J = (SSE_{Luma} + W_{chroma} \times SSE_{chroma} + \lambda \times B_{mode}) \quad (1)$$

where, SSE is the sum of squared transform coefficients, λ is the Lagrange multiplier, B_{mode} specifies bit cost to be considered for mode decision and W_{chroma} is the weighting parameter for chroma-specific decisions [13].

HEVC test model employs a fast intra prediction algorithm proposed by Zhao et al. [9] to reduce the number of tested modes in RDO stage. This pre-processing step, namely RMD, uses Hadamard transform instead of DCT to determine best candidate PMs. N best candidates in RMD stage are used in RDO process along with the Most Probable Modes (MPMs) from neighboring blocks. The number N depends on the PU size, for 8×8 and 4×4 block it is selected as 8 and for the rest of PU sizes it is 3. As a result, selecting appropriate CU, PU and TU sizes includes four stages. In the first stage, for each PU size, RMD is employed to select the best N candidate modes. In the next stage, the PMs of neighboring blocks are added to the selected N modes to be tested in RDO stage and the best PU size and prediction mode is determined. In the last stage (RQT) various TU sizes are tested for the best PU size and mode to select the most effective TU size(s).

IV. PROPOSED ALGORITHM

In the HEVC test model (HM15.0), CU size varies from 64×64 to 8×8 and careful inspecting the selected CU size for each region indicates that decoder splits the coded region to smaller CUs until achieves a homogeneous area. If 64×64 region to be a homogeneous area, it is coded as a 64×64 CU but if it is non-homogeneous it will be split to 32×32 CUs and the same approach is employed for coding the 32×32 CUs. HM15.0 tests the entire PMs, for each CU size by ignoring the aforementioned relation between CU size and homogeneity of its region. In this paper, we decide about coding of a CU according to the homogeneity of its region. In order to detect the homogeneity of a region, we consider the dominant direction for a 4×4 block is found by dividing it into four non overlapping 2×2 blocks and one 2×2 block in the center of 4×4 block (Fig. 2). The average of pixels in each 2×2 block is calculated as:

$$P_j = \left(I(x_j, y_j) + I(x_j + 1, y_j) + I(x_j, y_j + 1) + I(x_j + 1, y_j + 1)\right) >> 2 \quad (2)$$

where (x_j,y_j) denotes the starting position of k[th] 2×2 sub-block and $I(x,y)$ refers to the x row and y column pixel in the block.

The four edge directions are calculated by using the P_i values as:

$$d_1 = |P_1 - P_0| + |P_3 - P_2| \quad (3)$$

$$d_2 = |P_2 - P_0| + |P_3 - P_1| \quad (4)$$

$$d_3 = |P_4 - P_1| + |P_2 - P_4| \quad (5)$$

$$d_4 = |P_4 - P_0| + |P_3 - P_4| \quad (6)$$

where d_1 to d_4 indicate the prediction error when we use a specified direction among horizontal, vertical, 45^o and 135^o directions to predict the pixel values of the 4×4 block by the other pixels in the same block. As a result the direction which generates minimum prediction, which means minimum residual, is the dominant direction of the 4×4 block. In our method each 64×64 CU is divided to 4×4 non overlapping blocks and the direction for each 4×4 block is detected. In order to decide that the PMs of a CU should be tested or the CUs should be split to lower size CUs we employ the direction of 4×4 blocks in the CU (Fig. 3). The ratio of the number of blocks with majority direction to the total number of 4×4 blocks in the CU is called strength of direction in CU. If strength of direction of a CU is lower than 50%, we assume that the PMs in the CU should not to be tested and the CU should be split to lower size CUs. Otherwise the PMs are tested for the CU size.

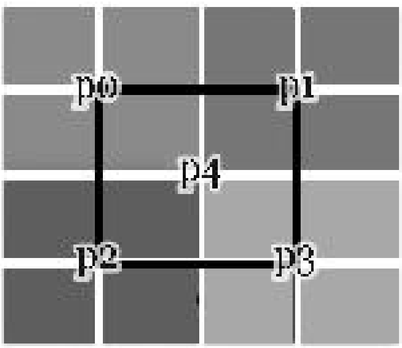

Fig. 2 Partitioning 4×4 block into five 2×2 sub-blocks

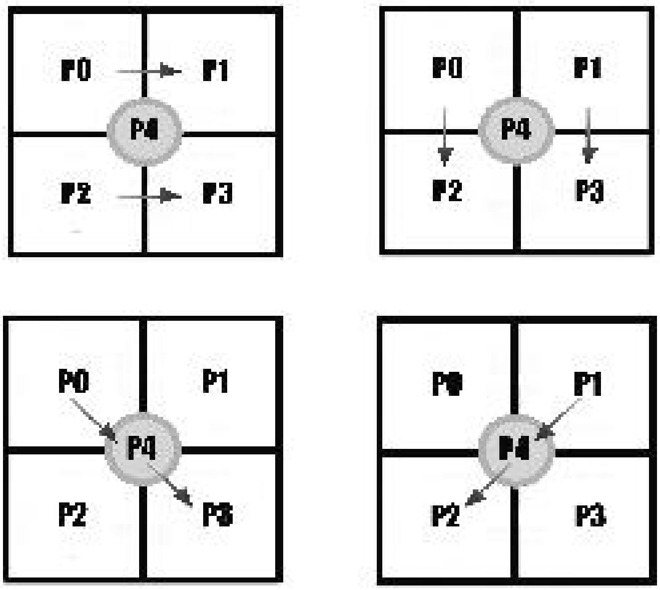

Fig. 3 detection of edge direction in a) horizontal b) vertical c) right diagonal d) left diagonal directions

We encoded a number of standard test video sequences by HM15.0 and Table I indicates the percentage of CUs that our assumption may fail. The negligible percentages reported in Table I indicate that our assumption can be used in CU decision with negligible negative impact on video coding efficiency.

TABLE I. PERCENTAGE OF CU IS SELECTED AS THE BEST CU WHEN STRENGTH OF DIRECTION OF CU S LOWER THAN 50%

Test sequences	Resolution	Percentage
Basketballpass	416×240	1%
Partyscene	832×480	2%
Fourpeople	1280×720	2%
Parkscene	1920×1080	1%
Peopleonstreet	2560×1600	2%

V. EXPERIMENTAL RESULTS

HEVC test model (HM15.0) is used to implement our proposed method. Nine standard video sequences from class A to class E which are proposed by JCT-VC group are used in our experiments. The experiment conditions are according to [14] and are as follows.

All frames are encoded as intra frames. Quantization Parameter (QP) are set to be 22, 27, 32, and 37. Sample Adaptive Offset (SAO) and Context Adaptive Binary Arithmetic Coding (CABAC) are enabled. Maximum CU size selected as 64×64 and maximum depth level set to 4. Fast intra prediction mode is enabled. More specially, the macro symbol FAST_UDI_USE_MPM is set to 1.

The efficiency of the proposed method is measured by the Bjontegaard-Distortion Rate (BD-Rate) [15] according to PSNR value. PSNR value for color videos including luma and chroma components is generated as:

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) \quad (7)$$

where the Mean Square Error (MSE) computed as:

$$MSE = \frac{6\times MSE_Y + MSE_U + MSE_V}{8} \quad (8)$$

where MSE_Y, MSE_U and MSE_V are MSE for Y, U and V components of coded color video, respectively [16]. The reduction in the coding time is employed as a measure indicating the improvement in the computational complexity. The relative coding time is defined as:

$$DT = \frac{Time_{proposed} - Time_{HM15.0}}{Time_{Hm15.0}} \times 100 \quad (9)$$

Table II tabulates the results of comparison of the proposed method with HM15.0 in the conducted experiments. The negative numbers in Table II indicate the reduction in the measured parameters with respect to HM15.0. According to the experimental results given in Table II, we conclude that the proposed method in comparison with the HEVC test model (HM15.0) can on average reduce the encoding time by 34.8% and in the same video qualities the average bitrate increase is about 1.03%. The maximum encoding time reduction is 46% in *Basketballdrill*, whereas the minimum encoding time reduction is 29% in *Partyscene* sequence. The comparison of coding time and efficiency of the proposed method with other methods is given in Table III. The results given in Table III indicate that the proposed method achieves high reduction in the coding time while imposing the minimum degradation in the compression performance with the other methods.

VI. CONCLUSION

This paper presents a low complexity CU size decision algorithm to reduce the computational complexity of the HEVC I-frame coding. The proposed algorithm is implemented on the HEVC test model (HM15.0). The experimental results show that the proposed method can significantly reduce the computational complexity of HEVC I-frame coding while maintaining almost the same coding efficiency as the HEVC test model. The performance of the proposed method is compared with other proposed methods in the literature. Comparison with other methods indicate that the proposed method achieves higher reduction in coding time with respect to other methods, whereas the other methods impose even higher increase in the rate of the coded video. By considering the experimental results and comparison with the other methods it can be deduced that the proposed method achieves high reduction in the coding time of the HEVC encoder with minimum increase in the rate of coded video.

TABLE II. COMPLEXITY AND ENCODING TIME OF THE PROPOSED METHOD WITH RESPECT TO HM15.0 IN ALL INTRA-MAIN CONFIGURATION

Test sequences	BD-Rate [%]	BD-PSNR [db]	DT [%]
Basketballpass	1.4	-0.08	-43
Blowingbubbles	1.5	-0.07	-32
Partyscene	2.1	-0.15	-29
Basketballdrill	0.6	-0.03	-46
Fourpeople	0.9	-0.05	-31
Kimono1	0.4	-0.01	-36
Parkscene	0.5	-0.02	-32
Peopleonstreet	1.3	-0.07	-33
Traffic	0.7	-0.03	-31
Average	**1.03**	**-0.06**	**-34.8**

TABLE III. COMPARISON OF CODDING TIME AND EFFICIENCY FOR DIFFERENT METHODS

Methods	BD-Rate [%]	BD-PSNR [db]	Time Saving [%]
Shen et al.[3]	1.74	-0.08	-21.1
Yan et al.[7]	1.30	---	-23.5
Yao et al.[12]	1.86	-0.11	-36
Proposed method	1.03	-0.06	-34.8

References

[1] G. J. Sullivan, J. Ohm, W.-J. Han, T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," IEEE Trans. Circuits and Syst. Video Technol., vol. 22, pp. 1649-1668, 2012.

[2] B. Bross, J Ohm, G. J. Sullivan, W.J. Han, T. Wiegand, "High efficiency video coding text specification draft 10. JCTVC-L1003, 2013.

[3] L. Shen, Z. Zhang, Z. Liu, "Effective CU size decision for HEVC intra coding," IEEE Trans. Image Processing, vol. 23, pp. 4232-4241, 2014.

[4] X. Shen, L.Yu, "CU splitting early termination based on weighted SVM," EURASIP journal on image and video processing, 2013, DOI 10.1186/1687-5281-2013-4.

[5] G. Tian, S. Goto, "Content adaptive prediction unit size decision algorithm for HEVC intra coding," in: IEEE Picture Coding Symposium (PCS), Krakow, pp.405-408, 2012.

[6] O. Kim, D. Jun, S-h Jung, J. S. Choi, 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.

[7] S. Yan, L. Hong, W. He, Q. Wang, "Group based fast mode decision algorithm for intra prediction in HEVC," in: IEEE International Conference on signal image technology and internet based systems, Naples, pp. 225-229, 2012.

[8] L.-L. Wang, W.-C. Siu, "Novel adaptive algorithm for intra prediction with compromised modes skipping and signaling processes in HEVC," IEEE Trans. Circuits and Syst. Video Technol., vol. 23, pp. 1686-1694, 2013.

[9] L. Zhao, I. Zhang, S. Ma, D. Zhao, "Fast mode decision algorithm for intra prediction in HEVC," in: IEEE International Conference on Visual communication and image processing, Tainan, Taiwan, pp.1-4, 2011.

[10] M. Ramezanpour Fini, F. Zargari, "A fast intra mode decision method based on reduction of the number of modes in HEVC standard," IEEE international Symposium on telecommunications (IST2014), pp.839-843, 2014.

[11] M. Ramezanpour Fini, F. Zargari, "Two stage fast mode decision algorithm for intra prediction in HEVC," Multimedia Tools and Applications. Vol. 74, No. 9, 2015, DOI: 10.1007/s11042-015-2675-5.

[12] Y. Yao, X. Li, Y. Lu, "Fast intra mode decision algorithm for HEVC based on dominant edge assent distribution. Multimedia Tools and Applications, 2014, DOI: 10.1007/s11042-014-2382-7

[13] I-K Kim, K McCann, K Sugimoto, B Bross, W-J Han, G Sullivan, “High efficiency video coding (HEVC) test model 15 (HM15) encoder description,” In: joint collaborative team on video coding (JCT-VC), Document of JCTVC-Q1002, 17thMeeting, Valencia, 2014.

[14] F. Bossen, Common test condition and software reference configurations, in: Joint Collaborative Team on Video Coding (JCT-VC), Document of JCTVC-L1100, 12th Meeting, Geneva, 2013.

[15] G. Bjontegaard, "Calculation of average PSNR differences between RD_curves,” in: 13th VCEG-M33 Meeting, Austin, Tx, 2001.

[16] G. Sullivan, K. Minoo, “Objective quality metric and alternative methods for measuring coding efficiency”, document JCTVCH0012, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), 8th Meeting: San Jose, CA, USA, PP. 1–10, 2012.