Accepted in IEEE 18th International Workshop on Multimedia Signal Processing (MMSP 2016), 2016

# Coding Unit Splitting Early Termination for Fast HEVC Intra Coding Based on Global and Directional Gradients

Mohammadreza Jamali and Stéphane Coulombe
École de technologie supérieure, Université du Québec, Montreal, Canada
mohammadreza.jamali.1@ens.etsmtl.ca, stephane.coulombe@etsmtl.ca

***Abstract*—High efficiency video coding (HEVC) doubles the compression ratio as compared to H.264/AVC, for the same quality. To achieve this improved coding performance, HEVC presents a new content-adaptive approach to split a frame into coding units (CUs), along with an increased number of prediction modes, which results in significant computational complexity. To lower this complexity with intra coding, in this paper, we develop a new method based on global and directional gradients to terminate the CU splitting procedure early and prevent processing of unnecessary depths. The global and directional gradients determine if the unit is predicted with high accuracy at the current level, and where that's the case, the CU is deemed to be *non-split*. Experimental results show that the proposed method reduces the encoding time by 52% on average, with a small quality loss of 0.07 dB (BD-PSNR) for *all-intra* scenarios, as compared to the HEVC reference implementation, HM 15.0.**

***Keywords—HEVC; intra coding; CU splitting; fast coding***

## I. INTRODUCTION

High efficiency video coding (HEVC) [1], [2] is the newest international standard for video compression, and was developed by the joint collaborative team on video coding (JCT-VC). HEVC provides an improved coding performance and achieves up to 50% bitrate reduction as compared to H.264/AVC [3] for the same perceptual video quality. This improved performance is attained through a higher flexibility of HEVC and the introduction of more coding tools than previous video coding standards. While H.264/AVC employs 16×16 macroblocks, HEVC introduces coding tree units (CTUs) with a 64×64 maximum size. The CTU may be split recursively and content-adaptively into coding units (CUs) in a quadtree-based manner, resulting in an efficient encoding of background regions and objects with various sizes and shapes. With mode decision, the number of intra modes is increased from 9 modes in H.264/AVC to 35 modes, including *dc*, *planar* and 33 angular modes, for each prediction unit (PU). The use of this encoding model provides a significant bitrate reduction and increases video quality. However, this comes at the cost of high computational complexity and long encoding times.

Fast HEVC intra coding is realized either by fast mode decision, i.e., a decrease in the number of modes to be checked by the high-complexity rate distortion optimization (RDO) process [4]-[7], or by CU splitting early termination, i.e., RDO is performed only for some selected CU levels. In [8], the depth information of the neighbouring CUs is exploited in order to take advantage of the spatial correlation in a frame. In [9], the authors propose a bottom-up approach to find an optimum depth, using regional video texture. In [10], a two-step early termination method is presented to divide a frame into homogenous and non-homogenous regions by employing the coefficient of variation (CV). In [11], the depth decision is made by skipping some specific CU levels which are seldom used in the previous frame and neighbouring CUs, along with thresholds on the RDO cost of the current level. A similar approach is proposed in [12] by the early termination of the CU splitting process using the RDO costs of the intra modes at the current depth. Using neighbouring CUs, as proposed by many works, results in a domino effect, i.e., a wrongly decided CU splitting pattern may be propagated, leading to a significant reduction in visual quality. On the other hand, utilizing variance or RDO cost thresholds results in a limited time reduction.

In this paper, we propose a new fast intra coding method based on the global and directional gradients to terminate the current CU splitting early and avoid performing the high-complexity RDO process for the next CU levels. The CU can be classified into a *non-split class,* where the CU is predicted with high accuracy at the current level, or a *split class,* when there is no intra mode to effectively predict the CU. By terminating the splitting process for the *non-split* class early, the encoder saves a considerable amount of time and computations since further splitting would require many RDO computations to find the optimum splitting pattern. Further, it increases the signalling cost while providing little prediction improvement. *Non-split* CUs include smooth blocks, which can be predicted using the *dc* or *planar* modes at the current depth, and non-smooth blocks, when the directional gradient along the best angular mode demonstrates an accurate prediction by that mode.

The rest of this paper is organized as follows. Section II introduces the CU splitting and the mode decision in HEVC. In Section III, the proposed fast method for HEVC intra coding is presented. Experimental results are shown in Section IV, and finally, Section V concludes the paper.

This work was funded by Vantrix Corporation and by the Natural Sciences and Engineering Research Council of Canada under the Collaborative Research and Development Program (NSERC-CRD 428942-11).

**Fig. 1:** Splitting of a frame into CUs and PUs using the HEVC quadtree structure with the *RaceHorses* sequence

## II. CU SPLITTING AND MODE DECISION IN HEVC

HEVC uses a quadtree structure based on the CTU to efficiently encode video sequences with various kinds of visual textures. The CTU is the root of the tree and is considered as the largest CU. It is set to 64×64 by default in the HEVC reference implementation HM [13]. Each CU, including the CTU, is split recursively into four equal size CUs. This adaptive splitting approach is especially useful for higher resolution videos (4k×2k (4K), 8k×4k (8K), etc.). The depth of the 64×64 CU is 0 and the maximum depth is 3, i.e., the size of the smallest CU is 8×8. To perform the prediction with intra coding, each CU is considered as a PU, except for the 8×8 CU, which could be predicted as an 8×8 PU or four 4×4 PUs [14]. Fig. 1 shows how a frame is split into CUs and PUs using the quadtree structure.

HEVC employs 35 modes for intra prediction, including the *dc* mode, which is used to predict homogeneous regions, and the *planar* mode, which is used to produce smooth sample surfaces and angular modes to predict directional blocks. This richness of modes (shown in Fig. 2) allows the efficient exploitation of the spatial correlation across a frame. Each PU inside a CU has its own best mode, which is the mode with the lowest RDO cost, which is computed as follows [15]:

$$Cost_{RDO} = (SSE_{luma} + \omega_{chroma} \times SSE_{chroma}) + \lambda_{Mode} \times R_{Mode} \quad (1)$$

where SSE (sum of squared errors) is the distortion between the original and reconstructed blocks for both *luma* and *chroma* components, $\omega_{chroma}$ is the *chroma* weight, $\lambda_{Mode}$ is the Lagrange multiplier, and $R_{Mode}$ is the number of bits required for coding the block.

The HEVC reference implementation, HM, applies an exhaustive approach, where all possibilities are tested, to find the optimum splitting pattern and best modes for CUs and PUs, based on the RDO cost. This approach requires a lot of time, but achieves the best coding efficiency. In the next section, we propose a method to improve this implementation, as well as other state-of-the art methods, by reducing the encoding time and computational complexity.

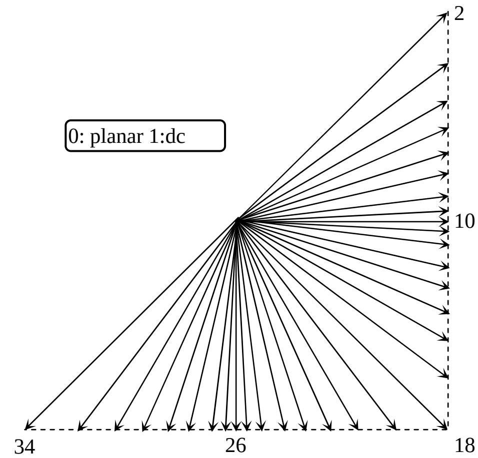


**Fig. 2:** HEVC intra angular prediction modes

## III. PROPOSED METHOD

The general idea behind CU splitting early termination is determining the splitting decision of the current CU by means that are less complex than performing the RDO process for the next depths. This is a binary classification problem in which $C=\{C_S, C_N\}$ is the set of the classes. $C_S$ and $C_N$ are the classes of *split* and *non-split* CUs, respectively. A classification problem is solved based on a feature vector $\boldsymbol{F} = \{f_1, f_2, f_3 \ldots\}$, in which the features need to be relevant and independent. Selecting relevant features can highly improve the CU classification process. The relevance between features and classes can be assessed using *mutual information (MI)* [16], defined as follows for two random variables *F* and *C*:

$$MI(F;C) = \int_C \int_F p(F,C) \log \frac{p(F,C)}{p(F)p(C)} dFdC \quad (2)$$

where $p(F,C)$ is the joint *probability density function (pdf)* of *F* and *C*. *MI* shows the shared information between *F* and *C*, i.e., it represents the amount of uncertainty that is reduced for *C* when *F* is known. Mathematically, the *MI* of *n* features as random variables and the selected class is expressed as:

$$\begin{aligned} & MI(F_1, F_2, \ldots F_n, C) \\ & = h(F_1, F_2, \ldots F_n) - h(F_1, F_2, \ldots F_n \mid C) \\ & = \iint p_{fc}(f_1, f_2, \ldots f_n, c) \log \frac{p_{fc}(f_1, f_2, \ldots f_n, c)}{p_f(f_1, f_2, \ldots f_n) p_c(c)} dfdc \end{aligned} \quad (3)$$

where $p_f(f_1, f_2, \ldots f_n)$, $p_c(c)$ and $p_{fc}(f_1, f_2, \ldots f_n, c)$ are the joint *pdf* of the *n* features, *pdf* of the selected class and the joint *pdf* of the features and the selected class, respectively. The function *h* shows the Shannon entropy. Since the encoder performs the intra prediction by using the neighbouring pixels at different directions, the features which are based on the CU's edges and CU's gradient maximize the mutual information between features and classes. Based on our extensive experiments with different video sequences, we selected two gradient-based features, which result in an excellent trade-off between computational complexity and coding efficiency.

### A. Splitting Early Termination by Global Gradient

When a high degree of pixel variation is present at the current level, the CU is split into four smaller CUs to allow it to perform prediction efficiently. The global gradient is an appropriate measure for evaluating this variation. The *mean of gradient amplitudes (MGA)* is given by:

$$MGA = \frac{1}{n}\sum_{i}\sum_{j}\left|G_X(i,j)\right| + \left|G_Y(i,j)\right| \quad (4)$$

where the sum is over all $n$ pixels across the CU. To compute the gradient components $G_X$ and $G_Y$ at each pixel, the *Sobel* operator is applied as a simple and low-complexity edge detector with 3×3 convolution masks. The sum of the absolute values of the vertical and horizontal components is used as an approximation of the gradient amplitude to avoid a resource-demanding *square root* operation. The CUs with larger *MGA* tend to split while those with smaller *MGA* are predicted at the current level. Fig. 3 shows the discrimination ability of this feature. It should be noted that when the *quantization parameter (QP)* is higher, the CU selects larger sizes as an optimum solution. From various simulations, we could observe that in the *MGA* graph, the effect of *QP* to discriminate *split* and *non-split* regions performed well when assumed linear. In view of this, to decide whether the CU splitting at the current level can be terminated, a measure is proposed as:

$$f_1 = \frac{MGA}{\alpha} - QP \quad (5)$$

where $\alpha$ is a coefficient assigned for each block size. The CU is of the *non-split* class if $f_1 < Th_1$ where the CU splitting process is terminated. Otherwise, we check whether the CU could be predicted effectively by an angular mode.

### B. Splitting Early Termination by Directional Gradient

If the global gradient of the CU is large, the CU cannot be predicted by the *dc* or *planar* modes at the current depth, but it may still be predicted by an angular mode provided that the directional gradient along the mode is small. To evaluate this gradient, the *mean of directional gradient amplitudes (MDGA)* is given as:

$$MDGA = \frac{1}{n}\sum_{i}\sum_{j}\left(\left|G_X(i,j)\right| + \left|G_Y(i,j)\right|\right)\times\cos\theta(i,j) \quad (6)$$

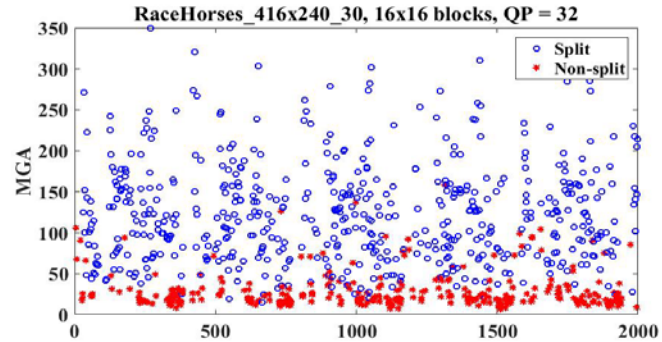


**Fig. 3:** *Split* and *non-split* CUs based on *MGA* feature

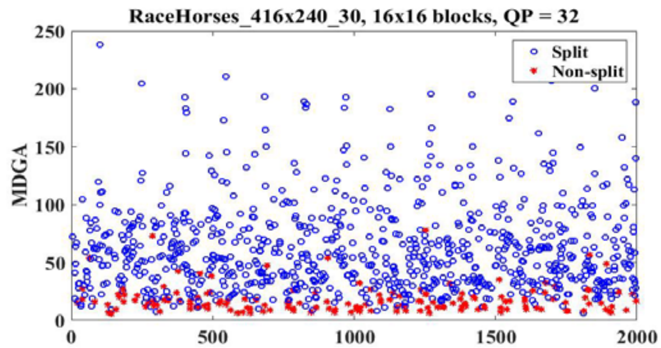


**Fig. 4:** *Split* and *non-split* CUs based on *MDGA* feature

where, again, the sum is over all $n$ pixels across the CU, $G_X$ and $G_Y$ are gradient components and $\theta$ is the angle between the gradient at each pixel and the best angular mode of the current CU. CUs with large *MGA* but small *MDGA* along the best angular mode can be predicted at the current level effectively. Fig. 4 shows the discrimination ability of the *MDGA*. To use the *MDGA* as a tool for classifying the CU, we need to first determine the best angular mode $m_{best\text{-}angular}$ of the CU. To that end, we utilize our previously proposed low-complexity intra mode decision approach [7] based on improved edge detection, consideration of most relevant modes from neighbouring blocks, and classification of SATD (sum of absolute transformed differences) costs permitting the elimination of several candidate modes prior to the RDO process. Once the best angular mode is known, the *MDGA* could be computed along this mode. Considering a linear effect of *QP*, as well as the block size, we have the following measure as the second feature:

$$f_2 = \frac{MDGA}{\beta} - QP \quad (7)$$

where $\beta$ is a coefficient assigned for each block size. The CU is of the *non-split* class if $f_2 < Th_2$, where the CU splitting process is terminated. Otherwise, the CU is split and the entire algorithm is repeated for the four newborn CUs.

Table I. Experimental results for various video sequences compared to HM 15.0

| | | Proposed Method | | | Shi et al. [9] | | | Oztekin et al. [10] | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Class*** | ***Video Sequences*** | ***TR(%)*** | ***BD-RATE(%)*** | ***BD-PSNR(dB)*** | ***TR(%)*** | ***BD-RATE(%)*** | ***BD-PSNR(dB)*** | ***TR(%)*** | ***BD-RATE(%)*** | ***BD-PSNR(dB)*** |
| A | Traffic | -52.5 | 1.30 | -0.061 | - | - | - | -46.17 | 5.0 | -0.09 |
| | PeopleOnStreet | -47.8 | 1.29 | -0.063 | -36.0 | 1.58 | -0.058 | -37.56 | 5.0 | -0.06 |
| B | Cactus | -50.8 | 1.49 | -0.049 | - | - | - | - | - | - |
| | Kimono | -59.1 | 0.85 | -0.028 | -40.5 | 0.18 | -0.018 | -60.13 | 3.5 | -0.04 |
| | ParkScene | -49.7 | 0.67 | -0.026 | -41.7 | 0.39 | -0.73 | -50.63 | 3.7 | -0.08 |
| | BasketballDrive | -62.2 | 2.35 | -0.059 | - | - | - | - | - | - |
| | BQTerrace | -50.1 | 0.87 | -0.041 | -37.4 | 1.89 | -0.054 | - | - | - |
| C | BQMall | -49.6 | 1.75 | -0.091 | -40.5 | 1.11 | -0.072 | -27.52 | 6.7 | -0.12 |
| | PartyScene | -39.2 | 1.14 | -0.076 | - | - | - | -28.67 | 5.6 | -0.14 |
| | RaceHorsesC | -47.2 | 0.76 | -0.043 | - | - | - | - | - | - |
| | BasketballDrill | -48.4 | 1.39 | -0.063 | -37.5 | 1.81 | -0.061 | - | - | - |
| D | RaceHorses(Training) | -41.7 | 1.15 | -0.065 | -27.0 | 1.31 | -0.051 | -29.47 | 4.5 | -0.11 |
| | BasketballPass | -54.1 | 2.07 | -0.110 | - | - | - | - | - | - |
| | BlowingBubbles | -40.5 | 1.00 | -0.052 | -25.6 | 1.61 | -0.077 | -29.37 | 3.6 | -0.09 |
| | BQSquare | -39.2 | 1.41 | -0.103 | - | - | - | - | - | - |
| E | Vidyo1 | -63.8 | 2.07 | -0.091 | -46.5 | 3.08 | -0.114 | - | - | - |
| | Vidyo3 | -63.8 | 2.94 | -0.141 | - | - | - | - | - | - |
| | Vidyo4 | -64.3 | 2.19 | -0.088 | -48.3 | 2.84 | -0.101 | - | - | - |
| Average | | -51.9 | 1.50 | -0.070 | -37.1 | 1.46 | -0.063 | -38.7 | 4.7 | -0.10 |

## IV. EXPERIMENTAL RESULTS

The proposed method is implemented in the HEVC test model HM15.0 using a PC equipped with an Intel® Core™ i7-4790 CPU @ 3.60 GHz and 32 GB of RAM. The first 100 frames of the test sequences recommended in [17] are used to conduct the experiments for the *all-intra* configuration and the *Main profile*. Parameter $\alpha$ is set to 1, 0.9, 0.4 and 0.3 for block sizes of 8×8, 16×16, 32×32 and 64×64, respectively, and parameter $\beta$ is set to 0.8, 0.7, 0.2 and 0.1 for the same block sizes. We select the smaller parameters for larger block sizes since the latter tend to split more than the smaller blocks. $Th_1$ and $Th_2$ are set to -5 and 0, respectively. By changing these thresholds, the proposed algorithm can provide an interesting trade-off between computational complexity and coding efficiency, which makes it a very flexible method for different applications. To obtain these thresholds, in our experiments, we used the *RaceHorses* sequence as a training sequence, and as a result, this sequence is excluded from the average results. To improve the results, in our next works, we are going to use an online training procedure to obtain parameters $\alpha$, $\beta$, $Th_1$ and $Th_2$ based on the frame characteristics. In this way the performance would improve as we would have higher time reduction and less quality loss. Bjontegaard delta bitrates (BD-RATE) and Bjontegaard delta peak signal-to-noise-ratios (BD-PSNR) [18] are used to compare our proposed method with the anchor HM15.0 by setting *QP* to 22, 27, 32 and 37.

The results are presented in Table I as the time reduction (TR), BD-RATE and BD-PSNR compared to HM 15.0. Further, a comparison is made between the proposed method and those in [9] and [10]. As the table shows, our proposed method provides a 52% time reduction, with a 0.07 dB quality loss, while [9] and [10] achieve time reductions of 37.1% and 38.7% with quality losses of 0.063 dB and 0.10 dB, respectively. The achieved time reduction is a very good trade-off for the negligible quality loss. While the loss is almost same compared to [9], the time reduction is higher and compared to [10], we have better results for both time reduction and quality loss.

## V. CONCLUSION

In this paper, we have proposed a CU splitting early termination algorithm for fast HEVC intra coding. The method is based on global and directional gradients. The global gradient determines whether the CU is smooth enough to be predicted by *dc* or *planar* modes at the current depth. The directional gradient is used for non-smooth CUs and classifies a CU as a *non-split* CU when it could be predicted efficiently by an angular mode at the current level; thus, the CU splitting process is terminated. The proposed method benefits from some controllable parameters, which allow a flexible trade-off between the compression rate and the encoder complexity. Based on the simulation results, on average, a 51.9% encoding time reduction is achieved using the proposed method, as compared to the reference HM 15.0, with a small BD-PSNR of 0.07 dB and 1.5% of BD-RATE.

## REFERENCES


[1] B. Bross, W. J. Han, J. R. Ohm, G. J. Sullivan, Y. K. Wang and T. Wiegand, “High Efficiency Video Coding (HEVC) Text Specification Draft 10”, *Document JCTVC-L1003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC),* 2013.

[2] G. J. Sullivan, J. R. Ohm, W. J. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard", *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 22, no. 12, 2012.

[3] T. Wiegand, G. J. Sullivan, G. Bjøntegaard and A. Luthra, "Overview of the H.264/AVC Video Coding Standard", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, 2003.

[4] L. Gao, S. Dong, W. Wang, R. Wang, W. Gao, "Fast Intra Mode Decision Algorithm Based on Refinement in HEVC", in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS),* 2015.

[5] L. Wang and W. Siu, "Novel Adaptive Algorithm for Intra Prediction with Compromised Modes Skipping and Signaling Processes in HEVC", *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 23, no. 10, 2013.

[6] M. Park and J. Jeong, "Fast HEVC Intra Prediction Algorithm with Enhanced Intra Mode Grouping Based on Edge Detection in Transform Domain", *Journal of Advances in Computer Networks,* vol. 3, no. 2, 2015.

[7] M. Jamali, S. Coulombe and F. Caron, "Fast HEVC Intra Mode Decision Based on Edge Detection and SATD Costs Classification", *in Proc. of the IEEE Data Compression Conference,* 2015.

[8] X. Shang, G. Wang, T. Fan and Y. Li, "Fast CU Size Decision and PU Mode Decision Algorithm in HEVC Intra Coding", in *Proc. of the IEEE International Conference on Image Processing (ICIP),* 2015.

[9] W. Shi, X. Jiang, T. Song and T. Shimamoto, "Edge Detector Based Fast Level Decision Algorithm for Intra Prediction of HEVC", *Journal of Signal Processing,* vol. 19, no. 2, 2015.

[10] A. Oztekin, E. Ercelebi, "An Early Split and Skip Algorithm for Fast Intra CU Selection in HEVC", *Journal of Real-Time Image Processing,* Springer, 2015.

[11] L. Shen, Z. Liu, X. Zhang, W. Zhao and Z. Zhang, "An Effective CU Size Decision Method for HEVC Encoders", *IEEE Transactions on Multimedia,* vol. 15, no. 2, 2013.

[12] L. Xiaobo, P. Zongju, J. Gangyi, Y. Mei and C. Fen, "CU Splitting Early Termination Algorithm in HEVC Intra Coding", *Journal of Signal Processing Research,* vol. 3, 2014.

[13] HEVC Test Model Software 15.0, https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/branches

[14] J. Lainema, F. Bossen, W. Han, J. Min and K. Ugur, "Intra Coding of the HEVC Standard", *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 22, no.12, 2012.

[15] V. Sze, M. Budagavi and G. J. Sullivan, "High Efficiency Video Coding (HEVC) Algorithms and Architectures", *Springer International Publishing, Switzerland,* 2014.

[16] E. Enriquez, A. Moreno, M. Pellon and F. Maria, "A Two-Level Classification-Based Approach to Inter Mode Decision in H.264/AVC", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 11, 2011.

[17] F. Bossen, "Common Test Conditions and Software Reference Configurations", *JCTVC-L1100*, 12th Meeting, Geneva, 2013.

[18] G. Bjontegaard, "Calculation of Average PSNR Differences Between R-D Curves", *VCEG-M33, Tech. Rep*, 2001.