Fast Intra Mode Decision Algorithm for HEVC Based on Block Textural Properties*

Ban Doan^[0000-0003-0900-6284] and Andrey Tropchenko^[000-0001-9812-7947]

ITMO University, Saint Petersburg, 197101, Russian Federation {bandoan, aatropchenko}@itmo.ru

Abstract. In order to achieve greater coding efficiency compared with the previous video coding standards, various advanced coding techniques are used in the High Efficiency Video Coding (HEVC) standard, such as a flexible partition and a large number of intra prediction modes. However, these techniques lead to much greater complexity that restricts HEVC from realtime applications. To solve this problem, a fast intra mode decision algorithm is proposed in this paper that uses the block's textural properties to determine the partition depth range and decide whether to split or skip smaller sizes of the coding unit. Besides that, the number of candidate modes for the rough mode decision process is also reduced depending on the block's property. Experimental results for the recommended test sequences by the JCT-VC show that the proposed algorithm can save an average of 44% encoder time with a slight loss in performance compared to the reference software HM-16.20.

Keywords: HEVC/H.265, Video Compression, Mode Decision, Intra-frame Prediction, Coding Tree Unit.

1 Introduction

The rapid development of social networks and the exchange of multimedia content sets many new requirements for transmission and storage systems. In terms of video compression, although H.264/AVC (Advanced Video Coding) [1] does a pretty good job of delivering compressed video to users, it has reached its limit of modernization, and new standards need to be developed to replace it. The international video coding standard H.265/HEVC (High Efficiency Video Coding) developed in 2013 [2] can reduce bitrate by 50% while ensuring equivalent quality compared to its predecessor H.264 [3]. This effectiveness is obtained through adaptive coding techniques, such as flexible partition structures, advanced and adaptive predictions, improved deblocking filter, and entropy coding. On the other hand, this advantage comes with highly increased computational complexity, so the development of complexity reduction schemes for HEVC is one of the most crucial research topics in the field of video compression.

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

^{*} Publication financially supported by RFBR grant 18-01-00569

2 B. Doan and A. Tropchenko

This paper focuses on improving the intra-frame coding - one important process in HEVC by early deciding to split the block and select the best suitable candidate modes. The rest of this paper is organized as follows. In Section 2, the intra-frame coding is briefly introduced. The proposed algorithm is presented in Section 3. Section 4 provides the experimental results, and Section 5 concludes this paper.

2 Intra prediction in HEVC

Instead of using macroblock as in H.264, HEVC uses the coding tree unit (CTU). Using the quad-tree partitioning technology, CTU can be split into coding units (CU) - the basic unit of separation, which has sizes from 64×64 to 8×8 , corresponding to a depth of 0 to 3 (Fig. 1a). Based on the partition depth, CU can be recursively split into four equal CUs. The optimal CTU partition is determined using the rate-distortion cost (RDcost). In intra-frame coding, each CU is divided into prediction units (PU), which will select the optimal predictions from 35 modes, including Planar mode (mode 0 - for a slowly changing area), DC mode (mode 1 - more suitable for a homogeneous area) and 33 angular modes (mode 2 - 34 for different texture directions) (Fig. 1b). Thus, the efficiency of intra coding increases, but with a significant computational load, even though the official reference software HM [4] uses a fast three-step encoding algorithm [5–7] with a combination of rough mode decision (RMD) and rate-distortion optimization (RDO) processes (Fig. 2) since for each PU, the encoder must calculate the RDcost for all 35 modes, while the maximum number of PU in a CTU is 341.



Fig. 1. Example of splitting CTU into CUs and corresponding quad-tree (a) and intra modes (b) in HEVC

Recently, much work has been published that offers various solutions to reduce the computational complexity of the intra mode decision process in HEVC. In general, they can be divided into two main types: optimization of the partitioning process of CU or PU [8–14] and adaptive selection of candidate modes for PU [9–11, 13, 19]. In these works, the coding information such as RDcost, Hadamard cost [9–12], or textural analysis [8] is used to decide the partition size on CU or PU. Decision trees [15], Bayesian



Fig. 2. Intra mode decision in HEVC Test Model

decisions [16], or other machine learning techniques are also used to predict the unit sizes. The number of input modes for the prediction process can be reduced by utilizing the gradient histogram distribution [16], the correlation of the corresponding blocks in the adjacent frames [18], or statistical analysis of mode distribution [19]. Whatever technique is used, these approaches have shown their effectiveness when efficiently reduce the coding complexity.

Based on the analysis of existing works, this paper proposes a fast algorithm for intra-frame coding in HEVC, which allows making an early decision on CTU splitting and adaptively choosing candidates for the prediction process.

3 Proposed fast intra mode decision algorithm

3.1 CTU depth range determination

CU can have a size from 64×64 to 8×8 , with the intra PU varying from 64×64 to 4×4 . The flexible structure in HEVC shows its effectiveness since the encoder can use CU in different sizes for different areas in the frame. To analyze the correlation between the quad-tree partition and the textural features of a region, Fig. 3 shows an enlarged part of the frame of the "BasketballDrive" sequence encoded by the HM. It can be shown that in homogeneous areas or areas with a smooth texture, CUs are always selected in large sizes, while small CUs are suitable for more complex areas with rich details. The optimal CU is determined for each area by checking all possible sizes, in other words, all separation depths. This process can cause redundancy and increase computational complexity since in some homogeneous areas, there is no need to check the small sizes of the CU. If the optimal CU size or partition depth range can be decided at the early stage, the wasteful search process of intra mode will be skipped, and thus a considerable amount of complexity can be significantly reduced in HEVC.

Based on the relationship between CU size and texture complexity, an algorithm will be proposed for quickly determining the depth of the CTU, according to which the encoder can collapse the search area to reduce the computational burden of the partitioning process.

The texture complexity of an area or block is strongly related to its samples' luminance value. The degree of variation in the luminance value in an $N \times N$ block is determined as follows:

$$L_{N \times N} = \frac{1}{N \times N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left(Y(i,j) - Y_a(i,j) \right)^2, \tag{1}$$

4 B. Doan and A. Tropchenko



Fig. 3. Quad-tree partition of the test sequence "BasketballDrive"

where Y(i, j) and $Y_a(i, j)$ are the luminance values at (i, j) and the average luminance value of the current block, respectively. $L_{N \times N}$ represents the different luminance parameter of the current block. The more sharply the different luminance values change, the larger value of $L_{N \times N}$ is obtained.

The value of $L_{N\times N}$ changes very quickly when the luminance value in the block changes. According to $L_{N\times N}$, we classify the block into homogeneous, middle texture, and complex texture. The encoder may terminate the split if the block is homogeneous. Otherwise, a split decision may be made early for blocks with the complex textural property.

$$\begin{cases} L \le P_1 & \text{homogeneous block;} \\ P_1 < L < P_2 & \text{middle texture block;} \\ L \ge P_2 & \text{complex texture block,} \end{cases}$$
(2)

where P_1, P_2 are lower and upper thresholds.

In our algorithm, to avoid calculating complexity for all block sizes, 16×16 CU is selected as the base block. A 32×32 CU is considered complex if at least one of the four 16×16 CUs inside it is complex. Similarly for 64×64 CU. In other cases, the texture of the 32×32 CU is determined by the value $L_{16 \times 16}^{Max}$ following Eq. 3, while Eqs. 1 and 2 are used for 32×32 CU, and the textural property of the block of size 64×64 is obtained from $L_{32 \times 32}^{Max}$.

$$\begin{cases} L_{N\times N}^{Max} \leq P_1 & \text{homogeneous block;} \\ P_1 < L_{N\times N}^{Max} < P_2 & \text{middle texture block;} \\ L_{N\times N}^{Max} \geq P_2 & \text{complex texture block,} \end{cases}$$
(3)

where $L_{N \times N}^{Max} = max\{L_{N \times N}^k\}, k = 0, ..., 3.$

Since a split decision was made for the CU, the depth of the CTU can be determined in advance, so that the encoder can skip unnecessary depths. Let D_{min} and D_{max} denote the minimum and maximum depth of the current CTU:

$$D_{min} = \begin{cases} 0 & \text{if } CU_{64\times64} \text{ homogeneous;} \\ 1 & \text{if } \exists CU_{32\times32} \text{ middle texture;} \\ 2 & \text{otherwise.} \end{cases}$$
(4)
$$D_{max} = \begin{cases} 0 & \text{if } CU_{64\times64} \text{ homogeneous;} \\ 1 & \text{if } 4 CU_{32\times32} \text{ homogeneous;} \\ 2 & \text{if } 16 CU \text{ homogeneous;} \end{cases}$$
(5)

$$\begin{array}{l}
2 & \text{if } 16 \ CU_{16 \times 16} \text{ homogeneous;} \\
3 & \text{if } \exists \text{ complex texture block } CU_{16 \times 16}.
\end{array}$$

Thus, based on texture complexity, the early decision to split or terminate splitting is made at the CU level, while the optimal depth range is made for the CTU. As a result, the computational cost of determining the optimal block is reduced.

3.2 Adaptive candidate modes selection

Statistical analysis of common test sequences shows that the optimal modes' distributions are not similar for all 35 intra modes. In most of the cases, modes 0 (Planar), 1 (DC), 10 (horizontal mode), and 26 (vertical mode) have larger proportions than other modes, where mode 0 has the highest proportion, and mode 1 has a second proportion, as many PU have relatively smooth textures after CU partition and most of them choose mode 0 or 1 as the optimum mode. Since the textural feature of most sequences has clear vertical or horizontal directions, the proportions of modes 10 and 26 are higher than other angle modes. Furthermore, the adjacent modes to those modes are relatively higher, while the proportions of the rest are insignificant. Hence, the number of input modes for intra prediction can be reduced, focusing on high probability modes to cut-off the complexity of the prediction process.

The original intra mode selection in HM will be modified as follows: instead of all 35, at the first stage of the prediction process, only most probability modes and angular modes in equal distances (*List1*) are tested using the RMD process to find modes with the lowest cost. After that, one more rough mode decision step (RMD2) was added to check a maximum of 4 more modes surrounding the best results of the first RMD. As a result, the minimum and the maximum number of modes that need to calculate the Hadamard costs are 11 and 15.

$$List1 = \{0, 1, 2, 6, 10, 14, 18, 22, 26, 30, 34\}.$$
 (6)

The block's texture properties can also be used to further reduce the total number of candidate modes for intra prediction. Let classify PUs into three categories depending on their sizes: low complex (homogeneous) (PU 64×64), middle complex (PU 32×32 , PU 16×16), and complex (PU 8×8 , PU 4×4). Table 1 shows the distribution of optimal modes for PU in low and middle complex groups. It can be seen that for homogeneous

6 B. Doan and A. Tropchenko

Homogeneous		Middle texture		
Mode	Proportion, %	Mode	Proportion, %	
0 and 1	82	0 and 1	42	
10	7	10 and its adjacent	28	
26	5	26 and its adjacent	20	
Other	6	Other	10	

Table 1. Distributions of optimum modes in different PUs

or middle complex PU, the number of candidate modes can be reduced to save the computational load.

Candidate modes for different PUs are given in Table 2, where for blocks of high complexity (PU 8×8 and PU 4×4) we will leave as in Eq. 6.

Table 2. Candidate modes for various block types

Block type	Candidates
Homogeneous	0, 1, 10, 26
Middle texture	0, 1, 6, 10, 14, 22, 26, 30
Complex texture	0, 1, 2, 6, 10, 14, 18, 22, 26, 30, 34

The overall algorithm, including early determination of the CTU depth range, early decision of the CU splitting and adaptive selection of candidate modes, is as follows:

Step 1. Classify CUs into a homogeneous, middle, and complex texture.

Step 2. Predict the depth range $[D_{min}, D_{max}]$ based on the texture properties of the CUs.

Step 3. Select a candidate list for PU in different areas. Calculate RMD.

Step 4. Calculate RMD2 for additional modes. Add most probable modes (MPM). **Step 5.** Calculate the RDO. Choose the best mode.

4 Experimental results

The proposed algorithm was implemented in HM-16.20, and several standard test sequences of different categories were selected for testing. All frames were encoded with the All Intra-Main configuration and four values of the quantization parameter (QP): 22, 27, 32, 37 [20]. All experiments are carried on macOS Mojave version 10.14.x with Intel[©] CoreTM i5-5257U @ 2.7GHz and 8GB of RAM.

The proposed algorithm's effectiveness is estimated using two metrics: Bjontegaard delta bitrate (BD-BR) and Bjontegaard delta peak-signal-to-noise (BD-PSNR) [21]. Encoding time-saving (ΔT) is used to evaluate computational complexity.

The comparative results of the proposed algorithm and the original HM-16.20 are shown in Table 3. The proposal can reach up to 63% of time-saving, while BD-BR increased on average by 2.54%, and the loss of BD-PSNR was 0.074dB.

Classes	Sequences	BD-BR, %	BD-PSNR, dB	$\Delta T, \%$
A	Traffic	1.58	-0.056	39.21
(2560×1600)	PeopleOnStreet	2.12	-0.061	33.31
	Kimono	0.62	-0.010	60.23
P	ParkScene	0.14	-0.046	49.58
(1020×1080)	Cactus	2.74	-0.054	37.65
(1920 × 1080)	BasketballDrive	1.71	-0.024	30.43
	BQTerrace	4.51	-0.057	35.16
	BasketballDrill	3.46	-0.077	60.82
C	BQMall	2.56	-0.088	32.00
(832×480)	PartyScene	1.31	-0.128	23.75
	RaceHorses	1.56	-0.073	31.08
	BasketballPass	4.07	-0.089	56.20
D	BQSquare	3.09	-0.145	63.72
(416×240)	BlowingBubbles	2.11	-0.109	46.53
	RaceHorses	2.69	-0.091	32.62
Е	FourPeople	4.36	-0.089	51.30
$\begin{bmatrix} \mathbf{E} \\ (1280 \times 720) \end{bmatrix}$	Johnny	3.79	-0.056	60.50
(1200×720)	KristenAndSara	3.41	-0.073	55.03
Average		2.54	-0.074	44.39

 Table 3. Comparison of coding efficiency between proposed intra prediction algorithm and HM-16.20

5 Conclusion

In order to reduce the computational complexity of the intra mode decision process in HEVC, this paper proposes a fast algorithm that uses the correlation between the textural property of an area and the CTU partitioning process. The variation degree of the luminance value is used to evaluate a block, from which the encoder can make a decision on CU splitting and determine the CTU depth range. Also, based on the PU sizes, the number of candidate modes for the selection process is significantly reduced. Experimental data show that the proposed algorithm can save about 44% of the encoding time while maintaining almost the same performance.

References

- H.264: Advanced video coding for generic audiovisual services, https://www.itu.int/rec/T-REC-H.264. Last accessed July 2020.
- Bross, B., Han, W.J., Ohm, J.R., Sullivan, G.J., Wang, Y.K., Wiegand, T.: High Efficiency Video Coding (HEVC) text specification draft 10 (for FDIS & Consent). Document JCTVC-L1003_v30, JCT-VC. Geneva, CH (2013).
- Ohm, J.R., Sullivan, G.J., Schwarz, H., Tan, T.K., Wiegand, T.: Comparison of the coding efficiency of video coding standards—including high efficiency video coding (HEVC). IEEE Transactions on Circuits and Systems for Video Technology, 22(12), 1669–1684 (2012) https://doi.org/10.1109/TCSVT.2012.2221192

- 8 B. Doan and A. Tropchenko
- 4. HEVC reference software, https://hevc.hhi.fraunhofer.de/. Last accessed July 2020.
- Lainema, J., Bossen, F., Han, W.-J., Min, J., Ugur, K.: Intra Coding of the HEVC Standard. IEEE Transactions on Circuits and Systems for Video Technology, 22(12), 1792–1802 (2012) https://doi.org/10.1109/TCSVT.2012.2221525
- Hosseini, E., Pakdaman, F., Hashemi, M.R., Ghanbari, M.: A computationally scalable fast intra coding scheme for HEVC video encoder. Multimed Tools Appl, 78(9), 11607–11630 (2019) https://doi.org/10.1007/s11042-018-6713-y
- 7. Piao, Y., Min, J., Chen, J.: Encoder improvement of unified intra prediction. Document JCTVC-C207, JCT-VC. Guangzhou, CN (2010).
- Min, B., Cheung, R.C.: A fast CU size decision algorithm for the HEVC intra encoder. IEEE Transactions on Circuits and Systems for Video Technology, 25(5), 892–896 (2014).
- Chen, Z.Y., Chang, P.C.: Rough mode cost-based fast intra coding for high-efficiency video coding. Journal of Visual Communication and Image Representation, 43, 77–88 (2017)
- Lim, K., Lee, J., Kim, S., Lee, S.: Fast PU skip and split termination algorithm for HEVC intra prediction. IEEE Transactions on Circuits and Systems for Video Technology, 25(8), 1335–1346 (2014).
- Yang, M., Grecos, C.: Fast intra encoding decisions for high efficiency video coding standard. Journal of Real-Time Image Processing, 13(4), 797–806 (2017).
- Gu, J., Tang, M. and Wen, J.: SATD based fast intra prediction for HEVC. 2017 Data Compression Conference (DCC) 442–442 (2017).
- Sun, X., Chen, X., Xu, Y., Xiao, Y., Wang, Y., Yu, D.: Fast CU size and prediction mode decision algorithm for HEVC based on direction variance. Journal of Real-Time Image Processing, 16(5), 1731–1744 (2019).
- Zhang, J., Kwong, S., Wang, X.: Two-stage fast inter CU decision for HEVC based on bayesian method and conditional random fields. IEEE Transactions on Circuits and Systems for Video Technology, 28(11), 3223–3235 (2017).
- Ruiz, D., Fernández-Escribano, G., Martínez, J.L., Cuenca, P.: 2019. A unified architecture for fast HEVC intra-prediction coding. Journal of Real-Time Image Processing 16(5), 1825– 1844 (2019).
- Hu, N. and Yang, E.H.: Fast mode selection for HEVC intra-frame coding with entropy coding refinement based on a transparent composite model. IEEE Transactions on Circuits and Systems for Video Technology 25(9), 1521–1532 (2015).
- Jiang, W., Ma, H., Chen, Y.: Gradient based fast mode decision algorithm for intra prediction in HEVC. 2nd international conference on consumer electronics, communications and networks (CECNet), 1836–1840 (2012).
- Motra, A.S., Gupta, A., Shukla, M., Bansal, P.: Fast intra mode decision for HEVC video encoder. 20th International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2012), 1–5 (2012).
- Doan, B., Tropchenko A.: Fast Intra Mode Decision for HEVC. 11th Majorov International Conference on Software Engineering and Computer Systems (MICSECS 2019). Saint-Petersburg, RU (2019).
- Bossen, F.: Common test conditions and software reference configurations. Document JCTVC-L1100, JCT-VC. Geneva, CH (2013).
- Bjontegaard, G.: Calculation of average PSNR differences between RD-curves. Document VCEG-M33, ITU-T. Austin, Texas (2001).