A REVIEW ON CODING UNIT DECISION IN THE CONTEXT OF HIGH EIFFICIENCY VIDEO CODING (HEVC)

Tasweer Ahmad
Nudrat Liaqat
Abdul Qadir
Syed Asghar Ali Shah
M. Umair Umer Lodhi
M. Ahsan Javaid
M. Imran Umar Lodhi

Abstract—

The High Efficiency Video Coding (HEVC) standard devised with the aim of providing double compression compared to its predecessors H.264 and MPEG-4 standards. This is primarily due to the variable sizes of blocks, called Coded Tree Unit (CTU) structure and numerous new prediction modes. However, this improved compression efficiency accompanies high computational complexity, which is primarily due to the recursive nature of the Coding Tree Unit structure and complex Rate-Distortion (RD) Optimization of Prediction Units (PU).

Keywords—Coding Tree Unit (CTU), Prediction Unit (PU), Transform Unit (TU)

1. Introduction

High Efficiency Video Coding (HEVC) is the latest video coding standard developed by Joint Collaborative Team on Video Coding (JCT-VC) established by the ISO and ITU-T on January 2013, which aims to achieve significant improvements in coding efficiency compare to H.264 ad MPEG-4 [2].

HEVC employs a highly flexible quadtree coding block partitioning structure as opposed to a Macro Block (MB) structure that was used in the previous H.264 and MPEG coding standards.

HEVC logically divides the picture into CTUs (Coding Tree Unit) as can be seen in Figure 1. All the CTUs in a video sequence can have any size of: 64×64, 32×32, or 16×16. CTU – Coding Tree Unit is a logical unit that usually consists of three blocks, namely one luma (Y) and two chroma samples (Cb and Cr), and associated syntax elements. Each block is called CTB (Coding Tree Block) [2, 3].

Each CTB still has the same size as CTU – 64×64, 32×32, or 16×16. Depending on a part of video frame, however, CTB may be too big to decide whether we should perform inter-picture prediction or intra-picture prediction [6]. Thus, each CTB can be differently split into multiple CBs (Coding Blocks) and each CB becomes the decision making point of prediction type i.e. inter-picture or intra-picture prediction [6]. For example, some CTBs are split to 16×16 CBs while others are split to 8×8 CBs. HEVC supports CB size all the way from the same size as CTB to as small as 8×8 [4, 5]. More precisely, the prediction type is coded in CU (Coding Unit). The CU is always square and the size of a CU is from 8x8 up to the size of the CTU. CU consists of three CBs (Y, Cb, and Cr) and associated syntax elements.

Likewise, each CB can be split to PBs differently depending on the partition mode. Once the prediction is made, DCT-like transformation needs to be applied on the residual (difference between predicted image and actual image) image. This is done through Transform Block. Therefore, each CB can be differently split into TBs (Transform Block).

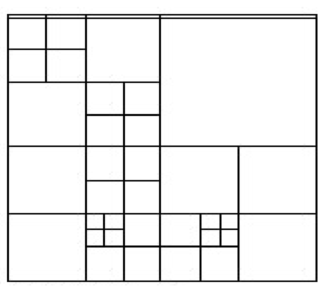

Figure 1. Example of CU Structure [5]

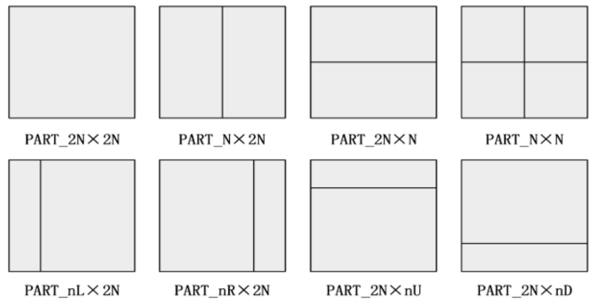

Figure 2. Partition modes for inter prediction unit [5]

Each CU may contain one or more PUs among the eight available prediction modes depending on partition mode mentioned in Figure 2. CU with PART_2Nx2N has one prediction unit while CU with PART_NxN has four prediction unit [7].

The most optimal selection of CU and PU is the peculiarity of HEVC encoder. One of possible solution is to incorporate Full search method but that appears too much computational complex. As can be seen below in Figure 3, the areas in the image that contain edges and details, the CTU is split into smaller units while in smooth areas CTU is not partitioned into smaller blocks.

Figure 3. CU representation of an image

The rest of this paper is organized as follows. A brief introduction of the concept of coding unit is presented in Section II. Then in Section III we give a detailed description of different fast coding unit selection approaches, and show the experimental results in Section IV. Section V concludes our work.

2. Review of Approaches

The HEVC encoder computes a large amount of RD (Rate Distortion) costs of various modes to select the best CU size, PU partition, and TU structure. At the same CU depth, the RD-cost of every prediction mode is calculated, and all the costs are compared to determine the best mode for the CU at this depth [6]. Next, the encoder compares the RD-costs of the best partitioned modes at different depths. The encoder spends a huge amount of computations on PUs and TUs in a CU quad-tree to identify the lowest RD cost [7]. Presently, a number of efficient methods to reduce the complexity are emerging.

A. CU-depth level decision

This section focuses on the recursive splitting and termination decisions of Coding Units. These Coding Units decisions are based on the observation that in temporal and spatial neighborhoods, the motion and texture characteristics of the picture patches are similar. The current CU depth can be predicted by checking the size of its neighbor CUs (spatial) and co-located CU (temporal). Figure 4 shows the relation between the referred neighboring CUs and the current coded CU. The co-located CU means that the previous frame CU has the same position as the current encoded CU. This algorithm executes recursively for the side length of CU equal to 64, 32, or 16 [4].

Splitting decision conditions are as follows:

- The co-located CU has smaller CUs.
- All neighboring CUs have smaller CUs.
- The current encoding frame is not I frame.

When the CU RD-analysis begins at the current depth and all the above conditions are satisfied, the PU mode search in the current depth will be skipped except for the 2N×N or N×2N inter modes, and then it jumps into the next depth directly.

The termination decision prevents the encoder from building a large tree with a lot of computational complexity owing to small CUs. The termination decision is determined by the following conditions:

- The co-located CU does not have any smaller CU.
- Three or more neighboring CUs do not have any smaller CU.
- The current encoding frame is not I frame.

When all the above conditions are satisfied, the mode selection process, whose depth is greater than the current depth, will not be conducted [4].

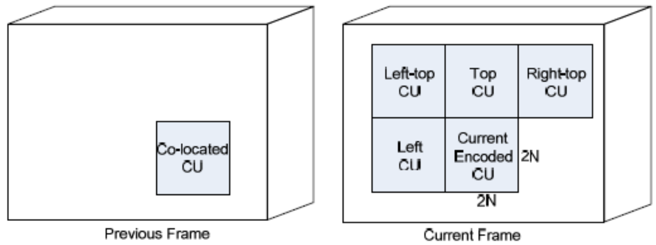

Figure 4. Reference CUs and the current encoded CU

Following techniques are used for CU depth selection

i) *Pixel Distribution Based Block Partitioning Algorithm*

In HEVC, each CU is partitioned into four sub-CUs called Coding Tree Units (CTUs), whereas Prediction Unit (PU) is used for the selection of Inter or Intra prediction. The encoder should compute the Rate Distortion (RD) cost for all possible CU and PU size that seems to be very time consuming for smooth areas which should not be further partitioned into smaller blocks. It is quite intuitive that in smooth areas, most pixel values are uniformly distributed. Therefore, a fast coding algorithm is proposed by highlighting smooth area of an image is mentioned in Figure 5:

1) Find out gray level distribution of pixels

a) Bin out the maximum range of pixels into 16 intervals, defined as

$$Qi = [16*i, 16*i+15], i = 0, 1\ldots15 \quad (1)$$

b) Define Ni as the number of pixels belonging to Qi in the current CU, then

$$Nmax = max (Ni) \quad , i= 0, 1, 2...15 \quad (2)$$

c) Scale is defined

$$Scale = Nmax/N \quad (3)$$

Where N defines max number of pixels in a CU.

2) Calculate the RD cost for 2Nx2N and skip mode

3) If Scale is not larger that threshold T1j(j = 0,1,2), calculate the RD cost for remaining inter PU

4) Calculate RD cost for Intra partition

5) Consider whether current CU needs to be further split if scale is not larger than threshold, otherwise skip it.

Threshold values are empirically found to be:

$$T1_j = \begin{cases} 0.70, & j = \{0,1\} \\ 0.65, & others \end{cases}$$

$$T2_j = \begin{cases} 0.80, & j = 0 \\ 0.75, & j = 1 \\ 0.70, & others \end{cases}$$

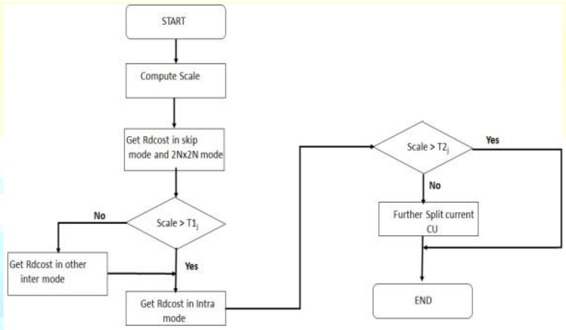

Figure 5. Flow Chart of Fast CU Depth Selection

HEVC supports 35 intra prediction modes, so a fast algorithm should be developed for Prediction Unit (PU) with optimal mode selection with minimum RD cost [4]. In HEVC, fast intra prediction has reduced its complexity drastically but its prediction mode selection is still very difficult process [4]. A fast intra-mode decision algorithm is developed by Jiajia et al. in [4], based on Hadamard transform which appears to be more accurate. The algorithm is discussed in Figure 6. The overall procedure of this algorithm is defined below:

1. define a Modeldx set it to be 0
2. Calculate J for all 35 modes and select two minimum J (J0 and J1)
3. If J0 is much smaller than J1, then choose dcost0 as best mode, otherwise, calculate RD cost of two candidate and chose smaller one as best mode

The formula for J is given as below:

$$J = \lambda_1.(D + J_{TU}) + \lambda_2.B_{mode} + N_{nz}$$

λ_1, λ_2 are weight factors where $\lambda_1 = 0.8$, while λ_2 is same as λ_{pred} mention in [9].

B_{mode} is bit cost for decision, N_{nz} specifies number of non-zero coefficients. J_{TU} is the cost for splitting TU. Then D depicts cost of current CU coded by current selected mode using SATD (Sum of Absolute Transform Differences).

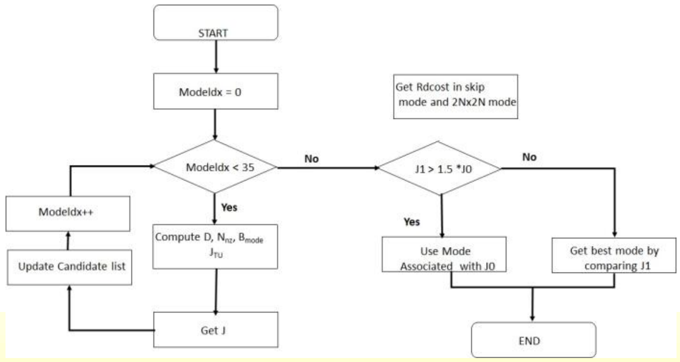

Figure 6. Fast Intra Mode Decision

It is depicted that proposed algorithm results in 45% time saving.

ii) Texture Based Fast CU Selection

In this technique, decsion of partition mode is made on the basis of texture of frame [8]. For plane texture, encoder chooses plane modes such as (PART_2Nx2N or PART_NxN) with one or four Prediction Units respectively [2]. The proposed fast algorithm is given:

- Divide 8-partition modes into two classes:
  A = {PART_2Nx2N, PART_NxN},
  B = {PART_2NxN, PART_Nx2N, PART_2NxnU, PART_2NxnD, PART_nLx2N, PART_nRx2N}
- Calculate the "Planeness" of the texture of CU with bi-linear interpolation
- If the value of planeness is less than threshold, the encoder skips the class B modes. The planeness of texture is calculated by

$$p = \frac{1}{[(\frac{n}{2}) - 2]^2} \sum_{x=1}^{(\frac{n}{2})-1} \sum_{y=1}^{(\frac{n}{2})-1} |org(x,y) - pred(x,y)|$$

If the value of p is less than threshold T_1, the encoder skips class B modes, where at a given depth threshold T_1 is given by

$$T_1[depth] = T_1[depth - 1] - k_2$$

Where depth is the depth of CU partitioning,

$$T_1[0] = k_3 * QP + k_4$$

QP is quantization parameter, empirically it is found that k_2 =5, k_3=1 and k_4=3.

Similarly, CU with same gradient directions are merged to into larger CU. CU merging algorithm is illustrated in Figure 7 given below. The algorithm is based on gradient direction, it computes gradient direction of smallest CU comprising of four (04) neighboring CUs of sizes 4x4, and merges them into one 8x8 CU. The gradient direction is computed by

$$grad\ (d) = \frac{1}{c}\sum_{x=0}^{n-1}\sum_{y=0}^{n-1} diff(d,x,y)$$

While

$$diff(d,x,y) = \sum_{i=1}^{x+ix_d<n,y+iy_d<n} |p(x,y) - p(x+ix_d, y+iy_d)|$$

Where d depicts direction in Table, n size of CU, c is the count number of calculation of diff. The encoder would merge smaller CU to larger one of grad (d) is smaller than threshold T2, given by

$$T_2[depth] = k_4 * T_2[depth-1]$$

And

$$T_2[0] = k_5 * QP + k_6$$

Where depth is CU merging, QP is quantization parameter, where k_4=k_5=0.5 and k_6 = 1.5. Fast CU and PU is discussed in Figure 8.

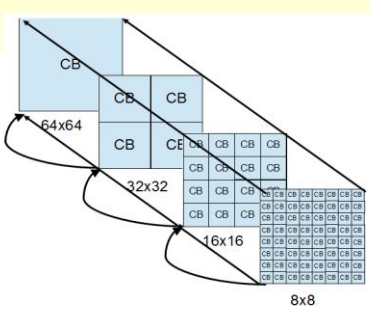

Figure 7. Sub-CU Merging into Large CU

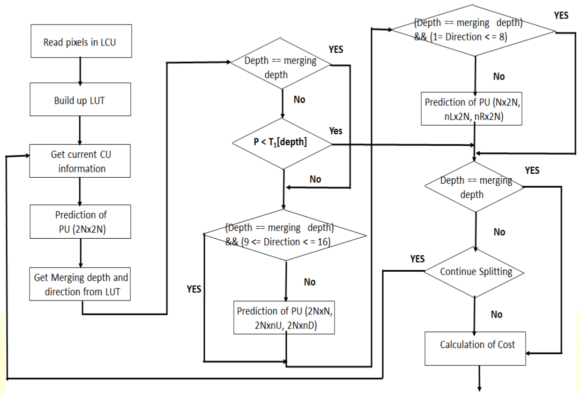

Figure 8. Fast CU and PU Selection

It is shown that the proposed algorithm has improved the average time saving of 52.8%.

iii) Early Termination Scheme for Fast Intra Prediction Algorithm

This algorithm employs Hadamard costs to readily decide intra mode process. Hadamard cost is computed by finding Sum of Absolute Transformed Difference (SATD). This algorithm includes three sub-algorithms

a) Early Termination of CU and TU Split
b) Early Termination of CU Cost Calculation

a) Early Termination of CU and TU Split

For a CU of size 2Nx2N, RD-cost is computed at current level n, then this RD-cost is compared with next level of four NxN sub-CUs. Then, CU with smallest RD-cost is selected as current level. The accumulated RD-cost after encoding k-th NxN sub-CU is given by

$$AccuC_{n+1,k} = \sum_{i=1}^{k} C_{n+1,i}$$

The early termination process would occur if following condition for predicted cost would hold

$$PredC_{n+1,k} > \alpha_k \, C_n$$

Where $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\} = \{1.5, 1.2, 1.1, 1\}$

The following condition is compared for CU splitting and early termination at (n+1)-th level

$$\min\left\{\frac{4}{k}, \quad \frac{AccuHC_{n+1,4}}{AccuHC_{n+1,k}}\right\} \sum_{i=1}^{k} C_{n+1,i} > \propto_k C_n$$

It depends on the value of k that a CU with n-th level 2Nx2N size is to be partitioned into (n+1)th level NxN sub-CU. If k=1, CU is not further divided into sub-CUs and splitting is skipped. On the other hand, if k=4, all four sub-CU are encoded.

TU splitting is same as that of CU splitting after CU size selection. TU size selection is decided on the basis of RD-cost by following relation

$$\frac{4}{k}\sum_{i=1}^{k} C_{n+1}, i > \beta_k C_n$$

Where $\{\beta_1, \beta_2, \beta_3, \beta_4\} = \{1.2, 1, 1, 1\}$

b) Early Termination of CU Cost Calculation

Hadamard Cost computed in previous section is modified to jump on to next CU level. Let $HC_{n,i}$ denote the Hadamard cost of each sub-CU within a CU. Scaled Mean Absolute Deviation (SMAD) is given by relation

$$MAD = \frac{\sum_{i=1}^{4} |HC_{ni} - \mu_{HC_n}|}{4\mu_{HC_n}}$$

Where μ_{HCn} represents mean of $HC_{n,I}$. If is found that if $SMAD_n > \gamma$, CU would be further divided into sub-CU otherwise it would not be partitioned.

// terminate CU splitting
For k = 1, 2, 3, 4
PredC = 0
for i = 1, …., k
PredC = PredC + $C_{n+1,i}$
End
rho = min {4/k, $AccuHC_{n+1,4}$ / $AccuHC_{n,k}$}
If PredC * rho > α_k * C_n
	// jump to the next CU depth //level calculate the variance of
	//Hadamard cost at CU level n
	
	If $\sigma_n > \gamma$
	Skip CU level n
	Jump to CU level n+1
	end

Break; end end	
// terminate TU splitting For k = 1, 2, 3, 4 PredC = 0 for i = 1, …., k PredC = PredC + $C_{n+1,i}$ End rho = 4/k If PredC * rho > β_k * C_n Break; end end	

3. Experimental Results and Comparison

The experimental results of given techniques are compared among each other. It is shown that Early termination CU selection technique which is to-date, appears to save 68% reduction in encoding time with a just 1.4% BD-rate increase. While Textual and Pixel Distribution techniques still performs well with 53.7% and 45.33% time reduction, while 3.7% and 4.32% BD-rate increase respectively.

Method	BD-Rate (%)	ΔT (%)
Pixel Distribution	4.32	-45.33
Textual based CU Selection	3.7	-53.7
SVM Based	1.3	-44.8
Early Termination CU Selection	1.4	-68

Table 1. CU Selection Techniques Comparison

4. Conclusion

In this paper, Coding Unit size selection approaches for HEVC are presented. The proposed algorithm employs important and computational-friendly features to help making a precise and fast selection on CU size. Our current work mainly focuses on the CU size selection, and the decision on TU size and mode selection are worthy of further investigations.

5. Acknowledgments

The authors would like to acknowledge the contribution of Dr. Nadeem Khan whose great mentorship was expedient for the project.

6. References

[1] Muhammad Usman Karim Khan, Muhammad Shafique, Jörg Henkel , "An adaptive complexity reduction scheme with fast Prediction Unit Decision For HEVC Intra Encoding," Karlsruhe Institute of Technology (KIT) , Germany, pp. 1578-1582, September 2013.

[2] Jiawen Qiu, Fan Liang*, Yonglin Luo , "A fast coding unit selection algorithm for HEVC", School of Information Science and Technology, Sun Yat-sen University, Guangzhou, July 2013.

[3] Wei-Jhe Hsu; Hsueh-Ming Hang, "Fast coding unit decision algorithm for HEVC," *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific* , vol., no., pp.1,5, Oct. 29 2013-Nov. 1 2013

[4] Jiajia Lu; Fan Liang; Liwei Xie; Yonglin Luo, "A fast block partition algorithm for HEVC," *Information, Communications and Signal Processing (ICICS) 2013 9th International Conference on* , vol., no., pp.1,5, 10-13 Dec. 2013

[5] Jiawen Qiu; Fan Liang; Yonglin Luo, "A fast coding unit selection algorithm for HEVC," *Multimedia and Expo Workshops (ICMEW), 2013 IEEE International Conference on* , vol., no., pp.1,5, 15-19 July 2013

[6] Hao Zhang; Zhan Ma, "Fast Intra Mode Decision for High Efficiency Video Coding (HEVC)," *Circuits and Systems for Video Technology, IEEE Transactions on* , vol.24, no.4, pp.660,668, April 2014

[7] Xiaolin Shen and Lu Yu "CU Splitting Early Termination Based on Weighted SVM" EURASIP Journal on Image and Video Processing *Vol*: 20 *No*: 1, January, 2013

[8] Xiaolin Shen; Lu Yu; Jie Chen, "Fast coding unit size selection for HEVC based on Bayesian decision rule," *Picture Coding Symposium (PCS), 2012* , vol., no., pp.453,456, 7-9 May 2012

[9] Il-Koo Kim, Ken McCann, Kazuo Sugimoto, Benjamin Bross, Woo-Jin Han.High Efficiency Video Coding (HEVC) Test Model 9 Encoder Description[S]. Shanghai, CN, 10–19 October 2012.

[10] http://codesequoia.wordpress.com/2012/10/28/hevc-ctu-cu-ctb-cb-pb-and-tb/