Bayesian Rule Based Fast TU Depth Decision Algorithm for High Efficiency Video Coding

Xiuzhe Wu [#§], Hanli Wang [#§*], Zhihua Wei [#§]

[#] *Department of Computer Science and Technology, Tongji University, Shanghai, China*

[§] *Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai, China*

314_wxz@tongji.edu.cn, hanliwang@tongji.edu.cn, zhihua_wei@tongji.edu.cn

***Abstract*—The latest video coding standard high efficiency video coding (HEVC) has made a significant progress in compression efficiency than previous standard H.264/advanced video coding (AVC) while it has led to a tremendous increase in encoding computations. Recently, a Bayesian model based transform unit (TU) depth decision approach has been designed to accelerate TU depth decision, which requires numerous variance computations. In this work, a novel relevant feature based Bayesian model is proposed for fast TU depth decision. Experimental results demonstrate that the best performance is achieved while the depths of upper TU, left TU and co-located TU are all taken into considerations. Moreover, as compared with previous research, the proposed algorithm reduces much more encoding computations while keeping the video quality and compression efficiency more or less intact.**

***Index Terms*—High efficiency video coding, transform unit, depth decision, reference features, Bayesian rule.**

I. Introduction

In order to provide better multimedia services and applications based on high definition video and ultra high definition video, the joint collaborative team on video coding (JCT-VC) has released the latest video coding standard called high efficiency video coding (HEVC) [1]. Due to a variety of advanced techniques such as the quadtree structure [2], angular intra prediction [3], sample adaptive offset (SAO) [4], the compression efficiency of HEVC is almost twice as much as that of the previous standard H.264/advanced video coding (AVC) [5] while those advanced strategies also result in enormous encoding computations.

Since the block partition process [2] accounts for plenty of encoding computations, a number of researches are designed to reduce the computational efforts in selecting the best encoding parameters for coding unit (CU) size (or depth), prediction unit (PU) encoding mode and transform unit (TU) depth. Yu *et al.* [6] propose an early termination algorithm for CU splitting which can skip checking unnecessary CU sizes. In [7], an all-zero block detection based method is developed to early decide the CU size. Ahn *et al.* [8] design a fast CU size decision algorithm by utilizing the information of SAO, motion vector and PU/TU size belonging to the reference frame and/or spatial neighbors. Teng *et al.* [9] define two TU processes denoted as 'TU Split' and 'TU Merge' for TU depth decision.

Meanwhile, early termination based on the Bayesian rule becomes popular in recent years. An early split test and an pruning test [10] have been developed for CU depth decision according to the Bayesian decision rule based on the rate distortion (RD) cost. In [11], Fang *et al.* applies different decision mechanisms for different regions of feature space to improve the Bayesian based classification for CU depth decision. In [12], the variance of the prediction errors by mode Inter$2N\times 2N$, the percentage of coefficient cost and the variance of sum of absolute transformed difference (SATD) of four sub-CUs are employed as reference features and then the risk threshold is derived by the Bayesian rule for CU depth decision. Similar to [12], Shen *et al.* [13] applies the same mechanism on TU splitting and the feature to be considered is the variance of residual coefficients. Although it achieves a significant acceleration, the calculation of variances brings extra computations as well.

In this work, the Bayesian rule is also applied as the basis of our proposed TU fast depth decision algorithm. Different from Shen's studies [13], we employ the correlations between a TU and its adjacent TUs, which are used as referential features for TU depth decision. As a result, the redundant computational complexity brought by variance computing in [13] is avoided. Moreover, in order to achieve a better coding efficiency, the performances of five alternative strategies are compared, which are all based on the Bayesian rule whereas the referred information is varied. The referential information includes the depth of the upper TU, left TU and co-located TU. Experiments show that when the depth of the upper TU, left TU and co-located TU are all regarded as references, the encoding performance reaches a peak. As compared to Shen's algorithm [13], our algorithm performs better with 20.36% in total time saving and 57.20% in TU splitting time saving, while the corresponding savings of 16.74% and 46.77% are obtained in [13]. The rest of this paper is organized as follows. At first, the proposed Bayesian rule based TU depth decision algorithm is detailed in Section II. In Section III, the comparative experimental results are presented. Finally, Section IV concludes this work.

II. Proposed TU Depth Decision

A. Overview of Quadtree Structure in HEVC

The elementary coding unit in HEVC is called coding tree unit (CTU). Each CTU can be divided into four CUs averagely and each CU can be divided reversely into four sub-CUs as

*Corresponding author (H. Wang).

VCIP 2016, Nov. 27 – 30, 2016, Chengdu, China

well. The size of CU flexibly varies from 64×64 to 8×8 aiming at diminishing the negative impact of unbalanced energy distribution in image. Each CU is further subdivided into PUs for intra/inter prediction. Then, transform is applied to the prediction residual with a nested residual quadtree (RQT) for each CU, *i.e.,* each CU is partitioned into TUs. TU has four kinds of depth 0, 1, 2 and 3, which stands for four sizes of 32×32, 16×16, 8×8 and 4×4, respectively. In fact, there exist strong spatial and temporal correlations in TU depths which are utilized to derive the proposed Bayesian rule based TU depth decision algorithm.

B. TU Depth Decision based on Bayesian Rule

In HEVC, the TU splitting can be classified as a two-class problem which can be defined as W given below.

$$W = \{w_S, w_N\}, \tag{1}$$

where the class w_S represents the choice that the current TU is split into four sub-TUs while the class w_N stands for not splitting into four sub-TUs. Considering there is a strong relationship between the current TU depth and its spatial-temporal neighboring TU depths, the TU depth information including that of upper TU, left TU and co-located TU is utilized as the features for the classification problem in Eq. (1). Given the depths of the current TU, upper TU, left TU and co-located TU denoted as D_{cur}, D_{up}, D_{left} and D_{col}, the last three variables are included in the feature vector F as the candidate elements.

The conditional probability of splitting and not splitting into four sub-TUs given the occurrence of the value of F can be defined as $P(w_S|F)$ and $P(w_N|F)$, respectively. It will lead to an increase in RD cost when making a wrong decision, and this increment of RD cost is recorded as $C_{i,j}$ which means the cost of making w_i as the current decision while the theoretical truth is w_j, where $i, j \in \{S, N\}$. Based on the above analysis, we can have the following formula.

$$C_{S,N} = \frac{RD_S - RD_N}{RD_N},\ C_{N,S} = \frac{RD_N - RD_S}{RD_S}, \tag{2}$$

where RD_i, $i \in \{S, N\}$, stand for the RD costs of splitting and not splitting into four sub-TUs.

Likewise, the risks of splitting and not splitting decisions $Risk(w_i|F)$, $i \in \{S, N\}$, are defined as follows.

$$Risk(w_S|F) = C_{S,N} \times P(w_N|F) + C_{S,S} \times P(w_S|F), \tag{3}$$

$$Risk(w_N|F) = C_{N,N} \times P(w_N|F) + C_{N,S} \times P(w_S|F). \tag{4}$$

Obviously, right decisions do not cause additional costs, therefore, we have

$$C_{S,S} = C_{N,N} = 0. \tag{5}$$

As a result, the Bayesian risks are derived as

$$Risk(w_S|F) = C_{S,N} \times P(w_N|F), \tag{6}$$

$$Risk(w_N|F) = C_{N,S} \times P(w_S|F). \tag{7}$$

According to the Bayesian theorem, the following conditional probabilities can be obtained as

$$P(w_N|F) = \frac{P(F|w_N) \times P(w_N)}{P(F)}, \tag{8}$$

$$P(w_S|F) = \frac{P(F|w_S) \times P(w_S)}{P(F)}. \tag{9}$$

Before calculating the RD cost of the current TU depth and further calculating the related RD costs of sub-TU depth, the first step is to evaluate the risk. Here lists the fundamental rules: if $Risk(w_N|F) \geq Risk(w_S|F)$, it is necessary to perform splitting on the current TU; otherwise, the examination process of the best TU depth can be early terminated. Taking all conditions and formulae mentioned above into considerations, the rule can be finally determined as

$$\begin{cases} \text{Split}, & if\ \ \frac{P(F|w_S)}{P(F|w_N)} \geq \frac{C_{S,N} \times P(w_N)}{C_{N,S} \times P(w_S)} \\ \text{Not split}, & if\ \ \frac{P(F|w_S)}{P(F|w_N)} < \frac{C_{S,N} \times P(w_N)}{C_{N,S} \times P(w_S)} \end{cases}. \tag{10}$$

In this work, we not only evaluate the performance of the proposed TU fast decision algorithm, but also analyze the different kinds of reference features in F used in the algorithm. As aforementioned, a strong correlation exists in the depth of the current TU and the depths of its spatially and temporarily adjacent TUs, especially the upper TU (D_{up}), the left TU (D_{left}) and the co-located TU (D_{col}) in the reference frame. Hence, five kinds of features are employed for evaluation aiming at achieving the best performance. To be clear, the depth information set is listed as follows.

$$F = \{f_{up}, f_{left}, f_{col}, f_{u+l}, f_{u+l+c}\}, \tag{11}$$

$$D = \{D_{up}, D_{left}, D_{col}\}, \tag{12}$$

where f_{up} means that only D_{up} is utilized, and the same definition to f_{left} and f_{col}. f_{u+l} indicates that both D_{up} and D_{left} are considered while f_{u+l+c} includes all of three depth information in D are employed.

Algorithm 1 Proposed TU depth decision algorithm.

Step 1 Check Bayesian data table with the given resolution and Q_p, and obtain the current Bayesian threshold $\frac{C_{S,N} \times P(w_N)}{C_{N,S} \times P(w_S)}$ labeled by $thres$ for simple.

Step 2 For each TU, after the current depth calculation, set the feature vector and calculate $\frac{P(F|w_S)}{P(F|w_N)}$ labeled by $prob$ for simple.

Step 3 Compare $thres$ with $prob$ and make decision according to Eq. (10). If splitting is permitted, then go on with **Step 4**. Otherwise stop exploring, choose another unexamined TU as the current TU and go to **Step 2**.

Step 4 Set TU depth as the current TU depth + 1 and one of sub-TUs as the current TU, then go to **Step 2** circularly until all candidate TUs are examined.

C. Proposed Overall Algorithm

Based on the above analysis, the proposed TU depth decision algorithm is presented in Algorithm 1. Specifically, the proposed optimization process is only implemented on TUs with the size of 32×32, 16×16 and 8×8, since it is no need to split 4×4-sized TUs. More importantly, the Bayesian threshold $\frac{C_{S,N}\times P(w_N)}{C_{N,S}\times P(w_S)}$ is initialized by collecting the statistics of five video sequences standing for five video classes (*i.e.*, Traffic for Class A, ParkScene for Class B, BQMall for Class C, BasketballPass for Class D and FourPeople for Class E) and stored in the Bayesian data table. The video resolutions of Class A, B, C, D and E are 2560×1600, 1920×1080, 832×480, 416×240 and 1280×720, respectively. Moreover, the proposed algorithm will be implemented with each individual feature vector in F to evaluate the corresponding performance shown in Section III.

TABLE I
VIDEO SEQUENCES FOR EXPERIMENTS.

Class	Sequence	Lowercase
A	PeopleOnStreet	a
	Traffic	b
B	BQTerrace	c
	Cactus	d
	ParkScene	e
C	BasketballDrill	f
	BQMall	g
	PartyScene	h
	RaceHorses	i
D	BasketballPass	j
	BlowingBubbles	k
	BQSquare	l
	RaceHorses	m
E	FourPeople	n
	Johny	o
	KristenAndSara	p
	Vidyo1	q
	Vidyo3	r

III. EXPERIMENTAL RESULTS

In order to evaluate the proposed TU depth decision algorithm, the HEVC reference software test model version HM16.0 [14] is used to carry out experiments. The configuration of random access is applied. Eighteen HEVC benchmark video sequences as listed in Table I are employed which are refereed to as a∼ r for short. Four Q_p values including 22, 27, 32 and 37 are used in our experiments to evaluate the performance at different bit rates.

Three criteria are utilized to study the performances of the comparative algorithms, including video quality degradation, bit rate increase and encoding time saving. Regarding the video quality and bit rate, they are presented with BDPSNR (dB) and BDBR (%) according to [15], which are very useful measurements in evaluating the average difference of Peak Signal-to-Noise Ratio (PSNR) and bit rate between the original encoder and a test approach. As far as the encoding speedup is concerned, the entire encoding time saving (ΔT_e) and the encoding time saving in TU depth decision process (ΔT_t) are presented, which have the form as $\Delta T = \frac{T_o - T_a}{T_o}\times 100\%$, where T_a and T_o stand for the corresponding encoding time of the original HEVC encoder with and without using the test algorithms.

TABLE II
COMPARISON OF BDPSNR (DB) PERFORMANCE.

Seq	Shen	Up	Left	Col	U+L	U+L+C
a	-0.034	-0.054	-0.055	-0.056	-0.056	-0.056
b	-0.002	-0.011	-0.011	-0.010	-0.012	-0.011
c	-0.005	-0.021	-0.021	-0.019	-0.020	-0.020
d	-0.001	-0.006	-0.007	-0.008	-0.008	-0.007
e	-0.006	-0.022	-0.022	-0.023	-0.024	-0.026
f	-0.006	-0.014	-0.015	-0.011	-0.014	-0.010
g	-0.010	-0.042	-0.042	-0.044	-0.043	-0.043
h	-0.011	-0.074	-0.075	-0.076	-0.075	-0.073
i	-0.040	-0.073	-0.074	-0.075	-0.075	-0.072
j	0.003	-0.023	-0.023	-0.030	-0.030	-0.030
k	-0.005	-0.019	-0.020	-0.025	-0.024	-0.027
l	-0.002	-0.048	-0.048	-0.052	-0.051	-0.060
m	-0.047	-0.072	-0.073	-0.077	-0.071	-0.075
n	-0.004	-0.003	-0.004	-0.003	-0.002	-0.003
o	-0.001	-0.001	-0.002	-0.002	-0.002	-0.002
p	-0.001	-0.009	-0.009	-0.009	-0.009	-0.009
q	-0.004	-0.006	-0.006	-0.006	-0.007	-0.006
r	-0.014	-0.015	-0.016	-0.015	-0.015	-0.015
Avg	**-0.010**	**-0.029**	**-0.029**	**-0.030**	**-0.030**	**-0.030**

The comparative experimental results are shown in Tables II-V. Meanwhile, Shen's algorithm [13] is also implemented for comparison with the same experimental configuration. As seen from the results, it is obvious that the proposed algorithm with our features is better than Shen's algorithm [13] to reduce redundant HEVC encoding computations while keeping the video quality and compression efficiency almost intact. And the best choice of referential feature is the combination of depths of upper TU, left TU and co-located TU (*i.e.*, f_{u+l+c} in Eq. (11)).

IV. CONCLUSION

In this paper, the Bayesian theorem is applied to reduce redundant computations in HEVC to early decide the best TU depth. The correlations in depth between the current TU and spatial-temporal adjacent TUs are studied as well. The experimental results demonstrate that the proposed algorithm is able to reduce computations significantly while maintaining the video quality and compression ratio.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China under Grants 61472281 and 61622115, the "Shu Guang" project of Shanghai Municipal Education Commission and Shanghai Education Development Foundation under Grant 12SG23, the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (No. GZ2015005), and the

TABLE III
COMPARISON OF BDBR (%) PERFORMANCE.

Seq	Shen	Up	Left	Col	U+L	U+L+C
a	0.79	1.25	1.27	1.29	1.29	1.30
b	0.04	0.35	0.37	0.32	0.38	0.36
c	0.24	1.00	1.02	0.91	0.99	0.98
d	0.10	0.32	0.34	0.40	0.40	0.37
e	0.20	0.74	0.76	0.79	0.81	0.88
f	0.14	0.37	0.39	0.29	0.37	0.27
g	0.23	1.00	1.02	1.04	1.02	1.03
h	0.21	1.62	1.64	1.67	1.63	1.60
i	1.05	1.93	1.95	1.97	1.98	1.91
j	-0.10	0.48	0.50	0.64	0.66	0.64
k	0.14	0.51	0.53	0.65	0.63	0.71
l	0.05	0.99	1.01	1.07	1.04	1.25
m	0.99	1.54	1.56	1.64	1.54	1.61
n	0.10	0.09	0.11	0.08	0.06	0.08
o	0.04	0.08	0.10	0.13	0.09	0.13
p	0.05	0.33	0.35	0.32	0.33	0.32
q	0.10	0.20	0.22	0.21	0.22	0.21
r	0.42	0.45	0.47	0.44	0.46	0.44
Avg	**0.27**	**0.74**	**0.76**	**0.77**	**0.77**	**0.78**

TABLE IV
COMPARISON OF ENTIRE ENCODING TIME SAVING (ΔT_e, %).

Seq	Shen	Up	Left	Col	U+L	U+L+C
a	15.52	21.62	21.48	21.79	22.14	22.36
b	17.04	20.15	19.83	19.54	20.12	20.19
c	15.68	18.97	18.85	18.55	19.47	19.59
d	17.28	20.31	20.21	19.75	20.82	20.90
e	15.19	19.38	19.22	18.81	19.92	20.13
f	18.36	20.69	20.67	19.87	20.79	20.82
g	16.61	21.87	21.88	21.04	21.98	22.13
h	12.94	22.36	22.28	21.81	22.46	22.86
i	15.46	21.59	21.63	21.16	21.72	22.05
j	17.09	18.39	18.41	17.64	18.70	19.26
k	14.76	17.02	16.93	16.74	17.26	17.96
l	15.37	16.85	16.71	16.20	17.02	17.55
m	14.41	18.69	18.57	18.80	19.06	19.64
n	19.47	20.24	20.20	19.61	20.41	20.21
o	18.87	20.00	19.97	19.49	20.22	20.05
p	18.73	19.86	19.68	19.17	19.90	19.76
q	19.23	20.09	20.03	19.65	20.20	20.20
r	19.35	20.62	20.63	20.22	20.75	20.73
Avg	**16.74**	**19.93**	**19.84**	**19.44**	**20.16**	**20.36**

Fundamental Research Funds for the Central Universities under Grants 0800219315.

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 (for FDIS & last call)," *Doc.JCTVC-L1003_v34*, Jan. 2013.

[2] I.-K. Kim, J. Min, T. Lee, W.-J. Han, and J. Park, "Block partitioning structure in the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1697–1706, Dec. 2012.

TABLE V
COMPARISON OF TU SPLITTING TIME SAVING (ΔT_t, %).

Seq	Shen	Up	Left	Col	U+L	U+L+C
a	42.48	59.02	58.75	60.66	60.08	60.75
b	50.69	58.97	58.95	58.72	59.32	59.39
c	47.18	55.83	55.49	55.64	56.70	56.90
d	48.64	56.57	56.81	56.34	57.76	57.49
e	45.51	57.50	57.11	56.57	57.97	58.51
f	50.61	57.59	57.57	57.11	58.78	58.27
g	44.49	58.59	57.91	57.32	57.54	57.84
h	33.68	56.87	55.90	56.20	57.23	57.70
i	41.82	58.07	58.20	57.69	58.79	58.75
j	45.56	48.67	50.38	49.65	51.08	53.16
k	39.89	47.16	47.55	47.26	48.74	48.25
l	42.93	45.47	47.72	46.74	47.02	48.47
m	35.17	48.88	48.45	47.70	47.57	50.68
n	58.15	60.52	60.38	59.98	61.22	59.45
o	58.62	60.89	61.63	61.13	62.26	61.33
p	57.52	61.39	60.93	59.83	61.98	61.13
q	58.02	60.85	61.85	60.39	61.33	60.77
r	56.90	60.49	61.08	59.84	61.25	60.72
Avg	**47.66**	**56.30**	**56.48**	**56.04**	**57.03**	**57.20**

[3] J. Lainema, F. Bosen, W.-J. 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, Dec. 2012.

[4] C.-M. Fu, E. Alshina, A. Alshin, Y.-W. Huang, C.-Y. Chen, C.-Y. Tsai, C.-W. Hsu, S.-M. Lei, J.-H. Park, and W.-J. Han, "Sample adaptive offset in the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1755–1764, Dec. 2012.

[5] "Advanced video coding for generic audiovisual services," *ISO/IEC 14496-10:2005(E) ITU-T Rec. H.264 (E)*, Mar. 2005.

[6] Q. Yu, X. Zhang, S. Wang, and S. Ma, "Early termination of coding unit splitting for HEVC," in *APSIPA ASC'12*, Dec. 2012, pp. 1–4.

[7] P.-T. Chiang and T. S. Chang, "Fast zero block detection and early CU termination for HEVC video coding," in *ISCAS'13*, May 2013, pp. 1640–1643.

[8] S. Ahn, M. Kim, and S. Park, "Fast decision of CU partitioning based on SAO parameter, motion and PU/TU split information for HEVC," in *PCS'13*, Dec. 2013, pp. 113–116.

[9] S.-W. Teng, H.-M. Hang, and Y.-F. Chen, "Fast mode decision algorithm for residual quadtree coding in HEVC," in *VCIP'11*, Nov. 2011, pp. 1–4.

[10] S. Cho and M. Kim, "Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 9, pp. 1555–1564, Sept. 2013.

[11] X. Fang, X. Zhu, L. Yu, and X. Shen, "Fast HEVC intra coding unit size decision based on an improved Bayesian classification framework," in *PCS'13*, Dec. 2013, pp. 273–276.

[12] X. Shen, L. Yu, and J. Chen, "Fast coding unit size selection for HEVC based on bayesian decision rule," in *PCS'12*, May 2012, pp. 453–456.

[13] 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 Commun.*, vol. 32, pp. 121–128, Mar. 2015.

[14] "JCT-VC. (2012). Subversion repository for the HEVC test model version HM16.0. [Online] available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.0/."

[15] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," *Doc. VCEG-M33*, Apr. 2001.