EARLY TERMINATION FOR TZSEARCH IN HEVC MOTION ESTIMATION

Zhaoqing Pan[1] *Yun Zhang*[1,2] *Sam Kwong*[1,⋆] *Xu Wang*[1] *Long Xu*[1,3]

[1] Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong
[2] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
[3] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China
{zqpan3-c, xuwang6-c}@my.cityu.edu.hk, {cssamk, yunzhang, lxu222}@cityu.edu.hk

ABSTRACT

The TZSearch algorithm was adopted in the high efficiency video coding reference software HM as a fast Motion Estimation (ME) algorithm for its excellent performance in reducing ME time and maintaining a comparable Rate Distortion (RD) performance. However, the multiple initial search point decision and the hybrid block matching search contribute a relatively high computational complexity to TZSearch. In this paper, based on the statistical analysis of the probability of median predictor to be selected as the final best point in the large Coding Units (CUs) (64×64, 32×32) and small CUs (16×16, 8×8) as well as the center-biased characteristic of the final best search point in ME process, we propose two early terminations for TZSearch. Experimental results show that the proposed early terminations can achieve 38.96% encoding time saving, while the RD performance degradation is quite acceptable.

Index Terms— Motion Estimation, TZSearch, Early Termination, HEVC

1. INTRODUCTION

High Efficiency Video Coding (HEVC) is an emerging next generation video coding standard, which is proposed by Joint Collaborative Team on Video Coding (JCT-VC) under ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) [1, 2]. The main goal of HEVC is to achieve 50% bit rate saving when compared to the exiting video coding standards [3], while this coding gain comes at the cost of the high computational complexity of the advanced coding tools, such as flexible size representation, large and asymmetric motion partitions, advanced motion vector prediction, new intra prediction methods and so on. The high computational complexity of HEVC limits it to be used in real-time applications.

⋆ Corresponding author

This work was supported in part by the Natural Science Foundation of China under Grants 61272289, 61102088 and 61202242.

It can be regarded that the flexible size representation technique contributes the largest proportion of encoding time to HEVC encoder. In order to reduce the computational complexity of HEVC encoder, many works were focus on optimizing the flexible size representation technique. Based on the Bayesian decision rule, a fast Coding Unit (CU) size selection algorithm was proposed for HEVC [4]. Based on the Rate Distortion (RD) cost of previous skipped CUs, Kim *et al.* proposed an adaptive CU early termination algorithm for HEVC [5]. Based on the differential motion vector and coded block flag, an early determination of mode decision was proposed for HEVC [6]. Choi *et al.* proposed an early transform unit decision method for video encoding in HEVC, which is based on the number of nonzero DCT coefficients [7].

The computational complexity of selecting the best CU and prediction unit is highly correlated with the Motion Estimation (ME) algorithm. In order to fast ME, Kim *et al.* proposed a sum of absolute differences-based selective Bi-prediction method for HEVC [8]. In [9], a novel fast two step sub-pixel ME algorithm was proposed for HEVC. In HEVC reference software HM, the TZSearch is used as the fast integer-pixel ME algorithm. It uses the multiple initial search point decision to locate an initial search point and hybrid block matching search to find the best matching block. However, the computational complexity is still relatively high due to these two techniques used.

In order to further reduce the computational complexity of HEVC ME. In this paper, we propose two early termination strategies for TZSearch, which is based on a diamond search and a hexagon search. The rest of this paper is organized as follows. The motivations and statistical analyses are given in Section 2. Then, the details of the proposed early termination algorithm are presented in Section 3. Experiment results are shown in Section 4. Finally, Section 5 concludes this paper.

2. MOTIVATIONS AND STATISTICAL ANALYSES

The TZSearch algorithm was adopted in the HEVC reference software as a fast ME algorithm for its excellent performance in reducing ME time and maintaining a comparable RD performance. The search procedure of TZSearch includes two steps, initial search point decision and block matching search, respectively. The first step is to determine the initial search point by using a set of predictors which includes Median Predictor (MP) [10], Left Predictor (LP), Above Predictor (AP), Above-Right Predictor (ARP) and (0,0). LP, AP and ARP are corresponding to the Motion Vector (MV) of left, top, top-right block of the current block, respectively. After the initial search point is determined, the hybrid block matching search, including multiple diamond/square search and raster search, are used to locate the best matching block which is with the minimum RD cost. However, the computational complexity of multiple initial search point decision and hybrid block matching search is still relatively high. If these two processes can be simplified, much more encoding time will be saved.

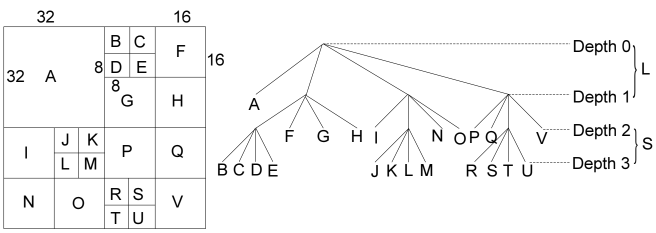

Fig. 1. An example of CU quad-tree structure in HEVC

In video sequences, there are a large number of blocks with static or quite slow motion activity. For these blocks, they have the largest probability to select the MP as the final best point in ME process. In addition, one of the most important innovation of HEVC is the quad-tree structure based CU. CU is the basic unit in HEVC encoding process, and it is defined by the Largest CU (LCU) and the hierarchical depth. Fig. 1 gives an example of the quad-tree structure based CU in HEVC, and its LCU and hierarchical depth equal to 64 and 4, respectively. One notably difference between HEVC and H.264/AVC is that HEVC supports larger CU, such as 64×64, 32×32 and so on. Hence, the final best search point in ME process for large CUs (the CUs in depth 0, 1, and denote as Depth L) and small CUs (the CUs in depth 2,3, and denote as Depth S) may have different characteristics. Four video sequences with various resolutions including BQMall (480P), Johnny (720P), Mobisode2 (240P) and Parkscene (1080P) are used to analyze the probability of selecting MP as the best point in TZSearch. The test conditions are tabulated in Table 1. Table 2 shows the statistical results of selecting MP as the best search point in CU with Depth L and S.

From Table 2, it can be observed that there are 45.41% to 67.41% MPs selected as the best point in CUs with Depth L. For CUs with Depth S, there are 56.39% to 89.03% MPs selected as the best point. There are 62.26% and 75.92% on average MPs selected as the best search point in CUs with Depth L and Depth S, respectively. If these MPs who are

Table 1. Test conditions

Search Range	64
CU size / depth	64 / 4
GOP size	4
Quantization Parameter (QP)	24, 28, 32,36
NO. of frames to be encoded	30

selected as the final best point can be early terminated, much more encoding time will be saved.

Table 2. Probability of selecting MP as the final best search point, Unit: (%)

CU	Sequence	24	28	32	36	Average
Depth L	BQMall	63.29	64.46	66.09	67.41	65.31
	Johnny	62.27	65.89	68.93	72.75	67.46
	Mobisode2	67.17	69.17	71.43	70.35	69.53
	Parkscene	45.41	45.89	46.97	48.69	46.74
	Average	59.54	61.35	63.36	64.80	**62.26**
Depth S	BQMall	72.42	74.67	77.33	79.30	75.93
	Johnny	74.18	76.34	78.38	81.60	77.63
	Mobisode2	87.58	89.03	88.50	86.92	88.01
	Parkscene	56.39	60.11	64.31	67.67	62.12
	Average	72.64	75.04	77.13	78.87	**75.92**

Another observation on Table 2, the probabilities of selecting MPs as the final best search point are different between CUs with Depth L and S. Compared to CUs with Depth S, the probability decreases about 13% in CUs with Depth L. Hence, when we optimize the TZSearch by using the MP, different methods should be considered for these two CUs.

3. PROPOSED EARLY TERMINATION ALGORITHM

It is well known that the best point of ME is center-biased [11]. Based on this characteristic, a diamond search with four points is employed to terminate early the ME process of CUs with Depth S. The candidates of the diamond search are defined as

$$\mathbf{MV}_{\mathrm{diamond}} = \{(\mathrm{MV_x}, \mathrm{MV_y}) | (\mathrm{MV_x}, \mathrm{MV_y}) = (\mathrm{x} \pm 1, \mathrm{y}), (\mathrm{x}, \mathrm{y} \pm 1)\}, \quad (1)$$

where (x, y) denotes the MV of the MP. Then the best point among the five points is determined as

$$\mathbf{MV}_{\mathrm{best}} = \arg\min\{J(\mathbf{MV}_{\mathrm{candidate}}, \lambda_{\mathrm{MOTION}})\}, \quad (2)$$

where J represents the RD cost function [12], $\mathbf{MV}_{\mathrm{candidate}} \in \{(\mathrm{x,y}), \mathbf{MV}_{\mathrm{diamond}}\}$. The TZSearch is terminated if

$$\mathbf{MV}_{\mathrm{best}} = \{(\mathrm{MV_x}, \mathrm{MV_y}) | (\mathrm{MV_x}, \mathrm{MV_y}) = (\mathrm{x,y})\}. \quad (3)$$

In addition, if the early termination is not satisfied, the $\mathbf{MV}_{\mathrm{best}}$ is set as an initial search point candidate for TZSearch.

For early termination of ME of CUs with Depth L, a diamond search and a hexagon search with six points are jointly used to detect whether MP is the final best search point. The candidates of the hexagon search are defined as

$$\mathbf{MV}_{\mathrm{hexagon}} = \{(\mathrm{MV_x}, \mathrm{MV_y})|(\mathrm{MV_x}, \mathrm{MV_y}) = (\mathrm{x} \pm 1, \mathrm{y} \pm 1), (\mathrm{x} \pm 2, \mathrm{y})\}, \quad (4)$$

where (x,y) represents the MV of the MP. Then the best point among the 11 candidate search points is determined as Eq. (2), and the $\mathbf{MV}_{\mathrm{candidate}} \in \{(\mathrm{x,y}), \mathbf{MV}_{\mathrm{diamond}}, \mathbf{MV}_{\mathrm{hexagon}}\}$. The TZSearch will be terminated if the $\mathbf{MV}_{\mathrm{best}}$ is satisfied to Eq. (3). Moreover, if the early termination is not satisfied, the $\mathbf{MV}_{\mathrm{best}}$ is set an initial search point candidate for TZSearch.

Table 3. Hit rate of the proposed early termination algorithm

HR	Sequence	24	28	32	36	*Average*
$\mathrm{HR_L}$	BQMall	97.91	98.10	98.21	98.02	98.06
	Johnny	97.04	97.79	98.09	98.50	97.86
	Mobisode2	97.67	96.66	96.89	97.08	97.08
	Parksence	95.59	96.24	96.07	96.51	96.10
	Average	97.05	97.20	97.32	97.53	**97.27**
$\mathrm{HR_S}$	BQMall	97.44	97.70	98.00	98.26	97.85
	Johnny	98.44	98.81	99.11	99.34	98.93
	Mobisode2	98.76	98.98	99.30	99.50	99.14
	Parksence	95.86	96.99	97.25	97.97	97.02
	Average	97.63	98.12	98.42	98.77	**98.23**

Table 4. Reject rate of the proposed early termination algorithm

RR	Sequence	24	28	32	36	*Average*
$\mathrm{RR_L}$	BQMall	19.45	18.49	17.04	16.26	17.81
	Johnny	24.21	22.73	20.43	18.26	21.41
	Mobisode2	12.98	10.99	11.70	12.78	12.11
	Parksence	22.31	20.91	19.02	17.31	19.89
	Average	19.74	18.28	17.05	16.15	**17.80**
$\mathrm{RR_S}$	BQMall	12.58	12.03	11.87	11.33	11.95
	Johnny	16.00	16.51	14.76	13.02	15.07
	Mobisode2	6.22	5.90	8.32	7.23	6.92
	Parksence	14.23	13.92	12.83	11.56	13.14
	Average	12.26	12.09	11.95	10.79	**11.77**

In order to evaluate the efficiency of the proposed early termination algorithm, the Hit Rate (HR) and Reject Rate (RR) are adopted. HR and RR indicate the accuracy and complexity of the early termination, respectively, and they are defined as

$$\begin{cases} \mathrm{P_{HR}}(\mathbf{A}|\mathbf{B}) = \dfrac{\mathrm{N}(\mathbf{A}|\mathbf{B})}{\mathrm{N}(\mathbf{B})} \times 100\% \ (\%), \\ \mathrm{Q_{RR}}(\mathbf{A}|\mathbf{B}) = \dfrac{N(\mathbf{C}|\mathbf{B})}{\mathrm{N}(\mathbf{B})} \times 100 \ \% \ (\%), \end{cases} \quad (5)$$

where $\mathrm{P_{HR}}(\mathbf{A}|\mathbf{B})$ and $\mathrm{Q_{RR}}(\mathbf{A}|\mathbf{B})$ denote the HR and RR, respectively; $N(\cdot)$ represents the number of total blocks of corresponding event, and the event **A** represents the MP is selected as the final best search point; **B** denotes the early termination condition; **C** be the event that the best MV is selected from the candidate MV of diamond/hexagon search. $\mathbf{A}|\mathbf{B}$ and $\mathbf{C}|\mathbf{B}$ are two conditional events. Four video sequences are used to test, and the test conditions are listed in Table 1. The detailed HR and RR are tabulated in Tables 3 and 4, where $\mathrm{HR_L}$ represents the HR of early TZSearch termination of CUs with Depth L, and $\mathrm{HR_S}$ denotes the HR of early TZSearch termination of CUs with Depth S; $\mathrm{RR_L}$ and $\mathrm{RR_S}$ represent the RR of CUs with Depth L and S, respectively.

From Table 3, it can be observed that the HR of the proposed early termination algorithm can achieve from 95.59% to 98.50%, 97.27% on average for CUs with Depth L, and from 95.86% to 99.50%, 98.23% on average for CUs with Depth S. These values demonstrate that the proposed algorithm can terminate early TZSearch accurately.

From Table 4, it can be observed that the RR of the proposed early terminated TZSearch algorithm is from 10.99% to 24.21%, 17.80% on average for CUs with Depth L, and from 5.90% to 16.51%, 11.77% on average for CUs with Depth S. Another observation from Table 4 is that the RR of CUs with Depth S is smaller than CUs with Depth L, this is because CUs with Depth L have large probability to select the best MV from the MVs which are far from the MP. These values demonstrate that the proposed early termination algorithm works efficiently, and most of the best point with MP can be early determined.

Based on above analysis, the proposed early terminated TZSearch algorithm is summarized and illustrated step-by-step as follows.

Step 1. Perform a diamond search on the MP, then compute the $\mathbf{MV}_{\mathrm{best}}$ according to Eq. (2), and the $\mathbf{MV}_{\mathrm{candidate}} \in \{(\mathrm{x,y}),\ (\mathrm{x} \pm 1, \mathrm{y}),\ (\mathrm{x}, \mathrm{y} \pm 1)\}$.

Step 2. If current CU belongs to depth 2 or 3, go to Step 3; else, go to Step 4.

Step 3. If Eq. (3) is true, go to Step 7; else, go to Step 6.

Step 4. Perform a hexagon search on the MP, then compute the $\mathbf{MV}_{\mathrm{best}}$ according to Eq. (2), and the $\mathbf{MV}_{\mathrm{candidate}} \in \{(\mathrm{x} \pm 1, \mathrm{y} \pm 1), (\mathrm{x} \pm 2, \mathrm{y})\}$.

Step 5. If Eq. (3) is true, go to Step 7; else, go to Step 6.

Step 6. Set the $\mathbf{MV}_{\mathrm{best}}$ as one initial search point candidate, then perform the original TZSearch for current CU.

Step 7. Return the best MV, and process the ME for the next block.

4. EXPERIMENTAL RESULTS

In order to evaluate the efficiency of the proposed algorithm, HEVC reference software HM8.0 is used as the software platform. The test conditions are listed in Table 1. The hardware platform is Intel Core 2 Duo CPU E5800 @ 3.16GHz and 3.17GHz, 4.00GB RAM with Microsoft Windows 7 64-bit operating system.

Table 5. Summary of encoding results, TZSearch vs Proposed Early terminated TZSearch

Sequence	QP	$\triangle$PSNR	$\triangle$BR	TS	Sequence	QP	$\triangle$PSNR	$\triangle$BR	TS
Mobisode2 (240P)	24	-0.025	0.21	-40.28	SlideShow (720P)	24	0.072	-1.32	-31.89
	28	0.024	0.85	-42.61		28	0.008	1.16	-34.49
	32	-0.031	1.97	-38.54		32	-0.015	1.14	-32.98
	36	-0.022	0.87	-33.75		36	-0.138	1.30	-38.19
	Average	-0.013	0.98	-38.79		Average	-0.018	0.57	-34.39
	BDPSNR/BDBR	-0.057/1.36				BDPSNR/BDBR	-0.073/0.94		
Flowervase (480P)	24	-0.015	-0.12	-31.34	Cactus (1080P)	24	-0.001	-0.11	-33.23
	28	0.001	-0.12	-33.49		28	-0.002	0.08	-37.82
	32	-0.012	0.31	-40.17		32	-0.002	0.08	-39.96
	36	0.011	-0.65	-37.95		36	0.001	0.56	-42.16
	Average	-0.003	-0.15	-35.74		Average	-0.001	0.15	-38.29
	BDPSNR/BDBR	-0.002/0.10				BDPSNR/BDBR	-0.005/0.18		
FourPeople (720P)	24	-0.008	-0.30	-38.22	Johnny (720P)	24	-0.005	0.22	-41.74
	28	-0.005	0.15	-42.34		28	-0.003	0.08	-42.78
	32	0.001	0.25	-46.00		32	-0.017	-0.41	-44.21
	36	0.002	0.60	-45.60		36	-0.019	-0.05	-45.31
	Average	-0.003	0.18	-43.04		Average	-0.011	-0.04	-43.51
	BDPSNR/BDBR	-0.012/0.26				BDPSNR/BDBR	-0.006/0.22		
Average PSNR/BR/TS		**-0.008/0.28/-38.96**							
Average BDPSNR/BDBR		**-0.026/0.51**							

We compare the coding performance of the proposed algorithm with the original TZSearch in terms of Peak-Signal-to-Noise Ratio (PSNR), Bit Rate (BR) and total encoding CPU time. The low delay configure file is used. The original TZSearch uses all initial search point predictors and refinement methods. Six video test sequences are used to compare the encoding performance. The experiment results are summarized in Table 5, where $\triangle$PSNR and $\triangle$BR represent the PSNR loss and BR increase, respectively; TS denotes the total encoding CPU time saving ratio. They are defined as

$$\begin{cases} \triangle\mathrm{PSNR} = \mathrm{PSNR_p} - \mathrm{PSNR_o}\ (\mathrm{dB}), \\ \triangle\mathrm{BR} = \dfrac{\mathrm{BR_p} - \mathrm{BR_o}}{\mathrm{BR_o}} \times 100\%\ (\%), \\ \mathrm{TS} = \dfrac{\mathrm{T_p} - \mathrm{T_o}}{\mathrm{T_o}} \times 100\%\ (\%), \end{cases} \tag{6}$$

where the subscript p represents the proposed algorithm; o denotes the original TZSearch algorithm in HM8.0. BDPSNR and BDBR are computed according to [13].

From Table 5, it can be observed that the total encoding time can be saved from 31.89% to 46.00%, 38.96% on average; meanwhile, the average PSNR degrades from 0.001 dB to 0.018 dB, 0.008 dB on average, and the average BR increases from -0.05% to 0.98%, 0.28% on average. The BDPSNR and BDBR between original TZSearch and proposed algorithm change from -0.073 dB to -0.003 dB, -0.026 dB on average, and from 0.10% to 1.36%, 0.51% on average, respectively. From the values of $\triangle$PSNR, $\triangle$BR, and TS, we can conclude that the proposed algorithm can reduce the computational complexity of TZSearch efficiently.

In order to demonstrate the RD performance of the proposed algorithm, we give the RD curves of five video sequences in Fig. 2. It can be observed that the proposed algorithm achieves a quite similar RD performance as compared with the original TZSearch.

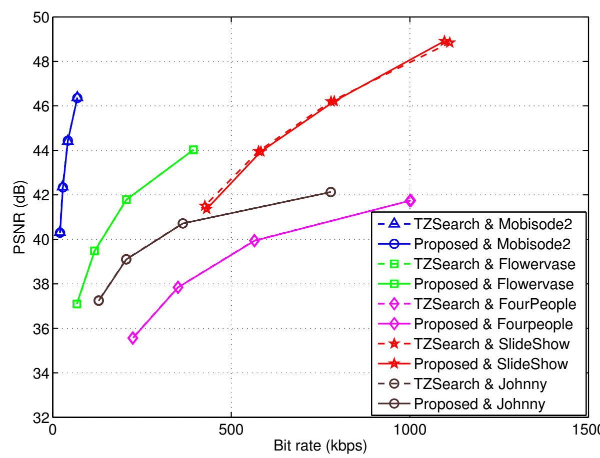

Fig. 2. RD Curves.

5. CONCLUSION

In this paper, we propose an early termination algorithm for TZSearch in HEVC. Based on the best point selection difference between large CUs and small CUs, a diamond search and a hexagon search are used to terminate early the TZSearch. Experimental results show that the proposed algorithm can achieve a quite promising coding performance in terms of RD performance and computational complexity saving.

6. REFERENCES

[1] B. Bross, W.-J. Han, J.-R. Ohm, G. Sullivan, T. Wiegand,"High efficiency video coding (HEVC) text specification draft 8," Document JCTVC-J1003, 10th Meeting: Stockholm, SE, 2012.

[2] K. Ugur et al., "High performance, low complexity video coding and the emerging HEVC standard," IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 12, pp.1688-1697, 2010.

[3] G.J. Sullivan, J.-R. Ohm, W.-J. Han, T. Wiegand,"Overview of the high efficiency video coding (HEVC) standard," IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 12, pp.1649-1668, 2012.

[4] X. Shen, L. Yu, J. Chen, "Fast coding unit size selection for HEVC based on Bayesian decision rule," in Proc. PCS'12, pp. 453-456, 2012.

[5] J. Kim, S. Jeong, S. Cho, J.S. Choi, "Adaptive coding unit early termination algorithm for HEVC," in Proc. ICCE'12, pp. 261-262, 2012.

[6] J. Kim, J. Yang, K. Won, B. Jeon, "Early determination of mode decision for HEVC," in Proc. PCS'12, 449-452, 2012.

[7] K. Choi, S. Jang, "Early TU decision method for fast video encoding in high efficiency video coding," Electronics Letters, vol. 48, no. 12, pp. 689-691, 2012.

[8] J. Kim, D. Jun, S. Cho, J.S. Choi, J. Kim, and C. Ahn, "An SAD-based selective Bi-prediction method for fast motion estimation in high efficiency video coding," ETRI Journal, vol. 34, no. 5,pp.753-758, 2012.

[9] W. Dai, Oscar C. Au, C. Pang, L. Sun, R. Zou, and S. Li, "A novel fast two step sub-pixel motion estimation algorithm in HEVC," in Proc. ICASSP'12, pp. 1197-1200, 2012.

[10] Z. Pan, S. Kwong, L. Xu, Y. Zhang, T. Zhao, "Predictive and distribution-oriented fast motion estimation for H.264/AVC," J Real-Time Image Proc, DOI 10.1007/s11554-012-0264-7, 2012.

[11] R. Li, B. Zeng, and M.L. Lio, "A new three-step search algorithm for block motion estimation," IEEE Trans. Circuits Syst. Video Technol., vol. 4, no. 4, pp. 438-443, 1994.

[12] K. Lim, "Text description of joint model reference encoding methods and decoding concealment methods," Document JVT-O079, 15th meeting, Busan, Korea, 2005.

[13] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," ITU-T VCEG, Document VCEG-M33, Thirteenth Meeting: Austin, Texas, USA, 2001.