Early Determination of Mode Decision for HEVC

Jaehwan Kim[1], Jungyoup Yang[2], Kwanghyun Won[3], and Byeungwoo Jeon[4]
School of Information and Communication Engineering
Sungkyunkwan University
300 Chunchun-dong, Jangan-gu Suwon, KOREA
[1]kjh4759@skku.edu, [2]binbak1@ece.skku.ac.kr, [3]wkh12345@skku.edu, [4]bjeon@skku.edu

Abstract—**In this paper, we propose a fast decision method scheme to reduce encoder complexity of high efficiency video coding. It is an early detection of SKIP mode in one CU-level based on the differential motion vector (DMV) and coded block flag (CBF). Experimental results show that the encoding complexity can be reduced by up to 34.55% in random access (RA) configuration and 36.48% in low delay (LD) configuration with only a little bit of rate increment compared to the high efficiency video coding test model (HM) 4.0 reference software.**

I. Introduction

Nowadays more need starts to arise for transmitting high quality and/or high resolution video contents over communication channels which still entail high cost [1]. In this context, the H.264/AVC standard [2, 3] which was developed for more than eight years ago may not be still the best to compress such high quality and/or high resolution video contents under limited bandwidth for transmission or storage. It motivated ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) to organize the Joint Collaborative Team on Video Coding (JCT-VC) in order to develop a next generation video coding standard, called High Efficiency Video Coding (HEVC) [4].

To achieve high coding efficiency, JCT-VC is developing many new coding tools, among which the generalized hierarchical structure utilizing a quad-tree structure is one of most important and effective innovations. Unlike the H.264/AVC only having a coding block of size 16x16 as macroblock, it employs a coding tree structure under which a CU, a transform unit (TU) and a prediction unit (PU) can have various block sizes. Furthermore, HEVC can achieve more accurate inter prediction by using competition and merging of motion vectors, known as AMVP (Advanced Motion Vector Prediction) technique [5]. However, its encoding complexity has significantly increased as compared to that of H.264/AVC.

To overcome this problem, a few methods have been already proposed to accomplish a fast inter mode decision in encoder [6, 7]. Gweon *et al.* [6] reduced complexity of the inter mode decision using an early termination scheme (it is called CBF fast method). This early termination utilizes the coded block flag (CBF) of luminance and chrominance. If the CBFs of luminance and chrominance are both zero, then the search process of next PU modes at the current-depth CU (denoted by CU^d) is not performed anymore. However, the CBF fast method scheme gives slight loss in coding efficiency (about 1%) since it only checks CBF for early termination of PU mode. To reduce complexity, thereby determining a fast CU depth decision, Choi *et al.* [7] proposes an early termination using the best PU mode of CU^d (it is called Early CU determination scheme). This method focuses on reducing division of CU using the PU mode. When the best PU mode of CU^d selects the SKIP mode, CU^d is not divided into sub-depth CUs (CU^{d+1}) anymore. This early CU determination utilizes the SKIP mode only for early termination of the CU depth search. Therefore, the HEVC encoder still needs a fast method decision scheme within one CU depth. It is noted that a fast method decision scheme [8-10] in that context has been already available in H.264/AVC reference software. The scheme is an early SKIP detection using SKIP mode conditions for fast method decision.

In this paper, an early detection of SKIP mode in one CU level is proposed to reduce the encoding complexity of HEVC. It simply checks the differential motion vector (DMV) and the coded block flag (CBF) after searching the best Inter 2Nx2N mode. Experimental results show significant complexity reduction of encoder at the cost of only a very little bitrate increment.

II. Overview of Mode Decision in HEVC

The current HEVC (HM4.0) is based on the generalized block structure having flexible coding unit (CU), prediction unit (PU), and transform unit (TU) [4]. The CU defines a basic unit which sub-partitions a picture (or slice) into multiple rectangular regions. The PU defines a basic unit used for intra and inter predictions, and TU specifies a basic unit used for transform and quantization. Their processing is as follows. After a picture is divided into CUs, each CU performs prediction to further define PUs. After a properly-sized PU is determined through testing various coding modes, an appropriate size of TU is determined for transform and quantization.

According to HM 4.0, a CU can be from 8x8 to 64x64 as in Fig. 1. The largest and the smallest CUs are respectively called as LCU (Largest CU) and SCU (Smallest CU). In inter slice, a PU can have a mode among 11 different possibilities (SKIP, inter 2Nx2N, inter Nx2N, inter 2NxN, inter 2NxnU, inter 2NxnD, inter nLx2N, inter nRx2N, intra 2Nx2N, intra NxN, intra PCM). The inter NxN and intra NxN are used for the SCU only. A TU can have one size from 4x4 to 32x32. In the case of inter prediction, it can use the non-square transforms (4x16,

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MEST) (No. 2011-001-7578).

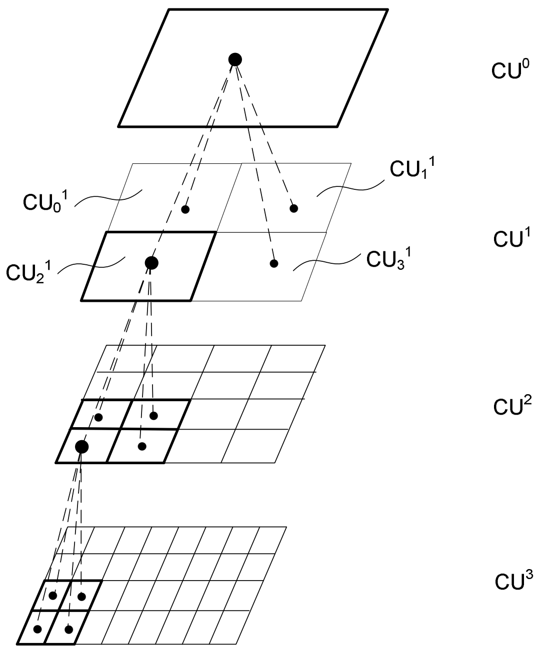

Figure 1. Hierarchical structure of HM4.0

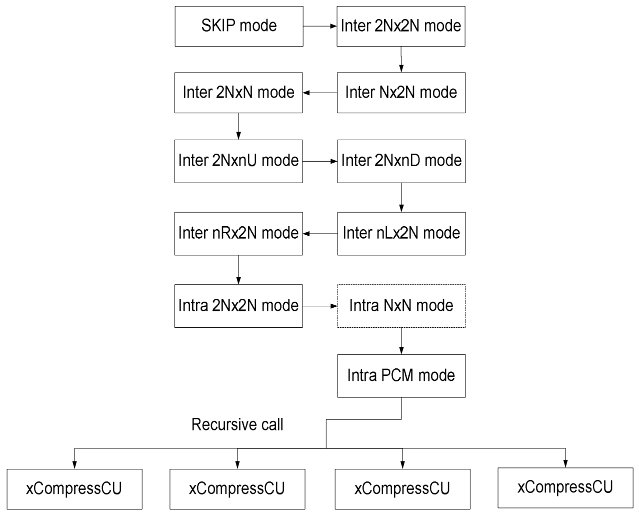

Figure 2. Mode decision process of HM4.0

16x4, 8x32, and 32x8) only if a PU has a non-square mode (inter Nx2N, inter 2NxN, inter 2NxnU, inter 2NxnD, inter nLx2N, inter nRx2N). A TU also can have a size identical or smaller than its PU in intra prediction, and identical or smaller than its CU size in inter prediction. To decide the best PU mode of a CU, the encoder typically investigates all the possible inter PU modes and intra PU modes by computing the rate-distortion (RD) cost [4]. The procedure is shown in Fig. 2. For example, for a given CU^d, RD cost is computed for the SKIP mode at the beginning. After the SKIP mode of PU is tested, the RD costs for all inter PU modes are computed from inter 2Nx2N to inter nRx2N. For the best mode among them, the coded block flag (CBF) is checked. If the CBF is zero, then all intra PU mode search is not executed. Otherwise, all intra PU modes are searched over. The best PU mode of CU^d is then selected, and a CU^d is divided into sub-depth CU (CU^{d+1}).

III. Proposed Method

According to the HM 4.0 [4], each CU is encoded by deciding its best PU mode(s). The RD cost is repeatedly computed for all the possible PU modes of CU^d. The computing process of RD cost needs a series of prediction, transform/inverse transform, quantization/dequantization, and entropy coding for each possible PU mode. Since each of the serial processing demands high computational complexity, it is practically very desirable if an encoder can decide the best PU mode at the earliest possible stage without exhaustively checking all other possible modes.

Especially, since the SKIP mode has a very high occurrence probability in many cases, early detection of SKIP mode will be very desirable. Table I shows occurrence probability of SKIP mode measured under a test conditions described in Table III. The random access (RA) configuration is used for the hierarchical B structure for coding. Under the low delay (LD) configuration, only the first picture is encoded as IDR (instantaneous decoding refresh) picture. As in Table I the occurring probability of the SKIP mode is about 0.82 in RA configuration and about 0.84 in LD configuration. These relative high probability values motivate us to design an early skip mode detection scheme.

The SKIP mode is a very efficient coding tool - it can represent a coded block data without motion search and residual coding since its transform coefficients are all quantized to zero and its motion vector(s) is same as its PMV. Unlike the SKIP mode of H.264/AVC, the SKIP mode in HEVC can utilize various PMVs and reference pictures although its basic concept of not performing residual coding and motion search is not changed.

Therefore, we propose a fast inter mode decision scheme based on an early detection of SKIP mode utilizing differential motion vector (DMV) and coded block flag (CBF) of inter

TABLE I. Occurrence Probability of SKIP mode

Sequence	HE(RA)	LC(RA)	HE(LD)	LC(LD)
BasketballDrive (1920x1080)	0.781	0.785	0.746	0.753
BQTerrace (1920x1080)	0.795	0.761	0.724	0.729
BasketballDrill (832x480)	0.847	0.859	0.843	0.841
BQMall (832x480)	0.846	0.847	0.795	0.795
BasketballPass (416x240)	0.803	0.803	0.761	0.763
BQSquare (416x240)	0.847	0.845	0.747	0.746
Vidyo1 (1280x720)			0.936	0.933
Vidyo3 (1280x720)			0.912	0.906
Vidyo4 (1280x720)			0.907	0.908
Average	**0.820**	**0.817**	**0.843**	**0.842**

HE(High Efficiency), LC(Low Complexity), RA(Random Access), LD(Low Delay)

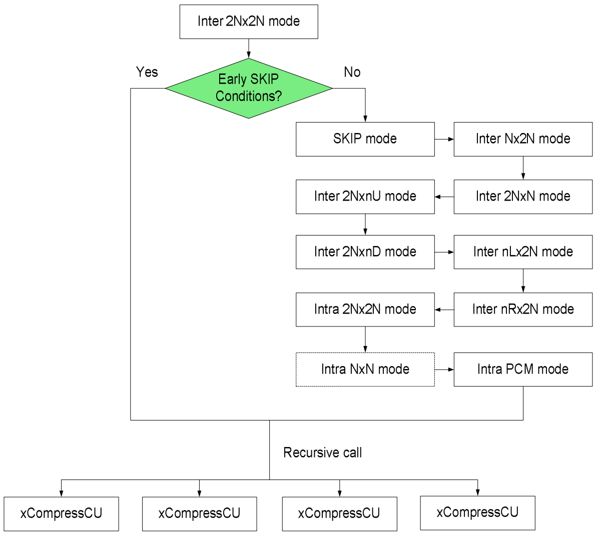

Figure 3. Mode decision process of the proposed method

2Nx2N mode. The flowchart of the proposed method is in Fig. 3. It shows that CU^d searches inter 2Nx2N modes (competition and merging) before checking the SKIP mode. Note that the HEVC HM4.0 has 2 motion vector coding methods, namely, competition-based and merging-based methods, and the SKIP mode can use only the merging-based method. After selecting the best Inter 2Nx2N mode having the minimum of RD cost, the proposed method checks its DMV and CBF. If the DMV and CBF of the best inter 2Nx2N mode are respectively equal to (0, 0) and zero (these two conditions together are called as "early SKIP conditions"), the best mode of CU^d is determined as the SKIP mode. In other words, the remaining PU modes are not searched for inter mode decision.

The SKIP mode must determine a merge index due to only using merging-based method. Therefore, when the early SKIP conditions are satisfied, the SKIP mode needs to select the best merge index. In case of the proposed method, if the best inter 2Nx2N mode is selected as competition-based method, the SKIP mode performs a merging-based method for selecting the best merge index of SKIP mode. On the other hand, if the best inter 2Nx2N mode is selected as merging-based method, then the best merge index of SKIP mode is determined as the merge index of the best inter 2Nx2N mode.

Since the proposed method dose not compute RD cost for the other modes, the encoding complexity is significantly reduced. This proposed process is summarized as follows:

$$\text{Best PU mode}=\begin{cases}\text{SKIP mode}, & \text{if CBF}=0\ \&\ \text{DMV}=(0,0)\\ \text{One of all modes}, & \text{Otherwise}\end{cases} \quad (1).$$

However, although the early SKIP conditions are satisfied, the true best PU mode of CU^d may not be the SKIP mode, thus resulting in a commission error. It can happen if the early SKIP conditions are incorrect, and the coding efficiency will be significantly dropped as a consequence. Therefore, to make sure the testing conditions are appropriate, we measured hit ratio of the SKIP mode in probability under the proposed early SKIP conditions in Table III. The detective probability of SKIP mode in Table II by the proposed method is seen to be quite

TABLE II. HIT RATIO OF THE SKIP MODE IN PROBABILITY

Sequence	HE(RA)	LC(RA)	HE(LD)	LC(LD)
BasketballDrive (1920x1080)	0.990	0.991	0.987	0.992
BQTerrace (1920x1080)	0.982	0.984	0.983	0.989
BasketballDrill (832x480)	0.988	0.989	0.989	0.990
BQMall (832x480)	0.981	0.984	0.979	0.985
BasketballPass (416x240)	0.978	0.982	0.979	0.987
BQSquare (416x240)	0.977	0.981	0.939	0.855
Vidyo1 (1280x720)			0.996	0.997
Vidyo3 (1280x720)			0.991	0.994
Vidyo4 (1280x720)			0.994	0.996
Average	**0.983**	**0.985**	**0.980**	**0.986**

high (~0.98) on average. Therefore, the proposed method is seen suitable for the fast inter mode decision.

IV. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed method, we simulated it on the HEVC test model (HM) 4.0 reference software [11] using test sequences comprising of a set of video sequences in four video resolutions. The experimental conditions shown in Table III are exactly aligned with the common test conditions recommended by the JCT-VC [12]. The performance of the proposed scheme is measured by the Bjontegaard difference bitrate (BDBR) [13] and the complexity reduction which is measured by average saving time (AST) given by:

$$\text{AST}[\%]=\frac{\text{Enc.Time(HM4.0)}-\text{Enc.Time(Proposed)}}{\text{Enc.Time(HM4.0)}}\times 100 \quad (2).$$

For comparative evaluation of the proposed method, the proposed method is compared with the CBF fast method [6] and the early CU determination [7]. Its experimental results in Table IV show that the proposed method can reduce the encoding time by about 34.55% in RA and 36.48% in LD configurations while its bit rate increment is mere 0.4% in RA

TABLE III. EXPERIMENTAL CONDITIONS

Sequence (Resolution & Total frames)	416x240: **BasketballPass (500), BQSquare (600)** 832x480: **BasketballDrill (500), BQMall (600)** 1280x720: **Vidyo1 (600), Vidyo3 (600), Vidyo4 (600)** 1920x1080: **Cactus (500), BasketballDrive (500)** 1920x1080: **BQTerrace (600)**
Max CU Depth	4
LCU size	64x64
QP	22, 27, 32, 37
GOP structure	Random Access (IBBP), Low Delay (IPPP)
Comparison	Anchor : HM4.0 Proposed method : HM4.0 + Proposed CBF fast mode (Ref [6]) : HM4.0 + Ref [6] Early CU determination (Ref [7]) : HM4.0 + Ref [7]

and 0.5% in LD. The time savings of the CBF fast method [6] and the early CU determination [7] are by about 40.70% in RA, 42.44% in LD, and 40.56 % in RA, 42.15 % in LD, respectively. The bit rate of the CBF fast method increases by 1.1% in RA, 1.3% in LD and its early CU determination causes increment of about 0.6% in RA, 0.3% in LD. When compared to the CBF fast method [6], the loss in coding efficiency of the proposed method is much less. Compared to [7], the proposed method is similar. However, it is to be noted that the proposed method is easily implementable together with the method [7] since two schemes are orthogonal to each other.

V. CONCLUSION

In this paper, we have proposed a fast method decision method for HEVC. It detects the SKIP mode at its earliest possible stage using DMV and CBF. Under the proposed scheme, the inter 2Nx2N mode (having both motion competition and merging) is searched before checking the SKIP mode. The proposed method can significantly reduce the encoding time with compromising only a very negligible loss in coding performance. Furthermore, since the proposed method can be applied as a pre-processing, it can be combined easily with other fast method decision schemes such as the CBF fast method [6] and the early CU determination [7] schemes, thus highly extensible. In this respect, our future work will combine other methods with the proposed fast method decision.

REFERENCES

[1] "Call for Evidence on High-performance Video Coding (HVC)," N10553, ITU-T Q6/16 Visual Coding and ISO/IEC JTC1/SC29/WG11 Coding on Moving Picture and Audio, Apr. 2009.

[2] Draft ITU-T recommendation and final draft international standard of joint video specification, ITU-T Rec. H.264/AVC/ISO/IEC 14496-10 AVC, JCT-G050 (2003).

[3] H.264/AVC Reference Software Version 18.3, http://iphome.hhi.de/suehring/tml, 2012.

[4] T. Wiegand, W.J. Han, B. Bross, and J. R Ohm, and G.J. Sullivan, "WD4: Working Draft 4 of High-Efficiency Video Coding," JCTVC-F803, Torino, IT, July 2011.

[5] K. McCann, B. Bross, S. Sekiguchi, and W.J. Han, "HM4: High Efficiency Video Coding(HEVC) Test Model 4 Encoder Description," JCTVC-F802, Torino, IT, July 2011.

[6] R. H. Gweon, J. L. Lee, and J. Lim, "Early Termination of CU Encoding to Reduce HEVC Complexity," JCTVC-F045, Torino, IT, July 2011.

[7] K. Choi, S. H. Park, and E. S. Jang, "Coding tree prunig based CU early termination," JCTVC-F092, Torino, IT, July 2011.

[8] B. Jeon, and J. Lee, "Fast method decision for H.264," JVT-J033, Hawaii, USA, Dec 2003.

[9] J. Lee, and B. Jeon, "Fast method decision for H.264," IEEE International Conference Multimedia and Expo (ICME), June 2004.

[10] I. Choi, J. Lee, and B. Jeon, "Fast coding mode selection with rate-distortion optimization for MPEG-4 Pare-10 AVC/H.264," IEEE Trans. On Circuits and System for Video Technology, vol. 16, no. 12, pp 1557-1561, Dec 2006.

[11] JCT-VC HEVC reference software version HM 4.0, available online at https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-4.0.

[12] F. Bossen, "Common test conditions and software reference configurations, " JCTVC-F900, Torino, IT, July 2011.

[13] G. Bjontegaard, "Calculation of average PSNR differences between RD-Curves," ITU-T SG16 Q.6 Document, VCEG-M33, Austin, April 2001.

TABLE IV. EXPERIMENTAL RESULTS

Method	Sequence	High Efficiency		Low Complexity	
		BDBR-Y[%]	*AST[%]*	*BDBR-Y[%]*	*AST[%]*
Proposed Method (RA)	Cactus	0.3	35.99	0.4	34.60
	BasketballDrive	0.4	31.69	0.4	38.87
	BQTerrace	0.4	38.49	0.5	41.15
	BasketballDrill	0.2	27.76	0.2	29.94
	BQMall	0.5	33.22	0.5	48.21
	BasketballPass	0.5	24.29	0.3	24.90
	BQSquare	0.5	35.67	0.4	38.90
	Average	**0.4**	**32.44**	**0.4**	**36.65**
Proposed Method (LD)	Cactus	0.4	33.04	0.4	31.79
	BasketballDrive	0.5	31.57	0.5	36.89
	BQTerrace	0.5	39.11	0.5	40.16
	Vidyo1	0.4	47.26	0.5	46.76
	Vidyo3	1.1	47.61	0.9	47.06
	Vidyo4	0.5	40.60	0.6	46.32
	BasketballDrill	0.2	25.64	0.2	24.68
	BQMall	0.6	41.75	0.5	39.74
	BasketballPass	0.3	22.16	0.3	22.25
	BQSquare	0.8	30.44	0.4	34.85
	Average	**0.5**	**35.92**	**0.5**	**37.05**

Method	Sequence	High Efficiency		Low Complexity	
		BDBR-Y[%]	*AST[%]*	*BDBR-Y[%]*	*AST[%]*
CBF fast mode [5] (RA)	Cactus	0.8	40.94	0.9	40.85
	BasketballDrive	1.1	37.53	1.2	45.43
	BQTerrace	0.9	43.22	1.0	47.21
	BasketballDrill	0.8	33.28	0.7	36.15
	BQMall	1.5	39.17	1.3	53.67
	BasketballPass	1.5	31.74	1.2	33.06
	BQSquare	1.1	41.72	0.8	45.84
	Average	**1.1**	**38.23**	**1.0**	**43.17**
CBF fast mode [5] (LD)	Cactus	1.0	38.01	0.9	36.83
	BasketballDrive	1.1	36.44	1.2	43.11
	BQTerrace	1.1	44.31	1.3	45.88
	Vidyo1	1.2	51.48	1.3	53.02
	Vidyo3	2.2	51.37	2.2	53.03
	Vidyo4	1.2	45.08	1.2	51.31
	BasketballDrill	0.8	30.13	0.5	30.43
	BQMall	1.5	48.22	1.4	46.78
	BasketballPass	1.4	28.43	1.1	28.67
	BQSquare	2.0	41.08	1.5	45.22
	Average	**1.4**	**41.46**	**1.3**	**43.43**

Method	Sequence	High Efficiency		Low Complexity	
		BDBR-Y[%]	*AST[%]*	*BDBR-Y[%]*	*AST[%]*
Early CU determina tion [6] (RA)	Cactus	0.9	42.37	0.7	43.20
	BasketballDrive	0.3	37.74	0.4	46.58
	BQTerrace	0.8	47.67	1.0	51.73
	BasketballDrill	0.3	29.98	0.3	40.28
	BQMall	0.7	38.13	0.7	55.05
	BasketballPass	0.5	24.18	0.5	25.58
	BQSquare	0.3	39.76	0.4	45.56
	Average	**0.5**	**37.12**	**0.6**	**44.00**
Early CU determina tion [6] (LD)	Cactus	0.5	41.97	0.4	36.27
	BasketballDrive	0.2	34.02	0.3	42.02
	BQTerrace	0.1	45.51	0.4	48.33
	Vidyo1	0.0	55.89	0.1	54.69
	Vidyo3	0.7	56.37	0.6	54.80
	Vidyo4	0.2	50.62	0.2	53.84
	BasketballDrill	0.3	25.52	0.1	26.52
	BQMall	0.3	46.83	0.3	46.72
	BasketballPass	0.1	20.19	0.2	31.76
	BQSquare	0.1	26.45	0.0	44.59
	Average	**0.3**	**40.34**	**0.3**	**43.95**