Fast Mode Decision Algorithm for Intra Prediction in HEVC

Liang Zhao[1], Li Zhang[2],
[1]Department of Computer Science and technology,
Harbin Institute of Technology,
Harbin, China
{lzhao, dbzhao}@jdl.ac.cn

Siwei Ma[2], Debin Zhao[1]
[2]Institute of Digital Media,
Peking University,
Beijing, China
{li.zhang, swma}@pku.edu.cn

***Abstract*—As the next generation standard of video coding, the High Efficiency Video Coding (HEVC) is intended to provide significantly better coding efficiency than all existing video coding standards. To improve the coding efficiency of intra frame coding, up to 34 intra prediction modes are defined in HEVC. The best mode among these pre-defined intra prediction modes is selected by rate-distortion optimization (RDO) for each block. If all directions are tested in the RDO process, it will be very time-consuming. To alleviate the encoder computation load, this paper proposes a new method to reduce the candidates in RDO process. In addition, the direction information of the neighboring blocks is made full use of to speed up intra mode decision. Experimental results show that the proposed scheme provides 20% and 28% time savings in intra high efficiency and low complexity cases on average compared to the default encoding scheme in HM 1.0 with almost the same coding efficiency. This algorithm has been proposed to HEVC standard and partially adopted into the HEVC test model**.

Index Terms- video coding, HEVC, mode decision, intra prediction

I. INTRODUCTION

Although the emerging HEVC standard [1] still belongs to block-based hybrid video coding framework, it provides a highly flexible hierarchy of unit representation which includes three block concepts: coding unit (CU), prediction unit (PU) and transform unit (TU). This separation of the block structure is helpful for each unit to be optimized. CU is a macroblock-like unit of region splitting. It is always square and its size can be from 8x8 luma samples up to the largest coding units (LCUs) . The CU concept allows recursive splitting into four equally sized blocks, starting from LCU. This process gives a content-adaptive coding tree structure comprised of CU blocks. The PU is used only for the CU which is the leaf node in quadtree structure. For intra prediction, two PU sizes are supported, which are $2N$x$2N$ and NxN. In addition to the CU and PU definitions, there is another transform and quantization related unit, TU, whose size can't exceed that of the CU.

Based on this recursive structure, encoder needs to exhaust all the combinations of CU, PU and TU to find the optimal solution, which is a very time-consuming process. What's more, intra prediction supports up to 34 directions to select the best direction. The encoder will not tolerate it if all the directions are employed in the rate-distortion optimization (RDO) process. To reduce the computational complexity of the encoder, a fast intra mode decision [4] was adopted in HM1.0 [5]. The unified intra in HM1.0 first determines the first N best candidate modes selected by a rough mode decision (RMD) process where all modes are tested by minimum absolute sum of Hadamard Transformed coefficients of residual signal (HSAD) and the mode bits in the rough mode decision. Instead of the total intra prediction modes decision, the RD optimization is only applied to the N best candidate modes selected by the rough mode decision where all modes are compared in this decision. However, computation load of the encoder is still very high. On the other side, the intra prediction modes are always correlated among the neighbors which are not considered in HM1.0. Therefore, there is still some space for further reducing the encoder complexity.

To further relieve the computation load of the encoder, it is important to reduce the candidates for RDO process and make full use of the information of its neighboring blocks. In this paper, we check less number of best RMD modes for RDO, and the most probable mode (MPM) is always included in the candidates for RDO.

The remainder of this paper is organized as follows. Section II presents an overview of intra prediction in HEVC. Section III gives a detailed description of the proposed fast intra mode decision algorithm. Experimental results are shown in Section IV. Finally, this paper is concluded.

II. OVERVIEW OF INTRA PREDICTION IN HEVC

Intra prediction is employed to remove the spatial redundancies within one image. In H.264/AVC, intra prediction of the target block is conducted in spatial domain by referring to the neighboring samples from left, up and top-right region. Although unified intra prediction is still conducted in spatial domain in current HEVC, boundary pixels from the left down region may be used as context pixels for prediction. What's the most important, it provides up to as much as 34 prediction modes for different PUs in emerging HEVC standard instead of only nine prediction modes being available for 4x4 luma blocks in H.264/AVC. Intra directions used for each PU size are demonstrated in Table I [1].

The prediction directions in the unified intra prediction have angles of +/- [0, 2, 5, 9, 13, 17, 21, 26, 32]/32. The angle is given by displacement of the bottom row of the PU and the reference row above the PU in case of vertical prediction, or displacement of the rightmost column of the PU and the reference column left from the PU in case of horizontal prediction ([2][3]). The pixels of the target block can be predicted by linearly extrapolating of the reference samples at 1/32th pixel accuracy for all block sizes instead of different precision for different PU sizes. Besides DC prediction mode, the 33 possible intra prediction directions are illustrated in Figure 1 below [1].

TABLE I: NUMBER OF INTRA DIRECTIONS FOR EACH PU SIZE

PU size	Number of Intra Directions
4x4	17
8x8	34
16x16	34
32x32	34
64x64	5

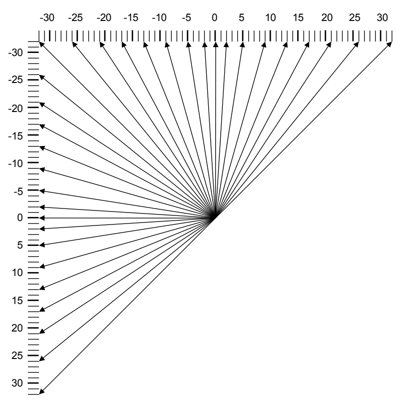

Figure 1. 33 intra prediction directions

Taking all the directions of each PU size into RDO procedure will cause a great burden to the encoder. Therefore, HM1.0 uses a combination of rough mode decision and RDO to boosting the speed of the process for making the final decision of intra prediction. However, the intra direction information of neighboring blocks has not been explored in the mode decision process. In fact, since the local image texture which has consistent orientation may cover several neighboring blocks, it is desirable to analyze the direction information of the neighboring blocks and make full use of it to accelerate the intra mode decision.

III. PROPOSED FAST INTRA MODE DECISION ALGORITHM

In this section, a fast mode decision algorithm for intra prediction is described, including motivating observations and implementation issues. We start with motivating observations, which provide useful guidelines for modeling the correlation between current optimal intra prediction and the prediction directions of its neighboring blocks. Therefore, this correlation is explored to accelerate the intra mode selection process.

A. Motivating Observations

In natural pictures, neighboring blocks usually hold similar textures. Consequently, the optimal intra prediction of current block may have strong correlation with its neighboring blocks. Based on this consideration, we estimate the conditional probabilities of the optimal intra direction of current block to be the most probable mode (MPM) of its neighboring blocks. Formally, we define

$$\begin{aligned} & P(M_{Curr} \mid (M_A, M_B)) \\ & = P(Mode_{Curr} = \min(M_A, M_B) \mid Mode_A = M_A, Mode_B = M_B) \end{aligned} \quad (1)$$

where $Mode_{Curr}$, $Mode_A$, $Mode_B$ are random variables that represent the RD optimal prediction modes of the current block and two neighboring blocks A and B as depicted in Figure 2. M_{Curr} , M_A and M_B are their possible values respectively.

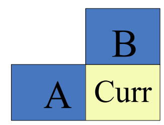

Figure 2. Neighboring blocks of current block

TABLE II: PERCENTAGES OF MPM IN HIGH EFFICIENCY TEST

	A	B	C	D	E
22	24.5%	28.0%	24.9%	20.5%	33.6%
27	27.7%	31.5%	27.5%	23.3%	37.1%
32	30.8%	36.7%	30.0%	27.5%	38.8%
37	34.7%	42.4%	34.8%	32.9%	40.8%

Table III: PERCENTAGES OF MPM IN LOW COMPLEXITY TEST

	A	B	C	D	E
22	34.0%	36.0%	34.9%	31.8%	42.2%
27	37.8%	40.1%	38.0%	35.4%	44.3%
32	40.9%	43.0%	40.6%	39.3%	45.6%
37	43.2%	45.6%	44.9%	43.4%	47.0%

In the statistics, we employ eighteen sequences in different resolutions from class A to class E with quantization parameters 22, 27, 32 and 37. Table II and Table III illustrate the percentage of MPM to be the optimal prediction mode for current block in high efficiency and low complexity test conditions, separately. From our statistic results, we find that the MPM of current block possesses a large ratio to be the best mode in current block in both test conditions and this ratio of MPM fluctuates only a little between different sequences. Consequently, the MPM of current block should be always employed as the candidate to compete for the best mode. However, this important information has not been explored in HM1.0 RDO process.

B. Implementation of Proposed Fast Intra Mode Decision

In our proposed algorithm, we still use the combination of the rough mode decision and RDO process to select the best intra direction. There are two differences between the intra mode decision in HM1.0 and our proposed method. First, we reduce the number of directions used for RDO process; second, the correlation between the neighboring blocks and the target block are used to remedy the loss of the coding efficiency produced from the first phase. Detailed algorithm is described as follows.

Firstly, we analyze the candidates selected in rough mode selection and reduce the number of candidates depending on the size of each PU. Based on lots of experiments, we observe that the candidates selected from rough mode decision render a descending trend to be the RDO-optimal mode according to their rank in candidates. In addition, first two candidates of all PU sizes present a majority ratio to be the RDO-optimal mode on average. This ratio differs a bit with different PU size, so we employ different number of candidates with different PU size for RDO process.

Then, we check whether most probable mode (MPM) is included in the candidates for each PU size. If MPM is not included in candidate set, $N + 1$ modes comprised of N best modes from rough mode decision and MPM will be employed in RDO process. Otherwise, only N best modes will be employed in RDO process.

Five settings (S1~S5) related to the number of N best modes in rough mode decision for RDO process are shown in Table IV. To further reduce the complexity of the encoder, four settings from S1 to S4 are employed. In S1, there are 3, 3, 2, 2 and 1 prediction modes for RDO process with the PU size of 4x4, 8x8, 16x16, 32x32 and 64x64. In S2, S3 and S4, we increase one prediction mode for the PU size of 4x4 and 8x8 in turn. To verify the coding efficiency of the MPM, there are 8, 8, 3, 3 and 3 prediction modes for RDO process with the PU size of 4x4, 8x8, 16x16, 32x32 and 64x64 in S5. Figure 3 depicts the architecture of the proposed method compared with the default HM1.0.

TABLE IV. DIFFERENT SETTINGS OF N IN PROPOSED SCHEME

Settings	Number of N
S1	$N = \begin{cases} 3, & 4\times4,\ 8\times8 \\ 2, & 16\times16,\ 32\times32 \\ 1, & 64\times64 \end{cases}$
S2	$N = \begin{cases} 4, & 4\times4,\ 8\times8 \\ 2, & 16\times16,\ 32\times32 \\ 1, & 64\times64 \end{cases}$
S3	$N = \begin{cases} 5, & 4\times4,\ 8\times8 \\ 2, & 16\times16,\ 32\times32 \\ 1, & 64\times64 \end{cases}$
S4	$N = \begin{cases} 6, & 4\times4,\ 8\times8 \\ 2, & 16\times16,\ 32\times32 \\ 1, & 64\times64 \end{cases}$
S5	$N = \begin{cases} 8, & 4\times4,\ 8\times8 \\ 3, & 16\times16,\ 32\times32 \\ 3, & 64\times64 \end{cases}$

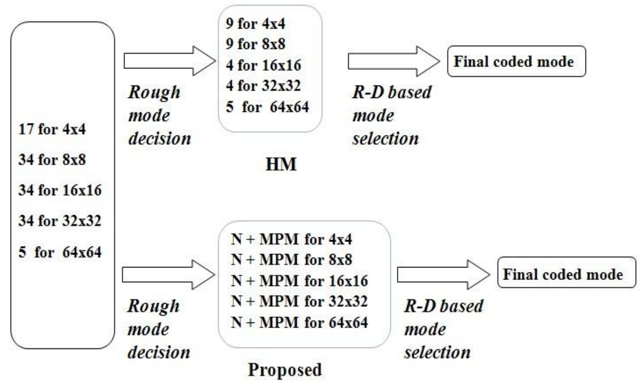

Figure 3. The proposed architecture compared with HM1.0

IV. EXPERIMENTAL RESULTS

To verify the performance of the proposed fast intra mode decision, it was implemented in the first test model HM1.0 of HEVC. Since we focus on the performance of intra coding, experiments are carried out for all I-frames sequences. According to the specifications provided in [6], two test conditions are conducted in our experiments.

The test platform used is Inter® Xeon® X5450-3.00GHz with eight cores, 8.0 GB RAM. A group of experiments were carried out on the recommended sequences with quantization parameters 22, 27, 32 and 37 as specified by [6]. Here the reference test model HM1.0 of HEVC is treated as the anchor. The "Time" in the following tables depicts the encoder time with proposed method compared to the anchor with default settings.

TABLE V. BD-RATES AND ENCODER COMPLEXITY OF S1

	All-Intra HE			All-Intra LoCo		
	Y	*U*	*V*	*Y*	*U*	*V*
Class A	0.15	-0.25	-0.22	0.27	-0.41	-0.43
Class B	0.00	-0.44	-0.42	0.15	-0.61	-0.63
Class C	0.18	-0.47	-0.49	0.42	-0.48	-0.48
Class D	0.26	-0.47	-0.49	0.49	-0.43	-0.46
Class E	0.04	-0.19	-0.28	0.24	-0.41	-0.40
All	0.12	-0.39	-0.40	0.31	-0.49	-0.50
Time	80%			72%		

TABLE VI. BD-RATES AND ENCODER COMPLEXITY OF S2

	All-Intra HE			All-Intra LoCo		
	Y	*U*	*V*	*Y*	*U*	*V*
Class A	0.06	-0.23	-0.16	0.13	-0.30	-0.33
Class B	-0.04	-0.37	-0.36	0.01	-0.47	-0.46
Class C	0.08	-0.36	-0.38	0.24	-0.32	-0.32
Class D	0.14	-0.38	-0.39	0.29	-0.32	-0.30
Class E	-0.02	-0.18	-0.27	0.12	-0.26	-0.30
All	0.04	-0.32	-0.33	0.16	-0.35	-0.35
Time	83%			76%		

TABLE VII. BD-RATES AND ENCODER COMPLEXITY OF S3

	All-Intra HE			All-Intra LoCo		
	Y	*U*	*V*	*Y*	*U*	*V*
Class A	0.01	-0.17	-0.19	0.03	-0.24	-0.26
Class B	-0.07	-0.33	-0.30	-0.08	-0.36	-0.36
Class C	0.03	-0.27	-0.28	0.13	-0.24	-0.23
Class D	0.07	-0.22	-0.27	0.18	-0.24	-0.23
Class E	-0.05	-0.12	-0.21	0.05	-0.20	-0.22
All	-0.00	-0.24	-0.26	0.06	-0.27	-0.27
Time	86%			80%		

TABLE VIII. BD-RATES AND ENCODER COMPLEXITY OF S4

	All-Intra HE			All-Intra LoCo		
	Y	*U*	*V*	*Y*	*U*	*V*
Class A	-0.02	-0.17	-0.14	-0.03	-0.19	-0.20
Class B	-0.09	-0.29	-0.27	-0.15	-0.29	-0.28
Class C	0.00	-0.21	-0.21	0.04	-0.16	-0.14
Class D	0.04	-0.18	-0.21	0.08	-0.18	-0.14
Class E	-0.06	-0.11	-0.23	-0.01	-0.14	-0.17
All	-0.03	-0.20	-0.22	-0.02	-0.20	-0.19
Time	88%			84%		

TABLE IX. BD-RATES AND ENCODER COMPLEXITY OF S5

	All-Intra HE			All-Intra LoCo		
	Y	*U*	*V*	*Y*	*U*	*V*
Class A	-0.12	-0.17	-0.19	-0.28	-0.22	-0.23
Class B	-0.19	-0.26	-0.27	-0.37	-0.27	-0.27
Class C	-0.07	-0.12	-0.13	-0.10	-0.11	-0.12
Class D	-0.05	-0.10	-0.14	-0.07	-0.09	-0.11
Class E	-0.20	-0.16	-0.26	-0.26	-0.24	-0.23
All	-0.13	-0.17	-0.20	-0.21	-0.18	-0.19
Time	97%			94%		

The BD-rate performance of Y, U and V components and encoder complexity compared to the default encoding scheme are shown in the above tables (Table V~IX). In Table V, our proposed method saves encoding time of 20% and 28% on average compared with HM1.0 in high efficiency and low complexity test conditions while the increase of BD-rates are only 0.12% and 0.31% for luma in the setting of S1.

We can also find that our proposed method with class D in Table V cause the BD-rates to an increase of 0.26% and 0.49% on average in high efficiency and low complexity test conditions which are the worst loss in all above five tables. However, from the RD curves of Figure 4 and Figure 5, we can observe that our proposed method performs almost the same coding efficiency from low to high bit-rate compared with the anchor.

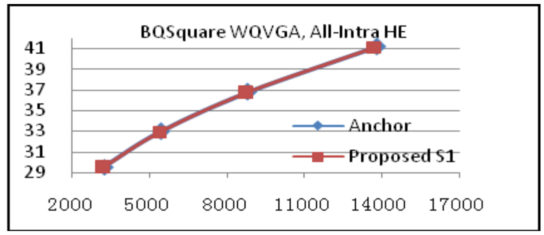

Figure 4. RD curves of BQSquare (Class D 416x240)

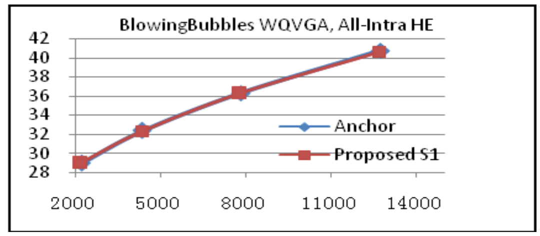

Figure 5. RD curves of BlowingBubbles (Class D 416x240)

V. CONCLUSIONS

This paper proposes a fast mode decision algorithm for intra prediction in the emerging HEVC standard. By making use of neighboring intra prediction information, we reduce the number of directions taking part in RDO process. This proposed method results in significant reduction of the encoder complexity. Experimental results show that our proposed algorithm has almost negligible performance loss compared to HM1.0. This algorithm was proposed to the HEVC standard [7], and has been partially adopted in HM2.0.

ACKNOWLEDGMENT

This work was partly supported by the National Science Foundation, 60833013, the National Science Foundation of China, 60736043, and the National Basic Research Program of China, 973 Program 2009CB320903.

REFERENCES

[1] JCT-VC, "WD1: Working Draft 1 of High-Efficiency Video Coding", JCTVC-C403, JCT-VC Meeting, Guangzhou, October 2010.

[2] J.-H. Min, S. Lee, I.-K. Kim, W.-J. Han, J. Lainema, and K. Ugur, "Unification of the Directional Intra Prediction Methods in TMuC," JCTVC-B100, Geneva, Switzerland, Jul 2010.

[3] J.-H. Min, S. Lee, I.-K. Kim, W.-J. Han, J. Lainema, and K. Ugur, "TE5: Results for Simplification of Unified Intra Prediction" JCTVC-C042, Guangzhou, Jul 2010.

[4] Yinji Piao, Junghye Min, Jianle Chen, "Encoder improvement of unified intra prediction," JCTVC-C207, Guangzhou, October 2010.

[5] https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware

[6] JCT-VC, "Common test conditions and software reference configurations", JCTVC-C500, JCT-VC Meeting, Guangzhou, October 2010.

[7] Liang Zhao, Li Zhang, Xin Zhao, Siwei Ma, Debin Zhao, Wen Gao, "Further encoder improvement for intra mode decision", JCTVC-D283, JCT-VC Meeting, Daegu, January 2011.