

A Hardware-Friendly Method for Rate-Distortion Optimization of HEVC Intra Coding

Weiwei Shen, Yibo Fan, Leilei Huang, Jiali Li, Xiaoyang Zeng
State Key Lab of ASIC and System, Fudan University
Shanghai, China
10110720024@fudan.edu.cn, fanyibo@fudan.edu.cn

Abstract—The next generation standard of coding, High Efficiency Video Coding (HEVC) aims to provide significantly improved compression performance in comparison with all existing video coding standards, such as MPEG-4, H.263, and H.264/AVC. The tool of rate-distortion optimization (RDO) mode decision has proven to be extremely important. Due to the increase in the number of intra prediction directions and large size of coding unit (CU), prediction unit (PU) and transform unit (TU), the number involved in RDO process rises dramatically, which is an extremely timing-and-computation-consuming process. In this paper, we propose a four-pixel-strip based ESAD and quantized TU coefficients based linear model to estimate the distortion and rate part of the RD cost function respectively. Experimental results show that the proposed RD cost function provides 85.8% area reduction and 1260% throughput improvement in hardware design, with negligible loss of bitrate and PSNR.

Keywords—hardware-friendly algorithm, intra prediction, rate-distortion optimization, HEVC.

I. INTRODUCTION

ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) formed a Joint Collaborative Team on Video Coding (JVT-VC) in 2010, and the next generation coding standard, High Efficiency Video Coding (HEVC), is now being developed. HEVC aims to reduce 50% bit rate in comparison with the existing H.264/AVC high profile, under the same visual quality.

Alike the previous video coding standard H.264/AVC, HEVC also adopted the prediction/transform block based hybrid coding framework [1] [2]. In the framework, a quadtree based coding structure is an important feature of the new standard.

Several new coding structures have been introduced in HEVC: coding unit (CU), prediction unit (PU) and transform unit (TU). The CU is the basic unit of region splitting used for intra/inter coding [3]. It can be split from the largest coding unit (LCU, can be large to 64x64 pixels) into the smallest coding unit (SCU, can be small to 8x8 pixels). Coupled with CU, PU carries the information related to the prediction processes. TU is used for transform and quantization, and it depends on PU partitioning modes.

A comparison between HEVC and H.264/AVC [3] shows that HEVC provides greater bit rate savings but also brings

higher encoding complexity. In the HEVC intra encoding process, the increase in the number of intra prediction directions and large size of coding unit (CU), prediction unit (PU) and transform unit (TU), lead to a dramatic rise in computational complexity.

In the official reference software HM9.0, the encoder needs to exhaust all the combinations of CU, PU and TU to find the optimal solution, which is a time-consuming process. The derivation of the best intra prediction mode of a given PU can be divided into three steps. Firstly, a candidate subset of all intra prediction modes is obtained by Rough Mode Decision (RMD) operation [4], which takes the Sum of Absolute Hadamard Transformed Difference (SATD) and mode bits into account. The number of the candidate subset is pre-set to be 8 for 4x4, 8x8 PUs, 3 for 16x16, 32x32, 64x64 PUs. Secondly, the Most Probable Modes (MPM) [5] from the neighboring block are added to the subset. Finally, the best prediction mode is selected by using rate distortion optimization (RDO) method, which increases the computational overhead as it exhausts a large number of multipliers in distortion part, and it requires the header and coefficient bits to be calculated by CABAC encoding in rate part. After the best PU mode obtained by the RDO process, the TU size decision is also involved in RDO process.

For an LCU in HEVC, there are one 64x64-PU, four 32x32-PUs, 16 16x16-PUs, 64 8x8-PUs, and 256 4x4-PUs. Different size PUs are assigned to different numbers of candidate subsets, as mentioned above. So that it requires 2623 times CABAC encoding for prediction mode decision in total for one LCU, which would bear great challenge for high-throughput encoder. And many available designs [6] [7] do not allow the access to CABAC encoding due to the real-time encoding requirements.

Many efforts have been made to investigate fast intra prediction algorithms to reduce computational complexity in H.264/AVC. Most commonly, the RD cost function just uses the Sum of Absolute Difference (SAD) or SATD, which leads to a high degradation in coding efficiency. In [8], a SATD-based RD cost function that uses the absolute transform coefficients variance to evaluate the coded rate. Another rate estimation method is proposed in [9], which is based on a linear relationship between rate and quantized DCT coefficients. However the linear cost function is not so accuracy and too complicated to be hardware implemented.

In this paper, we propose a hardware-friendly method for

RDO of HEVC intra coding. A four-pixel-strip based enhanced-SAD (ESAD) method is adopted to estimate the distortion part of the RD cost function, which decreases the number of multipliers to be 25%. Meanwhile, a quantized DCT coefficients based linear model is proposed to estimate the rate part of the RD cost function, which would avoid the access to CABAC encoding. The proposed innovative RD cost function can be applied to high-throughput HEVC encoder in hardware implementation.

The rest of this paper is organized as follows. In section II, the original RD cost function is briefly introduced. In section III, our proposed distortion and rate estimation method is described. In section IV, our experimental results are presented. Finally, in section V, some conclusions are drawn.

II. RD COST FUNCTION IN HEVC

In HEVC, the best CU, PU, TU combination solution is the one with minimum RD cost. The RD cost function is expressed as

$$J_{RD} = SSE + \lambda * R \quad (1)$$

where the SSE is the sum of squared difference between the current original block \mathbf{O} and the current reconstructed block \mathbf{R} . The SSE and R are the distortion and rate part of the RD cost function respectively. The SSE is expressed by

$$SSE = \sum_{i=1}^N \sum_{j=1}^N (O_{i,j} - R_{i,j})^2 \quad (2)$$

where $O_{i,j}$ and $R_{i,j}$ are the $(i, j)^{th}$ elements of the current original blocks \mathbf{O} and the current reconstructed block \mathbf{R} . N is the block size of the current TU. In Eq. (1), R is the true rate needed to encode the block and λ is an exponential function of the quantization parameter (QP). However this RDO process bears extremely high computational load. Hence, the cost function will make HEVC difficult to be realized in real-time applications.

III. PROPOSED DISTORTION AND RATE ESTIMATION METHOD

A. Distortion Estimation Method

In Eq. (2), the calculation of SSE is a multiplier-consuming process, which exhausts $N \times N$ multipliers. Here, a four-pixel-strip based ESAD is proposed to estimate the distortion part of the RD cost function. The SAD is given by

$$SAD = \sum_{i=1}^N \sum_{j=1}^N |O_{i,j} - R_{i,j}| = \sum_{i=1}^N \sum_{j=1}^N D_{i,j} \quad (3)$$

where D_{ij} is the $(i, j)^{th}$ elements of the absolute difference between current original block \mathbf{O} and the current reconstructed block \mathbf{R} , and we define it as a difference block \mathbf{D} .

It can be proved that

$$SAD^2 = SSE + \sum_{i=1}^N \sum_{j=1}^N D_{i,j} * (SAD - D_{i,j}) \quad (5)$$

It is reasonable to say that the values of absolute difference between the original pixels and the reconstructed pixels are close to each other in a TU (the reconstructed block is based on a TU size due to the RDO process in HEVC), meanwhile the value of D_{ij} is relatively small than SAD. Thus we assume that D_{ij} in Eq. (5) equals to the mean value of the block (SAD/N^2), i.e. ,

$$\begin{aligned} SAD^2 &\approx SSE + \sum_{i=1}^N \sum_{j=1}^N D_{i,j} * (SAD - SAD / N^2) \\ &= SSE + \sum_{i=1}^N \sum_{j=1}^N D_{i,j} * SAD - \sum_{i=1}^N \sum_{j=1}^N D_{i,j} * (SAD / N^2) \quad (6) \\ &= SSE + SAD^2 - SAD * (SAD / N^2) \end{aligned}$$

From Eq. (6), the proposed SSE can be described as

$$SSE_{proposed,I} = SSE \approx SAD * (SAD / N^2) \quad (7)$$

1	0	4	4	2	4	1	0
5	3	0	2	1	2	2	1
3	7	4	6	4	3	3	4
1	1	2	4	5	7	1	0
2	2	0	0	1	1	4	3
2	4	1	2	0	1	1	0
1	1	5	1	2	3	2	8
16	15	11	9	2	1	0	3

Fig. 1. Example of an 8x8 Difference Block \mathbf{D} .

From the experimental results, the performance of the proposed $SSE_{proposed,I}$ function is not so good. Considering the example in Fig. 1, which shows an 8x8 difference block \mathbf{D} , it is found that the most values of the block vary in a small range. However, the values labelled as red circles are quite higher compared to their neighbors. These exceptions weaken the assumption that D_{ij} in Eq. (5) to be the mean value of the block. In fact, the mean value deviates most D_{ij} due to the exceptions.

$$SSE_{proposed,II} = \sum_{i=1}^{N/4} \sum_{j=1}^N (SAD_{i,j}) * (SAD_{i,j} \gg 2) \quad (8)$$

$$SAD_{i,j} = D_{4*i-3,j} + D_{4*i-2,j} + D_{4*i-1,j} + D_{4*i,j} \quad (9)$$

On the other hand, values of four consecutive absolute difference between original blocks and the reconstructed blocks in a strip are close to each other. As shown in Fig.1, each strip contains four D_{ij} labeled in a green box. As a result, a four-pixel-strip based ESAD ($SSE_{proposed,II}$) is proposed to estimate the distortion part of the RD cost function, which is expressed in Eq. (8), and the $SAD_{i,j}$ is described in Eq. (9). With the proposed method, the number of multipliers needed for the calculation of the distortion part of RD cost function can be reduced to 25%, compared with the SSE.

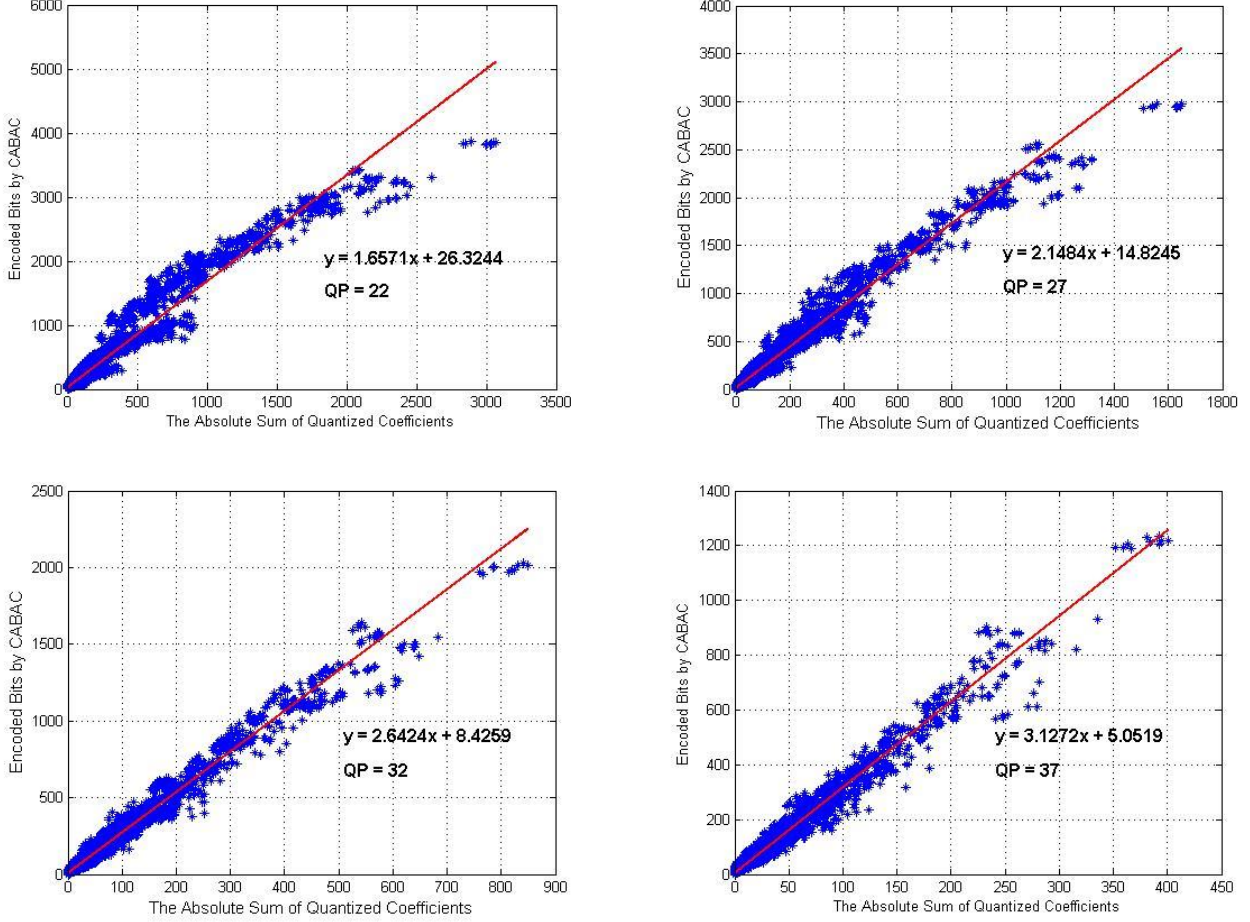


Fig. 2. The Absolute Sum of Quantized Coefficients vs Encoded Bits By CABAC.

B. Rate Estimation Method

In Eq. (2), the calculation of actual coded bits (R) for every TU involves in CABAC encoding in HM 9.0, which is a computationally complex task for real-time applications. Inspired by [10], the information of quantized TU coefficients are used to estimate the rate part of RD cost function.

Based on the above information, we plot the absolute sum of quantized coefficients of a given TU against the actual encoded bits by CABAC. From the experimental results shown in Fig. 2, we can see that there is a strong linear relationship between each other. We apply a linear regression model to these plots to simulate the linear trend. These plots are generated from five test-sequences applied by JCT-VC (Traffic(2560x1600), BasketballDrive(1920x1080), Johnny (1280x720), BQMall(832x480), BlowingBubbles(416x240)). Four of these plots from test-sequence BlowingBubbles (QP equals to 22, 27, 32, 37 respectively) are shown in Fig. 2. So far, a new rate model is proposed as:

$$Rate = a * \sum_{i=1}^N \sum_{j=1}^N |q_coeff_{i,j}| + b \quad (10)$$

where $q_coeff_{i,j}$ is the (i, j) th elements of quantized TU coefficients block, and a, b are the model parameters.

It is observed that the values of parameter a and b vary according to the QPs in Fig. 2. As QP goes up, the value of a becomes bigger and the value of b reverses the process, getting smaller. To approximate the relationship between parameter a, b and QP, a linear and exponential model, for the values of a and b respectively, are obtained by experimental results, shown Eq. (11) and (12). Along with Eq. (10), it is easy to find the QP's contribution to rate part of the RD cost function. For hardware implementation convenience, two look-up tables are utilized instead of Eq. (11) and (12), where QP is the integer index ranging from 0 to 51.

$$a = 0.06713 * QP + 0.13568 \quad (11)$$

$$b = 144.46 * e^{-QP/15.6} - 8.9349 \quad (12)$$

IV. EXPERIMENTAL RESULTS

In this section, the performance of the proposed method for distortion and rate estimation of the RD cost function is evaluated in terms of Δ Bitrate and Δ PSNR_YUV. The performance gain or loss is measured with respect to the HEVC reference software platform (HM9.0). The experiments are carried out for "ALL Intra-Main" setting by the common condition supplied by JCT-VC.

TABLE I. BITRATE AND PSNR_YUV INCREMENTS OF THE PROPOSED ALGORITHM

Class	Sequence	Δ Bitrate(%)	Δ PSNR_YUV(dB)
Class A	Nebuta	0.89%	-0.16
	SteamLocomotive	2.91%	-0.10
Class B	ParkScene	1.82%	-0.14
	BQTerrace	1.96%	-0.16
Class C	PartyScene	1.81%	-0.21
	RaceHorses	2.10%	-0.20
Class D	BlowingBubbles	1.41%	-0.20
	RaceHorses	3.43%	-0.19
Average		2.04%	-0.15

TABLE II. HARDWARE IMPLEMENTATION RESULTS

Purpose	Gate count (NAND2)	Process cycles (process a 4x4 block)
Distortion calculated by SSE	4552	1
Distortion calculated by $SSE_{proposed, II}$ (Eq. (8))	2631	1
Rate encoded by CABAC	27625	13.6
Rate calculated by proposed method (Eq. (10))	1947	1

TABLE III. HARDWARE IMPLEMENTATION COMPARISON

	Area Reduction (%)	Throughput Improvement (%)
Distortion part	42.2	0
Rate part	93	1260
RD cost function	85.8	1260

Table I shows the Δ Bitrate and Δ PSNR_YUV of our proposed method in comparison with HM9.0. The average bitrate increment is about 2.04%, and the average PSNR_YUV decrement is about 0.15dB, which are acceptable relative to area reduction and throughput improvement in hardware implementation, shown in Table II.

The proposed and original method of distortion and rate calculation in RD cost function are implemented in Verilog-HDL and synthesized under a timing constraint of 200MHz with TSMC 65-nm library. The number of Gate count and process cycles of every module are summarized in Table II. Because of the irregular process cycles needed for CABAC encoding a 4x4 block, here, we give an average process cycles 13.6 in hardware implementation.

The area of distortion part is reduced by 42.2% with our proposed four-pixel-strip based ESAD, and the rate part is reduced by 93% with the proposed linear model, showed in

Table III. Meanwhile, a 1260% increase in throughput of the rate part makes it meaningful to real-time application.

Combining the four-pixel-strip based ESAD with the rate linear model, the total area of the RD cost function is reduced by 85.8%, and the throughput is enhanced by 1260%. From the results, we can see that our proposed method is hardware-friendly for rate-distortion optimization of HEVC intra coding.

V. CONCLUSION

A four-pixel-strip based ESAD and quantized TU coefficients based linear model are proposed to estimate the distortion and rate part of the RD cost function respectively. Experimental results show that the proposed method provides 85.8% area reduction and 1260% throughput improvement in hardware design, with negligible loss of bitrate and PSNR, which is very suitable for real-time encoder application. Furthermore, this proposed method can be combined with any other fast intra algorithm technique to further reduce the complexity of intra coding.

ACKNOWLEDGEMENT

This paper was supported by National Natural Science Foundation of China (61306023), Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP, 20120071120021), STCSM (13511503400), National High Technology Research and Development Program (863, 2012AA012001).

REFERENCES

- [1] 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.
- [2] T. Wiegand, B. Bross, W.-J. Han, J.-R. Ohm, and G. J. Sullivan, "Working Draft 3 of High-Efficiency video Coding," JCTVC-E603, March 2011.
- [3] B. Li, G. J. Sullivan, J. Xu, "Comparison of Compression Performance of HEVC Draft 6 with AVC High Profile," JCTVC-I0409, April 2012.
- [4] Y. Piao, J.-H. Min, J. Chen, "Encoder improvement of unified intra prediction," JCTVC-C207, October 2010.
- [5] Liang Zhao, Li Zhang, Xin Zhao, Siwei Ma, Debin Zhao, Wen Gao, "Further Encoder Improvement of intra mode decision," JCTVC-D283, January 2011.
- [6] Young-Joon Jo, Jin-Su Jung, Hyuk-Jae Lee, "Fast pipeline schedule for an H.264 intra frame encoder with early termination," Signal Processing Systems, 2009. SiPS 2009. IEEE Workshop on , vol., no., pp.109,114, 7-9 Oct. 2009.
- [7] Yu-Kun Lin, Chun-Wei Ku, De-Wei Li, Tian-Sheuan Chang, "A 140-MHz 94 K Gates HD1080p 30-Frames/s Intra-Only Profile H.264 Encoder," Circuits and Systems for Video Technology, IEEE Transactions on , vol.19, no.3, pp.432,436, March 2009.
- [8] Yu-Ming Lee, Yu-Ting Sun, Yinyi Lin, "SATD-Based Intra Mode Decision for H.264/AVC Video Coding," Circuits and Systems for Video Technology, IEEE Transactions on , vol.20, no.3, pp.463,469, March 2010.
- [9] Johar, S.; Alwani, M., "Method for fast bits estimation in rate distortion for intra coding units in HEVC," Consumer Communications and Networking Conference (CCNC), 2013 IEEE , vol., no., pp.721,724, 11-14 Jan. 2013.
- [10] Zhihai He, Yong Kwan Kim, Mitra, S.K., "Low-delay rate control for DCT video coding via p-domain source modeling," Circuits and Systems for Video Technology, IEEE Transactions on , vol.11, no.8, pp.928,940, Aug 2001.