A Fast Intra Optimization Algorithm for HEVC

Shanshe Wang[1], Siwei Ma[2], Xiaolong Jiang[3], Juanting Fan[2], Debin Zhao[1], Wen Gao[2]

[1]School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

sswang@jdl.ac.cn; dbzhao@hit.edu.cn

[2]Institute of Digital Media, Peking University, Beijing, China

swma@pku.edu.cn; jtfan@jdl.ac.cn; wgao@pku.edu.cn

[3] University of Chinese Academy of Sciences , Beijing, China

xljiang@jdl.ac.cn

***Abstract*—High Efficiency Video Coding (HEVC) has drastically improved the coding efficiency, and can provide more than 23% bit rate reduction compared to its predecessor H.264/AVC for all intra (AI) configuration. However, the improvement on coding efficiency is obtained at the expense of much more intensive computation complexity. In this paper, a fast intra optimization scheme is proposed. Firstly, a shrinkage scheme for the number of mode prediction is proposed based on gradient variance. Then a novel estimation model for rate distortion cost is provided based on Hadamard transform and quantization to better determine the mode prediction for RQT. Experimental results demonstrate that, compared with the original HEVC reference encoder implementation in AI configuration, the proposed algorithms can achieve more than 26% coding complexity reduction on average with ignorable coding performance degradation.**

***Index Terms*— HEVC; intra coding; Hadamard transform; Gradient variance; rate distortion optimization**

I. INTRODUCTION

The dramatically increasing of high definition (HD) and beyond-HD (e.g. 4kx2k or 8kx4k) videos are creating stronger demand of high efficiency video coding technology, which are beyond the capabilities of the state-of-the-art video coding standard such as H.264/AVC [1]. Therefore, the ITU-T Video Coding Expert Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) formed the Joint Collaborative Team on Video Coding (JCT-VC) in April 2010 to develop the new coding standard which is formally published in 2013 with the name as high efficiency video coding (HEVC)[2] or H.265. Many new coding tools and coding structures are adopted or improved in HEVC, which bring great performance improvement. For all-intra (AI) configuration, HEVC achieved significant improvement in coding efficiency with more than 23% bit rate saving [3][4] for AI compared to its predecessor H.264/AVC. However, the improvement on coding efficiency is obtained at the expense of more intensive computation complexity. In [8][9], the authors analysed the implementation complexity and the coding efficiency of these advanced coding tools in HEVC.

For AI, HEVC still adopts the multi direction intra prediction and the number of prediction mode can be up to 35. For evaluating the compression efficiency of each prediction mode, the encoder usually employs the Lagrange multiplier optimization technique [6], which is expressed by

$$\min\{J\} \quad J = D + \lambda \cdot R \tag{1}$$

where D and R are the reconstruction distortion and entropy coding bits of a certain unit. J is the Lagrange rate-distortion (R-D) cost function to be minimized and λ is the Lagrange multiplier. The minimization process of the R-D cost is the well-known rate-distortion optimization (RDO). In general, to obtain accurate D and R, for each candidate, the encoder has to perform transform, quantization, entropy coding, inverse quantization, inverse transform, and pixel reconstruction, which makes the R-D cost calculation very time-consuming and brings great burden to the encoder implementation [7]. Thus, it is necessary to reduce the number of prediction mode for AI to achieve coding complexity reduction.

For intra mode decision optimization, many fast algorithms are proposed to reduce the coding complexity of intra mode decision. In [12], a partition depth prediction scheme is proposed based on the spatial correlation among adjacent PU. In order to improve the prediction accuracy, an adaptive threshold updating method is provided. Due to the fact that the gradient can represent some characteristics of video content, in [13], a gradient based optimization scheme for intra mode decision is proposed. In [14], according to the correlation between the texture of video and the optimal mode, fast intra mode decision method is presented. In [15], the authors proposed a fast intra mode decision for HEVC, and micro level and macro level decision algorithms were presented. At the micro level, Hadamard transform based R-D cost is utilized to reduce the candidate number and an early RDOQ skip method is also introduced to further reduce the coding complexity. At the macro level, an early coding unit (CU) split termination method is provided in terms of estimated R-D cost. In [16], a fast intra decision method is proposed to reduce the number of intra-prediction modes in rough mode decision (RMD) process based on direction information of the co-located neighbouring block of previous frame along with neighbouring blocks of current frame. In [17], according to analysing the distribution and changing trend of the costs generated by RMD, a fast intra decision scheme is presented to reduce the number of the candidates for the RDO process. In [12], Shen et al. proposed a fast scheme for intra coding in HEVC in order to determine the size of CU based on the depth

978-1-4799-6139-9/14/$31.00 ©2014 IEEE

of the surrounding CUs. Furthermore, a mode decision method is also provided for intra coding.

In this paper, a gradient variance based scheme related to the video content is utilized to the shrink the number of intra prediction mode for full RDO and a novel estimation model of rate distortion cost is proposed to better choose the prediction mode based on the Hadamard transform and quantization. The rest of the paper is organized as follows. Section II presents the proposed shrinkage scheme for the determination of the number of intra mode prediction. The proposed estimation model of rate distortion cost is shown in section III. Section IV presents the experimental results. Finally, the conclusion is presented in section V.

II. Gradient Variance based Shrinkage of the Number of Intra Prediction Mode

The determination of intra mode decision is to choose the coding mode with the minimal R-D cost, which is represented as follows,

$$\theta_{opt} = \arg\min_{\theta\in\Omega}\{D_\theta + \lambda \cdot R_\theta\} \quad (2)$$

where Ω denotes the set of all the possible intra prediction modes. D_θ and R_θ are the distortion and bit rate for the prediction mode θ, respectively. The mode with minimum R-D cost will be determined as the optimal mode. However, performing all the possible modes greatly increases the coding complexity. In [10], a fast RDO algorithm is provided to reduce the intra coding complexity for HEVC. The possible prediction modes to perform full RDO process is reduced to 8, 8, 3, 3, 3 together with the most probable modes (MPM) for PU with size of 4x4, 8x8, 16x16, 32x32 and 64x64, respectively. The scheme reduced the coding complexity by 6% with little performance improvement. However, the number of modes to perform RDO is still too much. And it does not consider the video content. TABLE I shows the percentages of the first 3,3,2,2 and 1 candidate prediction mode to be the best direction mode. It can be seen that most PUs will select the former ones to be the best prediction mode. Thus, we propose a shrinkage scheme for intra prediction to further reduce the number of the intra prediction mode considering the correlation between video content and possible number of prediction modes based on the gradient variance.

TABLE I

Percentages of the first 3,3,2,2 and 1 candidate direction to be the best direction

CU Size	Class A	Class B	Class C	Class D	Class E
64x64	58%	60%	51%	95%	65%
32x32	83%	81%	84%	86%	84%
16x16	84%	84%	85%	83%	88%
8x8	87%	87%	87%	86%	91%
4x4	83%	81%	79%	80%	86%

The proposed shrinkage scheme can be formulated as a problem to determine the number of prediction mode to perform RDO as follows.

$$\Psi_N(\theta) = \arg\min_{N,\ \theta\in N_h}\{D_{SATD,\theta} + \lambda \cdot R_\theta\} \quad (3)$$

where $\Psi_N(\theta)$ denotes the set of the chosen N prediction modes. $D_{SATD,\theta}$ is the SATD based R-D cost for the mode θ, and R_θ represents the bit rate for coding the index of mode.

For nature video sequence, texture information of different areas varies much, which results in distinct properties within a frame. In HEVC, as illustrated in Fig. 1, the area with rare texture information tends to be coded as a big coding unit. While the area with rich texture information tends to be partitioned into smaller coding units.

Fig. 1. The coding structure for a frame in RaceHoreses with HEVC

Furthermore, for the area with rare texture information, the distribution of Hadamard based R-D cost is in accordance with the SSE based R-D cost as illustrated in Fig. 2. The point with yellow color is the optimal mode for the PU. It can be observed the optimal prediction mode can be obtained from the less number of prediction mode with smaller Hadamard based R-D cost. Thus for the areas with rare texture information, less number of prediction mode is enough.

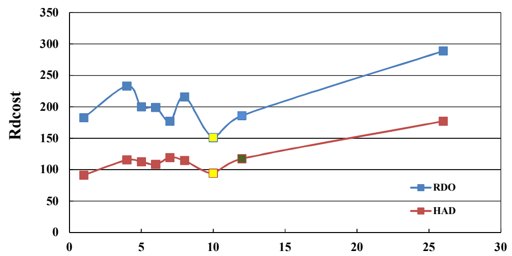


Fig. 2 The comparison of Hadamard based RD cost and the SSE based RD cost. The horizontal axis denotes the index of the intra prediction mode.

In this paper, we propose an adaptive scheme based on the gradient variance of the PU to determine the number of the candidate for full RDO. Firstly, for a given PU, the gradient of each pixel is calculated as follows,

$$Grad(I_{i,j}) = |I_{i,j} - I_{i,j+1}| + |I_{i,j} - I_{i+1,j}| \quad (4)$$

where $I_{i,j}$ denotes pixel value with the position as (i,j). For the PU with size of MxM, the gradient variance is calculated as follows,

$$Var_g = \frac{1}{M^2}\{\sum_{i=1}^{M}\sum_{j=1}^{M}(Grad(I_{i,j}) - \bar{G})^2\} \quad (5)$$

where $\bar{G}$ is the average gradient of each pixel. Considering the difference of PU size, the above formula is modified as:

$$Var_{g\_p} = \frac{1}{M^2} \cdot Var_g \quad (6)$$

Then the number of candidate mode can be represented as a function of *Varg_p*.

$$N = f(Var_{g\_p}) = [N_h \cdot \alpha + 0.5] \quad (7)$$

where α denotes the shrink factor for the number of prediction mode which is related to the *Varg_p* as illustrated in (8), and the function/notation [] denotes the rounding operation.

$$\alpha = \begin{cases} 1 & Var_{g\_p} > Th_1 \\ 3/4 & Th_1 \ge Var_{g\_p} > Th_2 \\ 2/4 & Th_2 \ge Var_{g\_p} > Th_3 \\ 1/4 & Var_{g\_p} \le Th_3 \end{cases} \quad (8)$$

III. Proposed Hadamard Transform Based Model for the Determination of Intra Mode Candidate

After determining the number of the prediction mode, the next step is to decide which prediction mode should be selected for full RDO. In HEVC, Hadamard transform based scheme is utilized to estimate the rate distortion cost for the determination of prediction mode as follows.

$$J_h(\theta) = SATD_\theta + \lambda \cdot R_{\theta,mode} \quad (9)$$

where $J_h(\theta)$ denotes the Hadamard transform based R-D cost, $R_{\theta,mode}$ indicates the bit rate for coding the prediction mode index. This method cannot reflect the true R-D cost especially when the texture information is rich for some coding unit. In this paper, based on the analysis of the Hadamard transform based coefficients, we propose an improved scheme for the determination of prediction mode for intra coding.

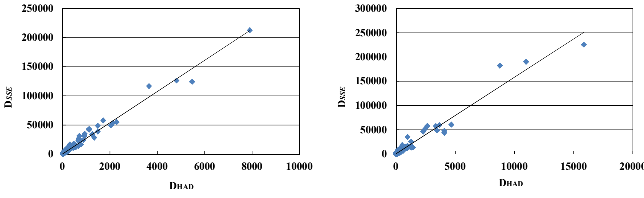
Firstly, the Hadamard transformed coefficients, $C(i,j)$, are quantized as follows.

$$C'(i,j) = C(i,j) >> ((QP-4)/6) \quad (10)$$

where $C'(i,j)$ denotes the quantized coefficients. QP is the quantization parameter.

Then the Hadamard based distortion, D_{HAD}, is defined as

$$D_{HAD} = \sum_{i=1}^{N}\sum_{i=1}^{N} C'(i,j) \quad (11)$$



(a) BlowingBubbles_416x240 (b) BQsqaure_416x240

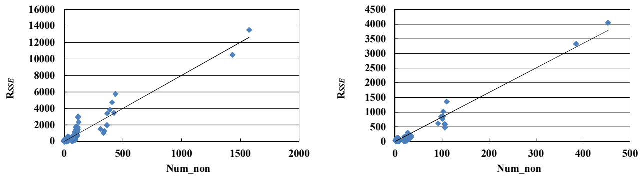
Fig.3 The correlation between SSE and Hadamard based distortion D_{HAD}

Fig.3 shows the correlation between the SSE and D_{HAD}. It can be seen that a linear correlation holds firmly between them with the intercept as zero. Thus the SSE can be represented in terms of D_{HAD} with a linear function as follows.

$$D_{SSE} = \beta \cdot D_{HAD} \quad (12)$$

In order to further present the correlation between the Hadamard transform and DCT. We also investigate the correlation between the true bit rate and the non-zero quantized Hadamard transformed coefficients as illustrated in Fig. 4. It can be observed that the true bit rate for a coding unit can be presented by a linear correlation approximately. Considering the influence of bit rate of mode index, the Hadamard transform based estimation mode of the bit rate is presented as (13).

$$R_{\theta,e} = \gamma.N_{HAD,nonzero} + R_{index} \quad (13)$$



(a) BlowingBubbles_416x240 (b) BQsqaure_416x240

Fig. 4 The correlation between bit rate and Hadamard transform based non-zreo coefficients

Combine (12) and (13), the proposed Hadamard transformed based estimation of R-D cost can be modelled as (14). Then J_{HAD} can be utilized for the determination of prediction candidate mode.

$$J_{HAD} = \mu \cdot D_{HAD} + \lambda_{mode} \cdot (R_{index} + \gamma.N_{HAD,nonzero}) \quad (14)$$

IV. Experimental Results

In order to verify the efficiency of the proposed R-D optimization scheme, we incorporate the proposed scheme into the HM10.0. The complexity reduction is presented as follows.

$$\Delta T = \frac{T_{anchor} - T_{pro}}{T_{anchor}} \times 100\% \quad (15)$$

where T_{anchor} and T_{pro} denote the encoding time of original HM anchor and our proposed algorithm respectively. When comparing the coding performance difference, we utilize the popular method proposed in [18] to calculate the difference between two R-D curves. In order to further show the efficiency of the proposed scheme, we propose an efficiency factor to represent the coding efficiency of the proposed scheme which is calculated as follows.

$$E = \frac{\Delta T}{BD\_rate} \quad (16)$$

where ΔT denotes the complexity reduction and the BD-rate indicates the R-D performance loss. The proposed E reflects the complexity reduction per BD-rate loss. With the same BD-rate, the scheme with higher complexity reduction is of high efficiency.

Firstly, experiments are conducted to verify the efficiency of the proposed estimation mode as in (14). The results are illustrated in TABLE II. The parameter μ in (14) is set as 0.5 and γ is set as $1/\lambda_{mode.}$ It can be seen that the proposed estimation mode can get better R-D performance under different number of prediction candidate modes. The average coding gain can be 0.3%, 0.4% and 0.3% with the candidate as one, two and three respectively. It indicates that the proposed scheme can contain the optimal prediction mode more accurately.

Secondly, experiments are implemented for the proposed intra optimization scheme including the gradient variance based shrinkage scheme of the number of candidate mode and the proposed estimation mode of rate distortion cost. The experimental results are provided in TABLE III. It can be observed that the proposed intra optimization scheme can improve the intra coding efficiently. The coding complexity reduction can be over 26% on average with only 0.52% BD-rate loss. TABLE III also provides the performance comparisons with other related works. It can be seen that the proposed scheme can achieve more complexity reduction and higher efficiency factor.

TABLE II

The coding performance of the proposed intra mode decision model [%]

Sequence	N=1			N=2			N=3		
	Y	U	V	Y	U	V	Y	U	V
ClassA	-0.4	-0.2	0.0	-0.4	-0.1	0.1	-0.3	0.0	0.1
ClassB	-0.4	-0.2	-0.1	-0.4	0.0	0.1	-0.3	0.0	0.1
ClassC	-0.3	-0.3	-0.2	-0.4	-0.2	-0.2	-0.3	-0.2	-0.1
ClassD	-0.3	-0.2	-0.1	-0.4	-0.1	-0.2	-0.3	-0.1	0.0
ClassE	-0.4	-0.2	-0.2	-0.4	-0.2	-0.1	-0.4	-0.1	0.0
Average	-0.3	-0.2	-0.1	-0.4	-0.1	-0.1	-0.3	-0.1	0.0

TABLE III

The coding performance of the proposed intra optimization scheme and performance comparisons with reference [16] and [17]

	Pro		ref[16]		ref[17]	
	BD-rate	ΔT	BD-rate	ΔT	BD-rate	ΔT
ClassA	0.44%	30.7%	0.55%	20.4%	0.68%	20%
ClassB	0.40%	25.9%	0.70%	20.5%	1.20%	21%
ClassC	0.49%	24.7%	0.74%	19.4%	0.98%	21%
ClassD	0.59%	26.8%	0.94%	19.5%	1.45%	19%
ClassE	0.72%	27.6%	0.84%	20.1%	1.05%	20%
Average	0.52%	26.6%	0.75%	19.98%	1.07%	20.20%
E	51.5		26.50		18.84	

V. CONCLUSIONS

In this paper, a fast intra optimization scheme is proposed to reduce the intra coding complexity for HEVC. Firstly, a shrinkage scheme for the number of prediction mode is proposed based on gradient variance. Then a novel estimation for rate distortion cost is provided based on Hadamard and quantization to better determine the prediction mode for RQT. Experimental results demonstrate that, compared with the original HEVC reference encoder implementation in AI configuration, the proposed algorithms can achieve more than 26% reduced encoding time on average with ignorable coding performance degradation.

ACKNOWLEDGEMENT

This work was supported in part by the National High-tech R&D Program of China (863 Program, 2012AA010805), National Science Foundation of China (61322106, 61390515), which are gratefully acknowledged.

REFERENCES

[1] ITU-T and ISO/IEC JTC 1, "Advanced Video Coding for generic audio-visual services," ITU T Rec. H.264 and ISO/IEC 14496-10 (AVC), May 2003.

[2] B. Bross, W.-J. Han, G. J. Sullivan, J.-R. Ohm and T. Wiegand, "High efficiency video coding (HEVC) text specification draft 9 (SoDIS)", JCTVC-K1003, Sep. 2012.

[3] G. J. Sullivan, J.-R. Ohm, W. J. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard", *IEEE Trans. on Circuits and Systems for Video Techno*logy, 2012, vol.22, no.12, pp. 1649-1668, 2012.

[4] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan and T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standards– Including High Efficiency Video Coding (HEVC)", I*EEE Transactions on Circuits and Systems for Video Technology,* , vol. 22, no. 12, pp. 1669-1684, 2012.

[5] I. K. Kim, J. Min, T. Lee, W. J. Han and J. Park, "Block Partitioning Structure in the HEVC Standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no.12, pp. 1697-1706, 2012.

[6] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74–90, Nov. 1998.

[7] J. Vanne, M. Viitanen, T. D. Hämäläinen and A. Hallapuro, "Comparative Rate-Distortion-Complexity Analysis of HEVC and AVC Video Codecs", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1885-1898, 2012.

[8] G. Correa, P. Assuncao, L. Agostini and L. Cruz, "Performance and Computational Complexity Assessment of High Efficiency Video Encoders", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no.12, pp. 1899-1909, 2012.

[9] F. Bossen, B. Bross, K. Suhring and D. Flynn, "HEVC Complexity and Implementation analysis", *IEEE Transactions on Circuits and Systems for Video Technology*, vol.22, no.12, pp. 1685-1696, 2012.

[10] L. Zhao, L. Zhang, X. Zhao, S. Ma, D. Zhao and W. Gao, "Further encoder improvement for intra mode decision", JCTVC-D283, Daegu, Jan. 2011.

[11] Y. Piao, J. Min, J. Chen, "Encoder improvement of unified intra prediction," JCTVC-C207, Guangzhou, Oct. 2010.

[12] L. Shen, Z. Zhang, P. An, "Fast CU Size Decision and Mode Decision Algorithm for HEVC Intra Coding", *IEEE Transactions on Consumer Electronics*, Vol. 59, No. 1, pp. 207-213, Feb, 2013.

[13] W. Jiang, H. Ma, Y. Chen, "Gradient Based Fast Mode Decision Algorithm for Intra Prediction in HEVC," CECNet, 2012.

[14] M. Zhang, C. Zhao, and J. Xu , "An adaptive fast intra mode decision in HEVC," ICIP, pp.221-224, 2012.

[15] H. Zhang, Z. Ma, "Fast Intra Mode Decision for High-Efficiency Video Coding (HEVC)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.24, no.4, pp.660-668, Apr. 2014.

[16] Motra A S, Gupta A, Shukla M, *et. al*. "Fast intra mode decision for HEVC video encoder," *International Conference on. IEEE Telecommunications and Computer Networks* (*SoftCOM*), pp.1-5, 2012.

[17] Y. Shi, O.C. Au, H. Zhang, *et al.* "Local saliency detection based fast mode decision for HEVC intra coding," MMSP, pp. 429-433, 2013.

[18] G. Bjontegaard, "Calculation of average PSNR difference between RD-curves," in Proc. ITU-T Q.6/SG16 VCEG 13th Meeting, Austin, TX, Apr. 2001, document VCEG-M33.