An All-zero Blocks Early Detection Method for High Efficiency Video Coding

Zhengguang Lv, Shengfu Dong*, Ronggang Wang, Xiaodong Xie, Huizhu Jia, Wenmin Wang and Wen Gao
Shenzhen Graduate School, Peking University, China

ABSTRACT

The High Efficiency Video Coding has a significant compression performance benefit versus previous standards. Thanks to the high efficiency prediction tools, blocks with all-zero quantized transform coefficients are quite common in HEVC. The computation load of transform and quantization can be remarkably reduced if the all-zero blocks can be detected prior to transform and quantization. Based on the theoretical analysis of the integer transform and quantization process in HEVC, we propose some SAD thresholds under which all-zero block can be detected. Simulation results show that with our proposed method, nearly 37% time saving for computation time of transform and quantization can be saved.

Keywords: All-zero Block, HEVC, video coding, early detection, discrete cosine transform (DCT), quantization, computational complexity, fast algorithm

I.INTRODUCTION

High Efficiency Video Coding standard is the most recent joint video project of the ITU-T VCEG and ISO/IEC MPEG standardization organizations, working together in a partnership known as the Joint Collaborative Team on Video Coding (JCT-VC) [1]. HEVC significantly outperforms previous standards such as H.264/AVC in term of coding efficiency. However, it also has a considerable increase in encoder complexity. Therefore, reducing the computational complexity of the encoder is vital for the application of this standard.

All-zero blocks (AZBs) are those blocks which all the transformed and quantized coefficients are all zeros. AZBs are quite common in low bit-rate video application [6]. The computations of transform and quantization need not be performed if AZBs can be detected prior to transform and quantization. Thus, early detection of AZBs can effectively reduce the computational complexity. Many works [3]-[10] have been done on H.264/AVC in this filed. For example, the earliest detection method for AZBs proposed in [3] defines a sufficient condition for quantizing all 8×8 floating-point transform coefficients to zero. [4]-[8] propose more precise conditions to detect 4×4 AZBs, in [8] an adaptive method is propose to detect the 4×4 AZBs, in [10] a fast and efficient mode decision algorithm based on all-zero blocks detection method is proposed. All of the zero quantized transform coefficients can effectively reduce the redundant computations.

Methods above are all based on the characteristics of DCT formula and quantization and the simulation platform are all on H.264/AVC. In this paper, all-zero blocks from 4×4 to 32×32 detection method is proposed for HEVC. Simulation results show that by utilizing our proposed method, the computations of transform and quantization in HEVC can be reduced by about 37% without coding performance loss.

Visual Information Processing and Communication V, edited by Amir Said, Onur G. Guleryuz,
Robert L. Stevenson, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 9029, 902902

The rest of the paper is organized as follows. Section II gives a review of the existing methods in H.264/AVC. In Section III , we derive a new AZBs early detection method for HEVC. Experimental results are presented in Section IV. Section V concludes this paper. And Section VI shows the acknowledge.

II.PREVIOUS METHOD FOR DETECTION AZB IN H.264/AVC

As we all know, in H.264/AVC, the integer transform is used in order to eliminate any mismatch problems between the encoder and the decoder. Thus, the conventional early detection method for the integer transform is derived in this section. Since it is based on Sousa's approach [3], we defined a 4×4 integer transform block e(x, y) as

$$E(u,v)=\sum_{x=0}^{3}\sum_{y=0}^{3} C(x,u)e(x,y)C(y,v) \qquad 0\le u,v\le 2$$

where

$$C(m,n)=\langle 2.5\cdot\sqrt{\frac{1}{2}}L(n)\cos\frac{(2m+1)n\pi}{8}\rangle \tag{1}$$

where $L(n)=1/\sqrt{2}$, for n = 0; and $L(n)=1$, otherwise. $\langle x\rangle$ denotes to round the operand x to the nearest integer. Given a transform coefficient $E(x,y)$ and a quantization parameter QP ranging from 0 to 51, the quantized coefficient

$$E(u,v)_Q=\text{sign}\{E(u,v)\}\times\frac{|E(u,v)|\cdot M\left(Q_p\%6,r\right)+f}{2^{qbits}} \tag{2}$$

where % denote the modular operator, qbits = 15+floor(Q_p/6), f = (2^{qbits})/6 for inter blocks and f = (2^{qbits})/3 for intra blocks. The quantization coefficient M (Q_p%6, r) in (2) is pre-defined for each frequency component and provided with a periodic table. It is formulated as follows [4]:

$$M\left(Q_p\%6,r\right)=\begin{bmatrix}5243 & 8066 & 12107\\ 4660 & 7490 & 11916\\ 4194 & 6554 & 10082\\ 3647 & 5825 & 9362\\ 3355 & 5243 & 8192\\ 2893 & 4559 & 7282\end{bmatrix}$$

where

$$r=2-(u\%2)-(v\%2) \tag{3}$$

III.PROPOSED ALGORITHM IN HEVC

In HEVC, we define some variables as follows [2]:

B = internal bit depth (as specified by InternalBitDepth in the common config files)

N = transform size

M = $\log_2(N)$

$Q[q_{rem}] = f(q_{rem}) = f(QP\%6)$,

where $f(x)=\{26214,23302,20560,18396,16384,14564\}, \quad x=0,\ldots,5$ (4)

Residual data is defined as e(x, y)

$$e(x,y)=f_1(x,y)-f_2(x,y) \tag{5}$$

In HEVC, there are core transform matrices of various sizes from 4×4 to 32×32. For all the blocks of residue data, an integer DCT is applied. Firstly, we deduce the SAD threshold for a 4×4 transform block. We can define a 4×4 transform unit e (x, y) like (1), because the coefficients of transform matrix in HEVC are nearly eight times of those in H.264/AVC. We replace 2.5 with $5\sqrt{2}$. The new integer transform is defined as

$$E(u,v)=\sum_{x=0}^{3}\sum_{y=0}^{3} C(x,u)\, e(x,y) C(y,v) \qquad 0\le u,v\le 3$$

where
$$C(m,n)=\left\langle 5\sqrt{2}\cdot\sqrt{\frac{1}{2}}L(n)\cos\frac{(2m+1)n\pi}{8}\right\rangle \tag{6}$$

where $L(n)=1/\sqrt{2}$, for n = 0; and $L(n)=1$, otherwise. $\langle x\rangle$ denotes to round the operand x to the nearest integer. And we have DCT transform formula like this:

$$E(u,v)=\left(\frac{2}{N}\right)k_u k_v\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} e(x,y)\cos\left(\frac{(2x+1)u\pi}{2N}\right)\times\cos\left(\frac{(2y+1)v\pi}{2N}\right) \tag{7}$$

where u, v = 0,1,...,N-1, and k_u, $k_v = 1/\sqrt{2}$, for u, v = 0, else k_u, k_v = 1. And we know the relationship of transform matrices between 4×4 to $N\times N$. So we can get the $N\times N$ integer transform formula as follows:

$$E(u,v)=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} C(x,u)\, e(x,y) C(y,v) \qquad 0\le u,v\le N-1$$

where
$$C(m,n)=\left\langle 5\sqrt{2}\cdot\sqrt{\frac{2}{N}}L(n)\cos\frac{(2m+1)n\pi}{2N}\right\rangle \tag{8}$$

Using the same way of (2), the quantization parameter (QP) ranges from 0 to 51, the quantized coefficient is calculated as follows:

$$E(u,v)_Q=\mathrm{sign}\big(E(u,v)\big)\times\big(|E(u,v)|\times Q[q_{rem}]+\mathrm{offset}\big)\gg qbit$$

where $$\text{qbit}=19+\frac{QP}{6}-M-(B-8) \tag{9}$$

$$\text{offset}=171\ll(\text{qbit}-9) \quad \text{for intra mode} \tag{10}$$

$$\text{offset}=85 \ll(\text{qbit}-9) \quad \text{for inter mode} \tag{11}$$

The absolute value of the coefficient E(x, y) is limited to

$$|E(u,v)|\leq\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}|e(x,y)||C(x,u)C(y,v)| \tag{12}$$

From (8) and (9), the following inequality is easily derived:

$$\left|E(u,v)_Q\right|\leq\left|\left(\left|\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}|e(x,y)||C(x,u)C(y,v)|\right|\times Q\left[q_{rem}\right]+\text{offset}\right)\gg\text{qbit}\right| \tag{13}$$

where $Q[q_{rem}] = f(q_{rem}) = f(QP\%6)$,

If $\left|E(x,y)_Q\right|\leq1$, then the upper bound of (13) is smaller than 1, and it is obvious that all transform coefficients are simultaneously quantized into zeros. Based on this fact, the sufficient condition is as follows:

$$\left|\left(\left|\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}|e(x,y)||C(x,u)C(y,v)|\right|\times Q\left[q_{rem}\right]+\text{offset}\right)\gg\text{qbit}\right|\leq1 \tag{14}$$

We then derive the following functions from formula (7):

$$|E(u,v)|_{\{u=0,v=0\}}\leq\langle\frac{5\sqrt{2}}{\sqrt{N}}\rangle\langle\frac{5\sqrt{2}}{\sqrt{N}}\rangle\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}|e(x,y)| \tag{15}$$

$$|E(u,v)|_{\{(u=0,v\neq0)\cup(u\neq0,v=0)\}}\leq\langle\frac{5\sqrt{2}}{\sqrt{N}}\rangle\langle\frac{10}{\sqrt{N}}\cos\frac{\pi}{2N}\rangle\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}|e(x,y)| \tag{16}$$

$$|E(u,v)|_{\{u\neq0,v\neq0\}}\leq\langle\frac{10}{\sqrt{N}}\cos\frac{\pi}{2N}\rangle\langle\frac{10}{\sqrt{N}}\cos\frac{\pi}{2N}\rangle\sum_{x=0}^{N}\sum_{y=0}^{N}|e(x,y)| \tag{17}$$

By using the theory of inequality zooming and (14), when N = 4, we get a SAD threshold of 4×4 block as follows:

$$SAD_{4\times4}=\sum_{x=0}^{3}\sum_{y=0}^{3}|e(x,y)|\leq\frac{1}{25}\times\frac{2^{\text{qbit}}-\text{offset}}{Q\left[q_{rem}\right]}=T_1 \tag{18}$$

where T_1 is the threshold of the 4×4 block. Note that the left-hand-side term of (18) is equal to the sum of absolute difference (SAD) for the 4×4 block. If $SAD_{4\times4}<T_1$, the 4×4 block is an all-zero block.

For other transform blocks 32×32, 16×16, and 8×8, we set $N \times N$ as the block size. The SAD thresholds for $N \times N$ can be approximately calculated as

$$\mathrm{SAD}_{N \times N} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |e(x,y)| \le \frac{N}{4} \times T_1, \qquad N=32,16,8,4 \tag{19}$$

Therefore, the threshold T_2 for 8×8 inter mode can be defined as $2T_1$, the threshold T_3 for 16×16 inter mode can be defined as $4T_1$ and the threshold T_4 for 32×32 inter mode can be defined as $8T_1$.

IV. EXPERIMENTAL RESULTS

In order to evaluate the performance of our proposed method, we integrate our method into the HEVC HM 10.0 encoder. RDOQ, adaptive quantization and the procedure of ScalingList are disabled. The test set has twenty sequences split in five classes depending on the resolution: class A(4K),class B(1080p),class C(WVGA) ,class D(WQVGA) and class E(720p). The coding configuration in common test conditions is low_delay (LD). The quantization parameter is set as QP = {32, 38, 42, 46}. We use our proposed method only for inter mode block.

TABLE I shows different measurements of the proposed method compare to the original HEVC. These measurements include TS(time saving), ∆RATE and ∆PSNR. Let subscript p denotes our proposed method, and subscript o denotes the original HEVC respectively. TQT is a representative run time of transform and quant. These measurements are defined as

$$TS = \frac{\mathrm{CodingTime}_o - \mathrm{CodingTime}_p}{TQT} \times 100\% \tag{20}$$

$$\Delta PSNR = PSNR_p - PSNR_o \tag{21}$$

$$\Delta RATE = \frac{RATE_p - RATE_o}{RATE_o} \times 100\% \tag{22}$$

TABLE I The overall performance of our proposed method

	TS[%]	**∆RATE[%]**	**∆PSNR-Y**	**∆PSNR-U**	**∆PSNR-V**
Class A	17.18%	-0.01%	-0.01	-0.02	-0.01
Class B	26.14%	-0.12%	0.00	0.00	0.00
Class C	11.01%	0.04%	0.00	0.01	-0.01
Class D	9.23%	-0.05%	0.00	-0.01	-0.01
Class E	57.50%	0.15%	0.00	0.05	0.06
Overall	37.34%	0.07%	0.00	0.02	0.03

From TABLE I we can see that the run time of transform and quantization can be saved nearly 37% without performance loss by our proposed method.

TABLE II Four parameters to measure the performance of our method

	RRAZB[%]	**DRAZB[%]**	**FAR[%]**	**FRR[%]**
Class A	93.96%	52.63%	0.45%	44.05%
Class B	93.87%	56.14%	0.55%	41.01%
Class C	89.16%	45.61%	1.02%	50.05%
Class D	85.59%	43.49%	0.98%	50.75%
Class E	97.99%	68.67%	0.56%	30.17%
Over-all	95.98%	60.65%	0.50%	37.11%

TABLE II shows the all-zero blocks detection rate of our method. We present the following four parameters to measure the performance of our method. The false rejection rate (FRR) and false acceptance rate (FAR) are provided to show the prediction capacity of AZBs for the AZB detection method enabled. The FRR and FAR are defined as

$$\mathrm{FRR}=\frac{N_{mz}}{N_z}\times 100\% \qquad \mathrm{FAR}=\frac{N_{mn}}{N_n}\times 100\% \tag{23}$$

N_{mz} is the number of AZBs being miss classified as non-AZBs, N_z is the total number of AZBs, N_{mn} is the number of non-AZBs being miss classified as AZBs, and N_n is the total number of non-AZBs. The smaller the FRR is, the more efficiently the approach detects AZBs. The smaller the FAR is, the less the video quality degrades. Thus, it is more desirable to have small FAR and FRR values for efficient prediction approaches. TABLE II shows the FAR and FRR, besides it show the other two parameters RRAZB and DRAZB. The real rate of AZBs (RRAZB) and the detection rate of AZBs (DRAZB) are also used to show the prediction capacity of AZBs for the AZB detection method enabled. They are defined as

$$\mathrm{RRAZB}=\frac{N_z}{N_z+N_n}\times 100\% \tag{24}$$

$$\mathrm{DRAZB}=\frac{N_{dz}}{N_z+N_n}\times 100\% \tag{25}$$

N_{dz} is the detected number of AZBs. The lager the RRAZB is, the more AZBs in the bit-streams. The lager the DRAZB is, the more AZBs are detected. We can see from TABLE II that we get reasonable results.

V. CONCLUSION

In this paper, an efficient algorithm is proposed to early detect all-zero blocks for HEVC. We perform a theoretical analysis of the transform and quantization used in HEVC and derive a sufficient condition to early detect all-zero blocks for various transforms. As a result, the proposed method can reduce redundant computations without performance loss.

VI. ACKNOWLEDGE

This work was partly supported by the grant of National Science Foundation of China 61370115, and Shenzhen Basic Research Program of JC201104210117A, JC201105170732A, and JCYJ2012061450301623.

REFERENCES

[1] I.-K. Kim, K. McCann, K. Sugimoto, B. Bross and W.-J. Han ,"High Efficiency Video Coding (HEVC) Test Model 10 (HM10) Encoder Description ,"Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11,12th Meeting: Geneva, CH, 14 – 23 Jan. 2013

[2] P. Topiwala , M. Budagavi , A. Fuldseth , R. Joshi and E. Alshina , "CE10: Core transform design," JCTVC-F910, 6th Meeting: Torino, IT, July 2011

[3] L. A. Sousa, "General method for eliminating redundant computations in video coding," Electron. Lett. , vol. 36, no. 4, pp. 306–307, Feb. 2000.

[4] Y. H. Moon, G. Y. Kim, and J. H. Kim, "An improved early detection algorithm for all-zero blocks in H.264 video encoding," IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 8, pp. 1053–1057, Aug. 2005.

[5] H. Wang, S. Kwong, and C. W. Kok, "Efficient prediction algorithm of integer DCT coefficients for H.264/AVC optimization," IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 4, pp. 547–552, Apr. 2006.

[6] Z. Xie, Y. Liu, J. Liu, and T. Yang, "A general method for detecting all-zero blocks prior to DCT and quantization," IEEE Trans. Circuits Syst. Video Technol., vol. 17, no. 2, pp. 237–241, Feb. 2007.

[7] H. Wang and S. Kwong, "Hybrid model to detect zero quantized DCT coefficients in H.264," IEEE Trans. Multimedia, vol. 9, no. 4, pp. 728–735, Jun. 2007.

[8] Maojun Zhang, Tao Zhou, and Wei Wang. "Adaptive Method for Early Detecting Zero Quantized DCT Coefficients in H.264 AVC Video Encoding," IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 19, NO. 1, JANUARY 2009

[9] Y. Zhao and Y. L. Yuan. "All-zero block detecting algorithm based on look-up threshold tables in H.264," in Proc. IEEE IDC, Aug. 2010, vol. 6, pp. 183-187.

[10] Guoyun Zhong, ling Lu, Niande Jiang. "Fast Mode Decision Based on All-zero block in H.264/AVC,"2011 Second International Conference on Digital Manufacturing & Automation

[11] JCT-VC, "High Efficiency Video Coding (HEVC) Test Model 10 (HM 10) Encoder Description", JCTVC-L1002, January 2013