Hybrid Zero Block Detection for High Efficiency Video Coding

Hongfei Fan, Ronggang Wang, Lin Ding, XiaodongXie, HuizhuJia, and Wen Gao, *Fellow, IEEE*

Abstract—**In this paper we propose an efficient hybrid zero block early detection method for high efficiency video coding (HEVC). Our method detects both genuine zero blocks (GZBs) and pseudo zero blocks (PZBs). For GZB detection, we use two sum of absolute difference bounds and one sum of absolute transformed difference threshold to decrease the GZB detection complexity. A fast rate-distortion estimation algorithm for HEVC is proposed to improve the PZB detection rate. Experimental results on HM platform show that the proposed method saves about 50% of the rate-distortion optimization time, with negligible Bjøntegaard delta bit rate loss. Our method is faster than other state-of-the-art ZB detection methods for HEVC by 10%~30%.**

Index Terms—**HEVC, RDO, zero block detection, genuine zero block, pseudo zero block, rate-distortion estimation**

I. INTRODUCTION

HIGH efficiency video coding (HEVC) is the most recent video compression standard issued jointly by the ISO/IEC Moving Picture Experts Group and the ITU-T Video Coding Experts Group [1]. HEVC doubles the coding efficiency of the previous standard, H.264/AVC, using new coding tools involving much higher computational complexity. Much work has been done on fast rate-distortion optimization (RDO) to accelerate the encoding process for HEVC. In this paper, we focus on the rapid detection of a special type of transform unit (TU) called zero block (ZB).

A TU is defined as a ZB if all of its transform coefficients are zeros. If the ZBs are detected early, some operations, such as transform, quantization, inverse quantization, inverse transform and entropy coding, can be skipped, so the encoding complexity is decreased. Two types of ZBs were defined in [16], i.e., genuine zero block (GZB) and pseudo zero block (PZB). A GZB is a TU in which all transform coefficients are quantized to zeros in the quantization process, while a PZB is a TU in which all transform coefficients are forced to zeros, to obtain better rate-distortion performance.

Many works have been done on GZB detection for H.264/AVC. In [2]-[11], the thresholds of the sum of absolute differences (SAD) were derived using various methods for H.264/AVC. In [14] and in our previous work [15], the SAD thresholds were extended from H.264/AVC to HEVC to accommodate large TUs of 16×16 and 32×32. Although the computational complexity of these SAD-based methods is low, their ZB detection rates are limited since the SAD cannot fully express the information in transform coefficients.

In [13], the Hadamard threshold was proposed for H.264/AVC using the relationship between the Hadamard transform and the integer discrete cosine transform (DCT). In [16], the SATD-based method was extended to HEVC for detecting both GZBs and PZBs, and the ZB detection rate was increased. However, as only 2×2, 4×4 and 8×8 Hadamard transforms were supported, multiple 8×8 Hadamard matrices were combined to simulate 16×16 and 32×32 Hadamard transforms. Therefore, the GZB detection for the 16×16 and 32×32 TUs was not as effective as the small TUs. Furthermore, empirical conditions of PZB detection were applied only for 16×16 and 32×32 TUs, the detection rates of 16x16 and 32x32 PZBs are limited, and no 4x4 or 8x8 PZB can be detected.

To further improve the ZB detection rate and decrease the encoding complexity, in this paper, we propose a new hybrid ZB detection method for HEVC. First, two bounds of SAD and one threshold of SATD are used to detect GZBs. Second, a fast rate-distortion cost estimation method for HEVC is proposed to detect PZBs. Experimental results on HM platform demonstrate that the proposed method can save about half of the RDO time with negligible Bjøntegaard delta bit rate (BD-BR) loss. Our method is faster than other state-of-the-art ZB detection methods for HEVC by 10%~30%.

Our method improves the state-of-the-art scheme in [16] in two ways. For GZB detection, a low bound of SAD is added to further reduce the GZB detection complexity, while for PZB detection, a fast rate-distortion estimation scheme is proposed to replace the empirical SATD-based detection conditions in [16], thus improving the PZB detection rate.

The remainder of this paper is organized as follows. Section II presents our proposed hybrid ZB detection method. Section III provides the experimental results to demonstrate the effectiveness of our method, and Section VI concludes the paper.

H. Fan, R. Wang and L. Ding are with Peking University Shenzhen Graduate School, Xili University Town Peking University Campus, Nanshan District, Shenzhen, 518055, China (e-mail: hffan@pku.edu.cn; rgwang@pkusz.edu.cn; ding.lin@pku.edu.cn).

H. Fan, L. Ding, X. Xie, H. Jia and W. Gao are with the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing100871, China (e-mail: donxie@pku.edu.cn; hzjia@pku.edu.cn; wgao@pku.edu.cn)

II. Proposed Hybrid ZB detection method

Inspired by our previous works and the work of [16], we propose a new hybrid ZB method for HEVC. Fig. 1 shows the pseudo codes of our method. In the first step, if the SAD of a TU is smaller than the first SAD threshold (SAD_{lower}), the TU is detected as a GZB; otherwise, if the SAD is larger than the second SAD threshold (SAD_{upper}), the TU is predicted as a non-ZB. In the second step, for the rest TUs, Hadamard thresholds are used to detect GZBs. In the third step, if the above conditions are not met, a fast rate-distortion cost estimation method is proposed to predict $Cost_{zero}$ and $Cost_{nonzero}$, and if $Cost_{zero}$ is smaller than $Cost_{nonzero}$ the TU is detected as a PZB; otherwise the TU is encoded as a non-ZB.

A. GZB detection

Although the GZB detection method in [16] effectively utilized the relationship between the Hadamard transform and the integer DCT, there are still some redundant computations. In our method, we propose adding an additional lower SAD threshold of SAD_{lower}, to early terminate the GZB detection. If the SAD of a TU is smaller than SAD_{lower}, it is directly detected as a GZB; otherwise, the GZB detection method in [16] is used as follows. If the SAD is larger than SAD_{upper}, the TU is detected as a non-ZB; otherwise the Hadamard thresholds in [16] are used to detect GZBs.

The quantization process of HEVC is defined as

$$QC = (C \times mult + offset) \gg shift, \qquad (1)$$

where C and QC denote the transform coefficient and quantized transform coefficient, respectively. $mult$, $offset$ and $shift$ are calculated respectively by

$$
\begin{aligned}
shift &= Q_{SHIFT} + D_{RANGE} + iPer - BitDepth - w \\
offset &= 85 \ll (shift - 9) \quad \text{for non-I slice} \\
&\quad\ 171 \ll (shift - 9) \quad \text{for I slice} \\
mult &= Q_{Array}[iRem],
\end{aligned}
$$

with

$$
\begin{aligned}
Q_{SHIFT} &= 14, \ D_{RANGE} = 15, iPer = QP/6, \\
w &= \log_2(W_{DCT}), iRem = QP\%6, \\
Q_{Array} &= \{26214, 23302, 20560, 18396, 16384, 14564\}, (2)
\end{aligned}
$$

where $BitDepth$ and W_{DCT} denote the pixel bit-depth and the TU width respectively.

In our previous work [15], we proposed the following SAD threshold for HEVC

$$SAD_{lower} = \frac{N}{100} \cdot \times \frac{2^{shift} - offset}{mult}, \qquad (3)$$

where N is the TU depth. As SAD_{lower} is derived from the nature of the HEVC integer DCT and quantization, almost all of the TUs with a SAD lower than SAD_{lower} are true ZBs, so we adopt it as the first SAD threshold.

To avoid redundant computations, an upper bound of SAD

```
if SAD<SAD_lower then
    GZB = 1;
else if SAD>SAD_upper then
    Non-ZB = 1;
else
    if NUM_hadmard = 0  then
        GZB = 1;
    else
        Fast RD cost Estimation
        if Cost_zero< Cost_nonzero then
            PZB = 1;
        else
            Non-ZB = 1;
            Update α
        end if
    end if
end if
```

Fig. 1. The proposed hybrid ZB detection.

was proposed in [16]

$$SAD_{upper} = \sqrt{\frac{C[W_{DCT}]}{k[W_{DCT}]} \times \sum UB[QP][W_{DCT}]^2}$$

with

$$UB[QP][W_{DCT}] = TH_{DCT}[QP][W_{DCT}] \times R_{UB} < |D - D'|, (4)$$

where $C[W_{DCT}]$ was obtained experimentally and set as 580, 148, 37 and 17 for TUs of 32×32, 16×16, 8×8 and 4×4, respectively. $|D - D'|$ denotes the quantization error. The value of $UB[QP][W_{DCT}]$ is set as 0.95*$|D - D'|$. R_{UB} is a constant value. $k[W_{DCT}]$ is a normalization factor and set as 1024, 256, 64, and 16 for 32×32, 16×16, 8×8 and 4×4 TUs respectively. Detailed explanation was given in [16].

Considering the similar frequency property of the Hadamard transform and the integer DCT, $TH_H[QP][W_{DCT}]$ was defined in [16] as the Hadamard threshold for a given QP and TU width, derived by

$$TH_H[QP][W_{DCT}] = TH_{DCT}[QP][W_{DCT}] \times m, \qquad (5)$$

where m equals 8, 2, 1/2 and 1/8 for the TUs of 32×32, 16×16, 8×8 and 4×4 respectively. The GZB detection condition is

$$if \ (NUM_{hadmard} = 0) \quad GZB = 1, \qquad (6)$$

where $NUM_{hadmard}$ denotes the number of non-zero Hadamard coefficients those are larger than $TH_H[QP][W_{DCT}]$.

As only 2×2, 4×4 and 8×8 Hadamard transforms were supported, DC Hadamard transform was used to detect 16×16 and 32×32 GZBs.

B. PZB detection

Even if a TU is not a GZB, it can still be encoded as a PZB to improve the rate-distortion performance. When the following condition is met the TU will be enforced as a PZB in RDO process.

$$Cost_{zero} < Cost_{nonzero}, \qquad (7)$$

Notation	Meaning
S	original signals
RS	reconstructed signals
P	predicted signals
R	residual signals
RR	reconstructed residual signals
D	DCT coefficients
RD	reconstructed DCT coefficients
T	DCT transform matrix
T^T	DCT inverse transform matrix

where $Cost_{zero}$ and $Cost_{nonzero}$ denote the rate-distortion costs to code the TU as ZB and non-ZB respectively.

The real value of $Cost_{nonzero}$ can only be calculated after transform, quantization, de-quantization, inverse transform and entropy coding. If PZB is detected earlier, some of the above operations can be skipped. The values of $Cost_{zero}$ and $Cost_{nonzero}$ for a TU are calculated by

$$Cost_{zero} = Dist_{zero} + \lambda \times Bits_{zero} \qquad (8)$$
$$Cost_{nonzero} = Dist_{nonzero} + \lambda \times Bits_{nonzero}, \qquad (9)$$

where $Dist_{zero}$ and $Dist_{nonzero}$ denote distortions to encode the TU as ZB and non-ZB, respectively. $Bits_{zero}$ and $Bits_{nonzero}$ represent the number of bits generated by coding this TU as ZB and non-ZB respectively. λ is a parameter to balance the distortion and bits in the rate-distortion cost, which can be derived using quantization parameter and coding structure [1].

To facilitate describing the derivation process of $Cost_{zero}$ and $Cost_{nonzero}$, a set of notations is defined in Table I.

The distortion is calculated by

$$Dist = \|S - RS\|^2 = \|(P + R) - (P + RR)\|^2$$
$$= \|R - RR\|^2. \qquad (10)$$

For $Dist_{zero}$, RR is zero. A ZB can be roughly coded using just one flag bit. Therefore,

$$\hat{C}ost_{zero} \approx Dist_{zero} + \lambda = \|R\|^2 + \lambda, \qquad (11)$$

where $\hat{C}ost_{zero}$ denotes the estimated value of $Cost_{zero}$.

The value of $Dist_{nonzero}$ is derived by

$$Dist_{nonzero} = \|T(R - RR)T^T\|^2$$
$$= \|T(R)T^T - T(RR)T^T\|^2$$
$$= \frac{\|D - RD\|^2}{k[W_{DCT}]} = \frac{\sum Dist_i^2}{k[W_{DCT}]}, \qquad (12)$$

where $k[W_{DCT}]$ has been given in (4).

For each transform coefficient, a visual description of the quantizer in (1) is shown in Fig. 2. First, a transform coefficient denoted by the blue point is shifted to the red point by adding the $offset$ in (1). Then, by subtracting the $discard$ with a rounding operation, the red point falling into

the quantization interval is moved to the black point (which denotes the reconstructed value). With (1), the distortion of a single coefficient can be approximated by

$$\widehat{D}ist_i = |D - RD|^2 = \left(\frac{|offset - discard|}{mult}\right)^2,$$

with

$$discard = (C \times mult + offset)\&[(1 \ll shift) - 1] \quad (13)$$

where $\widehat{D}ist_i$ denotes the estimated distortion of the ith coefficient and '&' denotes bitwise AND. The $Dist_{nonzero}$ is estimated by

$$\widehat{D}ist_{nonzero} = \frac{\sum \widehat{D}ist_i^2}{k[W_{DCT}]} = \frac{\sum|offset - discard|^2}{k[W_{DCT}] \cdot mult^2} \qquad (14)$$

In our previous work [17], a rapid method for estimating $Bits_{nonzero}$ was proposed for H.264/AVC. As new TU sizes were adopted in HEVC and some quantization differences between H.264/AVC and HEVC, we extend the method in [17] to accommodate multiple TU sizes and different quantization parameters. We also propose a different parameter updating method. The following self-information of the TU is used to estimate the rate

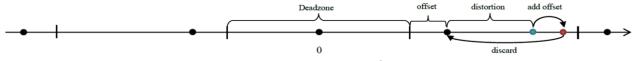$$r_{uv}^w = -\log_2 P\{\hat{C}_{uv}^w = \hat{x}\}, \qquad (15)$$

where w is the TU width, C_{uv}^w is the transform coefficient located at position (u, v) within one $w \times w$ TU and r_{uv}^w is the estimated bits. $\hat{C}_{uv}^w$ is the quantized coefficient of C_{uv}^w and $P\{\hat{C}_{uv}^w = \hat{x}\}$ denotes the probability of C_{uv}^w being quantized to $\hat{x}$.

We apply a zero-mean generalized Gaussian distribution to model the distribution of the transform coefficients. The rationale was given in our previous work [17]. As there are four different TU sizes in HEVC, we estimate the self-information by

$$\hat{r}_{uv}^w \approx \begin{cases} a_{uv} \cdot |\hat{x}|^{\eta_{uv}^w} + b_{uv}, & \hat{x} \neq 0 \\ a_{uv}^0 \cdot |f|^{\eta_{uv}^w} + b_{uv}^0, & \hat{x} = 0 \end{cases}'$$

with

$$\begin{cases} a_{uv}^w = \log_2 e \cdot \left[Q_{step} \cdot \alpha_{uv}^w(\eta_{uv}^w)/\sigma_{uv}^w\right]^{\eta_{uv}^w} \\ b_{uv}^w = -\log_2 \left[\frac{Q_{step} \cdot \eta_{uv}^w \alpha_{uv}^w(\eta_{uv}^w)}{(2 \cdot \sigma_{uv}^w \Gamma(1/\eta_{uv}^w))}\right] \\ a_{0uv}^w = \log_2 e \cdot \left[Q_{step} \cdot \frac{\alpha_{uv}^w(\eta_{uv}^w)}{\sigma_{uv}^w}\right]^{\eta_w} \\ b_{0uv}^w = -\log_2 \left[\frac{2(1 - f)Q_{step} \cdot \eta_{uv}^w \alpha_{uv}^w(\eta_{uv}^w)}{\left(2 \cdot \sigma_{uv}^w \Gamma\left(\frac{1}{\eta_{uv}^w}\right)\right)}\right] \end{cases}, (16)$$

where $\hat{r}_{uv}^w$ denotes the estimated value of r_{uv}^w and $\Gamma(\cdot)$ is the gamma function. The other variables in (19) are calculated as

Fig. 2. Quantization process of a single transform coefficient in HEVC.

$$Q_{step} = \frac{1 \ll shift}{mult}, f = \frac{offset}{1 \ll shift}$$

$$\alpha_{uv}^w(\eta_{uv}^w) = \sqrt{\frac{\Gamma(3/\eta_{uv}^w)}{\Gamma(1/\eta_{uv}^w)}}$$

$$\eta = R^{-1}\left(\frac{\left(\frac{1}{N}\sum_{i=1}^{N}|X_i|\right)^2}{\frac{1}{N}\sum_{i=1}^{N}X_i^2}\right), \quad \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}X_i^2}$$

$$R^{-1}(x) \approx \frac{0.2718}{0.7697 - x} - 0.1247, \qquad (17)$$

where X_i denotes the transform coefficient. In our experiment, η_{uv}^w and σ_{uv}^w are derived from the transform coefficients of the previous frame.

To avoid redundant computation, considering that most of the quantized DCT coefficients are smaller than 200 when PZB detection is activated, we accelerate the calculation using a look-up table. The value of $\hat{r}_{uv}^w$ for the quantized coefficient (smaller than 200) is pre-calculated and stored in the look-up table. The prediction accuracy is improved by updating η_{uv}^w and σ_{uv}^w dynamically during encoding. We update η_{uv}^w and σ_{uv}^w after each frame is encoded and the η_{uv}^w and σ_{uv}^w of the current frame are used to estimate the $Bits_{nonzero}$ for the next frame.

As self-information is the minimum number of bits needed to code a value theoretically, there may be some differences between the self-information and the actual number of bits. The estimated value of $Bits_{nonzero}$ is assumed to be linearly correlated with the number of actual bits. We estimate $Bits_{nonzero}$ by

$$\hat{B}its_{nonzero} \approx \alpha_w \cdot \left(\sum_u \sum_v \hat{r}_{uv}^w + \beta_w\right), \qquad (18)$$

where $\hat{B}its_{nonzero}$ represents the estimated value of $Bits_{nonzero}$, w denotes TU width, α_w and β_w are used to adjust the estimation accuracy of $Bits_{nonzero}$. $\hat{B}its_{nonzero}$ is 1 when all of the transform coefficients are zeros and the intercept is a constant value. Therefore, β_w is initialized to $-\sum_u \sum_v \hat{r}_{uv}^w$ when $\hat{x}$ in (16) is 0. We only update α_w during encoding to reduce the computational complexity, and α_w is initialized as 1. When a TU is detected as non-ZB, the encoder counts the actual $Bits_{nonzero}$ and α_w is updated accordingly. Considering that many TUs are non-ZBs, the α_w is updated often, so a low computational updating method is preferred. We propose the following updating method for α_w

$$\alpha_w^k = \frac{1}{T_{update}}\left(\sum_{T_{update}-1}^{i=1}\alpha_w^{k-i} + \frac{Bits_{nonzero}^{k-1}}{\hat{B}its_{nonzero}^{k-1}} \cdot \alpha_w^{k-1}\right) \quad (19)$$

Resolution	Sequence	Paper [17]				Proposed			
		4×4	8×8	16×16	32×32	4×4	8×8	16×16	32×32
1080P	BasketballDrive	2.17	4.99	24.48	38.04	1.73	3.78	11.48	21.37
	BQTerrace	3.96	14.73	69.70	43.70	3.12	7.67	31.61	23.00
	Cactus	3.26	30.68	98.67	152.57	2.71	15.63	38.21	72.27
	ParkScene	3.62	8.60	51.64	53.34	2.81	5.69	21.04	46.57
480P	BasketballDrill	2.19	6.66	19.87	16.16	1.50	3.71	9.77	10.05
	BQMall	3.76	16.09	55.89	47.90	2.86	9.72	33.50	35.74
	PartyScene	3.05	10.83	37.81	48.45	2.39	6.29	18.83	38.78
	RaceHorses	3.07	9.34	44.75	33.35	2.56	5.98	23.55	21.60
Average Difference		3.14	12.74	50.35	54.19	2.46	7.31	23.50	33.67

where α_w^k denotes the α_w updated with the $(k-1)$th non-ZB and is used to estimate the kth $Bits_{nonzero}$. $Bits_{nonzero}^{k-1}$ and $\hat{B}its_{nonzero}^{k-1}$ represent the actual value and the estimated value of $Bits_{nonzero}$ for the $(k-1)$th non-ZB, respectively. T_{update} is a variable used to control the updating frequency. If the T_{update} is too large, the estimation value does not reflect the changes. If the T_{update} is too small, the estimation accuracy of $Bits_{nonzero}$ tends to be affected by outliers. To achieve a balance, we set T_{update} to 10 in our experiments.

In our previous work [17], only 15 actual values of $Bits_{nonzero}$ were calculated by the entropy coder to estimate the subsequent 100 values of $Bits_{nonzero}$. The estimation accuracy is limited. Our proposed method in this paper handles this issue well by updating α_w once after coding each non-ZB. In Table II, the average absolute differences between the estimated value and the actual value of $Bits_{nonzero}$, using our proposed method and the method in [17], were compared. These results were obtained with different QPs (22, 27, 32 and 37) and two configurations (random access and low delay). The results showed that the accuracy of the bit-rate estimation is much improved with our proposed updating strategy over all of the test sequences.

As the residue of intra prediction is relatively large and the ratio of ZBs is low for intra TUs, the time saving for ZB detection is limited for intra TUs. Thus, similar to other state-of-the-art works [15][16], we apply our method only on inter TUs.

III. EXPERIMENTAL RESULTS

We integrated our proposed hybrid ZB detection method into the HEVC reference software HM 12.0. Both random access and low delay configurations were used. The related methods in references [16] and [15] were also implemented and compared with our method. As there were very few true PZBs when the RDOQ was turned on, in our experiments, we

TABLE III   NOTATION USED IN EVALUATING THE ERROR RATE

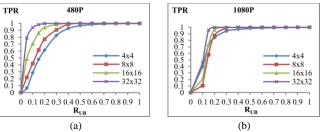| Notation	Full name	Definition
TP	True Positive	ZBs which are predicted as ZBs
TN	True Negative	Non-ZBs which are predicted as non-ZBs
FP	False Positive	Non-ZBs which are predicted as ZBs
FN	False Negative	ZBs which are predicted as non-ZBs
TPR	True Positive Rate	TP / (TP + FN)
FNR	False Negative Rate	FN / (TP + FN)
FPR	False Positive Rate	FP / (FP + TN)

(a)   (b)

Fig.3. The TPR of different R_{UB} for different resolutions (a) 480P (b) 1080P. $UB[QP][W_{DCT}]$ is used in both our method and the method in [16] for fair comparison.

first turned off RDOQ to show the effectiveness of the PZB detection. HM 12.0 was set as the anchor and the BD-BR performance was obtained with QPs of 22, 27, 32 and 37. In Table III we define the notations used in evaluating the ZB detection performance.

As the value of R_{UB} in formulation (4) was not given in [16], we tested different values of R_{UB} from 0 to 1, and the results are given in Fig. 3. Four curves representing four different TU sizes of 4×4, 8×8, 16×16 and 32×32 were given. Considering that SAD_{upper} was only used as a rough estimation, we paid more attention to the true positive rate (TPR) than the false positive rate (FPR), as non-ZBs misclassified as ZBs using condition (4) can be further corrected with the conditions in (6). Therefore, we choose R_{UB} which obtains a TPR greater than 0.95. In Fig. 3, the horizontal axis and vertical axis represent R_{UB} and the TPR, respectively. According to the experimental results shown in Fig.3, setting R_{UB} as 0.5 is reasonable. The value of

Table IV compares the proposed method with our previous work of [15]. The 'TS' column denotes the time saving of the RDO process. The 'BD' column denotes the BD-BR performance. 'LD' and 'RA' in 'CFG' denote low delay and random access configurations, respectively. All of the results were obtained with QPs of 22, 27, 32 and 37. Our method reduced the RDO time by 30% compared with the method in [15], under a similar BD-BR performance.

Table V shows the BD-BR and time savings for our method and the method in [16], with the RDOQ off. The test sequences recommended by JCT-VC were all tested. The 'GZB' and "All ZB" columns in Table VI list the results with only GZB detection and both GZB and PZB detections. The data were obtained by averaging the results with the QPs of 22, 27, 32 and 37. Compared with the method in [16], for GZB detection, the proposed method achieved a similar

TABLE V   THE AVERAGE RD PERFORMANCE AND TIME SAVING OF THE PROPOSED ALGORITHM AND THAT IN [16] FOR THE REST SEQUENCES

Resolution	Sequence	Random Access								Low Delay							
		Paper [16]				Proposed				Paper [16]				Proposed			
		GZB		All ZB		GZB		All ZB		GZB		All ZB		GZB		All ZB	
		TS	BD	TS	BD	TS	BD	TS	BD	TS	BD	TS	BD	TS	BD	TS	BD
Class A	People On Street	23.20%	0.47%	22.88%	0.57%	25.00%	0.54%	30.46%	0.53%	21.99%	0.32%	22.21%	0.60%	23.86%	0.40%	25.40%	0.61%
	Traffic	44.48%	0.33%	45.10%	0.63%	48.01%	0.42%	55.25%	0.77%	40.41%	0.10%	41.26%	0.57%	43.77%	0.16%	50.33%	0.94%
Class B	Kimono	11.07%	0.00%	11.27%	0.00%	11.17%	0.00%	9.11%	-0.01%	10.21%	-0.02%	10.80%	-0.02%	10.60%	-0.02%	10.44%	0.02%
	Basketball Drive	27.99%	0.28%	32.56%	0.43%	33.87%	0.45%	43.22%	0.61%	28.50%	0.32%	31.78%	0.47%	34.41%	0.62%	37.94%	0.76%
	BQ Terrace	41.74%	0.49%	47.94%	0.51%	47.60%	0.47%	61.45%	0.61%	40.56%	0.46%	42.68%	0.44%	45.69%	0.45%	53.87%	0.70%
	Cactus	34.30%	0.33%	39.92%	0.39%	40.20%	0.38%	54.06%	0.66%	31.67%	0.43%	33.93%	0.31%	36.60%	0.47%	44.19%	0.90%
	Park Scene	40.13%	0.47%	44.49%	0.52%	47.61%	0.51%	56.74%	0.68%	38.19%	0.51%	40.38%	0.55%	43.25%	0.57%	46.40%	0.89%
Class C	Basketball Drill	31.14%	0.16%	34.69%	0.35%	37.26%	0.25%	50.12%	0.42%	30.69%	0.14%	34.08%	0.24%	35.80%	0.10%	48.29%	0.62%
	BQ Mall	37.24%	0.47%	40.14%	0.54%	43.12%	0.41%	54.78%	0.59%	36.03%	0.58%	37.06%	0.75%	40.42%	0.54%	53.06%	1.00%
	Party Scene	30.57%	0.69%	32.97%	0.66%	34.58%	0.79%	46.10%	0.62%	27.56%	0.61%	34.09%	0.83%	29.60%	0.70%	41.13%	0.80%
	Racehorses	20.49%	0.51%	20.02%	0.52%	25.10%	0.50%	37.97%	0.38%	18.99%	0.55%	20.42%	0.62%	22.72%	0.56%	33.58%	0.43%
Class D	Basketball Pass	37.69%	0.22%	38.76%	0.51%	40.27%	0.28%	49.35%	0.68%	34.51%	0.50%	36.05%	0.86%	37.53%	0.79%	44.93%	0.80%
	Blowing Bubbles	31.49%	0.35%	33.30%	0.54%	33.76%	0.41%	47.98%	0.71%	25.73%	0.18%	28.74%	0.68%	28.92%	0.26%	37.85%	0.79%
	BQ Square	41.55%	0.41%	42.19%	0.72%	44.63%	0.29%	58.02%	0.58%	36.35%	0.46%	38.05%	0.80%	39.49%	0.43%	49.97%	0.48%
	Racehorses	18.21%	0.70%	19.72%	0.30%	19.65%	0.72%	28.18%	0.30%	15.78%	0.40%	17.82%	0.41%	18.10%	0.45%	20.33%	0.34%
Class E	Four People	53.61%	0.03%	56.27%	0.24%	58.85%	0.08%	67.81%	0.20%	49.66%	0.01%	52.86%	0.30%	54.90%	0.08%	64.09%	0.36%
	Johnny	54.93%	-0.03%	57.03%	0.37%	60.77%	0.20%	70.36%	0.42%	50.29%	0.22%	53.22%	0.57%	56.35%	0.25%	64.76%	0.46%
	Kristen and Sara	51.59%	0.18%	54.68%	0.37%	58.30%	0.32%	66.11%	0.39%	46.79%	0.04%	47.69%	0.26%	52.48%	0.11%	58.39%	0.33%
Class F	China Speed	21.78%	0.75%	20.65%	0.66%	25.42%	0.74%	31.37%	0.47%	22.35%	0.37%	22.87%	0.46%	26.10%	0.42%	30.34%	0.46%
	Slide Editing	54.98%	0.04%	56.43%	0.08%	61.26%	0.06%	67.34%	0.06%	54.94%	0.19%	56.98%	0.41%	61.29%	0.23%	69.11%	0.20%
	Slideshow	22.77%	-0.75%	22.37%	-0.28%	26.22%	-0.21%	25.31%	-0.65%	20.72%	0.10%	21.24%	0.42%	22.90%	0.16%	18.75%	1.01%
Average		34.81%	0.29%	36.83%	0.41%	39.17%	0.36%	48.15%	0.43%	32.47%	0.31%	34.49%	0.50%	36.42%	0.37%	43.01%	0.61%
All Average										33.64%	0.30%	35.66%	0.46%	37.80%	0.37%	45.58%	0.52%

Resolution	Sequence	Paper [15]		Proposed	
		TS	BD	TS	BD
1080P	Basketball Drive	13.47%	0.36%	40.58%	0.69%
	BQ Terrace	18.04%	0.20%	57.66%	0.66%
	Cactus	15.32%	0.06%	49.13%	0.78%
	Park Scene	15.21%	0.24%	51.57%	0.79%
480P	Basketball Drill	21.86%	0.16%	49.21%	0.18%
	BQ Mall	18.37%	0.27%	53.92%	0.48%
	Party Scene	13.97%	0.35%	43.62%	0.75%
	Racehorses	17.31%	0.39%	35.78%	0.53%
Average		16.69%	0.25%	47.68%	0.61%

Resolution	TU	CFG	Paper [16]			Proposed		
			FNR	FPR	Distance	FNR	FPR	Distance
1080P	4×4	RA	0.105	0.016	0.106	0.014	0.078	0.079
		LD	0.089	0.022	0.092	0.027	0.099	0.103
	8×8	RA	0.239	0.022	0.240	0.050	0.113	0.124
		LD	0.225	0.029	0.227	0.096	0.163	0.190
	16×16	RA	0.128	0.182	0.223	0.089	0.150	0.174
		LD	0.137	0.216	0.256	0.114	0.200	0.231
	32×32	RA	0.348	0.085	0.358	0.240	0.141	0.278
		LD	0.401	0.084	0.409	0.301	0.163	0.343
480P	4×4	RA	0.115	0.008	0.115	0.017	0.062	0.064
		LD	0.097	0.020	0.099	0.037	0.078	0.086
	8×8	RA	0.246	0.009	0.246	0.062	0.129	0.143
		LD	0.225	0.016	0.225	0.100	0.165	0.193
	16×16	RA	0.192	0.076	0.206	0.135	0.123	0.183
		LD	0.212	0.094	0.232	0.154	0.160	0.222
	32×32	RA	0.466	0.028	0.467	0.236	0.096	0.255
		LD	0.525	0.026	0.525	0.283	0.125	0.309
Average		RA	0.230	0.053	0.245	0.105	0.112	0.163
		LD	0.239	0.063	0.258	0.139	0.144	0.210
		All	0.235	0.058	0.252	0.122	0.128	0.187

BD-BR performance and saved 4.16% of the RDO time. For all of the ZB detection (both GZB and PZB), the proposed method obtained similar BD-BR performance and saved 9.92% of the RDO time. Table VI compares the FNR (false negative rate) and FPR (false positive rate) of ZB detection of these two methods. The FNR and FPR are the average values over all sequences and QP values. The 'Distance' column provides the Euler distance between (FNR, FPR) and (0, 0), to measure the overall detection performance. The smaller the 'Distance' is, the better detection performance that can be obtained. The FNR of our method is much lower than that of [16] for all TU sizes, especially for 16×16 and 32×32, which means that more PZBs are detected by our method. Although the FPR of our method is a little larger than the method in [16], the BD-BR performances of the two methods are still similar. Since for most of the non-ZBs being incorrectly detected as ZBs, the $Cost_{zero}$ and $Cost_{nonzero}$ of them are very similar and these blocks caused very little BD-BR performance loss. As shown in Table VI, the 'Distance' of our method is much smaller than the method in [16].

IV. CONCLUSIONS

There are two main technical contributions in our method. First, we introduce an additional lower bound of SAD to further decrease the computational complexity of GZB detection. Second, we propose a fast RD cost estimation for HEVC to detect PZBs, which considerably improves the PZB detection performance. Experimental results show that our proposed method is faster than other state-of-the-art methods with similar BD-BR performance from 10% to 30%.

ACKNOWLEDGMENT

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] J.F. Yang, S.C. Chang, C.Y. Chen, "Computation reduction for motion search in low rate video coders," IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, no. 10, pp. 948-951, Oct. 2002.

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

[4] H.L. Wang, S. Kwong, C.W. Kok, "Efficient prediction algorithm of integer DCT coefficients for H.264/AVC optimization," IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, no. 4, pp. 547–552, April 2006.

[5] H.L. Wang, S. Kwong, "Hybrid model to detect zero quantized DCT coefficients in H.264," IEEE Transactions on Multimedia. vol. 9, no. 4, pp. 728-735, June 2007.

[6] D. Wu, K.P. Lim, T.K Chiew, "An adaptive thresholding technique for the detection of all-zeros blocks in H.264," IEEE International Conference on Image Processing, vol. 5, pp. 329–332, Sept. 2007.

[7] C.Y. Su, Sh. L. Chang, "Adaptive Early Termination for Fast H.264 Video Coding," Ninth IEEE International Symposium on Multimedia. pp. 72-77, Dec. 2007.

[8] M.J. Zhang, T. Zhou, W. 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, pp. 103–107, Jan. 2009.

[9] Y. Zhao, Y.L. Yuan, "All-zero block detecting algorithm based on look-up threshold tables in H.264," The 6th International Conference on Digital Content, Multimedia Technology and its Applications. pp. 183-187, Aug. 2010.

[10] W.Y. Chiu, Y.M. Lee, Y.Y. Lin, "Advanced zero-block mode decision algorithm for H.264/AVC video coding," TENCON 2010 - 2010 IEEE Region 10 Conference, pp. 687-690, Nov. 2010.

[11] G.Y. Zhong, L. Lu, N.D. Jiang, "Fast Mode Decision Based on All-zero block in H.264/AVC," 2011 Second International Conference on Digital Manufacturing and Automation, pp. 535-538, Aug. 2011.

[12] C. Zhu, H.Z. Jia, J. Liu, X.H. Ji, "Multi-Level Low-Complexity Coefficient Discarding Scheme for Video Encoder," IEEE International Symposium on Circuits and Systems, pp. 5-8, Jun. 2014.

[13] H. Wang, S. Kwong, "Prediction of zero quantized DCT coefficients in H.264/AVC using Hadamard transformed information," IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 4, pp. 510–515, Apr. 2008.

[14] P.T. Chiang, T.S. Chang, "Fast Zero Block Detection and Early CU Termination for HEVC Video Coding," 2013 IEEE International Symposium on Circuits and Systems, pp. 1640-1643, May 2013.

[15] Z.G. Lv, S.F. Dong, R.G. Wang, X.D. Xie, "An all-zero blocks early detection method for high-efficiency video coding," Proc. SPIE Visual Information Processing and Communication V, vol. 9029, Feb. 2014.

[16] K. Lee, H.J. Lee, J. Kim, Y. Choi, "A Novel Algorithm for Zero Block Detection in High Efficiency Video Coding," IEEE Journal on Selected Topics in Signal Processing, vol. 7, no. 6, pp. 1124–1134, Dec. 2013.

[17] X. Zhao, J. Sun, S.W. Ma, W. Gao, "Novel Statistical Modeling, Analysis and Implementation of Rate-Distortion Estimation for H.264 AVC Coders," IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 5, pp. 647–660, May 2010.

Hongfei Fan received the B.S. degree in software from Shanghai Jiao Tong University, Shanghai, China, in 2013. From Sep. 2013, he is working toward the Ph.D. degree in the School of Electronics Engineering and Computer Science, Peking University, Beijing, China. His research interests include video coding and image processing.

Ronggang Wang is an Associate Professor with Peking University Shenzhen Graduate School. His research interest is on stereo video and mobile video processing. He has done many technical contributions to the China AVS standard, such as subpel motion compensation, background-predictive picture, and field coding. He has authored about 30 papers in international journals and conferences. He holds more than 20 patents in the field of video coding and processing. He serves as a reviewer of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, Signal Processing: Image Communication, and GlobalCom.

Lin Ding received the B.S. degree in computer science from University of Electronic Science and Technology of China, Chengdu, China, in 2013. From Sep. 2013, she is working toward the Ph.D. degree in the School of Electronics Engineering and Computer Science, Peking University, Beijing, China. Her research interests include feature coding and video coding.

Xiaodong Xie received his Ph.D. degree in electrical engineering, University of Rochester, USA. He was a Senior Scientist at Eastman Kodak Company, New York, USA, from 1994 to 1997; a Principal Scientist at Broadcom Corporation, California, USA, from 1997 to 2009.

He is currently a Professor with the Department of Electrical Engineering, Peking University, Beijing, China. His research interests include multimedia SoC design, embedded system. He holds 24 U.S. Patents.

Huizhu Jia received his Ph.D. degree in electrical engineering from the Chinese Academy of Sciences, Beijing, China, in 2007. He is currently with the National Engineering Laboratory for Video Technology, Peking University, Beijing. His research interests include: ASIC design, image processing, multimedia data compression and VR.

Wen Gao received the Ph.D. degree in electronics engineering from the university of Tokyo, Japan, in 1991. He is a professor in computer science at Peking University, China. Before joining Peking University, he was a professor at Harbin Institute of Technology from 1991 to 1995, and a professor at the Institute of Computing Technology of Chinese Academy of Sciences from 1996 to 2006. He has published extensively including five books and over 600 technical articles in refereed journals and conference proceedings in the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics.

Prof. Gao served or serves on the editorial board for several journals, such as IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Multimedia, IEEE Transactions on Autonomous Mental Development, EURASIP Journal of Image Communications, Journal of Visual Communication and Image Representation. He chaired a number of prestigious international conferences on multimedia and video signal processing, such as IEEE ICME and ACM Multimedia, and also served on the advisory and technical committees of numerous professional organizations. He is a member of Chinese Academy of Engineering.