J. Vis. Commun. Image R. 25 (2014) 1784–1790

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Early detection of all-zero 4 × 4 blocks in High Efficiency Video Coding

Hanli Wang [a,*], Han Du [a], Weiyao Lin [b], Sam Kwong [c], Oscar C. Au [d], Jun Wu [a], Zhihua Wei [a]

[a] Department of Computer Science & Technology, Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai 200092, PR China
[b] Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, PR China
[c] Department of Computer Science, City University of Hong Kong, Hong Kong, PR China
[d] Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, PR China

ARTICLE INFO

Article history:
Received 19 June 2013
Accepted 26 August 2014
Available online 3 September 2014

Keywords:
HEVC
Zero coefficients
DCT
Quantization
Fast encoding
Video quality
Compression efficiency
Early detection

ABSTRACT

Similar to previous video coding standards, transform and quantization are employed in the most recent video coding standard High Efficiency Video Coding (HEVC), and there exist a large number of transform coefficients being quantized to zeros. In order to reduce the computations in transform and quantization, a novel prediction algorithm is designed to early detect all-zero 4 × 4 blocks prior to transform and quantization. A theoretical analysis is performed on the 4 × 4 Discrete Cosine Transform (DCT) and the related quantization functions in HEVC; consequently, two sufficient conditions to early detect all-zero 4 × 4 blocks are derived and the proposed prediction algorithm is developed while balancing the detection efficiency and computational overheads. Experimental results demonstrate that the proposed algorithm is able to efficiently predict all-zero 4 × 4 blocks and reduce redundant DCT and quantization computations while maintaining the same coding performance in terms of video quality and bit rates.

1. Introduction

Owing to the ever-increasing demand for multimedia entertainment, the growing popularity of High Definition (HD) video and the emergence of Ultra High Definition (UHD) video (such as the resolution of 4 k × 2 k and 8 k × 4 k) are creating strong needs to better video coding efficiency than that of the state-of-the-art video coding standard H.264/Advanced Video Coding (AVC) [1]. To this aim, the Joint Collaborative Team on Video Coding (JCT-VC) designs a new video coding standard known as High Efficiency Video Coding (HEVC) [2,3] with one of the primary objectives being offering about two times the compression efficiency of H.264/AVC under the same video quality [4]. Besides the superior performance in compression efficiency, the targets of HEVC reside in other multiple-folds such as the implementability with parallel processing techniques [5].

The elementary difference between HEVC and previous video coding standards is that the quadtree structure [6] is employed in HEVC, which is a flexible mechanism to split a picture into different-size blocks not only for prediction but also for residual coding. The basic processing unit in HEVC is a Coding Tree Unit (CTU) which can be regarded as a generalization of a macroblock in H.264/AVC. A CTU can be set as one of the following block sizes: 64 × 64, 32 × 32 or 16 × 16, and it can be recursively subdivided into smaller Coding Units (CU) according to the quadtree structure. Each CU is further subdivided into Prediction Units (PU) for either intra or inter prediction. After the prediction residual is generated, transform is applied in HEVC with a nested Residual QuadTree (RQT) being utilized for each CU, *i.e.*, each CU is partitioned into Transform Units (TU). According to the HEVC specification, the TU transform size ranges from 4 × 4 to 32 × 32 for luma and from 4 × 4 to 16 × 16 for chroma. The determination of the block partitioning size including that of CU, PU and TU is usually made in the joint rate-distortion optimized manner [7]. In addition to the quadtree structure, HEVC adopts a number of other advanced coding techniques, such as an enriched set of angular intra prediction modes [8], sample adaptive offset processing [9], and in-loop filter [10], to further improve the compression efficiency.

Although the compression efficiency has been greatly improved, the computational complexity of HEVC is remarkably increased with a detailed analysis shown in [11]. A comparison of the encoding computations between H.264/AVC and HEVC is presented in [12], which indicates that the required amount of encoding computations of HEVC is much higher than that of H.264/AVC and it is highly desired to design fast encoding techniques to reduce the computational complexity of HEVC. To achieve this, a number of research works have been proposed, such

---

* Corresponding author.

E-mail addresses: hanliwang@tongji.edu.cn (H. Wang), 1233703@tongji.edu.cn (H. Du), hellomikelin@gmail.com (W. Lin), cssamk@cityu.edu.hk (S. Kwong), eeau@ust.hk (O.C. Au), wujun@tongji.edu.cn (J. Wu), Zhihua_Wei@tongji.edu.cn (Z. Wei).

http://dx.doi.org/10.1016/j.jvcir.2014.08.007
1047-3203/

as early termination of CU splitting [13], fast intra mode decision [14], fast motion estimation [15], fast PU size decision [16], early TU size determination [17], cloud-based optimal reference frame selection [18], parallelization of context-adaptive binary arithmetic coding [19], and single instruction multiple data-based encoder implementation [20].

In this work, we investigate another kind of optimization approaches, which is to early detect zero quantized transform coefficients (*i.e.*, the transform coefficients equal to zero after quantization) before implementing transform and quantization. As analyzed in [11], the transform and quantization in HEVC account for about 10–25% of the total encoding computations with different test configurations, therefore it is worth designing effective techniques to trim redundant computations in these two processes. To this aim, a novel prediction algorithm is proposed in this work to early detect all-zero 4 × 4 quantized blocks ahead of transform as well as quantization, and thus all the corresponding transform and quantization computations with these 4 × 4 blocks can be saved. Experimental results demonstrate that the proposed algorithm is able to efficiently predict all-zero 4 × 4 blocks without introducing any loss in compression efficiency.

The rest of this paper is organized as follows. A number of related works about predicting zero quantized transform coefficients are briefly described in Section 2, the HEVC transform and quantization with block size 4 × 4 are also introduced therein. In Section 3, the sufficient conditions to early detect zero quantized transform coefficients in HEVC are theoretically derived and the proposed algorithm for early detection of all-zero 4 × 4 blocks is presented. Experimental results are provided in Section 4. Finally, Section 5 concludes this paper.

2. Related work and preparation

In video coding, it is quite common that a substantial number of transform coefficients of the prediction residual are quantized to zeros. Considerable computations can be saved if these zero quantized transform coefficients can be predicted prior to transform and quantization, and a number of research works [21–30] have been proposed to early detect zero quantized Discrete Cosine Transform (DCT) coefficients in previous video coding standards MPEG-2/4, H.263, and H.264/AVC, including early detection of all-zero blocks [21–23] and zero sub-blocks/individual coefficients [24,25] for the 8 × 8 DCT-based video encoders such as MPEG-2/4 and H.263; early detection of all-zero blocks [26,27] and zero sub-blocks/individual coefficients [28–30] for the 4 × 4 DCT-based video encoder H.264/AVC. In these methods, the energy information of a residual block (such as the sum of absolute difference or SAD in short between the original video signal and the predicted video signal), which is available during intra/inter-prediction and ahead of DCT and quantization, is utilized to derive threshold-based conditions for early detecting zero coefficients or all-zero blocks. The derivation of these thresholds is usually achieved with theoretical analysis on the DCT and quantization functions.

In HEVC, the DCT with block sizes 4 × 4, 8 × 8, 16 × 16, and 32 × 32 can be applied to encode a TU [6]. In addition, the Discrete Sine Transform (DST) with block size 4 × 4 can be alternatively used to transform luma intra-prediction residuals [3]. In this work, we focus on investigating early detection of all-zero 4 × 4 DCT blocks in HEVC, and it will be our future research efforts to explore prediction methods to early detect zero quantized transform coefficients with other transform types. The presented work is extending [27] by performing a detailed analysis on the 4 × 4 DCT and quantization in HEVC, which is different from that of H.264/AVC. As a consequence, sufficient prediction conditions are derived to early detect all-zero 4 × 4 DCT blocks in HEVC. In the following, the 4 × 4 DCT and the related quantization functions employed by HEVC are briefly described, which is useful to elicit the proposed all-zero 4 × 4 DCT block prediction algorithm.

For a residual 4 × 4 block $e(x,y),\ 0 \leqslant x,\ y \leqslant 3$, a two-dimensional 4 × 4 transform is applied by employing one-dimensional transforms in the horizontal and vertical directions. The elements of the transform matrix are derived by approximating scaled DCT basis functions for the sake of limiting the necessary dynamic range for transform computation and maximizing the precision and closeness to orthogonality when the matrix entries are set to integer values [3]. For computational efficiency, only multiplication, addition and shift operations in integer arithmetic are involved in the transform as the following formula.

$$F(u,v) = \left\{ \sum_{x=0}^{3} C(u,x) \cdot \left[\left(\sum_{y=0}^{3} e(x,y) \cdot C^T(y,v) + 1 \right) \gg 1 \right] + 2^7 \right\} \gg 8, \tag{1}$$

where $F(u,v),\ 0 \leqslant u,\ v \leqslant 3$, is the transform coefficient block, C is the two-dimensional transform matrix specified in [3] as below.

$$C = \begin{bmatrix} 64 & 64 & 64 & 64 \\ 83 & 36 & -36 & -83 \\ 64 & -64 & -64 & 64 \\ 36 & -83 & 83 & -36 \end{bmatrix}. \tag{2}$$

And C^T is the transposed matrix of C, $\gg$ indicates the operation of binary shift right.

Given the transform coefficient $F(u,v)$ and a quantization parameter Q_p which ranges from 0 to 51 as defined in HEVC, the quantized transform coefficient $Z(u,v),\ 0 \leqslant u,\ v \leqslant 3$, is generated as

$$Z(u,v) = \mathrm{sign}(F(u,v)) \cdot [(|F(u,v)| \cdot m + r) \gg qbits], \tag{3}$$

where $qbits = 19 + \mathrm{floor}(Q_p/6)$. $r = \upsilon \ll (qbits - 9)$ in which $\ll$ stands for the operation of binary shift left and υ is an offset parameter for quantization which is 171 for intra coding and 85 for inter coding, respectively. m is the multiplication factor depending on Q_p as given in Table 1. The rationale behind Eq. (3) is that a scaling multiplication is integrated into the quantizer to avoid divisions for quantization so as to reduce quantization computations.

3. Proposed all-zero 4 × 4 block early detection algorithm

According to Eq. (3), the sufficient condition under which the transform coefficient $F(u,v)$ is quantized to zero can be expressed as

$$|F(u,v)| < \frac{2^{qbits} - r}{m}. \tag{4}$$

By inserting $F(u,v)$ as given in Eq. (1) into Eq. (4) and after manipulations, we can obtain

$$\left| \sum_{x=0}^{3} C(u,x) \cdot \left[\left(\sum_{y=0}^{3} e(x,y) \cdot C^T(y,v) + 1 \right) \gg 1 \right] \right| < \frac{2^{qbits} - r}{m} \cdot 2^8 - 2^7. \tag{5}$$

Considering the subadditivity property of absolute values, *i.e.*, $|a + b| \leqslant |a| + |b|$, the following inequation can be derived from Eq. (5).

Table 1
Multiplication factor m.

$Q_p\%6$	0	1	2	3	4	5
m	26,214	23,302	20,560	18,396	16,384	14,564

$$\left|\sum_{x=0}^{3} C(u,x)\cdot\left(\sum_{y=0}^{3} e(x,y)\cdot C^T(y,v)\right)\right| < \frac{2^{qbits}-r}{m}\cdot 2^9 - 2^8 - \left|\sum_{x=0}^{3} C(u,x)\right|. \tag{6}$$

From Eq. (6), we can further get

$$\sum_{x=0}^{3}\sum_{y=0}^{3}|C(u,x)|\cdot|e(x,y)|\cdot|C(v,y)| < \frac{2^{qbits}-r}{m}\cdot 2^9 - 2^8 - \left|\sum_{x=0}^{3} C(u,x)\right|. \tag{7}$$

In order to make Eq. (7) hold true at each (u,v) position, the two terms $|C(u,x)|$ and $|C(v,y)|$ on the left side of Eq. (7) are replaced with the element which has the largest absolute value in C and the term $\left|\sum_{x=0}^{3} C(u,x)\right|$ on the right side is set to R_{max} which is defined as

$$R_{max} = \max_{i\in\{0,1,2,3\}}\left|\sum_{x=0}^{3} C(i,x)\right|. \tag{8}$$

As a result, Eq. (7) can be reformulated as

$$\sum_{x=0}^{3}\sum_{y=0}^{3}|e(x,y)| < \frac{\phi}{C_{max}^2}, \tag{9}$$

where $\phi = \frac{2^{qbits}-r}{m}\cdot 2^9 - 2^8 - R_{max}$, C_{max} is the largest element of C in absolute value. According to Eq. (2), $C_{max} = 83$ and $R_{max} = 256$. Considering $\mathrm{SAD} = \sum_{x=0}^{3}\sum_{y=0}^{3}|e(x,y)|$, a sufficient condition to early detect all-zero 4×4 DCT blocks in HEVC is

$$\mathrm{SAD} < TS_1, \quad TS_1 = \frac{\phi}{C_{max}^2}. \tag{10}$$

As the threshold TS_1 can be pre-computed for each Q_p and stored in a lookup table, only one comparison operation is introduced to apply Eq. (10).

For further improving the prediction efficiency, a thorough analysis is performed to investigate the sufficient condition for each individual DCT coefficient $F(u,v)$ to be quantized to zero. To achieve this, the analysis is performed by studying the effects of the two terms $C(u,x)$ and $C(v,y)$ on the left side of Eq. (7) for each (u,v) position. To facilitate the following description, an intermediate 4×4 matrix $C_m^{(u,v)}$ is defined as

$$C_m^{(u,v)} = |C^T(u,:)\times C(v,:)|_{abs}, \tag{11}$$

where $C(i,:)$ indicates the ith row vector of C, and the operand $|\cdot|_{abs}$ is applied to a matrix to replace each element with its corresponding absolute value. According to Eqs. (7) and (11), we can derive

$$\sum_{x=0}^{3}\sum_{y=0}^{3} C_m^{(u,v)}(x,y)\cdot|e(x,y)| < \phi. \tag{12}$$

The sufficient condition for each DCT coefficient $F(u,v)$ to be quantized to zero can be obtained by studying the property of $C_m^{(u,v)}$ as indicated in Eq. (12).

In order to simplify the following discussion, we take the DCT coefficient $F(1,3)$, *i.e.*, $(u,v) = (1,3)$, as an example to illustrate the derivation on the sufficient condition for $F(1,3)$ being quantized to zero, which is expressed as

$$\mathrm{SAD} < \frac{\phi}{2988} + \frac{1692}{2988}\cdot S_3 - \frac{3901}{2988}\cdot S_2, \tag{13}$$

where the partial sum of absolute residuals S_i is calculated as

$$S_i = \sum_x\sum_y |e(x,y)|, \qquad \forall (x,y)\in R_i, \quad 1\leqslant i\leqslant 4, \tag{14}$$

with four regions R_i, $1\leqslant i\leqslant 4$, within a 4×4 block defined as

$$\begin{aligned} R_1 &= \{(x,y)|x=1,2, y=1,2\},\\ R_2 &= \{(x,y)|x=0,3, y=1,2\},\\ R_3 &= \{(x,y)|x=1,2, y=0,3\},\\ R_4 &= \{(x,y)|x=0,3, y=0,3\}. \end{aligned} \tag{15}$$

The derivation of the sufficient condition given in Eq. (13) is presented below. By inserting $(u,v) = (1,3)$ into Eq. (11), we can obtain

$$C_m^{(1,3)} = \begin{bmatrix} 2988 & 6889 & 6889 & 2988\\ 1296 & 2988 & 2988 & 1296\\ 1296 & 2988 & 2988 & 1296\\ 2988 & 6889 & 6889 & 2988 \end{bmatrix}. \tag{16}$$

Then, after inserting Eq. (16) into Eq. (12) and performing manipulations with the consideration of Eq. (14), we can finally get Eq. (13) to determine whether $F(1,3)$ is quantized to zero.

In a similar manner, the sufficient conditions for predicting zero quantized DCT coefficients at other positions are derived and presented in Table 2. According to Table 2, three thresholds Th_i, $1\leqslant i\leqslant 3$, can be derived for early detecting zero quantized DCT coefficients, which are

$$\begin{aligned} &Th_1 = \phi_a, && \forall (u,v)\in\mathcal{A}_1,\\ &Th_2 = \phi_b + \alpha\cdot\min\{S_1+S_2, S_3+S_4, S_1+S_3, S_2+S_4\}, && \forall (u,v)\in\mathcal{A}_2,\\ &Th_3 = \phi_c + \min\{\beta\cdot S_1-\gamma\cdot S_4, \beta\cdot S_3-\gamma\cdot S_2,\\ &\beta\cdot S_2-\gamma\cdot S_3, \beta\cdot S_4-\gamma\cdot S_1\}, && \forall (u,v)\in\mathcal{A}_3, \end{aligned} \tag{17}$$

where three position sets $\mathcal{A}_i$, $1\leqslant i\leqslant 3$, are defined as $\mathcal{A}_1 = \{(0,0),(0,2),(2,0),(2,2)\}$, $\mathcal{A}_2 = \{(0,1),(2,1),(0,3),(2,3),(1,0),(1,2),(3,0),(3,2)\}$ and $\mathcal{A}_3 = \{(1,1),(1,3),(3,1),(3,3)\}$, respectively; the other parameters are set as $\phi_a = \frac{\phi}{4096}$, $\phi_b = \frac{\phi}{5312}$, $\phi_c = \frac{\phi}{2988}$, $\alpha = \frac{3008}{5312}$, $\beta = \frac{1692}{2988}$, $\gamma = \frac{3901}{2988}$. Therefore, if $\mathrm{SAD} < Th_i$, then all the DCT coefficients belonging to $\mathcal{A}_i$ will be quantized to zeros. As a result, the sufficient condition to detect all-zero 4×4 DCT blocks can be summarized as

$$\mathrm{SAD} < TS_2, \quad TS_2 = \min\{Th_1, Th_2, Th_3\}. \tag{18}$$

It can be easily proved that $Th_1 > TS_1$ and $Th_2 > TS_1$ based on Eqs. (10) and (17). Regarding Th_3 in Eq. (17), it can be rewritten as

$$\begin{aligned} Th_3 = TS_1 + \min\{&\eta\cdot(S_2+S_3)+\varphi\cdot S_1, \eta\cdot(S_1+S_4)\\ &+\varphi\cdot S_3, \eta\cdot(S_1+S_4)+\varphi\cdot S_2, \eta\cdot(S_2+S_3)+\varphi\cdot S_4\}, \end{aligned} \tag{19}$$

where $\eta = \frac{3901}{6889}$ and $\varphi = \frac{5593}{6889}$. From Eq. (19), it can be seen that $Th_3 \geqslant TS_1$. Therefore, based on the above analysis we can derive

$$TS_2 \geqslant TS_1, \tag{20}$$

Table 2

Sufficient conditions for early detecting zero quantized DCT coefficients within a 4×4 block in HEVC.

Sufficient condition	Position (u, v)
$\mathrm{SAD} < \phi_a$	(0, 0), (0, 2), (2, 0), (2, 2)
$\mathrm{SAD} < \phi_b + \alpha\cdot(S_1+S_2)$	(0, 1), (2, 1)
$\mathrm{SAD} < \phi_b + \alpha\cdot(S_3+S_4)$	(0, 3), (2, 3)
$\mathrm{SAD} < \phi_b + \alpha\cdot(S_1+S_3)$	(1, 0), (1, 2)
$\mathrm{SAD} < \phi_b + \alpha\cdot(S_2+S_4)$	(3, 0), (3, 2)
$\mathrm{SAD} < \phi_c + \beta\cdot S_1 - \gamma\cdot S_4$	(1, 1)
$\mathrm{SAD} < \phi_c + \beta\cdot S_3 - \gamma\cdot S_2$	(1, 3)
$\mathrm{SAD} < \phi_c + \beta\cdot S_2 - \gamma\cdot S_3$	(3, 1)
$\mathrm{SAD} < \phi_c + \beta\cdot S_4 - \gamma\cdot S_1$	(3, 3)

$\phi_a = \frac{\phi}{4096}$, $\phi_b = \frac{\phi}{5312}$, $\phi_c = \frac{\phi}{2988}$, $\alpha = \frac{3008}{5312}$, $\beta = \frac{1692}{2988}$, $\gamma = \frac{3901}{2988}$.

which indicates that it is more efficient to apply Eq. (18) than Eq. (10) to detect all-zero DCT blocks in HEVC. However, as shown in Eqs. (17) and (18), a number of computational overheads are introduced to employ TS_2, including 10 addition operations, 9 comparison operations and 9 multiplication operations.

In order to balance the detection efficiency against the additional overheads about applying Eq. (18), a prediction algorithm is proposed with the following two steps.

Step 1. For a 4×4 block, if $SAD < TS_1$, this block is an all-zero quantized DCT block and the computations of DCT, Quantization (Q), Inverse Q (IQ) and Inverse DCT (IDCT) can be totally skipped; otherwise, *i.e.*, $SAD \geqslant TS_1$, go to Step 2.

Step 2. Computing TS_2 according to Eq. (18). If $SAD < TS_2$, this block is an all-zero quantized DCT block and the computations of DCT, Q, IQ and IDCT are totally skipped; otherwise, DCT, Q, IQ and IDCT are performed to all the coefficients.

In the following, we also make a discussion about why the proposed prediction algorithm is not applied for the other three types of DCT and quantization, including 8×8, 16×16 and 32×32.

In a similar manner to the 4×4 type of DCT and quantization, the transform coefficient and quantized transform coefficient for the other three types can be calculated and the following formulae can be generalized to represent transform coefficients $F(u,v)$ and quantized transform coefficients $Z(u,v)$ for all of the four types of DCT and quantization.

$$F_i(u,v) = \left\{ \sum_{x=0}^{2^{i+2}-1} C_i(u,x) \cdot \left[\left(\sum_{y=0}^{2^{i+2}-1} e(x,y) \cdot C_i^T(y,v) + 2^i \right) \gg (i+1) \right] + 2^{7+i} \right\} \gg (8+i), \tag{21}$$

$$Z_i(u,v) = \text{sign}(F_i(u,v)) \cdot [(|F_i(u,v)| \cdot m + r_i) \gg qbits_i], \tag{22}$$

where $i = 0,1,2,3$ correspond to the 4×4, 8×8, 16×16, and 32×32, respectively; $0 \leqslant u,\ v \leqslant 2^{i+2} - 1$; $qbits_i = 19 - i + \text{floor}(Q_p/6)$; $r_i = v \ll (qbits_i - 9)$. Considering the transform matrices C_i, they are specified in [3], *e.g.*, C_0 is shown in Eq. (2).

Similar to the sufficient condition derived in Eq. (10) which is used at the first step of the proposed algorithm, we can obtain the following general form to early detect the ith-type all-zero blocks.

$$SAD^i < TS_1^i, \tag{23}$$

where

$$TS_1^i = \frac{\frac{2^{qbits_i - r_i}}{m} \cdot 2^{9+2i} - 2^{8+2i} - 2^i \cdot R_{max}^i}{\left(C_{max}^i\right)^2}, \tag{24}$$

in which C_{max}^i is the largest element of transform matrix C_i in absolute value, and R_{max}^i is defined as

$$R_{max}^i = \max_{0 \leqslant k \leqslant 2^{i+2}-1} \left| \sum_{x=0}^{2^{i+2}-1} C_i(k,x) \right|. \tag{25}$$

Table 3
Threshold ratio $\frac{TS_1^i}{TS_1^0}$.

Inter coding			Intra coding		
$\frac{TS_1^1}{TS_1^0}$	$\frac{TS_1^2}{TS_1^0}$	$\frac{TS_1^3}{TS_1^0}$	$\frac{TS_1^1}{TS_1^0}$	$\frac{TS_1^2}{TS_1^0}$	$\frac{TS_1^3}{TS_1^0}$
1.6–1.7	2.8–3.4	3.8–6.8	1.6–1.7	2.6–3.4	2.9–6.8

Table 4
Video sequences for test.

Class	Seq	Lowercase
A	Traffic	a
	PeopleOnStreet	b
	Nebuta	c
	SteamLocomotive	d
B	Kimono	e
	ParkScene	f
	Cactus	g
	BasketballDrive	h
	BQTerrace	i
C	BasketballDrill	j
	BQMall	k
	PartyScene	l
	RaceHorses	m
D	BasketballPass	n
	BQSquare	o
	BlowingBubbles	p
	RaceHorses	q
E	FourPeople	r
	Johnny	s
	KristenAndSara	t

Table 5
Results of DR, ΔT_d and ΔT_e on video sequences of Class A.

Seq	Q_p	AM-1			AM-2		
		DR (%)	ΔT_d (%)	ΔT_e (%)	DR (%)	ΔT_d (%)	ΔT_e (%)
a	24	31.47	10.49	1.97	44.43	13.53	2.31
	28	37.94	13.96	2.10	46.87	16.28	2.41
	32	42.96	16.81	2.17	48.22	18.39	2.44
	36	45.19	18.52	2.25	48.60	19.64	2.48
	40	46.84	20.75	2.70	48.78	21.05	2.52
	Average	40.88	16.11	2.24	47.38	17.78	2.44
b	24	33.63	9.89	0.65	44.86	11.25	2.07
	28	38.61	11.71	1.88	46.78	13.55	2.22
	32	42.48	13.09	1.90	47.84	14.10	2.39
	36	44.38	14.53	2.34	48.17	15.39	2.40
	40	45.70	15.95	2.43	48.33	16.41	3.21
	Average	40.96	13.03	1.84	47.20	14.14	2.46
c	24	7.87	2.91	0.72	32.99	4.76	1.14
	28	16.48	5.00	1.09	36.23	7.05	1.42
	32	24.17	8.13	1.41	38.66	10.18	1.74
	36	30.90	11.50	1.53	44.96	15.29	1.76
	40	40.69	16.26	1.55	48.11	18.73	2.26
	Average	24.02	8.76	1.26	40.19	11.20	1.66
d	24	23.82	7.93	1.41	40.00	10.89	1.73
	28	33.37	11.82	1.52	45.43	14.77	1.92
	32	41.33	16.41	1.77	48.13	18.19	2.09
	36	45.24	18.79	2.14	48.76	19.82	2.15
	40	47.24	21.09	2.23	48.89	21.38	2.20
	Average	38.20	15.21	1.81	46.24	17.01	2.02
Average		36.01	13.28	1.79	45.25	15.03	2.14

According to Eq. (24), we can study the threshold ratio $\frac{TS_1^i}{TS_1^0}$ given all Q_p values (*i.e.*, $Q_p \in [0, 51]$). Based on the analysis, we find the following relation as shown in Table 3.

However, we believe the ideal threshold ratio should be proportional to the corresponding block size, *i.e.*, $\frac{TS_1^1}{TS_1^0} = 4$, $\frac{TS_1^2}{TS_1^0} = 16$ and $\frac{TS_1^3}{TS_1^0} = 64$. As compared with the ideal threshold ratio, the corresponding actual threshold ratio in Table 3 is much smaller. This indicates that it is not as efficient as for the 4×4 type to apply Eq. (23) to early detect all-zero 8×8, 16×16 and 32×32 blocks.

Furthermore, as aforementioned, regarding the second step of the proposed prediction algorithm, a number of computational overheads are introduced to employ TS_2 in Eq. (18) to early detect all-zero 4×4 blocks. When considering the other three types of DCT and quantization (*i.e.*, 8×8, 16×16 and 32×32), we need to consider more cases than that of 4×4 type (as shown in Table 2). As a result, much more computational overheads will be required to derive Eq. (18)-alike conditions for the 8×8, 16×16 and 32×32 types than that of 4×4 type, which will noticeably increase the encoding time.

Based on the above analysis, in this work we only employ the proposed prediction algorithm to early detect all-zero 4×4 blocks to reduce redundant DCT and quantization computations.

4. Experimental results

The HEVC software test model version HM8.0 [31] is used to evaluate the proposed early detection algorithm. The configuration of Random Access (RA) is employed in the experiment with the size of Group Of Pictures (GOP) equal to 8. The CTU size, minimum CU size, maximum TU size, minimum TU size and both maximum depths of transform trees of intra-coded CU and inter-coded CU are set equal to 64×64, 8×8, 32×32, 4×4, 3 and 3, respectively. A total of twenty HEVC benchmark video sequences are applied for evaluation. For the presentation convenience, the lowercase letters ‘a’–‘t’ are used to represent these twenty sequences as shown in Table 4, where the video resolutions of Class A, B, C, D and E are 2560×1600, 1920×1080, 832×480, 416×240 and 1280×720, respectively. For each of the sequences, a total of 32 frames (*i.e.*, 4 GOPs) are encoded for test. In order to examine the prediction performance at different bit rates, five Q_p values including 24, 28, 32, 36, and 40 are utilized in the experiments. All the experiments are running on a PC with Intel i3-2120 3.3 GHz CPU and 2 GB RAM. The operating system is Windows XP SP3 and the software platform is Visual Studio 2008 and only the C/C++ code is applied for experiments.

In order to evaluate the proposed prediction algorithm, two test conditions are used. Under the first condition, only the first step of the proposed algorithm is enabled to early detect all-zero 4×4 DCT blocks, *i.e.*, only Eq. (10) is enabled for prediction; and we denote this algorithm as AM-1. Under the second condition, both of the two steps of the proposed algorithm are applied as denoted as AM-2. Because sufficient conditions are provided by AM-1 and AM-2 to predict all-zero 4×4 DCT blocks, the performances of video quality and compression efficiency in terms of bit rates are kept the same as that of the original encoder, which are also verified by the experimental results. Therefore, it is not necessary to present the performances of video quality and compression efficiency achieved by AM-1 and AM2. Instead, for real-time performance evaluation, three criteria are applied in this work including Detection Rate (DR), computation reduction in the DCT/Q/IQ/IDCT procedures (ΔT_d) and entire computation reduction (ΔT_e), which are defined as

$$DR = \frac{N_{dec}}{N_{all}} \times 100\%,$$
$$\Delta T_d = \frac{T_d^{org} - T_d}{T_d^{org}} \times 100\%, \qquad (26)$$
$$\Delta T_e = \frac{T_e^{org} - T_e}{T_e^{org}} \times 100\%,$$

where N_{dec} is the number of all-zero 4×4 DCT blocks detected by AM-1 or AM-2, N_{all} is the total number of all-zero 4×4 DCT blocks

Table 6
Results of DR, ΔT_d and ΔT_e on video sequences of Class B.

Seq	Q_p	AM-1			AM-2		
		DR (%)	ΔT_d (%)	ΔT_e (%)	DR (%)	ΔT_d (%)	ΔT_e (%)
e	24	30.29	9.65	1.82	43.05	11.87	2.23
	28	37.86	11.98	1.88	45.74	13.64	2.30
	32	42.49	14.34	2.19	47.34	15.57	2.51
	36	44.78	16.73	2.23	48.10	17.45	2.60
	40	46.47	18.65	2.49	48.57	19.04	2.62
	Average	40.38	14.27	2.12	46.56	15.51	2.45
f	24	26.41	8.53	1.16	43.26	12.30	1.82
	28	35.09	12.58	1.72	46.16	15.52	1.89
	32	41.10	15.70	1.74	47.73	17.43	2.23
	36	44.14	18.15	1.83	48.32	19.57	2.30
	40	46.38	20.25	1.89	48.66	20.98	2.31
	Average	38.62	15.04	1.67	46.83	17.16	2.11
g	24	20.50	6.45	1.16	38.71	10.12	1.73
	28	31.60	11.17	1.70	44.76	14.14	2.04
	32	40.00	14.68	1.92	47.10	16.58	2.17
	36	43.33	17.10	1.95	47.87	17.96	2.22
	40	45.46	18.66	2.28	48.32	19.45	2.30
	Average	36.18	13.61	1.80	45.35	15.65	2.09
h	24	30.45	8.59	1.59	42.81	11.42	1.93
	28	38.76	13.56	1.97	45.69	14.02	2.09
	32	43.00	14.97	2.00	47.16	15.93	2.19
	36	44.82	16.62	2.09	47.72	16.98	2.24
	40	45.95	18.18	2.20	48.04	18.41	2.28
	Average	40.60	14.38	1.97	46.28	15.35	2.15
i	24	27.77	8.49	1.37	41.02	10.58	1.65
	28	33.93	12.12	1.47	44.50	14.84	1.76
	32	40.14	16.42	1.54	47.33	18.54	1.84
	36	43.77	19.22	1.78	48.42	20.70	1.94
	40	46.29	21.25	1.80	48.80	21.94	2.01
	Average	38.38	15.50	1.59	46.01	17.32	1.84
Average		38.83	14.56	1.83	46.21	16.20	2.13

Table 7
Results of DR, ΔT_d and ΔT_e on video sequences of Class C.

Seq	Q_p	AM-1			AM-2		
		DR (%)	ΔT_d (%)	ΔT_e (%)	DR (%)	ΔT_d (%)	ΔT_e (%)
j	24	29.36	8.84	1.84	43.10	11.17	2.07
	28	34.43	11.00	1.98	45.01	13.17	2.17
	32	38.99	13.23	2.07	46.61	14.58	2.29
	36	41.50	14.75	2.17	47.25	16.29	2.42
	40	43.57	17.09	2.56	47.73	17.49	2.59
	Average	37.57	12.98	2.12	45.94	14.54	2.31
k	24	29.44	8.90	1.48	43.62	11.30	1.62
	28	35.52	11.64	1.90	45.97	13.81	2.06
	32	40.23	14.03	1.91	47.31	15.61	2.09
	36	42.57	15.70	1.94	47.87	17.02	2.23
	40	44.55	17.52	2.01	48.21	18.57	2.25
	Average	38.46	13.56	1.85	46.60	15.26	2.05
l	24	19.06	5.24	1.06	37.14	7.44	1.45
	28	23.61	6.97	1.25	40.69	9.63	1.62
	32	29.93	9.65	1.38	44.09	12.18	1.65
	36	34.02	11.71	1.57	45.67	15.04	2.03
	40	38.10	14.36	1.58	46.91	16.97	2.25
	Average	28.94	9.59	1.37	42.90	12.25	1.80
m	24	24.61	6.09	1.25	39.39	7.51	1.57
	28	29.62	7.89	1.44	42.33	9.93	2.08
	32	34.58	10.20	1.87	44.71	11.84	2.17
	36	37.51	11.94	1.94	45.84	13.59	2.20
	40	40.46	13.73	1.96	46.76	15.16	2.30
	Average	33.36	9.97	1.69	43.81	11.61	2.06
Average		34.58	11.53	1.76	44.81	13.41	2.05

Table 8
Results of *DR*, ΔT_d and ΔT_e on video sequences of Class D.

Seq	Q_p	AM-1			AM-2		
		DR (%)	ΔT_d (%)	ΔT_e (%)	*DR* (%)	ΔT_d (%)	ΔT_e (%)
n	24	35.00	10.98	1.33	45.72	13.62	1.66
	28	38.22	13.74	1.80	46.63	15.80	1.93
	32	41.89	15.92	1.81	47.47	17.54	2.21
	36	43.65	18.05	2.02	47.87	19.36	2.28
	40	45.17	19.48	2.22	48.21	20.04	2.51
	Average	40.79	15.63	1.84	47.18	17.27	2.12
o	24	27.69	8.66	1.16	43.01	11.87	1.79
	28	32.91	11.83	1.38	44.99	14.66	1.95
	32	38.36	15.52	1.72	46.49	17.60	2.05
	36	41.26	17.08	2.18	47.58	19.38	2.27
	40	43.61	20.06	2.23	48.32	21.55	2.31
	Average	36.77	14.63	1.73	46.08	17.01	2.07
p	24	20.64	6.55	0.71	38.46	9.47	1.76
	28	26.02	7.86	0.76	42.01	11.43	1.80
	32	33.16	11.52	1.35	45.08	14.48	1.95
	36	37.31	13.05	1.43	46.64	16.41	2.16
	40	41.40	16.43	1.51	47.67	18.64	2.18
	Average	31.71	11.08	1.15	43.97	14.09	1.97
q	24	23.86	6.09	1.31	39.63	7.46	1.46
	28	27.69	7.91	1.68	42.10	9.92	1.96
	32	32.28	9.59	1.71	44.10	11.74	2.19
	36	35.34	11.60	1.76	45.44	13.52	2.21
	40	38.99	13.38	1.90	46.53	15.09	2.23
	Average	31.63	9.71	1.67	43.56	11.55	2.01
Average		35.22	12.76	1.60	45.20	14.98	2.04

generated during video encoding; T_d^{org} is the encoding time consumed by DCT/Q/IQ/IDCT of the original encoder including that of TU size of 4×4, 8×8, 16×16 and 32×32, T_d is the encoding time consumed by DCT/Q/IQ/IDCT of the modified encoder with AM-1 or AM-2 including that of TU size of 4×4, 8×8, 16×16 and 32×32; T_e^{org} is the entire encoding time of the original encoder, and T_e is the entire encoding time of the modified encoder with AM-1 or AM-2.

Table 9
Results of *DR*, ΔT_d and ΔT_e on video sequences of Class E.

Seq	Q_p	AM-1			AM-2		
		DR (%)	ΔT_d (%)	ΔT_e (%)	*DR* (%)	ΔT_d (%)	ΔT_e (%)
r	24	41.21	16.24	2.02	47.69	18.24	2.16
	28	44.66	18.76	2.20	48.50	19.96	2.34
	32	46.70	20.77	2.21	48.78	21.47	2.41
	36	47.53	21.38	2.58	48.86	21.64	2.49
	40	48.09	22.56	2.60	48.88	22.42	2.82
	Average	45.64	19.94	2.32	48.54	20.75	2.44
s	24	42.34	17.16	1.87	47.58	18.37	2.11
	28	45.55	19.62	2.06	48.57	20.68	2.16
	32	47.51	21.55	2.20	48.87	21.87	2.34
	36	48.18	22.53	2.31	48.92	22.70	2.42
	40	48.52	22.59	2.52	48.91	22.60	2.64
	Average	46.42	20.69	2.19	48.57	21.24	2.33
t	24	42.66	17.10	2.27	47.61	18.60	2.41
	28	45.23	19.11	2.41	48.42	20.20	2.43
	32	46.99	20.80	2.42	48.77	21.18	2.58
	36	47.79	21.74	2.44	48.88	21.89	2.73
	40	48.32	22.68	2.62	48.91	22.63	2.78
	Average	46.20	20.29	2.43	48.52	20.90	2.59
Average		46.09	20.31	2.31	48.54	20.96	2.45

From the definition, it can be seen that the larger the *DR* is, the more efficient the prediction algorithm can early detect all-zero 4×4 DCT blocks. The experimental results are shown in Tables 5–9 for the five classes of test video sequences, respectively.

From the results, several observations can be made. Firstly, both of the two algorithms AM-1 and AM-2 are able to early detect all-zero 4×4 DCT blocks and thus reduce redundant DCT and Q computations. Secondly, along with the increase of Q_p values, the performances of both of the two algorithms AM-1 and AM-2 have been improved. This is because the detection thresholds TS_1 and TS_2 as given in Eqs. (10) and (18) become larger with the increment of Q_p and hence more all-zero 4×4 DCT blocks can be detected leading to more encoding computations being reduced. Thirdly, by comparing the performances between AM-1 and AM-2, it can be easily seen that AM-2 is more effective than AM-1 to reduce redundant encoding computations, which indicates that it is worth utilizing the second step of the proposed AM-2 algorithm for early detection of all-zero 4×4 DCT blocks although a number of computational overheads are introduced to apply Eq. (18).

5. Conclusion and future works

In this work, a prediction algorithm is proposed to early detect all-zero 4×4 DCT blocks in HEVC. Theoretical analyses are performed to investigate the dynamic range of quantized DCT coefficients when the TU size is 4×4. As a consequence, two sufficient prediction conditions are mathematically deduced and the proposed prediction algorithm is developed while considering both the prediction efficiency and the computational overheads resulted from employing the two prediction conditions. Extensive experiments are implemented on HEVC benchmark video sequences and it is verified that the proposed algorithm is able to effectively detect all-zero 4×4 blocks and thus redundant DCT and Q computations are saved while introducing no loss in video quality and compression efficiency.

From the results, it can be observed that there is still much room to improve the all-zero block detection rate, especially when Q_p is small. To achieve this, it is worth investigating other techniques to further enhance the prediction efficiency, such as employing statistical models [24,28] to derive more efficient prediction conditions and to early detect zero quantized coefficients at individual coefficient level. Moreover, we will study efficient detection algorithms for other TU block sizes, including 8×8, 16×16, and 32×32 to further reduce the HEVC encoding computational complexity.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grants 61102059, 61472281 and 61471235, the "Shu Guang" project of Shanghai Municipal Education Commission and Shanghai Education Development Foundation under Grant 12SG23, the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning, the Fundamental Research Funds for the Central Universities under Grants 0800219158 and 0800219270.

References

[1] Advanced Video Coding for Generic Audiovisual Services, ISO/IEC 14496-10:2005(E) ITU-T Rec. H.264 (E), 2005.
[2] B. Bross, W.-J. Han, J.-R. Ohm, G.J. Sullivan, Y.-K. Wang, T. Wiegand, High Efficiency Video Coding (HEVC) Text Specification Draft 10 (for FDIS & Last Call), Doc. JCTVC-L1003_v34, 2013.
[3] G.J. Sullivan, J.-R. Ohm, W.-J. Han, T. Wiegand, Overview of the high efficiency video coding (HEVC) standard, IEEE Trans. Circ. Syst. Video Technol. 22 (12) (2012) 1649–1668.

[4] J.-R. Ohm, G.J. Sullivan, H. Schwarz, T.K. Tan, T. Wiegand, Comparison of the coding efficiency of video coding standards – including high efficiency video coding (HEVC), IEEE Trans. Circ. Syst. Video Technol. 22 (12) (2012) 1669–1684.

[5] C.C. Chi, M. Alvarez-Mesa, B. Juurlink, G. Clare, F. Henry, S. Pateux, T. Schierl, Parallel scalability and efficiency of HEVC parallelization approaches, IEEE Trans. Circ. Syst. Video Technol. 22 (12) (2012) 1827–1838.

[6] I.-K. Kim, J. Min, T. Lee, W.-J. Han, J. Park, Block partitioning structure in the HEVC standard, IEEE Trans. Circ. Syst. Video Technol. 22 (12) (2012) 1697–1706.

[7] T. Wiegand, M. Lightstone, D. Mukherjee, T.G. Campbell, S.K. Mitra, Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard, IEEE Trans. Circ. Syst. Video Technol. 6 (4) (1996) 182–190.

[8] J. Lainema, F. Bosen, W.-J. Han, J. Min, K. Ugur, Intra coding of the HEVC standard, IEEE Trans. Circ. Syst. Video Technol. 22 (12) (2012) 1792–1801.

[9] C.-M. Fu, E. Alshina, A. Alshin, Y.-W. Huang, C.-Y. Chen, C.-Y. Tsai, C.-W. Hsu, S.-M. Lei, J.-H. Park, W.-J. Han, Sample adaptive offset in the HEVC standard, IEEE Trans. Circ. Syst. Video Technol. 22 (12) (2012) 1755–1764.

[10] E. Wige, G. Yammine, P. Amon, A. Hutter, A. Kaup, In-loop noise-filtered prediction for high efficiency video coding, IEEE Trans. Circ. Syst. Video Technol. 24 (7) (2014) 1142–1155.

[11] F. Bossen, B. Bross, K. Sühring, D. Flynn, HEVC complexity and implementation analysis, IEEE Trans. Circ. Syst. Video Technol. 22 (12) (2012) 1685–1696.

[12] G. Corrêa, P. Assuncão, L. Agostini, L.A. da Silva Cruz, Performance and computational complexity assessment of high-efficiency video encoders, IEEE Trans. Circ. Syst. Video Technol. 22 (12) (2012) 1899–1909.

[13] L. Shen, Z. Liu, X. Zhang, W. Zhao, Z. Zhang, An effective CU size decision method for HEVC encoders, IEEE Trans. Multimedia 15 (2) (2013) 465–470.

[14] S. Cho, M. Kim, Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding, IEEE Trans. Circ. Syst. Video Technol. 23 (9) (2013) 1555–1564.

[15] M.E. Sinangil, V. Sze, M. Zhou, A.P. Chandrakasan, Cost and coding efficient motion estimation design considerations for high efficiency video coding (HEVC) standard, IEEE J. Sel. Top. Signal Process. 7 (6) (2013) 1017–1028.

[16] F. Sampaio, S. Bampi, M. Grellert, L. Agostini, J. Mattos, Motion vectors merging: low complexity prediction unit decision heuristic for the inter-prediction of HEVC encoders, in: Proc. ICME'12, 2012, pp. 657–662.

[17] K. Choi, E.S. Jang, Early TU design method for fast video encoding in high efficiency video coding, Electron. Lett. 48 (12) (2012) 689–691.

[18] B. Li, J. Xu, H. Li, F. Wu, Optimized reference frame selection for video coding by cloud, in: Proc. MMSP'11, 2011, pp. 1–5.

[19] V. Sze, M. Budagavi, Parallelization of CABAC transform coefficient coding for HEVC, in: Proc. PCS'12, 2012, pp. 509–512.

[20] K. Chen, Y. Duan, L. Yan, J. Sun, Z. Guo, Efficient SIMD optimization of HEVC encoder over X86 processors, in: Proc. APSIPA ASC'12, 2012, pp. 1–4.

[21] A. Yu, R. Lee, M. Flynn, Performance enhancement of H.263 encoder based on zero coefficient prediction, in: ACMMM'97, 1997, pp. 21–29.

[22] X. Zhou, Z. Yu, S. Yu, Method for detecting all-zero DCT coefficients ahead of discrete cosine transformation and quantization, Electron. Lett. 34 (19) (1998) 1839–1840.

[23] L.A. Sousa, General method for eliminating redundant computations in video coding, Electron. Lett. 36 (4) (2000) 306–307.

[24] I.M. Pao, M.T. Sun, Modeling DCT coefficients for fast video encoding, IEEE Trans. Circ. Syst. Video Technol. 9 (4) (1999) 608–616.

[25] H. Wang, S. Kwong, C.-W. Kok, Efficient predictive model of zero quantized DCT coefficients for fast video encoding, Image Vis. Comput. 25 (6) (2007) 922–933.

[26] Y.H. Moon, G.Y. Kim, J.H. Kim, An improved early detection algorithm for all-zero blocks in H.264 video encoding, IEEE Trans. Circ. Syst. Video Technol. 15 (8) (2005) 1053–1057.

[27] H. Wang, S. Kwong, C.-W. Kok, Efficient prediction algorithm of integer DCT coefficients for H.264/AVC optimization, IEEE Trans. Circ. Syst. Video Technol. 16 (4) (2006) 547–552.

[28] H. Wang, S. Kwong, Hybrid model to detect zero quantized DCT coefficients in H.264, IEEE Trans. Multimedia 9 (4) (2007) 728–735.

[29] M. Zhang, T. Zhou, W. Wang, Adaptive method for early detecting zero quantized DCT coefficients in H.264/AVC video coding, IEEE Trans. Circ. Syst. Video Technol. 19 (1) (2009) 103–107.

[30] J. Li, M. Gabbouj, J. Takala, Zero-quantized inter DCT coefficient prediction for real-time video coding, IEEE Trans. Circ. Syst. Video Technol. 22 (2) (2012) 249–259.

[31] JCT-VC, Subversion Repository for the HEVC Test Model Version HM8.0, 2012. <https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-8.0/>.