# Early Determination of Zero-Quantized 8 × 8 DCT Coefficients

Xiangyang Ji, Sam Kwong, Debin Zhao, Hanli Wang, *Member, IEEE,* C.-C. Jay Kuo, *Fellow, IEEE,* and Qionghai Dai, *Senior Member, IEEE*

***Abstract*—This paper proposes a novel approach to early determination of zero-quantized 8 × 8 discrete cosine transform (DCT) coefficients for fast video encoding. First, with the dynamic range analysis of DCT coefficients at different frequency positions, several sufficient conditions are derived to early determine whether a prediction error block (8 × 8) is an all-zero or a partial-zero block, i.e., the DCT coefficients within the block are all or partially zero-quantized. Being different from traditional methods that utilize the sum of absolute difference (SAD) of the entire prediction error block, the sufficient conditions are derived based on the SAD of each row of the prediction error block. For partial-zero blocks, fast DCT/IDCT algorithms are further developed by pruning conventional 8-point butterfly-based DCT/IDCT algorithms. Experimental results exhibit that the proposed early determination algorithm greatly reduces computational complexity in terms of DCT/IDCT, quantization, and inverse quantization, as compared with existing algorithms.**

***Index Terms*—Butterfly-based DCT, computational complexity, discrete cosine transform, sum of absolute difference, zero-quantized coefficients.**

## I. Introduction

Most video compression standards, such as MPEG-x and H.26x, are designed based on a hybrid video coding framework. In such a framework, video is first decorrelated by block-wise spatial prediction or temporal prediction with motion estimation (ME) and motion compensation (MC) and then, each prediction error block goes through DCT to exploit spatial redundancy. Finally, transform coefficients are quantized and fed into entropy encoder to yield desired video bit streams. At the decoder, the inverse sequential operations are performed to decode video for playback.

In video encoders such as H.263 [1] and MPEG-4 Part 2 [2], ME, $8 \times 8$ DCT, quantization (Q), inverse quantization (IQ), and inverse DCT (IDCT) are usually computationally intensive [3], [4]. Many fast ME algorithms [5]–[13] have been proposed to reduce the encoding complexity. Consequently, saving the computational complexity for DCT, Q, IQ, and IDCT becomes important. In [14], according to the execution time distribution for the *Foreman* sequence, DCT/IDCT and Q/IQ accounts for approximately 40% of the total encoding time when the XVID encoder is used for evaluation with the PMVFAST ME algorithm [13] employed. It is also claimed in [14] that the execution time distribution is similar for other test sequences. On the other hand, as stated in [15], average time cost in terms of DCT/IDCT and Q/IQ accounts for 34.8% of the total encoding time when a software H.263 encoder (based on H.263 TMN-5) is used for testing.

Typically, asignificant number of prediction error blocks will have all-zero coefficients after Q in low bit-rate video coding. Even in higher bit-rate video coding, some well-predicted blocks may have all-zero DCT coefficients after Q. If these blocks can be determined earlier, computations of DCT, Q, IQ, and IDCT can be completely skipped. Yu *et al*. [16] proposed a method of early predicting all-zero blocks by comparing SAD of the predicted error block with a predefined threshold to skip DCT, Q, IQ, and IDCT, where no additional SAD computations are required since the SAD values can be obtained after ME. However, the corresponding coding efficiency is highly dependent on the threshold selection mechanism and video quality can be degraded due to an improper threshold. Zhou *et al*. [17] performed a theoretical analysis on the dynamic range of DCT coefficients and proposed a sufficient condition to check whether a given prediction error block has all-zero coefficients based on the SAD of the prediction error. With this approach, DCT, Q, IQ, and IDCT can be completely skipped for all-zero blocks without video quality degradation. Hsueh *et al*. [18] derived a similar sufficient condition and applied a looser condition in the early determination of all-

Manuscript received March 17, 2008; revised February 2, 2009. First version published July 7, 2009; current version published December 1, 2009. This work was supported by Hong Kong RGB Competitive Earmarked Research Grant Projects 9041236 (CityU 114707), 9041353 (CityU 115408), and a Grant from the Major State Basic Research Development Program 973 of China, No. 2009CB320905. This paper was recommended by Associate Editor J. Boyce.

X. Ji is now with the Broadband Networks and Digital Media Laboratory, Automation Department, Tsinghua University, Beijing 100084, China. He was with both the Graduate School and the Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China (e-mail: xyji@tsinghua.edu.cn).

S. Kwong is with the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, China (e-mail: cssamk@cityu.edu.hk).

D. Zhao is with the Department of Computer Science, Harbin Institute of Technology, Harbin 150001, China (e-mail: dbzhao@vilab.hit.edu.cn).

H. Wang is with the Fern Universität in Hagen, Lehrgebiet Informationstechnik, Hagen 58084, Germany (hanli.wang@gmail.com).

C.-C. Jay Kuo is with the Department of Electrical Engineering and Integrated Media Systems Center, University of Southern California, Los Angeles, CA 90089 USA (e-mail: cckuo@sipi.usc.edu).

Q. Dai is with the Broadband Networks and Digital Media Laboratory, Automation Department, Tsinghua University, Beijing 100084, China (e-mail: qhdai@tsinghua.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2009.2026839

zero blocks. Sousa [19] developed a tighter sufficient condition than Zhou's model [17] by comparing sufficient conditions for different frequency components quantized to zeros. As a result, Sousa's model could detect all-zero blocks more efficiently and thus reduces the computational complexity furthermore. Jun *et al*. [20] proposed a new criterion for the early determination of all-zero blocks, where the sign of each predicted residual is considered. This method demands a higher computational overhead since it needs additional SAD calculation for the prediction error block (i.e., no SAD reuse in the ME stage) as well as a sign evaluation for each residual sample. All aforementioned models were designed for 8 × 8 DCT used in H.263 and MPEG-4 Part 2. More recently, Wang *et al*. [21] applied Sousa's model [19] to the DCT-like 4 × 4 integer transform in H.264/AVC. Kim *et al*. [22] proposed a novel all-zero blocks detecting algorithm for the DCT-like 4 × 4 integer transform in H.264/AVC. Recently, Wang *et al*. [23] derived more effective sufficient conditions to early determination of all-zero 4 × 4 blocks in H.264/AVC.

All methods reviewed above are only used to early determine all-zero blocks. However, although some of the prediction error blocks cannot be determined as all-zero ones under a sufficient condition, partial coefficients in those blocks can be quantized to zeros and determined by new sufficient conditions. For example, Wang *et al*. [4] proposed a new analytical model to eliminate redundant DCT, Q, IQ, and IDCT. By a finer analysis on the dynamic range of DCT coefficients, the model in [4] is capable of detecting zero-quantized coefficients at both block and individual frequency levels.

In this paper, we first perform a theoretical analysis on the dynamic range of 8 × 8 DCT coefficients at different frequency positions, and then derive a more precise sufficient condition than Sousa's model [19] for early determination of all-zero blocks. As compared with existing analytical models, our proposed sufficient condition is not based on SAD of the entire prediction error block but that of each row of the prediction error block. Second, several sufficient conditions for early determination of partial-zero blocks are derived. For partial-zero blocks, fast DCT/IDCT pruning algorithms are further developed based on fast 8-point butterfly-based DCT/IDCT algorithms to save the computational complexity efficiently.

The rest of this paper is organized as follows. Section II describes the proposed early determination of zero-quantized DCT coefficients in detail. In this section, Sousa's model [19] for 8 × 8 all-zero block determination is first introduced and a more precise sufficient condition is derived to determine all-zero blocks and new sufficient conditions are derived to determine partial-zero blocks. Then, the fast DCT/IDCT algorithms for detected partial-zero blocks are presented. Experimental results are given in Section III, where the computational complexity of the proposed algorithm is compared with existing methods. Finally, Section IV concludes this paper and presents future research directions.

## II. Proposed Early Determination Strategy

### A. Sousa's Model for All-Zero Block Determination

When applying 2-D DCT to a given block $f$ of $N \times N$ samples, each frequency component $F(u, v)$ can be calculated by

$$F(u, v) = C(u)C(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x, y)\cos\left[\frac{(2x+1)u\pi}{2N}\right] \times \cos\left[\frac{(2y+1)v\pi}{2N}\right] \tag{1}$$

with

$$C(n) = \begin{cases} \sqrt{\frac{1}{N}}, & \text{for } n = 0 \\ \sqrt{\frac{2}{N}}, & \text{for } n > 0. \end{cases}$$

For an 8 × 8 inter-block with motion-compensated prediction, its prediction error signals can be computed by

$$f(x, y) = s(x, y) - p(x, y), \quad \text{for } x, y = 0, 1, ..., 7. \tag{2}$$

Here, $s$ and $p$ represent the original block and the motion-compensated prediction block, respectively.

In encoding, DCT coefficients of each prediction error block are quantized to represent them in a reduced range of values. For a uniform scalar quantization in H.263 [1] and MPEG-4 Part 2 [2], the quantized coefficient $F_Q(u, v)$ can be obtained by

$$F_Q(u, v) = \text{sign}(F(u, v))\left\lfloor\frac{|F(u, v)| - (Q_p/2)}{2Q_p}\right\rfloor \tag{3}$$

where $Q_p$ is the quantization scalar and the symbol $\lfloor\rfloor$ denotes rounding to the nearest integer. It can be observed for an 8 × 8 block that when

$$\begin{aligned} |F(u, v)| &\le C(u)C(v)\textstyle\sum_{x=0}^{7}\sum_{y=0}^{7}|f(x, y)| \\ &\quad \times \max\left|\cos\left(\tfrac{(2x+1)u\pi}{16}\right)\cos\left(\tfrac{(2y+1)v\pi}{16}\right)\right| \\ &= C(u)C(v) \times SAD \\ &\quad \times \max\left|\cos\left(\tfrac{(2x+1)u\pi}{16}\right)\cos\left(\tfrac{(2y+1)v\pi}{16}\right)\right| \\ &< \tfrac{5}{2}Q_p. \end{aligned} \tag{4}$$

$F(u, v)$ will be quantized to zero. According to Sousa's model [19], all DCT coefficients of an 8 × 8 block will be quantized to zeros if

$$SAD < 10Q_p/\cos^2\left(\frac{\pi}{16}\right). \tag{5}$$

### B. New Sufficient Condition for All-Zero Block Determination

Sousa's model provides a sufficient condition to early determining zero-quantized coefficient before transform and quantization. The inequality (4) is achieved by analyzing the dynamic range of each frequency component based on the SAD of an entire block. In fact, within a predicted error block, the energy of predicted error signals in different parts is usually inhomogeneous and thus, it is possible to derive a more precise sufficient condition to early determining all-zero block-based on the SADs of partial blocks.

The condition in (4) for DCT coefficient $F(u, v)$ quantized to zero can be further rewritten as follows:

$$|F(u,v)| \le S^{\mathrm{I}}(u,v) \le C(u)C(v)\times SAD \times \max\left|\cos\left(\tfrac{(2x+1)u\pi}{16}\right)\cos\left(\tfrac{(2y+1)v\pi}{16}\right)\right| \tag{6}$$

with

$$S^{\mathrm{I}}(u,v) = C(u)C(v)\sum_{x=0}^{7}\left|\cos\left(\frac{(2x+1)u\pi}{16}\right)\right| SAD_x \times \max\left|\cos\left(\frac{(2y+1)v\pi}{16}\right)\right|$$

and

$$SAD_i = \sum_{y=0}^{7}|f(i,y)|\,.$$

For each row of $S^{\mathrm{I}}(u,v)$ with row index u′, namely $S^{\mathrm{I}}(\mathrm{u}',v)$, its maximum item is

$$\begin{aligned}\max\left(S^{\mathrm{I}}(\mathrm{u}',v)\right) &= \max\left(C(\mathrm{u}')C(v)\times\max\left(\left|\cos\left(\tfrac{(2y+1)v\pi}{16}\right)\right|\right)\times\textstyle\sum_{x=0}^{7}\left|\cos\left(\tfrac{(2x+1)\mathrm{u}'\pi}{16}\right)\right|\times SAD_x\right)\\ &= S^{\mathrm{I}}(\mathrm{u}',v)\big|_{v=1,3,5,7}\\ &= \tfrac{1}{2}C(\mathrm{u}')\cos\left(\tfrac{\pi}{16}\right)\textstyle\sum_{x=0}^{7}\left|\cos\left(\tfrac{(2x+1)\mathrm{u}'\pi}{16}\right)\right|\\ &\quad\times SAD_x.\end{aligned} \tag{7}$$

According to (7), in each row with $u=\mathrm{u}'$, $S^{\mathrm{I}}(\mathrm{u}',v)\big|_{v=1,3,5,7}$ are the same and they should be no less than $S^{\mathrm{I}}(\mathrm{u}',v)\big|_{v\neq 1,3,5,7}$. Consequently, the maximum item in all $S^{\mathrm{I}}(u,v)$ can be found from $S^{\mathrm{I}}(u,1)$. Let us first take care of items $S^{\mathrm{I}}(u,1)\big|_{u=1,3,5,7}$. Apparently, the maximum item in $S^{\mathrm{I}}(u,1)\big|_{u=1,3,5,7}$ can be obtained by several comparisons among them. For example, to compare $S^{\mathrm{I}}(1,1)$ with $S^{\mathrm{I}}(3,1)$, we have to calculate

$$\begin{aligned}&S^{\mathrm{I}}(1,1)-S^{\mathrm{I}}(3,1)\\ &=\tfrac{1}{4}\cos\left(\tfrac{\pi}{16}\right)\Big\{\cos\left(\tfrac{\pi}{16}\right)\left(\textstyle\sum_{x=0,7}SAD_x-\sum_{x=2,5}SAD_x\right)\\ &\quad+\cos\left(\tfrac{3\pi}{16}\right)\left(\textstyle\sum_{x=1,6}SAD_x-\sum_{x=0,7}SAD_x\right)+\cos\left(\tfrac{5\pi}{16}\right)\\ &\quad\left(\textstyle\sum_{x=2,5}SAD_x-\sum_{x=3,4}SAD_x\right)\\ &\quad+\cos\left(\tfrac{7\pi}{16}\right)\left(\textstyle\sum_{x=3,4}SAD_x-\sum_{x=1,6}SAD_x\right)\Big\}.\end{aligned} \tag{8}$$

However, this comparison is still complicated in computation. For further simplification, a new definition

$$S^{\mathrm{II}}(u,v)=\frac{1}{4}\cos\left(\frac{\pi}{16}\right)\times\left(\cos\left(\frac{\pi}{16}\right)\sum_{x\in X_0}SAD_x+\cos\left(\frac{5\pi}{16}\right)\sum_{x\in X_1}SAD_x\right) \tag{9}$$

is introduced, in which for different sets $X_0$ and $X_1$, we have definitions in (11) and (12).

It is obvious that $S^{\mathrm{I}}(u,1)$ is not larger than $S^{\mathrm{II}}(u,1)$ in (11) and (12). According to the appendix, a sufficient condition for all-zero block detection can be derived as follows:

$$\begin{aligned}&\max(S^{\mathrm{II}}(u,1)|_{u=1,2,3,5,6,7})\\ &=\tfrac{1}{4}\cos\left(\tfrac{\pi}{16}\right)\left(\cos\left(\tfrac{\pi}{16}\right)\textstyle\sum_{x\in X_0}SAD_x+\cos\left(\tfrac{5\pi}{16}\right)\sum_{x\in X_1}SAD_x\right)\\ &\le\tfrac{1}{4}\cos\left(\tfrac{\pi}{16}\right)\left(\cos\left(\tfrac{\pi}{16}\right)\textstyle\sum_{x\in X_0}SAD_x+\tfrac{4}{7}\cos\left(\tfrac{\pi}{16}\right)\sum_{x\in X_1}SAD_x\right)\\ &=\tfrac{1}{7}\cos^2\left(\tfrac{\pi}{16}\right)\left(\tfrac{7}{4}\textstyle\sum_{x\in X_0}SAD_x+\sum_{x\in X_1}SAD_x\right)\\ &<\tfrac{5}{2}Q_p\\ &\Rightarrow SAD+\tfrac{3}{4}\textstyle\sum_{x\in X_0}SAD_x\\ &=SAD+\left(\textstyle\sum_{x\in X_0}SAD_x-\left(\left(\sum_{x\in X_0}SAD_x\right)>>2\right)\right)\\ &<\tfrac{35}{2}Q_p/\cos^2\left(\tfrac{\pi}{16}\right).\end{aligned} \tag{10}$$

Here, sets $X_0$ and $X_1$ are related to (11) and (12), shown at the bottom of the page and dependent on which one of $S^{\mathrm{II}}(u,1)|_{u=1,2,3,5,6,7}$ is the maximum. To reduce computations

$$\begin{cases}S^{\mathrm{II}}(1,1)=\frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\cos\left(\frac{\pi}{16}\right)\sum\limits_{x\in X_0=(0,1,6,7)}SAD_x+\cos\left(\frac{5\pi}{16}\right)\sum\limits_{x\in X_1=(2,3,4,5)}SAD_x\right)\\ S^{\mathrm{II}}(3,1)=\frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\cos\left(\frac{\pi}{16}\right)\sum\limits_{x\in X_0=(0,2,5,7)}SAD_x+\cos\left(\frac{5\pi}{16}\right)\sum\limits_{x\in X_1=(1,3,4,6)}SAD_x\right)\\ S^{\mathrm{II}}(5,1)=\frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\cos\left(\frac{\pi}{16}\right)\sum\limits_{x\in X_0=(1,3,4,6)}SAD_x+\cos\left(\frac{5\pi}{16}\right)\sum\limits_{x\in X_1=(0,2,5,7)}SAD_x\right)\\ S^{\mathrm{II}}(7,1)=\frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\cos\left(\frac{\pi}{16}\right)\sum\limits_{x\in X_0=(2,3,4,5)}SAD_x+\cos\left(\frac{5\pi}{16}\right)\sum\limits_{x\in X_1=(0,1,6,7)}SAD_x\right)\end{cases} \tag{11}$$

and

$$\begin{cases}S^{\mathrm{II}}(2,1)=\frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\cos\left(\frac{\pi}{16}\right)\sum\limits_{x\in X_0=(0,3,4,7)}SAD_x+\cos\left(\frac{5\pi}{16}\right)\sum\limits_{x\in X_1=(1,2,5,6)}SAD_x\right)\\ S^{\mathrm{II}}(6,1)=\frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\cos\left(\frac{\pi}{16}\right)\sum\limits_{x\in X_0=(1,2,5,6)}SAD_x+\cos\left(\frac{5\pi}{16}\right)\sum\limits_{x\in X_1=(0,3,4,7)}SAD_x\right)\end{cases} \tag{12}$$

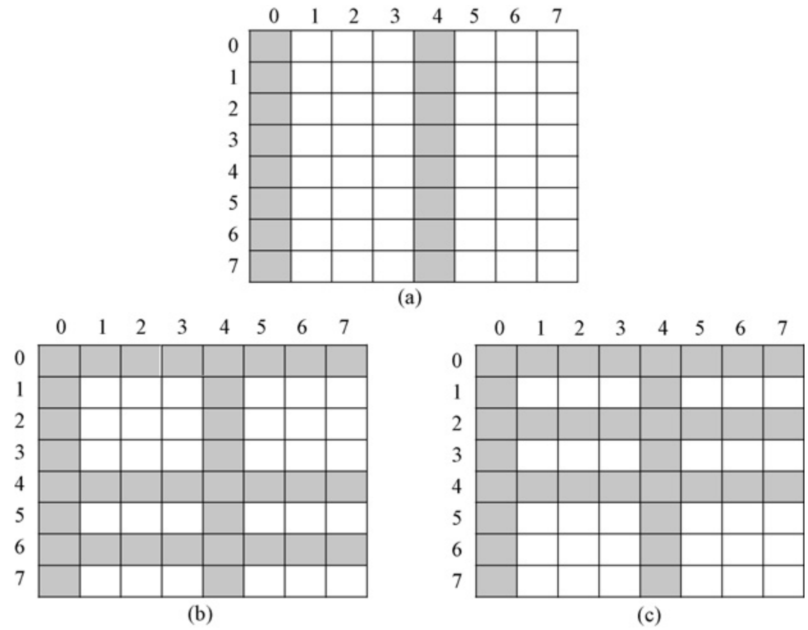



Fig. 1. Partial-zero block patterns (nonzero and zero-quantized frequency positions are indicated with blank and gray blocks, respectively). (a) Type-III. (b) Type-II with SI(2, 1) = SI(6, 1). (c) Type-II with SI(2, 1) < SI(6, 1).

TABLE I

DIFFERENT THRESHOLDS FOR EARLY DETERMINING ZERO-QUANTIZED COEFFICIENTS IN AN 8 × 8 BLOCK

| Type | Conditions | Number of Zero-quantized Coefficients |
|---|---|---|
| I (all-zero block) | $SAD + \frac{3}{4}\sum_{x\in X_0} SAD_x < T_1 = \frac{35}{2} Q_p / \cos^2\left(\frac{\pi}{16}\right)$ | 64 |
| II | $SAD < T_2 = 10\sqrt{2} Q_p / \cos\left(\frac{\pi}{16}\right)$ | 34 |
| III | $SAD + \frac{3}{4}\sum_{x\in X_0} SAD_x < T_3 = \frac{35}{\sqrt{2}} Q_p / \cos\left(\frac{\pi}{16}\right)$ | 16 |
| IV (Normal block) | Otherwise | 0 |

in (11) and (12), during calculating SAD in ME, $\sum_{x=0,7} SAD_x$, $\sum_{x=1,6} SAD_x$, $\sum_{x=2,5} SAD_x$, $\sum_{x=3,4} SAD_x$ should be first computed and temporarily stored and then

$SAD = \sum_{x=0,7} SAD_x + \sum_{x=1,6} SAD_x + \sum_{x=2,5} SAD_x + \sum_{x=3,4} SAD_x$. As a result, in (10), three addition and one shift operations are needed to get $SAD + \frac{3}{4}\sum_{x\in X_0} SAD_x$ additionally. It should be noted that $\frac{35}{2} Q_p / \cos^2\left(\frac{\pi}{16}\right)$ is only dependent on quantization parameter $Q_p$ and thus, can be pre-computed.

Alternatively, the sufficient condition for all-zero block detection in (10) can also be equivalently derived as

$$\begin{aligned} &SAD + \tfrac{3}{4}\textstyle\sum_{x\in X_0} SAD_x < \tfrac{35}{2} Q_p / \cos^2\left(\tfrac{\pi}{16}\right) \\ &\quad = \tfrac{7}{4} SAD - \tfrac{3}{4}\textstyle\sum_{x\in X_1} SAD_x < \tfrac{35}{2} Q_p / \cos^2\left(\tfrac{\pi}{16}\right) \\ &\quad \Rightarrow SAD < \tfrac{3}{7}\textstyle\sum_{x\in X_1} SAD_x + 10 Q_p \cos^2\left(\tfrac{\pi}{16}\right). \end{aligned} \tag{13}$$

Compared with (5), obviously, this sufficient condition for all-zero block detection is more precise than Sousa's model.

### *C. Sufficient Conditions for Partial-zero Blocks Determination*

Furthermore, to reduce the computational complexity for partial-zero blocks, new conditions are derived to check whether DCT coefficients in a specified subgroup are all quantized to zeros. First, if $S^{\mathrm{II}}(u, v)|_{u=1,2,3,5,6,7,v=0,4}$ are defined as $\frac{\sqrt{2}}{8}\left(\cos\left(\frac{\pi}{16}\right)\sum_{x\in X_0} SAD_x + \cos\left(\frac{5\pi}{16}\right)\sum_{x\in X_1} SAD_x\right)$, similar to (10), a sufficient condition for zero-quantized $F(u, v)\,|_{u=1,2,3,5,6,7,v=0,4}$ can be derived as

$$\begin{aligned} &\max(S^{\mathrm{II}}(u, v)|_{u=1,2,3,5,6,7,v=0,4}) \\ &\quad = \tfrac{\sqrt{2}}{8}\left(\cos\left(\tfrac{\pi}{16}\right)\textstyle\sum_{x\in X_0} SAD_x + \cos\left(\tfrac{5\pi}{16}\right)\sum_{x\in X_1} SAD_x\right) \\ &\quad \le \tfrac{\sqrt{2}}{8}\left(\cos\left(\tfrac{\pi}{16}\right)\textstyle\sum_{x\in X_0} SAD_x + \tfrac{4}{7}\cos\left(\tfrac{\pi}{16}\right)\sum_{x\in X_1} SAD_x\right) \\ &\quad = \tfrac{\sqrt{2}}{14}\cos\left(\tfrac{\pi}{16}\right)\left(\tfrac{7}{4}\textstyle\sum_{x\in X_0} SAD_x + \sum_{x\in X_1} SAD_x\right) \\ &\quad < \tfrac{5}{2} Q_p \\ &\quad \Rightarrow SAD + \tfrac{3}{4}\textstyle\sum_{x\in X_0} SAD_x < \tfrac{35}{\sqrt{2}} Q_p / \cos\left(\tfrac{\pi}{16}\right). \end{aligned} \tag{14}$$

Similar to (28) in the Appendix, it can be proven that the larger one of $S^{II}(2, v)\,|v = 0, 4$ and $S^{II}(6, v)\,|v = 0, 4$ is not less than $S^{I}(u, v)|u, v = 0, 4$. Thus, if the criterion in (14) is satisfied, all coefficients of the columns indexed with 0 and 4 in an $8 \times 8$ block, as shown in Fig. 1(a), will be quantized to zeros.

In addition, as compared with (14), we have the following inequality:

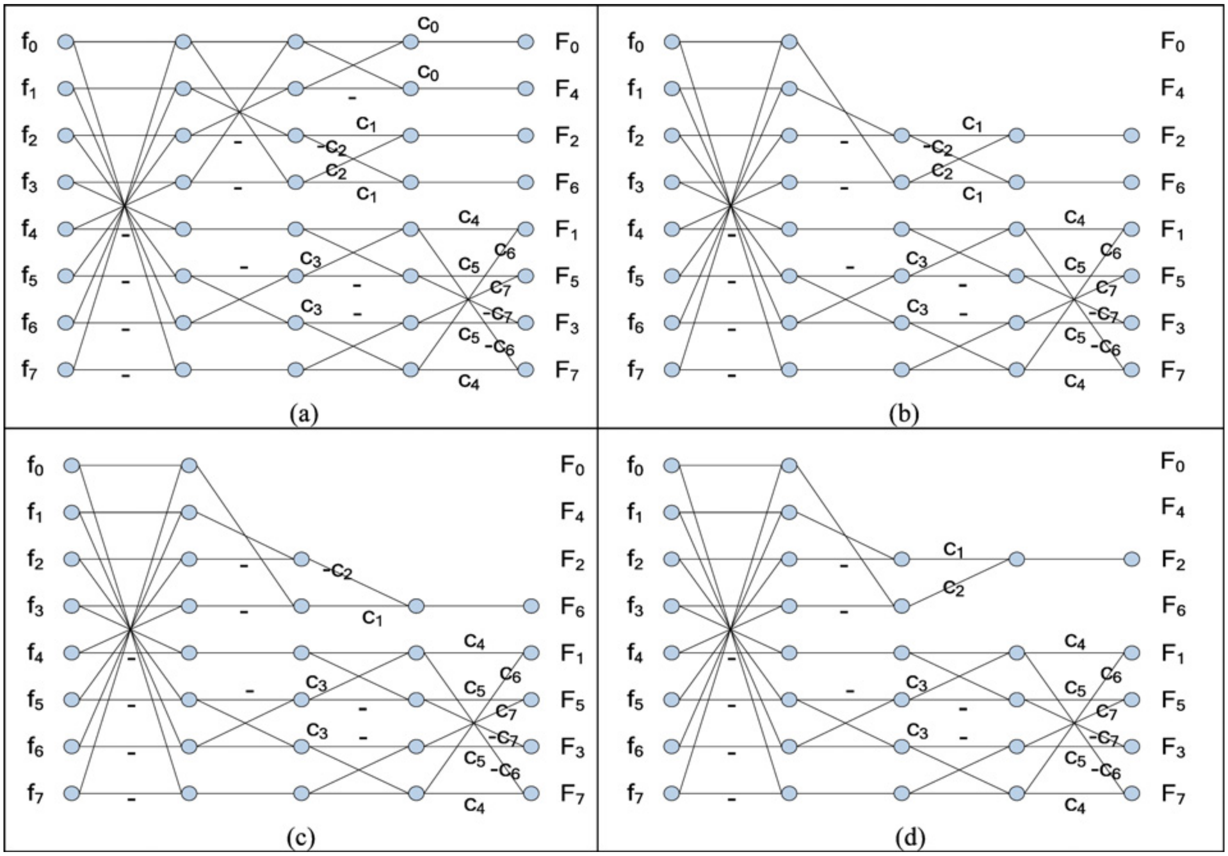



Fig. 2. Chen's 8-point DCT butterfly algorithm and its pruning ones.

$$S^{\mathrm{I}}(u,1)|_{u=0,4} = \frac{\sqrt{2}}{8}\cos\left(\frac{\pi}{16}\right) SAD \geq \frac{\sqrt{2}}{8}\left(\cos\left(\frac{\pi}{16}\right)\sum\nolimits_{x\in X_0} SAD_x + \frac{4}{7}\cos\left(\frac{\pi}{16}\right)\sum\nolimits_{x\in X_1} SAD_x\right) \tag{15}$$

and according to the condition

$$S^{\mathrm{I}}(2,1)+S^{\mathrm{I}}(6,1)-2S^{\mathrm{I}}(u,1)\Big|_{u=0,4} = \frac{1}{4}\cos\left(\frac{\pi}{16}\right)\sum\nolimits_{x=0}^{7} \left(\cos\left(\frac{\pi}{8}\right)+\cos\left(\frac{3\pi}{8}\right)-\sqrt{2}\right) SAD_x \leq 0 \tag{16}$$

the smaller one of $S^{\mathrm{I}}(u,1)\big|_{u=2,6}$ is not larger than $S^{\mathrm{I}}(u,1)|_{u=0,4}$. As a result, if $S^{\mathrm{I}}(u,1)|_{u=0,4} < \frac{5}{2}Q_p \Rightarrow SAD < 10\sqrt{2}Q_p/\cos\left(\frac{\pi}{16}\right)$, all coefficients indicated by the blank box as shown in Fig. 1(b) or (c) will be quantized to zeros. Note that the comparison of $S^{\mathrm{I}}(u,1)\big|_{u=2,6}$ can be directly derived by the comparison of $S^{\mathrm{II}}(u,1)\big|_{u=2,6}$ by (23) in the Appendix and thus, no additional computation is needed. Finally, according to different sufficient conditions to determine zero-quantized coefficients, an $8 \times 8$ block can be classified into four categories as shown in Table 1.

### *D. Fast DCT/IDCT Pruning Algorithms for Partial-Zero Blocks*

As mentioned above, for Type-I, all DCT coefficients in an $8 \times 8$ block can be early determined as zero-quantized and thus, DCT, Q, IQ, and IDCT are not needed. For Type-II and Type-III, only partial DCT coefficients in an $8 \times 8$ block can be early determined as zero-quantized and thus, DCT/IDCT cannot be completely skipped. For these two types, fast DCT/IDCT pruning algorithms are further proposed to reduce the computational complexity.

The 2-D $N \times N$ DCT transform as described in (1) can also be expressed as

$$\mathrm{Y} = \mathrm{AXA}^{\mathrm{T}} \tag{17}$$

where X and Y represent the $N \times N$ input signal matrix and DCT coefficients matrix, respectively. A is an orthogonal $N \times N$ transform matrix and AT is the transpose matrix of A. Each element in A can be expressed by

$$\mathrm{A}(i,j) = C(i)\cos\left(\frac{(2j+1)i\pi}{2N}\right)$$

with

$$C(i) = \sqrt{\frac{1}{N}}, \quad \text{for } i = 0 and C(i) = \sqrt{\frac{2}{N}}, \quad \text{for } i \neq 0.$$

Such a 2-D DCT can be obtained by applying a 1-D DCT to each row of the input coefficients matrix X and followed by a 1-D DCT to each column. Its corresponding Chen's 8-point butterfly-based 1-D DCT [25] is illustrated in Fig. 2(a). For Type-III, as illustrated in Fig. 1(a), DCT coefficients of two columns indexed with 0 and 4 in an $8 \times 8$ block are all quantized to zeros and thus, the 8-point DCT butterfly pruning algorithm as illustrated in Fig. 2(b) can be applied to each row of X. Subsequently, the 1-D DCT transform for the columns indexed with 0 and 4 can be completely skipped and

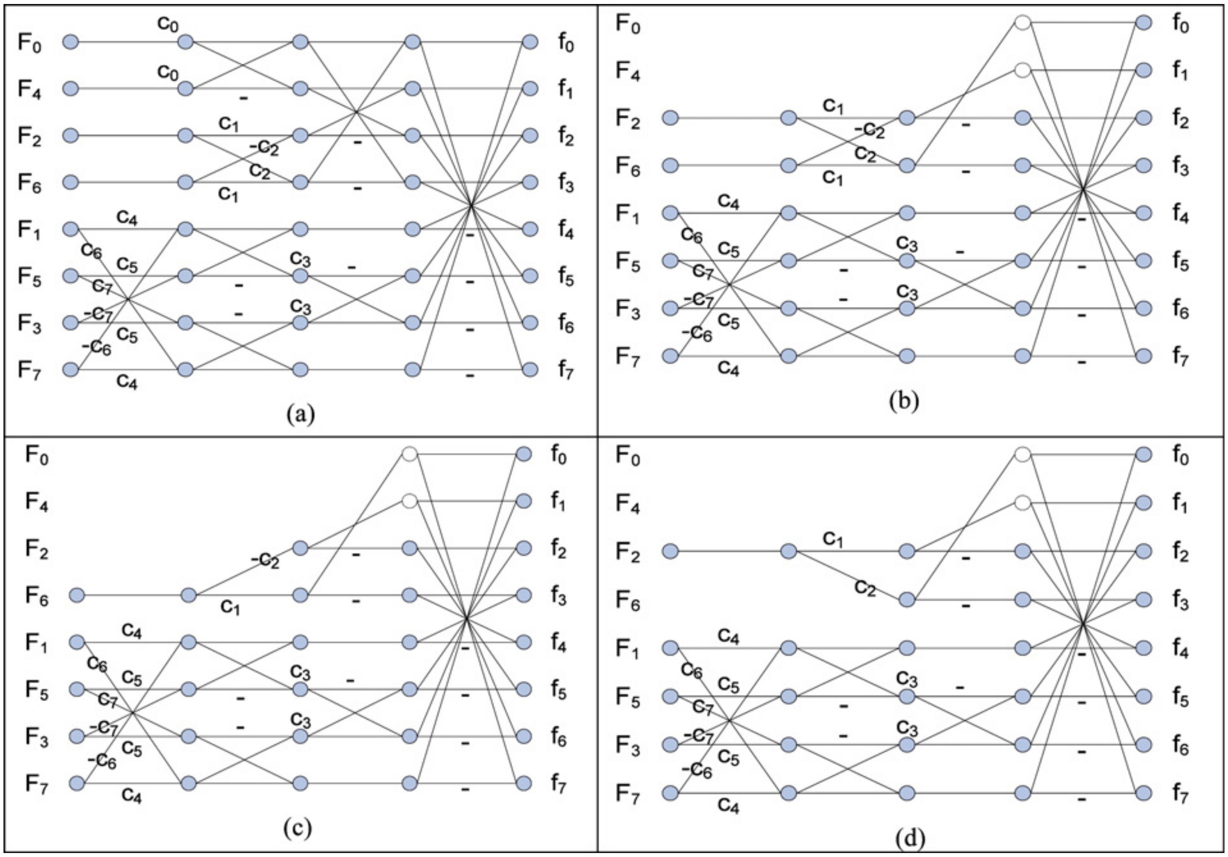


Fig. 3. 8-point Chen–Wang's IDCT butterfly algorithm and its pruning ones.

the normal 8-point butterfly-based 1-D DCT is used for the remaining columns. Correspondingly, for Type-II, the 8-point DCT pruning algorithm as illustrated in Fig. 2(b) is applied to each row of X and then, the 1-D DCT for the columns indexed with 0 and 4 can be completely skipped and the 8-point DCT pruning algorithm, as illustrated in Fig. 2(c) or (d), is used for the remaining columns. However, such a pruning algorithm is still not effective enough. According to the matrix transpose property, (17) can also be expressed as

$$\mathrm{Y} = \mathrm{AXA}^{\mathrm{T}} = (\mathrm{AX}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}})^{\mathrm{T}}. \tag{18}$$

Through (18), for Type-II, the 8-point DCT butterfly pruning algorithm as illustrated in Fig. 2(c) or (d) can be used for each row of the matrix $\mathrm{X}^{\mathrm{T}}$and then, the 1-D DCT for the columns indexed with 0, 4 and 2 or 6 can be completely skipped and meanwhile, the 8-point DCT pruning algorithm as illustrated in Fig. 3(b) is used for the remaining columns. Obviously, this approach has a higher computational complexity saving for Type-II.

On the other hand, to get the reconstructed prediction error block, a 2-D $N \times N$ IDCT transform will be carried out on reconstructed DCT coefficients via

$$\hat{\mathrm{X}} = \mathrm{A}^{\mathrm{T}}\hat{\mathrm{Y}}\mathrm{A}, \tag{19}$$

where '^' is used to represent the reconstructed signal.

In terms of IDCT, the classical Chen–Wang's 8-point IDCT butterfly algorithm [25], [26] is illustrated in Fig. 3(a). For Type-II, the 1-D DCT transform for the rows indexed with 0, 4 and 2 or 6 of $\hat{\mathrm{Y}}$ can be completely skipped and the 8-point IDCT pruning algorithm as illustrated in Fig. 3(c) or (d) will be used for each column. Note that in Fig. 3(b), (c), and (d) circles with the dashed line are used to indicate that no operations are needed. For Type-III, for saving more computational complexity, IDCT will be carried out on matrix $\hat{\mathrm{Y}}^{\mathrm{T}}$ based on

$$\hat{\mathrm{X}} = (\mathrm{A}^{\mathrm{T}}\hat{\mathrm{Y}}^{\mathrm{T}}\mathrm{A})^{\mathrm{T}}. \tag{20}$$

As a result, 1-D IDCT for rows indexed with 0 and 4 of $\hat{\mathrm{Y}}^{\mathrm{T}}$can be completely skipped and the normal 8-point IDCT butterfly algorithm as illustrated in Fig. 3(a) is still performed on the remaining rows of $\hat{\mathrm{Y}}^{\mathrm{T}}$. Subsequently, the fast 8-point IDCT pruning algorithm as illustrated in Fig. 3(b) is used for each column.

Finally, to summarize our above discussion, Fig. 4 depicts the procedure for the early determination of zero-quantized coefficients for an $8 \times 8$ inter-block.

## III. Experimental Results

To verify the efficiency of the proposed method in terms of the computational complexity saving, XVID 1.1.0 [27] is used for testing, which is an MPEG-4 Advanced Simple Profile (ASP) compliant video codec. PMVFAST in [13] is enabled for fast ME. The test sequences include *Foreman, Silent, Table_tennis, News, Akiyo* and *Container* in *CIF*(352

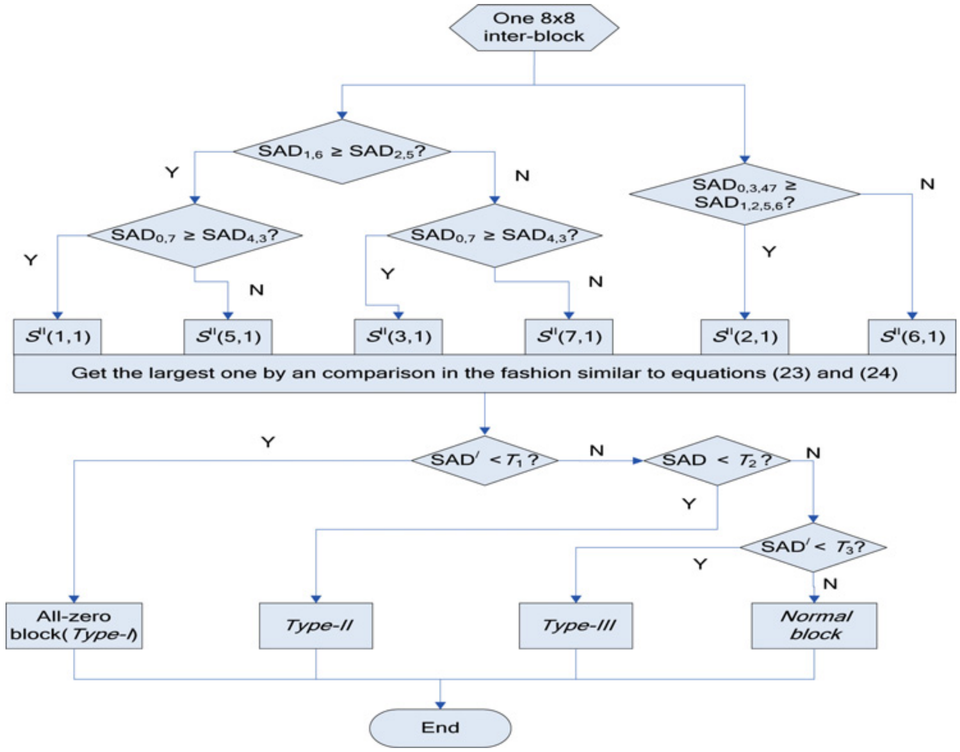



Fig. 4. Proposed early determination algorithm ($SAD' = SAD + \frac{3}{4}\sum_{x\in X_0} SAD_x$).

TABLE II

NUMBER OF OPERATIONS SAVED FOR DCT AND IDCT OF AN 8 × 8 BLOCK AND ADDITIONAL OVERHEAD FOR DIFFERENT TYPES

| | Type | DCT | | | IDCT | | | Overhead | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Add | MUL | Shift | Add | MUL | Shift | Add | Shift | CMP |
| Proposed | All-zero (TYPE-I) | 512 | 256 | 96 | 632 | 176 | 240 | 3 | 1 | 5 |
| | TYPE-II | 186 | 82 | 24 | 195 | 41 | 70 | 3 | 1 | 6 |
| | TYPE-III | 104 | 48 | 16 | 116 | 22 | 40 | 3 | 1 | 7 |
| | Normal | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 1 | 7 |
| Wang's | All-zero (TYPE-I) | 512 | 256 | 96 | 632 | 176 | 240 | 0 | 0 | 1 |
| | TYPE-II | 304 | 136 | 48 | 264 | 68 | 64 | 0 | 0 | 3 |
| | TYPE-III | 148 | 64 | 28 | 164 | 28 | 52 | 0 | 0 | 4 |
| | TYPE-IV | 140 | 60 | 28 | 146 | 22 | 52 | 0 | 0 | 5 |
| | TYPE-V | 44 | 16 | 12 | 38 | 6 | 8 | 0 | 0 | 6 |
| | TYPE-VI | 12 | 4 | 4 | 10 | 0 | 4 | 0 | 0 | 6 |
| | Normal | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |

× 288)@30 Hz and the frame number for each sequence for testing is 300. Only I and P-frames are used for coding and quantization parameters ($Q_p$) used for testing include 14, 21, and 28.

### A. *Definition for Basic Operations*

Chen's 8-point butterfly-based DCT and Chen–Wang's 8-point butterfly-based IDCT algorithms are used for 2-D 8 × 8 DCT and IDCT in XVID [27], respectively. To avoid floating-point computation, they are implemented by integer algorithm with proper scaling. Table II gives the number of basic operations saved by the pruned fast butterfly-based DCT and IDCT algorithms for each type in the proposed model and Wang's model [4], in which Type-I represents the all-zero block determination and the other types are used for the partial-zero blocks determination. In Table II, *Add*, *MUL*, and *CMP* stand for the addition, multiplication, and comparison operations, respectively. The required computational overhead for the determination of zero-quantized coefficients in different types is also listed in Table II. For the proposed method, thresholds in Table I can be precalculated since they are constant for each $Q_p$ and no additional computations are required. It can

TABLE III

NUMBER OF OPERATIONS PER SAMPLE IN (I)Q

| | | Add | MUL | Shift | CMP |
|---|---|---|---|---|---|
| Q | Zero | 1 | | | 2 |
| | Nonzero | 1 | 1 | 1 | 2 |
| IQ | Zero | | | | 1 |
| | Nonzero | 1 | 1 | | 3 |

TABLE IV

*CSR*s FOR DIFFERENT TESTING METHODS

| Sequences | $Q_p$ | CSR (Proposed versus Original) | | | | CSR (Proposed versus Sousa) | | | | CSR (Proposed versus Wang) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Add | MUL | Shift | CMP | Add | MUL | Shift | CMP | Add | MUL | Shift | CMP |
| *Foreman* | 14 | 34.37% | 18.45% | 31.14% | 39.33% | 14.22% | 6.31% | 11.39% | 16.01% | 8.82% | 4.36% | 6.84% | 8.14% |
| | 21 | 47.47% | 25.95% | 46.93% | 53.09% | 17.39% | 7.14% | 15.49% | 19.29% | 10.82% | 4.93% | 9.46% | 9.58% |
| | 28 | 55.76% | 30.72% | 57.65% | 61.64% | 18.85% | 7.14% | 18.06% | 20.82% | 11.73% | 4.91% | 11.03% | 10.13% |
| *Silent* | 14 | 30.27% | 16.28% | 29.47% | 32.75% | 13.04% | 5.91% | 11.38% | 13.42% | 8.16% | 4.12% | 6.74% | 6.86% |
| | 21 | 43.21% | 23.32% | 43.89% | 47.55% | 19.81% | 8.39% | 18.47% | 22.07% | 12.79% | 5.88% | 11.43% | 12.32% |
| | 28 | 54.00% | 29.43% | 56.56% | 59.53% | 25.75% | 10.34% | 25.80% | 29.63% | 16.96% | 7.27% | 16.84% | 17.18% |
| *Table_tennis* | 14 | 33.79% | 18.12% | 31.98% | 37.75% | 12.66% | 5.43% | 10.37% | 13.74% | 8.13% | 3.84% | 6.12% | 7.51% |
| | 21 | 50.30% | 27.40% | 50.39% | 56.04% | 23.37% | 9.73% | 21.80% | 26.75% | 14.97% | 6.77% | 13.97% | 14.47% |
| | 28 | 59.39% | 32.84% | 61.86% | 65.53% | 21.22% | 7.93% | 21.11% | 23.69% | 13.04% | 5.40% | 13.16% | 11.19% |
| *News* | 14 | 50.65% | 27.68% | 51.00% | 56.23% | 19.80% | 7.94% | 18.15% | 22.13% | 12.34% | 5.44% | 11.06% | 11.06% |
| | 21 | 58.43% | 32.21% | 61.02% | 64.43% | 18.87% | 6.81% | 18.23% | 20.89% | 11.85% | 4.69% | 11.02% | 10.17% |
| | 28 | 64.00% | 35.42% | 68.16% | 70.18% | 21.67% | 7.41% | 22.37% | 24.28% | 13.61% | 5.08% | 13.76% | 11.59% |
| *Akiyo* | 14 | 63.87% | 35.18% | 67.50% | 70.44% | 26.77% | 9.54% | 27.64% | 31.43% | 17.07% | 6.51% | 17.75% | 16.68% |
| | 21 | 70.13% | 39.03% | 76.08% | 76.70% | 24.21% | 7.68% | 27.59% | 27.85% | 15.52% | 5.35% | 17.96% | 13.60% |
| | 28 | 74.35% | 41.44% | 81.34% | 81.12% | 25.33% | 7.30% | 30.52% | 29.62% | 16.14% | 5.03% | 19.64% | 13.57% |
| *Container* | 14 | 43.55% | 23.55% | 43.38% | 48.29% | 19.52% | 8.28% | 17.74% | 21.88% | 12.50% | 5.78% | 10.99% | 11.94% |
| | 21 | 57.24% | 31.45% | 59.63% | 63.09% | 22.81% | 8.69% | 22.46% | 25.84% | 14.42% | 5.97% | 14.10% | 13.29% |
| | 28 | 64.32% | 35.72% | 68.90% | 70.24% | 21.40% | 7.40% | 22.65% | 23.33% | 13.07% | 5.01% | 13.89% | 10.15% |

TABLE V

COMPARISON OF RATIOS FOR DIFFERENT TYPES IN THE PROPOSED MODEL AND WANG'S MODEL

| | $Q_p$ | Proposed | | | | Wang's | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Type-I | Type-II | Type-III | Normal | Type-I | Type-II | Type-III | Type-IV | Type-V | Type-VI | Normal |
| *Foreman* | 14 | 38.54% | 5.56% | 9.23% | 46.67% | 30.76% | 2.24% | 2.50% | 8.62% | 2.88% | 12.36% | 40.64% |
| | 21 | 52.14% | 5.10% | 7.77% | 35.00% | 44.32% | 2.29% | 2.48% | 8.15% | 2.40% | 10.49% | 29.87% |
| | 28 | 60.38% | 4.60% | 7.33% | 27.68% | 53.45% | 2.17% | 2.07% | 7.31% | 2.18% | 10.08% | 22.73% |
| *Silent* | 14 | 30.06% | 5.21% | 9.90% | 54.83% | 23.91% | 1.65% | 1.96% | 7.76% | 3.01% | 14.42% | 47.29% |
| | 21 | 42.08% | 7.24% | 11.56% | 39.12% | 34.12% | 2.00% | 2.30% | 10.90% | 3.75% | 16.14% | 30.80% |
| | 28 | 53.97% | 6.86% | 11.44% | 27.73% | 43.75% | 2.84% | 3.13% | 11.10% | 3.44% | 15.54% | 20.18% |
| *Table_tennis* | 14 | 35.02% | 5.20% | 11.62% | 48.16% | 30.13% | 1.27% | 1.48% | 7.35% | 3.24% | 16.36% | 40.18% |
| | 21 | 53.11% | 6.84% | 8.45% | 31.60% | 42.04% | 3.07% | 3.37% | 11.48% | 2.77% | 10.26% | 27.00% |
| | 28 | 65.10% | 4.19% | 6.17% | 24.54% | 56.74% | 2.89% | 2.49% | 7.16% | 1.83% | 8.89% | 20.00% |
| *News* | 14 | 53.96% | 6.47% | 7.52% | 32.04% | 45.92% | 2.31% | 2.55% | 9.66% | 2.89% | 9.17% | 27.49% |
| | 21 | 62.64% | 5.00% | 7.36% | 25.00% | 56.81% | 1.72% | 1.95% | 7.16% | 2.42% | 8.98% | 20.96% |
| | 28 | 68.40% | 5.17% | 6.12% | 20.32% | 62.06% | 1.82% | 1.83% | 7.85% | 2.07% | 7.75% | 16.62% |
| *Akiyo* | 14 | 67.24% | 6.16% | 7.74% | 18.87% | 58.37% | 3.07% | 2.40% | 9.55% | 3.09% | 7.57% | 15.94% |
| | 21 | 75.80% | 3.58% | 5.89% | 14.74% | 68.89% | 1.97% | 2.18% | 6.34% | 1.78% | 7.07% | 11.77% |
| | 28 | 80.03% | 4.20% | 4.73% | 11.04% | 74.16% | 1.60% | 1.92% | 6.56% | 1.46% | 5.57% | 8.74% |
| *Hall_monitor* | 14 | 53.80% | 7.51% | 9.89% | 28.80% | 40.16% | 4.56% | 4.47% | 12.14% | 3.33% | 12.13% | 23.20% |
| | 21 | 66.24% | 4.62% | 7.70% | 21.45% | 59.84% | 2.09% | 2.11% | 6.83% | 2.37% | 9.80% | 16.96% |
| | 28 | 71.99% | 4.58% | 6.26% | 17.17% | 66.13% | 1.73% | 1.79% | 6.92% | 1.92% | 8.18% | 13.34% |
| *Container* | 14 | 44.09% | 6.97% | 10.49% | 38.46% | 35.63% | 2.23% | 2.65% | 10.55% | 3.39% | 14.01% | 31.54% |
| | 21 | 60.32% | 6.08% | 7.65% | 25.96% | 51.89% | 2.56% | 2.86% | 9.08% | 2.68% | 9.24% | 21.69% |
| | 28 | 69.41% | 4.60% | 4.84% | 21.16% | 62.55% | 2.22% | 2.08% | 7.15% | 1.78% | 6.04% | 18.17% |

be clearly observed in Fig. 4 that several extra comparison operations are required in early determination of zero-quantized coefficients. Additionally, three *add* and one *shift* operations to calculate $SAD + \frac{3}{4}\sum_{x\in X_0} SAD_x$ in (10) is required by calculating $SAD + \sum_{x\in X_0} SAD_x - ((\sum_{x\in X_0} SAD_x) >> 2)$. Note that the calculation of $SAD + \frac{3}{4}\sum_{x\in X_0} SAD_x$ in (14) is not needed any more, since it is the same as the one in (10). In addition, Table III enumerates the number of basic operations per sample in Q and IQ for zero-quantized and nonzero-quantized coefficients in the original algorithm [27]. Obviously, Q and IQ for each determined zero-quantized sample in different detection types of blocks can be skipped. On the other hand, since the early determination condition is sufficient but not necessary for all-zero block determination, the block not belonging to Type-I, after Q, still needs a check to see if it is an all-zero block. If a block has zero summation, IQ and IDCT will be skipped; otherwise, the normal or pruned IQ and IDCT algorithms will be further applied.

### *B. Computational Complexity Comparison*

To compare the computational complexity in terms of DCT, Q, IQ, and IDCT, the computational saving ratio (CSR) between the two different methods is defined by

$$CSR = \frac{TC_1 - TC_0}{TC_1}\% \tag{21}$$

where $TC_0$ and $TC_1$ represent the total calculations for two different methods. Table IV gives the comparisons of the proposed method against other existing methods in terms of the computational complexity based on the basic operation numbers listed in Tables II and III. In Table IV, it can be observed that the proposed method significantly saves the computational complexity as compared with the original algorithm. Especially, for the *Akiyo* sequence, up to 74.35% *add*, 81.34% *Shift*, and 81.12% *CMP* operations can be saved when $Q_p$ is 28. On the other hand, it also can be found that with larger $Q_p$, *CSR* is also much higher since more DCT coefficients will be early determined as zero-quantized at low bit-rate coding.

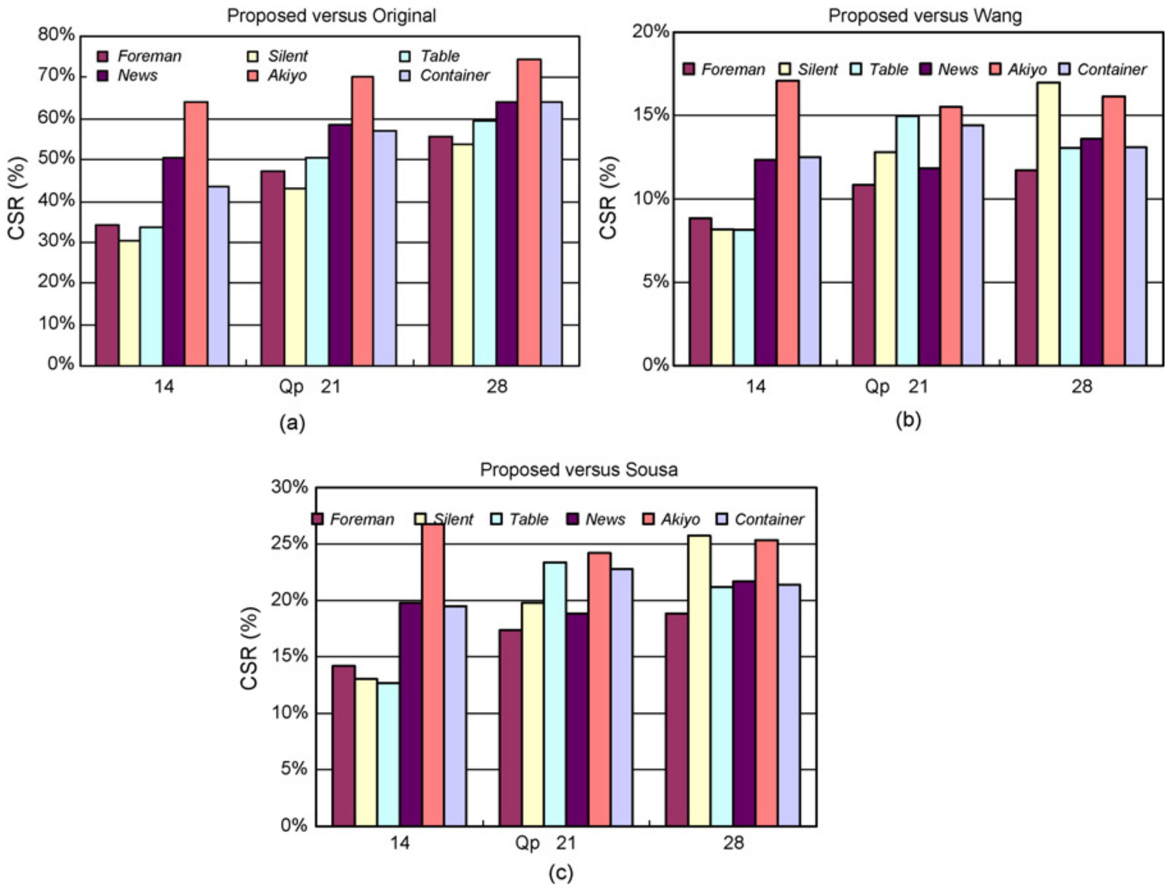



Fig. 5. Comparison of CSR of the proposed algorithm against the different methods in terms of the Add operation number for several test testing sequences.

This tendency can be clearly observed in Fig. 5(a), which gives *CSR* in terms of the number of *Add* operations.

Furthermore, the proposed method is compared with Sousa's [19] and Wang's [4] models. Sousa's model provides a sufficient condition for early determination of all-zero blocks and, Wang's model use the same condition as Sousa's model for all-zero blocks determination and further gives several sufficient conditions for partial-zero block determination. As compared with Sousa's model and Wang's model, our model provides a more precise sufficient condition for early determination of all-zero blocks. Table V gives the ratios for different types in the proposed model and Wang's model. It can be seen that the percentage of all-zero blocks (Type-I) in the proposed model is higher than Wang's model and Sousa's model. As shown in Table V for sequence *Table_tennis* at $Q_p = 21$, only 42.04% blocks are determined as all-zero ones for Wang's mode while 53.11% for the proposed model. On the other hand, as compared with Wang's model for partial-zero blocks detection, Wang's model has more types for the partial-zero blocks determination by finer categorization. In fact, the proposed method can also adopt a similar fine categorization scheme. However, since some of types such as Type-V and Type-VI only save a limited amount of computational complexity, the proposed method only uses two types for partial-zero blocks detection. It can be observed that, although the proposed model has a higher percentage for normal blocks, as shown in Table V, it can still save the computational complexity more efficiently than Wang's model. The detailed *CSR* comparisons of the proposed model and Sousa's model and Wang's models are given in Table IV and Fig. 5.

## IV. Conclusion and Future Work

A novel and efficient early determination scheme for zero-quantized DCT coefficients has been proposed in this paper. It provides more precise sufficient conditions for the early determination of all-zero and partial-zero blocks as compared with existing methods through a thorough theoretical analysis of the dynamic range of DCT coefficients based on row-based SAD. Furthermore, for partial-zero blocks, fast 8-point butterfly-based DCT and IDCT algorithms with pruning have been introduced to save more computational complexity. Experimental results demonstrated that the computational complexity of DCT, Q, IQ, and IDCT is further saved by the proposed approach against the original algorithm and other existing fast algorithms. Thus, the proposed model is more applicable for fast video encoding applications, especially for mobile wireless applications with low target bitrates.

It should be pointed out that in this paper, the proposed fast DCT/IDCT pruning algorithms for partial-zero blocks is implemented based on Chen's DCT in XVID codec. Certainly, in practical applications, other faster DCT algorithms [28], [29] can also be employed and even further implemented with quantization together.

On the other hand, the proposed model can also be combined with other statistical models such as [24], [30], and [31] to further reduce DCT, Q, IQ, and IDCT computations.

## Appendix

In (11), the maximum item among $S^{\mathrm{II}}(u,1)\big|_{u=1,3,5,7}$ can be obtained according to

$$\begin{cases} S^{\mathrm{II}}(1,1) - S^{\mathrm{II}}(3,1) = S^{\mathrm{II}}(5,1) - S^{\mathrm{II}}(7,1) \\ = \frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\left(\frac{\pi}{16}\right) - \cos\left(\frac{5\pi}{16}\right)\right) \\ \left(\sum_{x=1,6} SAD_x - \sum_{x=2,5} SAD_x\right) \\ S^{\mathrm{II}}(1,1) - S^{\mathrm{II}}(5,1) = S^{\mathrm{II}}(3,1) - S^{\mathrm{II}}(7,1) \\ = \frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\left(\frac{\pi}{16}\right) - \cos\left(\frac{5\pi}{16}\right)\right) \\ \left(\sum_{x=0,7} SAD_x - \sum_{x=3,4} SAD_x\right). \end{cases} \tag{22}$$

In (12), the larger one of $S^{\mathrm{II}}(u,1)\big|_{u=2,6}$ can be obtained by

$$S^{\mathrm{II}}(2,1) - S^{\mathrm{II}}(6,1) = \frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\cos\left(\frac{\pi}{16}\right) - \cos\left(\frac{5\pi}{16}\right)\right) \times \left(\sum_{x=0,3,4,7} SAD_x - \sum_{x=1,2,5,6} SAD_x\right). \tag{23}$$

For simplicity, assume the maximum item is $S^{\mathrm{II}}(1,1)$ after the comparisons in (22). $S^{\mathrm{II}}(1,1)$ can be further compared with the larger one of $S^{\mathrm{II}}(u,1)\big|_{u=2,6}$ by

$$S^{\mathrm{II}}(1,1) - S^{\mathrm{II}}(2,1) = \frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\cos\left(\frac{\pi}{16}\right) - \cos\left(\frac{5\pi}{16}\right)\right) \times \left(\sum_{x=1,6} SAD_x - \sum_{x=3,4} SAD_x\right) \tag{24}$$

or

$$S^{\mathrm{II}}(1,1) - S^{\mathrm{II}}(6,1) = \frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\cos\left(\frac{\pi}{16}\right) - \cos\left(\frac{5\pi}{16}\right)\right) \times \left(\sum_{x=0,7} SAD_x - \sum_{x=2,5} SAD_x\right). \tag{25}$$

Note that the similar derivation procedure can be applied when the maximum item is $S^{\mathrm{I}}(3,1)$, $S^{\mathrm{I}}(5,1)$ or $S^{\mathrm{I}}(7,1)$ after the comparisons by (21).

On the other hand, when comparing $S^{\mathrm{II}}(u,1)\big|_{u=2,6}$ with $S^{\mathrm{I}}(u,1)\big|_{u=0,4}$, we have

$$\begin{aligned} S^{\mathrm{II}}(2,1) - S^{\mathrm{I}}(u,1)\big|_{u=0,4} &= \frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\left(\cos\left(\frac{\pi}{16}\right) - \frac{\sqrt{2}}{2}\right)\sum_{x=0,3,4,7} SAD_x \right. \\ &\quad \left. + \left(\cos\left(\frac{5\pi}{16}\right) - \frac{\sqrt{2}}{2}\right)\sum_{x=1,2,5,6} SAD_x\right) \end{aligned} \tag{26}$$

and

$$\begin{aligned} S^{\mathrm{II}}(6,1) - S^{\mathrm{I}}(u,1)\big|_{u=0,4} &= \frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\left(\cos\left(\frac{5\pi}{16}\right) - \frac{\sqrt{2}}{2}\right)\sum_{x=0,3,4,7} SAD_x \right. \\ &\quad \left. + \left(\cos\left(\frac{\pi}{16}\right) - \frac{\sqrt{2}}{2}\right)\sum_{x=1,2,5,6} SAD_x\right). \end{aligned} \tag{27}$$

As a result, the following inequality is satisfied:

$$S^{\mathrm{II}}(2,1) + S^{\mathrm{II}}(6,1) - 2S^{\mathrm{I}}(u,1)\big|_{u=0,4} = \frac{1}{4}\cos\left(\frac{\pi}{16}\right)\left(\cos\left(\frac{\pi}{16}\right) + \cos\left(\frac{5\pi}{16}\right) - \sqrt{2}\right)\sum_{x=0}^{7} SAD_x \geq 0. \tag{28}$$

This inequality implies that the larger one of $S^{\mathrm{II}}(2,1)$ and $S^{\mathrm{II}}(6,1)$ is always no less than $S^{\mathrm{I}}(u,1)\big|_{u=0,4}$.

According to the above derivation, all DCT coefficients will be quantized to zeros if

$$\begin{aligned} |F(u,v)| &\leq S^{\mathrm{I}}(u,v)\big|_{v=1,3,5,7} \\ &\leq \max\left(S^{\mathrm{II}}(u,1)\big|_{u=2,3,4,5,6,7}\right) \\ &< \tfrac{5}{2}Q_p. \end{aligned} \tag{29}$$

## References


[1] *Video Coding for Low Bit Rate Communication (ITU-T Rec. H.263)*, Feb., 1998.

[2] *Coding of Audio-visual Objects: Part 2. Visual*, ISO/IEC 14496-2 (MPEG-4 Part 2), Jan. 1999.

[3] K. Yu, J. Lv, J. Li, and S. Li, "Practical real-time video codec for mobile devices," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, vol. 3. Jul. 2003, pp. 509–512.

[4] H. Wang, S. Kwong, C.-W. Kok, and M.-Y. Chan, "Zero-quantized discrete cosine transform prediction technique for video encoding," in *Proc. IEE Vis. Image Signal Process.*, vol. 153, no. 5. Oct. 2006, pp. 677–683.

[5] L. M. Po and W. C. Ma, "A novel four-step search algorithm for fast block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 313–317, Jun. 1996.

[6] J. Chalidabhongse and C.-C. J. Kuo, "Fast motion vector estimation using multiresolution-spatio-temporal correlations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 3, pp. 477–488, Jun. 1997.

[7] Z. He and M. L. Liou, "A high performance fast search algorithm for block matching motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 5, pp. 826–828, Oct. 1997.

[8] J. Y. Tham, S. Ranganath, M. Ranganath, and A. A. Kassim, "A novel unrestricted center-biased diamond search algorithm for block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 4, pp. 369–377, Aug. 1998.

[9] S. Zhu and K. K. Ma, "A new diamond search algorithm for fast block matching motion estimation," *IEEE Trans. Image Process.*, vol. 9, no. 2, pp. 287–290, Feb. 2000.

[10] C. Zhu, X. Lin, and L. P. Chau, "Hexagon-based search pattern for fast block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 5, pp. 349–355, May 2002.

[11] C. Zhu, X. Lin, L. P. Chau, and L. M. Po, "Enhanced hexagonal search for fast block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 10, pp. 1210–1214, Oct. 2004.

[12] R. Li, B. Zeng, and M. L. Liou, "A new three-step search algorithm for block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, no. 4, pp. 438–442, Aug. 1994.

[13] A. M. Tourapis, O. C. Au, and M. L. Liou, "Predictive motion vector field adaptive search technique (PMVFAST): Enhancing block-based motion estimation," in *Proc. Vis. Commun. Image Process. Conf. (VCIP)*, vol. 4310. 2001, pp. 883–892.

[14] H. Wang, S. Kwong, C.-W. Kok, and M. Chan, "Zero-quantized discrete cosine transform prediction technique for video encoding," in *Proc. IEE Vis. Image Signal Process.*, vol. 153, no. 5. Oct. 2006, pp. 677–683.

[15] I. E. G. Richardson and Y. Zhao, "Video CODEC complexity management," in *Proc. Picture Coding Symp. (PCS)*, Apr. 2001, pp. 647-649.

[16] A. Yu, R. Lee, and M. Flynn, "Early detection of all-zero coefficients in H.263," in *Proc. Picture Coding Symp. (PCS)*, 1997, pp. 159–164.

[17] X. Zhou, Z. Yu, and S. Yu, "Method for detecting all-zero DCT coefficients ahead of discrete cosine transformation and quantization," *Electron. Lett.*, vol. 34, no. 19, pp. 1839–1840, Sep. 1998.

[18] *Intelligent Pre-Quantization in Motion Compensated Video Coding*, document Q15-D-35.doc, ITU-T SG 16 VCEG, Apr. 1998.

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

[20] S. Jun and S. Yu, "Efficient method for early detection of all-zero DCT coefficients," *Electron. Lett.*, vol. 27, no. 3, pp. 160–161, Feb. 2001.

[21] Y. Wang, Y. Zhou, and H. Yang, "Early detection method of all-zero integer transform coefficients," *IEEE Trans. Consum. Electron.*, vol. 50, no. 3, pp. 923–928, Aug. 2004.

[22] G. Y. Kim, Y. H. Moon, 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.

[23] 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.

[24] I. M. Pao and M. T. Sun, “Modeling DCT coefficients for fast video encoding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 4, pp. 608–616, Jun. 1999.

[25] W. H. Chen, C. H. Smith, and S. C. Fralick, “A fast computational algorithm for the discrete cosine transform,” *IEEE Trans. Commun.*, vol. 25, no. 9, pp. 1004–1009, Sep. 1977.

[26] Z. Wang, “Fast algorithm for the discrete W transform and for the discrete Fourier transform,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 32, no. 4, pp. 803–816, Aug. 1984.

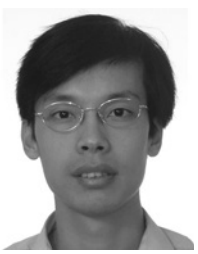
[27] XVID codec. (2005). [Online]. Available: http://www.xvid.org

[28] N. I. Cho and S. U. Lee, “Fast algorithm and implementation of 2-D DCT,” *IEEE Trans. Circuits Syst.*, vol. 38, no. 3, pp. 297–305, Mar. 1991.

[29] P. Duhamel and C. Guillemot, “Polynomial transform computation of the 2-D DCT,” in *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Apr. 1990, pp. 1515–1518.

[30] H. Wang, S. Kwong, and C.-W. Kok, “Efficient predictive model for zero-quantized DCT coefficients for fast video encoding,” *Image Vis. Comput.*, vol. 25, no. 6, pp. 922–933, Jun. 2007.

[31] 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.



**Xiangyang Ji** received the B.S. in materials and M.S. degrees in computer science from Harbin Institute of Technology, Harbin, China, in 1999 and 2001 respectively, and the Ph.D. degree in computer science from both the Graduate School and the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2008.

He is now with Broadband Networks and Digital Media Laboratory, Automation Department, Tsinghua University, Beijing, China. He has authored or coauthored over 40 conference and journal papers. His research interests include video/image coding, video streaming, and multimedia processing.

**Sam Kwong** received the B.Sc. and M.A.Sc. degrees in electrical engineering from the State University of New York, Buffalo, and the University of Waterloo, Waterloo, ON, Canada, in 1983 and 1985, respectively, and the Ph.D. degree in computer science from the University of Hagen, Germany, in 1996.

From 1985 to 1987, he was a Diagnostic Engineer with Control Data Canada. He later joined Bell Northern Research Canada as a member of scientific staff and he then joined the City University of Hong Kong, Kowloon, Hong Kong, China as a Lecturer in the Department of Electronic Engineering, in 1990. He is currently an Associate Professor in the Department of Computer Science, City University of Hong Kong. His research interests include video coding and evolutionary algorithms.

**Debin Zhao** received the B.S., M.S., and Ph.D. degrees in computer science from Harbin Institute of Technology, Harbin, China, in 1985, 1988, and 1998, respectively.

He joined the Department of Computer Science, Harbin Institute of Technology as an Associate Professor, in 1993. Currently, he is a Professor at Harbin Institute of Technology and Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. He has been a Research Fellow in the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, China. His research interests include multimedia compression and its related applications. He has authored or coauthored over 100 publications.

Dr. Zhao has obtained the National Science and Technology Progress Award of China (Second Prize) three times. He also obtained an Excellent Teaching Award from Baogang Foundation.

**Hanli Wang** (M’08) received the B.Sc. and M.Sc. degrees, both with the Best Graduate Honor in electrical engineering from Zhejiang University, Hangzhou, China, in 2001 and 2004, respectively, and the Ph.D. degree in computer science, with the Outstanding Research Thesis Award, from the City University of Hong Kong, Kowloon, Hong Kong, China, in 2007.

From November 2007 to August 2008, he was a Visiting Scholar at Stanford University, Stanford, CA, invited by Professor Charles K. Chui. Since November 2008, he has been a Research Engineer in Precoad Inc., a start-up company in Menlo Park, CA. His current research interests include digital video coding, multimedia communication, video tracking, evolutionary computation, and pattern recognition.

Dr. Wang won the Alexander von Humboldt Research Fellowship in March 2008. From November 2007 to November 2008, he was a Research Fellow at the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, China. He is currently a Humboldt Research Fellow with Fern Universität in Hagen, Lehrgebiet Informationstechnik, Hagen, Germany.

**C.-C. Jay Kuo** (F’99) received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1980, and the M.S. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1985 and 1987, respectively.

He is the Director of the Signal and Image Processing Institute and a Professor of electrical engineering, computer science and mathematics at the Department of Electrical Engineering and Integrated Media Systems Center, University of Southern California, Los Angeles. His research interests include digital image/video analysis and modeling, multimedia data compression, communication and networking, and biological signal/image processing. He is the coauthor of about 150 journal papers, 780 conference papers and 9 books.

Dr. Kuo is a Fellow of The International Society for Optical Engineers (SPIE).

**Qionghai Dai** (SM’05) received the B.S. degree in mathematics from Shanxi Normal University, Xi’an, China, in 1987, and the M.E. and Ph.D. degrees in computer science and automation from Northeastern University, Shenyang, Liaoning, China, in 1994 and 1996, respectively.

Since 1997, he has been with the faculty of Tsinghua University, Beijing, China, and is currently a Professor and the Director of the Broadband Networks and Digital Media Laboratory, Department of Automation, Tsinghua University. His research areas include signal processing, broadband networks, video processing, and communication.