

ANALYTICAL MODEL OF ZERO QUANTIZED DCT COEFFICIENTS FOR VIDEO ENCODER OPTIMIZATION

Hanli Wang¹, Sam Kwong¹, Chi-Wah Kok²

¹Department of Computer Science, City University of Hong Kong, Hong Kong, China

E-mail: wanghl@cs.cityu.edu.hk, cssamk@cityu.edu.hk

² Hong Kong Applied Science & Technology Research Institution Co. Ltd., Hong Kong, China

E-mail: tedkok@astri.org

ABSTRACT

This paper proposes a novel analytical model to predict zero quantized DCT coefficients for fast video encoding. The dynamic range of quantized DCT coefficients are analyzed and a threshold scheme is derived in order to determine DCT and quantization computations to be skipped without video quality degradation. The proposed model is compared with other models in the literature. Experimental results demonstrate that the proposed analytical model can greatly reduce the computational complexity of video encoding without any performance degradation, and outperforms other models.

1. INTRODUCTION

Most high quality video encoders, such as MPEG-4 [1], H.263 [2] and H.264 [3], use the discrete cosine transform (DCT), motion estimation (ME), quantization (Q), inverse quantization (IQ) and inverse DCT (IDCT) as the building blocks. Such video encoders are computationally intensive. In fact, there is a significant interest and research in reducing these computations. Previously, the efforts to reduce the computations of video encoding are mainly focused on fast motion estimation algorithms. However, as the motion estimation becomes optimized, we also need to optimize other functions to further speed up video encoding.

In digital video coding, it is quite common that a substantial number of DCT coefficients of the prediction difference are quantized to zeros. Therefore, considerable computations may be saved if there is a method which can early detect Zero Quantized DCT (ZQDCT) coefficients, i.e., the DCT coefficients equal to zero after Q, before implementing DCT and Q. Yu *et al.* [4] propose to compare the sum of absolute difference SAD with the product of the quantization parameter Q_p and a predetermined threshold T . If $SAD < T \times Q_p$, then DCT and Q computations can be skipped, and the quantized DCT coefficients are all set to zeros. This model is shown to be effective in reducing the computational complexity of the H.263 encoder. However, the quality of the encoded video is heavily dependent on the threshold T , where to define a suit-

able value is not trivial. In order to reduce the degradation of video quality, Yu *et al.* [5] decrease the threshold value experimentally to detect all-zero DCT blocks. In [6], Zhou *et al.* perform theoretical analyses on the range of DCT coefficients and derive the same threshold as [5] to skip DCT and Q computations. Zhou's model [6] is further refined by Sousa [7] and Kim [8] where tighter sufficient conditions are derived to obtain more reductions of the computational complexity without video quality degradation.

However, all the models mentioned above only consider to detect all-zero DCT blocks and do not consider the more general case to predict ZQDCT coefficients in the individual coefficient level. Therefore, this paper proposes a novel analytical model to skip redundant DCT and Q computations without quality degradation through a more theoretical analysis on the dynamic range of DCT coefficients. The proposed analytical model can not only detect all-zero DCT blocks, but also predict individual ZQDCT coefficients. As a result, higher prediction efficiency and more savings in the computational complexity can be achieved by the proposed model.

The rest of this paper is organized as follows. The ZQDCT coefficients are analyzed in Section 2. In Section 3, the novel analytical model is proposed and compared with other models theoretically. The experimental results are presented in Section 4. Finally, Section 5 concludes this paper and gives some future research directions.

2. ANALYSIS OF ZQDCT COEFFICIENTS

We will firstly analyze the sufficient condition for the quantized DCT coefficients to be zeros. In this paper, we mainly focus on the 8×8 DCT which is widely used in MPEG-4 [1] and H.263 [2] standards and will consider the 4×4 DCT in our future research. We define $f(x, y)$, $0 \leq x, y \leq 7$, as the 8×8 motion-compensated pixel block, such that

$$f(x, y) = I(x, y) - I_m(x, y), \quad 0 \leq x, y \leq 7 \quad (1)$$

where $I(x, y)$ is the image block and $I_m(x, y)$ is the best-matched block predicted from the reference frame. The best-matched block is obtained in the motion estimation stage to

Table 1. Thresholds for prediction of ZQDCT coefficients.

Threshold	DCT Coefficients (u, v)
$T_1 = \frac{4\alpha Q_p}{\cos^2(\frac{\pi}{16})}$	$u = 1, 3, 5, 7, v = 1, 3, 5, 7$
$T_2 = \frac{4\alpha Q_p}{\cos(\frac{\pi}{16})\cos(\frac{\pi}{8})}$	$u = 1, 3, 5, 7, v = 2, 6$ $u = 2, 6, v = 1, 3, 5, 7$
$T_3 = \frac{4\alpha Q_p}{\cos^2(\frac{\pi}{8})}$	$u = 2, 6, v = 2, 6$
$T_4 = \frac{4\sqrt{2}\alpha Q_p}{\cos(\frac{\pi}{16})}$	$u = 0, 4, v = 1, 3, 5, 7$ $u = 1, 3, 5, 7, v = 0, 4$
$T_5 = \frac{4\sqrt{2}\alpha Q_p}{\cos(\frac{\pi}{8})}$	$u = 0, 4, v = 2, 6$ $u = 2, 6, v = 0, 4$
$T_6 = 8\alpha Q_p$	$u = 0, 4, v = 0, 4$

minimize the sum of absolute difference SAD given by

$$SAD = \sum_{x=0}^7 \sum_{y=0}^7 |f(x, y)| \quad (2)$$

The two-dimensional 8×8 DCT coefficients $F(u, v)$, $0 \leq u, v \leq 7$, are computed by

$$F(u, v) = \frac{C(u)C(v)}{4} \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos\left(\frac{(2x+1)u\pi}{16}\right) \cos\left(\frac{(2y+1)v\pi}{16}\right) \quad (3)$$

where $C(u), C(v) = 1/\sqrt{2}$, for $u, v = 0$, and $C(u), C(v) = 1$, otherwise. $F(u, v)$ are quantized for compression, and will be zeros if the following condition holds true

$$F(u, v) < \alpha Q_p \quad (4)$$

where Q_p is the quantization parameter and α is related to the quantization method applied. For example, the quantization performed in H.263 and MPEG-4 inter mode follows

$$L(u, v) = \text{sign}(F(u, v)) \times \left\lfloor \frac{|F(u, v)| - \frac{Q_p}{2}}{2Q_p} \right\rfloor \quad (5)$$

where $L(u, v)$ is the quantized DCT coefficient. The DCT coefficients are quantized to zeros if $|L(u, v)| < 1$. Therefore, when $|F(u, v)| < 2.5Q_p$, the coefficients $F(u, v)$ will be quantized to zeros. As a result, α should be chosen as $\alpha = 2.5$.

3. PROPOSED ANALYTICAL MODEL

3.1. Sufficient Condition for ZQDCT Prediction

From (2) and (3), the DCT coefficient $F(u, v)$ is bounded by

$$F(u, v) \leq \frac{C(u)C(v)}{4} \max_{x, y} \left\{ \left| \cos\left(\frac{(2x+1)u\pi}{16}\right) \right| \left| \cos\left(\frac{(2y+1)v\pi}{16}\right) \right| \right\} \times SAD \quad (6)$$

We start our discussion on the range of $F(u, v)$ by considering the case of $u = 4$ and $v = 2$ such that

$$F(4, 2) \leq \frac{1}{4} \max_{x, y} \left\{ \left| \cos\left(\frac{(2x+1)\pi}{4}\right) \right| \left| \cos\left(\frac{(2y+1)\pi}{8}\right) \right| \right\} \times SAD \quad (7)$$

Considering

$$\max_{0 \leq x \leq 7} \left| \cos\left(\frac{(2x+1)\pi}{4}\right) \right| = \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2} \quad (8)$$

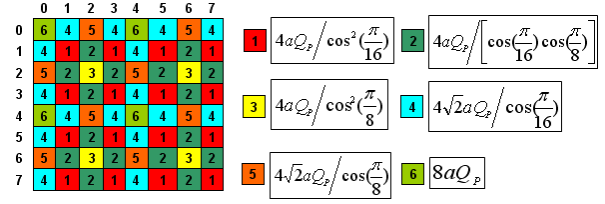


Fig. 1. Thresholds to determine quantized DCT coefficients in (u, v) to be zeros, where label i is that for T_i listed in Table 1.

Table 2. Implementation of DCT, Q, IQ and IDCT.

Type	Condition	Strategy for implementation
Skip	$SAD < T_1$	Not performed
1	$T_1 \leq SAD < T_2$	Performed to 16 coefficients
2	$T_2 \leq SAD < T_3$	Performed to 32 coefficients
3	$T_3 \leq SAD < T_4$	Performed to 36 coefficients
4	$T_4 \leq SAD < T_5$	Performed to 52 coefficients
5	$T_5 \leq SAD < T_6$	Performed to 60 coefficients
8×8	$T_6 \leq SAD$	Performed to all the 64 coefficients

$$\max_{0 \leq y \leq 7} \left| \cos\left(\frac{(2y+1)\pi}{8}\right) \right| = \cos\left(\frac{\pi}{8}\right) \quad (9)$$

Equation (7) is further given as

$$F(4, 2) \leq \frac{\sqrt{2}}{8} \cos\left(\frac{\pi}{8}\right) \times SAD \quad (10)$$

So $F(4, 2)$ can be predicted as zero if

$$SAD < \frac{4\sqrt{2}\alpha Q_p}{\cos(\frac{\pi}{8})} \quad (11)$$

Therefore, we can predict $F(4, 2)$ as zero by comparing SAD with the threshold $T = \frac{4\sqrt{2}\alpha Q_p}{\cos(\frac{\pi}{8})}$. Similarly, other DCT coefficients are bounded depending on the frequency position that affects the maximum values of the two cosine functions. As a result, the thresholds that determine the quantized DCT coefficients to be zero-valued are listed in Table 1, and also illustrated in Fig. 1, where the numeric labels are the threshold labels as listed in Table 1.

3.2. Implementation and Comparison with Other Models

Based on the thresholds in Table 1, we develop an algorithm to perform different types of DCT, Q, IQ and IDCT computations. This algorithm is explicitly shown in Table 2: If $SAD < T_i$, compute all the DCT coefficients labelled by j with $j < i$ in Fig. 1, and $0 < i, j < 7$. The case for $i = 1$ implies the block is an all-zero DCT block, and thus all the DCT, Q, IQ and IDCT computations are skipped. Similarly, the case when $SAD \geq T_6$ would require all the DCT coefficients and the associated quantizations to be computed.

Regarding the DCT/IDCT implementation, we make use of the row-column approach [9] and optimize DCT/IDCT in conformity to the different types shown in Table 2. Since

Table 3. Comparison results of FRR.

Q_p		Foreman (QCIF)	Silent (CIF)	News (QCIF)	Table Tennis (CIF)
7	[6]	89.89%	88.97%	83.63%	86.38%
	[7]	89.40%	88.47%	82.82%	85.63%
	AM	85.35%	84.03%	75.59%	80.30%
14	[6]	76.80%	76.79%	69.70%	70.28%
	[7]	75.78%	75.84%	68.80%	69.55%
	AM	68.96%	68.94%	61.09%	63.12%
21	[6]	65.97%	67.04%	60.07%	59.52%
	[7]	64.73%	65.69%	58.90%	57.59%
	AM	58.08%	56.71%	51.20%	47.64%
28	[6]	58.34%	57.68%	56.11%	45.03%
	[7]	57.09%	56.08%	55.20%	43.01%
	AM	50.64%	46.04%	46.86%	35.72%

we can predict some DCT coefficients as zeros in advance, it is no need to calculate such coefficients and the row-column DCT/IDCT butterfly-flow structure can be optimized accordingly. For Q/IQ, the elements which are predicted as zeros are directly set to zeros to save computations.

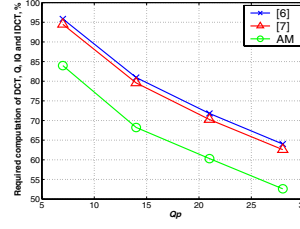
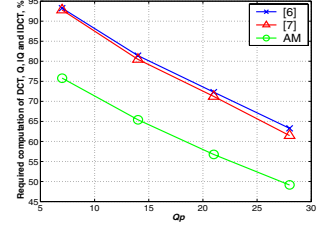
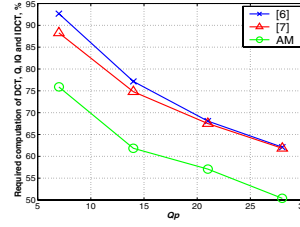
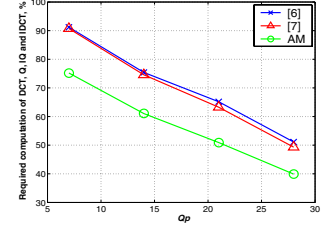
Compared with the models in [6] and [7], which are also based on 8×8 DCT and only consider the case of detecting all-zero DCT blocks, our proposed minimum threshold $T_1 = \frac{4\alpha Q_p}{\cos^2(\frac{\pi}{16})}$ is equal to the threshold proposed in [7], and larger than the threshold $4\alpha Q_p$ in [6]. More importantly, besides the *skip* type (Table 2), the proposed model also considers other prediction types, hence can achieve more computational complexity reductions than the models in [6] and [7].

4. EXPERIMENTAL RESULTS

The XVID codec [10] is implemented for experiments, which is an MPEG-4 compliant video codec. To avoid introducing any biasing factors, the encoder fixes the quantization parameter Q_p during video encoding. We have tested many benchmark video sequences with different resolutions, due to the space limit, herein only present four benchmark video sequences: Foreman and News with QCIF format, Silent and Table Tennis with CIF format. Because the model in [8] is proposed for H.264 codec (4×4 DCT), so in the current experiments, we mainly compare the proposed analytical model with the models in [6] and [7].

4.1. False Acceptance Rate and False Rejection Rate

The false acceptance rate (FAR) and false rejection rate (FRR) are provided to compare the ZQDCT prediction capacity of the proposed analytical model with the models in [6] and [7]. The smaller the FAR is, the less the video quality degrades. The smaller the FRR is, the more efficiently the model can detect ZQDCT coefficients. Therefore, it is more desirable to have small FAR and FRR for an efficient predictive model of ZQDCT coefficients.

**Fig. 2.** Required computation of DCT, Q, IQ and IDCT, *Foreman* (QCIF).**Fig. 3.** Required computation of DCT, Q, IQ and IDCT, *Silent* (CIF).**Fig. 4.** Required computation of DCT, Q, IQ and IDCT, *News* (QCIF).**Fig. 5.** Required computation of DCT, Q, IQ and IDCT, *Table Tennis* (CIF).

From the results, all of the models observe zero FAR, which indicates that there is no false acceptance of ZQDCT coefficients as we expected since all of the three models are derived based on mathematically verified analysis. So we only list the FRR results in Table 3 for Q_p equal to 7, 14, 21 and 28, where AM denotes the proposed analytical model. As evident in Table 3, our analytical model can achieve smaller FRR than those of the models in [6] and [7] indicating that the proposed model is more efficient to predict ZQDCT coefficients and hence reduce more computations related to DCT, Q, IQ and IDCT manipulations.

4.2. Computation Reduction of DCT, Q, IQ and IDCT

To further demonstrate the proposed analytical model can efficiently reduce the computations of video encoding, the comparisons of computational complexity about DCT, Q, IQ and IDCT between the test models and the original MPEG-4 encoder are illustrated in Figs. 2-5. In these figures, the required computational complexity of DCT, Q, IQ and IDCT for the test model is defined as

$$C = \frac{T_d}{T_d^o} \times 100\% \quad (12)$$

where T_d is the encoding time of DCT, Q, IQ and IDCT for the test model, and T_d^o is the encoding time of these four procedures in the original encoder. From these figures, it is obvious that the proposed analytical model can obtain better performance in reducing the computational complexity of DCT, Q, IQ and IDCT than the other models in [6] and [7]. It reveals that the proposed model can effectively eliminate redun-

Table 4. Comparison results of encoding time performance.

Q_p		Foreman (QCIF)	Silent (CIF)	News (QCIF)	Table Tennis (CIF)
7	[6]	1.58%	3.13%	3.55%	3.73%
	[7]	2.57%	3.41%	4.72%	4.15%
	AM	5.87%	10.71%	9.76%	10.84%
14	[6]	7.40%	8.58%	10.59%	11.29%
	[7]	8.74%	8.95%	12.22%	11.91%
	AM	12.70%	14.87%	16.41%	17.18%
21	[6]	13.22%	12.65%	16.06%	16.05%
	[7]	13.90%	12.79%	16.41%	16.59%
	AM	17.28%	18.59%	20.53%	21.72%
28	[6]	15.42%	16.57%	18.44%	22.39%
	[7]	15.89%	17.53%	18.52%	23.16%
	AM	19.08%	21.92%	23.39%	26.90%

dant computations which are impossible to detect in [6] and [7]. In general, for different Q_p values and different video sequences, the average computations of DCT, Q, IQ and IDCT have been reduced by about 40 percent as compared with the original encoder when the proposed model is applied.

4.3. Encoding Time and Video Quality

Finally, we will study the entire encoding time and video quality performances. The video quality is objectively measured by the Peak Signal to Noise Ratio (PSNR). From the experimental results, no PSNR drop is observed for all the three evaluated models when compared with the original encoder. This is because that all of these three models are proposed on mathematical analysis of the quantized DCT coefficients, and the PSNR results are consistent with the FAR results as stated previously. In the following, we will compare the entire encoding time performance about these three models via Table 4, where the encoding time improvement (ΔT , %) is presented as

$$\Delta T = \frac{T_{org} - T}{T_{org}} \times 100\% \quad (13)$$

where T_{org} and T are the entire encoding time of the original MPEG-4 encoder and the test model, respectively.

From Table 4, it can be seen that the real-time performance based on our analytical model is better than the other two models for all the cases. This validates that the proposed model can reduce the computational complexity of video encoding more efficiently, and is more practical to be applied in real-time applications.

5. CONCLUSION AND FUTURE WORKS

In this paper, a novel ZQDCT coefficient prediction model is proposed to skip redundant DCT, Q, IQ and IDCT computations. We have analyzed the ZQDCT coefficients and derived an sufficient condition for each quantized DCT coefficient to be zero. Based on these theoretical analysis, a precise analytical model is proposed. The experimental results demon-

strate that the proposed analytical model can achieve higher encoding efficiency than other theoretically analyzed models without any degradation of video quality.

In the future, we will apply the proposed analytical model to the H.264 [3] encoder after modifications since the 4×4 DCT and combination of DCT and Q are employed in this standard. In addition, from the experiments, we notice that although our analytical model can achieve smaller FRR values than other theoretical models, the FRR results are still noticeable. In order to further improve the ZQDCT coefficients prediction efficiency, some statistical techniques such as [11] may be combined with the proposed analytical model. Thus, we can study the trade-off between the real-time performance and video quality for the potential models.

Acknowledgments

The authors would like to acknowledge the City University of Hong Kong Strategic Grant 7001697 for financial support.

6. REFERENCES

- [1] "MPEG-4 Video Verification Model Version 18.0", *ISO/IEC JTC1/SC29/WG11 N3908*, Pisa, Italy, Jan. 2001.
- [2] "Video Coding for Low Bit Rate Communication", *ITU-T Rec. H.263*, Feb. 1998.
- [3] "Advanced Video Coding for Generic Audiovisual Services", *ITU-T Rec. H.264*, Mar. 2005.
- [4] A. Yu, R. Lee and M. Flynn, "Early Detection of All-Zero Coefficients in H.263", *Proc. Coding Symp.*, pp. 159-164, 1997.
- [5] A. Yu, R. Lee and M. Flynn, "Performance Enhancement of H.263 Encoder Based on Zero Coefficient Prediction", *Proc. the Fifth ACM Int. Multimedia Conf.*, pp. 21-29, Nov. 1997.
- [6] X. Zhou, Z. Yu, and S. Yu, "Method for Detecting All-Zero DCT Coefficients Ahead of Discrete Cosine Transformation and Quantization", *Electro. Lett.*, vol. 34, no. 19, pp. 1839-1840, Sept. 1998.
- [7] L. A. Sousa, "General Method for Eliminating Redundant Computations in Video Coding", *Electron. Lett.*, vol. 36, no. 4, pp. 306-307, Feb. 2000.
- [8] G. Y. Kim, Y. H. Moon, and J. H. Kim, "An Early Detection of All-Zero DCT Blocks in H.264", *Proc. IEEE Int. Conf. Image Processing*, pp. 453-456, Oct. 2004.
- [9] W. Chen, C. H. Smith, and S. C. Fralick, "A Fast Computational Algorithm for The Discrete Consine Transform", *IEEE Trans. Commun.*, vol. COMM-25, pp. 1004-1009, Sept. 1977.
- [10] XVID Team, [Online] Available: <http://www.xvid.org>, 2005.
- [11] 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.