Zero-quantised discrete cosine transform prediction technique for video encoding

H. Wang, S. Kwong, C.-W. Kok and M.-Y. Chan

Abstract: A new analytical model to eliminate redundant discrete cosine transform (DCT) and quantisation (Q) computations in block-based video encoders is proposed. The dynamic ranges of the quantised DCT coefficients are analysed, then a threshold scheme is derived to determine whether the DCT and Q computations can be skipped without video quality degradation. In addition, fast DCT/inverse DCT (IDCT) algorithms are presented to implement the proposed analytical model. The proposed analytical model is compared with other comparable analytical models reported in the literature. Both the theoretical analysis and experimental results demonstrate that the proposed analytical model can greatly reduce the computational complexity of video encoding without any performance degradation and outperforms other analytical 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) and motion compensation (MC), quantisation (Q), inverse Q (IQ) and inverse DCT (IDCT) as the building blocks. Such video encoders are computationally intensive. Recently, there is a significant interest and research in reducing these computations. Previously, the efforts to reduce the computations of video encoding were mainly focused on fast ME algorithms. However, as the ME becomes optimised, we also need to optimise other functions such as the DCT to further speed up video encoding. For example, the 8×8 DCT operating on MPEG2 MP@ML video (720×480 pixels per frame, 30 fps) demands almost 1 Giga operations (equivalent to 16-bit additions) per second. As a result, lowering the DCT complexity is very important, especially for battery-operated applications where the energy budget is tight.

In digital video coding, especially in the very low bit-rate coding, it is quite common that a substantial number of DCT coefficients of the prediction difference are quantised to zeros. Therefore considerable computations may be saved if there is a method to detect early zero-quantised DCT (ZQDCT) coefficients, that is the DCT coefficients equal to zero after Q, before calculating DCT and Q. Chen *et al.* [4] propose to compare the signal energy with a threshold T, and set the DCT output to be a block of zero values if the signal energy is less than the threshold. In a similar manner, Yu *et al.* [5] propose to compare the sum of absolute difference (SAD) available from ME with a predetermined threshold T. If SAD $< T$, then DCT and Q computations can be skipped, and the quantised 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 suitable value is not trivial. In order to reduce the degradation of video quality, Yu *et al.* [6] decrease the threshold value experimentally to detect all-zero DCT blocks. Pao *et al.* [7] propose a Laplacian distribution-based statistical model to reduce redundant DCT and Q computations. The statistical model [7] has shown improvement over previous models to reduce the computational complexity of video encoding. However, the video-quality degradation has been observed due to the trade-off nature of this model.

In a different manner, Zhou *et al.* [8] perform the theoretical analysis on the dynamic range of DCT coefficients and propose a sufficient condition for all the DCT coefficients to be simultaneously quantised to zeros. By checking the motion compensated pixel block against the derived sufficient condition, an early detection of all-zero DCT blocks can be identified. Thus, the redundant DCT and Q computations can be avoided without video-quality degradation. Zhou's [8] model is further refined by Sousa [9] and Kim [10], where tighter sufficient conditions are derived to obtain more reductions in the computational complexity without video-quality degradation. Wang *et al.* [11] follow the model of Sousa [9] and apply it to the H.264 standard. However, those analytical models [8–10] only consider early detection of all-zero DCT blocks before DCT and Q 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 in video encoding without quality degradation through a more theoretical analysis on the dynamic range of DCT coefficients. The proposed analytical model not only detects all-zero DCT blocks, but also predicts individual ZQDCT coefficients. As a result, higher prediction efficiency and more savings in the computational complexity can be achieved by the proposed analytical model. In addition, the proposed analytical model can be applied with fast butterfly-based DCT/IDCT algorithms after our made-to-order optimisation of the row–column-based algorithm [12].

IEE Proceedings online no. 20050317

doi:10.1049/ip-vis:20050317

Paper first received 18th October 2005 and in revised form 7th February 2006

H. Wang and S. Kwong are with the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, People's Republic of China

C.-W. Kok and M.-Y. Chan are with the Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology, Hong Kong, People's Republic of China

E-mail: cssamk@cityu.edu.hk

2 DCT and Q

In this paper, we mainly focus on the 8 × 8 DCT, which is widely used in MPEG-4 [1] and H.263 [2] standards. We define $f(x, y)$, $0 \le x, y \le 7$, as the 8 × 8 motion-compensated pixel block, such that

$$f(x, y) = I(x, y) - I_m(x, y), \quad 0 \le x, y \le 7 \tag{1}$$

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 ME stage to minimise the SAD given by

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

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

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

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

$$F(u, v) < aQ_p \tag{4}$$

where Q_p is the Q parameter which is usually equal to half of the Q step size and ranges from 1 to 31. The parameter a is related to the Q method applied. For example, the Q performed in H.263 and MPEG-4 inter mode follows

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

where $L(u, v)$ are the quantised DCT coefficients. The DCT coefficients $F(u, v)$ are quantised to zeros if

$$L(u, v) < 1 \tag{6}$$

Therefore when $|F(u, v)| < 2.5Q_p$, the coefficients $F(u, v)$ will be quantised to zeros. As a result, a should be chosen as $a = 2.5$ such that the prediction of ZQDCT coefficients in (4) will not result in video-quality degradation.

3 Proposed analytical model for ZQDCT prediction

3.1 Sufficient condition for ZQDCT prediction

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

$$F(u, v) \le \frac{C(u)C(v)}{4} \max_{x,y} \left\{ \left| \cos\left(\frac{(2x+1)u\pi}{16}\right) \right| \times \left| \cos\left(\frac{(2y+1)v\pi}{16}\right) \right| \right\} \times \text{SAD} \tag{7}$$

We start our discussion on the range of $F(u, v)$ by considering the case of $u = v = 0$, such that

$$F(0, 0) \le \frac{1}{8}\text{SAD} \tag{8}$$

As a result, the DC term $F(0, 0)$ will be quantised to zero if

$$F(0, 0) \le \frac{\text{SAD}}{8} < aQ_p \Rightarrow \text{SAD} < 8aQ_p \tag{9}$$

Therefore $F(0, 0)$ can be determined to be quantised to zero by comparing SAD with the threshold $T = 8aQ_p$. Take $u = 4$ and $v = 2$ for another example

$$F(4, 2) \le \frac{1}{4} \max_{x,y} \left\{ \left| \cos\frac{(2x+1)\pi}{4} \right| \left| \cos\frac{(2y+1)\pi}{8} \right| \right\} \times \text{SAD} \tag{10}$$

Considering

$$\max_{0 \le x \le 7} \left| \cos\frac{(2x+1)\pi}{4} \right| = \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2} \tag{11}$$

$$\max_{0 \le y \le 7} \left| \cos\frac{(2y+1)\pi}{8} \right| = \cos\left(\frac{\pi}{8}\right) \tag{12}$$

Equation (10) is further given as

$$F(4, 2) \le \frac{\sqrt{2}}{8} \cos\left(\frac{\pi}{8}\right) \times \text{SAD} \tag{13}$$

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

$$\text{SAD} < \frac{4\sqrt{2}aQ_p}{\cos(\pi/8)} \tag{14}$$

Therefore we can predict $F(4, 2)$ as zero after Q by comparing SAD with the threshold $T = (4\sqrt{2}aQ_p)/\cos(\pi/8)$. Similarly, other DCT coefficients can be bounded depending on the frequency position that affects the maximum values of the two cosine functions. As a result, six thresholds in ascending order for SAD that determine the quantised DCT coefficients to be zero-valued are illustrated in Fig. 1 where the numeric labels in the figure are corresponding to the threshold index.

3.2 Implementation of the proposed analytical model

On the basis of the thresholds in Fig. 1, we develop an algorithm to perform different types of DCT, Q, IQ and IDCT computations. This algorithm is explicitly shown in Table 1. If SAD < T_i, only the DCT coefficients labelled by j with $j < i$ in Fig. 1 are required for computation, the calculation of the other coefficients are skipped. The case of $i = 1$ implies the block is an all-zero DCT block, and thus all the DCT, Q, IQ and IDCT computations are avoided. The case of SAD ≥ T_6 would require all the DCT coefficients and the associated Q to be computed. To illustrate the proposed algorithm clearly, Figs. 2–4 are shown for the reference of the prediction types 1–3,

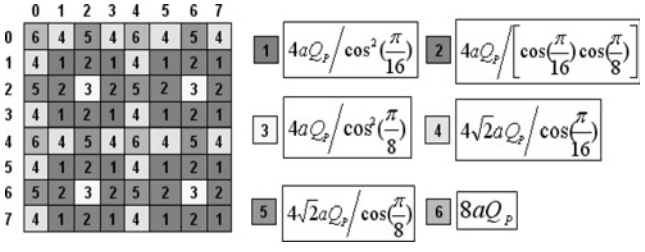

Fig. 1 *Threshold values for SAD to determine the quantised DCT coefficients to be zero-valued, where the cell label i is related to the threshold T_i*

Table 1: Threshold scheme for DCT, Q, IQ and IDCT implementation

Type	Condition	Strategy for implementation
Skip	SAD $< T_1$	not performed
1	$T_1 \leq$ SAD $< T_2$	performed to 16 coefficients with label = 1 in Fig. 1 (also Fig. 2)
2	$T_2 \leq$ SAD $< T_3$	performed to 32 coefficients with label < 3 in Fig. 1 (also Fig. 3)
3	$T_3 \leq$ SAD $< T_4$	performed to 36 coefficients with label < 4 in Fig. 1 (also Fig. 4)
4	$T_4 \leq$ SAD $< T_5$	performed to 52 coefficients with label < 5
5	$T_5 \leq$ SAD $< T_6$	performed to 60 coefficients with label < 6
8 × 8	$T_6 \leq$ SAD	performed to all the 64 coefficients

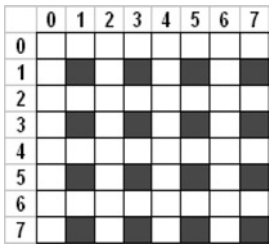

Fig. 2 *Prediction type 1*

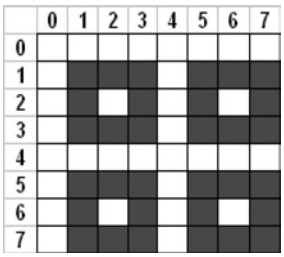

Fig. 3 *Prediction type 2*

where DCT coefficients in blank positions are predicted as ZQDCT coefficients whose calculations are skipped, whereas the shadow ones are non-ZQDCT coefficients which require computations. Regarding DCT/IDCT implementation, we make use of the fast row–column algorithm [12] and optimise DCT/IDCT in conformity to the different prediction types of the proposed analytical model. As we can predict some quantised DCT coefficients as zeros in advance, these is no need to calculate such coefficients and DCT/IDCT can be optimised accordingly.

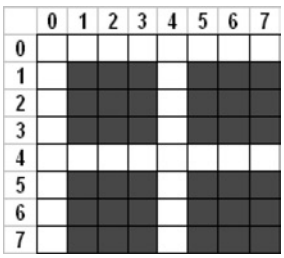

Fig. 4 *Prediction type 3*

In the following, we will take the prediction type 1 of our analytical model as an example to illustrate how to apply fast custom-built DCT algorithm to our analytical model. The fast row–column DCT algorithm [12] can be done by one-dimensional butterfly-based DCT [13] on each row followed by one-dimensional butterfly-based DCT on each column. Fig. 5 shows the one-dimensional traditional butterfly-based DCT flow graph where C_i ($1 \leq i \leq 7$) are cosine factors. From Fig. 5, it can be seen that 26 additions and 16 multiplications are required for each one-dimensional DCT operation. Therefore totally 416 additions and 256 multiplications are required to implement the two-dimensional row–column-based DCT algorithm. As far as the prediction type 1 of our analytical model is concerned, 48 coefficients are predicted as zeros. Thus, we can optimise the fast DCT algorithm [12] accordingly. In the row–column-based fast DCT algorithm, the column transform depends on outputs of the row transform. As the coefficients in the even rows and even columns are predicted as zeros, there is no need to calculate the coefficients in even positions such as Y[0], Y[2], Y[4] and Y[6] in Fig. 5 during the row transform. Similarly, the calculation of coefficients in even positions can be skipped while applying the column transform to the odd columns such as column 1, 3, 5 and 7. In addition, the column transform can be totally skipped when the even columns, column 0, 2, 4 and 6, are considered. The optimised one-dimensional butterfly-based DCT flow graph associated with the prediction type 1 of the proposed analytical model is shown in Fig. 6. Only 168 additions and 120 multiplications are required to implement DCT for the proposed analytical model type 1 after optimising the row–column algorithm [12]. For other prediction types of our analytical model, we can optimise the DCT flow structure accordingly. The computational complexity of DCT for each prediction type of the proposed analytical model against the row–column algorithm [12] is given in Table 2. The IDCT algorithm can be optimised in a similar manner to the optimisation of DCT algorithm. Most of the ideas presented here can also be applied to other fast DCT/IDCT approaches. For Q/IQ, the elements which are predicted as zeros are directly set to zeros to save computations.

4 Experimental results

In this work, the XVID codec [14] is implemented for experiments, which is an MPEG-4 compliant video codec and applies the advanced PMVFAST ME algorithm [15]. Fig. 7 presents the execution time distribution for the original encoder. The percentage distribution is very similar for all test sequences, so only the results for the 'Foreman' sequence are shown here, where we can see that DCT/IDCT/Q/IQ consist of approximately 40% of the total encoding time. Thus, it is important to reduce the computational complexity of these four procedures. To avoid introducing any biasing factors in our experiments, the encoder fixes the Q parameter Q_p during video encoding. Four benchmark video sequences are used for comparisons, each of which is of CIF format (352 × 288) and have 300 frames. All the simulations are running on a PC with Intel Pentium 3.0 GHz CPU and 512 Mbytes of RAM. The software platform is Visual C++ 6.0 and only the C code is applied for experiments. Because the model in the work of Kim *et al.* [10] is proposed for H.264 codec (4 × 4 DCT), in the current experiments, we mainly compare the proposed analytical model with the models in the work of Zhou *et al.* [8] and Sousa [9].

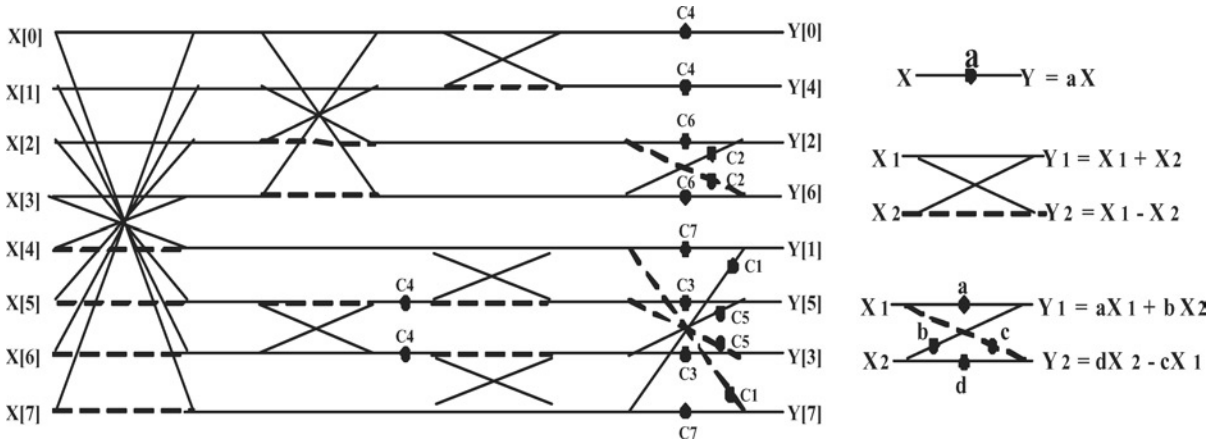

Fig. 5 *One-dimensional traditional fast butterfly-based DCT flow graph [13]*

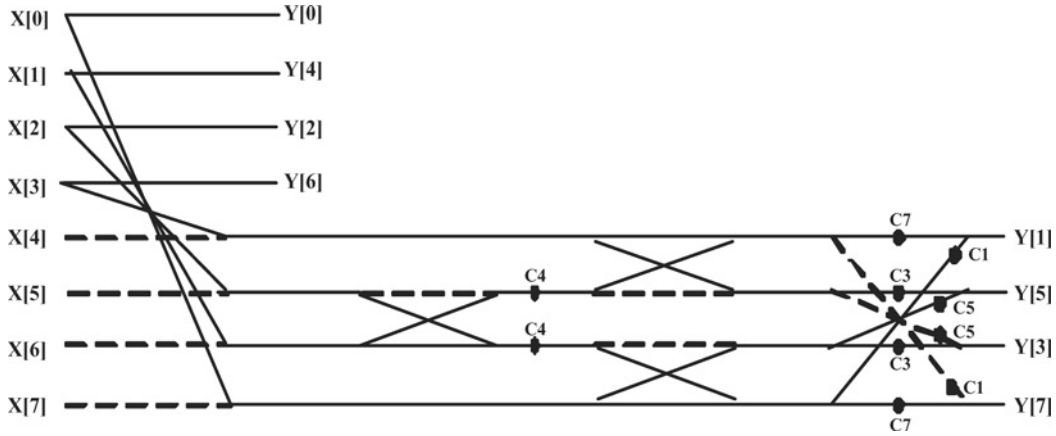

Fig. 6 *One-dimensional optimised fast butterfly-based DCT flow graph, which is applied to row transforms and odd column transforms in prediction type 1 of the proposed analytical model*

Table 2: Computational complexity of DCT used in the proposed analytical model against that of Rao and Yip [12]

	Rao and Yip [12]	Type 1	Type 2	Type 3	Type 4	Type 5
Add (+)	416	168	292	308	384	408
Mul (×)	256	120	188	196	240	252

4.1 Occurrences of different types of the proposed analytical model

In the first place, we study the number of occurrences of different types proposed by our analytical model, the results are shown in Tables 3–6 where Q_p equals to 7, 14, 21 and 28, respectively. The cell value of the tables indicates the ratio of the number of blocks corresponding to the different types according to Table 1 proposed by our analytical model.

From Tables 3–6, it can be seen that each type presented in Table 1 does occur according to our analytical model. Especially, when Q_p becomes larger, the total number of blocks associated with the Skip type and types 1–5 also increases, which implies that more computations can be saved. It indicates that the classification of different types derived from the proposed analytical model is certainly useful for ZQDCT coefficients prediction.

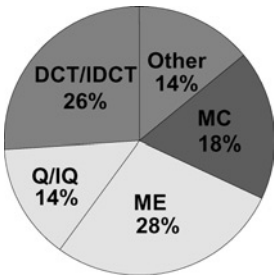

Fig. 7 *Execution time distribution for 'Foreman'*

Table 3: Ratio of blocks (%) against different types, $Q_p = 7$

	Foreman	Silent	News	Table tennis
Skip	14.66	11.17	25.44	13.72
Type 1	1.42	1.26	3.08	1.82
Type 2	1.39	1.24	3.01	1.72
Type 3	5.08	4.71	9.33	5.61
Type 4	2.07	1.93	3.39	1.99
Type 5	10.69	10.79	14.70	8.76
8 × 8	64.69	68.91	41.06	66.38

4.2 False acceptance rate and false rejection rate

The false acceptance rate (FAR) and false rejection rate (FRR) are provided to compare the ZQDCT prediction

Table 4: Ratio of blocks (%) against different types, $Q_p = 14$

	Foreman	Silent	News	Table tennis
Skip	30.76	23.91	45.92	30.13
Type 1	2.24	1.65	2.31	1.27
Type 2	2.50	1.96	2.55	1.48
Type 3	8.62	7.76	9.66	7.35
Type 4	2.88	3.01	2.89	3.24
Type 5	12.36	14.42	9.17	16.36
8×8	40.64	47.29	27.49	40.18

Table 5: Ratio of blocks (%) against different types, $Q_p = 21$

	Foreman	Silent	News	Table tennis
Skip	44.32	34.12	56.81	42.04
Type 1	2.29	2.00	1.72	3.07
Type 2	2.48	2.30	1.95	3.37
Type 3	8.15	10.90	7.16	11.48
Type 4	2.40	3.75	2.42	2.77
Type 5	10.49	16.14	8.98	10.26
8×8	29.87	30.80	20.96	27.00

Table 6: Ratio of blocks (%) against different types, $Q_p = 28$

	Foreman	Silent	News	Table tennis
Skip	53.45	43.75	62.06	56.74
Type 1	2.17	2.84	1.82	2.89
Type 2	2.07	3.13	1.83	2.49
Type 3	7.31	11.10	7.85	7.16
Type 4	2.18	3.44	2.07	1.83
Type 5	10.08	15.54	7.75	8.89
8×8	22.73	20.18	16.62	20.00

capacity of the proposed analytical model with the models in Zhou *et al.* [8] and Sousa [9]. The FAR and FRR are defined as

$$\text{FAR} = \frac{N_{mn}}{N_n} \times 100\%, \quad \text{FRR} = \frac{N_{mz}}{N_z} \times 100\% \qquad (15)$$

where N_{mn} is the number of non-ZQDCT coefficients being miss classified as ZQDCT coefficients, N_n is the total number of non-ZQDCT coefficients, N_{mz} is the number of ZQDCT coefficients being miss classified as non-ZQDCT coefficients and N_z is the total number of ZQDCT coefficients. The smaller the FAR is, the less the video quality will degrade. And 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.

From the results, all of the three models observe zero FAR, which indicates that there is no false acceptance of ZQDCT coefficients as we expected, as all of the three models are derived on the basis of mathematically verified analysis. So we only list the FRR results in Table 7,

Table 7: Comparison results of FRR

Q_p		Foreman, %	Silent, %	News, %	Table tennis, %
7	Zhou *et al.* [8]	85.19	88.97	75.16	86.38
	Sousa [9]	84.59	88.47	73.78	85.63
	AM	79.79	84.03	64.79	80.30
14	Zhou *et al.* [8]	70.05	76.79	54.94	70.28
	Sousa [9]	68.83	75.84	53.66	69.55
	AM	61.02	68.94	45.57	63.12
21	Zhou *et al.* [8]	56.88	67.04	44.12	59.52
	Sousa [9]	55.40	65.69	42.96	57.59
	AM	47.94	56.71	36.69	47.64
28	Zhou *et al.* [8]	47.75	57.68	38.89	45.03
	Sousa [9]	46.36	56.08	37.80	43.01
	AM	39.58	46.04	31.30	35.72

where AM denotes the proposed analytical model. As evident in Table 7, our analytical model can achieve smaller FRR than those of the models in Zhou *et al.* [8] and Sousa [9] meaning that the proposed model is more efficient to predict ZQDCT coefficients and hence reduce more computational complexities related to DCT, Q, IQ and IDCT manipulations.

4.3 Computation reduction of DCT, Q, IQ and IDCT

To further demonstrate, the proposed analytical model can greatly reduce the computational complexity of video encoding, the comparisons of computational complexity about DCT, Q, IQ and IDCT between the test predictive models and the original MEPG-4 encoder are illustrated in Figs. 8–11. In these figures, the required computational complexity of DCT, Q, IQ and IDCT for the test predictive model is defined as

$$C = \frac{T_d}{T_d^o} \times 100\% \qquad (16)$$

where T_d is the encoding time of DCT, Q, IQ and IDCT for the test predictive model, and T_d^o is the encoding time of these four procedures in the original encoder. Note that the execution time of comparing SAD with thresholds to

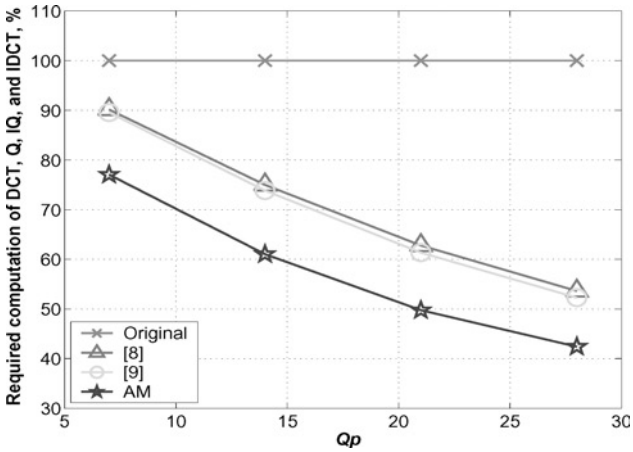

Fig. 8 *Required computation of DCT, Q, IQ and IDCT, Foreman*

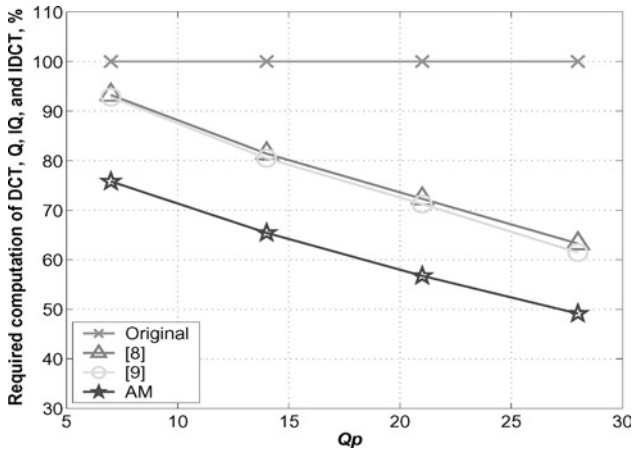

Fig. 9 *Required computation of DCT, Q, IQ and IDCT, Silent*

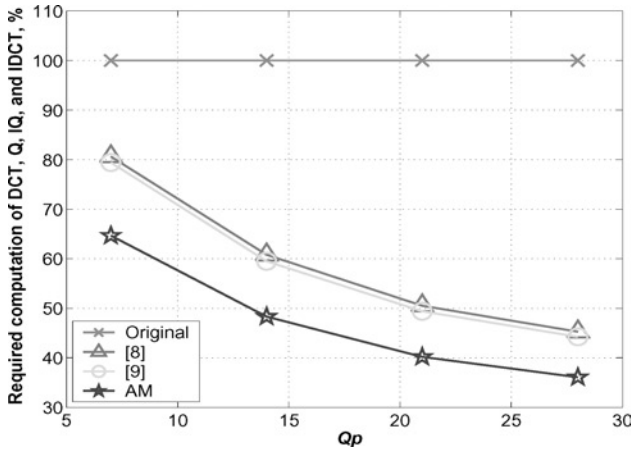

Fig. 10 *Required computation of DCT, Q, IQ and IDCT, News*

calculate the prediction type of the proposed model is already included in T_d. 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 Zhou *et al.* [8] and Sousa [9]. It reveals that the proposed model can effectively eliminate redundant computations which are impossible to detect in [8] and [9]. Compared with the works of Zhou *et al.* [8] and Sousa [9], which only consider the case of detecting all-zero DCT blocks,

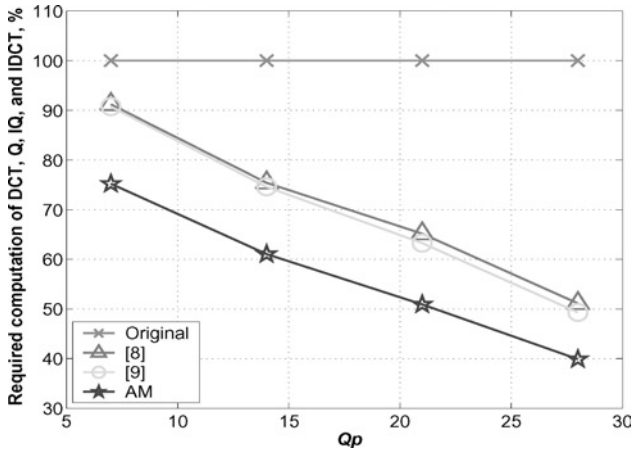

Fig. 11 *Required computation of DCT, Q, IQ and IDCT, Table tennis*

our proposed minimum threshold $T_1 = (4aQ_p)/(\cos^2(\pi/16))$ is equal to the threshold proposed by Sousa [9], and larger than the threshold $4aQ_p$ in Zhou *et al.* [8]. More importantly, besides the *Skip* type (Table 1), the proposed analytical model also considers other prediction types, and hence can achieve more computational complexity reductions than those in Zhou *et al.* [8] and Sousa [9]. In general, for different Q_p values and different video sequences, the average computations of DCT, Q, IQ and IDCT have been decreased by about 40% when compared with the original encoder when the proposed analytical model is applied.

4.4 Encoding time and video quality

Finally, we will study the total encoding time and video-quality performances for the proposed analytical model. 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. This is because that all of these three models are proposed on mathematical analysis of the quantised DCT coefficients, and the PSNR results are consistent with the FAR results as stated previously. In the following, we will compare the encoding time performance about these three models using Table 8, where the encoding time improvement (ΔT, %) is presented as

$$\Delta T = \frac{T_{org} - T}{T_{org}} \times 100\% \tag{17}$$

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

From Table 8, it can be seen that the real-time performance based on our analytical model is better than the

Table 8: Comparison results of encoding time performance

Q_p		Foreman, %	Silent, %	News, %	Table tennis, %
7	T_{org}, s	7.803210	6.998934	6.459169	7.132042
	Zhou *et al.* [8]	4.02	3.13	9.99	3.73
	Sousa [9]	4.29	3.41	10.76	4.15
	AM	9.17	10.71	17.32	10.84
14	T_{org}, s	7.937127	7.286926	6.586075	7.296515
	Zhou *et al.* [8]	10.55	8.58	20.26	11.29
	Sousa [9]	10.92	8.95	20.99	11.91
	AM	15.65	14.87	26.67	17.18
21	T_{org}, s	8.034191	7.462873	6.709746	7.389291
	Zhou *et al.* [8]	15.53	12.65	25.13	16.05
	Sousa [9]	16.48	12.79	25.55	16.59
	AM	20.64	18.59	29.53	21.72
28	T_{org}, s	8.098949	7.604948	6.809327	7.462950
	Zhou *et al.* [8]	19.60	16.57	27.11	22.39
	Sousa [9]	20.06	17.53	27.91	23.16
	AM	23.55	21.92	31.31	26.90

other two models for all the cases. This validates that the proposed analytical model can reduce the computational complexity of video encoding more efficiently and is superior to the other two models.

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 analysed the ZQDCT coefficients, and derived sufficient conditions for each quantised DCT coefficient to be zero. On the basis of these theoretical analysis, a more efficient analytical model is proposed to predict ZQDCT coefficients before DCT and Q. The proposed model can be implemented with fast row–column-based DCT/IDCT algorithms. The experimental results demonstrate that the proposed analytical model can achieve higher encoding efficiency than other theoretically analysed models without any degradation of video quality. In the future, we will apply the proposed analytical model to the H.264 [3] standard after some modifications, as the 4×4 DCT and combination of DCT and Q are adopted in this standard. We definitely believe that the application of the proposed analytical model to H.264 can efficiently predict ZQDCT coefficients when compared with the model in the work of Kim *et al.* [10], as the model [10] only considers the prediction of all-zero DCT blocks without considering individual level prediction.

6 Acknowledgment

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

7 References

1 'MPEG-4 Video Verification Model Version 18.0', Standard ISO/IEC JTC1/SC29/WG11 N3908, January 2001, Pisa, Italy
2 'Video Coding for Low Bit Rate Communication', Standard ITU-T Rec. H.263, February 1998
3 'Advanced Video Coding for Generic Audiovisual Services', Standard ITU-T Rec. H.264, March 2005
4 Chen, H.T., Wu, P.C., Lai, Y.K., and Chen, L.G.: 'A multimedia video conference system: using region base hybrid coding', *IEEE Trans. Consum. Electron.*, 1996, **42**, pp. 781–786
5 Yu, A., Lee, R., and Flynn, M.: 'Early detection of all-zero coefficients in H.263'. Proc. Coding Symp., 1997, pp. 159–164
6 Yu, A., Lee, R., and Flynn, M.: 'Performance enhancement of H.263 encoder based on zero coefficient prediction'. Proc. 5th ACM Int. Multimedia Conf., November 1997, pp. 21–29
7 Pao, I.M., and Sun, M.T.: 'Modeling DCT coefficients for fast video encoding', *IEEE Trans. Circuits Syst. Video Technol.*, 1999, **9**, (4), pp. 608–616
8 Zhou, X., Yu, Z., and Yu, S.: 'Method for detecting all-zero DCT coefficients ahead of discrete cosine transformation and quantization', *Electron. Lett.*, 1998, **34**, (19), pp. 1839–1840
9 Sousa, L.A.: 'General method for eleminating redundant computations in video coding', *Electron. Lett.*, 2000, **36**, (4), pp. 306–307
10 Kim, G.Y., Moon, Y.H., and Kim, J.H.: 'An early detection of all-zero DCT blocks in H.264'. Proc. IEEE Int. Conf. Image Processing, October 2004, pp. 453–456
11 Wang, Y., Zhou, Y., and Yang, H.: 'Early detection method of all-zero integer transform coefficients', *IEEE Trans. Consum. Electron.*, 2004, **50**, (3), pp. 923–928
12 Rao, K.R., and Yip, P.: 'Discrete cosine transform: algorithms, advantages, applications' (Academic, New York, 1990)
13 Chen, W., Smith, C.H., and Fralick, S.C.: 'A fast computational algorithm for the discrete cosine transform', *IEEE Trans. Commun.*, 1977, **COM-25**, (9), pp. 1004–1009
14 XVID Team. Available online at http://www.xvid.org, 2005
15 Tourapis, A.M., Au, O.C., and Liou, M.L.: 'Predictive motion vector field adaptive search technique (PMVFAST) – enhancing block based motion estimation'. Proc. SPIE Conf. on VCIP'01, January 2001, pp. 883–892