

PREDICTION OF DISCRETE COSINE TRANSFORMED COEFFICIENTS IN RESIZED PIXEL BLOCKS

Jin Li¹ Weiwei Chen¹ Moncef Gabbouj¹ Jarmo Takala¹ Hexin Chen²

¹Tampere University of Technology
Tampere, Finland

¹FirstName.LastName@tut.fi

²Jilin University
Changchun, China

²chx@jlu.edu.cn

Abstract— A hybrid model was developed to predict the zero-quantized discrete cosine transform (ZQDCT) coefficients for intra blocks in our previous work. However, the complicated overhead computations seriously degrade its performance in complexity reduction. This paper proposes a new prediction algorithm with less overhead operations. First, each $N \times N$ pixel block at the input of transform is resized to a $N/2 \times N/2$ block. Then, this downsized block is decomposed into a mean value and a residual pixel block. Finally, the $N \times N$ 2-D DCT is computed from the mean value and this residual pixel block. Experimental results show that the proposed method reduces more redundant computations than the competing techniques and better real-time performance can be expected. This is particularly suitable for low-power processors with video sequences or great number of images to encode.

1. INTRODUCTION

Visual communication via mobile devices is becoming more and more popular. However, current video coding standards are not initially designed for this type of application area. Traditionally, high compression efficiency is usually achieved at the expense of increasing computations. As the newest standard, H.264/AVC [1] significantly outperforms the others in terms of coding efficiency, the computational complexity is however greatly increased. The high complexity limits its application on mobile devices with low computational power and battery constraint as well as the real-time performance of the encoder. Currently, the emerging standard H.265 [2] claims an aggressive goal to reduce the encoding complexity by 50% with the 25% better visual quality compared to H.264.

When reducing redundant computations due to the ZQDCT coefficients, two types of prediction methods are considered: the prediction for the ZQDCT coefficients of the residual pixel blocks and the prediction of intra blocks. A lot of efforts have been done regarding the residual pixel blocks. In [3], Pao *et al.* proposed a Laplacian distribution based model for prediction of ZQDCT coefficients. Wang *et al.* [4] proposed new rate-quantization models for H.264 to detect the all-zero-quantized DCT blocks. In 2005, an improved detection algorithm [5] was proposed to reduce the complexity of H.264. Using the same implementation method as [3] comparable results are achieved by the Gaussian based thresholds [6] based on XVID codec [7]. When these prediction methods were applied to H.264 in [8] and [9], good computational savings were also obtained. In 2007, the Gaussian distribution based model was optimized in [10] to further reduce the redundant computations for inter transform and quantization.

Reducing the redundant calculations for intra DCT and quantization has not been studied actively. Nishida proposed a zero-value prediction for fast DCT calculation in [11]. Although it reduces the total computations of DCT by 29% and quantization by 59% when applied to MPEG-2, the video quality is degraded by 1.6 dB on the average. A sufficient condition based ZQDCT prediction method [12] is proposed for intra DCT to speed up the encoding process

without video quality degradation. However, the prediction efficiency need to be further improved.

In our previous work [13], we extend Pao's [3] and Wang's [6] results to intra DCT and quantization to simplify the complexity with minimal video quality degradation. Experiments show that this hybrid prediction method can significantly reduce the multiplications of transform and quantization. In addition, it performs better particularly at low bitrates.

However, the overall complexity reduction is still seriously degraded due to its complicated overhead operations. This drawback is particularly evident for addition operations. For instance, at high bitrates the required number of additions is even increased by 20% compared to the original method. Therefore, it is of great importance to simplify the overheads for further complexity reduction.

To solve this problem, a new prediction algorithm is proposed in this paper. First, each $N \times N$ pixel block is resized to a $N/2 \times N/2$ block. Then, this downsized block is decomposed into a mean value and a residual pixel block. Finally, the $N \times N$ 2-D DCT is computed from the mean value and the residual block. Since the overheads are completed in a smaller size of block, the required operations is significantly decreased. As a result, high prediction efficiency and good computational savings are achieved by the proposed model.

The rest of the paper is organized as follows. A new prediction algorithm in a resized pixel block is mathematically analysed and proposed in Section 2. The experimental results and discussions are presented in Section 3. Finally, Section 4 concludes the paper.

2. PROPOSED PREDICTION FOR $N \times N$ 2-D DCT

2.1 Mathematical Analysis of 2-D DCT

The 2-D DCT of a $N \times N$ data block f is defined as

$$F(u, v) = \frac{2}{N} c(u)c(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \quad (1)$$

where $0 \leq x, y < N$, $0 \leq u, v < N$, $c(i) = 1/\sqrt{2}$, for $i = 0$, and $c(i) = 1$, otherwise. The values $F(u, v)$, $0 \leq u, v < N$, are the DCT coefficients. Practically, N is usually selected as powers of 2.

If \bar{f} is the mean value of the $N \times N$ data block and $f'(x, y)$ is the residual pixel, we define

$$\bar{f} = \frac{1}{NN} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y), \quad f'(x, y) = f(x, y) - \bar{f} \quad (2)$$

Then, each DCT coefficient can be respectively computed by the mean value \bar{f} and the residual value $f'(x, y)$ as (3). The deduction process is similar to (10) in [13].

TABLE I REQUIRED NUMBER OF OVERHEADS

	[13]	DM	PM
ADD	$4N^2 - 2$	$2N^2 - 1$	$N^2/2 - 1$
MUL	$2N$	2	2
COM	$[1, (N - 1)^2]$	1	1

From (2) and (3), the DC coefficient is only relevant to the mean value \bar{f} and the AC coefficients can be directly calculated from the residual pixels $f'(x, y)$. Therefore, if we can efficiently predict the ZQDCT coefficients for the AC coefficients without the real transform and quantization, lots of computations will be saved. More details about the prediction algorithms can be found in [13].

2.2 Brief Review of Gaussian Distribution Based Modelling

The experiments show that the distribution of the residual pixel $f'(x, y)$ after above decomposition for $f(x, y)$ can be well modeled by a Gaussian distribution with a significant peak at zero. In the following, a brief description of the approach [6] is given.

If SAD is defined as the sum of absolute difference in a $N \times N$ residual block as

$$SAD = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |f'(x, y)| \quad (5)$$

Since the residual pixel $f'(x, y)$ is approximated by a Gaussian distribution with zero mean and the variance, σ , we obtain

$$\sigma \approx \sqrt{\frac{\pi SAD}{2 N^2}} \quad (6)$$

As the residual pixels have a zero mean value, the variance of the (u, v) th DCT coefficient can be expressed as

$$\sigma_F^2(u, v) = \sigma^2 [ARA^T]_{u,u} [ARA^T]_{v,v} \quad (7)$$

where A is the matrix in (2) and $[\cdot]_{u,u}$ is the (u, u) th component of a matrix. The matrix R is defined as [14]

$$R = \begin{bmatrix} 1 & \rho & \dots & \rho^{N-1} \\ \rho & 1 & \dots & \rho^{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \dots & 1 \end{bmatrix}$$

and in this work, the parameter ρ is selected according to [3], [6] and [13] as $\rho = 0.6$.

By the central limit theorem, the DCT coefficient $F(u, v)$ will be quantized to zero with a probability controlled by the confidence parameter γ as

$$\gamma \sigma_F(u, v) < \alpha Q_p \quad (8)$$

where $u, v = 0, 1, \dots, N - 1$ and Q_p is the quantization parameter for the DCT coefficients of the residual pixel blocks.

Derived from (5), (6) and (7), a criterion for the ZQDCT

$$F(u, v) = \begin{cases} \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) & \text{for } u = 0, v = 0 \\ \frac{2c(u)c(v)}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f'(x, y) \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} & \text{otherwise} \end{cases} \quad (3)$$

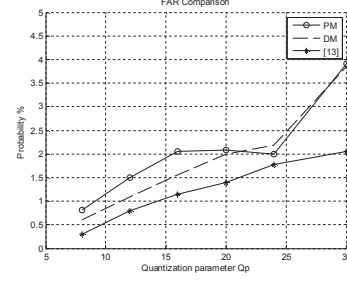


Fig.1 FAR comparison among the proposed method (PM) and the direct method (DM) and [13] based on eight 4CIF sequences

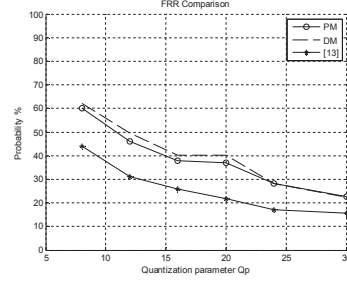


Fig.2 FRR comparison between the proposed method (PM) and the direct method (DM) and [13] based on eight 4CIF sequences

coefficient with high probability [6] is

$$SAD < \beta_G(u, v) \alpha Q_p(u, v) \quad (9)$$

where

$$\beta_G(u, v) = \frac{\sqrt{2} N^2}{\gamma \sqrt{\pi [ARA^T]_{u,u} [ARA^T]_{v,v}}} \quad (10)$$

Based on the above analysis, the Gaussian distribution based model with multiple thresholds is developed to reduce the inter DCT and quantization computations. For $N = 8$, the Gaussian distribution based thresholds β_G can be found in Table I in [6]. If $SAD < \beta_G(0,1) \alpha Q_p(0,1)$, the residual pixel block $f'(x, y)$ will be predicted as an all-zero-quantized DCT block. Therefore, only the DC coefficients need to be calculated and all the AC coefficients will be directly set to zeros without transform and quantization.

2.3 Overhead computations in resized pixel blocks

However, directly calculation of SAD consumes lots of overhead operations, which consist of the decomposition of pixel block f and the computations regarding the residual pixel block f' . Therefore, it is of great importance to simplify the overhead computations. In this paper, this is achieved by operating the overheads in a downsized block.

Firstly, each $N \times N$ pixel block f is downsized into a new $N/2 \times N/2$ pixel block p , which is defined as

$$p(x, y) = f(2x, 2y) \quad (11)$$

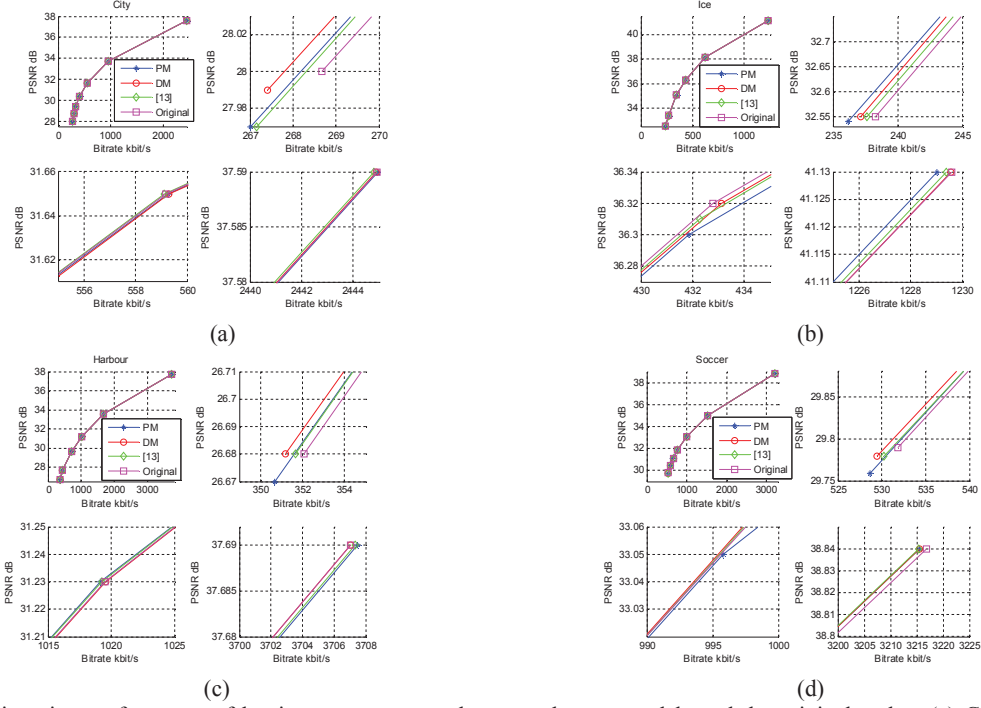


Fig. 3 Rate-Distortion performance of luminance component between the test models and the original codec, (a) City, (b) Ice, (c) Harbour and (d) Soccer. Results for each sequence consist of an overall R-D performance and three locally magnified comparisons.

where $0 \leq x, y < N/2$.

Then, this new pixel block $p(x, y)$ is decomposed into a mean value \bar{p} and a $N/2 \times N/2$ residual pixel block $p'(x, y)$ in accordance with (2).

Since this new block \mathbf{p} is resized from \mathbf{f} , the sum of absolute difference SAD^p of the residual pixel block \mathbf{p}' can be approximated as

$$SAD^p \approx SAD/4 \quad (12)$$

Combined with (9) the prediction algorithm for the ZQDCT coefficient in a $N \times N$ pixel block can be modified as

$$SAD^p < \beta_G(u, v)\alpha Q_p(u, v)/4 \quad (13)$$

Given the above analysis, we propose a model to predict the ZQDCT coefficients for intra DCT and quantization by utilizing the Gaussian distribution based prediction and simplified overhead computations in downsized pixel.

- The $N \times N$ pixel block \mathbf{f} is first resized to a $N/2 \times N/2$ pixel block \mathbf{p} as (11);
- SAD^p is calculated in the new downsized block as (2);
- The coefficient $F(0,0)$ in the original pixel block \mathbf{f} is calculated as (3);
- If $SAD^p < \beta_G(0,1)\alpha Q_p(0,1)/4$, all the AC coefficients in the $N \times N$ pixel block \mathbf{f} are set to zeros without DCT and quantization. Otherwise, do the transform and quantization as normal.

By computing the overheads in resized pixel blocks, the required number of additional operations is significantly decreased compared to the direct calculation in original blocks and the reference [13]. For a $N \times N$ pixel block \mathbf{f} , the comparisons of required number of

overheads are shown in Table I with different methods, i.e. direct overhead calculation from the original $N \times N$ pixel block (DM), the proposed method (PM) and [13]. Take $N = 8$ for instance, the proposed method requires 34 operations. However, totally 130 operations are introduced if the overheads are calculated in the original pixel block. As for [13], the number of overhead is between 271 to 319, which is much more complicated than the proposed method

3. EXPERIMENTAL RESULTS AND DISCUSSION

In order to evaluate the performance of the proposed model, a series of experiments were carried out, using C-code, based on XVID codec against the competing methods. We take $N = 8$ throughout the experiments. Eight sequences are tested to verify the proposed algorithm with various quantization Q_p . Four sequences (City, Ice, Harbour and Soccer) are used in the report for discussions and the codec was compiled with Microsoft Visual Studio 2008.

3.1 Performance analysis based on individual coefficients

The falsely acceptance ratio (FAR) and the falsely rejection ratio (FRR) are compared among the proposed algorithm, the direct method and [13]. The definition of FAR and FRR can be referred to (28) in [13]. The average results based on eight sequences are shown in Fig.1 and Fig.2, respectively.

Experimental results show that the prediction in the downsized blocks achieves very close results as predicted in the original blocks. In addition, [13] achieves the best results than the proposed method and the direct method in terms of both FAR and FRR.

However, it is unfair to determine the performance of the test models only by comparing FAR and FRR or FAR versus FRR. Since the DCT coefficients are not computed individually, the complexity reduction is also related to the implementation structure. On the other hand, prediction errors have a negative effect on the video quality but

positive on the volume of bitrates. In addition, the magnitudes of these errors also affect the PSNR performance at the decoding side.

3.2 Comparisons of video quality

The rate-distortion (R-D) performance is also studied in the experiments. Fig. 3 shows the comparisons between the test models and the original codec. According to the results, all the three test models achieve comparable R-D performance and the resulted quality degradations are negligible compared to the original codec. In addition, based on the results in the magnified local areas it is hard to tell which model performs best in terms of video quality and each model is not necessarily better or worse than the others. The main reason is that prediction errors could cause PSNR degradation at the decoding side. However, it also generates more zero valued coefficients, which in turn results in a lower volume of bitrate. Experimental results also show that the variations of R-D performance are like random values, i.e. noise.

3.3 Comparison of operation number

Finally, the complexity reduction in terms of operation number is measured. The default 2-D DCT in the XVID codec requires 12 multiplications and 32 additions for each 1-D transform [7]. All overheads in the test models are taken into account for measurement. If an $N \times N$ intra block is predicted as an all-zero-quantized DCT block by the proposed algorithm, the DC term is still required to be calculated as (3). Therefore, the additions of $(N^2 - 1)$ and one multiplication are still performed. For DM, this can directly be calculated from (2) and only a scaling operation is introduced. The complexity reduction is defined as

$$OP = \frac{OP_{rdct+q} + OP_{overhead}}{OP_{dct+q}} \times 100\% \quad (14)$$

where OP_{rdct+q} denotes the required operation number of DCT and quantization in the test models, OP_{dct+q} is the total operation number of DCT and quantization in the original codec, and $OP_{overhead}$ is the overheads in the test models. The results are shown in Fig. 4.

According to the results, the proposed method outperforms the others in terms of operation numbers. Although [13] achieves the best reduction on transform and quantization only, it practically consumes more operations due to the complicated overhead computations. In addition, experiments show that the proposed method has a similar performance concerned the complexity reduction of transform and quantization only. Thus, its advantage also mainly exists in the overhead operations.

4. CONCLUSION

This paper proposes an improved prediction method to reduce the complexity of DCT and quantization. Compared to the competing techniques, the proposed model can recognize the all-zero-quantized DCT blocks with less overhead computations. Experiments show that the proposed model significantly reduces the redundant computations. The contribution of the proposed algorithm mainly exists in its simplified overhead computation. In addition, the resulted video quality is comparable to the references. The proposed method is particularly valuable for low-power processors.

5. REFERENCE

[1] H. Kalva, "The H.264 Video Coding Standard," *IEEE Trans. Multimedia*, Vol. 13, No. 4, pp. 86-90, Oct. 2006.
 [2] [online] <http://www.h265.net/>
 [3] I. M. Pao and M. T. Sun, "Modeling DCT Coefficients for Fast Video Encoding," *IEEE Trans. Circuits Systems Video Technology*, Vol. 9, No.4, pp. 608-616, Jun. 1999.

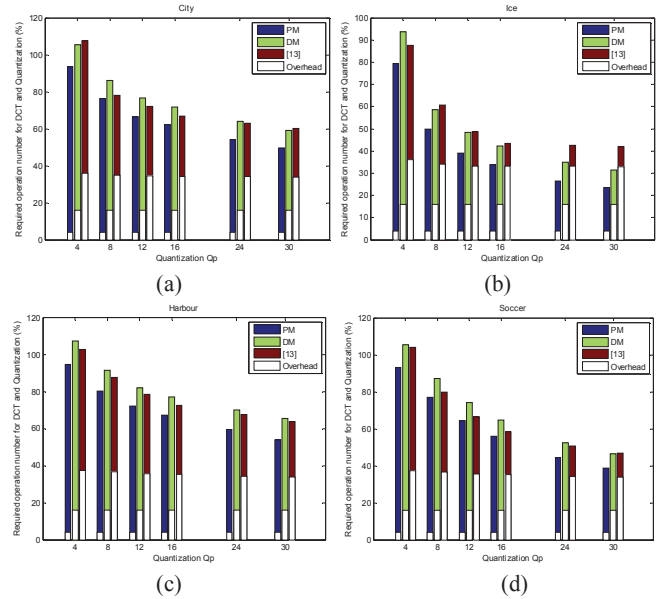


Fig.4 Complexity reduction of operation number. Each bar is composed of two parts: the "WHITE" area means the overhead operations and the upper part only considers the DCT and quantization.

[4] H. Wang, S. Kwong and C. W. Kok, "Early Detection Method of All-zero DCT Blocks in H.264," *IEEE Int. Conf. Image Processing*, pp. 453-456, Singapore, Oct. 2004.
 [5] Y. H. Moon, G. Y. Kim, and J. H. Kim, "An Improved Early Detection Algorithm for All-Zero Blocks in H.264 Video Coding," *IEEE Trans. Circuits Systems Video Technol.*, Vol. 15, No.8, pp. 1053-1057, 2005.
 [6] H. Wang, S. Kwong and C. Kok, "Fast Video Coding Based on Gaussian Model of DCT Coefficients," *IEEE ISCAS*, pp. 1073-1076, Island of Kos, Greece, May 2006.
 [7] [online] <http://www.xvid.org/>
 [8] H. Wang, S. Kwong and C. Kok, "Efficient Prediction Algorithm of Integer DCT Coefficients for H.264/AVC Optimization," *IEEE Trans. Circuits Systems Video Technol.*, Vol. 16, No.4, pp. 547-552, Apr. 2006.
 [9] H. Wang and S. Kwong, "Hybrid Model to Detect Zero Quantized DCT Coefficients in H.264," *IEEE Transactions on Multimedia*, Vol. 9, No. 4, pp. 728-735, Jun. 2007.
 [10] H. Wang, S. Kwong and C. W. Kok, "Efficient Predictive Model of Zero Quantized DCT Coefficients for Fast Video Encoding," *Image and Vision Computing*, Volume 25, Issue 6, pp. 922-933, 2007.
 [11] Y. Nishida, K. Inoue and V. G. Moshnyaga, "A Zero-Value Prediction Technique for Fast DCT Computation," *IEEE Workshop on Sig. Proc. Syst.*, pp. 165-170, Aug. 2003.
 [12] J. Li, J. Takala, M. Gabbouj and H. Chen, "A Detection Algorithm for Zero Quantized DCT Coefficients in JPEG," *IEEE Int. Conf. Acoustic, Speech and Signal Processing*, pp. 1189-1192, Las Vegas, USA, Mar.30-Apr.4, 2008.
 [13] J. Li, M. Gabbouj and J. Takala, "Hybrid Modeling of Intra DCT Coefficients for Real-Time Video Encoding," *EURASIP Journal on Image and Video Processing*, Volume 2008 (2008), Article ID 749172, 13 pages, doi:10.1155/2008/749172.
 [14] A. K. Jain, "Fundamentals of Digital Image Processing," Englewood Cliffs, NJ: Prentice Hall, 1989.