

BIT RATE ESTIMATION FOR COST FUNCTION OF 4x4 INTRA MODE DECISION OF H.264/AVC

Mohammed Golam Sarwer, and Lai Man Po

Department of Electronic Engineering, City University of Hong Kong
Kowloon, Hong Kong SAR, China
sarwer@ee.cityu.edu.hk, eelmpo@cityu.edu.hk

ABSTRACT

H.264/AVC is a newest international video coding standard that can achieve considerably higher coding efficiency than previous standards. This comes at the cost of the complex mode decision procedure using the rate-distortion optimization, which makes real-time encoding difficult. To reduce the complexity of rate-distortion cost, we propose a bit rate estimation technique to avoid the entropy coding method during mode decision of intra prediction. The estimation method is based on the properties of context-based variable length coding (CAVLC). Simulation results demonstrate that the proposed estimation method achieves up to 53 % reduced encoding time of intra coding with ignorable degradation of coding performance.

1. INTRODUCTION

The appearance and development of various new multimedia services have need for higher coding efficiency. The ITU-T/ISO/IEC Joint Video Team established the newest video coding standard known as H.264/AVC [1]. H.264/AVC offers a significant performance improvement over previous video coding standards such as H.263++ and MPEG-4 [2,3]. New and advanced techniques are introduced in this new standard, such as intra prediction for I-frame encoding, multi-frames inter prediction, small block-size transform coding, context-adaptive arithmetic entropy coding, de-blocking filtering, etc. These advanced techniques make this new standard provides approximately 50% bit rate saving for equivalent perceptual quality relative to the performance of prior standards.

To achieve the highest coding efficiency, H.264/AVC uses rate-distortion optimization (RDO) technique to get the best coding result in terms of maximizing coding quality and minimizing bit rates. This means that the encoder has to code the video by exhaustively trying all the mode combinations including the different intra and inter prediction modes. Therefore, the complexity and computation load of video coding in H.264/AVC increase drastically compared to any previous standards.

To reduce computational complexity of H.264/AVC, a number of efforts have been made to explore the fast algorithm in motion estimation, intra mode prediction and inter mode prediction for H.264/AVC video coding [4-6]. In 4x4 intra mode decision, 9 prediction modes are used. To reduce the complexity, H.264/AVC reference software suggested [7] sum of absolute difference (SAD) and sum of absolute transform difference (SATD) based cost functions. These two cost functions reduce computation significantly but performance of rate-distortion characteristics is not good enough. To improve the rate-distortion performance, a new cost function for intra 4x4 mode decision is proposed in [8]. In this cost function, sum of absolute integer transform difference (SAITD) is used in distortion part and a rate prediction algorithm is used in rate part. The major drawback of this cost is that the bit estimation method can not give the very good estimation.

In this paper, we propose a shortcut way to get the number of entropy coded bits as soon as the transform coefficients are quantized. A method for estimation of rate for cost function of 4x4 intra mode decision is proposed. This method is based on the properties of context-based variable length coding (CAVLC) and observation of VLC tables.

2. COST FUNCTION FOR INTRA 4X4 MODE DECISION OF H.264/AVC

To get a richer set of intra-prediction patterns, H.264/AVC offers 9 prediction modes for 4x4 luma blocks and 4 prediction modes for 16x16 luma blocks. The best mode is the one having minimum rate-distortion cost and this cost is expressed as

$$J_{RD} = SSD + \lambda \cdot R \quad (1)$$

where, the SSD is the sum of squared difference between the original block and the reconstructed block, and R is the true bits needed to encode the block and λ is an exponential function of the quantization parameter (QP).

In order to compute RD cost for each mode, same operation of forward and inverse transform/quantization and variable length coding is repetitively performed. All of these

processing explains the high complexity of RD cost calculation.

To reduce the computational complexity, H.264/AVC provides another SAD based cost function:

$$J_{SAD} = SAD + \lambda_1 \cdot 4P \quad (2)$$

where, SAD is sum of absolute difference between the original block and the predicted block. The λ_1 is also approximate exponential function of the QP which is almost the square of λ , and the P equal to 0 for the most probable mode and 1 for the other modes. This SAD-based cost function could save a lot of computations as the distortion part is based on the differences between the original block and the predicted block instead of the reconstructed block. In addition, the rate part is pre-defined by constants either equal 4 or 0. Thus, the variable length coding using CAVLC or CABAC can also be saved. However, the expense of the computation reduction usually comes with quite significant degradation of coding efficiency.

To achieve better rate-distortion performance, JM reference software also provided an alternative SATD-based cost function:

$$J_{SATD} = SATD + \lambda_1 \cdot 4P \quad (3)$$

where, SATD is sum of absolute Hadamard-transformed difference between the original block and the predicted block. Experimental results show that the J_{SATD} could achieve better rate-distortion performance than the J_{SAD} , however, the overall rate-distortion performance is still lower than the optimized J_{RD} .

An enhanced cost function for H.264/AVC 4x4 intra decision is proposed with use of an integer transformed difference for the distortion part and a rate-predictor for the rate part [8]. In this cost function, rate is estimated as follows

$$R_e = 4T_c - T_o + 4P \quad (4)$$

where T_c is the number of non-zero coefficients and T_o is the number of +/- trailing ones of transformed residual block.

Experimental results show that this enhanced cost function could achieve better rate-distortion performance than both J_{SAD} and J_{SATD} . The major drawback of the J_{SATD} is that it requires performing the true integer transform. Even though fast transformation algorithm was proposed to perform SATD but the overall complexity of J_{SATD} is still quite high. It is also shown that in this cost function the bit estimation method cannot give the very good estimation. In this paper a rate prediction scheme is introduced based on the properties of CAVLC entropy coding method. Before we propose the new estimation method, review of the CAVLC encoding method is described at following section.

3. REVIEW OF CONTEXT BASED ADAPTIVE VARIABLE LENGTH CODING (CAVLC)

CAVLC is the method used to encode residual, zig-zag ordered blocks of quantized transform coefficients. The block is encoded by five syntax elements. These elements are described as follows:

Coeff_token: The first VLC, coeff_token, encodes both the total number of non-zero coefficients (T_c) and the number of trailing +/- 1 (T_o) value. Four VLC tables are used for the coeff_token syntax elements. The choice of table depends on the parameter N which is calculated from the number of non-zero coefficients in upper and left handed previously coded blocks [9].

Sign of trailing +/- 1: For each T_o (trailing +/-1), a single bit encodes the sign (0=+, 1=-).

Level: The level (sign and magnitude) of all remaining non-zero coefficients in the block is encoded in reverse order, starting with the highest frequency and working back towards the DC coefficients. Seven level_VLC tables are used to encode the level. The choice of table is adapted in a way described in reference [9].

Total_zeros: Total_zero is the total number of zero coefficients between the DC and last non-zero coefficients.

Run_before: The number of zero coefficients preceding each non-zero coefficients in reverse order.

For example, if the coefficients in the zig-zag order of a 4x4 block are [0,3,0,1,-1,-1,0,1,0,0...]. Here total number of non-zero coefficients (T_c) is 5, total number of zero before the last non-zero coefficients (T_z) is 3, number of trailing one's (T_o) is 3 in fact there are 4 trailing ones but only 3 can be encoded as a "special case".

4. PROPOSED RATE ESTIMATION

From the observation of VLC tables [10], it is clearly shown that

- Total number of bits to encode a block is increasing with number of non-zero coefficients (T_c) and decreasing with number of trailing +/- 1 value (T_o).
- Total number of bits is increasing with the absolute value of level (non-zero coefficients except the trailing ones).
- Total number of bits is also influenced with total number of zero coefficients before the last non-zero coefficients (T_z).

Based on this idea, the proposed rate estimator for cost function of 4x4 intra mode decision of H.264/AVC is defined as follows:

$$R_{est} = aT_c - bT_o + cSAT_L + dT_z + 4P \quad (5)$$

where, SAT_L is the sum of absolute values of all levels of quantized transform residual block, T_z is the total number of zeros before the last non zero coefficients. T_c , T_o and P are same as (4). In addition, a, b, c, d are the positive

constants such that $a > b$. By experiments, we find the better coding performance can be achieved if we choose $a = 3$ and $b = c = d = 1$.

5. EXPERIMENTAL RESULTS

To verify the proposed technique, JM 8.3 [7] reference software is used in simulation. Three well-known video sequences are used as test materials. It should be noted that the test sequences are representative in the sense that they range from sequences of low motion activity such as “Akiyo” to sequences with very high motion such as “Stefan”. Each sequence is encoded with GOP structure of only I frames and only 4x4 intra mode is enabled. The test conditions are as follows:

- a) Hadamard transform is used
- b) RD optimization is enabled
- c) CAVLC is enabled
- d) The number of frame in a sequence is 100
- e) Frame rate is 30

Comparison results were produced based on the percentage of difference of coding time (ΔT %), the PSNR difference between original JM encoder and encoder with proposed rate estimation (ΔP_{snr}) and percentage of the bit rate difference (ΔBit %). ΔBit % is calculated as follows:

$$\Delta Bit \% = \frac{Bit_rate_{original} - Bit_rate_{proposed}}{Bit_rate_{original}} \times 100 \quad (6)$$

where $Bit_rate_{original}$ and $Bit_rate_{proposed}$ are the total bit rate of original JM encoder and encoder with proposed rate estimator, respectively.

In order to evaluate complexity reduction, ΔT (%) is defined as follows

$$\Delta T = \frac{T_{original} - T_{proposed}}{T_{original}} \times 100\% \quad (7)$$

where, $T_{original}$ denotes the total encoding time of the JM 8.3 encoder with rate distortion optimization and $T_{proposed}$ is the total encoding time with proposed fast rate estimation technique.

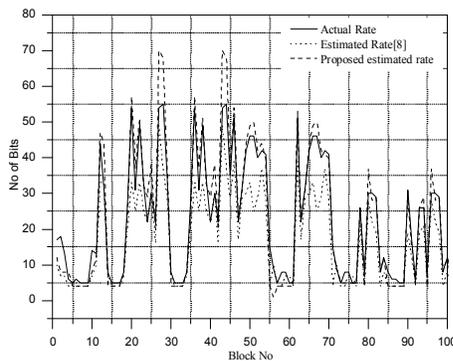
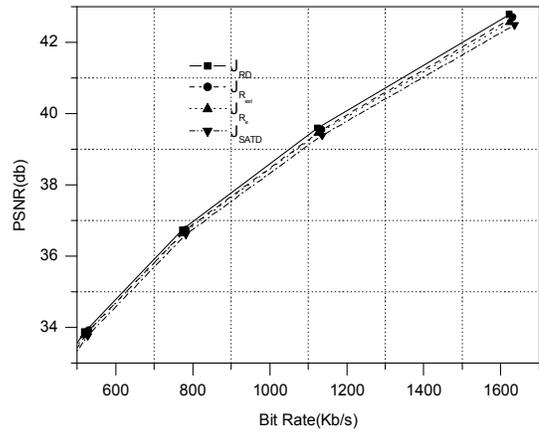
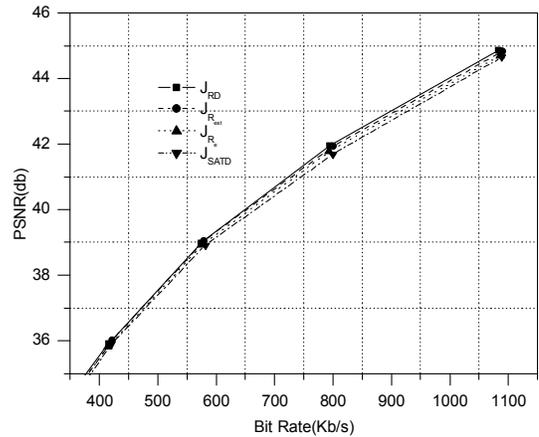


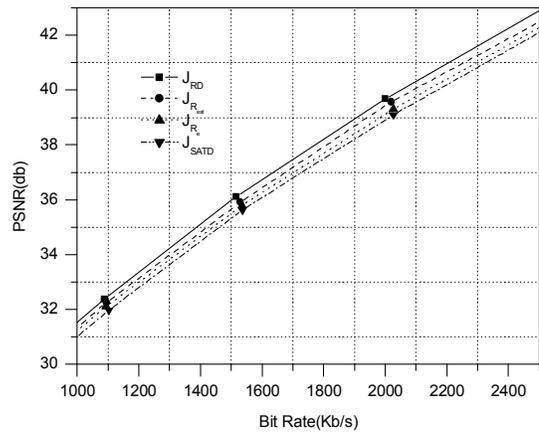
Fig. 1 comparison of our proposed method with rate estimation method described in [8]



(a) Akiyo (QCIF)



(b) Foreman (QCIF)



(c) Stefan (QCIF)

Fig. 2 Rate distortion performance of proposed method

Figure 1 shows the comparison of our proposed method with rate predictor described in [8]. Data is collected from first 100 4x4 blocks of foreman sequence. It is shown that the proposed technique is good enough as compared to R_e .

Table I shows the experimental results of proposed rate estimation method with compared to original encoder with rate-distortion cost $J_{RD}(=SSD + \lambda R)$. It is shown than the proposed method reduces up to 53% of computation time without negligible degradation of bit rate and quality.

TABLE I
Experimental results for proposed estimation method
with compared to J_{RD}

Sequence	QP	ΔP_{snr}	$\Delta \text{Bit}\%$	$\Delta T\%$
Akiyo QCIF	28	+0.02	+1.74	40.35
	32	+0.04	+2.88	39.21
	36	+0.06	+3.33	36.16
Foreman QCIF	28	+0.00	+1.29	50.00
	32	+0.04	+2.88	39.21
	36	+0.07	+3.16	41.17
Stefan QCIF	28	-0.21	+1.99	53.08
	32	-0.16	+0.55	50.00
	36	-0.08	+1.37	47.61

Table I also indicates that percentage of complexity reduction is decreased with increasing the QP values. This is because there is large number of non-zero coefficients at small QP values. If true encoding process is used large entropy coding time is spent at small QP values whereas in our proposed method no entropy coding is required during mode-decision process.

To evaluate the performance of RD characteristics four different cost functions, $J_{RD}(=SSD + \lambda R)$, $J_{R_{est}}(=SSD + \lambda R_{est})$, $J_{R_e}(=SSD + \lambda R_e)$ and $J_{SATD}(=SATD + 4P\lambda)$ were used in intra 4x4 mode decision for Akiyo, Foreman, and Stefan sequences. Figure 2 indicates the RD performance of four cost functions of akiyo, foreman and stefan video sequences respectively. It is shown that cost function using the proposed rate estimator is very close to RD optimized curve.

6. CONCLUSION

In this paper, simple and fast bit rate estimation method of 4x4 blocks of H.264/AVC is presented. This method is based on the VLC tables used in CAVLC entropy coding method. The experimental results verified that the proposed technique is suitable for intra mode decision of H.264/AVC. With the proposed scheme, entropy coding can be skipped during the mode decision. The proposed technique reduced up to 53% of encoding time of intra frame coding of H.264/AVC.

ACKNOWLEDGMENT

The work described in this paper was substantially supported by a grant from City University of Hong Kong, Hong Kong SAR, China. [Project No.7001796].

7. REFERENCES

- [1] Joint Video Team(JVT) of ISO/IEC MPEG and ITU-T VCEG, "Joint Video Specification (ITU-T Rec. H.264/ISO/ISE 14496-10)- Joint Committee Draft," Doc. JVT-G050r1.doc, Dec. 2002.
- [2] P. Topiwala, G. Sullivan, A. Joch, F. Kossentini, "Performance evaluation of H.26L TML 8 vs. H.263++ and MPEG4," Document VCEG-O42, 15th Meeting, ITU-T Q.6/SG16, Video Coding Experts Group, Pattaya, Thailand, December 2001.
- [3] B. Erol, M. Gallant, G. Cote, F. Kossentini, "The H.263+ video coding standard: complexity and performance," in: IEEE Data Compression Conference, Snowbird, Utah, March 1998, pp. 259-268.
- [4] Changsung Kim, Hsuan-Huei Shih and C-C- Kuo, "Fast H.264 Intra-prediction mode selection using joint spatial and transform domain features," J. Vis. Commun. Image R. 17 (2006) 291-310.
- [5] Chun-Ling Yang, Lai Man Po and Wing-Hong Lam, "A fast H.264 Intra Prediction algorithm using Macroblock Properties," ICIP-2004, pp 461-464.
- [6] Feng Pan, Xiao Lin, Susanto Rahardja, Keng Pang Lim, Z.G. Li, Dajun Wu, and Si WU, "Fast Mode Decision Algorithm for Intraprediction in H.264/AVC Video Coding," IEEE Trans. on Circuits and Sys. For Video Tech., vol. 15, no. 7, July 2005, pp-813-822.
- [7] Joint Video Team (JVT) reference software version 8.3, http://iphome.hhi.de/suehring/tml/download/old_jm/
- [8] C.H. Tseng, H.M. Wang, and J. F. Yang, "Enhanced Intra 4x4 Mode Decision for H.264/AVC Coders," IEEE Trans. on Circuits Syst. Video Tech., vol. 16, no. 8, August 2006, pp.1027-1032.
- [9] Iain E. G. Richardson, H.264 and MPEG-4 Video Compression - video coding for next generation multimedia, John Wiley & Sons, 2003, pp. 198-207.
- [10] ITU-T Rec. H.264/ISO/IEC 11496-10, "Advanced Video Coding", Final Committee Draft, Documents JVT-E022, September 2002.