# ADAPTIVE RATE ESTIMATION FOR H.264/AVC INTRA MODE DECISION

*Ke-Ying Liao*[1]*, Jar-Ferr Yang*[1]*, and Ming-Ting Sun*[2]

Institute of Computer and Communication Engineering, Department of Electrical Engineering,
National Cheng Kung University, Tainan, Taiwan[1]
Department of Electrical Engineering, University of Washington, Seattle, USA[2]
jfyang@ee.ncku.edu.tw

## ABSTRACT

In this paper, a modified bit-rate estimation method is proposed to reduce the computation for 4x4 intra mode decision of H.264/AVC video encoder. The number of coded bits is modeled by a linear combination of existing coding parameters, which are highly related to the entropy coding of H.264/AVC. Furthermore, to improve the accuracy of the estimation, the proposed scheme is made adaptive to the information obtained from previously coded blocks. Comparing to the original rate distortion optimized (RDO) encoding process, which needs to calculate the actual encoded bits of H.264/AVC for each coding mode, the proposed adaptive rate estimation can save about 28% and 21% of the total encoding time for QCIF and VGA sequences, respectively. For the coding performance, the proposed method achieves nearly no loss in visual quality with only slight bit-rate increases.

***Index Terms—*** H.264/AVC, Intra Rate Estimation, Rate-Distortion Optimization, Intra Rate Prediction, Adaptive Bit-rate Estimation

## 1. INTRODUCTION

Nowadays, high quality and low rate multimedia transmission and storage has become increasingly important in modern communications and entertainments. Several video compression standards with different computation complexity have been developed to support these multimedia applications. Comparing to other existing video coding standards, H.264/AVC achieves an excellent coding performance due to several adopted new coding techniques, such as inter/intra prediction, integer transform, multiple reference frames, variable block-size motion estimation, in-loop deblocking filter, highly efficient entropy coding, and rate-distortion optimization, etc. The rate distortion optimization (RDO) is often used to greatly increase the performance gain in both intra and inter mode decisions.

+ The research was supported by National Science Council under Grant Number NSC 95-2221-E-006 -223, Taiwan.

However, the RDO procedure in the H.264/AVC reference software also greatly increases the computational complexity. Thus, the development of fast RDO computation algorithms for H.264/AVC encoders has gained significant attentions recently. Several methods to speed up the intra/inter mode decision process of H.264/AVC encoders have been developed in [2] and [3]. Intra/inter rate estimation [4]-[6] have also been proposed to reduce the encoding time in entropy coding. In this paper, we present an adaptive bit-rate estimation method to predict the number of encoded bits.

The rest of this paper is organized as follows. In Section 2, we briefly review the concept of rate-distortion optimization and introduce some related fast computation methods. In Section 3, we model the coding bit-rate as a linear combination of coding parameters, which are related to entropy coding, such that we can avoid the actual encoding process to save the computation. In Section 4, simulation results with brief discussions to verify the proposed method are then presented. Finally, several conclusions about this paper are given in Section 5.

## 2. RELATED WORKS

The H.264/AVC JM reference encoder adopts the RDO technique for the mode decision to achieve the best coding performance. The RDO based on a Lagrange cost function is expressed as

$$J_{RD} = D_{SSD} + \lambda R_{Act}, \tag{1}$$

where the distortion measure $D_{SSD}$ denotes the sum of squared difference of the current block and the reconstructed block, and $R_{Act}$ is the actual number of coded bits resulted from the residual block. In (1), the Lagrange multiplier $\lambda$ is expressed as a function of the quantization parameter, QP. In order to determine the optimal coding mode of each block, all inter/intra modes for each coding block have to go through the entire encoding process, including forward and inverse transforms, quantization, and entropy coding. The encoder tries to find the minimum Lagrange cost by trying all possible inter/intra modes for each block. Thus, it results in high computational cost. To speed up the computation of Lagrange costs, Tu et. al in [5] proposed a **T**ransform-**D**omain bit-rate estimation scheme, which models the source rate as a linear combination of the number of nonzero quantized transform coefficients $N_{nz}$, and the sum of absolute values of all nonzero quantized coefficients $E_{QTC}$ as:

$$R_{TD} = \alpha N_{nz} + \beta E_{QTC} \tag{2}$$

where $\alpha$ and $\beta$ depend on the quantization parameter QP. The rate estimation can avoid the actual encoding process of H.264/AVC in the mode decision, which is very computationally expensive.

To reduce the computation, the JM reference encoder also suggests another cost function $J_{SATD}$ as

$$J_{SATD} = D_{SATD} + \lambda' 4P \tag{3}$$

where the distortion measure $D_{SATD}$ is the sum of absolute Hadamard-transformed differences between the original block and the predicted block selected by the intra prediction. The most probable mode is first determined by the coded intra modes of the upper and left blocks. During the intra 4x4 mode decision, $P$ = 0 if the candidate prediction mode is the most probable mode, otherwise $P$ = 1. The $D_{SATD}$ can be directly obtained after intra prediction. Thus, the simplified cost function stated in (3) does not need to perform the actual encoding/decoding processes for the residual components. Although this approach can save computation, the quality degradation is significant compared to that using the rate-distortion optimization in (1).

An enhanced rate estimation for the 4x4 intra mode decision for H.264/AVC is proposed in [2]. The rate estimation further considers the coding parameters of context-adaptive variable length coding (CAVLC), i.e., the number of nonzero coefficients $T_c$ and the number of trailing ones $T_o$ after thresholding the transformed residual block, to improve the precision of bit-rate estimation. The **CAVLC**-based cost function proposed in [4] is expressed by

$$J_{CAVLC} = D_{SSD} + \lambda(3T_c - T_o + E_{QTC} + T_z + 4P) \tag{4}$$

where $D_{SSD}$ is the sum of squared difference, and $T_z$ is the number of total zeros in the coding block.

It is clear that the closer the source bit-rate model is to the characteristics of the VLC tables used in entropy coding, the more accurate the bit-rate prediction will be. For the CAVLC-based methods, the entropy coding is not required in the mode decision after intra prediction in the H.264/AVC encoding process. However, simulations show that the uses of the properties of CAVLC and VLC tables may not be good enough to accurately estimate the number of source bits of 4x4 intra block of H.264/AVC.

## 3. ADAPTIVE BIT-RATE ESTIMATION

According to the previous rate estimation functions and the uses of VLC tables of H.264/AVC, we investigate a modified cost function for the 4x4 intra mode prediction as:

$$J_{proposed} = D_{SSD} + \lambda B \tag{5}$$

where $B$ is the estimated number of bits for the 4x4 quantized block, and is expressed by

$$B = \alpha N_{nz} + \beta E_{QTC} + \gamma N_{zc} + \mu T_z + 4P \tag{6}$$

where $\alpha, \beta, \gamma,$ and $\mu$ are weighting factors used in bit rate estimation function, and $N_{zc}$ is the number of times of non-zero zero-run-length after the zigzag scan. Thus, $N_{zc}$ can be expressed by

$$N_{zc} = \sum_{k=1}^{N_{nz}} \phi_k, \quad \phi_k = \begin{cases} 1, & i_{QTC(k)} - i_{QTC(k-1)} > 1 \\ 0, & i_{QTC(k)} - i_{QTC(k-1)} = 1 \end{cases} \tag{7}$$

where $i_{QTC(k)}$ is the index of $k_{th}$ non-zero quantized coefficients after zigzag scan. Therefore, the best intra 4x4 mode is considered by finding minimum cost in all probable intra 4x4 modes. The comparisons between the actual numbers of bits and the estimated numbers of bits for the Foreman sequence are shown in Fig.1. We found that although most of time it performs well, there are some significant prediction errors between the actual numbers of coded bits and the predicted numbers of bits by using (6). This indicates that the parameters in the prediction suggested in (6) contain insufficient information for making accurate estimations.

To further improve the prediction, we suggest an adaptive bit-rate estimation, which effectively collects information from the quantized coefficients involved in (6) from several previous reconstructed blocks. In our experiments, we use the statistics up to the previous 15 blocks to increase the accuracy of the prediction as follows:

$$\begin{aligned} B' = {} & \alpha'\left(\sum_{p=1}^{15} N_{nz}{}^{(P)} + N_{nz}\right) + \beta'\left(\sum_{p=1}^{15} E_{QTC}{}^{(P)} + E_{QTC}\right) \\ & + \gamma'\left(\sum_{p=1}^{15} N_{zc}{}^{(P)} + N_{zc}\right) + \mu'\left(\sum_{p=1}^{15} T_z{}^{(P)} + T_z\right) - \sum_{p=1}^{15} B^{(P)} \end{aligned} \tag{8}$$

where parenthesis superscript $p$ denotes the $p^{\text{th}}$ previously-quantized block which is coded as the 4x4 intra mode. By incorporating the information from the previously coded blocks, it helps the algorithm adapt to the local variations of the statistics in the actual video sequences. By training the first 30 frames of Forman, Mobile, and News sequences in QCIF@30fps with different QPs, we choose the prediction coefficients as $\alpha' = 2.952, \beta' = 0.55, r' = 1.395$, and $\mu' = 0.$ 818 for the estimation model depicted in (8) and use the same Lagrange multiplier $\lambda$ as that in the original RDO cost function described in (5).

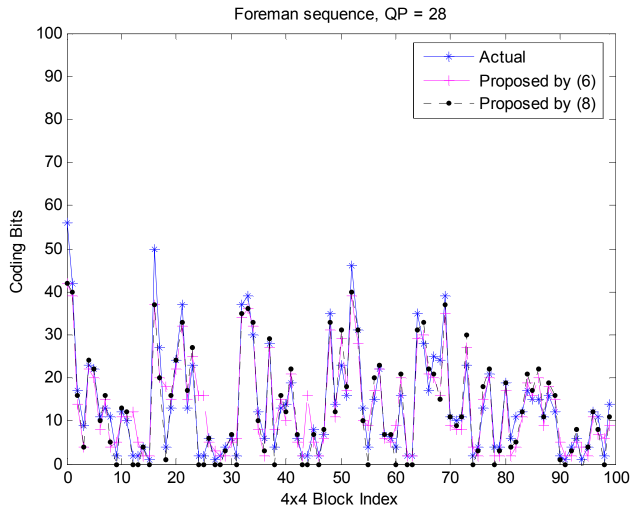



FIG 1. COMPARISON OF ACTUAL CODING BITS AND ESTIMATED CODING BITS PREDICTED BY (6), AND (8) FOR FORMAN SEQUENCE (QCIF)

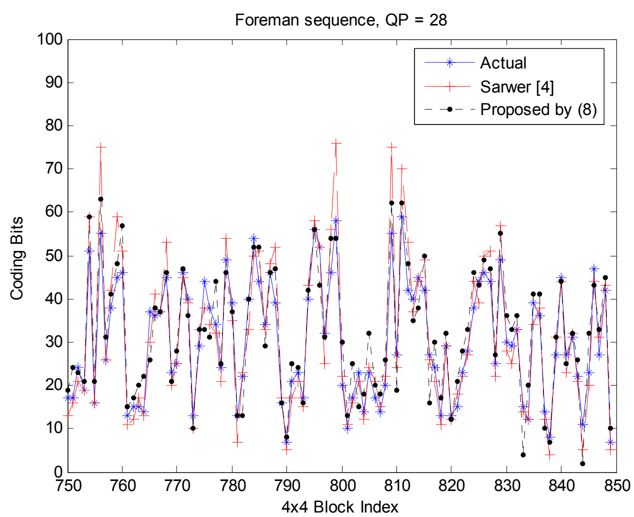



FIG 2. COMPARISONS OF THE EXISTED AND THE PROPOSED BIT-RAT ESTIMATION METHODS FOR FORMAN SEQUENCE IN QCIF

By using the proposed adaptive bit-rate estimation function, the predicted bit-rate curve depicted in Fig. 1 shows that the rate estimation using (8) achieves better prediction, where the estimated numbers of bits are closer to the actual numbers of coded bits than the method using (6). Fig. 2 shows the estimated results by the proposed method and the method suggested in [4] for some 4x4 coded blocks. The mean squared error (MSE) performance of the mode-decision using the rate estimation given by (4), (6), and (8) for different sequences is given by Table I. The simulations of Fig. 2 and Table I illustrate that the proposed adaptive rate predictor accomplishes a good estimation of the numbers of encoded bits for the 4x4 intra prediction.

## 4. EXPERIMENTAL RESULTS

To show the performance of using the proposed model for RDO, we apply the bit-rate estimation function to the RDO process in H.264/AVC reference software JM version 10.2 with Microsoft Visual C++ 6.0. The experiments are conducted on an Intel P4 3.4 GHz personal computer with 1 GB RAM. In our experiments, only the 4x4 block size is utilized for the intra prediction of the mode decision. Some encoding parameters are shown in Table II. The ΔPSNR and ΔBit-Rate (%) represent the PSNR and bit rate differences between the actual and the proposed rate calculation, respectively. $\Delta T_{Total}$ (%) is defined as the percentage of time saving of total encoding time.

In the simulations, 300 frames of several test video sequences such as Foreman and Stefan in QCIF @ 30 fps are used. Fig. 3 shows the R-D curves achieved by the different cost functions mentioned above. Simulations show that the suggested algorithm only suffers slight performance degradation compared to the rate-distortion optimization of the JM encoder. From Table III, the suggested cost function can reduce about 28% of the total encoding time compared to the conventional H.264/AVC with the RDO on. The performance reduction is acceptable by human perception because the PSNR only decreases less than 0.1 dB in average.

In Table IV, we also show the coding performances of the proposed scheme and RDO without intra 4x4 mode (i.e. only used the intra 16x16 mode) with respect to the original RDO method using 30 frames of video in VGA resolution @ 30 fps. It can be seen that the intra 16x16 mode can reduce about 73% of the total encoding time compared to the RDO with full mode on, but the PSNR decreases by 0.43 dB approximately and the bit-rate increases by 19.8% in average. On the other hand, although the encoding time is only reduced by about 21% in average, the RD performance achieved by the proposed scheme is much closer to the original RDO solution.

## 5. CONCLUSION

Rate distortion optimization plays an important role in optimizing H.264/AVC coding performance. However, the computation is extremely high since each block needs to be encoded and decoded for each mode to determine the best inter/intra prediction mode. In this paper, we first propose a modified bit-rate estimation function for the 4x4 intra prediction, which effectively uses existed coding information provided in the CAVLC coder. In addition, we further propose an adaptive source bit-rate estimation method for better prediction. The proposed adaptive method utilizes information from the previous coded blocks. Experimental results show that the proposed adaptive rate estimation method is accurate for the 4x4 intra mode decision, and can reduce the coding time by about 28% and

21 % compared to the original method in the H.264/AVC JM reference encoder for QCIF and VGA sequences.

## 6. REFERENCES


[1] Joint Video Team, ‘Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC),’ Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-G050, March 2003.

[2] C. H. Tseng, H. M. Wang, and J. F. Yang, “Enhanced Intra-4×4 Mode Decision for H.264/AVC Coders,” *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 16, no. 8, pp. 1027–1032, Aug. 2006.

[3] K. P. Lim, S. Wu, D. J. Wu, S. Rahardja, X. Lin, F. Pan, and Z. G. Li, “Fast inter mode selection,” *in Joint Video Team (JVT)*, Doc. JVT-I020, Sept. 2003..

[4] M.G. Sarwer, and L.M. Po, “Bit Rate Estimation for Cost Function of 4×4 Intra Mode Decision of H.264/AVC,” *in International Conference on Multimedia & Expo*, Beijing, China, pp. 1579-1582, July, 2007.

[5] Y. K. Tu, J. F. Yang, and M. T. Sun, “Efficient Rate-Distortion Estimation for H.264/AVC Coders,” *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 16, no. 5, pp. 600–611, May. 2006.

[6] D. K. Kwon, M. Y. Shen, and C.-C. J. Kuo, “Rate Control for H.264 Video With Enhanced Rate and Distortion Models,” *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 16, no. 8, pp. 1027–1032, Aug. 2006.


TABLE I. MSE OF THE PROPOSED RATE ESTIMATION MODELS AND THE METHOD IN [4] FOR QCIF SEQUENCES WITH QP = 28

| Sequences | Sarwer [4] | Proposed (6) | Proposed (8) |
|---|---|---|---|
| Foreman | 43.78 | 28.21 | 13.26 |
| Mobile | 124.13 | 56.99 | 30.21 |
| Coastguard | 19.78 | 22.25 | 13.28 |
| Average | 62.6 | 35.8 | 18.9 |

TABLE II. ENCODING PARAMETERS USED IN THE SIMULATIONS

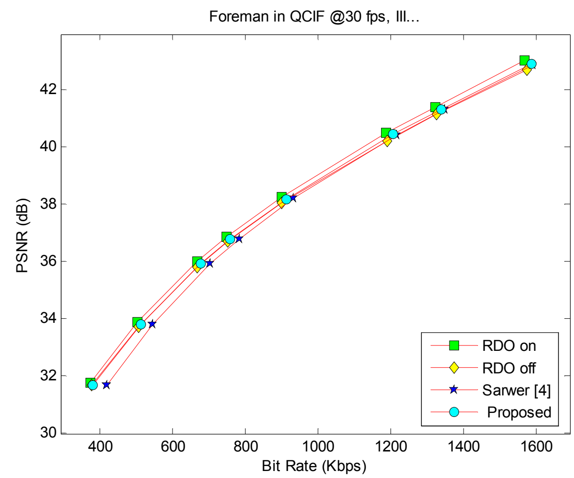
| Functions | Parameters |
|---|---|
| Profile | Baseline |
| Sequence type | III… |
| Entropy Coding | CAVLC |
| Frame Rate | 30 fps |
| RD Optimization | Off (Hadamard Transform) /On |

TABLE III. CODING PERFORMANCE OF THE PROPOSED RATE ESTIMATION SCHEME COMPARED TO R-D OPTIMIZATION FOR QCIF SEQUENCES

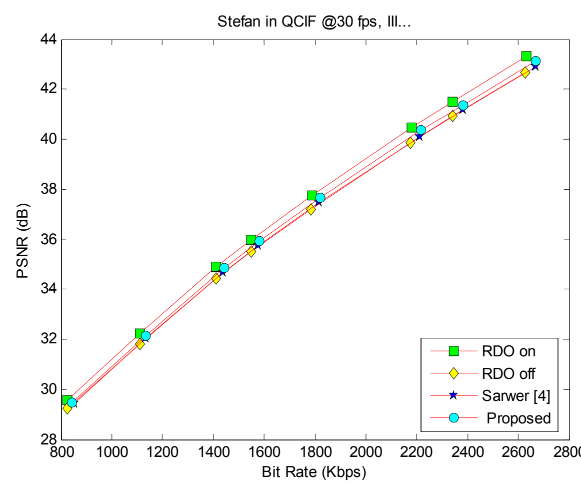
| Sequences | QP | $\Delta T_{Total}$ (%) | ΔPSNR (dB) | ΔBit-Rate (%) |
|---|---|---|---|---|
| Foreman | 26 | 33.9 | -0.07 | 1.41 |
| | 28 | 31.36 | -0.06 | 1.51 |
| | 29 | 29 | -0.06 | 1.8 |
| | 32 | 24.74 | -0.07 | 1.71 |
| Coastguard | 26 | 32.71 | -0.07 | 2.02 |
| | 28 | 33.48 | -0.07 | 2.18 |
| | 29 | 28.05 | -0.06 | 2.42 |
| | 32 | 23.46 | -0.08 | 2.27 |
| Mother & Daughter | 26 | 25.95 | -0.1 | 1.91 |
| | 28 | 27.51 | -0.11 | 1.91 |
| | 29 | 23.37 | -0.08 | 2.22 |
| | 32 | 21.92 | -0.09 | 1.63 |
| Average | | 27.954 | -0.0767 | 1.916 |

TABLE IV. CODING PERFORMANCE OF THE PROPOSED RATE ESTIMATION SCHEME AND RDO WITHOUT INTRA 4X4 MODE COMPARED TO R-D OPTIMIZATION FOR VGA SEQUENCES

| Sequences | QP | Intra 16x16 mode only | | | Proposed | | |
|---|---|---|---|---|---|---|---|
| | | $\Delta T_{Total}$ (%) | ΔPSNR (dB) | ΔBit-Rate (%) | $\Delta T_{Total}$ (%) | ΔPSNR (dB) | ΔBit-Rate (%) |
| Race | 26 | 74.119 | -0.386 | 18.405 | 25.755 | -0.093 | 5.775 |
| | 28 | 73.553 | -0.36 | 20.139 | 26.357 | -0.134 | 7.256 |
| | 29 | 73.971 | -0.373 | 20.697 | 27.227 | -0.15 | 8.31 |
| | 32 | 72.967 | -0.373 | 21.723 | 21.483 | -0.22 | 10.5 |
| Ballroom | 26 | 73.415 | -0.387 | 17.524 | 18.32 | -0.053 | 3.451 |
| | 28 | 73.482 | -0.371 | 18.367 | 19.163 | -0.062 | 4.226 |
| | 29 | 73.14 | -0.388 | 18.553 | 16.956 | -0.07 | 4.752 |
| | 32 | 72.971 | -0.394 | 18.394 | 17.171 | -0.098 | 6.771 |
| Flamenco | 26 | 73.519 | -0.515 | 20.601 | 24.341 | -0.341 | 12.611 |
| | 28 | 73.37 | -0.513 | 21.384 | 28.708 | -0.453 | 14.74 |
| | 29 | 73.222 | -0.53 | 21.311 | 13.643 | -0.509 | 16.214 |
| | 32 | 73.196 | -0.552 | 20.571 | 17.426 | -0.7 | 20.128 |
| Average | | 73.41 | -0.429 | 19.806 | 21.379 | -0.24 | 9.561 |




(A) FOREMAN




(B) STEFAN

FIG. 3. THE PERFORMANCE COMPARISONS RESULT FOR SEQUENCES: (A) FOREMAN; (B) STEFAN IN 100 INTRA FRAMES @30FPS