Fast Intra-Prediction Model Selection for H.264 Codec

Chang-sung Kim, Qing Li and C.-C. Jay Kuo

Integrated Media Systems Center and Department of Electrical Engineering
University of Southern California, Los Angeles, CA 90089-2564

ABSTRACT

We investigate the encoding speed improvement for H.264 with a special focus on fast intra-prediction mode selection in this work. It is possible to adopt the rate-distortion (RD) optimized mode in H.264 to maximize the coding gain at the cost of a very high computational complexity. To reduce the complexity associated with the intra-prediction mode selection, we propose a two-step fast algorithm. In the first step, we make a course-level decision to split all possible candidate modes into two groups: the group to be examined further and the group to be ignored. The sizes of these two groups are adaptively determined based on the block activities. Then, in the second step, we focus on the group of interest, and consider an RD model for final decision-making. It is demonstrated by experiment results that the proposed scheme performs 5 to 7 time faster than the current H.264 encoder (JM5.0c) with little degradation in the coding gain.

Keywords: H.264/JVT, Intra prediction, encoder complexity, Intra mode decision.

1. INTRODUCTION

Various video compression standards have been developed in the last decade for a wide range of applications. The diversity of applications and its different flavor of technological demand have resulted in two families of standards: the ITU H.26x and the ISO MPEG families. Recently, an emerging video coding standard, known as H.264/JVT, is jointly developed by ITU-T and ISO MPEG. H.264 has significantly improved the coding performance in both low and high bit rates as compared with previous coding standards such as H.263, MPEG-2 and MPEG-4 [1,7]. Primary technical objectives of H.264 include the following [2]: (i) significant coding efficiency (the average bitrate saving up to 50% as compared with H.263+ and MPEG-4 Simple Profile); (ii) adaptation to delay constraints (the low delay mode); (iii) error robustness; and (iv) network friendliness (NAL). To accomplish these objectives, many features are included in H.264 such as the 4x4 integers transform, the deblocking filter, variable block sizes and multiple reference frames for motion compensation, enhanced Intra prediction, and context based adaptive binary arithmetic code (CABAC).

Among these features, enhanced Inter and Intra prediction techniques are key factors to the success of H.264. To achieve high coding efficiency, H.264 employs the rate-distortion optimization (RDO) technique to get the best result in terms of visual quality and coding rates. In order to perform RDO, the encoder encodes video by exhaustively searching the best mode in the RD sense among various predefined modes [3]. As a result, the computational complexity of the H.264 encoder is dramatically increased, which makes it difficult for practical applications such as real time video communication. Several attempts have been made to explore fast algorithms in motion estimation for H.264 video coding [4-6]. What seems to be lacking is the fast algorithm in the intra-mode prediction for H.264. Only few attempts have so far been made. Pan *et al.* [3] proposed a fast mode decision scheme based on pre-processing, which measures the average edge direction of a given block so as to reduce the number of probable modes to achieve complexity reduction. The overall performance is 20~30% faster than the RDO method at the cost of 2% extra bits and 55~65% faster at the cost of 5% extra bits. In terms of the speedup factor, there is still a huge gap between the desired encoding speed and the actual one. Note also that the work in [3] shows the potential that the optimal performance in Intra prediction can be achieved by using some *a priori* information. That is, the *a priori* information can be used to reduce the complexity by detecting most probable modes at an early stage. The following two questions arise naturally: (i) what kind of a *priori* information can be utilized to obtain most probable modes and (ii) what method can replace RDO, which is most time consuming task.

In this work, we present a simple yet effective fast mode decision algorithm for H.264 Intra prediction using a two-level decision process. At the first level, some unlikely modes are filtered out. At the second level, the best mode is chosen among the remaining candidates. For further complexity reduction, an adaptive RD cost computation procedure is adopted. This procedure switches between our proposed RD model and the RDO procedure given in the original H.264 reference codes based on the block activities. The two-level mode decision scheme is implemented and integrated with H.264 JM5.0c codes. It is compared with RDO in some performance metrics such as the computational cost, the average PSNR and the coding bit-rate for all the sequences recommended in [10]. Simulation results demonstrate an excellent compression performance of the proposed fast algorithms for a wide range of bit-rates. For the computational speed, the proposed algorithm is five to seven times faster than the H.264 RDO method with little performance degradation.

The rest of this paper is organized as follows. After a brief overview of H.264 intra mode decision, we introduce the proposed two-level fast mode decision framework in Section 2. A rate distortion model for mode decision is proposed and an adaptive RD cost computation procedure is presented in Section 3. In Section 4, we provide experimental results to show the performance of the proposed scheme in terms of the rate distortion tradeoff and the speedup factor. Concluding remarks are given in Section 5.

2. FAST INTRA PREDICTION MODE DECISION

2.1 Intra-mode Decision

This section reviews the H.264 Intra mode decision scheme and analyzes the computational complexity of the exhaustive search scheme. Intra prediction in H.264 exploits the spatial correlation between the adjacent macroblocks. In JVT, the current macroblock is predicted by adjacent pixels in the upper and the left macroblocks that are decoded earlier. Then, the residual between the current macroblock and its prediction is transformed, quantized and entropy coded. Roughly speaking, the smaller the difference is, the fewer the coding bits are demanded for the current macroblock. To get a richer set of prediction patterns, H.264 offers 9 prediction modes for 4×4 luma blocks and 4 prediction modes for 16×16 luma blocks. For the chrominance components, there are 4 prediction modes applied to the two 8×8 chroma blocks (U and V).

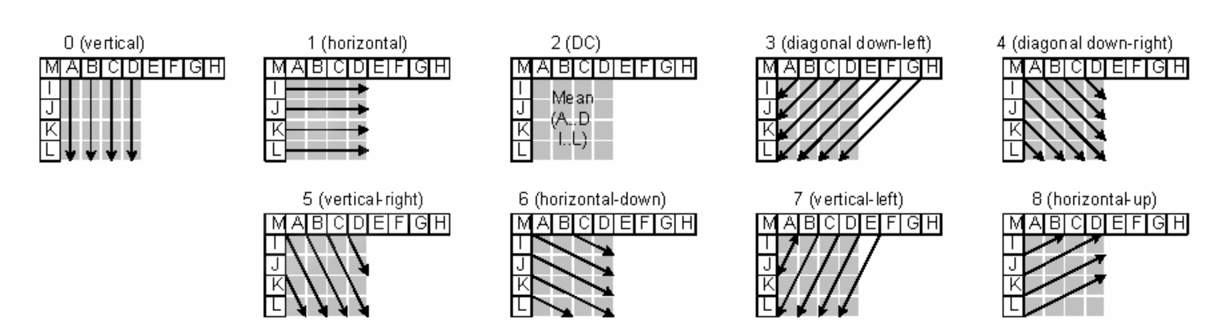


Figure 1: The nine intra prediction modes for the 4x4 luminance block.

The nine intra prediction modes for 4x4 luminance blocks are illustrated in Figure 1, which include the DC prediction (Mode 2) and eight directional modes, labeled 0 thru 8 [8]. The arrows in Figure 1 indicate the direction of prediction in each mode. For modes 3-8, the predicted samples are formed from a weighted average of the prediction samples A-M [9]. For example, if mode 4 is selected, the top-right sample of gray 4x4 block (cross point of D and I) is predicted by round(*B*/4+ *C*/2 + *D*/4). Four prediction modes for 16x16 macroblock are vertical (mode 0), horizontal (mode 1), DC (mode 2) and plane prediction (mode 3). Basically, the 16x16 intra-prediction is chosen for regions with less spatial details such as the flat region. The 4 prediction modes for each 8x8 chroma component of a macroblock are very similar to the 16x16 luminance prediction modes. Note that if any of the 8x8 blocks of the luminance component are coded in the intra mode, both chroma blocks for U and V are intra-coded using the same intra-prediction mode. H.264 is developed based on the rate distortion optimization. That is, the encoder has to select the best combination of prediction modes for each macroblock to obtain the optimal RD performance [3]. If the RDO is chosen, the mode decision for a macroblock is made by minimizing the Lagrangian functional.

The RDO procedure to encode one macroblock, denoted by *s*, in an I-frame is given below.

(a) Given the last decoded frames and the macroblock quantization parameter *QP*, the Lagrangian multiplier is given by $l_{MODE} = 0.85 \cdot 2^{QP/3}$ [18].

(b) Select the best 4x4 intra prediction mode from nine intra 4x4 macroblock modes by minimizing the following functional:

$$J(s, c, MODE \mid QP, l_{MODE}) = SSD(s, c, MODE \mid QP) + l_{MODE} \cdot R(s, c, MODE \mid QP),$$

where *QP* is the macroblock quantizer, l_{MODE} is the Lagrange multiplier for mode decision, MODE indicates a mode chosen from 9 intra 4x4 prediction modes, *SSD* is the sum of squared differences between the original 4x4 block luminance signal *s* and its reconstruction *c*, and $R(s, c, MODE \mid QP)$ represents the number of bits associated with the chosen *MODE*. It includes the bits needed for coding the intra prediction mode and DCT-coefficients for the 4x4 luminance block.

(c) Determine the best 16x16 intra prediction mode by choosing the mode that results in the minimum *SATD* (Sum of Absolute Transformed Difference).

(d) Compare the RD cost for the two best modes that are the 4x4 mode from Step (b) and the 16x16 mode from Step (c), and choose the better one as the macroblock prediction mode.

According to the above procedure of intra prediction in H.264, the number of mode combinations for luma and chroma blocks in a macroblock is *N8´ (N4´16+N16)*, where *N8*, *N4* and *N16* represent the number of modes for 8×8 chroma blocks, 4×4 and 16×16 luma blocks, respectively. This means that, for the intra coding of a macroblock in H.264, it has to perform 592 different RDO calculations to determine the optimal RDO mode [3]. As a result, the complexity of the encoder is extremely high. To reduce the encoding complexity with little RD performance degradation, the two-level mode decision is proposed in the next section.

2.2 The Proposed Mode Decision Scheme: An Overview

RDO guarantees the best mode in the RD sense since it exhaustively searches the best mode by measuring the RD cost based on the actual rate and distortion after entropy coding and reconstruction, respectively. As mentioned in the previous section, a total of 592 possible modes should be processed for the intra mode decision. This is too complex to be implemented for practical applications. It is desirable to find the best mode or the nearly best mode using some simplified method [3].

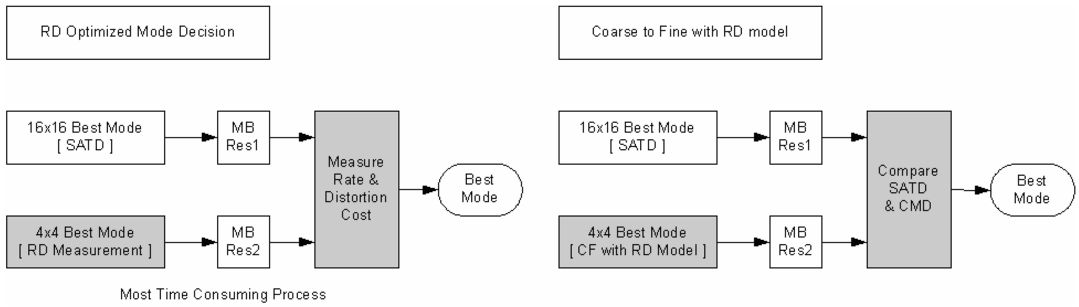


Figure 2: The framework of intra mode decision: the original intra mode decision in H.264 (left) and the proposed intra mode decision scheme (right).

In this work, we propose an efficient method to improve the encoding speed without much sacrifice at the RD performance. The original and the proposed intra mode selection schemes are shown in Figure 2. In the original H.264 Intra mode decision, the best 16x16 mode is selected by choosing the mode whose SATD (sum of absolute transform differences) value is the minimum while the best 4x4 mode is selected by choosing the one that has the minimum RD measurement. *MB res*1 and *MB res*2 in this figure indicate the MB residuals for the selected optimal 16x16 and 4x4 modes, respectively. For the final decision, the Lagrangian RD cost is computed for these two best modes, and the one with the minimum cost is chosen to be the final mode for the current macroblock. In the proposed scheme, we first reduce the complexity by replacing the RD measurement for 4x4 blocks with two-level mode decision and the RD cost computation for the final two candidates is further simplified. Comparing the original and the proposed mode decision schemes, there are two major modifications as indicated by the gray blocks.

2.3 The Two-level Mode Decision

Figure 3 gives the block-diagram for the two-level mode decision scheme applied to 4x4 blocks. The idea of the two-level (or the coarse-to-fine) mode decision scheme can be described as follows. First, a coarse–level decision is made to fast detect the most probable modes based on some transform domain feature. Then, for the fine-level mode decision, we search the best mode with an RD model. More details will be given in Sections 2.3.1 and 2.3.2.

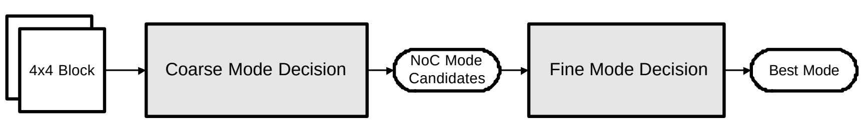


Figure 3: Two-level mode decision for 4x4 blocks.

2.3.1 The Coarse-level Mode Decision

The coarse mode decision (CMD) is used to filter out unlikely modes to decrease the number of candidates to be considered for the fine mode decision process that is more complicated. It has been observed [11] that the sum of absolute transform differences (SATD) has strong correlation with the rate-distortion performance so that it can be used as a feature to detect most probable modes. The CMD process is shown in Figure 4.

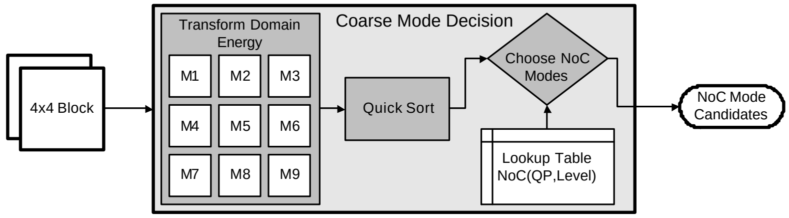


Figure 4: The block-diagram of the coarse mode decision (CMD) process.

The CMD process consists of the following two steps:

(i) Compute the approximate SATD value in all modes, and sort these modes from the smallest to the largest.

(ii) Choose a certain number of modes, which have smaller SATD values according to the block activity and the quantization parameter, where the block activity is measured in terms of the variance of residuals and the number of probable modes to be selected is specified by a predefined table called the NOC (number of candidates) table.

To compute SATD, we use the Hadamard transform to approximate DCT due to its simplicity. (Note that the Hadamard transform can be implemented with only addition and shift operations.) Then, all prediction modes are sorted according to their SATD values. This can be efficiently implemented using the quick sort algorithm of complexity $O(n \log(n))$ [12]. The next step is to threshold the SATD values to select the most probable candidates for the second-level (or the fine-level) decision making.

The number of candidates (NoC) for selection is critical in complexity reduction, since the fewer modes are chosen, the higher the complexity is saved. The NoC should be adaptive to the block activity. For example, some residual block has more details while others are relatively flat. Also, the quantization parameter should be taken into account in choosing NoC since a smaller NoC is needed for a larger QP. In this work, the block activity is measured in terms of the standard deviation of residual coefficients [15, 16, 17]. The NoC with different block activities and different quantization parameters can be implemented as a lookup table.

2.3.2 The Fine-level Mode Decision

The objective of the fine mode decision (FMD) is to search the best mode among the most probable modes that have passed the CMD process as shown in Figure 5. Now, the best mode is the one that has minimum rate-distortion cost defined as:

$$Best\,Mode = \arg\min_{k}\left(D_k + l \cdot R_k\right)$$

To reduce the RD performance degradation while keeping the complexity low, we choose one of the two RD methods dynamically. As shown in Figure 6, the FMD process can switch between the proposed RD model and the RDO process given by H.264. Generally speaking, the RD model contains some modeling errors, which lead to performance degradation. It is observed that the RD model error is negligible when the block residual is small. Thus, the RD model can be used. On the other hand, the RDO process should be utilized when the block residual is high to reduce the RD model error.

For a given block residual standard deviation and QP, the block activity can be measured for NoC selection as discussed in Section 2.3.1. The RD model is adopted when the NoC takes a small value such as 1, 2 and 3. The RD model used in this work will be presented in Section 3.

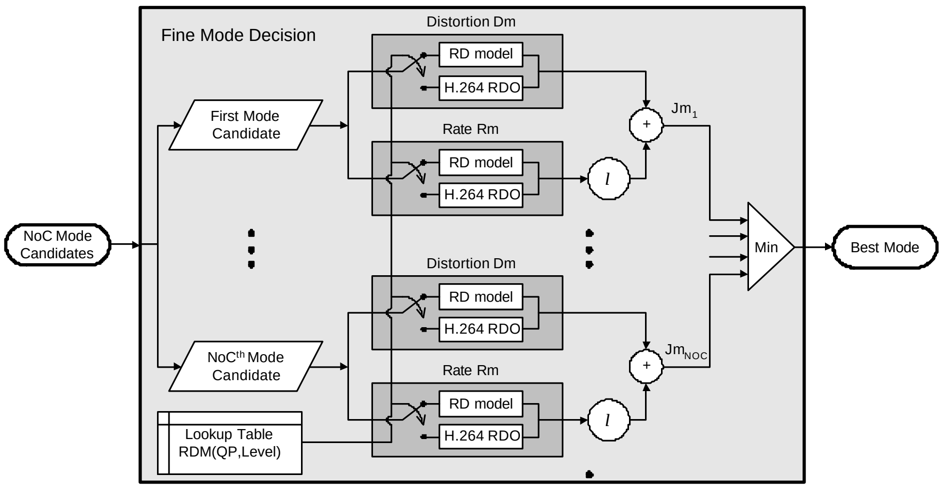


Figure 5: The block-diagram of the fine mode decision (FMD) process.

3. RATE DISTORTION MODEL FOR MODE DECISION

Several macroblock-based RD models were proposed for rate control and bit allocation in video coding [11,16,19,20, 21]. The normalized rate-distortion model [11] was based on the normalized variance of 8x8 blocks for different QP values. This model requires the training of model coefficients for different sequences, which may not be practical in real-time video coding. Yang and Jacquin [19] proposed a RD model based on the DC (mean) and the residual variance of 8x8 blocks. This process is fairly complex due to the search of the Lagrange multiplier. Besides, it is not applicable to 4x4 blocks. The quadratic rate distortion model [20] also demands the training of model coefficients. This model is not useful in our current context, since the modeling error may be large if the statistics of the input video sequence is different with that of training sequences. He and Mitra [16] proposed a r domain RD model that estimates the rate and the distortion of a macroblock by the percentage of nonzero coefficients. The model accuracy heavily depends on the non-zero percentage and parameters. The frame level rate estimation is of sufficient accuracy. However, for the rate estimation of 4x4 blocks for H.264 intra prediction, the block-based statistics is non-stationary, and the parameters obtained from the previously coded blocks are not accurate for the current block.

Thus, a more suitable RD model is needed for 4x4 blocks in our current application. We propose a new RD model to predict the rate and distortion of the 4x4 block and then refine it by the relationship between the predicted RD and the real RD. This proposed RD model is heuristic since the relationship is obtained by fitting the curve of observed data. However, it is shown by simulation that our model works well with a reasonable complexity.

3.1 Logarithmic Affine Distortion Model

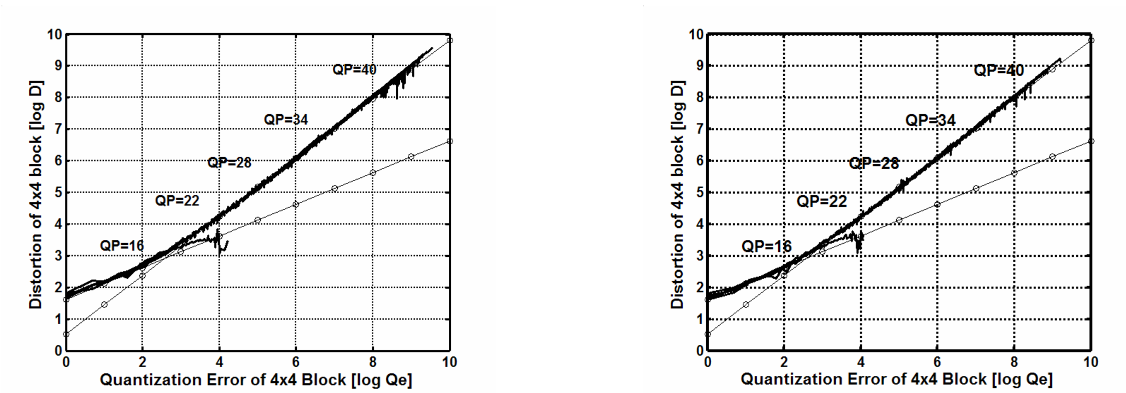
The distortion between the original and the reconstructed 4x4 blocks depends on the quantization error. Thus, the proposed distortion model is based on the quantization error (rather than the quantization parameter or the variance of the block). Figure 6 shows that the relationship between the actual distortion and the total quantization error. The relationship is approximately linear in the natural logarithmic domain. This leads to a refined RD model called the logarithmic affine distortion model, which can be written as:

$$y = \begin{cases} m_1 \cdot x + h_1, & x \le c_0 \\ m_2 \cdot x + h_2, & \text{otherwise} \end{cases}, \text{ where } y = \ln(D),\ x = \ln(Qe)$$

$$\ln(D) = \begin{cases} m_1 \cdot \ln(E_Q) + h_1, & x \le c_0 \\ m_2 \cdot \ln(E_Q) + h_2, & \text{otherwise} \end{cases}$$

$$D = e^{m_K \cdot \log(E_Q) + h_K}$$

In above, E_Q denotes the quantization error, and m_K, h_K are parameters to characterize the piecewise linear model as shown in Figure 6. The piecewise linear model for the distortion can be obtained by linear regression [22]. With this model, the distortion can be accurately estimated without reconstruction.



(a) Akiyo, QCIF

(b) Stefan, QCIF

Figure 6: The two line segments labeled by solid thick lines give the distortion model as a function of the quantization error.

3.2 Parametric Rate Model

Generally speaking, the actual coding bit rate depends on the entropy coding scheme. For the JVT baseline, the entropy codec is CAVLC (Context adaptive variable length code). It encodes 5 different types of symbols (or called tokens): (i) the coefficient token (the number of coefficients, the number of trailing ones), (ii) the sign of trailing ones, (iii) the level of nonzero coefficients, (iv) the total number of zeros before the last coefficients, and (v) the run of zeros. It uses four lookup tables for coefficient token and seven lookup tables for the level of nonzero coefficients. Different lookup tables are adaptively chosen based on the context. For this reason, it is difficult to predict the coding bit rates for 4x4 blocks.

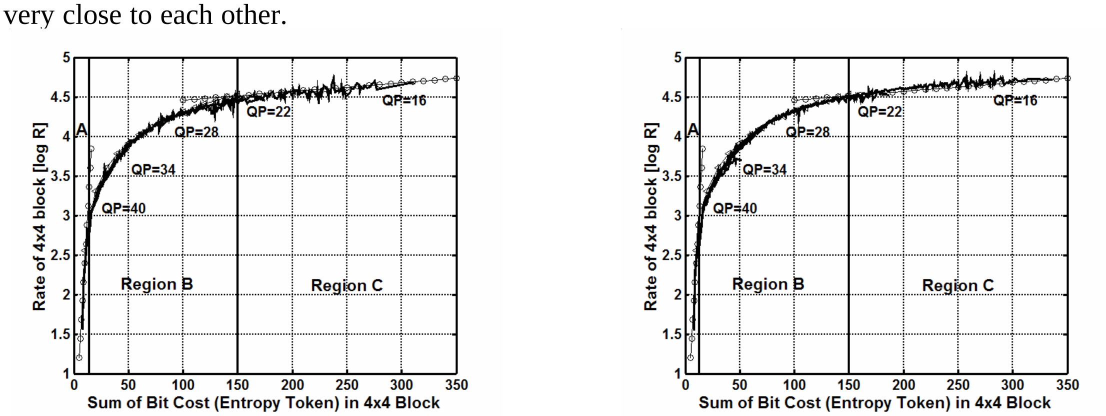
Here, we propose a rate model that predicts the rate of a 4x4 block using the entropy token used in CAVLC. Then, the predicted rate is refined furthermore using the relationship between the actual and the predicted rates. The proposed bit cost is the sum of the bit cost spent for each token within the given block.

$$Bit\ Cost_{4x4} = \sum_{x \in 4\times 4 block} C_b(x) = \sum_{x \in 4\times 4 block} (w_x \cdot I(x) + a_x)$$

where x denotes the encoding token, $C_b(x)$ is the bit cost function for x, and w_x, a_x are the weight and constant for the encoding token, and $I(x)$ denotes a linear term. To approximate the rate of a given 4x4 block, parameters w_x, a_x are obtained empirically based on observations for various test sequences. The actual bit rate is shown in Figure 7. The curve fitting method is used to approximate the mapping function. It is easier to get fitting functions in three different regions. The three regions are shown in Figure 8 and can be written as:

$$R = \begin{cases} e^{k_1 \cdot BitCost_{4x4}} & \text{, Region } A \\ e^{k_2 \cdot \ln(BitCost_{4x4} - k_3) + k_4} & \text{, Region } B \\ e^{k_5 \cdot BitCost_{4x4} + k_6} & \text{, Region } C \end{cases}$$

The parameter vectors $K = [k_1, k_2, \cdots, k_6]$ in regions A and C are obtained by linear regression and those in region B are obtained by minimizing MSE. Figure 7 compares the predicted bit rate using our rate model. They are very close to each other.



(a) Akiyo, QCIF (b) Stefan, QCIF

Figure 7. The predicted rate (circled lines) model and the actual bit rate (solid curves).

In video clips with slow motion and relatively smooth texture, the proposed two-level mode decision with the RD model gives almost the same performance as the RDO of H.264. However, its performance is degraded for fast motion and rough texture sequences. Thus, the RD computation method is changed adaptively based on the block activity as described in Section 2.4.

4. EXPERIMENTAL RESULTS

In the experiment, the proposed algorithm was integrated within the JVT reference software JM5.0. The test sequences were 15 MPEG sequences of classes A, B, and C. The system platform is the Intel Pentium 4 Processor of speed 1.8GHz, 512MB DDR RAM, and Microsoft Windows XP. The parameters of the RD model are fixed by analyzing all test sequences via an off-line computation.

To compare the rate distortion performance and the computational complexity of the proposed scheme with RDO of H.264, the PSNR and the bit rate (per frame) are measured at different QP from 10 to 40. The maximum number of candidates is limited to 4. Figures 8 (a)-(c) show the RD performance and the computational complexity for 3 different sequences. The left one shows the RD performance and the right one gives the computational complexity, which is measured in terms of the encoding time. The circled line is the RDO result and the squared line is the proposed approach. In the legend, VNoC means variable NoC and VM means the variable RD computing method according to its block activity and QP.

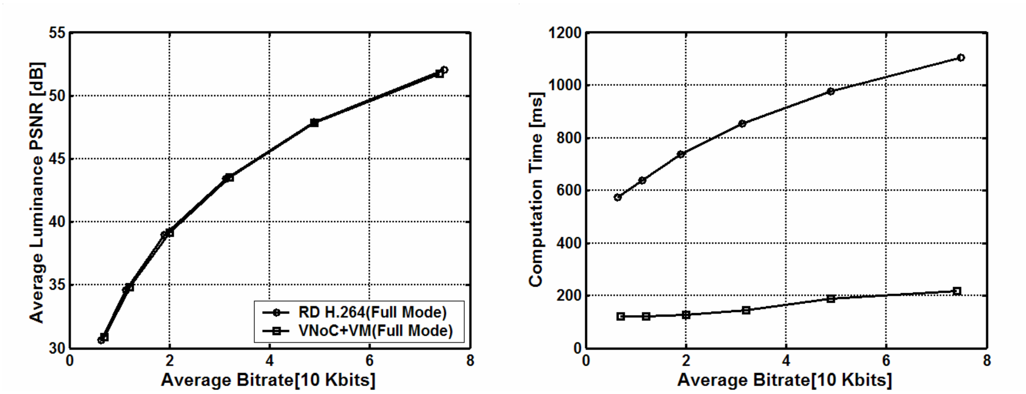


Figure 8-(a) Akiyo.QCIF [Class A]

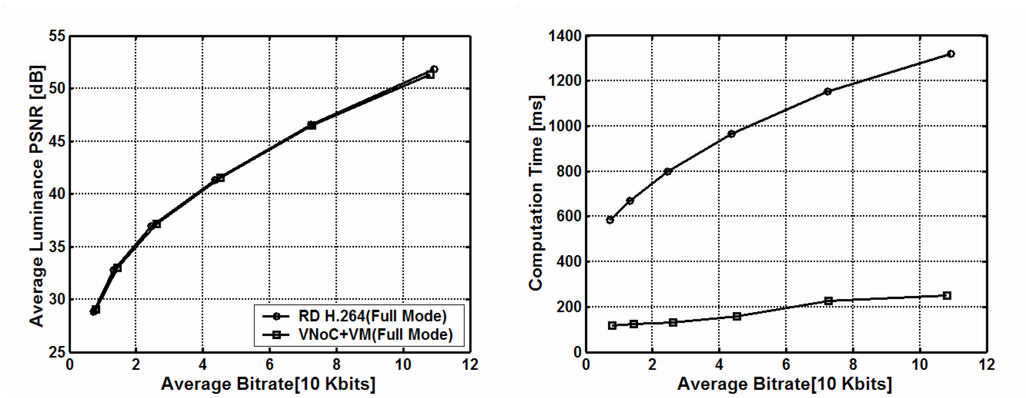


Figure 8-(b) Foreman.QCIF [Class B]

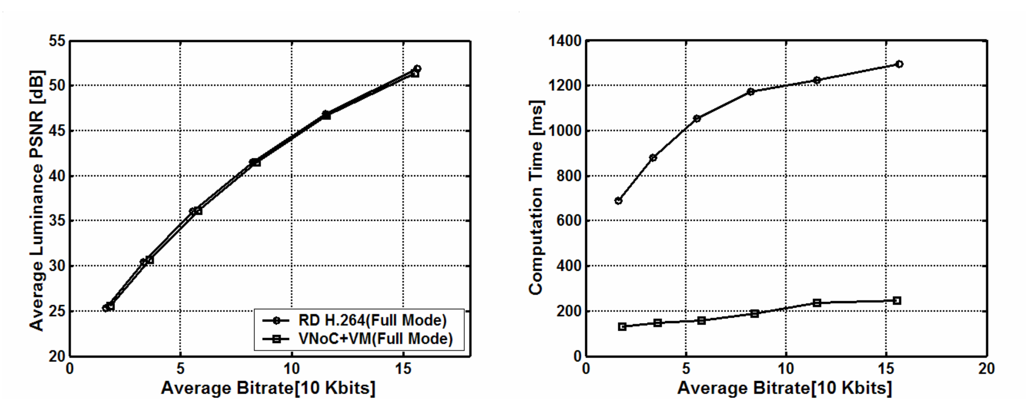


Figure 8-(c) Stefan.QCIF [Class C]

From Figure 8, we see that the proposed fast intra-mode decision scheme gives almost identical RD performance while providing a speed-up factor of 5-7. The RD performance of the proposed scheme is slightly degraded for the Stefan sequence as shown in Figure 8(c), which has a large number of macroblocks with a high block activity. Since the error for the coarse-level mode decision increases with the block activity, the rate distortion performance of the proposed scheme degrades.

The speedup factor is the ratio of the encoding time using the RDO technique and the proposed scheme. If the scale in Figure 9 is 6, then the proposed algorithm is 6 times faster than RDO mode decision. As shown in Figure 9, the proposed approach is approximately 5 to 6.5 times faster than RDO mode decision. Also, the speed factor is different from one sequence to the other. Especially, the higher the percentage of high activity blocks, the higher the speedup factor. The can be explained by the fact that the complexity of RDO increases in proportion to the percentage of high activity blocks due to the computational load to measure the bitrate and the distortion. In contrast, the complexity of the proposed scheme is limited by the RD model and the NoC, which is bounded by 4. As a result, the gap between RDO and our approach is larger.

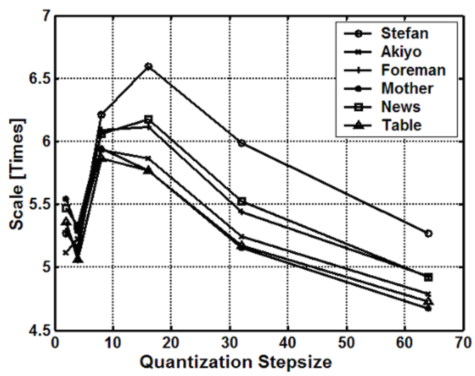


Figure 9. Variation in the time complexity and the speedup factor [QCIF]

5. CONCLUSION AND FUTURE WORK

In this research, we proposed a fast mode decision scheme for intra prediction used in H.264 encoding. With this algorithm, we can speed up the JVT reference software JM5.0 by a factor 5 to 7 times in this module with little RD performance degradation. The actual speedup performance is related with the average block activity and the quantization parameter. Future research topics include the fast mode decision for the baseline profile and further improvement of the RD model.

REFERENCES

1. 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, Dec. 2001.
2. Emerging H.26L Standard: Overview and TMS320C64x Digital Media Platform Implementation, White paper, UB video Inc., Feb 2002.
3. Feng Pan, Xiao Lin, *et al*, "Fast Mode Decision for Intra Prediction", *ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6*, JVT 7th Meeting Pattaya II, Thailand, 7-14, March 2003.
4. Xiang Li, Guowei Wu, "Fast Integer Pixel Motion Estimation", JVT-F011, 6th Meeting, Awaji Island, Japan, December 5-13, 2002.
5. Zhibo Chen, Peng Zhou, Yun He, "Fast Integer Pel and Fractional Pel Motion Estimation for JVT", JVT-F017, 6th Meeting, Awaji Island, Japan, December 5-13, 2002.
6. Hye-Yeon Cheong Tourapis, Alexis Michael Tourapis, Pankaj Topiwala, "Fast Motion Estimation within the JVT Codec", JVT-E023, 5th Meeting, Geneva, Switzerland, October 9-17 2002.
7. B. Erol, M. Gallant, G. Côté and F. Kossentini, "The H.263+ Video Coding Standard: Complexity and Performance", IEEE Data Compression Conference, Snowbird, Utah, pp. 259-268, March 1998.
8. Iain E. G. Richardson, "H.264/MPEG-4 Part 10 White Paper: Intra Prediction", www.vcodex.com, May 2003.
9. ITU-T Rec. H.264/ISO/IEC 11496-10, "Advanced Video Coding", H.264/JVT Working Draft, Document WD2R8, April 2002.
10. JVT Test Model Ad Hoc Group, "Evaluation Sheet for Motion Estimation", Draft version 4, Feb. 19, 2003.
11. Kyeong Ho Yang; Jacquin, A.; Jayant, N.S., "A normalized rate-distortion model for H.263-compatible codecs and its application to quantizer selection", Image Processing, 1997. Proceedings., International Conference on , vol.2, pp. 41 -44, Oct 1997
12. Thomas H.Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein, " Introduction to Algorithms *second edition*", The MIT Press & McGraw-Hill Book Company, 2001.
13. C. A. Gonzales, E. Viscito, "Motion video adaptive quantization in the transform domain", IEEE Trans. CSVT, Vol. 1, No. 4, Dec. 1991.
14. de Queiroz, R.L.; Eschbach, R., "Segmentation of compressed documents", Proceedings International Conference on Image Processing 1997, Volume: 3 , pp. 70 -73, Oct 1997.
15. V. Darmstaedter, et al, "Low cost spatial watermarking", Computer & Graphics, Vol. 22, No. 4, pp. 417-424, 1988.
16. Zhihai He, Sanjit K. Mitra, " A Unified Rate Distortion Analysis Framework for Transform Coding", IEEE Trans. CSVT, Vol. 11, No. 12, Dec. 2001.
17. Mester, R.; Franke, U., "Spectral entropy-activity classification in adaptive transform coding", Selected Areas in Communications, IEEE Journal on , Volume: 10 Issue: 5 pp. 913 -917, Jun 1992
18. Thomas Stockhammer, Dimitrios Kontopodis, Thomas Wiegand, "Rate-Distortion Optimization for JVT/H.26L Video Coding in Packet Loss Environment", Munich University of Technology, International Packet Video Workshop, 2002

19. Kyeong Ho Yang, Arnaud Jacquin, "Real time implementation of rate-distortion optimized coding mode selection for H.263 video coders", ICIP 98 Volume: 2, pp. 387 -390, Oct 1998.

20. Tihao Chiang, Y. Q. Zhang, "A new rate control scheme using quadratic rate distortion model", IEEE Trans. Circuits Syst. Video Technol., Feb. 1997.

21. Hwangjun Song, Jongwon Kim, and C. –C. Jay Kuo, "Hybrid DCT/Wavelet I-Frame Coding for Efficient H.263+ Rate Control at Low Bit Rate", 1998 IEEE International Symposium on Circuits and Systems, Monterey, California, May 31-June 3, 1998.

22. Zhihai He, Sanjit K. Mitra, " A Unified Rate Distortion Analysis Framework for Transform Coding", IEEE Trans. Circuits Syst. Video Technol., Vol. 11, No. 12, Dec. 2001