Complexity Reduced Mode Selection of H.264/AVC Intra Coding

Mohammed Golam Sarwer[1,2], Lai-Man Po[1], Jonathan Wu[2]
[1]Department of Electronic Engineering City University of Hong Kong Kowloon, Hong Kong
[2]Department of Electrical and Computer Engineering University of Windsor Windsor, ON, Canada N9B3P4
sarwer@uwindsor.ca, eelmpo@cityu.edu.hk, jwu@uwindsor.ca

Abstract

One of the new features in the H.264/AVC encoder is the use of Lagrangian Rate-Distortion Optimization (RDO) method during mode decision at the macroblock level. The RDO technique has been employed in H.264/AVC for intra prediction mode selection to achieve better coding efficiency. But the computation complexity of mode decision algorithm is extremely high. To reduce the complexity of mode decision, we propose an efficient and fast 4x4 intra prediction mode selection scheme. The proposed method reduces the candidate of the prediction modes based on the Sum of Absolute Transformed Difference (SATD) between the original block and the intra predicted block. Rank of each mode is obtained based on the SATD value. The candidate modes are further reduced by using the combination of rank and most probable mode. Simulation results demonstrate that the proposed mode decision method reduces about 95 % of mode decision time of intra coding with ignorable degradation of coding performance.

1. Introduction

H.264/AVC is the newest international video coding standard [1]. H.264/AVC offers very high compression ratios as compared to previous standards. To achieve the highest coding efficiency, H.264/AVC uses rate distortion optimization (RDO) technique. However, the computation load of H.264/AVC increase drastically compared to any previous standards. To reduce the complexity of intra 4x4 mode decision, H.264/AVC reference software suggested [2] 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 (RD) characteristics is not good enough. Several fast intra mode-decision approaches proposed in [3-6] focused on how to eliminate unnecessary modes. Meng et al. [3] selected optimal mode by computing a partial cost for down-sampled pixels instead of 4x4 block. To find the best intra mode, the information of the edge map of the whole frame is proposed in [5]. A fast intra mode decision method based on sum of absolute transform differences (SATD) between actual and predicted block is presented in [6] but the complexity reduction is not high. In this paper, we propose a fast mode decision scheme for intra 4x4 prediction in H.264/AVC based on SATD. Natural video sequences are highly spatially correlated. So the probability of mode of upper or left block to be the best mode of current block is high. Based on these observations, the proposed method reduces the number of candidate modes to either 1 or 2.

2. Overview of 4x4 intra prediction

2.1 Cost function for Intra 4x4 mode decision

To take the full advantages of all modes, the H.264/AVC encoder can determine the mode that meets the best RD tradeoff using RDO mode decision scheme. The best mode is the one having minimum rate-distortion cost and this cost is expressed as

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

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

To reduce the computational complexity, H.264/AVC reference software provides an SATD-based cost function:

$$J_{SATD} = SATD + \lambda_1 \cdot 4P \qquad (2)$$

where SATD is sum of absolute transformed difference between the original block and the predicted block , P is 0 for most probable mode and 1 for other modes.

2.2 Most probable mode

The choice of intra prediction mode for each 4x4 luma block must be signaled to the decoder and this could potentially require a large number of bits. For this reason, the best mode is not directly encoded into the compressed bit stream. Intra modes for neighboring 4x4 blocks are highly correlated. For current block, a mode is predicted based on the modes of upper and left blocks. This mode is defined as most probable mode. The most probable mode is inferred according to the following rules [7]. If the left neighboring block or the up neighboring block is unavailable, the most probable mode is set to 2(DC). Otherwise, the most probable mode is set to the minimum of the prediction mode of left neighboring block and up neighboring block.

3. Proposed mode decision scheme

3.1 Reducing the number of candidate modes

In general, exhaustive cost calculation for all possible modes is required to find an RD optimal mode. However in some cases, we can ignore some modes that are almost not used because they have a very little effect on encoding efficiency. It is reasonable to say that a good prediction should also produce a small value of the sum of absolute transform differences (SATD). A mode selected by SATD-based method can be a good indicator that the candidate mode is highly probable that it also can be selected by RDO method [6]. The mode which has the smallest SATD value has rank 1 and which has highest SATD value has rank 9. Another important thing mentioned in the previous section is that, when the prediction modes of the top and left block are known, we have a probability list of the mode to be chosen for present block. By using the H.264/AVC reference software JM9.6, we have done some experiments and analyze the mode distribution. The quantization parameter was set to 28. The results are shown in Table 1. In these experiments rank of mode is obtained when best mode is not same as the most probable mode. From the results, we observed that about 93% (on average) modes are either most probable mode or mode with rank 1 or mode with rank 2. So the percentages of remaining modes (from rank 3 to 9) are very low. Therefore, we can eliminate the modes with rank 3-9 because they have a very little effect on encoding efficiency.

Table 1. Percentage of mode distribution

Sequences	Most Proba ble mode	Rank		
		1	2	3 to 9
Akiyo (QCIF)	73.03	20.14	4.09	2.73
Foreman (QCIF)	58.68	28.70	7.00	5.62
Stefan (QCIF)	56.21	24.91	8.31	10.58
Carphone (QCIF)	68.16	22.46	5.35	4.04
Claire (QCIF)	80.97	13.80	3.18	2.04
Container (QCIF)	79.80	13.04	3.65	3.52
Mobile (QCIF)	37.32	36.38	12.3	13.94
Salesman (QCIF)	53.33	31.98	7.67	7.02
Paris (CIF)	67.27	21.33	6.01	5.39
Table_tennis (CIF)	64.72	20.34	6.50	8.44
Tempete (CIF)	48.22	28.73	10.4	12.58
Average	62.52	23.80	6.78	6.9

3.2 Algorithm of proposed mode decision

To reduce the further computation we use the early termination algorithm. If the best coding mode can be determined at an early stage of mode decision, significant time saving can be achieved. Based on most probable mode and rank, three different cases are described as follows.

3.2.1 Most probable mode is the mode associated with rank 1. It can observed that if most probable mode and mode with rank 1 are same, there is high chance of mode associated with rank 1 to be the best intra mode. To verify the above observation, extensive experiments have been conducted on different video sequences and at different quantization parameters (QP) to find out the statistics of coding modes in test video sequences. One way is examining the conditional probability of 9 different modes with different rank given that most probable mode is same as the mode with rank 1. It is called conditional probability because the condition "most probable mode is same as mode with rank 1" is given. In Fig. 1 X-axis represents the rank of mode and Y-axis shows their corresponding conditional probability. It is shown that conditional probability of event "best mode equal to mode with rank 1" is about 98%. From this statistics, it can be concluded that if most probable mode is the mode with rank1 then best intra mode is the mode associated with rank 1.

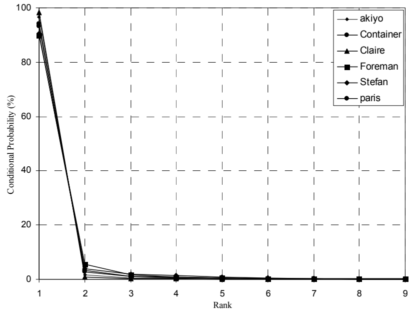

Fig 1. Conditional probability based on case 1

3.2.2 Most probable mode is the mode associated with rank 2. In this case most of the modes are within mode associated with either 1 or 2. Fig. 2 shows the conditional probability of 9 different modes with different rank. From Fig. 2 it is shown that only little numbers of modes are within rank from 3 to 9. So we can eliminate the modes associated with rank 3 to 9. Only two modes are selected as candidate mode.

To reduce further computation, we have used thresholding technique based on relative SATD values between mode with rank 1 and 2. It is reasonable to say that if the difference between the SATD value of mode with rank 1 and that of most probable mode (rank 2) is relatively small, the probability of the most probable mode (rank 2) to be the RDO mode is higher. The relative SATD value between most probable mode (rank 2) and mode with rank 1 is calculated as follows:

$$RSATD_{2->1} = \frac{SATD_2 - SATD_1}{SATD_2} \times 100\% \qquad (3)$$

where, $SATD_2$ and $SATD_1$ are the SATD value of mode with rank 2 and mode with rank 1, respectively. In this case, the algorithm of proposed mode decision scheme is given as follows:

if($RSATD_{2->1} < T_1$) Best mode = Mode with rank2;
else Best mode = Mode with rank1;

Now the factor is how to find out the threshold (T_1) value. In order to justify the proposed algorithm and to find out the threshold value, we have done some experiments with different video sequences at different QP values and observed the accuracy of mode decision with different values of threshold. By varying the value of threshold from 0 to 100, we have calculated the probability that a wrong mode is chosen for the current block. This probability is called as conditional error probability. Here first condition is the most probable mode is same as mode with rank 2 and another condition is the value of threshold is given. Fig. 3 shows variation of the probability of wrong mode decision with threshold value (T_1) for difference video sequences at QP=24. It is shown that for all of the sequences minimum error is found with threshold value of around 15 and 20. From similar experiments of different QP factors, it is shown that threshold values for minimum error are within range 15 to 25 for all QP factors and within this range variation of conditional error probability is very low and around the minimum value. Therefore, the threshold value (T_1) of this algorithm is any value within the range 15-25.

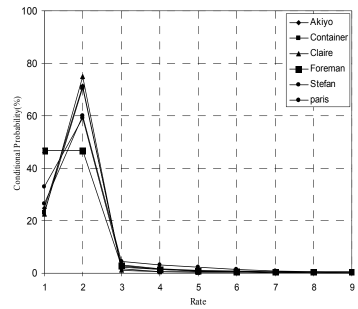

Fig 2. Conditional probability based on case 2

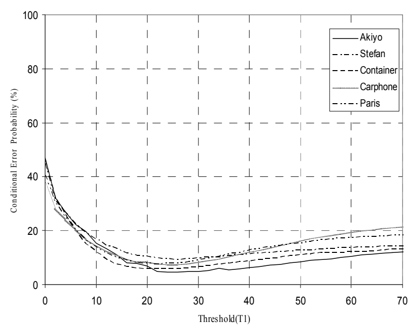

Fig 3. Variation of conditional error probability with threshold value (T_1) for case 2

3.2.3 Most probable mode is the mode associated with rank 3 to 9. Fig. 4 shows the conditional probability of different modes with the condition *"Most probable mode is the mode associated with rank 3 to 9"*. In this figure, most probable mode is represented as rank 0. It is clear that most of the modes are within rank 0 to 2. In order to further reduce the complexity, we have used some early termination algorithms. If SATD of rank 2 is much larger than SATD of rank1, then probability of the event *"Mode with rank1 is the best mode"* is higher. The relative SATD between mode with rank 1 and 2 is calculated in (3). It is reasonable to say that if the SATD value of mode with rank 1, 2 and most probable mode are more similar, the chance of most probable mode to be the best mode is higher. To measure the similarity, we have used deviation of SATD values between most probable mode, mode with rank 1 and mode with rank 2. The deviation (D) is defines as follows:

$$D = \sum_{rank=0}^{2} \left| SATD_{rank} - \mu \right| \quad (4)$$

where, μ is the mean SATD of mode with rank 1, 2 and most probable mode. Rank 0 represents the most probable mode.Algorithm of mode decision for case 3 is summarized as follows:

if($RSATD_{2->1}$> T_2) Best mode = Mode with rank1;
else if(D< T_3) Best mode = Most probable mode;
else Best mode = Mode which has minimum RD cost between mode with rank1 and 2;

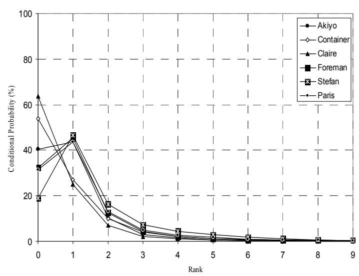

Fig 4. Conditional probability of case 3

In order to set the threshold T_2 and T_3, we have done several experiments for three different types of video sequences at different QP values. We observed that T_2 is quietly independent on both of the type of video sequence and QP values and better rate distortion performance was found at T_2=35. Threshold T_3 is also independent on the type of video sequence but depends on the QP values. Fig. 5 shows the variation of selected threshold T_3 with QP values of three different video sequences. A generalized threshold curve is formed by averaging the threshold values of all three sequences for each QP which is shown in Fig. 5. By using the polynomial fitting technique, the generalized threshold value T_3 is approximated as follows:

$$T_3 = 5.41 - 1.2QP + 0.06QP^2 \quad (5)$$

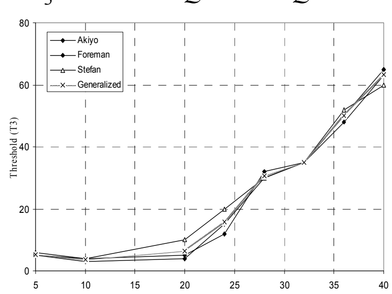

Fig 5. Threshold (T_3) of different QP

4. Simulation results

The proposed fast mode decision algorithm was tested using the first 100 frames of different video sequences with QCIF format. The experiment was carried out in the JVT JM9.6 [2] encoder and the test parameters are listed as follows: CABAC is enabled; Frame rate is 30; All frames are I frame; Intra 4x4 prediction only; Threshold (T_1) is 17;

The rate-distortion and complexity comparisons between the original encoder and the proposed mode decision method and J_{SATD} are tabulated in Table 2. Δ Psnr means the difference of PSNR between original encoder and encoder with proposed method. The positive values mean increments whereas negative values mean decrements. The average value of three different QPs is presented. ΔT % is the percentage of reduction of mode decision time as compared to original encoder. In case of J_{SATD} the average PSNR reduction is about 0.26 db and average bit rate increment is about 2.7%. Whereas in our proposed method, the average PSNR reduction is about 0.05 db and average bit rate increment is about 1.3%. So the coding results of proposed method are very similar to actual RDO method.

In the proposed method RD cost calculation is not required during case 1, case 2 and part of case 3. So in these conditions, computation of DCT, quantization, entropy coding, inverse quantization and inverse DCT have saved. The proposed algorithm reduced about 95% of mode decision time as compared to rate-distortion optimized cost function. For simple sequences (i.e. akiyo and Claire), the computation of proposed method is very similar with J_{SATD}. This is because these types of sequences has high spatially correlation. So most of the modes are fall within case 1 and case 2. For complex sequences (Stefan, mobile), the proposed algorithm reduces about 89% of mode decision time whereas J_{SATD} reduces about 96% of mode decision time. However, the rate-distortion performance of proposed method is much better as compared with J_{SATD}. Fig. 6 shows the RD curves of Claire sequence. It is shown that RD curve of proposed method close match with original RD curve. Similar results were found for other video sequences.

Table 2. Experimental results of proposed method

Sequence	J_{SATD}			Proposed		
	Δ	Δ	ΔT	Δ	Δ	ΔT

	Bit%	P_{snr}	(%)	Bit%	P_{snr}	(%)
Akiyo	2.56	-0.16	96.5	1.54	0.04	92.8
Claire	4.76	-0.12	96.1	2.38	0.11	93.9
Foreman	3.52	-0.17	96.2	2.08	0.01	92.2
Container	3.17	-0.19	96.3	1.95	-0.03	93.0
Stefan	0.78	-0.47	96.8	-0.3	-0.24	89.7
Mobile	1.27	-0.44	96.4	0.02	-0.21	89.3

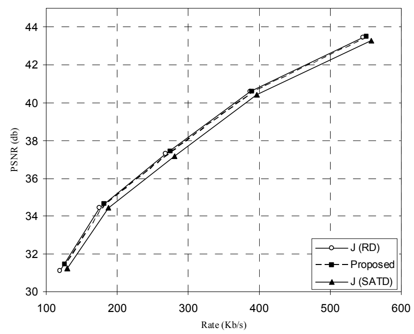

Fig 6. RD curves of Claire

Table 3. Comparison with fast mode decision method [5]

Sequence	PSNR increment in db	Bit-rate reduction (%)	Complexity reduction (%)
Akiyo	0.05	4.07	74.6
Claire	0.13	7.29	77.3
Foreman	0.06	0.39	70.7
Container	0.01	2.36	73.6
Stefan	-0.04	1.53	65.2
Mobile	-0.02	1.30	61.3

Table 3 shows the comparison of the proposed method with fast intra mode decision method described in [5]. The proposed method reduced the bit rate of about 2.8% and increases the PSNR of about 0.03 db on average. It is shown that for high motion sequences such as Stefan and Mobile, PSNR of proposed method is smaller than fast mode decision method in [5]. However, bit rate is also reduced; the resulting RD performance is very closed. The proposed method saves complexity of about 70% as compared to fast mode decision method [5].

5. Conclusions

In this paper, an efficient SATD based intra mode decision algorithm is presented. This algorithm is motivated by the fact that there is strong correlation between RDO cost and SATD. Also natural video sequences are highly spatially correlated. So the probability of mode of upper or left block to be the best mode of current block is high. For RD performance very close to the standard, the proposed scheme affords significant time saving as compared to H.264/AVC. The proposed technique reduces mode decision time by about 95% during intra frame coding of H.264/AVC.

6. Acknowledgment

The work described in this paper was substantially supported by a grant from Hong Kong SAR Government with CERG Projects 9041251.

7. References

[1] "Information Technology-Coding of audio-visual objects-part 10: Advanced video coding," Final Draft International Standard, ISO/IEC FDIS 14496-10, Dec. 2003.

[2] Joint Video Team (JVT) reference software version 9.6, http://iphome.hhi.de/suehring/tml/download/old_jm/

[3] B. Meng, O. Au, C. Wang, anfd H. Lam, "Efficient intra-prediction mode selection for 4x4 blocks in H.264." proceeding of ICME 2003, pp. III521-III524, Baltimore, USA.

[4] C. Kim, H. Shih, and C. Kuo, " Feature-based intra prediction mode decision for H.264," Proceeding of ICIP 2004, pp. 769-772, Oct. 2004, Singapore.

[5] Feng Pan, Xiao Lin, Susanto Rahardja, Keng Pang Lim, Z.G. Li, Dajun Wu, and Si WU, " Fast Mode Decision Algorithm for Intra-prediction in H.264/AVC Video Coding," IEEE Trans. Circuits Syst. Video Tech.., vol. 15, no. 7, pp-813-822, July 2005.

[6] Yeping Su, Jun Xin, Anthony Vetro, Hui fan sun, "Efficient MPEG-2 to H.264/AVC Intra Transcoding in Transform-domain", ISCAS 2005, pp-1234-1237.

[7] Iain E. G. Richardson, H.264 and MPEG-4 Video Compression – video coding for next generation multimedia, John Wily & Sons, 2003, pp. 180-181.