Most probable mode-based fast 4 × 4 intra-prediction in H.264/AVC

Conference Paper · January 2009

DOI: 10.1109/ICSPCS.2008.4813752 · Source: IEEE Xplore

ELectronics EXpress

Fast 4×4 intra-prediction based on the most probable mode in H.264/AVC

Byeongdu La[a], Jinwoo Jeong[b], and Yoonsik Choe[c]
Department of Electrical and Electronics Engineering, Yonsei University,
134 Sinchon-dong, Seodaemun-gu, Seoul, 120–749, Korea
a) *bdla6077@yonsei.ac.kr*
b) *691@yonsei.ac.kr*
c) *yschoe@yonsei.ac.kr*

Abstract: In this paper, we analyze that the best mode for intra-prediction can be changed by the RD cost related with the bit to signal the prediction mode according to whether the most probable mode (MPM) and the prediction mode are same or not. With this understanding, we propose a fast 4×4 intra-prediction based on the MPM in the H.264/AVC. This algorithm uses a defined RD cost including the minimum bit to signal the prediction modes excluded the MPM as a threshold. Experimental results show that the proposed algorithm is capable of reducing the overall encoding time by 36% and the overall intra-prediction time by 47% compared to the full search, with negligible loss of quality.

Keywords: H.264, intra prediction, mode probable mode
Classification: Integrated circuits

References

[1] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, "Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264 and ISO/IEC 14496-10 AVC)," March 2003.
[2] J. C. Wang, J. F. Wang, J. F. Yang, and J. T. Chen, "A fast mode decision algorithm and its VLSI design for H.264/AVC Intra-Prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 10, pp. 1414–1422, Oct. 2007.
[3] Y. N. Sairam, N. Ma, and N. Sinha, "A Novel Partial Prediction Algorithm for Fast 4x4 Intra Prediction Mode Decision in H.264/AVC," DCC 2008, 25–27 March 2008, pp. 232–241.
[4] G. Bjontegaard, "Calculation of average PSNR difference between RD-Curve," ITU-T Q.6/16, Doc. VCEG-M33, April 2001.

DOI: 10.1587/elex.5.782
Received July 31, 2008
Accepted August 25, 2008
Published October 10, 2008

IEICE Electronics Express, Vol.5, No.19, 782–788

1 Introduction

H.264/AVC [1] is the latest video coding standard developed by the Joint Video Team (JVT), which is an organization of the ISO Moving Picture Experts Group (MPEG) and the ITU-T Video Coding Experts Group (VCEG). Compared to the previous standards, H.264/AVC has many advanced characteristics such as 4×4 integer transformation, spatial intra-prediction, quarter-pixel motion compensation, and so on. These added characteristics permit H.264/AVC to achieve higher coding efficiency than other standards, especially by using intra-prediction and variable block size motion compensation. To maximize the coding efficiency, H.264/AVC uses the rate distortion optimization (RDO) technique. The RDO method checks inter/intra-predictions for I/P-pictures to choose the best coding mode. The computational complexity is increased by using RDO calculation in all possible modes of inter- and intra-predictions.

To reduce this complexity, a number of fast algorithms for inter- and intra-predictions have been proposed. In this paper, we focus only on intra-prediction. One of approaches to reducing complexity is based on the minimization of the candidate modes by using preprocessing. Wang et al. proposed a fast mode decision algorithm based on the dominant edge strength (DES) [2]. According to the DES, a subset of the prediction modes is chosen for RDO calculation. Sairam et al. proposed a 4×4 intra-mode decision algorithm exploiting inherent symmetry of the spatial prediction modes lately [3]. By using a cost measure derived from the symmetry in all prediction modes, one or 3 modes instead of 9 are searched for best mode. The preprocessing to minimize the candidate modes is an additional computation step that reduces the improvement for fast intra-prediction.

In this paper, we propose a fast 4×4 intra-prediction algorithm that uses the RD cost of the MPM. We analyze that the best mode for intra-prediction can be changed by the RD cost that is affected according to whether the MPM and the prediction mode are same or not. With this understanding, a defined RD cost with a specified bit is used as a threshold for fast intra-prediction. When the RD cost of the MPM is less than the threshold, the MPM is determined with the best mode. If not, the full search continues through eight prediction modes excluding the MPM (hereafter referred to as the non-MPM). The intra-prediction is terminated when any of the non-MPMs satisfies the threshold.

In Section 2, we present an overview of the MPM on the 4×4 intra-prediction. In Section 3, we introduce a fast 4×4 intra-prediction algorithm. We present the performance evaluations of the algorithm in Section 4 and our conclusions are given in Section 5.

2 MPM on the 4×4 intra-prediction in H.264/AVC

To select the best mode for the 4×4 intra-prediction, the residual between the current block and its predicted block according to the nine prediction modes is transformed, quantized, entropy-coded, and reconstructed. The mode that

DOI: 10.1587/elex.5.782
Received July 31, 2008
Accepted August 25, 2008
Published October 10, 2008

IEICE Electronics Express, Vol.5, No.19, 782–788

has the minimum RD cost among all the prediction modes is selected as the best mode. The MPM is used to reduce coding bits for coding efficiency in the 4×4 intra-prediction. The encoder estimates the MPM for the current block based on the availability of neighboring blocks and the magnitude of the best modes of two neighboring blocks. If one of the two neighboring blocks of the current block is not available, the MPM is decided as the DC (Non-edge) mode. If the two neighboring blocks of the current block are available, the least number of the two best modes is chosen as the MPM. For intra-prediction according to each prediction mode, the encoder uses the condition of the MPM with a flag to signal the prediction mode. If the MPM is the same as the prediction mode, the flag is set to "1" and only one bit is needed to signal the prediction mode. When the MPM and the prediction mode are different, the flag is set to "0" and an additional signal called the remaining mode selector is sent to indicate the prediction mode. To describe the all of prediction modes, four bits are required, one to indicate the MPM flag and three others to indicate the eight prediction modes except for the MPM.

3 Proposed fast mode decision algorithm

3.1 Method 1: Fast intra-prediction using the relation between the MPM and the non-MPMs

The coding efficiency for 4×4 intra-prediction is affected by the relationship between the MPM and the prediction mode. By examining this relationship, we can obtain the information required for fast intra-prediction. To select the best mode for intra-prediction, H.264/AVC uses the RDO method as a measure. The general equation for this is

$$J(MODE|QP) = D(MODE|QP) + \lambda \times R(MODE|QP) \tag{1}$$

where J is the RD cost, D is a measure of the distortion between the original 4×4 block and the reconstructed block for each prediction mode, $MODE$ is one of nine prediction modes selected for intra-prediction, and R is the number of bits. The λ is the Lagrange parameter, which is described by

$$\lambda = 0.85 \times 2^{(QP-12)/3} \tag{2}$$

The best intra-prediction mode has the least RD cost among the nine prediction modes including the MPM as previously noted. The value of RD cost changes depending on whether the MPM and the prediction mode are the same when the distortion is constant. This means that the best mode for intra-prediction can be changed depending on whether the MPM and the prediction mode are the same. When the MPM and the prediction mode are the same, the RD cost can be described as Eq. (3) by using Eq. (1). And for the bitstream, at least 2 bits are required, 1 to indicate the prediction mode and the other to denote the minimum value 0 of the residual (In this case, $R_{residual} = 1$) that the difference between the original 4×4 block and the

DOI: 10.1587/elex.5.782
Received July 31, 2008
Accepted August 25, 2008
Published October 10, 2008

predicted block is quantized.

$$J_{MPM} = D_{MPM} + \lambda \times (1 + R_{residual}),\ if(MPM = PedictionMode) \quad (3)$$

When the MPM and the prediction mode are different (denoted as non-MPM), at least 5 bits are required for the bitstream, 4 to indicate the 8 possible prediction modes except for the MPM as mentioned previously, and 1 to indicate the minimum value 0 of the residual. So for the case in which the MPM and the prediction mode are different, the minimum RD cost is expressed using $D = 0$ and $R = 5$, and Eq. (1) becomes

$$J_{Min\ of\ non\_MPM} = 0 + \lambda \times 5,\ if(MPM \neq PedictionMode) \quad (4)$$

This means that if the RD cost of any of the nine prediction modes is less than Eq. (4), the mode is acceptable as the best mode because the distortion is zero in the case of the non-MPM, or at least smaller than that of the non-MPM in the case of the MPM. Therefore, the MPM satisfying this condition can be determined with the best mode. By applying the minimum RD cost of the non-MPMs as a threshold, we can expect to save time on the intra-prediction while the RD performance of the full search will not change regardless of any video sequences. We define this threshold with a bit of 5, which is the minimum value for the non-MPM in Threshold 1 and it can be described as

$$Best\ mode = MPM,\ if\ J_{MPM} \langle\ (J_{Threshold1} = \lambda \times 5) \quad (5)$$

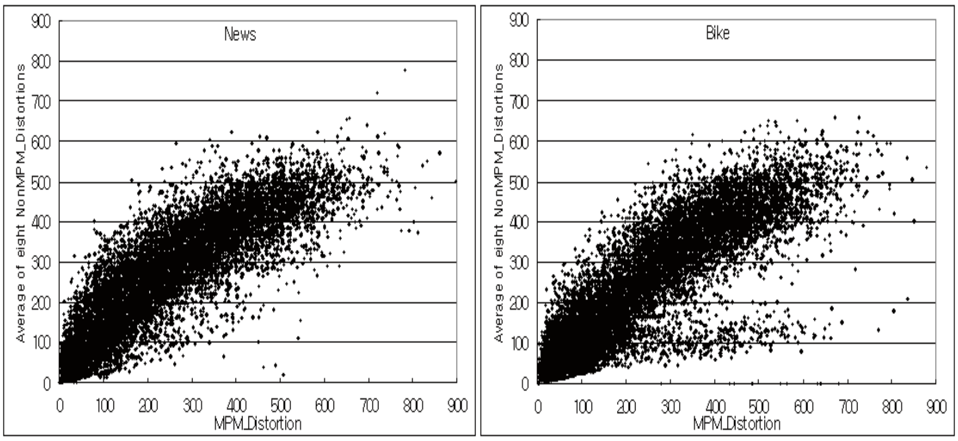

Fig. 1. Distribution of the distortion for the MPM and the average distortion of the non-MPMs on the 4×4 block (IIII type, QCIF Sequence, Frame rate: 30, QP: 28).

3.2 Method 2: An additional threshold for fast intra-prediction based on the distortion of the MPM and the non-MPMs

The best mode for intra-prediction is selected with the prediction mode that has the least distortion when the bit to signal the prediction mode is constant as describe in Eq. (1). Thus, the relationship between the distortion

DOI: 10.1587/elex.5.782
Received July 31, 2008
Accepted August 25, 2008
Published October 10, 2008

IEICE Electronics Express, Vol.5, No.19, 782–788

of the MPM and the distortion of the non-MPM is important in determining the reliability of distortion of the MPM when using the distortion of the MPM as an additional threshold for fast intra-prediction. The distribution between the distortion of the MPM (X-coordinate) and the average distortion of the non-MPMs (Y-coordinate) according to each block on the two video sequences using the full search is shown in Fig. 1. In the figure, the average distortion of the non-MPMs is measured among all values of bit that the non-MPMs have on each block and one point indicates a 4×4 block. As the distortion of the MPM increases, the average distortion of the non-MPMs also increases. Therefore, we can use the distortion of the MPM as an additional threshold for fast mode decision by using the linear characteristics between two distortions. To control the distortion of the MPM for fast intra-prediction, we define an additional parameter as sigma (σ). By using this sigma, we can control the distortion of the MPM according to the desired performance. In the Eq. (1), the bit related with the RD cost varies according to the variety of prediction modes when the distortion is constant. This means that the RD cost can be affected with prediction modes and the decision for best mode can be miss-selected when the bit is not fixed for fast intra-prediction. Therefore, to describe this sigma with the RD cost equation for fast intra-prediction, we fix the bit at 5 because the minimum bit that the non-MPMs can have is 5 as used in Method 1. With this definition, the threshold equation can be described as

$$J_{Threshold} = D_{MPM} \times \sigma + \lambda \times 5 \tag{6}$$

As the value of sigma increases, the number of block for fast mode decisions also increases and cause more degradation of performance accompanied by more saved time. The sigma value can be changed in accordance with the desired performance. In this paper, we set the value of sigma at 0.9; these values were determined according to experiments with a variety of video sequences. Using a sigma of 0.9 as an additional threshold is defined as Threshold 2. Therefore by applying the sigma value for fast mode decision as an additional threshold, we can rewrite Eq. (6) as

$$Best\ mode = MPM, if\ J_{MPM} \langle\ (J_{Threshold2} = D_{MPM} \times \sigma + \lambda \times 5,\ \sigma = 0.9) \tag{7}$$

Unlike the full search of JM reference software, which searches the nine prediction modes from zero through eight, this algorithm starts the prediction with the MPM. If the RD cost of the MPM is satisfied with the early termination condition (i.e., the threshold), the intra-prediction is terminated with the MPM as the best mode. If not, the full search continues through eight prediction modes excluding the MPM. The intra-prediction is terminated when any of the eight prediction modes satisfies the threshold for early termination.

DOI: 10.1587/elex.5.782
Received July 31, 2008
Accepted August 25, 2008
Published October 10, 2008

Table I. Evaluation of the QCIF/CIF sequence (IIII type).

Res.	Sequence	DES[2]				Sairam et al. [3]				Method 1				Method 2			
		ΔT_E	ΔT_I	Δ PSNR	ΔBit	ΔT_E	ΔT_I	Δ PSNR	ΔBit	ΔT_E	ΔT_I	Δ PSNR	ΔBit	ΔT_E	ΔT_I	Δ PSNR	ΔBit
QCIF	News	-23.7	-36.2	-0.06	8.0	-40.8	-49.8	-0.05	3.6	-26.1	-34.9	0.00	0.0	-34.2	-44.6	-0.06	0.1
	Bike	-25.8	-40.0	-0.07	7.1	-32.3	-45.9	-0.07	5.0	-32.7	-44.0	0.00	0.0	-43.6	-60.3	-0.07	0.2
	Carphone	-24.6	-37.7	-0.05	10.0	-30.1	-39.9	-0.07	5.6	-28.4	-35.9	0.00	0.0	-38.0	-47.4	-0.03	0.4
	Salesman	-24.6	-37.6	-0.06	7.8	-21.6	-36.2	-0.08	3.3	-21.9	-29.1	0.00	0.0	-30.3	-40.9	-0.06	0.1
	Container	-22.8	-35.2	-0.04	6.9	-33.8	-50.9	-0.07	2.7	-27.4	-38.5	0.00	0.0	-35.6	-46.7	-0.04	0.1
	Foreman	-24.3	-38.7	-0.07	11.6	-24.1	-34.8	-0.16	12.9	-24.6	-31.1	0.00	0.0	-33.6	-43.5	-0.12	0.7
	Table tennis	-23.7	-36.3	-0.06	4.1	-28.7	-44.8	-0.09	3.0	-26.5	-36.7	0.00	0.0	-35.1	-46.3	-0.04	0.0
	Average	-24.2	-37.4	-0.06	7.9	-30.2	-43.2	-0.08	5.2	-26.8	-35.7	0.00	0.0	-35.8	-47.1	-0.06	0.2
CIF	Paris	-23.2	-34.0	-0.05	9.0	-28.6	-39.9	-0.07	3.0	-22.3	-26.4	0.00	0.0	-30.1	-39.4	-0.05	0.1
	Hall	-22.1	-33.5	-0.04	13.6	-35.8	-49.5	-0.04	3.9	-29.5	-37.2	0.00	0.0	-40.9	-53.8	-0.06	0.2
	Mobile	-22.3	-32.9	-0.08	4.6	-20.7	-28.0	-0.10	2.5	-18.4	-20.0	0.00	0.0	-22.7	-27.6	-0.03	0.0
	Salesman	-23.4	-35.3	-0.05	9.6	-25.9	-37.9	-0.05	4.2	-23.1	-26.2	0.00	0.0	-33.2	-42.5	-0.07	0.2
	Suzie	-21.6	-31.7	0.00	5.2	-30.6	-41.2	0.01	3.3	-35.0	-43.6	0.00	0.0	-45.8	-60.4	-0.06	0.3
	Bream	-0.7	-0.2	-0.05	2.7	-55.4	-73.6	-0.06	4.1	-52.4	-71.6	0.00	0.0	-57.4	-78.6	-0.10	0.1
	Football	-23.7	-34.9	-0.08	4.5	-16.4	-21.3	-0.09	3.0	-19.1	-22.1	0.00	0.0	-26.6	-32.5	-0.05	0.0
	Average	-19.6	-28.9	-0.05	7.0	-30.5	-41.6	-0.06	3.4	-28.5	-35.3	0.00	0.0	-36.7	-47.8	-0.06	0.1

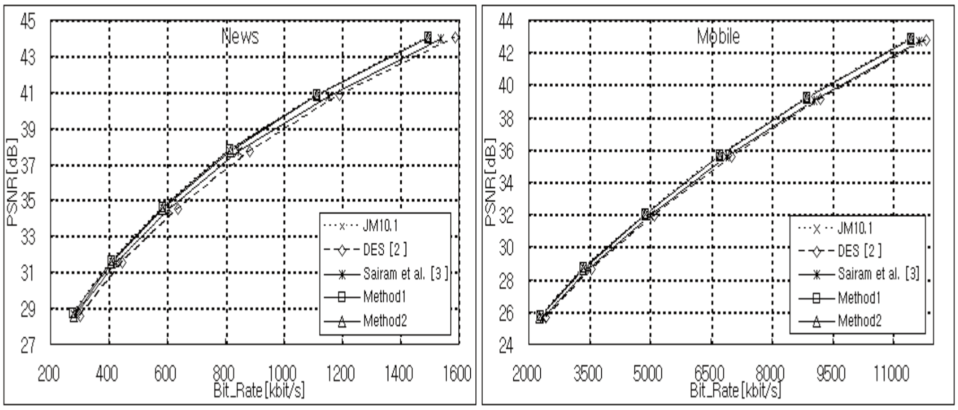

Fig. 2. RD curve for the QCIF "News" sequence and CIF "Mobile" sequence.

4 Performance evaluations

For evaluation, the proposed algorithm, the previously proposed DES-based algorithm [2], the previously proposed Sairam et al.'s algorithm [3], and the full search were simulated on QCIF and CIF sequences in the JM10.1 reference software. For the DES-based algorithm, we used only 4×4 intra-prediction to compare its performance with other algorithms on the same environment. We used the Bjonteggard delta peak signal-to-noise ratio (PSNR) and Bjonteggard delta bit rates [4] to compare the performance. All of the test sequences were coded in intra-only sequence type (denoted as IIII type) for six quantization values (20, 24, 28, 32, 36, and 40), and one hundred frames were coded in every sequence. A negative value of ΔT_E and ΔT_I indicates the percent of time saved during the entire encoding and entire intra-prediction compared to the full search by each.

The experimental results are shown in Table I. With Method 1, the proposed method reduces the intra-prediction time by 35% on average, while PSNR and bit rates are same with the full search. For Method 2, the intra-prediction time and PSNR decreased by 47% and 0.06 dB respectively, while the bit rates only increased by about 0.2%. Our proposed algorithm outper-

DOI: 10.1587/elex.5.782
Received July 31, 2008
Accepted August 25, 2008
Published October 10, 2008

forms the two previously proposed algorithms in time saving and bit rates while maintaining same visual quality. Figure 2 shows the RD curve for the QCIF "News" sequence and the CIF "Mobile" sequence with QP values of 20, 24, 28, 32, 36, and 40. These results indicate that Methods 2 maintains the full search performance with slight loss of quality or increase in bitrate and outperforms the two proposed algorithms.

5 Conclusions

In this paper, we proposed a fast 4×4 intra-prediction mode decision algorithm based on the MPM. This algorithm selects the best mode using the RD cost of the MPM and thus requires no additional computation. Simulation results show that the MPM-based algorithm using Method 2 reduces the intra-prediction time by about 47% over the full search of JM reference software on average, with almost same level of bit rates. Also Method 1 can be combined with other algorithms without any quality loss.

DOI: 10.1587/elex.5.782
Received July 31, 2008
Accepted August 25, 2008
Published October 10, 2008