FAST MODE DECISION FOR INTRA PREDICTION IN H.264/AVC ENCODER

Byeongdu La, Minyoung Eom, Yoonsik Choe
School of Electrical and Electronics Engineering, Yonsei University, Seoul, Korea
bdla6077@yonsei.ac.kr, emy0606@yonsei.ac.kr, yschoe@yonsei.ac.kr

ABSTRACT

The H.264/AVC video coding standard uses the rate distortion optimization (RDO) method to improve the compression performance in the intra prediction. Whereas the computational complexity is increased comparing with previous standards due to this method, even though this standard selects the best coding mode for the current macroblock. In this paper, a fast intra mode decision algorithm for H.264/AVC encoder based on dominant edge direction (DED) is proposed. The algorithm uses the approximation of discrete cosine transform (DCT) coefficient formula. By detecting the DED before intra prediction, 3 modes instead of 9 modes are chosen for RDO calculation to decide the best mode in the 4x4 luma block. For the 16x16 luma and the 8x8 chroma, instead of 4 modes, only 2 modes are chosen. Experimental results show that the computation time of the proposed algorithm is decreased to about 71% of the full search method in the reference code with negligible quality loss.

Index Terms— H.264, Intra-prediction, Mode Decision

1. INTRODUCTION

H.264/AVC was developed by JVT (Joint Video Team), VCEG (Video Coding Expert Group) of ITU-T and MPEG (Moving Picture Expert Group) of ISO/IEC [1]. This standard contains several new techniques, such as spatial prediction in intra coding, adaptive block size motion compensation, 4x4 integer transformation, multiple reference pictures and content adaptive binary arithmetic coding (CABAC) [2]. The video coding standard achieves considerably higher coding efficiency than previous standards such as MPEG-4 by using intra prediction and variable block size motion compensation. To minimize the temporal and spatial redundancy, H.264 uses the RDO technique. The RDO technique checks inter/intra prediction for I/P-picture to choose the best coding mode. The rate-distortion (RD) cost computation for intra mode is conducted according to the block sizes and the edge modes about chroma and luma. The RD cost computation for motion prediction is also conducted according to each block size and each reference frame. The computational complexity by using RDO calculation in all possible modes of inter and intra predictions is the critical issue point for real time encoding. With this demand, a number of fast algorithms of inter and intra predictions have been proposed up to date. In this paper, intra prediction is only focused. Pan et al. proposed an intra prediction using edge detection based on Sobel operator [3]. This algorithm calculates the edge direction histogram for luma 4x4, luma 16x16 and chroma 8x8 respectively, and accumulates the edge direction by each pixel in a block. Then the maximum accumulation direction becomes the best edge mode. Wang et al. proposed another algorithm using edge histogram descriptor [4]. This algorithm detects the dominant edge strength (DES) by using edge histogram descriptor, and a subset of the prediction modes is chosen for RDO calculation. These two algorithms reduce the computation time about 55~60% of the full search time. Although these two algorithms reduce much complexity, intra mode decision requires still much computation.

In this paper, a fast mode decision algorithm for intra prediction by detecting dominant edge direction that approximates the DCT coefficient is proposed. The number of search modes for RDO calculation is reduced from nine to three for luma 4x4 in this algorithm. For luma 16x16 and chroma 8x8, the search modes are reduced from four to two as Wang's algorithm. Compared with Pan's algorithm and Wang's algorithm, our DED-based algorithm requires the shortest computation time at about 71% reduction of the full search time. The PSNR degradation and Bit/Rate increase are acceptable. The edge direction calculation may be so simple that it can be implemented in VLSI hardware architecture.

The rest of this paper is composed as follows. In Section 2, overview of intra prediction performed in H.264 is described. In Section 3, we introduce a fast intra mode decision algorithm based on the dominant edge direction. The Performance evaluations of algorithm are demonstrated in Section 4 and conclusions are described in Section 5.

2. OVERVIEW OF INTRA PREDICTION

H.264 uses the intra mode decision for intra frame coding as a new technique. The luma intra prediction has the 4x4 prediction mode and the 16x16 prediction mode. The 4x4 prediction mode predicts each 4x4 luma block of sixteen

"This research was supported by the MIC (Ministry of Information and Communication), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment)" (IITA-2006-(C1090-0603-0011))

4x4 luma blocks in a macroblock. The intra 16x16 prediction searches the entire 16x16 luma blocks. For chroma prediction, chroma 8x8 mode is only used. Fig.1 shows a block that has its neighboring pixels and 4x4 intra mode that has the nine directions. The a-p samples of a block are predicted using samples A-M. The number 0, 1, 3,…, 8 indicate the vertical, horizontal, diagonal down-left,

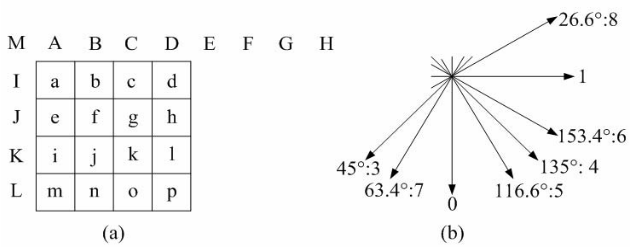

Fig. 1. (a). A 4x4 block and its neighboring pixels. (b) nine 4x4 intra prediction modes including DC.

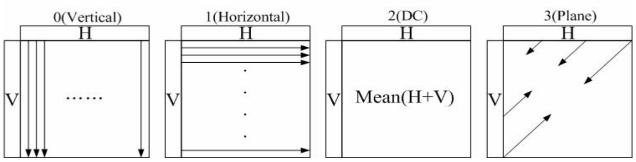

Fig. 2. 16x16 luma prediction modes

vertical-right, horizontal-down, vertical-left and horizontal-up, respectively. And the number 2 indicates the DC prediction. Fig. 2 shows four directions of intra 16x16 mode. The chroma intra prediction has 8x8 prediction mode. The chroma U and V mode are predicted by one mode. The decision for chroma edge direction is very similar to the intra 16x16 prediction except that the order of mode is different. DC, horizontal, vertical and plane are represented by the number 0, 1, 2, 3, respectively.

3. FAST MODE DECISION ALGORITHM

AC coefficients of DCT reflect variations in certain direction [5][6]. Using this property, we propose a fast algorithm based on the dominant edge direction. The DCT formula is as below. M and N indicate the vertical size and the horizontal size, respectively.

$$AC_{u,v} = C_u C_v \frac{2}{\sqrt{MN}} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(i,j) \cos\frac{(2i+1)u\pi}{2M} \cos\frac{(2j+1)v\pi}{2N} \quad (1)$$

$$C_u, C_v = \begin{cases} \sqrt{\frac{1}{2}}, for \quad u, v = 0 \\ 1, otherwise \end{cases}$$

From (1), $AC_{0,1}$ and $AC_{1,0}$ can be described as (2) and (3).

$$AC_{0,1} = C_0 C_1 \frac{2}{\sqrt{MN}} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \cos\frac{(2j+1)\pi}{2N} f(i,j) \quad (2)$$

$$AC_{1,0} = C_1 C_0 \frac{2}{\sqrt{MN}} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \cos\frac{(2i+1)\pi}{2M} f(i,j) \quad (3)$$

For 4x4 block, the formula can be represented as below.

$$AC_{0,1} = \frac{1}{2\sqrt{2}} \sum_{i=0}^{3} \sum_{j=0}^{3} \cos\frac{(2j+1)\pi}{8} f(i,j) \quad (4)$$

$$AC_{1,0} = \frac{1}{2\sqrt{2}} \sum_{i=0}^{3} \sum_{j=0}^{3} \cos\frac{(2i+1)\pi}{8} f(i,j) \quad (5)$$

For simplification, (4) and (5) can be approximated as below.

$$AC_{0,1} = \sum_{i=0}^{3} f(i,0) - \sum_{i=0}^{3} f(i,3) + \sum_{i=0}^{3} f(i,1) - \sum_{i=0}^{3} f(i,2) \quad (6)$$

$$AC_{1,0} = \sum_{j=0}^{3} f(0,j) - \sum_{j=0}^{3} f(3,j) + \sum_{j=0}^{3} f(1,j) - \sum_{j=0}^{3} f(2,j) \quad (7)$$

a	b	c	d
e	f	g	h
i	j	k	l
m	n	o	p

A= a+b+c+d
B= e+f+g+h
C= i+j+k+l
D= m+n+o+p
I = a+e+i+m
J = b+f+j+n
K = c+g+k+o
L = d+h+l+p

Fig. 3. Pixel summation in 4x4 prediction.

The formula (6) and (7) can be described as shown in Fig.3. $AC_{0,1}$ and $AC_{1,0}$ are expressed as (8), (9),(10) and (11).

$$AC_{0,1} = I - L + J - K \quad (8)$$

$$AC_{1,0} = A - D + B - C \quad (9)$$

$$AC\_MUL = AC_{0,1} * AC_{1,0} \quad (10)$$

$$abs(AC_{0,1}) = |I - L + J - K|, \quad abs(AC_{1,0}) = |A - D + B - C| \quad (11)$$

$$Condition \ : abs(AC_{0,1}) > abs(AC_{1,0}) \quad AC\_DIV01 = abs(AC_{0,1}) / abs(AC_{1,0}) \quad (12)$$

$$Condition \ : abs(AC_{1,0}) > abs(AC_{0,1}) \quad AC\_DIV10 = abs(AC_{1,0}) / abs(AC_{0,1}) \quad (13)$$

AC coefficients reflect variations in certain direction as mentioned above. Especially $AC_{0,1}$ and $AC_{1,0}$ only depend on the vertical and horizontal intensity difference, respectively. Therefore, by using the ratio of $AC_{0,1}$ and $AC_{1,0}$ as (12) and (13), the dominant edge direction can be extracted. AC_DIV01 and AC_DIV10 can be considered as the dominant edge directions (DED) in the current block.

Fig. 4. The proposed algorithm of 4x4 block.

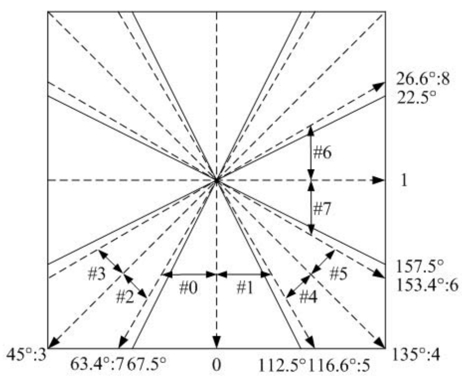

Fig. 5. The relations between dominant edge direction and 4x4 intra modes.

Table 1. Search modes by the DED of 4x4 block.

DED	Search Mode	DED	Search Mode
#0	0,2,7	#4	2,4,5
#1	0,2,5	#5	2,4,6
#2	2,3,7	#6	1,2,8
#3	2,3,8	#7	1,2,6

After deciding the dominant edge direction, RD cost calculation for three remaining modes can be conducted to select the best mode. The 4x4 intra mode decision algorithm in Fig. 4 can be extracted using the relations between the dominant edge direction and intra mode in Fig 5. The value 1.997 comes from the value of $\tan 63.4°$. Table 1 shows the search modes by DED.

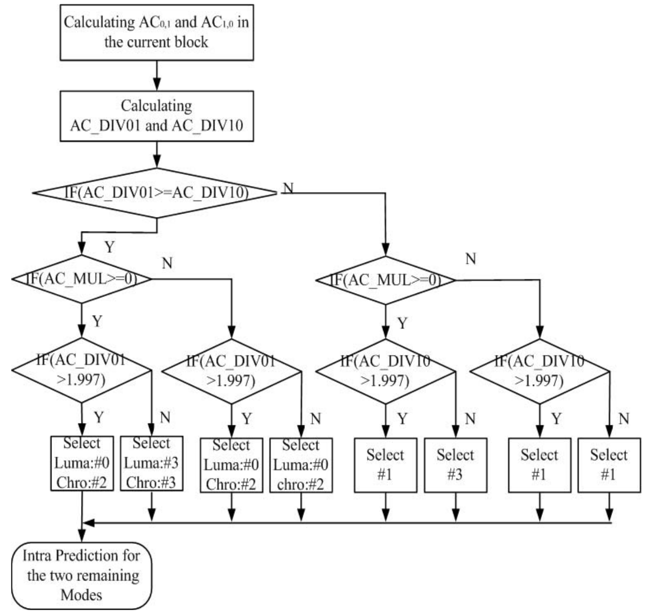

Fig. 6. The proposed algorithm of 16x16 luma block and 8x8 chroma block.

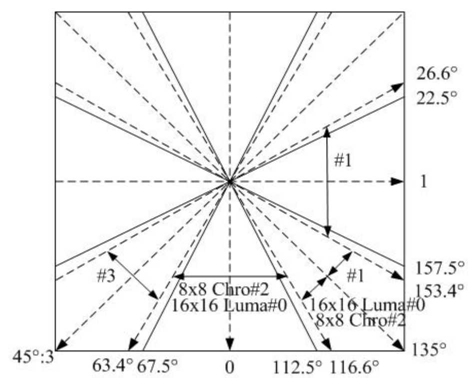

Fig. 7. The relations between dominant edge direction and intra modes (16x16 luma and 8x8 chroma).

Table 2. Search modes by the DED of 8x8 and 16x16 block.

Luma DED	16x16 luma Search Mode	Chroma DED	8x8 chroma Search Mode
#0	0,2	#1	0,1
#1	1,2	#2	0,2
#3	2,3	#3	0,3

For 16x16 luma and 8x8 chroma, the decision method is similar to the 4x4 block. After selecting the DED of their block size, RD cost calculation for two remaining modes is conducted to select the best intra mode as Table 2. The 16x16 and 8x8 intra mode decision algorithm in Fig. 6 can be extracted using the relations between the dominant edge direction and intra mode in Fig 7.

4. THE PERFORMANCE EVALUATIONS

The proposed algorithm was implemented on JM10.1 provided by JVT. The test conditions are as follows.

1). Test Image: CIF, QCIF 300frames
QCIF : Foreman, News,Container, Silent, Coastguard
CIF : Paris, Mobile, Tempete, Stefan
2). RDO: On, Rate Control: Off, Intra Period: 1, QP: 28
3). ΔPSNR : Comparing by full-search(+/- : Inc./Dec.)
4). $\Delta B / R(\%) = \frac{Bit-rate(proposed) - Bit-rate(full\_search)}{Bit-rate(full-search)} \times 100$
5). $\Delta Time\ (\%) = \frac{Time\ (proposed) - Time\ (full-search)}{Time\ (full-search)} \times 100$

Table 3 and 4 list the evaluation results of the proposed and Wang's DES algorithm by comparing with full search method in PSNR, Bit-rate and computation time. The proposed algorithm (DED) saves about 71% of computation time of full search method. The proposed outperforms the previous algorithm in computational time with a negligible loss of PSNR and Bit-rate. The dominant edge direction method of proposed algorithm uses only three modes that are smaller than the DES-based algorithm in the 4x4 block. The rate distortion curves for the image sequences "News" and "Mobile" are described in Fig.8 and 9.

Table 3. Evaluation Results for QCIF Sequences.

Sequence	DES[4]			DED(Proposed Algorithm)		
	ΔPSNR [dB]	ΔB/R [%]	ΔTime [%]	ΔPSNR [dB]	ΔB/R [%]	ΔTime [%]
Foreman	-0.031	2.136	-59.33	-0.088	4.847	-71.32
News	-0.045	3.823	-58.99	-0.130	5.550	-71.03
Container	-0.044	3.583	-59.95	-0.090	4.437	-71.76
Silent	-0.041	2.982	-60.47	-0.122	5.441	-71.21
Coastguard	-0.033	3.648	-62.76	-0.088	5.405	-72.29
Average	-0.039	3.235	-60.30	-0.104	5.136	-71.52

Table 4. Evaluation Results for CIF Sequences.

Sequence	DES[4]			DED(Proposed Algorithm)		
	ΔPSNR [dB]	ΔB/R [%]	ΔTime [%]	ΔPSNR [dB]	ΔB/R [%]	ΔTime [%]
Paris	-0.046	2.754	-62.25	-0.102	5.407	-72.66
Mobile	-0.069	1.657	-59.31	-0.211	3.321	-72.83
Tempete	-0.073	2.275	-61.38	-0.185	4.019	-72.75
Stefan	-0.077	2.545	-61.79	-0.191	4.226	-72.89
Average	-0.066	2.308	-61.18	-0.172	4.2432	-72.780

5. CONCLUSIONS

In this paper, a fast mode decision algorithm for intra prediction in H.264 is proposed. The proposed algorithm uses the dominant edge direction induced from DCT formula. By using DED detection, the number of candidates for RDO calculation is decreased from 592 to 100. The evaluation results show that the proposed algorithm saves 71% of computational time of the full search method in the reference code with a negligible loss of quality and Bit-rate. By the reduction of computational complexity, this algorithm can be used for real time encoding or hardware design.

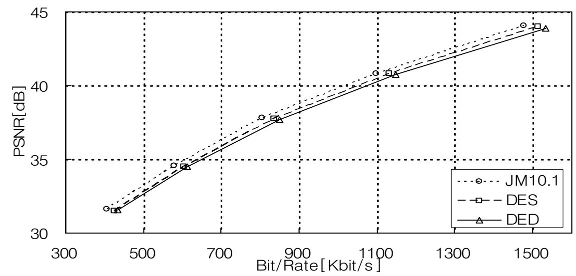

Fig. 8. The RD curve of the QCIF sequence "News"

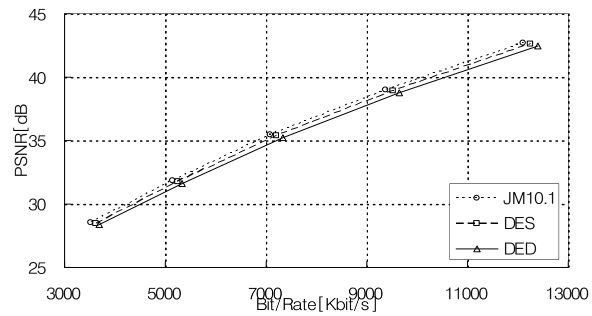

Fig. 9. The RD curve of the CIF sequence "Mobile"

6. REFERENCES

[1] Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra, "Overview of the H.264/AVC Video Coding Standard," IEEE Transaction on Circuits and System for Video Technology, Vol. 13, No.7, pp. 560-576, July 2003.

[2] Iain E.G. Richardson, "H.264 and MPEG-4 Video Compression," The Robert Gordon University, Aderdeen, 2004.

[3] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. Wu, and S. Wu, "Fast mode decision algorithm for intra prediction in H.264/A VC video coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, no. 7, pp. 813-822, July 2005.

[4] Jhing-Fa Wang, Jia-Ching Wang, Jang-Ting Chen, An-Chao Tsai, and Anand Paul, "A Novel Fast Algorithm for Intra Mode Decision in H.264/AVG Encoders," ISCAS2006, pp.3498-3501, July 2006.

[5] Seong-Whan Lee, Young-Min Kim, and Sung Woo Choi, "Fast Scene Change Detection using Direct Feature Extraction from MPEG Compressed Videos," IEEE Transactions On Multimedia, Vol. 2, No. 4, pp.240-254, Dec 2000.

[6] Minyoung Eom, and Yoonsik Choe, "Fast Extraction of Edge Histogram in DCT Domain based on MPEG7," Transactions On Engineering, Computing and Technology V9, pp. 209-212, Nov 2005.