

H.264 FAST INTRA MODE SELECTION ALGORITHM BASED ON DIRECTION DIFFERENCE MEASURE IN THE PIXEL DOMAIN

Li-Li Wang and Wan-Chi Siu

Department of Electronic and Information Engineering
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

ABSTRACT

In this paper, a fast mode decision algorithm for Intra prediction in the H.264/AVC is proposed. We use the characteristics of each directional prediction mode to compute the strength of directional differences in the original pixel domain to find the minimal direction error. This is the first time reported in the literature that the intrinsic differences between the real-data and the predictors of modes are used to form an algorithm for mode decision. The approach allows us to select several better candidate modes for evaluation instead of using the full search. Experimental results show that the proposed method can achieve more than 80% reduction in computation with negligible degradation in rate-distortion performance, and the results are better than other algorithms available in the literature.

Index Terms - H.264/AVC, intra prediction, fast mode decision, RDO, prediction error.

1. INTRODUCTION

The H.264 is a current international standard for advanced video coding [1]. It has gained much attention and allows many applications. A number of advanced coding techniques are available in the H.264, including context-based arithmetic coding (CABAC), integer cosine transform (ICT), and various modes and mode types for intra prediction in I-MB coding. Its coding efficiency is significantly better than the previous standards. Among these new features, inter prediction with fast motion estimation algorithms [2-4] and the intra prediction techniques [5-8] are considered as the main factors that contribute to the success of H.264/AVC. In order to select the best prediction mode, H.264 employs the Lagrangian rate distortion optimization (RDO) method, while it is very complex and time-consuming. To reduce the computational complexity, many algorithms have been proposed. In [8], the authors proposed spatial filters to compute the DES (dominant edge strength), which is used to eliminate the unlikely modes, only a small part of intra prediction modes were chosen for RDO calculation, using this idea, much computational time is saved. We have realized many algorithms[5,6,8], Ref.8 gives relatively good results besides our approach.

In this paper, we propose some fast and efficient techniques to form an algorithm to simplify the selection process of intra prediction. In the new method, we compute the strength of direction differences of each intra direction prediction mode according to the characteristics of intra prediction modes. The mode with the minimum direction difference strength will be selected as one of the best four candidate modes. Compared with previous algorithms, the advantage of our algorithm has more accurate direction decision strategy, which can be demonstrated by experimental results.

The rest of the paper is organized as follows: In section 2, we introduce the H.264 intra prediction technique. Then fast techniques for the intra prediction are proposed in section 3. Experimental results are shown in section 4. Finally, we conclude the paper in section 5.

2. INTRA PREDICTION IN H.264/AVC

H.264 introduces the mode decision for intra MB coding. Intra prediction in H.264 exploits the spatial correlation between adjacent macroblocks. The H.264 makes intra prediction for both luma block and chroma block. The luma intra prediction of H.264 in baseline profile has two prediction types: I16MB type and I4MB type. As for the chroma intra prediction, there is only one block size: 8x8 block.

In I4MB type [1], there are a total of nine optional prediction modes for each 4x4 luma block, which are shown in fig.1. In I16MB type, the whole macroblock is predicted with 4 possible modes, which are vertical (mode 0), horizontal (mode 1), DC (mode 2) and plane (mode 3). For chroma intra prediction, it also includes four modes, DC (mode 0), horizontal (mode 1), vertical (mode 2) and plane (mode 3).

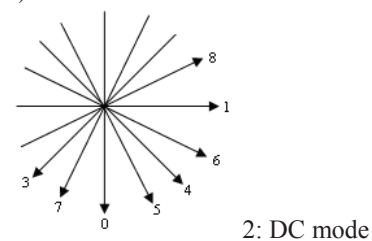


Figure 1: Nine intra prediction modes for 4x4 luminance block

Among all the intra prediction modes, in order to select the best one, H.264 has to implement 592 different RDO calculations. As a result, the complexity of the intra mode decision is especially high, which makes it difficult to achieve real-time implementations.

3. FAST INTRA PREDICTION ALGORITHM

In order to simplify all the intra prediction modes selection process, in this part, we propose a new direction detection algorithm based on the intrinsic characteristics of each intra prediction mode.

3.1 To decide candidate prediction modes for I4MB type

In I4MB type, for each 4x4 subblock, there exist nine 4x4 prediction modes, including eight directional modes and one DC mode. For each directional mode, in our new algorithm, we need to compute the average of the sum of absolute difference between two adjacent original pixels, whose predicted values, interestingly, are equal according to the current intra 4x4 prediction mode defined in H.264 standard. We define this value as the strength of direction differences for this mode. According to the equation of each intra 4x4 prediction mode, we can indicate the locations of pixels with equal predicted values as shown in fig. 2. For pixels with the same value, we label them using the same color (gray, red, yellow, green, blue and light-blue, except white).

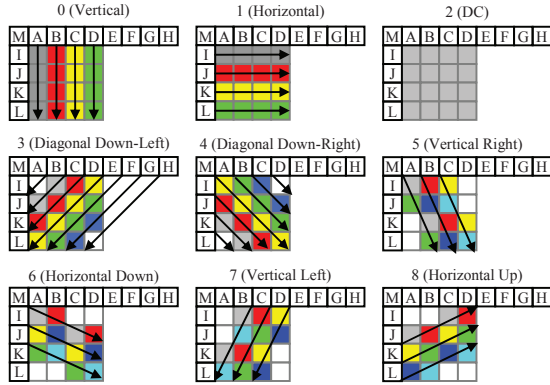


Figure 2: Distribution of “equal pixels” in 4x4 Luma intra prediction modes

Mode 0: Vertical Prediction Mode

For the vertical prediction mode, the upper samples A, B, C, D are used to predict all the pixels in the same columns in the current 4x4 subblock, respectively. It means that the vertical prediction mode has smooth columns, which is depicted in fig.2 (Vertical). Hence we use $\text{DiffVer}_{4 \times 4}$ to estimate the strength of direction differences of this kind of smooth columns in the original subblock, $\text{DiffVer}_{4 \times 4}$ is defined as follows:

$$\text{DiffVer}_{4 \times 4} = \frac{1}{12} \times \sum_{x=0}^3 \sum_{y=0}^3 |p(x, y+1) - p(x, y)| \quad (1)$$

where (x, y) is the position of the pixel $p(x, y)$ in the current 4x4 block, and the weighting factor 12 is the number of absolute differences.

Similarity, for other modes, we use the following equations to estimate their strengths of direction differences in the original subblock.

Mode 1: Horizontal Prediction Mode

$$\text{DiffHor}_{4 \times 4} = \frac{1}{12} \times \sum_{x=0}^3 \sum_{y=0}^3 |p(x+1, y) - p(x, y)| \quad (2)$$

Modes 3: Diagonal Down-Left

$$\text{DiffDDL}_{4 \times 4} = \frac{1}{9} \times \left(\begin{aligned} &|p(1,0) - p(0,1)| + |p(2,0) - p(1,1)| + |p(1,1) - p(0,2)| \\ &+ |p(3,0) - p(2,1)| + |p(2,1) - p(1,2)| + |p(1,2) - p(0,3)| \\ &+ |p(3,1) - p(2,2)| + |p(2,2) - p(1,3)| + |p(3,2) - p(2,3)| \end{aligned} \right) \quad (3)$$

Modes 4: Diagonal Down-Right

$$\text{DiffDDR}_{4 \times 4} = \frac{1}{9} \times \left(\begin{aligned} &|p(0,2) - p(1,3)| + |p(0,1) - p(1,2)| + |p(1,2) - p(2,3)| \\ &+ |p(0,0) - p(1,1)| + |p(1,1) - p(2,2)| + |p(2,2) - p(3,3)| \\ &+ |p(1,0) - p(2,1)| + |p(2,1) - p(3,2)| + |p(2,0) - p(3,1)| \end{aligned} \right) \quad (4)$$

Mode 5: Vertical-Right Mode

$$\text{DiffVR}_{4 \times 4} = \frac{1}{6} \times \left(\begin{aligned} &|p(0,0) - p(1,2)| + |p(1,0) - p(2,2)| + |p(2,0) - p(3,2)| \\ &+ |p(0,1) - p(1,3)| + |p(1,1) - p(2,3)| + |p(2,1) - p(3,3)| \end{aligned} \right) \quad (5)$$

Mode 6: Horizontal-Down Mode

$$\text{DiffHD}_{4 \times 4} = \frac{1}{6} \times \left(\begin{aligned} &|p(0,0) - p(2,1)| + |p(1,0) - p(3,1)| + |p(0,1) - p(2,2)| \\ &+ |p(1,1) - p(3,2)| + |p(0,2) - p(2,3)| + |p(1,2) - p(3,3)| \end{aligned} \right) \quad (6)$$

Mode 7: Vertical-Left Mode

$$\text{DiffVL}_{4 \times 4} = \frac{1}{6} \times \left(\begin{aligned} &|p(1,0) - p(0,2)| + |p(2,0) - p(1,2)| + |p(3,0) - p(2,2)| \\ &+ |p(1,1) - p(0,3)| + |p(2,1) - p(1,3)| + |p(3,1) - p(2,3)| \end{aligned} \right) \quad (7)$$

Mode 8: Horizontal-Up Mode

$$\text{DiffHU}_{4 \times 4} = \frac{1}{6} \times \left(\begin{aligned} &|p(2,0) - p(0,1)| + |p(3,0) - p(1,1)| + |p(2,1) - p(0,2)| \\ &+ |p(3,1) - p(1,2)| + |p(0,3) - p(2,2)| + |p(3,2) - p(1,3)| \end{aligned} \right) \quad (8)$$

After all eight $\text{DiffXX}_{4 \times 4}$ have been computed, the mode with the minimal $\text{DiffXX}_{4 \times 4}$ (MinDiff) will be selected as the candidate mode. Beside this detected mode, to achieve good prediction performance in a smoother block or a boundary block, the DC mode is always chosen. In our algorithm, in order to obtain better performance, we also select two adjacent 22.5 degree modes of the above detected mode with the minimal $\text{DiffXX}_{4 \times 4}$ value according to fig.1 as two more candidate modes for the current 4x4 subblock, say for example, if the current MinDiff is mode 8, then the candidate modes are modes 3, 8, 1 and DC. Hence, the required number of candidate modes, which are required to perform the RDO process to select the best one, is reduced from nine to four for a 4x4 luma subblock.

3.2 To decide candidate prediction modes I16MB type

(1) Modes 0 and 1: Vertical and Horizontal Prediction Modes

Similar to I4MB type, we use eqns.9 and 10 to estimate the strengths of direction differences of the two modes.

For mode 0 (the vertical prediction mode),

$$DiffVer_{16 \times 16} = \frac{1}{240} \times \sum_{x=0}^{15} \sum_{y=0}^{14} |p(x, y+1) - p(x, y)| \quad (9)$$

For mode 1 (the horizontal prediction mode),

$$DiffHor_{16 \times 16} = \frac{1}{240} \times \sum_{y=0}^{15} \sum_{x=0}^{14} |p(x+1, y) - p(x, y)| \quad (10)$$

(2) Mode 3: Plane Mode

For this prediction mode, all the pixels in the macroblock are predicted as in fig.3. A linear 'plane' function, which is defined in eqn.11, is fitted to the upper and left-hand samples H and V [1, 10].

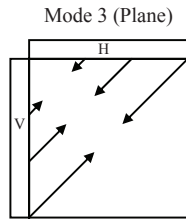


Figure 3: The plane mode in I16MB type

$$pred_{16 \times 16}[x, y] = Clip((a + b \times (x - 7) + c \times (y - 7) + 16) \gg 5) \quad (11)$$

where (x, y) is the position of the pixel p[x, y] in the current macroblock; Clip(•) is the operator to make the prediction pixel value not smaller than zero and not bigger than 255; a, b and c are constants [1,10]. In this prediction mode, we define two pseudo blocks, PseudoBlock_V and PseudoBlock_H, as follows.

$$PseudoBlock_V(x, y) = |p(x, y+1) - p(x, y)| \quad (12)$$

$$PseudoBlock_H(x, y) = |p(x+1, y) - p(x, y)| \quad (13)$$

Combined eqn.11, we can see that PseudoBlock_H and PseudoBlock_V have smooth rows and smooth columns, respectively. So, for the plane prediction mode, we use the following equation to compute the strength of direction differences of the "equal pixels" in the current macroblock.

$$DiffPlane_{16 \times 16} = \frac{1}{2} \times (GradPseB_V + GradPseB_H) \quad (14)$$

where,

$$DiffPseB_V = \frac{1}{224} \times \sum_{x=0}^{15} \sum_{y=0}^{13} |PseudoBlock_V(x, y+1) - PseudoBlock_V(x, y)| \quad (15)$$

Note that 224 is the normalization factor, for which 224=16*14, where 16 is the number of columns in the pseudo PseudoBlockV and 14 is the number of rows in the pseudo PseudoBlockV minus one. Furthermore,

$$DiffPseB_H = \frac{1}{224} \times \sum_{x=0}^{13} \sum_{y=0}^{15} |PseudoBlock_H(x+1, y) - PseudoBlock_H(x, y)|$$

(16)

Similar to the I4MB type, we have to compute all three DiffXX_{16x16} values, and select the one with the minimum value and DC mode as two candidate modes.

3.3 To decide candidate prediction modes for 8x8 Chroma prediction block

For the two Chroma subblocks (U and V) of each MB, the prediction modes are almost the same as those of intra Luma 16x16 macroblock except the differences in indexing. We use the same algorithm proposed in section 3.2 to select two better candidate modes.

Mode 1: Horizontal Prediction Mode

$$DiffHor_{8 \times 8} = \frac{1}{56} \times \sum_{x=0}^6 \sum_{y=0}^7 |p(x+1, y) - p(x, y)| \quad (17)$$

Mode 2: Vertical Prediction Mode

$$DiffVer_{8 \times 8} = \frac{1}{56} \times \sum_{x=0}^7 \sum_{y=0}^6 |p(x, y+1) - p(x, y)| \quad (18)$$

Mode 3: Plane Prediction Mode

$$DiffPlane_{8 \times 8} = \frac{1}{2} \times (GradPseB_V + GradPseB_H) \quad (19)$$

where

$$DiffPseB_V = \frac{1}{48} \times \sum_{x=0}^7 \sum_{y=0}^5 |PseudoBlock_V(x, y+1) - PseudoBlock_V(x, y)| \quad (20)$$

$$DiffPseB_H = \frac{1}{48} \times \sum_{x=0}^5 \sum_{y=0}^7 |PseudoBlock_H(x+1, y) - PseudoBlock_H(x, y)| \quad (21)$$

4. EXPERIMENTAL RESULTS

The proposed intra prediction algorithms have been realized in the H.264 reference software JM 12.2[9] encoder provided by JVT. All frames of the sequences employ intra coding, and the values of QP are 22, 27, 32 and 37. CABAC is available, only 4x4 transform is used, the number of frames to be encoded is 300. Frame rate is 30 HZ. We use nine sequences for our testing, including five QCIF sequences: News, Silent, Foreman, Container, Akiyo and four CIF sequences: Stefan, Tempete, Waterfall, and Paris. Table 1 and Table 2 list the evaluation results of the proposed algorithm and the algorithm in ref.8 by comparing with the full search method in terms of PSNR, bit rate and time. It can be seen from the two tables that our algorithm always achieves more timesaving than algorithm in ref.8. We also can see that our algorithm can achieve more than 80% timesaving of computational complexity compared with that of the full search in JM12.2 software. The average loss of PSNR is about 0.05dB for QCIF sequence and 0.1dB for CIF sequence, and there is slight increase in bit rate of about 2.4% for QCIF sequence and 1.7% for CIF sequence. Fig.4 shows the performance of RD for "Akiyo" sequence. Fig.5 shows the performance of the computational complexity of sequences "Akiyo". From these figures, we also can see that the proposed algorithm gives almost identical RD performance while providing a

drastic decrease in computational complexity compared with the full search in JM12.2 software.

5. CONCLUSIONS

In this paper, we propose one strategy to accelerate the process of intra prediction. The strategy is to make full use of the characteristics of each prediction mode, which is defined in H.264 standard, to select the best candidate prediction mode more accurately by the computation of direction difference of each prediction mode. This strategy can efficiently simplify the process of intra prediction mode selection. Experimental results show that our new algorithm can achieve more than 80% computation reduction (i.e. a speed up of 5 times) of JM12.2, with negligible loss in PSNR and increase in bit rate, which are better than other algorithms reported in the literature.

Table 1. Evaluation results for QCIF sequences

sequence	[8]			The proposed algorithm		
	Δ PSNR [dB]	Δ BR [%]	Δ TIME [%]	Δ PSNR [dB]	Δ BR [%]	Δ TIME [%]
News	-0.000	3.631	77.038	-0.033	2.756	80.572
Silent	-0.045	2.631	76.820	-0.080	2.327	80.475
Foreman	-0.015	3.335	76.640	-0.050	2.438	80.160
Container	-0.020	3.003	76.263	-0.058	2.016	80.543
Akiyo	-0.008	3.405	77.455	-0.028	2.664	80.456
Average	-0.018	3.201	76.843	-0.050	2.440	80.441

Table 2. Evaluation results for CIF sequences

sequence	[8]			The proposed algorithm		
	Δ PSNR [dB]	Δ BR [%]	Δ TIME [%]	Δ PSNR [dB]	Δ BR [%]	Δ TIME [%]
Stefan	-0.050	2.609	77.815	-0.118	2.113	81.008
Tempete	-0.050	2.253	78.049	-0.118	1.730	81.151
Waterfall	-0.080	1.454	77.150	-0.145	1.157	80.791
Pairs	-0.000	2.368	79.066	-0.045	1.966	81.776
Average	-0.045	2.171	78.020	-0.107	1.742	81.182

Acknowledgment

This work is supported by the Centre for Multimedia Signal Processing, Hong Kong Polytechnic University and the Research Grant Council of the Hong Kong SAR Government (PolyU 5267/07E).

REFERENCES

- [1] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, "Draft ITU-T Recommendation and Final Draft international Standard of Joint Video Specification (ITU-T Rec. H.264| ISO/IEC 14496-10 AVC)", JVT-G050, March 2003.
- [2] Yiu-Lam Chan and Wan-Chi Siu, "New Adaptive Pixel Decimation for Block Motion Vector Estimation," IEEE Transactions on Circuits and Systems for Video Technology, pp.113-118, vol.6, no.1, Feb. 1996.
- [3] Ko-Cheung Hui, Wan-Chi Siu and Yui-Lam Chan, "New Adaptive Partial Distortion Search using Clustered Pixel Matching

Error Characteristic," IEEE Transactions on Image Processing, vol.14, no.5, pp.597-607, May 2005.

[4] Ying Zhang, Wan-Chi Siu and Tingzhi Shen, 'Yet a Faster Motion Estimation Algorithm with Directional Search Strategies', pp.475-478, Proceedings, 15th International Conference on Digital Signal Processing (DSP'2007), July 2007, Cardiff, UK.

[5] An-Chao Tsai, Anand Paul, Jia-Ching Wang, Jhing-Fa Wang, "Intensity Gradient Technique for Efficient Intra-Prediction in H.264/AVC," IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 5, pp.694-698, MAY 2008.

[6] Hongliang Li, King Ng Ngan, and Zhenyu Wei, "Fast and Efficient Method for Block Edge Classification and Its Application in H.264/AVC Video Coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 6, pp.756-768, JUNE 2008.

[7] Chao-Hsuing Tseng, Hung-Ming Wang, and Jar-Ferr Yang, "Enhanced Intra-4x4 Mode Decision for H.264/AVC Coders", IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, no. 8, pp. 1027-1032. AUGUST 2006.

[8] 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.

[9] JVT H.264/AVC Reference Software version JM 12.2, http://iphome.hhi.de/suehring/ttml/download/old_jm

[10] Iain E.G. Richardson, "H.264 and MPEG-4 Video Compression", pp.180, 2003.

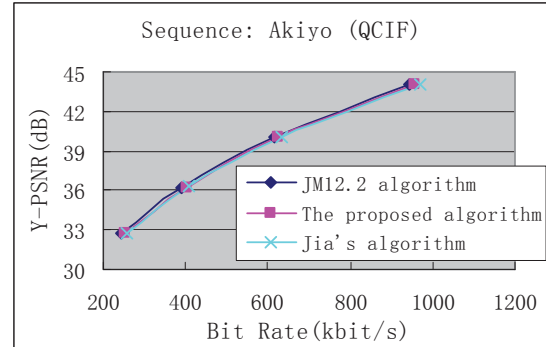


Figure 4: R-D performance for Akiyo sequence

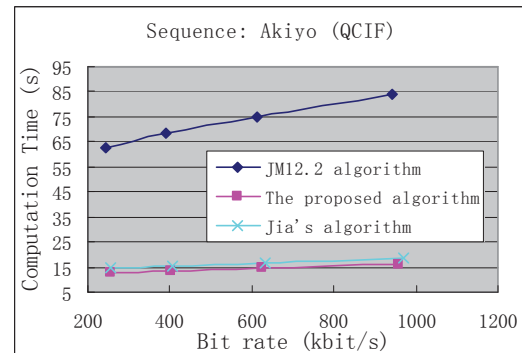


Figure 5: Computational complexity for Akiyo sequence