제27회 한국정보처리학회 춘계학술발표대회 논문집 제14권 제1호 (2007. 5)

A Hierarchical Mode Decision Method for H.264 Intra Image Coding

Jiantan Liu, Kook-yeol Yoo
Depts. Info. & Comm. Eng., Graduate School, Yeungnam University
kyoo@yu.ac.kr

Abstract

Due to its impressive compression performance, the H.264 video coder is highlighted in the video communications industry, such as DMB (Digital Multimedia Broadcasting), PMP (Portable Multimedia Player), etc. The main bottleneck to use the H.264 coder lays in the computational complexity, i.e. five times more complex than the market leading MPEG-4 Simple Profile codec. In this paper, we propose the hierarchical mode decision method for intraframe coding for the reduction of the computation complexity of the encoder. By determining the mode group early, the propose algorithm can skip the computationally demanding computation in the mode decision. The proposed algorithm is composed of three steps: 16x16 mode decision, 4x4 mode-group decisions, and final mode decision among the selected mode group. The simulation results show that the proposed algorithm achieves 20% to 50% reduction in the computational complexity compared with the conventional algorithm.

1. Introduction

Nowadays the H.264 video compression standard [1] has gained the huge interest from the video, game, broadcasting industries. The H.264 achieves about 50% performance improvement in compression performance over the conventional video compression standards, such as MPEG-4 Advanced Simple Profile codec [2]. H.264 uses several new techniques to achieve the highest coding efficiency, such as variable block types in intra and inter prediction, multi-frame inter prediction, Lagrangian Rate-Distortion optimization technique to decide the coding mode [3] of each MB, etc.

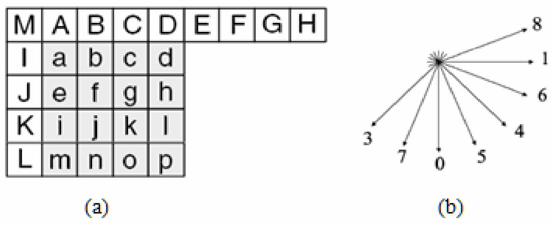
The consumer electronics companies have deployed the digital video products with H.264 video codec, such PMP's from Apple, Sony, Iriver, etc, and the DMB (Digital Multimedia Broadcasting), which is the mobile handy digital TV in South Korea, adopts the H.264 as a video codec. Though the H.264 codec has successfully become a market leading video compression standard in the short period of time, its computational complexity, especially at the encoder side is great bottleneck in the implementation of H.264. Specifically the encoding computational complexity is about five times more complex than the MPEG-4 Simple Profile encoder [2]. For this reason, some products use the H.264 for the decoder and MPEG-4 Simple Profile for the encoder and decoder. It means that the H.264 only uses the distribution of the pre-encoded video contents, limiting its possible use of the bi-directional communications.

There are many literatures to reduce the computational complexity by simplifying the mode decision procedures [4][5]. Two improved methods in luminance and chrominance component are proposed in [4]. And for luminance component, [4] choose the best intra_16x16 mode and adds several correlative modes as the candidate modes for intra_4x4 prediction. In this paper, we investigate on the reduction method of the computational complexity of H.264 encoders, especially on the Intraframe mode decision for luminance component and this paper is organized as follows. Section 2 is the overview of the intra prediction mode decision algorithm in the JVT reference software. Section 3 is a fast mode decision algorithm proposed by other researchers. And Section 4 is our proposed algorithm based on the reference paper. Simulation results and comparison are shown in Section 5. Finally, conclusions are presented.

2. Overview of Intra Prediction in H.264/AVC

Intra coding removes to the spatial redundancies within a video picture by using transform. For the further compression, the transform-domain DPCM is performed in the previous standards, such as H.263 Annex I and MPEG-4. Intra prediction in H.264 is, however, conducted in the spatial domain, by referring to neighboring samples of previously coded blocks, before applying transform. There are two intra prediction types in H.264, Intra_16x16 and Intra_4x4 types. In addition, the Intra_16x16 has the four intra prediction modes while nine modes for Intra_4x4 type.

The nine prediction modes for 4x4 luma block are shown in Fig. 1. And the I4MB prediction is conducted for samples a-p of a block using samples A-M. There are in total eight prediction directions and one DC prediction mode for I4MB prediction [1]:



(Fig. 1) (a) I4MB prediction by neighboring pixels
(b) Eight "prediction directions" for I4MB prediction.

As an alternative to I4MB prediction type described above, the entire MB may be predicted by I16MB type. This is well suited for smooth image areas where a uniform prediction is performed for the whole luma component of an MB and four prediction modes are supported in the type.

Each 8x8 chroma component of MB is predicted from above and left neighboring chroma samples that have previously been encoded and reconstructed. The four chroma prediction modes are very similar to that of the I16MB prediction except that the order of mode numbers is different. The same prediction mode is always applied to both chroma blocks.

H.264 uses the RDO Rate Distortion Optimization technique to achieve the best coding performance. This means that the encoder has to encode the intra block searching all the mode combinations and choose the one that gives the best performance in the sense of rate-distortion. Since the choice of prediction modes for chroma components is independent to that of luma components, thus, there should be four different chroma prediction modes. Therefore, the total number of mode combinations for luma and chroma components in an MB is M8*(M4*16+M16). It means that, for each MB, it has to perform 4*(9*16+4) = 592 different RDO calculations before encoding the MB. As a result, the complexity and computational load of the encoder is extremely high.

3. Previous Algorithm on Prediction Mode Decision

3.1 Full Search Algorithm of Intra Prediction in JM 10.1

In the Full Search algorithm, all the possible intra prediction modes are checked to find the best modes. The major steps for the process can be summarized as follows:

Step 1: Find the Best Intra_16x16 Prediction Modes

1.1: Generate four prediction blocks for the four Intra_16 x16 modes
1.2: Calculate the RD cost for each prediction block
1.3: Find the mode with the smallest cost as the best Intra_16x16 prediction mode

Step 2: Find the Best Intra_4x4 prediction Modes

2.1: Divide the macroblock into sixteen 4x4 blocks
2.2: For each 4x4 block
2.2.1: Generate nine 4x4 prediction blocks for the nine Intra_4x4 prediction modes
2.2.2: For each mode calculate the RD cost and find the mode with the smallest cost as the best Intra_4x4 prediction mode
2.3: Every 4x4 block in the macroblock has one best Intra_4x4 prediction mode

Step 3: Find the best Intra Prediction Modes

Compute the overall cost for this macroblock using the best Intra_16x16 prediction mode and the best sixteen Intra_4x4 prediction modes. The mode with the smallest cost will be chosen.

3.2 A Conventional Fast Intra Mode Decision Algorithm [5]

The full search algorithm for the H.264 Intraframe coding requires the huge computational complexity. To reduce the complexity, several fast algorithms are presented in the literature.

Yang [5] used macroblock properties; the Intra_16x16 prediction type performs prediction for the whole 16x16 macroblock, it should be more suitable for encoding of very smooth areas of a picture, since the Intra_4x4 mode handles each 4x4 block separately, it is more suitable for encoding the region with significant detail.

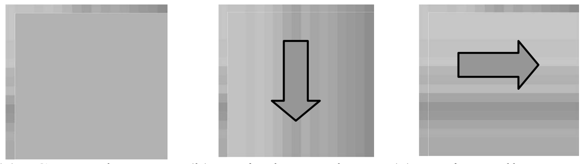
In order to decide the smoothness of each macroblock, the directional smoothness is used as show in Fig.2.

A. Smoothness decision by DC prediction

if MAD_{DC} is smaller than a predefined threshold T_{DC}, the macroblock is DC smooth.

$$MAD_{DC} = \frac{1}{16*16}\sum_{y=0}^{15}\sum_{x=0}^{15}\left|p(x,y)-m\right| \quad (1)$$

where $p(x, y)$ is the pixel value of luminance component of Macroblock at the location of (x, y) and m is the mean of the component.



(a)DC smooth MB (b)Vertical smooth MB (c) Horizontally smooth MB

(Fig. 2) Directional smoothness

B. Vertical smoothness decision

if MAD_V is smaller than a predefined threshold T_V, the macroblock is vertical smooth.

$$MAD_V = \frac{1}{16*16}\sum_{x=0}^{15}\sum_{y=0}^{15}\left|p(x,y)-m_x\right| \quad (2)$$

$$m_x = \frac{1}{16}\sum_{y=0}^{15} p(x,y) \quad (3)$$

where $p(x,y)$ is the pixel value of luminance component of Macroblock at the location of (x, y) and m_x is the mean of the component.

C. Horizontal smoothness decision

if MAD_H is smaller than a predefined threshold T_H, the macroblock is horizontal smooth.

$$MAD_H = \frac{1}{16*16}\sum_{x=0}^{15}\sum_{y=0}^{15}\left|p(x,y)-m_y\right| \quad (4)$$

$$m_y = \frac{1}{16}\sum_{x=0}^{15} p(x,y) \quad (5)$$

where $p(x, y)$ is the pixel value of luminance component of Macroblock at the location of (x, y) and m_y is the mean of the component.

The method by Yang [5] is summarized as follows:

Step 1: Smoothness decision

If the Macroblock is decided as smooth one, go to Step 2, otherwise go to Step 3

Step 2: Find the best mode among the four Intra_16x16 prediction Modes

Step 3: Find the best mode among the nine Intra_4x4 predictionModes for each 4x4 block in the Macroblock

The Full Search algorithm uses two intra prediction types, Intra_16x16 and Intra_4x4 to find the best prediction modes for each macroblock. But in this algorithm only one type, Intra_16x16 or Intra_4x4, is used to find the best prediction modes, which can save much computation, especially for the sequences with more smooth area.

4. Proposed Hierarchical Mode Decision Algorithm for H.264 Intraframe Coding

In this section, a hierarchical fast intra mode decision algorithm is proposed, which is extension of Yang's method [5] and block classification by spatial characteristics.

<Table 1> Proportion of Intra_4x4 mode for several sequences

Sequences (176*144)	Smallest MAD	Mode0 (V)	Mode1 (H)	Mode2 (DC)	Mode3 (D_dl)	Mode4 (D_dr)
Foreman	DC	11.0%	26.0%	33.8%	5.2%	9.4%
	V	8.2%	8.6%	17.1%	6.3%	22.5%
	H	11.5%	18.4%	19.2%	4.8%	21.4%
Claire	DC	0.0%	0.0%	0.0%	0.0%	0.0%
	V	34.4%	5.8%	24.3%	6.5%	5.2%
	H	15.7%	14.5%	32.6%	4.4%	6.7%
Carphone	DC	15.6%	9.3%	35.3%	14.5%	9.9%
	V	32.5%	9.7%	15.3%	6.0%	7.6%
	H	15.6%	34.1%	14.2%	5.2%	6.0%
New	DC	2.0%	5.2%	50.0%	7.3%	7.3%
	V	43.7%	13.3%	9.8%	4.2%	4.8%
	H	15.3%	46.6%	13.9%	2.9%	3.8%

<Table 1> Proportion of Intra_4x4 mode for several sequences (continue)

Sequences (176*144)	Smallest MAD	Mode5 (V_r)	Mode6 (H_d)	Mode7 (V_l)	Mode8 (H_u)
Foreman	DC	4.2%	5.5%	1.7%	3.2%
	V	11.8%	8.7%	8.9%	7.8%
	H	5.3%	1.6%	3.0%	14.8%
Claire	DC	0.0%	0.0%	0.0%	0.0%
	V	8.3%	3.3%	9.6%	2.5%
	H	3.0%	10.2%	4.5%	8.5%
Carphone	DC	2.9%	4.8%	3.0%	4.9%
	V	13.0%	3.6%	6.0%	6.2%
	H	5.2%	5.9%	3.0%	10.8%
New	DC	6.0%	0.4%	20.9%	0.9%
	V	7.5%	5.0%	8.4%	3.1%
	H	3.0%	5.0%	3.1%	6.3%

By using the full search algorithm in JM 10.1, the proportion of Intra_4x4 mode for sixteen 4x4 blocks in one Macroblock are shown in the Table 1. If the Macroblock has the smallest MAD in vertical direction (see Fig.2-(b)) in the step1 of Yang's method, the 4x4 blocks in the Macroblock have high possibility to be decided the modes 0, 1, 7, 5, 2 among the Intra_4x4 modes. In other words, the pixels in the Macroblock have high correlation in the vertical directions.

Based on the results from above simulation, discuss the validity of the block classification shown in Table 2.

<Table 2> Candidates of 4x4 modes groups

Candidate group	Base mode	Candidate 4x4 mode group
V	Mode 0 (vertical)	Mode 0, 1, 5, 7, 2
H	Mode 1 (horizontal)	Mode 0, 1, 6, 8, 2
DC	Mode 2 (DC)	Mode 0, 1, 3, 4, 2

The steps of proposed algorithm:

Step 1: Smoothness decision [5]

if $MAD_{DC} \le T_{DC}$, go to Step2
else if $MAD_V \le T_V$, go to Step2
else if $MAD_H \le T_H$, go to Step2
else go to Step3

Step 2: Smooth Macroblock, Intra_16x16 mode

Select the best Intra_16x16 mode and terminate the mode decision

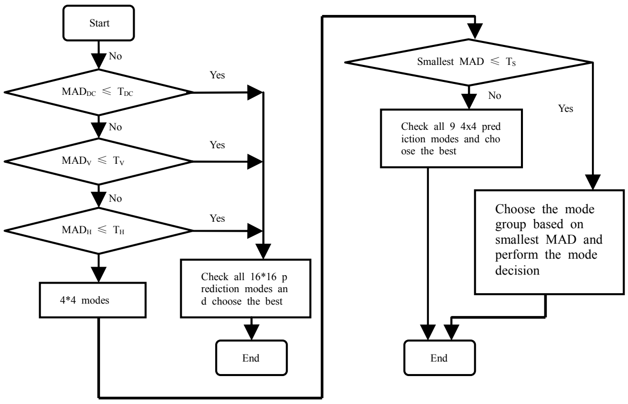
Step 3: Detail Macroblock, Intra_4x4 mode

if $\min\{ MAD_{DC}, MAD_V, MAD_H \} < T_S$,
go to Step3.1, otherwise Step3.2

Step3.1: Search the mode among the candidate group, g.

$g = \min\{ MAD_g \}$, where g = DC, V, H
Note that candidate modes are shown in Table 1

Step3.2: Search the best mode for all the Intra_4x4 modes



(Fig. 3) Flowchart of the proposed algorithm

Since for most cases, the four or five modes are searched for Intra_4x4 modes, instead of nine modes, it means that the computational complexity is reduced, compared with the reference method, which will be shown by the simulation results in the next section.

5. Simulation Results and comparison

All the simulations are performed based on JM10.1 software provided by JVT with a P4 3.0GHz personal computer.

5.1 Experiments on Sequence Claire, Carphone and Forman

Experimental simulations are performed on the first 100 frames of some well-known QCIF sequences, which are Claire, Carphone and Forman. The simulation results in terms of PSNR, bit-rate and encoding time using full searching algorithm, in [4] of luminance component, [5] and proposed algorithm are shown in Table 3.

In Table 3, PSNR are the Peak Signal to Noise Ratio of luminance component and the average number of bit for encoding 100 I frames of each sequence, respectively. Bit Increase is the percent increase in bit using the proposed algorithm and reference algorithm [4] and [5] relative to full searching algorithm. Time is the CPU time and Time saving is the percent of time saving using proposed algorithm and reference algorithm, relative to full searching algorithm.

<Table 3> Comparison of three algorithms (Average of 100 I frames, QP = 28)

Sequences (176*144)	Method	PSNR	Bits/Pic	Bit increase	Times (ms)	Time saving
Claire	JM 10.1	40.80	12928.48		408.62	
	Ref [4]	40.74	13548.08	4.79%	339.53	16.91%
	Ref [5]	40.83	13396.48	3.62%	240.16	41.23%
	Proposed	40.82	13557.36	4.86%	212.50	48.00%
Carphone	JM 10.1	38.34	20059.92		451.08	
	Ref [4]	38.29	21179.60	5.58%	357.24	20.80%
	Ref [5]	38.32	20463.76	2.01%	345.40	23.43%
	Proposed	38.32	20639.92	2.89%	304.19	32.56%
Foreman	JM 10.1	36.76	25670.96		496.72	
	Ref [4]	36.71	27558.32	7.35%	398.68	19.74%
	Ref [5]	36.75	25838.48	0.65%	425.11	14.42%
	Proposed	36.72	26109.12	1.71%	387.47	21.99%

The results in Table 3 show that the proposed algorithm can achieve quite good computation reduction while maintaining almost the same PSNR and a little bit rate increase. The performance of bit rate increase and time saving are both better than reference [4] algorithm. And the computation time saving of proposed algorithm is 20% to 50%, nearly 8% more than reference [5] algorithm.

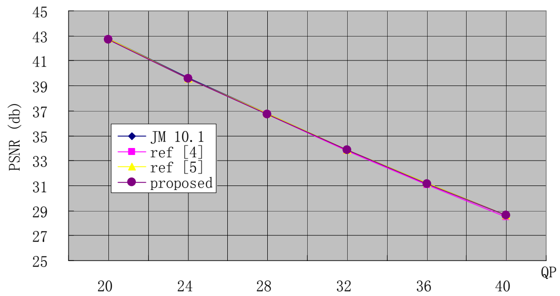
5.2 Experiments on Foreman Sequence with Different Quantization Parameters

In this sequence, 100 I frames are used. A group of experiments were carried out in foreman.qcif with quantization parameters 20, 24, 28, 32, 36, 40. The results of four algorithms are listed in Table 4.

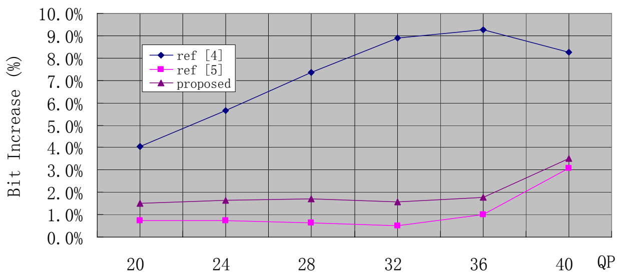
<Table 4> Experiments of different Quantization Parameters (Average of 100 I frames)

QP	Method	PSNR	Bits/Pic	Bit increase	Times (ms)	Time saving
20	JM 10.1	42.77	54000.88		622.37	
	Ref [4]	42.72	56277.76	4.04%	489.96	21.27%
	Ref [5]	42.76	54391.60	0.72%	529.65	14.90%
	Proposed	42.73	54808.64	1.50%	476.09	23.50%
24	JM 10.1	39.60	37426.32		560.77	
	Ref [4]	39.55	39538.40	5.64%	443.00	21.00%
	Ref [5]	39.59	37706.48	0.75%	460.51	17.88%
	Proposed	39.57	38040.24	1.64%	426.28	23.98%
28	JM 10.1	36.76	25670.96		496.72	
	Ref [4]	36.71	27558.32	7.35%	398.68	19.74%
	Ref [5]	36.75	25838.48	0.65%	425.11	14.42%
	Proposed	36.72	26109.12	1.71%	387.47	21.99%
32	JM 10.1	33.88	17114.24		431.18	
	Ref [4]	33.80	18639.76	8.91%	345.66	19.83%
	Ref [5]	33.88	17200.24	0.50%	371.56	13.83%
	Proposed	33.85	17384.64	1.58%	346.97	19.53%
36	JM 10.1	31.17	11315.28		388.80	
	Ref [4]	31.05	12364.24	9.27%	312.58	19.60%
	Ref [5]	31.17	11427.92	0.99%	337.00	13.32%
	Proposed	31.14	11515.92	1.77%	317.06	18.45%
40	JM 10.1	28.65	7795.44		369.84	
	Ref [4]	28.50	8438.24	8.25%	299.06	19.14%
	Ref [5]	28.65	8034.40	3.07%	313.71	15.18%
	Proposed	28.64	8068.88	3.51%	294.21	20.45%

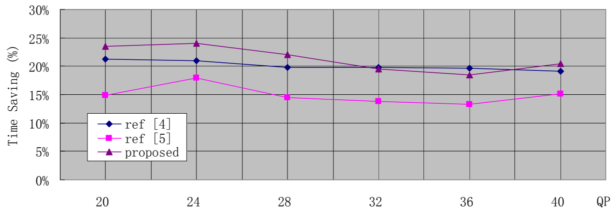
Fig.4 shows that for a given quantization step size, the PSNRs are almost the same. The Fig.5 shows that the proposed algorithm produces about 2.5~7.5% less bits than [4] and 0.6~1.1% more bits than [5].



(Fig. 4) PSNR of three algorithms for the various QP



(Fig. 5) Bit increase of Reference [4], [5] and proposed algorithms with various QP



(Fig. 6) Time Saving of Reference [4], [5] and proposed algorithms with various QP

Compared with [4], the proposed algorithm reduces a little computational complexity about 1~4% and sometimes is less than [4], but the proposed algorithm produces much less bits about 2.5~7.5%.

However, the proposed algorithm reduces the 6~11% computational complexity. In overall, the proposed algorithm gives about 10% reduction of computational complexity at the cost of 1% increase in the bit rate compared with [5] as show in Fig.6.

6. Simulation Results and comparison

In this paper, we proposed a hierarchical mode decision method for the reduction of computational complexity in the H.264 Intraframe coding for luminance component. From the statistical analysis, the Intra_4x4 mode can be classified into the several candidate groups. By apply mode decision only to the selected candidate group, we can reduce computational complexity by 10% or 20~50%, compared with the conventional and JM10.1 methods, respectively.

According to the reference [4] and [7], there is also much computation complexity residue in chrominance component. And we will perform the further researches of fast mode decision in chrominance component based on our proposed algorithm in this paper.

References

[1] Information Technology—Coding of Audio-Visual Objects—Part 10: Advanced Video Coding, Final Draft International Standard, ISO/IEC FDIS 14 496-10, Dec. 2003.

[2] MPEG-4 Video Verification Model version 16.0-Coding of Moving Pictures and Audio, ISO/IEC JTC1/SC29/WG11 N3312, March. 2000.

[3] G. Sullivan, T.Wiegand, and K.-P. Lim, “Joint model reference encoding methods and decoding concealment methods”, presented at the 9th JVT Meeting (JVT-I049d0), San Diego, CA, Sep. 2003.

[4] Jun-Sung Park and Hyo-Jung Song, “Selective intra prediction mode decision for H.264/AVC encoders”, Transactions on Engineering, computing and technology volume 13 May 2006 ISSN 1305-5313.

[5] Chun-Ling YANG, Lai-Man PO, Wing-Hong LAM, “A fast H.264 intra prediction algorithm using macroblock properties” Image Processing, 2004. ICIP '04. 2004 International Conference on Volume 1, 24-27 Oct. 2004Page(s):461 - 464 Vol. 1.

[6] Yong-dong Zhang, Feng Dai, Shou-xun Lin, "Fast 4×4 intra-prediction mode selection for H.264", Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference on Volume 2, 27-30 June 2004 Page(s):1151- 1154 Vol.2.

[7] Fangwen Fu, Xinggang Lin, Lidong Xu, “Fast intra prediction algorithm in H.264/AVC”, Signal Processing, 2004. Proceedings. ICSP '04. 2004 7th International Conference on Volume 2, 31 Aug.-4 Sept. 2004Page(s):1191 - 1194 vol.2