IEICE TRANS. INF. & SYST., VOL.E95–D, NO.11 NOVEMBER 2012

LETTER

Selective Intra Block Size Decision and Fast Intra Mode Decision Algorithms for H.264/AVC Encoder

Karunanithi BHARANITHARAN[†a)], ***Member***, **Jiun-Ren DING**[††], **Bo-Wei CHEN**[†††], ***and*** **Jhing-Fa WANG**[†††], ***Nonmembers***

SUMMARY In H.264/AVC intra frame coding, the rate-distortion optimization (RDO) is employed to select the optimal coding mode to achieve the minimum rate-distortion cost. Due to a large number of combinations of coding modes, the computational burden of Rate distortion optimization (RDO) becomes extremely high in intra prediction. In this paper, we proposed an efficient selective intra block size decision (SIB) that selects the appropriate block size for intra prediction, further proposed fast intra prediction algorithm reduces a number of modes required for RDO that significantly reduces the encoder complexity. Experimental results show that the proposed fast mode decision algorithm reduces the encoding time by up to 68% with negligible video quality degradation.
key words: *intra prediction, edge detection, RDO, advanced video coding, inter prediction*

1. Introduction

The H.264/AVC is a new video coding standard, which is developed by Joint Video Team (JVT) formed by ITU-T VCEG and ISO/IEC MPEG [1]. The new video coding standard is based on a traditional hybrid coding scheme which incorporates many state-of-the-art techniques to achieve outstanding coding performance [2], [3]. Compared to the previous MPEG-2 and MPEG-4 video standards, the latest H.264 video coding figured with many new coding tools, could improve coding efficiency by up to 50% [2]. One of the novel features in H.264 video coding is the use of 7 different macroblock (MB) encoding modes such as SKIP, INTER 16×16, INTER 16×8, INTER 8×16, INTER 8×8, INTRA 16×16, and INTRA 4×4 to best present the temporal and spatial details in a macroblock (MB). To select the best mode, rate-distortion optimization (RDO) is usually employed to optimize the coding efficiency. All the MB modes are tried and the one that leads to the least RD cost is selected to achieve the best trade-off between rate and distortion [3], [4]. However, the RDO technique dramatically increases the computational complexity of the H.264 encoder. Therefore, the development of more efficient algorithms to reduce the computation of intra and inter prediction is necessary. Several approaches have been proposed to achieve fast intra prediction. Tseng et al. [5] proposed an enhanced rate-distortion cost function, which combines the sum of absolute integer-transformed differences combines the sum of absolute integer-transformed differences (SAITD) and a rate predictor for H.264/AVC intra 4×4 mode decision. The algorithm proposed by Kim et al. [6] adopts a multi-stage sequential mode decision process that uses joint spatial and transform domain features to filter out unlikely candidate modes. Pan et al. [7] suggested a mode prediction method to reduce the encoding time by about 50%, where Sobel's operator is applied to every pixel in the block to detect the most probable texture edge. There are a number of ways to get the local edge directional information, such as using an edge direction histogram that's based on a simple edge detection algorithm [8]. Li et al. [9] proposed a fast mode decision method using difference properties from three coefficients in the non-normalized Haar transform (NHT). Meng et al. [10] proposed a threshold to terminate the early computation of the most probable mode. Wang et al. [11] suggested that the irregular structures of computational masks that partially involve the multiplications of 21/2 make the implementation of VLSI more difficult. For diagonal edges, the MPEG-7-based edge detection method in some cases produced incorrect results.

In this paper, we propose an efficient selective intra block size decision that selects the suitable intra mode for an MB, further, proposed fast intra mode prediction algorithm which reduces the candidate modes that are required for RDO. Therefore, a number of prediction modes are reduced from 9 to 4 for 4×4 luma blocks and from 4 to 2 for 16×16 luma and 8×8 chroma blocks. Simulation results show that the proposed algorithm achieves lower PSNR and less bit-rate degradation than that of Wang's algorithm [11]. K.B. Tharan et al. [12] proposed a low complexity fast intra mode decision algorithm. The proposed detection algorithm effectively reduces the number of the candidate modes, such that we can dramatically avoid unnecessary computation for intra prediction.

The rest of the paper is organized as follows. Section 2 is an overview of the intra prediction suggested in H.264. Section 3 explains the proposed algorithm in detail. Section 4 uses experimental results to evaluate the proposed algorithm as well as the existing methods. Finally, Sect. 5 addresses the conclusions of the paper.

2. An Overview of H.264 Intra Prediction

The conventional encoder employs both intra 16×16 and in-

Manuscript received December 14, 2011.
[†]The author is with the Feng Chia University, R.O.C.
[††]The author is with the ITRI, Taiwan.
[†††]The authors are with the National Cheng Kung University, Taiwan.
a) E-mail: dharan@ieee.org
DOI: 10.1587/transinf.E95.D.2720

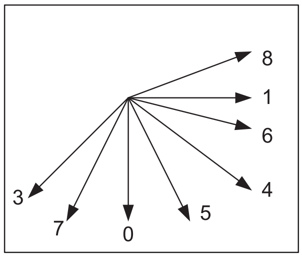

Fig. 1 Intra_4 × 4 mode.

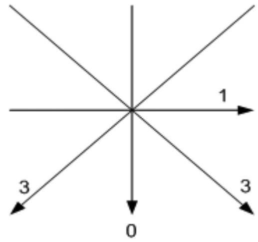

Fig. 2 Intra_16 × 16 mode.

tra 4×4 to an MB and finally a block size that gives the least RD cost is selected. Along various texture directions, H.264 offers a rich set of prediction patterns, called modes, for intra prediction. The original intra prediction needs to try 9, 4, and 4 prediction modes for 4 × 4 luma, 16 × 16 luma, and 8×8 chroma blocks, respectively. Each mode has its own direction of prediction and the predicted samples are obtained from a weighted average of decoded values of neighboring pixels [2]. Generally, Intra_4 × 4 is well suited for an MB with detailed information, while Intra_16×16 is appropriate for smooth MBs. Figure 1 shows eight directional prediction modes, which are represented by Modes 0, 1, 3, 4, 5, 6, 7, and while Mode 2 stands for the DC mode in Intra_4 × 4 prediction. For Intra_16 × 16 luma and Intra_8 × 8 chroma blocks, the prediction consists of vertical, horizontal, DC, and plane prediction modes represented by Modes 0, 1, 2, and 3, respectively. The decision modes in chroma prediction are similar to Intra_16×16 except for the different block size. Figures 1 and 2 show the prediction modes used in Intra_4 × 4 and Intra_16 × 16, respectively.

3. Proposed Fast Intra Mode Decision Algorithm

The conventional H.264/AVC encoder complexity is greatly increased since it employs both intra 4 × 4 and 16 × 16 to an MB. However, only one block type is chosen to encode an macroblock. Hence, it increases both computational complexity as well as encoding time. In our proposed algorithm, we use variance of the current macroblock to decide the appropriate intra block type i.e intra 4 × 4 or intra 16 × 16. Therefore, the conventional brute force search is avoided. We not only reduce the computational complexity of the intra block size type selection but also a number of candidate mode required for rate distortion optimization. In order to do so, we proposed a pixel based direction decision algorithm that computes the differences between two neighboring pixels, which are corresponding to four major horizontal, vertical, diagonal-down-left, and diagonal-down-right texture directions. The following Fig. 3 shows the proposed algorithm's operational blocks, in which SIB block chooses an appropriate intra block size type and the succeeding block reduces a number of candidate mode of the selected intra block type. Finally, the selected mode use to encode an macroblock.

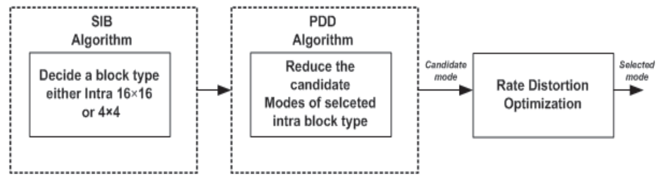

Fig. 3 Functional units of the proposed algorithm.

3.1 Selective Intra Block Size (SIB) Algorithm

In this section, we employ variance technique to classify the intra block types i.e. intra 4×4 and intra 16×16. In general, intra 16×16 block is used in less texture and similar motion region where variance is low whereas intra 4 × 4 block type is applied in texture region where variance is high, By using this phenomena, we categorized blocks using pre-defined threshold. However, in some region of block is not clearly identified by variance method which leads degradation in bit rate. In order to avoid the degradation, we used two thresholds that is explained in Step 2.

Step 1: Compute the mean and variance of the current macroblock using Eqs. (1) and (2).

$$\mu = \frac{1}{|R|} \sum_{(i,j)\in R} x(i, j) \tag{1}$$

where $x(i, j)$ denotes the pixel value and $|R|$ represents the total number of pixels in the region.

$$\sigma^2 = \frac{1}{|R|} \sum_{(i,j)\in R} (x(i, j) - \mu)^2 \tag{2}$$

Step 2: Select both intra 16×16 and 4×4 for RDO if Th1 is less than variance and Th2 is greater than variance.

Step 3: Select only intra 16×16 block size if variance is less than the predefined threshold (Th1).

Step 4: Select only intra 4×4 block size if variance is greater than the predefined threshold (Th2).

Step 5: Apply the proposed PDD algorithm on the selected intra block type to reduce a number of candidate modes.

3.2 Proposed Pixel Based Direction Decision (PDD) Algorithm

In each 4 × 4 block, $f(x, y)$, we can directly compute all the differences between two neighboring pixels, which are corresponding to four major horizontal, vertical, diagonal-down-left, diagonal-down-right texture directions. Therefore, as shown in Fig. 4, the four major pixel direction error strengths can be expressed as

$$d^{0^\circ} = |f(x, y + 1) - f(x, y)| \tag{3}$$

$$d^{90^\circ} = |f(x + 1, y) - f(x, y)|, \tag{4}$$

$$d^{45^\circ} = |f(x + 1, y - 1) - f(x, y)|, \tag{5}$$

$$d^{135^\circ} = |f(x + 1, y + 1) - f(x, y)|, \tag{6}$$

where x and y represent horizontal and vertical positions of pixel $f(x, y)$, respectively. Thus, the block direction error strengths by averaging all the possible pixel direction error strengths in each 4×4 block can be expressed by:

$$D^{0^\circ} = \left(\sum_{i=0}^{11} d_i^{0^\circ}\right) \Big/ 12 \tag{7}$$

$$D^{90^\circ} = \left(\sum_{i=0}^{11} d_i^{90^\circ}\right) \Big/ 12 \tag{8}$$

$$D^{45^\circ} = \left(\sum_{i=0}^{8} d_i^{45^\circ}\right) \Big/ 9 \tag{9}$$

$$D^{135^\circ} = \left(\sum_{i=0}^{8} d_i^{135^\circ}\right) \Big/ 9 \tag{10}$$

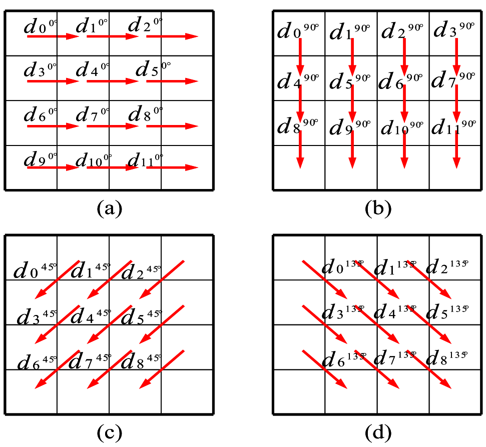

Fig. 4 Major pixel direction error strengths in (a) horizontal, (b) vertical, (c) diagonal-down left, and (d) diagonal-down right directions.

For horizontal texture, of course, the block horizontal difference of pixels stated in (7) will be smaller than block direction error strengths of the other directions. Thus, the block texture direction can be correctly detected as horizontal. For vertical, diagonal-down left, and diagonal-down right texture directions, we can find similar conclusions. Thus, the pixel-based direction detection can easily detect one of four major directions by computing all four major block direction error strengths from (7) to (10). Moreover, we use the linear interpolation to compute the minor direction error strength, for example, the block direction error strength for 67.5°, $D^{67.5^\circ}$, for detecting Mode 7. where the linear interpolation of D^{90° and D^{45° gives

$$D^{67.5^\circ} = \left(D^{90^\circ} + D^{45^\circ}\right) \big/ 2. \tag{11}$$

The remaining minor block direction error strengths can be obtained in a similar way. According to the above computations, we will choose three possible candidate modes, which possess the smallest direction error strengths. To retain the prediction in smoother block, the DC mode is always chosen in the RD optimization. Therefore, by using the proposed PDD method, there are only 4 modes, instead of 9, will be chosen for intra prediction.

For 16×16 luma and 8×8 chroma blocks, we divided them into 16 subblocks with 4×4 size and 2×2 size, respectively. The corresponding pixel values of subsampled 4×4 blocks can be computed by the average of subblocks as

$$a_k = \left(\sum_{i=0}^{3} \sum_{j=0}^{3} f(x_k + i, y_k + j)\right) \Big/ 16 \tag{12}$$

$$c_k = \left(\sum_{i=0}^{1} \sum_{j=0}^{1} f(x_k + i, y_k + j)\right) \Big/ 4 \tag{13}$$

Table 1 Experimental results of the proposed method.

Sequences	**[11]**			**SIB+ [12]**			**Proposed**		
QCIF	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**
News	-0.348	4.451	-58.03	-0.310	3.971	-63.21	-0.285	3.439	-68.02
Container	-0.293	4.440	-57.32	-0.210	3.081	-64.89	-0.214	2.976	-67.74
Silent	-0.255	4.580	-60.66	-0.280	5.091	-63.89	-0.232	4.013	-69.46
Coastguard	-0.225	4.034	-57.78	-0.161	3.031	-65.78	-0.181	2.820	-69.26
Average	**-0.280**	**4.376**	**-58.48**	**-0.240**	**3.79**	**-64.44**	**-0.228**	**3.324**	**-68.62**
CIF	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**
Mobile	-0.237	2.871	-58.07	-0.248	2.981	-66.67	-0.250	2.860	-68.89
Paris	-0.274	3.678	-58.78	-0.291	2.678	-65.67	-0.288	2.891	-67.83
Stefan	-0.301	3.889	-58.56	-0.289	3.736	-68.90	-0.290	3.730	-67.04
Tempete	-0.254	3.735	-56.86	-0.250	4.980	-67.89	-0.252	3.601	-69.53
Average	**-0.266**	**3.543**	**-58.06**	**-0.269**	**3.593**	**-67.28**	**-0.270**	**3.270**	**-68.32**

for 16 × 16 luma and 8 × 8 chroma blocks, respectively. In (12) and (13), (x_k, y_k) indicates the starting position of luma or chroma subblocks for $k = 0, 1, \ldots, 15$. Similarly, the three major block directional strengths can be obtained from (7) to (10). In 16 × 16 luma and 8 × 8 chroma predictions, the directional prediction mode merely includes Mode 0, 1, 2, and 3. Therefore, only the smallest block directional strength and DC Mode are chosen for the RD optimization.

4. Experimental Results

The proposed fast mode decision algorithms were implemented on JM10.0 provided by JVT [13]. In our experiments all the frames are encoded using I-frame coding and each contained 300 frames and the RD optimization is enabled. To compare the rate distortion performance and computational complexity of the proposed algorithms, the PSNR and bit rate (BR) are measured by using Bjontegaard's method [14] for quantization parameters (QP) of 28, 32, 36, 40. The simulation results of the proposed method and the method suggested in [11] are shown in Table 1, where ΔPSNR denotes the average difference of peak signal to noise ratio, ΔBR indicates the average bit rate increasing, and ΔTIME indicates the average time saving for four QP in coding process.

The following time saving calculation is defined to evaluate the time saving performance between the proposed method and JM [13]. In which T_{method} and T_{ref} refers to the encoding time of the proposed method and the reference software JM [13].

$$\Delta TIME = \frac{T_{method} - T_{ref}}{T_{ref}} \times 100\% \qquad (14)$$

5. Conclusion

In this research, we proposed an efficient fast intra mode decision algorithm for fast H.264/AVC intra prediction. The simulation results demonstrate that the proposed algorithm can reduce the encoding time by up to 68% with negligible degradation in video quality. Therefore, the proposed algorithm can be used as a pre-processing unit for intra prediction.

Acknowledgment

This work is supported by National Science Counsil of Taiwan under the grant NSC 101-2221-E-035-038.

References

[1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITUT Rec. H.264/ISO/IEC 14496-10 AVC), March 2003.
[2] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol., vol.13, pp.560–576, July 2003.
[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, Sept. 2003.
[4] C.L. Yang and L.-M. Po, "A fast H.264 intra prediction algorithm using macroblock properties," Proc. ICIP, vol.1, pp.461–464, Oct. 2004.
[5] C.H. Tseng, H.M. Wang, and J.F. Yang, "Enhanced intra-4x4 mode decision for H.264/AVC coders," IEEE Trans. Circuits Syst. Video Technol., vol.16, no.8, pp.1027–1032, Aug. 2006.
[6] C. Kim, H.-H. Shih, and C.-C.J. Kuo, "Feature-based intra-prediction mode decision for H.264," Proc. IEEE ICIP, vol.2, pp.769–772, Oct. 2004.
[7] 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/AVC video coding," IEEE Trans. Circuits Syst. Video Technol., vol.15, no.7, pp.813–822, July 2005.
[8] A.K. Jain and A. Vailaya, "Image retrieval using color and shape," Pattern Recognit., vol.29, pp.1233–1244, Aug. 1996.
[9] H. Li and K.-N. Ngan, "Fast and efficient method for block edge classification," Proc. ACM IWCMC, vol.2, pp.67–72, July 2006.
[10] B. Meng and O.C. Au, "Fast intra-prediction mode selection for 4x4 blocks in H.264," Proc. IEEE ICASSP, vol.3, pp.389–392, March 2003.
[11] 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 1411–1422, Oct. 2007.
[12] K. Bharanitharan, B.D. Liu, J.F. Yang, and W.C. Tsai, "A low complexity detection of discrete cross differences for fast h.264/AVC intra prediction," IEEE Trans. Multimedia, vol.10, no.7, pp.1250–1260, Nov. 2008.
[13] H.264/AVC Reference Software JM10. Available: http://iphome.hhi.de
[14] G. Bjontegaard, "Calculation of average PSNR differences between RD-urves," presented at the 13th VCEG-M33 Meeting, Austin, TX, April 2001.