Region based Complexity Reduction Algorithm for H.264/AVC Encoder

K. Bharanitharan, *Member, IEEE,*
Department of Electrical Engineering
Korea University,Seoul, Korea (South)
E-mail: dharan@korea.ac.kr

***Abstract*— H.264, MPEG-4 Part 10, is the latest digital video coding standard that achieves very high data compression by using several new coding features. One of the key features is the intra prediction based on the rate distortion optimization (RDO) procedure, which requires a true encoding of the intra block by trying all intra prediction modes. The computational burden of intra prediction with the RDO procedure is extremely high for real time applications. In this paper, we propose a classified block type (CBT) algorithm that selects the appropriate intra block size and further our proposed bi-region similarity detection algorithm(BSD) selected block that reduces a number of candidate mode required for RDO which significantly reduces the encoder complexity. Experimental results show that the proposed fast mode decision algorithm reduces the encoding time by about 68%, with negligible loss of video quality.**

Keywords- Intra prediction, fast algorithm, edge detection, H.264/AVC

I. 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]. The improvements are mainly obtained from various employments of improved intra and inter predictions, variable block motion estimation, and multiple reference frames. According to these features, enhanced inter and intra prediction techniques are key elements to the success of H.264. To achieve the best coding efficiency, H.264 also employs the rate-distortion optimization (RDO) technique to significantly improve the coding performance by finding the coding mode which has minimum rate-distortion (RD) cost by using Lagrangian coefficient. For a 4×4 intra block, the RD cost function is shown as

$$RD\cos t = D + \lambda \times R, \qquad (1)$$

where D denotes the sum of square difference (SSD) between the current block $\boldsymbol{A}$ and the reconstructed block $\hat{A}$. The details about the RD costs can be found in [4]. However, the computational complexity of the H.264/AVC encoder is dramatically increased by this optimization technique. The encoder has to encode the target block by searching all possible modes exhaustively for the best mode in the intra frame coding process. To deal with the problem described above, it is thus highly desirable to develop an efficient intra prediction mode selection method and VLSI architecture to speed up the encoding time in software and hardware realizations of H.264.

There are many faster prediction methods have been proposed to successfully reduce the computation for H.264 intra prediction. Huang et al. [5] developed a fast algorithm which performed the context-based adaptive skipping of unlikely prediction modes and simplification of matching operations to save about 50% of the encoding time.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 Harr transform (NHT). Unlike the methods mentioned above, other methods concentrate on the edge block classification in discrete cosine transform [10] – [15]. Meng *et al.* [16] proposed a threshold to terminate the early computation of the most probable mode. Wang *et al.* [17] 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 [17].
K. Bharanitharan *et al.* [18] proposed a low complexity fast intra mode decision algorithm. The proposed detection algorithm effectively reduces the number of the candidate modes.

The rest of the paper is organized as follows. Section II is a brief overview of the intra prediction suggested in H.264. Section III explains the proposed algorithm in detail. Section IV uses experimental results to evaluate the proposed algorithm as well as the existing methods. Finally, Section VI addresses the conclusions of the paper.

II. BRIEF INTRODUCTION OF H.264 INTRA PREDICTION

The luminance (luma) intra prediction of H.264 has two prediction strategies, intra 4×4 and intra 16×16, respectively. The intra 4×4, which predicts each 4×4 luma block individually, is well suited for coding parts of an image with significant details. The intra 16×16, which predicts the entire 16×16 luma block, is more suitable for coding smooth regions of an image. As for the chrominance (chroma) intra prediction, the block type is 8×8. The prediction method is similar to the intra luma 16×16 prediction.The conventional encoder employs both intra 16× 16 and intra 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. Fig. 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. Figs. 1 and 2 show the prediction modes used in Intra_4×4 and Intra_16×16, respectively

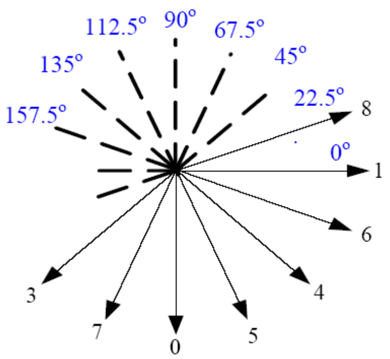

Fig. 1. Intra_4x4 mode directions

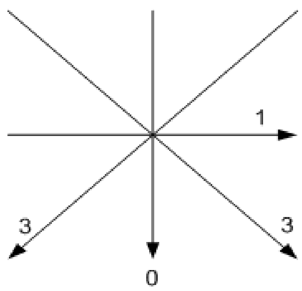

Fig. 2. Intra_16x16 mode directions

III. 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 sub-block 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.1 shows the proposed algorithm's operational blocks, in which part-1 block chooses an appropriate intra block size type and the preceding block i.e. part-2 reduces a number of candidate mode of the selected intra block type. Finally, the selected mode use to encode a macroblock.

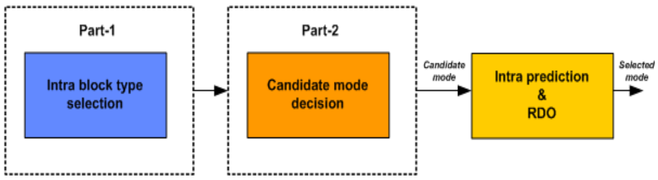

Fig. 3. Block diagram of the proposed algorithms

A. *Proposed Classified Block Type (CBT)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).

$$\mu_k = \frac{1}{|R_k|}\sum_{(i,j)\in R_k} x(i,j) \quad \text{for } k = 1 \text{ and } 2 \qquad (2)$$

where $x(i, j)$ denotes the pixel value and $|R_k|$ represents the total number of pixels in the k^{th} region.

$$\sigma_k^2 = \frac{1}{|R_k|}\sum_{(i,j)\in R_k} (x(i,j)-\mu_k)^2 \quad \text{for } k = 1 \text{ and } 2 \qquad (3)$$

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 BSDD algorithm on the selected intra block type to reduce a number of candidate modes.

B. *Proposed Bi-region Similarity Detection(BSD) Algorithm*

In the proposed algorithm, pixels properly divided by an edge line into two regions in the block normally have similar values. By exploiting this fact, a good mode detection method can be developed if we properly divide the pixels into two regions by a directional edge line. By knowing the detected edge direction, we can reduce the unlikely candidate modes for RDO computation. Instead of using a pixel-based convolution approache, we develop a region-based detection method. First, we divide the pixels of each 4×4 block into two regions for each prediction mode, as shown in Fig. 3, where Regions 1 and 2 are marked by shaded and crossed pixels, respectively. For I4×4 and I16×16 intra predictions, as shown in Fig.4, we design a pair of specific regions, which are fitted in the directions. For example, if we choose a diagonal pattern, the split regions should be diagonal. The proposed fast mode decision algorithm then computes the mean and variance of the regions for edge detection. If a 4×4 block has diagonal edges, for example, only the matched bi-region pattern, i.e., the 45^{o} pattern in Fig. 3, can achieve the smallest variances in both regions. Any unmatched pattern will result in a larger sum of variances in the two divided regions.

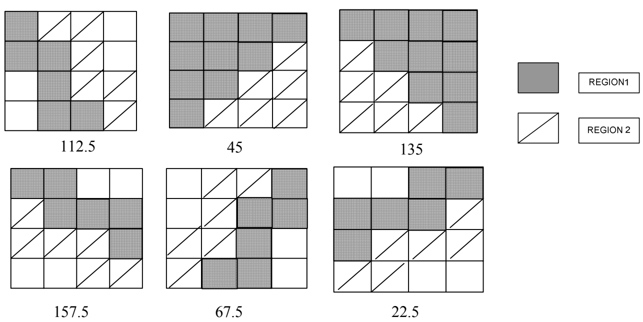

Fig. 3. Six patterns for Intra 4× 4 prediction

The steps of the proposed algorithm are shown below. The candidate modes are selected based on the analysis results and the proposed bi-region patterns which were explained previously.

Step 1: Divide the 4×4 block into Regions 1 and 2 for each mode, as shown in Fig.3.

Step 2: Compute the means for Regions 1 and 2 using:

$$\mu_k = \frac{1}{|R_k|} \sum_{(i,j)\in R_k} x(i,j) \quad \text{for } k = 1 \text{ and } 2 \qquad (4)$$

where $x(i, j)$ denotes the pixel value and $|R_k|$ represents the total number of pixels in the k^{th} region.

Step 3: Compute the variances of Regions 1 and 2 using:

$$\sigma_k^2 = \frac{1}{|R_k|} \sum_{(i,j)\in R_k} (x(i,j) - \mu_k)^2 \quad \text{for } k = 1 \text{ and } 2 \qquad (5)$$

Step 4: Add the variances of Regions 1 and 2 together to obtain the total roughness using:

$$r_{mode} = \sigma_1^2 + \sigma_2^2 \qquad (6)$$

Step 5: Repeat Step 1 to Step 4 for all six modes, i.e., 6 designed patterns.

Step 6: Select the mode with the minimum roughness as the candidate mode.

Step 7: Always choose the Modes 0, 1, and 2 along with the selected candidate mode.

Step 8: Reduce the number of candidate modes from nine to three for the best case.

Step9: Apply RDO calculation to the selected candidate modes.

C. BSD Algorithm for 16×16 and 8×8 blocks

We use the same algorithm for 16×16 chroma and 8×8 luma intra prediction. The main differences in these prediction algorithms are the block size and the mode numbers, though the number of prediction modes is the same. The MB was divided into sixteen 4×4 blocks and the average of the each block was found using (7) The designed patterns (0^{o} , 90 ^{o}) as shown in Fig.9, along with 45 ^{o} as shown in Fig.4 were then applied to these blocks.

$$I_{16\times 16} = \sum_{k=0}^{3}\sum_{l=0}^{3}(a+k, b+l)/16 \qquad (7)$$

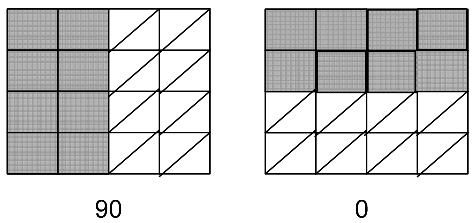

Fig.4. Two designed patterns for I16×16

In 16×16 luma, the edge prediction modes, including Modes 0, 1, 3, and DC, directly correspond to the detected edges. Therefore, the selected mode and the DC mode were chosen for further calculation of the RD cost. The decision modes used for chroma prediction were similar to those used for 16×16 luma. The number of modes is reduced from four to two, which are used as candidates for searching for both 16×16 luma and 8×8 chroma prediction.

IV. Experimental Results and Discussions

The proposed fast mode decision algorithms were implemented on JM10.0 provided by JVT [20]. 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 [21] for quantization parameters (QP) of 28, 32, 36, 40. The simulation results of the proposed method and the method suggested in [7] and [18] are shown in Table I, 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 proposed algorithms achieve the greatest bit-rate reduction with only negligible coding loss in PSNR. The experimental results demonstrate that the proposed algorithms can effectively speed up the intra prediction since the average reduction of total encoding time is around 68%.

V. Conclusion

In this research, we proposed classified block type and sub sampled direction detection algorithms for fast H.264/AVC intra prediction. The simulation results demonstrate that the proposed algorithm can reduce the encoding time by up to 68% as compared to those of Pan's and Wang's algorithms. The picture degradation and bit rate increase are insignificant. Therefore, the proposed algorithm can be used as a pre-processing unit for intra prediction.

Acknowledgment

The achievement is supported by a Korea University Grant Special Research Expense.

$$\Delta\text{TIME} = \frac{T_{method} - T_{ref}}{T_{ref}} \times 100\% \qquad (10)$$

(19)

Table I Experimental results of the proposed algorithm

Sequences	[17]			[18]			Proposed		
QCIF	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**
News	-0.348	4.451	-58.03	-0.194	2.484	-56.35	-0.199	2.710	-65.76
Container	-0.293	4.440	-57.32	-0.191	2.911	-55.83	-0.262	3.092	-66.75
Silent	-0.255	4.580	-60.66	-0.220	4.016	-53.54	-0.179	3.170	-66.97
Coastguard	-0.225	4.034	-57.78	-0.082	1.447	-57.89	-0.129	2311	-66.52
CIF	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**	**ΔPSNR [dB]**	**ΔBR [%]**	**ΔTIME [%]**
Mobile	-0.237	2.871	-58.07	-0.213	2.582	-56.89	-0.219	2.690	-68.43
Paris	-0.274	3.678	-58.78	-0.179	2.395	-55.54	-0.196	2.611	-69.67
Stefan	-0.301	3.889	-58.56	-0.211	2.694	-59.98	-0.249	3.231	-66.57
Tempete	-0.254	3.735	-56.86	-0.212	3.177	-56.45	-0.218	3.181	-66.68
Average	**-0.266**	**3.543**	**-58.06**	**-0.203**	**2.712**	**-57.21**	**-0.220**	**2.905**	**-67.83**

REFERENCES

[1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITUT Rec. H.264/ISO/IEC 14496-10 AVC), Mar. 2003.

[2] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264 / AVC Video Coding Standard," *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 13, pp. 560–576, Jul. 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, Sep. 2003.

[4] 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.

[5]Y. W. Huang, B. Y. Hsieh, T. C. Chen, and L. G. Chen, "Analysis, fast algorithm, and VLSI architecture design for H.264/AVC intra frame coder," *IEEE Trans. Circuits and Syst. Video Technol.,* vol. 15, no. 3, pp.378–401, Mar. 2005.

[6] C. Kim, H.-H Shih, C.-C. J. Kuo "Feature-based intra-prediction mode decision for H.264", in *Proc. IEEE ICIP*, Oct. 2004, vol. 2, pp.769 – 772.

[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, Jul. 2005.

[8] A. K. Jain and A. Vailaya, "Image retrieval using color and shape," *Pattern Recogn*., vol. 29, pp. 1233–1244, Aug. 1996.

[9] H. Li and K.- N. Ngan, "Fast and efficient method for block edge classification," in *Proc. ACM IWCMC*, Jul. 2006, vol. 2, pp. 67-72.

[10] S.-W. Lee, Y.-M. Kim, and S. W. Choi, "Fast scene change detection using direct feature extraction from mpeg compressed videos," *IEEE Trans. Multimedia*, vol. 2, no. 4, pp. 240–254, Dec. 2000.

[11] H. S. Chang and K. Kang, "A compressed domain scheme for classifying block edge patterns," *IEEE Trans. Image Processing*, vol.14, no. 2, pp.145–151, Feb. 2005.

[12] H. Wang, S. Kwong, and C.-W. Kok, "An efficient mode decision algorithm for H.264/AVC encoding optimization," *IEEE Trans. Multimedia*, vol. 9, no. 4, pp.882-888, Jun. 2007.

[13] P. A. A. Assuncao and M. Ghanbari, "Transcoding of MPEG-2 video in the frequency domain," in *Proc. IEEE ICASSP*, Apr. 1997, vol. 4, pp. 2633-2636.

[14] T. Shanableh, and M. Ghanbari, "Hybrid DCT/pixel domain architecture for heterogeneous video transcoding," *Signal Processing: Image Commun.*, vol.18, no. 8, pp. 601-620, Sep. 2003.

[15] G. Hwang, J. Park, B. Jung, K. Choi, Y. Joo, Y. Oh, and B. Jeon, "Efficient fast intra mode decision using transform coefficients," in *Proc. IEEE ICAT*, Feb.2007. vol.1, pp. 399-402.

[16] B. Meng and O. C. Au, "Fast intra-prediction mode selection for 4x4 blocks in H.264", in *Proc. IEEE ICASSP*, Mar. 2003, vol. 3, pp. 389-392.

[17] 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.

[18] 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

[19]A.-C. Tsai, J.-F. Wang, J.-F. Yang, and W.-G. Lin, "Effective Subblock-Based and Pixel-Based Fast Direction Detections for H.264 Intra Prediction", *IEEE Trans. Circuits Syst. Video Technol.,*vol. 18, no. 7, pp. 975-982, Jul. 2008.

[20]H.264/AVC Reference Software JM10.Available: ttp://iphome.hhi.de

[21] G. Bjontegaard, "Calculation of average PSNR differences between RD- urves," presented at the 13th VCEG-M33 Meeting, Austin, TX, Apr. 2001.