2019 IEEE 4th International Conference on Signal and Image Processing

A Novel Fast Mode Decision Algorithm for AVS2 Intra Coding

Xiao Liu[1], Wei Yan[1*], Guoqing Xiang[2] , Lei Cheng[1] , Yunyao Yan[2]
[1] School of Software and Microelectronics, [2] National Engineering Laboratory of Video Technology,
Peking University
Beijing 100871, P.R. China
e-mail: 1601210409@pku.edu.cn, yanwei@ss.pku.edu.cn*, gqxiang@jdl.ac.cn, leicheng@pku.edu.cn,
yunyao_yan@163.com

Abstract—**AVS2 is the latest generation of video coding standard independently developed by China. Compared with the most popular international standard HEVC/H.265, AVS2 achieves the same level of compression efficiency, which doubled in compression efficiency compared to the previous generation coding standard AVC/H.264 or AVS. However, the increased compression efficiency is accompanied by a significant increase in coding complexity. In order to improve the coding speed, this paper proposes a fast intra mode decision algorithm based on texture information and neighbor spatial-temporal information. Firstly, the Sobel operator to obtain the gradient information of the current PU (Prediction Unit) is utilized to determine available candidate intra predicted modes. Furthermore, the predicted candidate modes are updated by the most probable modes (MPMs) obtained from the neighbor spatial-temporal correlation encoded blocks. Experimental results are implemented on RD17.0, and about 36.03% encoding time is reduced on average with negligible coding performance loss.**

Keywords-AVS2; intra prediction; spatial-temporal MPM; texture information; Sobel operator.

I. INTRODUCTION

As the latest generation of audio and video coding standards developed by the Audio Video Coding Standard (AVS) Working Group of China, AVS2 consists of audio and video data. Compared with the previous generation coding standard H.264/AVC [1] or AVS [2], AVS2 has more than doubled coding efficiency under the same reconstructed video quality, which achieves the same level of performance as H.265/HEVC [3]. Especially in the compression efficiency of the scene-like video (such as surveillance videos), AVS2 has doubled compression efficiency than HEVC and 4 times compression efficiency than H.264/AVC.

Compared with the previous generation of encoders, AVS2 has made major adjustments in types of frame, coding structure, intra and inter prediction techniques and so on. The AVS2 frame types introduce the F frame type (a bidirectional forward reference frame) in addition to the I frame type, the B frame type and the P frame type. Similar to H.265/HEVC standard, AVS2 adopts block-based coding, prediction and transform coding techniques with the quad-tree partition

structure. With the quad-tree method the largest coding unit (LCU) can be recursively split into four equal-sized smaller CUs (Coding Units) until the CU size is equal to the smallest coding unit. To achieve the best prediction unit (PU) for each CU, one PU can be further split into smaller units with square or non-square shapes. The partition structures of different CUs and PUs are shown in Fig. 1. In conventional video coding process, the final optional prediction and transform modes for each CU mainly depends on rate distortion optimization (RDO) decision procedure. Namely all of kind's CU depth and PU partition modes are required to traverse the RDO to determine the optimal coding results least with the minimum rate distortion cost. The I frame is used as the first and key frame of the video sequence, which is referenced by other type frames, which has great significance for coding efficiency and coding quality. However, there are up to 33 available prediction modes in intra coding for AVS2 and 9 candidate prediction modes are selected to RDO for all PUs with size from 4x4 to 64x64. Obliviously, the intra coding complexity is too large for application.

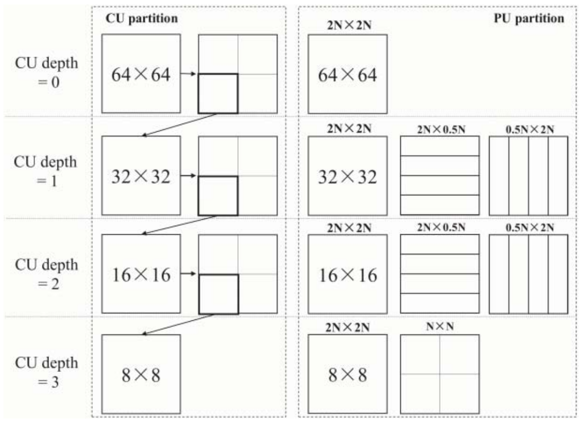

Figure 1. CU and PU partitions in AVS2

In order to reduce the computational complexity of intra coding in AVS2, quality ensured fast algorithms are researched recently. Several algorithms are proposed to reduce the complexity of intra coding, which focus on the CU partition or luminance prediction mode decision. In this paper, we mainly focus on the prediction modes reduction. Specially, for one certain PU, efficient fast intra prediction modes decision algorithm will help to save the time to choose the best mode for it. A fast rough mode decision method based on edge detection for intra coding is introduced by one histogram mapping method to achieve the computation time reduction [4]. Lu etc. [5] has presented a simple approach to reduce the number of candidate prediction modes to no more than 6 for AVS2 all intra configuration. A fast algorithm is proposed to adaptively reduce the number of candidate intra mode for RDO by utilizing the mode activity defined by analyzed the rough costs of all modes [6]. Even though these methods can improve compression efficiency, only spatial features are taken into account for estimate candidate prediction modes without necessary temporary considerations, and they only validate their performance within all the intra circumstance.

This paper proposes a novel fast algorithm of intra mode decision which optimizes rough mode decision (RMD) and RDO respectively. For the RMD process, the texture information of the image is obtained by the Sobel operator, and 9 candidate modes are selected from the prediction modes to form a first-level candidate list. Besides, with the spatial and temporal correlation the candidate list is updated. With the fast algorithm, the number of modes of updated candidates of different PU sizes for RDO traversal process can be optimized to reduce its complexity with guaranteed performance. Experimental results demonstrates that the proposed fast algorithm can achieve more than 36.03% time saving while the performance loss is negligible.

The rest of the paper is organized as follows. In section II, the intra prediction process in AVS2 is introduces. Section III presents the proposed fast algorithm of intra mode decision. The experiment results are shown in section IV while the conclusions are given in Section V.

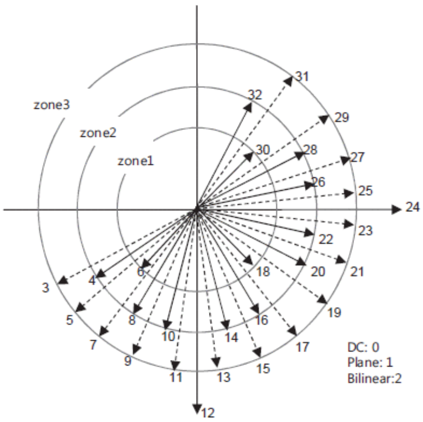

Figure 2. Intra prediction modes in AVS2.

II. AVS2 Intra Prediction Process

AVS2 adapts a flexible video coding framework including 3 different logical partition structures, coding unit, prediction unit and transform unit [7]. The maximum size of the CU is 64x64 while the minimum size of the CU is 8x8. The CU is further split into PUs and TUs as a basic unit of prediction and transform/quantization. According to the size of the CU the minimum size of PU can be split into 4x4. When the CU size is 32x32 or 16x16, the non-square PU are introduced to intra prediction, which is called short-distance intra prediction technology [8]. Non-square partitioning can bring higher accuracy to short-distance predictions, but it also brings higher computational complexity. In each luminance PU prediction, AVS2 utilizes 33 prediction modes as shown in Fig 2, including DC, Plane, Bilinear, and 30 kinds of angle modes, for which utilize the 1/32 precision sub-pixel interpolation technique with a 4-tap linear filter to get the prediction blocks for different angular.

In the AVS2 reference software, intra coding mainly includes two processes of RMD and RDO. Above of all, 9 candidate modes are chosen by the minimum J_{HAD} cost as

$$J_{HAD} = SATD + \lambda \cdot R_{mode} \qquad (1)$$

where SATD represents the sum of absolute Hadamard transformed coefficients of the residual signals between the CU and PU. λ is the Lagrange multiplier and R_{mode} is the estimated number of bits to encode the intra mode information. Then the RDO procedure will be traversed for all candidate modes to select the optimal prediction mode.

III. Proposed Fast Intra Mode Decision Algorithm

In AVS2, the optimal prediction mode is selected for the current PU in a combination of RMD and RDO. The RMD process requires full traversal of 33 prediction modes, and 9 candidate modes are obtained by equation (1). Then, traversal of RDO is performed on the 9 candidate modes, and the optimal prediction mode is selected. The fast algorithm we proposed in this paper optimizes the intra prediction complexity by two steps. Firstly, the Sobel operator is used to obtain the gradient information of the current PU. Compared with the 30 angle modes, 12 optimal angular modes are selected. Combined with the 3 non-angular modes, 9 candidate modes are selected by J_{HAD} calculation and comparison. Then, with the temporal and spatial correlation MPMs, the candidate list obtained in the first step is updated, and different numbers are selected according to different PU sizes for RDO. A large number of experimental statistics show that more than 96.9% of the optional mode will be selected from the top five of the candidate list, and CM0 represents the probability statistics of optional mode selected from the first of candidate list which is shown as below. The numbers are shown as Table I. It can be seen that the prediction modes of all PUs with different size for RDO are reduced from 9 to selected numbers, which means the complexity of intra RDO can be reduced. Detailed process of the fast algorithm are presented below.

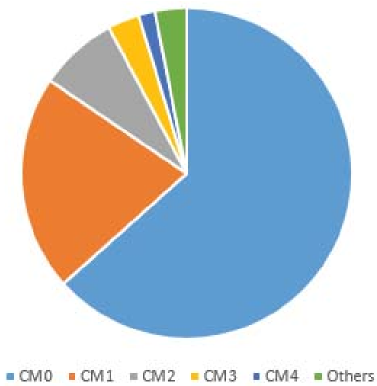

Figure 3. Probability statistics of optional mode selected from candidate list

TABLE I. SELECTED PREDICTION MODES FOR RDO

PU size	Number of PU modes in RDO
4x4	5
8x8	5
16x16	3
32x32	3
64x64	3

A. *First Level Candidate List Obtained with Sobel Operator*

As analyzed above, the texture information are utilized by the gradient of each pixel in current PU to obtain the first-level candidate modes list. In the paper, we use Sobel operator to get the gradient information. By the gradient several angular prediction modes are selected for the following RDO. The description in detail of the first step is shown as follows:

Firstly, two convolution masks of 3x3 metric are utilized to calculate the gradient values of each pixel in both horizontal and vertical of the current PU as shown as,

$$Gx_{ij} = p_{i+1,j-1} + 2 \times p_{i+1,j} + p_{i+1,j+1} - p_{i-1,j-1} \\ -2 \times p_{-1i,j} - p_{i-1,j+1} \quad (2)$$

$$Gy_{ij} = p_{i-1,j-1} + 2 \times p_{i,j-1} + p_{i+1,j-1} - p_{i-1,j+1} \\ -2 \times p_{i,j+1} - p_{i+1,j+1} \quad (3)$$

where the $p_{i,j}$ represents a pixel in a position of (i, j) and the Gx_{ij} and Gy_{ij} represents the horizontal and vertical gradient magnitude of the pixel, respectively. Then the prediction angular should be calculated with equation (4), which reflects intra modes' angular. Through the estimated angular θ , the intra modes corresponding to the 12 nearest angular are chosen into the RMD procedure.

$$\theta = \arctan \frac{\sum_{i=1,j=1}^{n} Gy_{ij}}{\sum_{i=1,j=1}^{n} Gx_{ij}} \quad (4)$$

Combined with the 3 non-angular prediction 9 candidate prediction modes are selected with RMD to form the first level list by the equation (1).

B. *Candidate List Upated Based on Temporal-spatial MPMs*

As we all know, usually frames in one video possess temporal and spatial correlations. The current PU is similar to the neighboring left and up PU in texture information, and in addition, the same position of its previous frame has significance for current PU. Hence, based on the analyzation above, the previous candidate prediction modes can be furtherly updated according to the information of left, up and previous PUs. Specially, the neighbor spatial-temporal PU's MPM modes are jointed with the current candidate modes together to find more optimal modes and reduced numbers. The decision procedure are shown as follows.

Firstly, the best prediction mode of left neighboring PU are selected to compare with modes in the candidate list. If there is the mode is not listed in the first level list, then update the list by add the mode. If there is a mode in the list which is the same to the mode, then unchanged the list. Similar to the first step, the procedure will be repeated for the best modes of the up and previous frame PUs. The overall flow cart of the proposed fast algorithm are shown as Fig. 4. LF represents the optional prediction mode of the same position PU in the last frame, which is used to update the previously obtained candidates. Through a large number of experimental statistics, we found that the optional intra prediction mode will be selected from the top several candidate modes by RMD probably. We adapt different numbers for RDO as shown in Table 1 according to the balance of performance and reduced time.

It can be seen that, the top 3 or 5 candidate modes are selected finally for different PUs. For example, for the PU 4x4, only 5 intra modes are utilized for RDO to deicide the best modes. It should be noted that, the LF is the optimal candidate intra mode for the previous co-located CU. When the previous co-located CU's best prediction mode is intra mode, the LF mode is set as the best intra prediction mode. While the corresponding prediction mode is inter mode, namely the PU mode for the previous co-located CU is skip mode, direct mode or other inter prediction mode, the LF value will be chosen as the intra mode with the closet rate distortion cost to the best inter mode's cost.

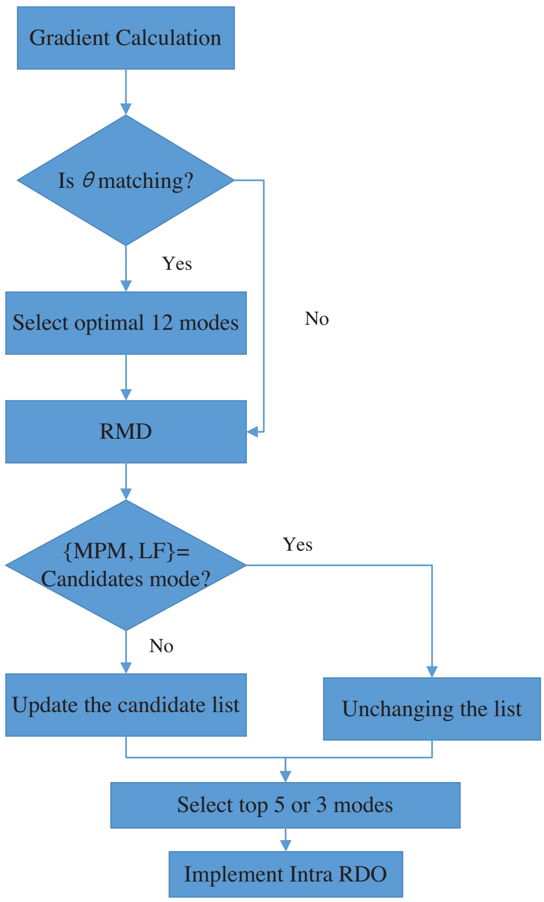

Figure 4. The overall flow of the proposed algorithm.

IV. EXPERIMENT RESULT

In order to evaluate the performance of the proposed algorithm, the experiment is implemented on RD17.0, reference software of AVS2 with the low-delay P configuration. The intra period is 60 with 30 fps. The proposed algorithm is applied for all intra coding modes, namely all the CUs in I type frame and the CUs with intra prediction modes in P type frame. The testing condition contains 12 sequences with different resolutions, with QPs set as 20, 26, 32 and 38. The performance of the proposed algorithm is evaluated against the RD17.0 reference software as anchor without any modification in term of encoding time-reduction as equation (5) and BD-Rate results, which the larger with positive, the more bitrate increase with the same PSNR (Peak Signal Nosie Ratio)constraint.

$$\triangle\,\text{Time} = \frac{\text{Time}_{proposed} - \text{Time}_{RD17.0}}{\text{Time}_{RD17.0}} \times 100\% \qquad (5)$$

The experimental results as shown in Table II indicate the proposed algorithm of intra prediction is faster and more efficient than the RD17.0 reference software of AVS2. We can see that the encoding time of intra has obviously reduction with negligible bitrate increase under the same PSNR quality. From the Table II, it's easy to find that the most encoding time-reduction is achieved for the sequence of *Snowscene* with the minimum BD-Rate increase from the testing sequences while the minimum encoding time-reduction is obtained for *KristenAndSara* with the most BD-Rate increase. We also make the performance comparisons with the method in [6]. Compared with it [6], we achieve 36.03% time reductions on average, which is more than 15% time saving than the method in [6] while for both of them the BD-Rate increase is very low. Thus our algorithm outperforms the other one with better compression efficiency. Detail performance comparison is shown as Fig.5 with the sequence *Snowscene* as the example. Obliviously, the performance is very close to the anchor, which means loss for proposed method is negligible.

TABLE II. TIME SAVING AND BD-RATE COMPARISON

Resolution	Sequences	BD-Rate		△ Time (%)	
		proposed	[6]	proposed	[6]
1920x1080	BQTerrace	0.3	0.3	-37.47	-19.72
	Cactus	0.5	0.5	-39.65	-25.02
	Snowscene	0.2	0.3	-39.43	-26.99
832x480	RaceHorses	0.8	0.2	-35.7	-21.76
	PartyScene	0.3	0	-38.41	-30.2
	BasketballDrill	1.1	0.2	-36.62	-25.39
416x240	BasketballPass	0.8	0.3	-33.93	-24.45
	BQSquare	0.2	0.1	-37.52	-33.21
	RaceHorses	0.7	0.1	-37.32	-28.65
1280x720	FourPeople	1	0.7	-35.62	19.79
	Johnny	1.2	1.1	-31.27	-15.67
	KristenAndSara	1.2	0.9	-29.47	-17.08
Average	-	0.7	0.4	-36.03	-20.7

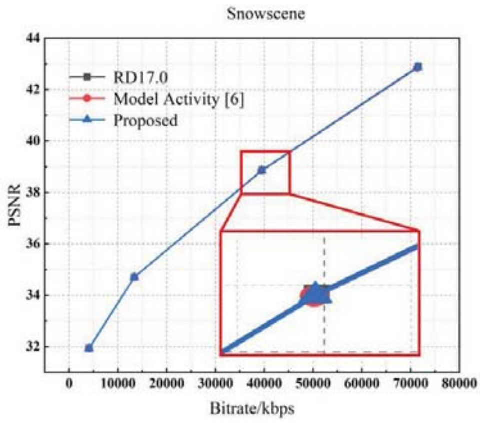

Figure 5. Performance comparison for sequence *Snowscene* for proposed method and mode activity [6].

V. CONCLUSION

The paper proposes a fast algorithm for intra mode decision to optimize RMD and RDO, respectively. In the process pf RMD, we utilize gradient obtained by Sobel operator to achieve the selection of intra prediction modes. Furthermore, we update the candidate list through temporal-spatial correlated prediction modes for RDO. Experiments are implemented in RD17.0 reference software of AVS2 with 12 sequences in 4 different QPs. The results show that the fast algorithm achieves 36.03% time reduction, with 0.7% bitrate increase, which is negligible performance loss. The algorithm can be also efficiently applied in other international standards of video coding platforms such as HEVC/H.265 etc. In the future, we will explore more efficient fast algorithms, for example, the better fast CU or PU split algorithm to improve the compression efficiencies furtherly.

ACKNOWLEDGMENT

This work is partially supported by National Science and Technology Major Project under contract No.2016YFC0801001 and National Engineering Laboratory of Video Technology.

REFERENCES

[1] T. Wiegand, G. J. Sullivan, G. Bjontegaard and A. Luthra, "Overview of the H.264/AVC video coding standard," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-576, July 2003.

[2] Liang Fan, Siwei Ma and Feng Wu, "Overview of AVS video standard," *2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763)*, Taipei, 2004, pp. 423-426 Vol.1.

[3] G. J. Sullivan, J. Ohm, W. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012.

[4] Chung, Byungjin, Joohyeok Kim, and Changhoon Yim. "Fast rough mode decision method based on edge detection for intra coding in HEVC." Consumer Electronics (ISCE 2014), The 18th IEEE International Symposium on. IEEE, 2014.

[5] Y. Lu and H. Zhao, "Simple approach to reduce the number of selected intra luminance prediction modes for RDO computation in AVS2," *2017 IEEE 2nd Advanced Information Technology, Electronic*

and Automation Control Conference (IAEAC), Chongqing, 2017, pp. 911-915.

[6] L. Yu, F. Ge, J. He, B. Sun and F. Dai, "Mode activity based adaptive fast intra mode decision for HEVC/H.265," *2015 8th International Congress on Image and Signal Processing (CISP)*, Shenyang, 2015, pp. 153-157.

[7] Z. He, L. Yu, X. Zheng, S. Ma and Y. He, "Framework of AVS2-video coding," *2013 IEEE International Conference on Image Processing*, Melbourne, VIC, 2013, pp. 1515-1519.

[8] Y. Piao, J. Chen, S. Lee and I. Kim, "Intra coding of AVS2 Video Coding Standard," *2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, Chengdu, 2014, pp. 1-5.