IEICE TRANS. INF. & SYST., VOL.E102–D, NO.3 MARCH 2019

LETTER

Fast Intra Prediction and CU Partition Algorithm for Virtual Reality 360 Degree Video Coding

Zhi LIU[†a)], **Cai XU**[†], ***Nonmembers***, **Mengmeng ZHANG**[†b)], ***Member***, ***and*** **Wen YUE**[††], ***Nonmember***

SUMMARY Virtual Reality (VR) 360 degree video has ultra-high definition. Reducing the coding complexity becomes a key consideration in coding algorithm design. In this paper, a novel candidate mode pruning process is introduced between Rough Mode Decision and Most Probable Mode based on the statistical analysis of the intra-coding parameters used in VR 360 degree video coding under Cubemap projection (CMP) format. In addition, updated coding bits thresholds for VR 360 degree video are designed in the proposed algorithm. The experimental results show that the proposed algorithm brings 38.73% and 23.70% saving in average coding time at the cost of only 1.4% and 2.1% Bjontegaard delta rate increase in All-Intra mode and Randomaccess mode, respectively.

key words: *360 degree video, CMP format, HEVC, fast intra coding*

1. Introduction

Virtual Reality (VR) 360 degree video is a spherical video that cannot be coded directly by traditional video coding framework. Joint Video Exploration Team (JVET) proposes a coding scheme to project the 360 degree image to a 2D plane one and utilizes traditional video coding framework to complete the rest of the coding task [1].

For the projecting process, many projection formats have been proposed, such as Equirectangular projection (ERP) and Cubemap projection (CMP) [2]. CMP format projects the 360 degree video to the circumscribed cube of the sphere and brings almost no stretch to the content located in any area of the sphere, thereby outperforming the ERP format [3]. However, the CMP format brings discontinuous phenomenon at the neighboring area of the edges of the six planes, which means some of the angular modes are excluded naturally. In addition, some objective quality metrics for spherical video have been proposed, i.e., WS-PSNR [4].

For the coding process, the latest coding scheme High Efficiency Video Coding (HEVC) is suggested. A video frame is divided into several coding tree units (CTUs) during coding. In the intra coding process, 35 prediction modes are defined to remove the strong spatial correlation within each frame of video, including a Planar mode, a DC mode, and 33 angular modes. The Rough Mode Decision (RMD) and Most probable Mode (MPM) are adopted to reduce the complexity of Rate-Distortion Optimization (RDO). 35 prediction modes are checked during RMD by calculating Sum of the Absolute Transformed Difference (SATD) first, and 8 modes are selected for 4×4 and 8×8 PUs and 3 modes for 16×16, 32×32 and 64×64 PUs [5]. Typical early coding unit (CU) partition determination algorithms using coding bits as metric to guide the partition [6]. The coding bits thresholds used in these algorithms are derived from traditional video and are inefficient for VR 360 degree video in our experiments. In this study, optimal thresholds for VR 360 degree video are designed from the perspective of statistics.

D. Zhang et al. in [7] proposed a fast intra mode decision algorithm for HEVC based on texture characteristic, which is judged from analyzing of the relation between the results of Rough Mode Decision (RMD) and Most probable Mode (MPM). Y. Wan et al. in [8] proposed a fast intra prediction algorithm for VR 360 degree video in ERP format and focused on the distortion in ERP format. Different from them, in this work, based on the statistical features of the relation between the candidate modes after RMD and the final best mode, a candidate mode pruning process is introduced to resolve the problem in CMP format.

The remaining chapters of this paper are arranged as follows. The second chapter introduces the proposed fast intra prediction and CU partition algorithm under CMP format. The experiment results and conclusion are described in last two chapters respectively.

2. Proposed Algorithm

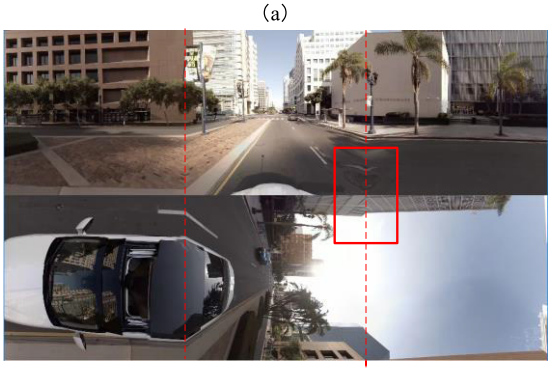
2.1 Candidate Modes Pruning

For CMP 3×2 frame packing, there is discontinuous phenomenon in the packed planes, as shown in Figs. 1 (a) and (b). Some completely unrelated contents are arranged together, which leads to the so-called step effects; due to the projecting limitation and arranging requirement, some continuous contents are split during projecting. Step effects are brought after arranging, which means some of the intra prediction modes are naturally excluded for the blocks in these areas, as shown in Fig. 2. However, the encoding process in HEVC still enumerates all 35 intra prediction modes to find the optimum one, which wastes coding time.

Focused on this problem, a large number of experiments has been conducted in this work to find the relation between the candidate modes after RMD and the final best

Manuscript received May 8, 2018.
Manuscript revised October 6, 2018.
Manuscript publicized December 18, 2018.
[†]The authors are with North China University of Technology, Beijing, China.
[††]The author is with China University of Geosciences, Beijing, China.
a) E-mail: lzliu@ncut.edu.cn
b) E-mail: muchmeng@126.com
DOI: 10.1587/transinf.2018EDL8098

(a)



(b)

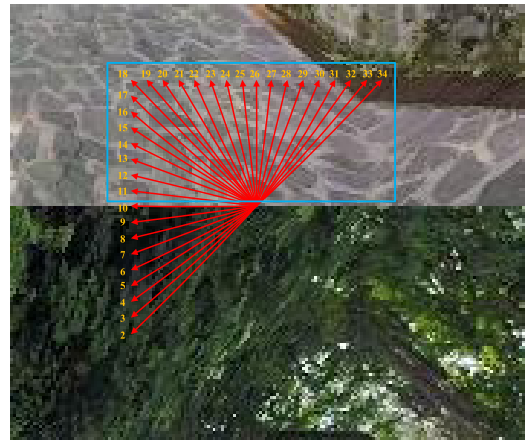
Fig. 1 The problem after arranging in CMP 3×2 frame packing (the areas marked by red rectangle).

mode. The probability of the first candidate mode in the candidate list after RMD to be the final best mode is more than 50%. One of the first two candidate modes is more than 60% (the final best mode). The high probability of the first two candidate modes to be the final best mode indicates that there is a chance to reduce candidate modes and to save coding time in the CMP 3×2 frame packing.

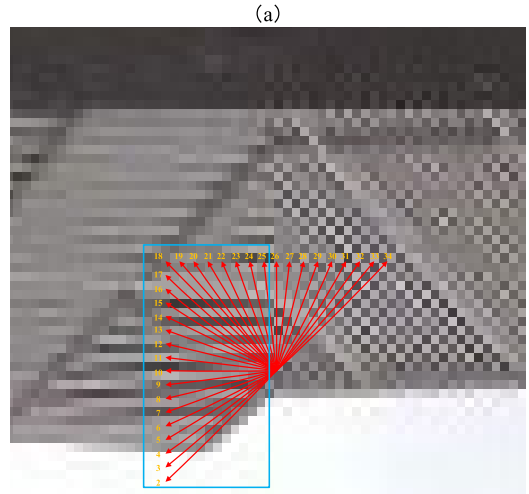
Moreover, two particular modes, DC and Planar, are employed in blocks with smooth content or extremely complicated texture. Step effect areas always have an extremely complicated texture, DC and Planar modes are considered in this kind of areas.

The entire candidate mode pruning process proposed after RMD is described as follows. The prediction unit (PU) depth is classed into two groups: less than or equal to 2 (64×64, 32×32, or 16×16 PU) and greater than 2 (8×8 or 4×4 PU). To distinguish the candidate list before and after the pruning process, the candidate list before pruning is called the original candidate list, and the candidate list after pruning is called the final candidate list.

1) If the first candidate mode (FC) in the original candidate list is an angular mode, then the second angular candidate mode (SAM) is judged further. There are two sub-situations, as follows: FC and SAM are neighboring modes (If the absolute value of two angular modes is less than or equal to 3, the two modes are defined as the neighboring modes.); and FC and SAM are non-neighboring modes. If the FC and SAM in the list are neighboring modes, it is possible that the PU is located at an area with continuous texture. Therefore, the first mode in the original list and four neighboring modes of the first mode (two modes in the left and two modes in the right) are selected to build the final candidate list. If the FC and SAM are non-neighboring modes, it is probable that the PU is located at the step effect areas. In this case, only FC and SAM are select as final candidate list.
2) If the FC in the original candidate list is DC or Planar mode, it probably means that the PU is located at the step effect area, and angular modes are not available. Then, only DC and Planar modes are selected to build the candidate list.



(a)



(b)

Fig. 2 The examples of the blocks located in the neighboring plane edges naturally exclude some of the intra prediction modes (the excluded modes marked by blue rectangle).

Based on the analysis of the statistical data, SC and SAM are judged to be neighboring modes if they are less than or equal to 3. Considering that the big-sized PUs (64×64, 32×32, or 16×16 PU) probably have more texture information, there is no change if FC is angular mode. If the FC in the original candidate list is DC or Planner, MPM can

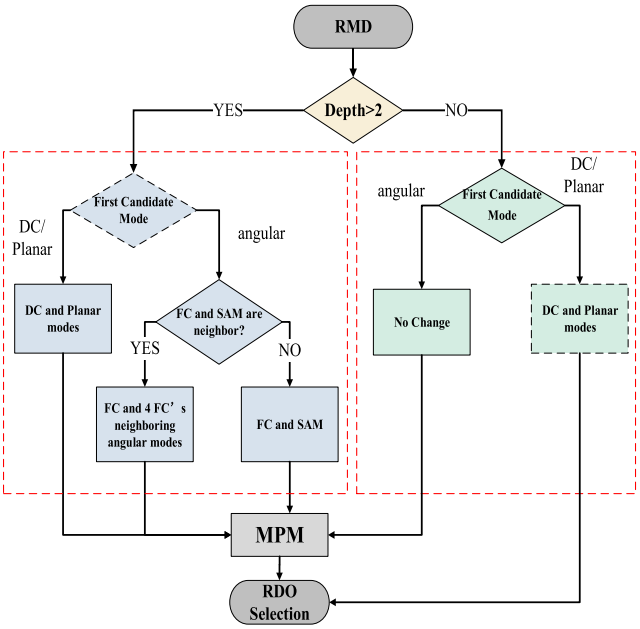


Fig. 3 CMP intra-coding candidate mode pruning process flow chart

Table 1 The CU partition threshold.

QP \ CU size	64x64	32x32	16x16
22	816	384	115
27	480	192	77
32	192	96	43
37	96	48	34

be omitted because it is probable that the PU is located at an area with discontinuous texture. Figure 3 presents the flow chart.

2.2 Fast CU Partition

A close relationship exists between RD-cost and CU coding bits. The larger the RD-cost, the greater the coding bits values, and vice versa. Therefore, many typical early CU partition determination algorithms using coding bits as metric for partition decision. The coding bits thresholds used in these algorithms are derived from traditional video. Our evaluation results show that the coding bits thresholds used for traditional video are inefficient for VR 360 degree one. Hence, we design optimal thresholds for VR 360 degree video from the perspective of statistics data (AerialCity for 4 K, Broadway for 6 K, ChairliftRide for 8 K).

Table 1 shows the new coding bits thresholds designed for VR 360 degree video.

3. Experiment Result

We tested proposed algorithm in All-intra and Randomaccess configuration on HM16.16 plus 360Lib 4.0. And 13 standard test sequences (three sequences for threshold training are removed) provided by GoPro [10], InterDigital [11], Nokia [12], and Letin VR [13] are used, which are the recommended test sequences. And we just tested 60 frames per sequence for time limitation. The test computer is Dell OptiPlex 7050 with Intel (R) Core (TM) i7-7700 CPU @ 3.6 GHz, 8 GB RAM and Windows 10 Professional.

Table 2 The results of candidate modes pruning.

Candidate Modes Pruning only					
Class	Sequence name	All-Intra Mode		Randomaccess Mode	
		BD rate_Y(%)	ΔT(%)	BD rate_Y(%)	ΔT(%)
4K	DrivingInCity	0.5	15.42	0.2	2.54
	PoleVault_le	0.8	15.36	0.5	2.89
	DrivingInCountry	0.5	17.30	0.2	2.99
6K	Balboa	0.4	14.57	0.4	2.82
	BranCastle2	0.5	18.20	0.2	2.37
	Landing2	0.4	15.52	0.3	2.19
8K	Gaslamp	0.7	15.71	0.7	2.63
	Harbor	0.7	18.97	0.5	2.71
	KiteFlite	0.6	16.81	0.4	3.00
	SkateboardInLot	0.7	11.09	0.3	2.73
	SkateboardTrick	0.5	16.08	0.3	2.43
	Train_le	0.3	7.15	0.4	2.74
	Trolley	0.4	5.75	0.4	2.38
Average		0.5	14.46	0.4	2.65

Table 3 The results of fast CU partition.

Fast CU Partition only					
Class	Sequence name	All-Intra Mode		Randomaccess Mode	
		BD rate_Y(%)	ΔT(%)	BD rate_Y(%)	ΔT(%)
4K	DrivingInCity	1.9	48.07	2.9	20.45
	PoleVault_le	0.4	35.74	1.1	24.66
	DrivingInCountry	0.6	34.72	3.3	25.18
6K	Balboa	1.0	45.97	3.1	28.80
	BranCastle2	0.1	19.67	1.8	16.17
	Landing2	0.9	38.29	4.4	7.30
8K	Gaslamp	2.1	37.39	0.4	26.04
	Harbor	1.9	38.69	0.6	24.39
	KiteFlite	0.3	20.88	1.5	20.20
	SkateboardInLot	1.3	44.66	3.7	22.51
	SkateboardTrick	1.0	37.56	1.8	21.37
	Train_le	0.3	32.86	0.1	25.78
	Trolley	0.3	20.75	0.3	24.30
Average		0.9	35.02	1.9	22.09

The proposed candidate modes pruning algorithm and fast CU partition algorithm are tested separately and Table 2 and Table 3 show the experiments results. Table 4 shows the combined experiment results in All-Intra and Randomaccess configuration. Time reduction is calculated by the following Eq. (1), and the objective quality metrics is WS-PSNR:

$$\Delta T = \frac{T_{HM\,16.16} - T_{proposed}}{T_{HM\,16.16}} * 100\% \quad (1)$$

It can be found that each part of the proposed algorithm has performance improvements, and the combined results are better than each one. For the All-Intra mode, the proposed algorithm reduces 38.73% average coding time with just 1.4% Bjontegaard delta rate (BD-rate) increase. Not only B frame is to be encoded in Randomaccess mode, thus, the proposed algorithm reduces 23.70% average coding time with 2.1% BD-rate increase.

Table 4 The combined experiment results.

Class	Sequence name	All-Intra Mode		Randomaccess Mode	
		BD-rate_Y(%)	ΔT(%)	BD-rate_Y(%)	ΔT(%)
4K	DrivingInCity	2.2	51.75	3.0	27.29
	PoleVault_le	1.0	34.20	1.5	25.93
	DrivingInCountr	0.9	39.25	3.5	27.52
6K	Balboa	1.5	49.11	3.3	26.84
	BranCastle2	0.6	22.63	1.8	18.99
	Landing2	1.3	45.49	3.8	25.71
8K	Gaslamp	2.8	44.99	1.0	26.53
	Harbor	2.5	42.09	1.0	23.57
	KiteFlite	1.0	23.63	1.9	15.86
	SkateboardInLot	1.9	47.68	3.4	20.91
	SkateboardTrick	1.5	42.74	2.0	17.04
	Train_le	0.6	36.99	0.5	25.80
	Trolley	0.8	22.98	0.7	26.16
Average		1.4	38.73	2.1	23.70

4. Conclusion

In this paper, a fast intra prediction and CU partition algorithm under CMP format for the VR 360 degree video is proposed based on the statistical data of this kind of video. In the proposed algorithm, a candidate mode pruning process is designed between RMD and MPM for CMP 3×2 frame packing, it reduces coding time during intra mode prediction. Moreover, the CU partition coding bits thresholds are designed for VR 360 degree video to reduce the computational complexity from a statistical perspective. As demonstrated, our algorithm greatly reduces the coding time without significantly degrading coding quality.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No.61370111), Beijing Municipal Natural Science Foundation (No.4172020), Great Wall Scholar Project of Beijing Municipal Education Commission (CIT&TCD20180304), Beijing Youth Talent Project (CIT&TCD 201504001), and Beijing Municipal Education Commission General Program (KM201610009003).

References

[1] Y. He, X. Xiu, Y. Ye, V. Zakharchenko, et al., "AHG8: Algorithm description of projection format conversion in 360Lib," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC 1/SC 29/WG 11, 5th Meeting, document JVET-E0084, Geneva, Switzerland, Jan. 2017.

[2] M. Coban, G. Van der Auwera, and M. Karczewicz, "AHG8: Adjusted cubemap projection for 360-degree video," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC 29/WG 11, 6th meeting, document JVET-F0025, Hobart, Australia, April 2017.

[3] M. Zhou, "AHG8: A study on compression efficiency of cube projection," Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC 29/WG11, 4th Meeting, document JVET-D0022, Chengdu, China, Oct. 2016.

[4] Y. Sun, A. Lu, and L. Yu, "AHG8: WS-PSNR for 360 degree video objective quality evaluation," Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 4th Meeting, document JVET-D0040, Chengdu, China, Oct. 2016.

[5] J. Lainema, F. Bossen, W.-J. Han, J. Min, and K. Ugur, "Intra coding of the HEVC standard," IEEE Trans. Circuits Syst. Video Technol., vol.22, no.12, pp.1792–1801, 2012.

[6] M. Zhang, H. Zhang, and Z. Liu, "A Novel Algorithm to Decrease the Computational Complexity of HEVC Intra Coding," 2016 Data Compression Conference (DCC), Snowbird, UT, pp.639–639, 2016.

[7] D. Zhang, Y. Chen, and E. Izquierdo, "Fast intra mode decision for HEVC based on texture characteristic from RMD and MPM," 2014 IEEE Visual Communications and Image Processing Conference, Valletta, pp.510–513, 2014.

[8] Y. Wang, Y. Li, D. Yang, and Z. Chen, "A fast intra prediction algorithm for 360-degree equirectangular panoramic video," 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, pp.1–4, 2017.

[9] H. Bai, C. Zhu, and Y. Zhao, "Optimized Multiple Description Lattice Vector Quantization for Wavelet Image Coding," IEEE Trans. Circuits Syst. Video Technol., vol.17, no.7, pp.912–917, July 2007.

[10] A. Abbas and B. Adsumilli, "AHG8: New GoPro test sequences for virtual reality video coding," Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 4th Meeting, document JVET-D0026, Chengdu, China, Oct. 2016.

[11] E. Asbun, Y. He, Y. He, and Y. Ye, "AHG8: InterDigital test sequences for virtual reality video coding," Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 4th Meeting, document JVET-D0039, Chengdu, China, Oct. 2016.

[12] S. Schwarz and A. Aminlou, "Tampere pole vaulting sequence for virtual reality video coding," Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 4th Meeting, document JVET-D0143 Chengdu, China, Oct. 2016.

[13] W. Sun and R. Guo, "Test sequences for virtual reality video coding from letinvr," Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 4th Meeting, document JVET-D0179, Chengdu, China, Oct. 2016.