Available online at www.ijpe-online.com

International Journal of Performability Engineering

vol. 14, no. 8, August 2018, pp. 1842-1847
DOI: 10.23940/ijpe.18.08.p21.18421847

Low Complexity Texture Video Coding for Multi-View Video System Reliability

Shuru Liu, Shuaichao Wei, and Qiuwen Zhang*

College of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou, 450002, China

Abstract

3D extension of high efficiency video coding (3D-HEVC) has become a video compression international standard for multi-view video system applications. The mode prediction process must search for optimal prediction modes and depth levels with a minimum rate-distortion cost for the coding unit (CU). New coding techniques with improved texture video coding efficiency have been introduced; however, they also increase computational burden, which obstructs 3D-HEVC encoder reliability applications. Therefore, the development of a novel texture video coding is necessary to reduce 3D-HEVC computational complexity. In this paper, the proposed algorithm will exploit an adjusted method of DE and ME based on motion homogeneity of texture videos. The algorithm includes early SKIP/Merge mode decision and adaptive motion search range adjustment. As shown by the simulation results, the proposed fast texture video coding method can reduce computational complexity by approximately 56.6% on average while resulting in a tiny loss of quality.

Keywords: 3D-HEVC; texture video; SKIP mode; search range

(Submitted on May 20, 2018; Revised on June 19, 2018; Accepted on July 26, 2018)

© 2018 Totem Publisher, Inc. All rights reserved.

1. Introduction

With the advancement of 3D television (3DTV) technology, 3D displays have received much attention in the electronics market. The multi-view video (MVV) technique has been developed for 3D video applications. MVV is captured via multiple synchronized cameras of the same scene, offering a better experience than traditional 2D video since it can be watched from different viewing angles interactively. However, the MVV system requires huge information involved with the multi-view content and a high transmission rate, which has high computational burden of the transmission coding. To address this issue, an advanced multi-view video coding (MVC) is standardized by the joint video team (JVT) [1]. MVC adopts the inter-view component to improve MVV system compression capability [2], which also supports spatial and temporal prediction with the H.264/AVC [3]. By doing this, a reduction in MVV bitrate related to independent coding of the multi-view is obtained without the loss of compression quality.

Recently, fast algorithms were presented in terms of computation burden reduction of HEVC in the literature. A fast mode decision and CU size decision algorithm are presented in [4] and possess a reduced set of depth levels for the HEVC intra prediction process. A fast CU depth decision algorithm [5] focuses on encoding time saving by utilizing the connection of depth information within the spatio-temporal treeblock and the current treeblock. A mode selection and complexity allocation method is reported in [6] based on linear programming among the frames. A novel zero block (ZB) detection algorithm is presented in [7] to check pseudo ZBs while attempting to shun the enhancement of the coding time for original encoder. A CU selection algorithm [8] can realize and early skip the specific inter CUs that utilize the pyramid motion divergence in HEVC. A fast CU size decision method that is temporally and spatially dependent is also explained in our previous work [9] that achieved computation complexity reduction for the HEVC encoder. All these fast CU size prediction methods will reduce computational complexity effectively with small performance loss. However, the aforementioned enumerative methods are not suitable for 3D-HEVC encoders, which do not exploit the inter-view advanced residual prediction and inter-component correlation to achieve time-saving.

* Corresponding author.
E-mail address: zhangqwen@126.com

Basic research was conducted on 3D-HEVC in previous papers [10-11]. However, the advantages of the texture motion homogeneity in the 3D-HEVC have not been fully exploited. Therefore, a low complexity texture video coding algorithm based on motion homogeneity correlation is proposed. Our algorithm is comprised of an early SKIP/Merge mode decision and adaptive motion search range adjustment. Experimental results prove that the coding time of the MVV system can be saved by using our low complexity texture video algorithm while maintaining a negligible RD performance loss.

2. Proposed Low Complexity Texture Video Coding Algorithm

2.1. Early SKIP/Merge Mode Decision

The SKIP/Merge mode has the advantages of good performance and low computational burden, and the compensated block is easy to search by utilizing the motion vector predictor directly. Therefore, different size ME and DE computations will be omitted if the SKIP/Merge mode is pre-decided. However, the aforementioned decision procedure is delayed since the computation of RD costs in the original encoder need to fully search the candidate mode and then choose the SKIP/Merge mode with the least costs. Usually, many texture video treeblocks are contained in a homogenous region and have a motionless object, and these blocks are most likely to choose the SKIP/Merge mode. Based on the above analysis, it is essential to develop an early SKIP/Merge mode decision method with a static homogenous region and then skip the full variable-size ME and DE search procedure.

For texture video treeblocks with static homogenous regions, only the SKIP and Merge mode can be used as the best mode and terminate the 3D-HEVC mode decision early. Therefore, the motion of the static homogenous region can be easily extracted using the SKIP/Merge mode, and this process will have lower energy residual with RD calculation compared to other prediction modes. In addition, it is unnecessary to further decide the mode decision of 3D-HEVC.

To evaluate the proposed decision algorithm, a simulation is performed and the results are shown in Table 1. After the full mode decision under the CTC conditions, we analyze the prediction mode distributions of texture video treeblocks with static homogenous regions. As shown in Table 1, the probabilities of a homogenous region choosing the SKIP mode and the Merge mode are 93.5% and 3.5%, respectively. Meanwhile, the probabilities of all the other modes are less than 0.9%. Based on the above analysis, the optimal mode to select SKIP/Merge can reach about 97.0%, and the other modes will omit to check. Therefore, the proposed early SKIP/Merge decision will reduce unnecessary ME and DE.

Table 1. Prediction mode distribution for treeblocks with static homogenous regions

Sequences	SKIP Mode(%)	Merge Mode(%)	Inter 2N×2N(%)	Intra 2N×2N(%)	Inter 2N×N(%)	Inter N×2N(%)	Intra N×N(%)	Small inter modes(%)
Kendo	94.7	3.5	0.6	0.3	0.2	0.2	0.1	0.4
Balloons	95.4	3.2	0.4	0.2	0.3	0.1	0.1	0.3
Newspaper	92.9	4.1	0.8	0.4	0.4	0.3	0.2	0.9
Shark	89.1	4.8	1.8	1.4	0.8	0.6	0.3	1.2
Undo_Dancer	88.9	5.2	1.5	1.1	0.8	0.8	0.4	1.3
GT_Fly	94.2	3.2	0.9	0.3	0.4	0.3	0.1	0.6
Poznan_Street	95.1	2.6	0.6	0.2	0.3	0.3	0.1	0.8
Poznan_Hall2	97.8	1.5	0.2	0.1	0.1	0.1	0.0	0.2
Average	93.5	3.5	0.9	0.5	0.4	0.3	0.2	0.7

2.2. Adaptive Motion Search Range Adjustment

Motion search, which is important to the 3D-HEVC system, is the process of searching for an optimal matched treeblock in temporal frames that have large computation complexity. The search range is defined as the motion information between the current treeblock and the matched treeblock. Therefore, a larger motion search window will bring a higher computational burden. Otherwise, a smaller motion search window will result in worsened video quality since the matched blocks are inaccurate. A suitable motion search window will have a tradeoff between low computational burden and good RD performance. In 3D-HEVC, the search range is fixed and large to check for each block in texture video. Based on the above analysis, a large fixed motion search range has a higher texture video coding efficiency but can also lead to a huge computational burden. Because the optimal motion search range is highly dependent on the treeblock motion homogeneity characteristic, it is unnecessary to maintain the fixed motion search window in 3D-HEVC. Therefore, the fixed motion search window is unnecessary to inherit from the original 3D-HEVC. Based on the above analysis, we skip some specific range levels that rarely emerge in the static homogenous region.

In Table 2, *"SR1"*, *"SR2"*, *and "SR3"* represent the windows of $\left[SR/16\times SR/16\right]$, $\left[SR/4\times SR/4\right]$, and $\left[SR/2\times SR/2\right]$ respectively. Based on the advantage of motion homogeneity, the motion search distribution for different texture video treeblocks are shown in Table 2. The texture video treeblocks in static homogenous region that select the $\left[SR/4\times SR/4\right]$ window can reach up to 97.0%. Therefore, if the search range is up to $\left[SR/4\times SR/4\right]$, almost all texture motion vectors are covered. Meanwhile, texture video treeblocks in motion complex regions select $\left[SR/16\times SR/16\right]$, $\left[SR/4\times SR/4\right]$, and $\left[SR/2\times SR/2\right]$ windows only about 48.5%, 73.7%, and 82.7% of the time, respectively, and thus the full motion search range may not skip. Based on the results of Table 2, a dynamical motion adjust range in the 3D-HEVC texture coding process is defined by

$$Search\ range=\begin{cases} SR/4\times SR/4, & Treeblock\in static\ homogenous\ region \\ SR\times SR, & Treeblock\in motion\ complex\ region \end{cases} \tag{1}$$

Table 2. Motion search distributions for two types of treeblocks in texture video coding

Sequences	Treeblocks in static homogenous region			Treeblocks in motion complex region		
	*SR*1 (%)	*SR*2 (%)	*SR*3 (%)	*SR*1 (%)	*SR*2 (%)	*SR*3 (%)
Kendo	82.2	97.1	97.5	50.2	77.3	84.1
Balloons	80.4	96.3	97.2	48.3	74.8	82.7
Newspaper	83.7	97.8	98.3	51.7	79.2	86.1
Shark	79.6	95.2	96.5	35.8	58.6	72.4
Undo_Dancer	78.9	96.1	97.1	34.7	56.3	71.6
GT_Fly	84.1	97.5	98.2	54.2	79.8	87.2
Poznan_Street	82.4	96.6	97.3	52.7	78.6	85.3
Poznan_Hall2	88.6	99.2	99.4	60.2	84.9	91.8
Average	82.5	97.0	97.7	48.5	73.7	82.7

With the proposed motion search range adjustment method, most texture video treeblocks of 3D-HEVC can omit useless candidate ranges. Based on the above three methods, the depth level range is selectively enabled, the number of candidate modes in texture video is reduced, and the search range can be adaptively decided. Finally, select the best mode from texture video candidate modes in 3D-HEVC encoders.

3. Experimental Results

For evaluating the availability of the computational complexity reduction on the low complexity texture video coding algorithm, the experiment is performed with the reference software (HTM ver.16.0). It adopts eight sequences that are published by JCT-3V, in which the variation of motion and texture for "Undo_Dancer" and "Shark" is rich. Meanwhile, "Kendo", "Balloons", "Newspaper", "GT_Fly", and "Poznan_Street" have a smaller variation of motion and texture, and "Poznan_Hall2" has the smallest motion and texture variation. The experiments are performed under common test conditions (CTC) configuration. Test conditions: 3 view of center-left-right (in coding order), P-I-P inter-view, group of pictures (GOP) in temporal prediction is 8, intra 24 frames rate, texture video quantization parameter (QP) values for independent view: 40, 35, 30, 25, depth map QP values: 45, 42, 39, 34. The "VSRS-1D-Fast" is used. As shown in Tables 3-5, our algorithm outperforms both the classical 3D-HEVC encoder and the fast methods [6, 9]. The results of video quality are measured by the signal-to-noise ratio (PSNR) and bitrates, and runtime is used to measure computation complexity reduction. The Bjontegaard Delta PSNR (BDPSNR) denotes PSNR gain, Bitrate (BDBR) shows bitrates increase, and "*Dtime* (%)" is the time variation:

$$Dtime=\frac{Time_{proposed}-Time_{origional}}{Time_{origional}}\times 100\% \tag{2}$$

Where $Time_{proposed}$ and $Time_{origional}$ denote the sum of the running time of the proposed algorithm and 3D-HEVC encoders, respectively. “Texture video” represents the results for texture views. In this condition, only the PSNR and bitrate of texture video are evaluated. The “Texture video BDBR / BDPSNR” shows the results of the views of coded texture video. The texture video PSNR is the mean value of the three views. The bitrate of the three views added together can obtain the texture video bitrate. Furthermore, the way to measure coding efficiency improvement is to consider the virtual views, while the intermediate views are synthesized among each view. The Rendered PSNR can be obtained by calculating the variation between the coded view and original rendered view. The Rendered PSNR can be calculated by exploiting a mean PSNR of all synthesized views. Then, the Rendered bitrate can be extracted by adding the bitrate of the three views of coded texture video and depth map.

3.1. Performance Evaluation of the Individual Algorithms

Tables 3 and 4 show the individual algorithms, including the early SKIP/Merge mode decision (ESMMD) and adaptive motion search range adjustment (AMSRA), when they are utilized alone or with the initial 3D-HEVC encoder. For the ESMMD algorithm, 17.2% of encoding time on average is saved with a maximum of 25.3% in “Poznan_Hall2” and a minimum of 12.4% in “Undo_Dancer”. However, there is a performance loss of about 0.39% bitrate increase or 0.01 dB PSNR drop of texture video, and 0.14% bitrate increase or 0.01 dB PSNR drop for rendered views. This result indicates that ESMMD can reduce useless ME and DE on small CU sizes. The proposed AMSRA algorithm achieves coding time reduction for all test sequences, with an average of 31.3% saved encoding time, a maximum of 39.6% in “Shark”, and a minimum of 24.5% in “Poznan_Hall2”. Meanwhile, the coding efficiency maintains almost the same performance, with 0.24-1.58% bitrate increase (0.01 dB-0.05 dB PSNR drop) for texture videos and 0.06-0.92% bitrate increase (0.00 dB-0.03 dB PSNR drop) for rendered views. Thus, the AMSRA can realize computational complexity reduction and has a similar performance of coding efficiency with the initial encoder.

Table 3. Coding results of ESMMD compared to original 3D-HEVC encoders

Sequences	Texture video		Rendered view		
	BDBR (%)	BDPSNR (dB)	BDBR (%)	BDPSNR (dB)	Dtime (%)
Kendo	0.29	-0.01	0.13	-0.01	-17.6
Balloons	0.25	-0.01	0.11	-0.01	-17.0
Newspaper	0.42	-0.01	0.18	-0.01	-15.4
Shark	0.83	-0.03	0.28	-0.01	-12.9
Undo_Dancer	0.75	-0.03	0.25	-0.01	-12.4
GT_Fly	0.24	-0.01	0.10	-0.01	-18.3
Poznan_Street	0.21	-0.01	0.06	-0.00	-19.1
Poznan_Hall2	0.09	-0.00	0.02	-0.00	-25.3
Average	0.39	-0.01	0.14	-0.01	-17.2

Table 4. Coding results of AMSRA compared to original 3D-HEVC encoders

Sequences	Texture video		Rendered view		
	BDBR (%)	BDPSNR (dB)	BDBR (%)	BDPSNR (dB)	Dtime (%)
Kendo	0.59	-0.02	0.18	-0.01	-28.3
Balloons	0.53	-0.02	0.13	-0.01	-27.8
Newspaper	0.91	-0.03	0.43	-0.02	-31.2
Shark	1.58	-0.05	0.92	-0.03	-39.6
Undo_Dancer	1.47	-0.05	0.85	-0.03	-38.7
GT_Fly	0.88	-0.03	0.23	-0.01	-29.8
Poznan_Street	0.48	-0.02	0.11	-0.01	-30.3
Poznan_Hall2	0.24	-0.01	0.06	-0.00	-24.5
Average	0.84	-0.03	0.36	-0.02	-31.3

3.2. Performance Evaluation of the Overall Algorithm

The coding performance of the overall algorithm, which incorporates ESMMD and AMSRA approaches, is illustrated in Table 5. In Table 5, the overall algorithm will clearly realize time-saving for these selected sequences compared to 3D-HEVC. The time reduction can reach up to 36.5%, with a minimum of 32.4% for “Poznan_Hall2” and a maximum of 38.7% for “Undo_Dancer”. For sequences of “Shark” and “Undo_Dancer” with large global motion, the algorithm saves more than 38.5% computational complexity. As the results show, time reduction is achieved by skipping the unnecessary depth level and mode decision with significant complexity in 3D-HEVC. Meanwhile, the average bitrate increase is 1.00% (or a 0.03 dB PSNR drop) for texture video, and a 0.41% bitrate increase (or 0.02 dB PSNR drop) for rendered views is negligible. Therefore, the proposed overall algorithm achieves a significant reduction in computational burden without hampering RD performance.

Table 5. Coding results of the overall algorithm compared to original 3D-HEVC encoders

Sequences	Texture video		Rendered view		
	BDBR (%)	BDPSNR (dB)	BDBR (%)	BDPSNR (dB)	Dtime (%)
Kendo	0.42	-0.02	0.31	-0.02	-38.1
Balloons	0.75	-0.03	0.16	-0.01	-36.2
Newspaper	1.02	-0.03	0.18	-0.01	-37.4
Shark	2.11	-0.05	1.11	-0.03	-38.5
Undo_Dancer	1.89	-0.04	1.02	-0.03	-38.7
GT_Fly	1.04	-0.03	0.25	-0.01	-34.2
Poznan_Street	0.45	-0.02	0.14	-0.01	-36.1
Poznan_Hall2	0.35	-0.02	0.10	-0.01	-32.4
Average	1.00	-0.03	0.41	-0.02	-36.5

3.3. Performance Comparison with the State-of-the-Art Fast 3D-HEVC Methods

The overall algorithm and other two fast methods (fast encoder decision for texture coding (FEDTC) [10], and online-learning-based complexity reduction scheme (OLCRS) [11]) have a comparison based on the “CTC” condition. Figure 1 shows the coding speed for the proposed algorithm and the original 3D-HEVC, FEDTC and OLCRS methods. From the pictures, the proposed overall algorithm will realize the encoding runtime reduction for all sequences. Furthermore, the whole algorithm has saved much time for the large global motion sequences of “Shark” and “Undo_Dancer”, with more run time-saving. Compared with FEDTC and OLCRS, it achieves the maximum reduction among all the sequences, reaching about 38.7%. All results prove that the low complexity texture video coding for test sequences is efficient and outperforms the recently fast methods with a quicker coding speed performance in the 3D-HEVC encoder.

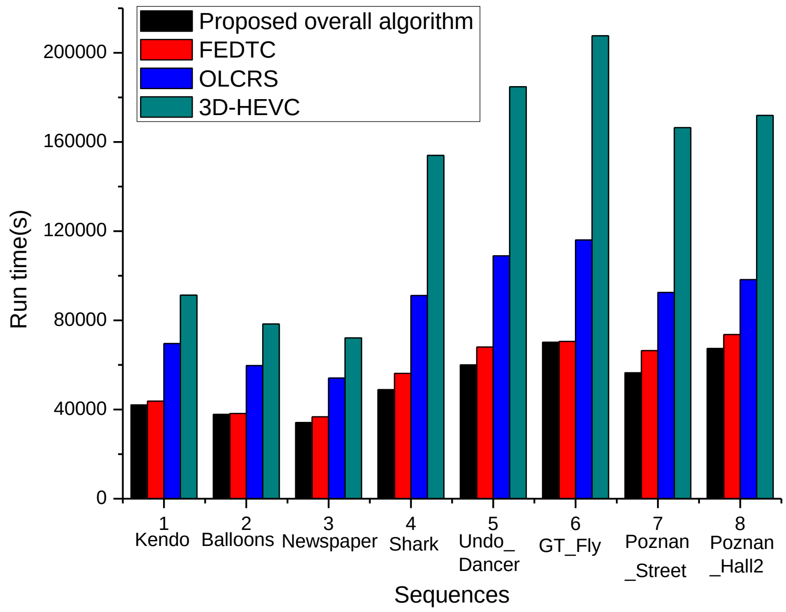

Figure 1. Coding speed of the proposed overall algorithm, FEDTC and OLCRS methods compared to 3D-HEVC encoders.

4. Conclusions

In this paper, a fast texture video coding algorithm that takes advantage of motion homogeneity is proposed. It can achieve computational complexity reduction of the 3D-HEVC encoder, which consists of two individual approaches, i.e., early SKIP/Merge mode decision and adaptive motion search range adjustment. The proposed algorithm employs a 3D-HEVC reference model for testing. The experimental results demonstrated that our algorithm could realize about 36.5% encoding time, while maintaining RD performances with the initial encoder. Furthermore, our method also outperforms some other fast approaches for 3D-HEVC and reaches the highest complexity reduction.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (No. 61771432 and 61302118), Scientific Project (No. 182102210156 and 182102210610), Innovation Talents (No. 17HASTIT022), Young Key Teacher (No. 2016GGJS-087), and Education Department Project (No. 18B510019 and 17B510011).

References

1. H. Urey, K. V. Chellappan, E. Erden, and P. Surman, "State of the Art in Stereoscopic and Autostereoscopic Displays," *Proceedings of the IEEE,* Vol. 99, No. 4,pp. 540-555, 2011
2. Y. Chen and A. Vetro, "Next-Generation 3D Formats with Depth Map Support," *IEEE MultiMedia,* Vol. 21, No.2, pp. 90-94, 2017
3. G. J. Sullivan, J. M. Boyce, C. Ying, J. -R. Ohm, C. A. Segall, and A. Vetro, "Standardized Extensions of High Efficiency Video Coding (HEVC)," *Journal of Selected Topics in Signal Processing*, Vol. 7, No. 6, pp. 1001-1016, 2013.
4. L. Shen, Z. Zhang, and P. An, "Fast CU Size Decision and Mode Decision Algorithm for HEVC Intra Coding," *IEEE Transactions on Consumer Electronics*, Vol. 59, No. 1, pp. 207-213, 2013
5. C. Zhou, F. Zhou, and Y. Chen, "Spatio-Temporal Correlation-based Fast Coding Unit Depth Decision for High Efficiency Video Coding," *Journal of Electronic Imaging*, Vol. 22, No. 4, pp. 043001, 2013
6. T. Zhao, Z. Wang, and S. Kwong, "Flexible Mode Selection and Complexity Allocation in High Efficiency Video Coding," *Journal of Selected Topics in Signal Processing*, Vol. 7, No. 6, pp. 1135-1144, 2013
7. K. Lee, H. Lee, J. Kim, and Y. Choi, "A Novel Algorithm for Zero Block Detection in High Efficiency Video Coding," *Journal of Selected Topics in Signal Processing*, Vol. 7, No. 6, pp. 1124-1134, 2013
8. J. Xiong, H. Li, Q. Wu, and F. Meng, "A Fast HEVC Inter CU Selection Method based on Pyramid Motion Divergence," IEEE Transactions on Multimedia, Vol. 16, No. 2, pp. 559-564, 2014
9. Q. Zhang, J. Zhao, X. Huang, and Y. Gan, "A Fast and Efficient Coding Unit Size Decision Algorithm based on Temporal and Spatial Correlation," *Optik-International Journal for Light and Electron Optics*, Vol. 126, No. 21, pp. 2793-2798, 2015
10. H. R. Tohidypour, M. T. Pourazad, and P. Nasiopoulos, "Online-Learning-based Complexity Reduction Scheme for 3D-HEVC," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 26 , No. 10, pp. 1870-1883, 2016
11. N. Zhang, D. Zhao, Y. Chen, J. Lin, and W. Gao, "Fast Encoder Decision for Texture Coding in 3D-HEVC," *Signal Processing*: *Image Communication*, Vol. 29, No. 9, pp. 951-961, 2014