A NOVEL BIT ALLOCATION ALGORITHM IN MULTI-VIEW VIDEO

Tao Yan[1], In-Ho Ra[2], Deyang Liu[3], Jiajun Hong[1], Jingyu Feng[1] and Shaojie Hou[1]

[1]School of Information Engineering, Putian University, Putian, China
[2]School of Computer, Information and Communication Engineering, Kunsan National University, Gunsan, South Korea
[3]School of Computer and Information, Anqing Normal University, Anqing 246000, China

ABSTRACT

The difficulty of rate control for Multi-view video coding(MVC) is how to allocate bits between views. The results of our previous research including the bit allocation among viewpoints uses the correlation analysis among viewpoints to predict the weight of each viewpoint. But when the scene changes, this prediction method will produce a lot of errors. Therefore, this article avoids this situation happening through scene detection. The core of the algorithm is to first divide all images into 6 types of encoded frames according to the structural relationship between disparity prediction and motion prediction, and improve the binomial rate distortion model, and then perform inter-view, frame layer, and basic unit based on the encoded information. Layer bit allocation and code rate control. In this paper, a reasonable bit rate is allocated between viewpoints based on the encoded information, and the frame layer bit rate is allocated using frame complexity and time-domain activity. Experimental simulation results show that the algorithm can effectively control the bit rate of MVC, while maintaining efficient coding efficiency, compared with the current MVC using JVT with fixed quantization parameters.

KEYWORDS

MVC, Quantization parameters, Bit allocation, Rate distortion model, Basic unit layer.

1. INTRODUCTION

3DTV / FTV system has broad application prospects in many aspects such as digital entertainment, virtual reality, 3D reconstruction, 3D monitoring, expo exhibition, medical treatment, education and so on. However, there are still many key technical problems in signal processing of 3DTV / FTV systems. The relevant international standards have not yet been developed, which is a very challenging and pioneering research field.[1-2] .

The multi-view video is compressed and sent to the channel for transmission and then decoded at the decoder. It is stored or displayed according to user needs. In the entire process of encoding → transmission → decoding, how to solve the compressed multi-view video data Adapt to the coding rate control problem of the channel. If this problem cannot be solved, the coded stream is directly sent to the channel for transmission, which will cause the channel or congestion or idle, which will greatly reduce the channel utilization.

David C. Wyld et al. (Eds): ITCSE, NLCA, ICAIT, CAIML, ICDIPV, CRYPIS, WiMo - 2020
pp. 137-144, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.100512

Previous video compression standards such as MPEG-2, MPEG-4, H.263, H.264 [3-6], etc. have been given code rate control models. However, the multi-view video encoding reference software JMVC of JVT currently has no effective rate control algorithm [7]. The MVC code rate control algorithm must not only reasonably allocate the code rate in time to prevent buffer overflow, but also perform reasonable code rate allocation among the various viewpoints to ensure the video quality balance between the viewpoints. Many scholars at home and abroad have begun to study the rate control of multi-view video coding[8-14]. Multi-view video coding is very complicated. I can first optimize multi-view video coding and extend the constraint relationship between the sports field and parallax field in stereo video to multi-view video coding. The acquisition cameras are corrected and synchronized, and they are arranged in parallel on the same horizontal straight line at equal intervals. In this way, only parallax in the horizontal direction appears in the system, which can better simulate the stereo vision system of the human eye.

Therefore, this paper first analyzes the problems of existing video rate control algorithms and proposes a rate control algorithm for multi-view video. Experimental simulation results show that the algorithm in this paper can maintain efficient encoding efficiency while effectively controlling the bit rate of multi-view video encoding.

2. RATE CONTROL ALGORITHM FOR MULTI-VIEW VIDEO

Multi-view video coding sets more B-pictures in order to improve the coding efficiency. The MPEG organization provides relevant requirements for multi-view video coding. The main requirements are: higher compression rate, lower coding and decoding complexity, and reasonable code rate control. Taking this as the guiding direction of MVC research, researchers have carried out a lot of work in recent years, and some results or suggestions have been submitted to the MPEG organization. Therefore, the bit-rate control of multi-view video coding should increase the bit-rate control of B-frames. Based on previous studies [12], [13], the key point of this paper is how to make a reasonable bit rate allocation among various views according to the requirements of multi-view video coding to ensure the balance of video quality between views. The key steps of the algorithm are as follows:

In the rate control algorithm in this paper, let $T_{GGOP}(sn_{i,0})$ denote the total number of bits allocated to the ith GGOP, and use the weight w_k to indicate the importance of the viewpoint k. A larger w_k indicates that the viewpoint is more important. The total number of bits allocated to the GOP_k at the k-th viewpoint in GGOP is given by equation (1)

$$T_{GOP}(\mathrm{n}_{k,0}) = T_{GGOP}(sn_{i,0}).w_k \quad (1)$$

According to the number of remaining bits $T_{GOP}{}^{*}(k-1)$ of the previous GOP_{k-1}, the number of bits finally allocated by the current GOP_k is:

$$T_{GOP}(\mathrm{n}_{k,0}) = T_{GGOP}(\mathrm{sn}_{i,0}).w_k + T_{GOP}{}^{*}(\mathrm{k}-1) \quad (2)$$

$w_k(k=1,2,\mathrm{L}\ ,N_{view})$ initial value is set to 1,

$$w_k = \frac{\frac{1}{N_{view}} \cdot \sum_{j=0}^{N_{view}-1} C(S_j, S_k)}{\sum_{k=0}^{N_{view}-1} \frac{1}{N_{view}} \cdot \sum_{j=0}^{N_{view}-1} C(S_j, S_k)} \quad j,k \in \{0,1,...,N_{view}-1\} \tag{3}$$

$$w_0 = max\{ w_0, w_1, w_2 \}, \ w_1 = \min\{ w_0, w_1, w_2 \}, \ w_2 = mid\{ w_0, w_1, w_2 \} \tag{4}$$

Frames with smaller time-domain activities require smaller bits; frames with larger time-domain activities require more bits. Using the time-domain activity of the previous frame, the MAD of the previous frame, and the time-domain activity of the current frame to predict the complexity of the current frame. The target bit of the current frame is shown in formula (5):

$$T_r(j) = T'_r(j-1) \cdot \frac{FD(j)}{\frac{1}{j-1} \sum_{k=1}^{k=j-1} \Theta \cdot FD(k)} \tag{5}$$

In the formula, $FD(j)$and $FD(j-1)$represent the time-domain activity of the j-th and j-1 frames, respectively.

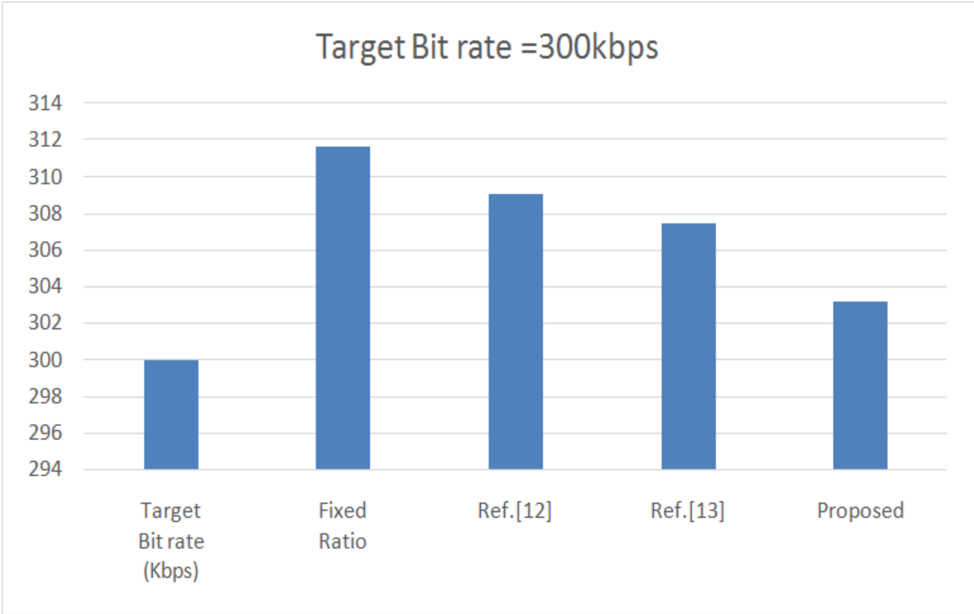
A new GoPencoding will begin after the scene switch frame. This paper refers to previous research results. GoP target bit TGoP is set in the method as follow:

$$T_{GoP} = \chi \cdot N_{GoP} \times \left(R_{picAvg} + \frac{R_{picAvg} \times N_{coded} - R_{coded}}{SW} \right) \tag{6}$$

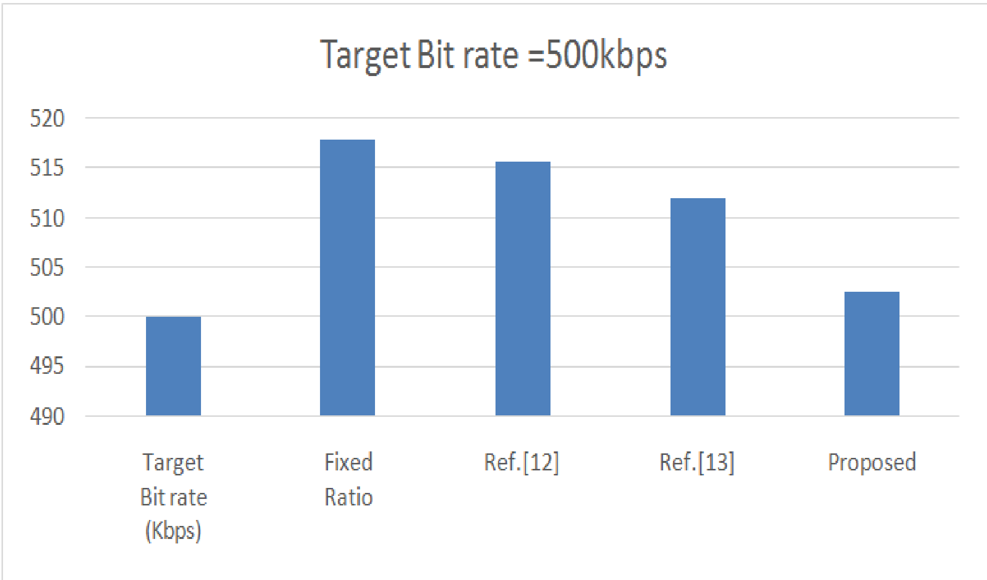
In the formula: NGop represents the size of GoP; RPicAvgis the target bit rate; Ncoded is the number of encoded frames; SW is the size of the smoothing window; Rcoded is the number of bits consumed by the encoded frame. The remaining number of bits of the GoPterminated early and the number of bits consumed by the coding scene switching frame need to be calculated.

3. EXPERIMENTAL RESULTS

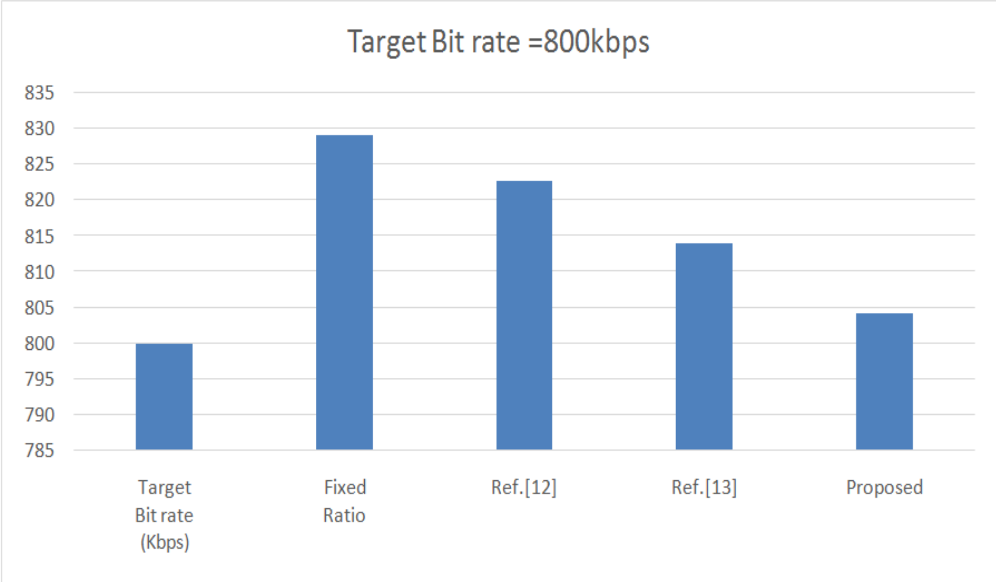
In order to verify the performance advantage of the algorithm in terms of coding efficiency, this paper uses the test sequences Rena, Ballroom, Exit, Flamenco2, and Vassar provided by MERL, KDDI and Nagoya University / Tanimoto Lab for experimental analysis. Newspaper-balloons sequence is a combination of Newspaper sequence and balloons sequence. The other sequences are also synthesized by two sequences. The Newspaper-balloons' sequence is obtained by resampling the Newspaper-balloons sequence. The other sequences are similar.


Target Bit rate =300kbps
314
312
310
308
306
304
302
300
298
296
294
Target Bit rate (Kbps)
Fixed Ratio
Ref.[12]
Ref.[13]
Proposed

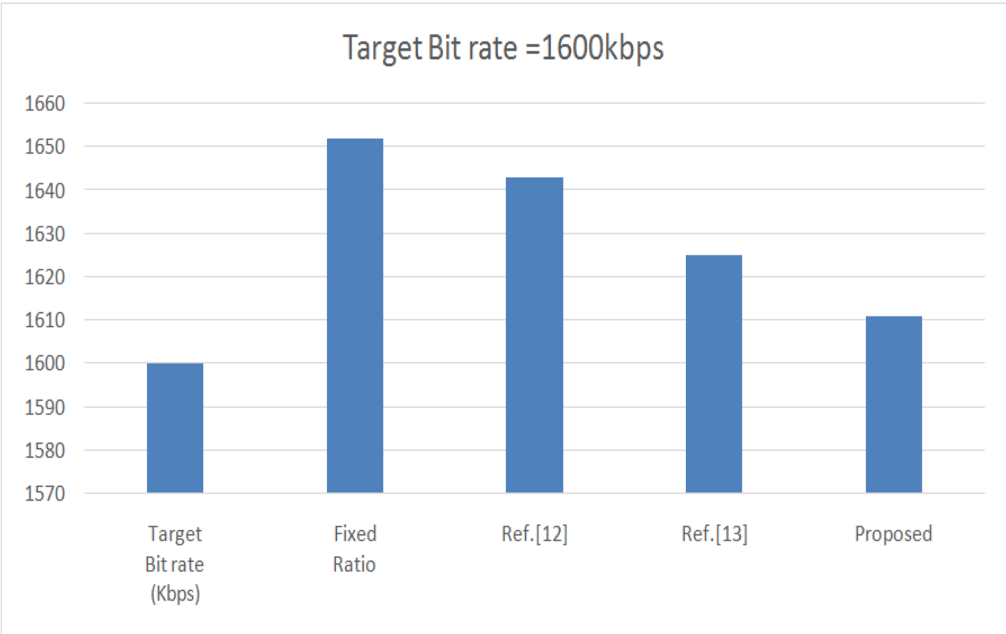
(a)


Target Bit rate =500kbps
520
515
510
505
500
495
490
Target Bit rate (Kbps)
Fixed Ratio
Ref.[12]
Ref.[13]
Proposed

(b)



(c)



(d)

Fig. 1 The experimental results of the sequence Rena-Ballroom.

Table. 1 Simulation results

Sequence		Target Bit rate (Kbps)	Actual generated bits (kbps)				Rate control error (%)			
			Fixed Ratio	Ref.[12]	Ref.[13]	Proposed	Fixed Ratio	Ref.[12]	Ref.[13]	Proposed
VGA	Rena-Ballroom'	300	311.67	309.09	307.5	303.18	3.89	3.03	2.5	1.06
		500	517.95	515.6	511.95	502.55	3.59	3.12	2.39	0.51
		800	829.2	822.72	814	804.16	3.65	2.84	1.75	0.52
		1600	1652	1642.88	1624.96	1611.04	3.25	2.68	1.56	0.69
	Exit-Flamenco2'	300	309.09	308.79	309.3	306.33	3.03	2.93	3.1	2.11
		500	514.3	514.1	509	501.4	2.86	2.82	1.8	0.28
		800	826.32	825.52	819.84	810	3.29	3.19	2.48	1.25
		1600	1659.84	1643.36	1633.44	1614.56	3.74	2.71	2.09	0.91
	Ballroom-Exit'	300	309.15	306.57	303.36	302.04	3.05	2.19	1.12	0.68
		500	516.4	514.25	513	508	3.28	2.85	2.6	1.6
		800	823.12	818.96	814.88	804.56	2.89	2.37	1.86	0.57
		1600	1651.84	1647.04	1635.2	1609.12	3.24	2.94	2.2	0.57
	Vassar-Flamenco2'	300	310.71	309.99	307.71	304.38	3.57	3.33	2.57	1.46
		500	518.15	513.85	511.7	507.35	3.63	2.77	2.34	1.47
		800	820.48	819.04	815.76	809.92	2.56	2.38	1.97	1.24
		1600	1651.2	1642.24	1634.72	1632.64	3.2	2.64	2.17	2.04
HD	Newspaper-balloons'	300	311.49	308.97	306.63	307.23	3.83	2.99	2.21	2.41
		500	517.05	514.85	512	509.2	3.41	2.97	2.4	1.84
		800	826.16	825.04	817.84	811.04	3.27	3.13	2.23	1.38
		1600	1650.08	1650.08	1638.24	1617.12	3.13	3.13	2.39	1.07
Average							3.32	2.85	2.19	1.18

As can be seen from Fig. 1, experimental results with target bits of 300, 500, 800, and 1600. As can be seen from Table 1, rate control for multi-view video coding based on Ref. [12], which is previous research results. Although it can maintain high coding efficiency, the code rate control deviation is relatively large and the average code rate the control error is 3.32%. The method in this paper controls the output bit rate of multi-view video coding more accurately. In most cases, the error between the actual bit rate and the target bit rate can be controlled at about 1.18% or lower. Among them, Flamenco2 and Ballroom sequence code rate control deviation is relatively large, the main reason is that the scene is relatively fierce, the use of coded information for inter-view bit allocation, easily lead to inaccurate bit allocation between views. As can be seen from Table 1 and Figure 2, compared with JMVC, the decoded image PSNR can be improved by 0.05-0.12dB; and compared with Ref. [12], the decoded image PSNR can be improved by 0.07-0.16dB.

4. CONCLUSION

This paper proposes a bit rate control algorithm for multi-view video coding based on a binomial model. Based on the encoded information, a reasonable bit rate allocation is made between viewpoints, and the frame layer bit rate is allocated using frame complexity and time domain activity. Establish a framework for continuous encoding of multiple viewpoints to achieve continuous encoding of multiple viewpoints: The multi-view video encoding reference model JMVM (Joint Multiview Video Model) provided by JVT (Joint Video Team) implements encoding from viewpoint to viewpoint, that is After the encoding of the current viewpoint is completed, the parameter configuration needs to be performed again when encoding the next viewpoint. Obviously, the reference model cannot perform simultaneous encoding of multiple viewpoints. Experimental results show that the actual bit rate can track the target bit rate while maintaining the coding efficiency.

ACKNOWLEDGMENTS

This work was supported by Natural Science Foundation of China (Grants No. 61741111, 61801006); This work was supported by Program for New Century Excellent Talents in Fujian Province University; This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIP) (No.2016R1A2B4013002); in part by Natural Science Foundation of Fujian (Grants No. 2019J01816); in part by Natural Science Foundation of Jiangxi (Grants No. 20181BAB202011); Putian University's Initiation Fee Project for Importing Talents for Scientific Research (Grants No. 2019003); This work was supported in part by the Key Project on Anhui Provincial Natural Science Study by Colleges and Universities under Grant KJ2018A0361.

REFERENCES

[1] E. M. El-Bakary, W. El-Shafai, S. El-Rabaie, O. Zahran, M. El-Halawany, F. E. Abd El-Samie, Proposed enhanced hybrid framework for efficient 3D-MVC and 3D-HEVC wireless communication[J], Multimedia Tools and Applications, Vol. 78, No. 11, pp. 14173–14193, June 2019

[2] Jianjun Lei, Jinhui Duan, Feng Wu, Nam Ling, and Chunping Hou. Fast mode decision based on grayscale similarity and Inter-view correlation for depth map coding in 3D-HEVC[J], IEEE Transactions on Circuits and Systems for Video Technology, Vol. 28, No. 3, 706-718, March 2018

[3] Jingyao Xu, Mai Xu, Yanan Wei, Zulin Wang, and Zhenyu Guan, Fast H.264 to HEVC Transcoding: A Deep Learning Method [J], IEEE Transactions on Multimedia, Vol. 21, No. 7, 1633-1645, July 2019

[4] Pan F, Li Z G, Lim K. A study of MPEG-4 rate control scheme and its implementations. IEEE Transactions on Circuits and System for Video Technology[J], 2003, 13(5):440-446.

[5] Corbera, J., R. and Lei, S., M., “A frame-layer bit allocation for H.263+,” IEEE Trans. Circuits Syst. Video Technol, 10(7), 1154-1158 (2000).

[6] R. Atta and M. Ghanbari, “Low-Complexity joint temporal-quality scalability rate control for H.264/SVC,” IEEE Transactions on Circuits and Systems for Video Technology, Vol.28, No. 9, pp. 2331-2344, September 2018

[7] Joint Video Team of ITU-T VCEG and ISO/IEC MPEG. WD 1 Reference software for MVC (JMVC) 1.0, Doc. JVT-AA212, CH, Geneva, April, 2012.

[8] Naito S, Matsumoto S. 34/35 Mbps 3D-HDTV digital coding scheme using a modified motion compensation with disparity vectors, VCIP, SPIE, USA, 1999, pp: 1082-1089.

[9] Woo W, Ortega A. Optimal blockwise dependent quantization for stereo image coding. IEEE Transactions on Circuits and Systems for Video Technology[J], 1999, 9(6):861-867.

[10] Z Zhao, L Shen, Q Hu, Z Zhang. "Algorithm optimization based on the complexity of HEVC rate control," Chinese Journal of Optoelectronics Laser, Vol.25, No.9, pp. 1715-1720, September 2014

[11] S. Park, D. Sim. An efficient rate-control algorithm for multi-view video coding. The 13th IEEE International Symposium on consumer Electronics[C], Kyoto, Japan, 2009: 115-118.

[12] T Yan, P An, L Shen, Q Zhang, and Z Zhang, "Rate control algorithm for multi-view video coding based on correlation Analysis", the International Symposium on Photonics and Optoelectronics (SOPO), pp. 1-4, Wuhan, China, August 2009.

[13] T Yan, P An, L Shen, Z Li, H Wang, and Z Zhang, "Frame-layer bit allocation for multi-view video coding based on frame complexity estimation," Journal of Shanghai university, vol.14, no. 1, pp. 50-54, February 2010.

[14] J. Ribas-Corbera, S. M. Lei, A frame-layer bit allocation for H.263+, IEEE Trans. Circuits Syst. Video Technol[J], vol.10, no.7, 2000: 1154-1158.