

# RATE CONTROL BASED ON SIMILARITY ANALYSIS IN MULTI-VIEW VIDEO CODING

Tao Yan<sup>1</sup>, In-Ho Ra<sup>2</sup>, Deyang Liu<sup>3</sup>, Deli Chen<sup>1</sup>,  
Yu Youhao<sup>1</sup>, and Shaojie Hou<sup>1</sup>

<sup>1</sup>School of Information Engineering, Putian University, Putian 351100, China

<sup>2</sup>School of Computer, Information and Communication Engineering, Kunsan  
National University, Gunsan 54150, South Korea

<sup>3</sup>School of Computer and Information, Anqing Normal University,  
Anqing 246000, China

## **ABSTRACT**

*As with previous single-view video coding, in order to promote the application of multi-view video coding, multi-view video coding also needs to study the appropriate rate control algorithm. Bit rate control has always been a difficult research topic in multi-view video coding. In multi-view video coding, there are 6 types of coded image, so the previous rate control algorithm for single-view video cannot be directly applied to multi-view video coding. The bit rate control is divided into a four-layer structure for bit rate control of multi-view video coding. Among them, the frame-layer code rate control considers the layer B frame and other factors to allocate the code rate, and the basic unit-layer code rate control uses different quantization parameters according to the content complexity of the macroblock. The average error between the actual bit rate and the target bit rate of the bit rate control algorithm can be controlled by 2.39%.*

## **KEYWORDS**

*Multi-view video coding, Rate control, Bit allocation, Basic unit layer*

## **1. INTRODUCTION**

The main work of 3DAV in the early days was the analysis of 3DAV requirements, including the definition of 3DAV application scenarios, including panoramic video, multi-view video and interactive 3D video. Among them, the multi-view video is taken by multiple cameras from different angles. It contains multiple viewpoint images in the same scene, and the viewpoints are highly correlated. Therefore, in addition to usually using spatial-temporal redundancy to achieve coding purposes as in 2D video, it is more important to eliminate spatial prediction redundancy between different viewpoints. Multi-view video coding often requires more complex and efficient coding technology and higher compression efficiency than single-view video coding such as H.264 / AVC, MPEG-2. 3DTV / FTV ( Three-dimensional television / Free-view television) will become the next-generation television technology after high-definition television (HDTV) [1-2]. But multi-view video coding (MVC) is still immature, and it is still a long way from the

application. The bit allocation and bit rate control of MVC, in particular, are key issues for transmission and applications.

In practical applications, bit rate control is one of the main key technologies. It is one of the most important technologies in video coding. Any video compression standard that leaves the bit rate control will be limited in its application. The previous video compression standards such as MPEG-2, MPEG-4, H.263, H.264, MVC, etc. have given a rate control model [3-6]. Compared with the previous video compression standards, the rate control of MVC The bit rate control is mainly to allocate a reasonable bit rate between the viewpoints to ensure the balance of the video quality between the viewpoints.

As with previous single-view video coding, in order to promote the application of multi-view video coding, multi-view video coding also needs to study the appropriate rate control algorithm. Bit rate control has always been a difficult research topic in multi-view video coding. In multi-view video coding, there are 6 types of coded image, so the previous rate control algorithm for single-view video cannot be directly applied to multi-view video coding. There have been many research results on bit rate control[7-9]. At present, the multi-view video coding rate control algorithm not only needs to reasonably allocate the bit rate in each frame in time to prevent buffer overflow, but also needs to allocate a reasonable bit rate between the viewpoints to ensure the balance of video quality between viewpoints. Therefore, an efficient multi-view video encoding bit rate control algorithm needs to be developed.

To this end, based on JMVC, a multi-view video rate control module is added in this article. The core of the algorithm is based on the analysis of the existing video bit rate control algorithms, according to the characteristics of multi-view video coding and the requirements of its bit rate control, to improve the traditional quadratic rate distortion model, and propose a multi-view video-oriented basic unit Layer rate control algorithm. Experimental simulation results show that compared with JMVC [10] which currently uses fixed quantization parameters, the algorithm in this paper can effectively control the bit rate of multi-view video coding while maintaining efficient coding efficiency.

## **2. MULTI-VIEW VIDEO CODING RATE CONTROL ALGORITHM**

The rate control of single-view video coding is, in essence, the rate control of I and P frames, and the rate control of B frames is not performed. The QP value of B frames is simply based on the adjacent I frames and P The QP value of the frame or two adjacent P frames is determined. When the number of B frames between adjacent I frames and P frames or two adjacent P frames is large, the accuracy of the rate control will change. Very poor. For compatibility with H.264, the bit allocation and bit rate control proposed in this paper is based on the bit rate control algorithm of JVT-G012. The bit rate model adopts our previous research results [11]. The MVC code rate control algorithm is divided into four layers. Among them, GOP code rate control is mainly based on the correlation between viewpoints to allocate reasonable code rates among various views. Frame layer code rate control mainly considers the impact of hierarchical B frames Allocation code rate, basic unit layer code rate control uses different quantization parameters according to the content complexity of the macroblock. The main algorithm is as follows:

The I frame and the first  $T(j)$  frame of the first GOP in each GGOP, and the first P frame of other GOPs are coded with  $QP_0$ . There is no need to allocate target bits, and other B frames (or P frames) target allocate bits B is weighted by  $T_r$  and  $T_{buf}$ ,

$$T(j) = \beta \cdot T_r(j) + (1 - \beta) \cdot T_{buf}(j) \quad (1)$$

Where  $\beta$  is a fixed value, its typical value is 0.5 when there are no B frames, and 0.9 when there are B frames; where the allocation of  $T_{buf}$  for the current frame depends on the current target buffer overflow degree  $Tbl$ , similar to the previous algorithm. The methods are:

$$T_{buf}(j) = \gamma \cdot (Tbl(j) - B_c(j-1)) + \frac{u(j)}{F_r} \cdot \left\{ \frac{\sum_{l=1}^L W(l) \cdot 2^n}{\sum_{l=1}^L \Theta \cdot W(l) + \sum_{l=1}^L W_B(l) \cdot (2^n - 1)} - 1 \right\} \quad (2)$$

Where  $\gamma$  is a fixed value, its typical value is 0.75 when there is no B frame, otherwise it is 0.25. Represents the time layer where the current frame is located,  $w(l)$  represents the weight of the complexity of each frame, and  $W_B(l)$  represents the weight of the B frame. For the value method, see Ref. [13].

The allocation of  $T_r$  in the current frame in formula (6) depends on the number of bits remaining in the current GOP and at a certain time layer in FIG. 1.

$$T_r(j) = T_{GOP}(n_{k,j}) \cdot W_b(n_{k,j}) \quad (3)$$

In the formula,  $T_{GOP}(n_{k,j})$  is the remaining bits of the current GOP, and  $W_b(n_{k,j})$  is the weight of the current frame. The initial value of  $W_b(n_{1,j})(j=1,2,\dots,N(i))$  is set to 1. After each GOP is encoded, it needs to be refreshed at the post-encoding stage. Let  $A_{GOP}(n_{k-1,0})$  be the number of bits actually used to encode the  $k-1$ th GOP of the  $i$ th GGOP.  $A(j)$  represents the number of bits actually used in the  $j$  frame encoding, and the weight  $W_b(n_{k-1,j})(j=1,2,\dots,N(i))$  of the current frame is given by equation (8)

$$W_b(n_{k-1,j}) = \frac{A(j)}{A_{GOP}(n_{k-1,0})} \quad (4)$$

$W_b(n_{k,j})$ 's linear prediction model is:

$$W_b(n_{k,j}) = \beta_{k1} \cdot W_b(n_{k-1,j}) + \beta_{k2} \quad (5)$$

Among them,  $\beta_{k1}$  and  $\beta_{k2}$  are univariate regression coefficients, and the initial values are set to 1 and 0. After each GOP encoding is completed, it needs to be refreshed in the post-encoding stage.  $T_{GGOP}(sn_{i,j})$  represents the number of remaining bits of the GGOP after the  $i$  frame of the  $j$  GGOP is encoded. This paper uses existing algorithms. The main ideas are as follows: Using H.264 / AVC Visa Baseline Tool to perform similarity analysis on video images to obtain the consistency of macroblocks is only possible using statistical features, and DCT coefficients are more suitable for image similarity than pixel brightness Sexual requirements. The macro block contains a part of the nose and lips. We refer to the results of previous research. It can be seen that if most of the subblocks obtain the same SAC value (dark part), the two blocks can match. Therefore, SAC is suitable for statistical matching of images, that is, the results are consistent with the error

concealment strategy of this paper. According to the statistical matching conditions of the sub-blocks, the similarity measure  $\Delta f_k$  for a certain block is defined as:

$$\Delta f_k = \sum_k \theta \cdot (SAC_k^{I_{i-1}} - SAC_k^{I_i}), I_i \in Z \quad (6)$$

Among them,  $\delta(\cdot)$  is the  $\delta$  function and  $I_i$  is the current frame. The similarity  $\Delta f_k$  represents the number of sub-blocks with the same statistical characteristics, and has the following conclusions. If  $\Delta f_i < T_{AC}(= \delta \times T_{\max})$ , then  $I_i$  and  $I_{i-1}$  are not similar; otherwise,  $I_i$  and  $I_{i-1}$  are similar. Therefore, it can be determined that the current frame  $I_i$  is a scene switching frame; if  $I_i$  and  $I_{i-1}$  are not similar, there is a scene switching frame between  $I_i$  and  $I_{i-1}$ ,  $I_i$  is not a scene switching frame, and if there is no switching frame between them,  $I_i$  is a scene switching frame for the scene.  $\sigma$  is a certain similarity measure.

### 3. EXPERIMENTAL TESTS

In order to verify the effectiveness of the proposed multi-view video coding bit allocation and rate control algorithm, we used the test sequences which include Kendo-Dancer, Poznanstreet-Gtfly, Gtfly-Kendo, Poznanstreet-Dancer, and Dancer-Race1. The test platform is Intel (R) Core (TM) Duo CPU 2.66GHz (dual), 1.96GB memory. The Gtfly-Kendo' sequence is obtained by resampling the Gtfly-Kendo sequence. The other sequences are similar. Table 1 shows the experimental results of the proposed algorithm compared with the two fixed bit allocation algorithms.

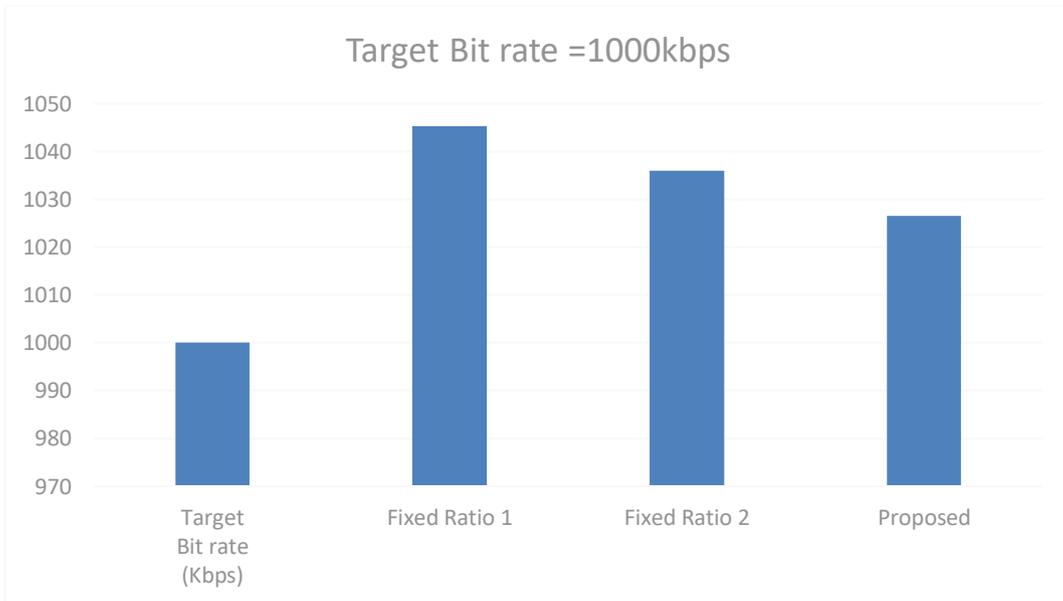


Table 1 Simulation Results

Sequence		Target Bit rate (Kbps)	Actual generated bits ( kbps)			Rate control error (%)		
			Fixed Ratio 1	Fixed Ratio 2	Proposed	Fixed Ratio 1	Fixed Ratio 2	Proposed
HD	Gtfly-Kendo'	1000	1045.30	1036.00	1026.60	4.53	3.60	2.66
		2000	2096.40	2075.80	2028.20	4.82	3.79	1.41
		3000	3148.20	3076.50	3063.60	4.94	2.55	2.12
	Kendo-Dancer'	1000	1046.30	1044.00	1030.10	4.63	4.40	3.01
		2000	2082.40	2070.00	2033.60	4.12	3.50	1.68
		3000	3143.70	3107.40	3073.50	4.79	3.58	2.45
	Poznanstreet- Gtfly'	1000	1036.90	1027.20	1015.80	3.69	2.72	1.58
		2000	2091.00	2068.00	2054.00	4.55	3.40	2.70
		3000	3119.10	3097.80	3062.10	3.97	3.26	2.07
	Poznanstreet- Dancer'	1000	1049.30	1035.70	1031.60	4.93	3.57	3.16
		2000	2091.40	2068.80	2045.40	4.57	3.44	2.27
		3000	3125.40	3098.10	3082.20	4.18	3.27	2.74

Table 1 shows the experimental results of the proposed algorithm compared with the two fixed bit allocation algorithms. The JMVC algorithm without bit rate control uses a fixed quantization value for multi-view video encoding. The bit rate obtained when the fixed quantization value is used is the target bit rate of the bit rate control algorithm.

Table 1 shows the simulation results of code rate control for multi-view video coding. As can be seen from Table 1, compared with the two fixed bit allocation algorithms, rate control algorithm in this paper has a smaller code rate deviation, and the code rate error can be controlled at 2.39%. The main reason for this is that we not only perform reasonable bit allocation between viewpoints, but the basic unit layer code rate control uses different quantization parameters for code rate control according to the content complexity of the macroblock.

#### 4. CONCLUSION

In the future, we consider the use of the relationship between the distortion of the video and the bit rate and the complexity of the video frame to optimize the allocation of the bit rate between the frames. In view of the large number of scene switches that may occur in the video frame, a scene switch detection based on non-connected points. The algorithm, combined with improved adaptive GOP packet technology, further improves the rate control performance. In videos where scenes change, the adaptive GOP grouping technology is used to eliminate the effects of scene switching and improve coding performance.

#### ACKNOWLEDGEMENTS

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] P Ren, X Zhang, H Bi, H Sun, and N Zheng, "Toward an Efficient Multiview Display Processing Architecture for 3DTV," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol.64, no. 6, pp. 705 - 709, June 2017
- [2] C Yang, P An, L Shen, and D Liu, "Adaptive Bit Allocation for 3D Video Coding," *Circuits, Systems, and Signal Processing*, vol. 36, no. 5, pp. 2102-2124, May 2017
- [3] Jinag M Q, Yi X Q, Ling N. Improved frame-layer rate control for H.264 using MAD ratio. *Proceedings of IEEE International Symposium on Circuits and Systems[C]*, Vancouver, Canada, 2004. 813-816.
- [4] 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
- [5] Li Z G, Pan F, Lim K P, Feng G N. Adaptive basic unit layer rate control for JVT. *JVT-G012,7Th meeting[C]*, Pattaya, Thailand, 2003 ,7-14.
- [6] L. Q. Shen, K. Li, G. R. Feng, P. An , and Z.Liu, "Efficient intra mode selection for depth-map coding utilizing spatiotemporal, inter-component and inter-view correlations in 3D-HEVC," *IEEE Transactions on Image Processing*, Vol. 27, No. 9, pp. 4195-4206, September 2018
- [7] 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.
- [8] Y. Z Zhang and C. Lu, "High-performance algorithm adaptations and hardware architecture for HEVC intra encoders," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.29, No. 7, pp. 2138-2145, July 2019.

- [9] 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, 2012: 115-118.
- [10] 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
- [11] T Yan, P An, L Shen, Z Li, H Wang, and Z Zhang, "Rate control algorithm based on frame complexity estimation for MVC", *SPIE International Conference on Visual Communications and Image Processing*, pp. 77-85, Huang Shan, An Hui, China, July 2010
- [12] J. Yang, X. Z. Fang and H. K. Xiong. A joint rate control scheme for H.264 encoding of multiple video sequences. *IEEE Transaction on Consumer Electronics*, Vol. 51, No. 2, May, 2005.