A Bit Allocation Method Based on Proportional-Integral-Derivative Algorithm for 3DTV

Tao Yan[1], In-Ho Ra[2*], Deyang Liu[3], and Qian Zhang[4]
[1] School of Information Engineering, Putian University, Fujian
Putian, China, 351100
[e-mail: yantao@ptu.edu.cn]
[2] School of Computer, Information and Communication Engineering, Kunsan National University
Gunsan, 54150, South Korea
[e-mail: ihra@kunsan.ac.kr]
[3] School of Computer and Information, Anqing Normal University
Anqing, 246000, China
[e-mail: liudeyang@163.com]
[4] School of Information and Electromechanical Engineering, Shanghai Normal University,
Shanghai, 200234, China
[e-mail: qianzhang@shnu.edu.cn]
*Corresponding author: In-Ho Ra

Abstract

Three-dimensional (3D) video scenes are complex and difficult to control, especially when scene switching occurs. In this paper, we propose two algorithms based on an incremental proportional-integral-derivative (PID) algorithm and a similarity analysis between views to improve the method of bit allocation for multi-view high efficiency video coding (MV-HEVC). Firstly, an incremental PID algorithm is introduced to control the buffer "liquid level" to reduce the negative impact on the target bit allocation of the view layer and frame layer owing to the fluctuation of the buffer "liquid level". Then, using the image similarity between views is used to establish, a bit allocation calculation model for the multi-view video main viewpoint and non-main viewpoint is established. Then, a bit allocation calculation method based on hierarchical B frames is proposed. Experimental simulation results verify that the algorithm ensures a smooth transition of image quality while increasing the coding efficiency, and the PSNR increases by 0.03 to 0.82dB while not significantly increasing the calculation complexity.

Keywords: Multiview High Efficiency Video Coding, Bit Allocation, PID, Image Similarity, Frame Complexity

1. Introduction

In a 3DTV/FTV system, the type of method used for the multiview acquisition part and the 3D stereo display part is inconsequential. Various key technologies of three-dimensional (3D) video systems are receiving extensive attention and research is being conducted on them owing to the continuous development of computer vision theory, video capture technology, and stereo display technology, especially the rapid development of naked-eye display technology in recent years. The development of 3DTV has introduced a new generation of video technology. A strong sense of perspective and immersion can be experienced through 3D video. The field of image processing has always been a hotspot [1-5]. Compared with traditional two-dimensional (2D) video, 3D video contains 2D information from multiple angles for the reconstruction of the depth information of the third dimension to produce a 3D feeling [6-9]. When scene switching occurs in the video, the time-domain redundancy of the video at the switching point disappears. At this time, if the adjacent video frames at the switching point are still used as the same coding unit, a larger error occurs in the rate control of the 3D video coding, thereby considerably reducing the video compression performance. There is a significant amount of scene switching in the actual video; hence, it is necessary to detect scene switching. This paper proposes two algorithms based on the incremental proportional-integral-derivative (PID) algorithm and similarity estimation between viewpoints to improve the bit allocation and rate control algorithm for multiview high-efficiency video coding (MV-HEVC).

The remainder of this paper is structured as follows. Section 2 presents our review of previous work on rate control. Section 3 introduces the PID rate control model and provides details of bit allocation algorithms for MV-HEVC based on PID. Section 4 describes the extensive experiments conducted to evaluate the performance of the proposed method. Finally, conclusions are drawn in Section 5.

2. Related Work

The bit allocation method is an important part of the video encoding process. It controls the encoding parameters to ensure that the generated bitstream meets the application requirements and maintains the subjective quality between the frame images in the sequence without significant fluctuations while maintaining high encoding quality. The bit rate fluctuation of the bitstream generated after encoding is very large. A buffer must be added at the output of the encoder to smoothen the bit rate fluctuation, but the buffer still faces the possibility of overflow and underflow. The bit rate control algorithm must verify that the buffer overflow level is used to adjust the encoder output bit stream. Bit rate control also takes into account the high quality of the encoded video while maintaining a constant (smooth change) quality. The average PSNR of the video sequence must be improved. Moreover, the subjective quality between the front and back images in the sequence must not fluctuate significantly. The fluctuation of the decoded image with time has a significant impact on the subjective quality experienced by the audience. The latest bit rate control needs to consider obtaining a higher objective quality and also the human subjective visual effects.

At present, most domestic and foreign research on rate control is based on H.264, HEVC, MVC, and 3D-HEVC [10-14]. The research on MV-HEVC is not as good as that on the other rate control technologies. However, the prediction characteristics according to the parallax between each viewpoint are allocated by bits. In the rate control algorithm of stereoscopic

video coding, Vizzotto BB and others only considered the case of two left and right viewpoints [15], but MV-HEVC generally has more than five viewpoints, and there are many types of image coding. The accuracy of the target bit allocation method based on TM5 will continue to deteriorate, resulting in very high bit allocation errors and difficulty in controlling the bit rate. Seanae Park et al. considered the effect of using layered B frames in the temporal coding layer in multiview video coding [16].

Since the introduction of HEVC, many scholars at home and abroad have been devoted to researching HEVC based multiview video coding rate control. Shao et al. used a fixed allocation ratio to assign a coding rate to the viewpoint and depth map [17]. This algorithm was simple and effective, but it did not allow for different sequences to obtain the best coding efficiency. Xiao et al. proposed the use of scalable bit rate allocation algorithms for different bandwidths to meet the actual application scenarios [18]. Pan Geng et al. [19] proposed a depth-based 3D-HEVC code rate control algorithm that uses a fixed color and depth code rate ratio of 4:1, but this algorithm could not achieve the best virtual viewpoint rendering quality. Fang et al. [20] estimated a virtual visual distortion model based on the coding information. The model could accurately allocate bits, but the complexity was higher.

Xiao Jianmin et al. [21] proposed a depth and texture grading rate control algorithm. Reference [22] established an R-D model between viewpoints by combining texture and virtual viewpoint distortion functions; however, this method did not use the correlation between viewpoints, and the rate control error was relatively large, with an average code rate error of 1.56%. Wang et al. [23] proposed a bit allocation rate control algorithm for 3D-HEVC. Because the algorithm directly used the traditional H.264 bit allocation model, the R-D model prediction was not accurate, thereby resulting in an unacceptable rate control accuracy. In 2018, Lei J et al. [24] proposed a bit allocation method based on the R-λ model. The experimental results were ideal, but the correlation among viewpoints was not considered, thereby leading to high computational complexity. Therefore, further research on the bit allocation optimization model for MV-HEVC is needed.

PID control theory has achieved considerable success in the field of mechanical automation. Many scholars have introduced PID controllers in 2D video coding rate control. Sun et al. [25] proposed a bit allocation method for MPEG-4 based on PID as early as 2004. This method enabled precise control of the buffer and achieved good video quality, buffer control, and quality fluctuation. Xu et al. [26] proposed rate control in the linear domain. A PID controller was used at the macroblock level. Simultaneously, they considered the impact of detecting scene changes and adopted a simple method to detect scene changes. Excessive output code flow may block the transmission channel, resulting in delayed video transmission or frame loss. The compression ratio is high, and the output code flow is too small; thus, the transmission channel resources cannot be fully utilized, and video quality degradation will be caused at the decoding end.

The main method to solve this problem is adopting rate control technology, that is, the size of the output code stream at the video coding end is adjusted by reasonably distributing the number of video coding bits so that the video coding can reach the best quality under the conditions of full video transmission. Zhou Yimin et al. [27] adopted an exponential code rate control model and adopted a PID controller for frame rate code allocation, but they used a fixed PID in the process of PID control, as in the aforementioned study. The control parameters affect the control accuracy, and the control has no flexibility.

Considering the limitations of the bit rate allocation of the frame layer and the GOP layer, Shen et al. [28] proposed a frame-level target bit allocation scheme based on incremental PID and frame complexity. First, they introduced incremental PID to control the buffer "liquid

level" to reduce the negative impact on the target bit allocation of the frame layer owing to the fluctuation of the buffer "liquid level." Next, a target bit allocation scheme based on frame complexity was proposed in which the MAD of the previous frame, time-domain activity of the previous frame, and time-domain activity of the current frame were used to predict the current frame complexity of the frame.

The aforementioned algorithms are all 2D video coding rate control algorithms based on the PID control theory, in which PID theory has not yet been introduced into 3D video coding rate control. This paper proposes a 3D video coding rate control algorithm (MV-HEVC) based on an incremental PID algorithm and similarity estimation between viewpoints. First, we introduce incremental PID to control the "liquid surface" of the buffer, and then, we use image correlation to analyze inter-view correlation and frame complexity to perform bit allocation and coding between the viewpoint layer, frame layer, and basic unit layer rate control. The experimental results verify that the proposed algorithm is consistent with the current popular multiview texture and depth. Compared with the algorithm, under the same coding bit rate, the objective quality of the virtual viewpoint and the basic viewpoint improved by an average of 0.45 dB; hence, it has a higher coding efficiency.

3. Proportional-Integral-Derivative Controller for MV-HEVC

The PID controller is one of the most popular feedback controllers. PID control is particularly suitable for complex control environments. We cannot predict the next frame in video coding; hence, this practical environment is particularly suitable for control with a PID controller. The PID controller has the characteristics of simple operation and high efficiency. These advantages are precisely why most engineers favor the PID controller. Video code rate control has achieved very successful results [28-29]. PID control uses the ratio of the error signal (proportional), integral, and derivative to generate control signals so that the system obtains the required output. **Fig. 1** is a schematic diagram of a typical PID controller. The entire system is controlled by a PID controller. And controlled devices.
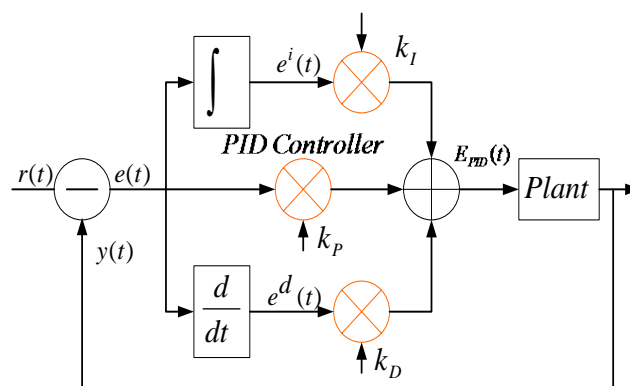

Fig. 1. PID control principle

The goal of PID control is to ensure that the difference between the output $y(t)$ and the demand signal $R(t)$ of the controlled device is as small as possible. In order to perform PID control, an error signal $e(t)$ is required as an input. Generally, the error signal is the demand signal $R(t)$ and the feedback signal The difference between $y(t)$. After the error signal is

obtained, the integral and derivative of the error signal are calculated, and then the three signals are k_p 、 k_i 、 k_d -weighted to form the control signal $E_{PID}(t)$ of the controlled module.

Where E_P, E_I, and E_D are the proportional, integral, and derivative control parameters, respectively, and they can be computed by

$$E_I(t) = k_I \cdot \int_0^t e(t)dt = k_I \cdot \sum_{i=0}^t e(t) \tag{1}$$

$$E_P(t) = k_P \cdot e(t) \tag{2}$$

$$E_I(t) = k_I \cdot \int_0^t e(t)dt = k_I \cdot \sum_{i=0}^t e(t) \tag{3}$$

$$E_D(t) = k_D \cdot \frac{de(t)}{dt} \approx k_D \cdot [e(t) - e(t-1)] \tag{4}$$

In the formula, $e(t)$ 、 $\sum_{i=0}^t e(i)$ and $\frac{de(t)}{dt}$ are respectively expressed as the error signal itself, as well as the integral and derivative of the error signal. When there is only proportional control, there is a steady-state error in the system output. The value of k_p determines the speed of proportional adjustment. When the k_p factor is increased, the adjustment speed of the regulator will be accelerated. If k_p is too large, then the control system will oscillate or overshoot. Increasing k_i can suppress the impact of accidental factors on the system, reduce oscillations or reduce fluctuations. Increasing k_d can increase the sensitivity to speed of error response, but it should be avoided to cause unnecessary fluctuations.

The error function $e(t)$ represents the difference between the demand function and the feedback functions. The output of the proportional controller is proportional to the input error signal, while the output of the integral controller is proportional to the integral of the input error signal. This causes the controller's output to increase and the steady-state error further decreases until it equals zero. The output of the differential controller is directly proportional to the rate of change of the error of the input error signal. The differential term can predict the trend of error changes, thereby avoiding serious overshoot of the controlled variable. This can improve the dynamic characteristics of the control system during the adjustment process.

In contrast to the above PID, the incremental PID control focuses on the impact of error accumulation on the system. The output of the controller is only the increment $\Delta E_{PID}(t)$ of the control amount. The control formula is:

$$\Delta E_{PID}(t) = k_P \cdot \Delta e(t) + k_I \cdot e(t) + k_D \cdot [\Delta e(t) - \Delta e(t-1)] \tag{5}$$

In the formula: k_p 、 k_i 、 k_d is the proportional, integral and differential coefficients respectively; $e(t)$ is the error signal. And $\Delta e(t)$ can be expressed as follows (6):

$$\Delta e(t) = e(t) - e(t-1) \tag{6}$$

Compared with the PID algorithm, the incremental PID algorithm has the following four advantages: (1) Each output signal of the PID algorithm is related to the entire previous PID controller state. The calculation formula uses the accumulated value $\sum_{i=0}^{t} e(t)$ of the past deviation, which is prone to generate large cumulative errors. The incremental method only needs to calculate the increment. When there is a calculation error or insufficient accuracy, the influence on the calculation of the control amount is small. Therefore, the reliability of the incremental PID algorithm is relatively high. (2) The incremental PID algorithm is only related to the past three samplings, and the calculation amount is relatively low. (3) When the control conditions change, because the incremental PID algorithm only outputs the increment, it will not have much impact on the system. (4) Incremental PID control is more conducive to eliminating the effect of error accumulation on the system.

Therefore, as shown in this section, we use incremental PID control. Determination of the control increment $\Delta E_{PID}(t)$ is only related to the last three samplings, and it is easier to obtain a better control effect through weighting processing.

4. Rate Control Algorithm for MV-HEVC

4.1 Bit allocation based on incremental PID for the view layer

According to the previous analysis, it is known that the target bit allocation in the GOP layer rate control of the results obtained in previous research is unreasonable. The problem with MV-HEVC bit rate control is that it is difficult to allocate bits between viewpoints, and it is impossible to assign a reasonable bit to each viewpoint accurately [30]. The weight, w_k, represents the proportion of all viewpoints from viewpoint k. A larger w_k indicates that the viewpoint is the main viewpoint. For different situations, we use similarities between viewpoints to determine a reasonable rate control and bit allocation. The viewpoint with the highest similarity to other viewpoints has a correspondingly higher weight; hence, that viewpoint should be given more bits. As a reference viewpoint for other viewpoints, the benefit of this viewpoint encoding is directly related to the entire video encoding result; therefore, a greater allocation of bits is needed. The total number of bits allocated to the ith viewpoint in the encoded GGOP picture group is given by (7).

$$T_{Gop}(i) = \begin{cases} \dfrac{u(i)}{F_r} \times N_{gop}(i) \times N_{view} * w_0 & i = 0, initialization \\ min(T_{Gop}(i-1) + \Delta T_{Gop}(i), \dfrac{u(i)}{F_r} \times N_{gop}(i) \times N_{view} * w_i + \Delta T_{buffer}) & i \neq 0 \end{cases} \quad (7)$$

where N_{view} represents the number of multi-view video viewpoints, F_r represents the channel bit rate, $u(i)$ represents the frame rate, and $N_{gop}(i)$ indicates the encoded GOP length. $\Delta T_{Gop}(i)$ represents the remaining bits of the last GGOP. ΔT_{buffer} represents the remaining bits of the last buffer.

$$w_k = \frac{\dfrac{1}{N_{view}} \cdot \sum\limits_{j=0, j \neq k}^{N-1} S(V_j, V_k)}{\sum\limits_{k=0}^{N_{view}-1} \dfrac{1}{N_{view}-1} \cdot \sum\limits_{j=0, j \neq k}^{N_{view}-1} S(V_j, V_k)} \tag{8}$$

where $S(V_j, V_k)$ is the similarity between viewpoints V_j and V_k. This paper uses a bilinear similarity measurement algorithm to test the similarity of two viewpoints. This algorithm is superior to the traditional distance measurement to test the similarity of two features, and there are no restrictions such as distance measurement. The algorithm has been successfully used in the field of image retrieval. Among them, S_j^d, S_k^d are the feature vectors of the two images, respectively.

$$S(V_j, V_k) = \frac{S_j^k \cdot S_k^d}{\left| S_j^d \right| \cdot \left| S_k^d \right|} \tag{9}$$

4.2 Bit allocation and rate control based on incremental PID for frame layer

From the above analysis, it can be observed that the target allocation bits of P frames are allocated by T_p and T_{buf} weighting. T_{buf} allocation depends on the current target buffer overflow, frame rate, channel bit rate, and true buffer overflow. When the scene is switched or the scene is moving violently, the predicted QP is not accurate. The actual number of bits generated by the current frame is much higher or lower than the number of allocated bits, which makes the real buffer "liquid level" far away. Moved the target buffer overflow. In order to meet the fullness of the target buffer, this will inevitably affect the allocation of target bits in the next frame or frames, causing fluctuations in PSNR. Therefore, incremental PID is used to control the "liquid level" of the buffer to balance the negative impact of the buffer fluctuation on the target bit allocation of the frame.

The definition of the error function is the key to the incremental PID controller, and the error function is defined on the basis of the demand function and the feedback function. This paper uses (10) as the error function of the buffer incremental PID control:

$$e(j) = Tbl(j) - B_c(j) \tag{10}$$

In the formula: $Tbl(j)$ and $B_c(j)$ are the buffer target overflow degree and the actual buffer occupancy of the first P frame of the current GOP. The error increment is:

$$\Delta e(j) = e(j) - e(j-1) \tag{11}$$

Next, the incremental PID algorithm is used to control the buffer. The output of the controller (increment of T_{buf}) ΔT_{buf} is as shown in (12).

$$\Delta T_{buf, PID} = k_P \cdot \Delta e(t) + k_I \cdot e(t) + k_D \cdot [\Delta e(t) - \Delta e(t-1)] \tag{12}$$

The parameters of the incremental PID controller in the formula are determined using the PID trial and error method; hence, based on experience, a set of controller parameters is temporarily and tentatively determined. Different sequences are tested, and the PSNR and buffer overflow curves are observed. If the control quality is not achieved, then the parameters of the controller are changed according to the influence of each parameter on the control process. Many tests are conducted until this requirement is satisfied.

According to the number of remaining bits, the number of remaining B frames in each layer, the scaling factor of the current GOP, and the incremental PID controller, determine the number of bits allocated to the current frame as

$$T_{l,i}^{B} = \chi_1 \cdot \frac{\omega_l B_l(i)}{\sum\limits_{k=1}^{d-1} \omega_k N_B^i(k)} + \eta_1 \cdot \Delta T_{buf,PID} \tag{13}$$

$$T^{P/I} = \chi_2 \cdot \frac{\omega^{P/I} B_{tot}}{\omega^{P/I} + \sum\limits_{k=1}^{d-1} \omega_k N_B(k)} + \eta_2 \cdot \Delta T_{buf,PID} \tag{14}$$

Among them, $T^{P/I}$ indicates the total number of bits allocated to the P picture or I picture, $T_{l,i}^{B}$ indicates the number of bits allocated to the i th B picture of the l th layer of the current GOP, and $B_l(i)$ indicates the i th B of the l th layer of the current GOP when encoding The number of remaining bits in the picture, $N_B^i(k)$ N indicates that the remaining bits have not been encoded in the k th layer. $\Delta T_{buf,PID}$ indicates that the remaining bits of the buffer are based on the PID controller. B_{tot} P represents the bits allocated by the current GOP, and $\omega^{P/I}$ and ω_k represent the scaling factor and k th score of the P picture or I picture, respectively. This is the scaling factor of the layer B image. $\chi_1, \chi_2, \eta_1, \eta_2$ respectively represent the weight coefficients.

5. Experimental Test

The six synthetic sequences in this study, verify the effectiveness of this algorithm. Among them, 6 sequences Balloons-PoznanHall2, PoznanStreet-Balloons, Kendo-Newspaper, BookArrival-PoznanStreet, BookArrival-PoznanStreet, Newspaper-Balloons, PoznanHall2-BookArrival with scene switching, they are respectively stitched by two different sequences (Balloons-PoznanHall2, PoznanStreet-Balloons, Kendo-Newspaper, BookArrival-PoznanStreet, Newspaper-Balloons, PoznanHall2-BookArrival). Balloons-Kendo-Newspaper is a new sequence synthesized by Balloons, Kendo and Newspaper. The synthesis of the above sequence requires resampling and preprocessing. As a result, each scene has a different complexity. The main coding parameters were as follows: coding structure I-P-B-B, GOP length of 12, each composite sequence encoded 480 frames or 720 frames, the entropy coding used CABAC, RDO was set to enable, the reference frame was 2, the search method was Fast Search, and the search range was 64 pixels. The test platform was an Intel(R) Core(TM) i7-5500U Duo CPU with 2.40 GHz (Double) and 8.00 GB of memory.
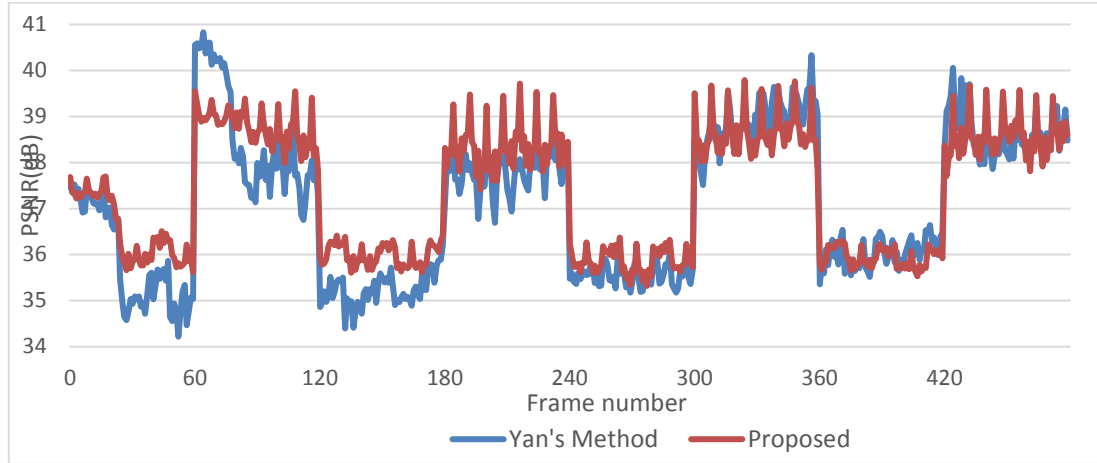
(a) Experimental PSNR results of PoznanStreet-Balloons sequence
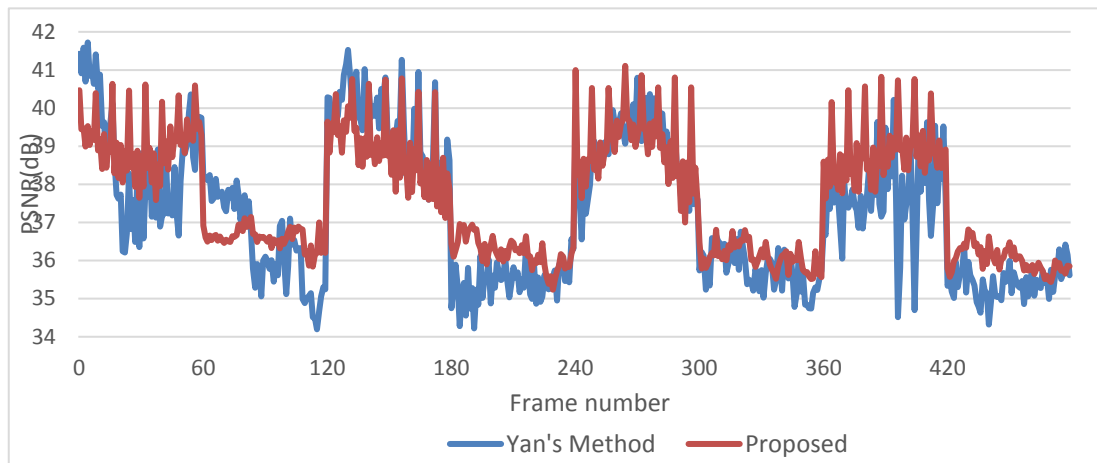
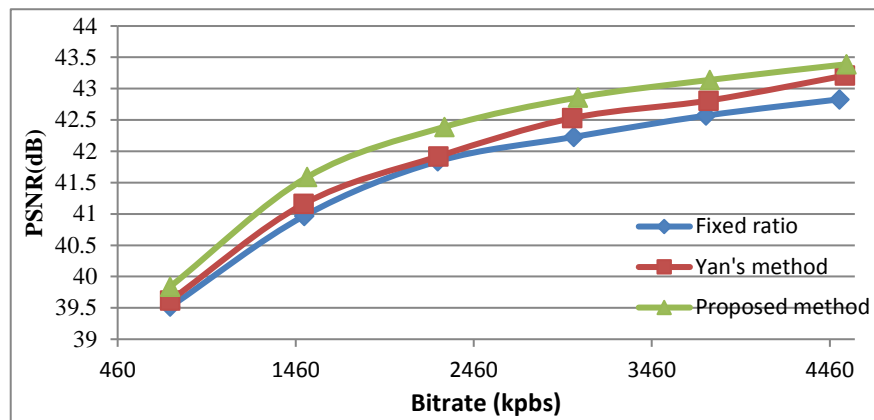(b) Experimental PSNR results of Kendo-Newspaper sequence

Fig. 2. Objective quality fluctuation of sequences

In this study, we compare the proposed algorithm based on frame complexity and incremental PID with our previous algorithm. **Fig. 2** presents a PSNR fluctuation chart of the PoznanStreet-Balloons and Kendo-Newspaper sequences. It can be observed from **Fig. 2** that the PSNR change in the proposed rate control algorithm has small fluctuations, thereby ensuring a smooth transition of image quality, and no obvious image quality changes are visually observed. The complexity of the rate model is relatively high, and the error is relatively large. Scene switching frames consume too many bits, and the "liquid level" of the buffer suddenly increases. For nonswitching gentle viewpoints and frames, fewer bits need to be allocated. The image quality based on the ρ-domain rate control algorithm is low and fluctuates considerably.

Table 1. Simulation results

Sequences	Fixed Ratio		Yan's Method		Proposed Method		Improved PSNR	
	Total Rate	Avg. PSNR	Total Rate	Avg. PSNR	Total Rate	Avg. PSNR	vs. Fixed Ratio	vs. Yan
Balloons-PoznanHall2	755.62	39.52	755.81	39.62	755.22	39.84	0.32	0.22
	1508.54	40.97	1505.32	41.16	1525.88	41.59	0.62	0.43
	2257.72	41.84	2261.45	41.92	2295.33	42.39	0.55	0.47
	3023.09	42.23	3013.70	42.53	3046.42	42.86	0.63	0.33
	3765.12	42.57	3779.47	42.81	3786.16	43.14	0.57	0.33
	4514.60	42.83	4545.77	43.21	4554.22	43.39	0.56	0.18
PoznanStreet-Balloons	751.01	37.18	752.46	37.07	751.54	37.29	0.11	0.22
	1502.14	38.97	1509.68	39.02	1526.98	39.46	0.49	0.44
	2259.36	40.04	2256.94	40.02	2269.46	40.49	0.45	0.47
	3022.54	40.58	3033.49	40.83	3021.70	41.08	0.50	0.25
	3778.71	40.98	3775.25	41.22	3814.67	41.57	0.59	0.35
	4515.40	41.32	4526.95	41.72	4549.45	41.92	0.60	0.20
Kendo-Newspaper	499.32	37.48	501.31	37.26	497.87	37.60	0.12	0.34
	988.31	39.77	994.74	39.81	1010.35	40.25	0.48	0.44
	1493.99	40.99	1484.83	41.02	1517.39	41.18	0.19	0.16
	1991.81	41.67	1988.94	41.92	2015.40	42.29	0.62	0.37
	2492.98	42.17	2475.92	42.48	2520.55	42.85	0.68	0.37
	3490.59	42.78	3035.99	43.08	3018.15	43.29	0.51	0.21
BookArrival-PoznanStreet	748.50	37.03	753.73	36.78	759.48	37.08	0.05	0.30
	1506.24	38.48	1500.56	38.36	1524.49	38.79	0.31	0.43
	2258.54	39.28	2256.19	39.27	2263.30	39.55	0.27	0.28
	3010.15	39.72	3017.19	39.82	3054.50	39.94	0.22	0.12
	3782.71	40.02	3770.20	40.18	3784.52	40.36	0.34	0.18
	4531.49	40.28	4554.30	40.47	4537.80	40.62	0.34	0.15
Newspaper-Balloons	502.45	37.47	504.5	37.58	509.2	37.64	0.17	0.06
	1005.6	39.58	1012.05	39.78	1014.43	40.19	0.61	0.41

	1509.51	40.72	1508.64	40.93	1530.27	41.28	0.56	0.35
	2021.4	41.39	2017.95	41.77	2029.25	42.21	0.82	0.44
	2511.6	41.88	2523.19	42.28	2535.4	42.67	0.79	0.39
	3024.2	42.32	3020.26	42.86	3020.3	43.13	0.81	0.27
	748.40	39.38	754.30	39.32	755.98	39.65	0.27	0.33
	1496.81	40.48	1508.42	40.52	1525.64	40.94	0.46	0.42
PoznanHall2-BookArrival	2258.46	41.07	2270.13	41.16	2294.33	41.45	0.38	0.29
	3002.62	41.38	3035.78	41.52	3044.11	41.73	0.35	0.21
	3775.49	41.62	3796.17	41.78	3783.01	41.92	0.30	0.14
	4538.80	41.79	4541.61	41.98	4541.66	42.03	0.24	0.05
	521.64	37.20	496.80	37.80	497.50	37.83	0.63	0.03
	1121.90	39.91	1001.70	40.50	987.90	40.60	0.69	0.10
Balloons-Kendo-Newspaper	1696.02	40.96	1500.90	41.32	1491.30	41.45	0.49	0.13
	2141.18	42.22	2001.11	42.50	1989.01	42.61	0.39	0.11
	2801.01	42.54	2500.90	43.00	2485.00	43.21	0.67	0.21
	3449.54	42.86	2999.60	43.40	3313.90	43.55	0.69	0.15
AVERAGE	2299.41	40.56	2262.96	40.75	2279.74	41.02	0.47	0.27

(a)
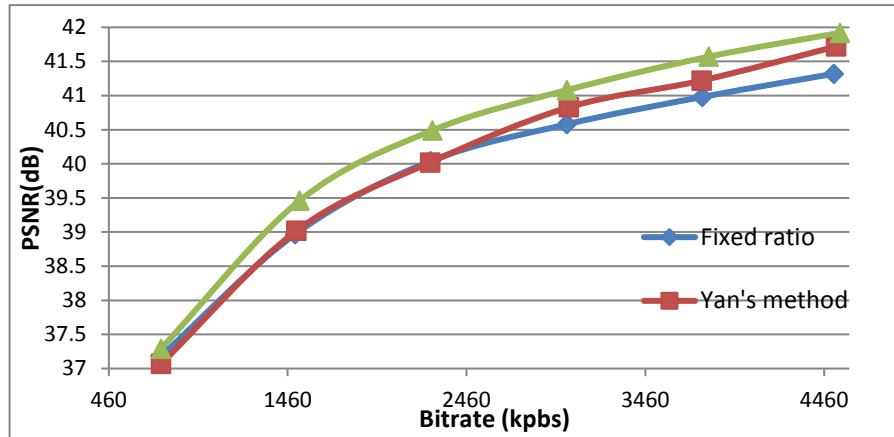
(b)
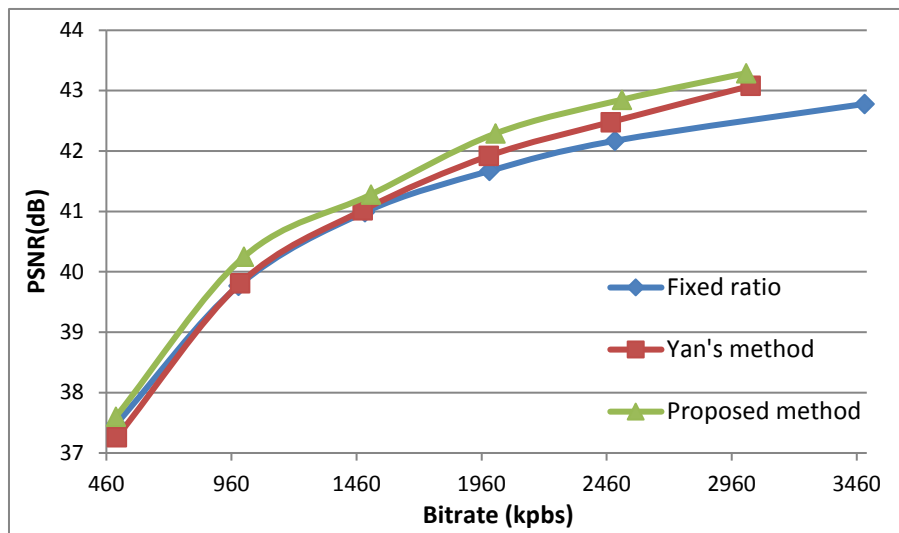
(c)

Fig. 3. Experimental results of sequences: **(a)** ″**Balloons-PoznanHall2**″,
(b) ″**PoznanStreet-Balloons**″, and **(c)** ″**Kendo-Newspaper**″

   Table 1 and **Fig. 3** present the coding efficiency of the three rate control algorithms. We previously proposed a rate control algorithm based on the ρ domain for MV-HEVC, but its coding efficiency was relatively low. Compared with the fixed ratio algorithm, the average PSNR gain was 0.47 dB. Compared with our previous algorithm, the PSNR of the algorithm proposed this paper is improved, and the average gain of PSNR is 0.27 dB. This is mainly because the scene switching frame consumes too many bits, and the "liquid level" of the buffer is suddenly increased. If a fixed bit allocation rate between the viewpoints is used, then the rate control algorithm of the fixed ratio is used to meet the full overflow of the buffer "liquid level" target. The bits allocated in the following frames must be significantly reduced, resulting in a sudden drop in the encoding quality of the following frames.

In summary, the rate control algorithm proposed in this paper is more accurate than the algorithm we previously proposed. The rate deviation is smaller, and the PSNR fluctuation is relatively small, thereby improving the visual effect. The scheme adopts incremental PID control for the "liquid level" of the buffer, which can distribute the impact of excessive bit consumption of scene switching frames to subsequent frames, to avoid large fluctuations in video quality. Experimental results verify that the algorithm can accurately control the output bit rate and significantly improve the overall coding efficiency of a video sequence.

6. Conclusions

Excessive output code flow may block transmission channels, resulting in delayed video transmission or frame loss. The compression ratio is large, and the output code flow is too small. Thus, transmission channel resources cannot be fully utilized, and the video quality is degraded at the decoding end. The main method to solve this problem is adopting rate control technology, in which by reasonably distributing the number of video coding bits, the size of the output code stream at the video coding end is adjusted so that the video coding can achieve the best quality under the conditions of full video transmission.

As the current bit allocation method for MV-HEVC has not been studied in depth, after an extensive review of related research, we found that the existing rate control is insufficient and proposed a rate control algorithm based on the PID control theory. The algorithm performs effective code rate control at the viewpoint layer and the frame layer. Experimental results verified that the introduction of PID theory to 3D video coding bit allocation achieves very good results, but the bit allocation between viewpoints needs to be improved. In the future, the correlation between viewpoints will be studied further to improve the rate control algorithm.

Acknowledgements

References

[1] H. Yuan, S. Kwong, X. Wang, W. Gao, and Y. Zhang, "Rate Distortion Optimized Inter-View Frame Level Bit Allocation Method for MV-HEVC," *IEEE Transactions on Multimedia*, vol. 17, no. 12, pp. 2134-2146, Dec. 2015. Article (CrossRef Link)

[2] Y. Xiao and G. Gao, "Digital Watermark-Based Independent Individual Certification Scheme in WSNs," *IEEE Access,* vol. 7, pp. 145516-145523, 2019. Article (CrossRef Link)

[3] G. Gao, Z. Cui, and C. Zhou, "Blind Reversible Authentication Based on PEE and CS Reconstruction," *IEEE Signal Process Letters*, vol. 25, no. 7, pp. 1099-1103, July 2018. [Article (CrossRef Link)](#)

[4] H. Yang, L. Shen, X. Dong, Q. Ding, P. An, and G. Jiang, "Low-Complexity CTU Partition Structure Decision and Fast Intra Mode Decision for Versatile Video Coding," *IEEE Transactions on Circuits Systems for Video Technology,* vol. 30, no. 6, pp. 1668-1682, June 2020. [Article (CrossRef Link)](#)

[5] G. Gao, S. Tong, Z. Xia, B. Wu, L. Xu, and Z. Zhao, "Reversible data hiding with contrast enhancement and tamper localization for medical images," *Signal Processing*, vol. 178, Jan. 2021. [Article (CrossRef Link)](#)

[6] L. Shen, K. Li, G. 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, Sep. 2018. [Article (CrossRef Link)](#)

[7] J. Xu, M. Xu, Y. Wei, Z. Wang, and Z. Guan, "Fast H.264 to HEVC Transcoding: A Deep Learning Method," *IEEE Transactions on Multimedia*, vol. 21, no. 7, pp. 1633-1645, July 2019. [Article (CrossRef Link)](#)

[8] J. Lei, J. Duan, F. Wu, N. Ling, and C. Hou, "Fast Mode Decision Based on Grayscale Similarity and Inter-View Correlation for Depth Map Coding in 3D-HEVC," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 28, no. 3, pp. 706-718, Mar. 2018. [Article (CrossRef Link)](#)

[9] T. Yan, I. Ra, H. Wen, M. Weng, Q. Zhang, and Y. Che, "CTU Layer Rate Control Algorithm in Scene Change Video for Free-Viewpoint Video," *IEEE Access*, vol. 8, pp. 24549-24560, 2020. [Article (CrossRef Link)](#)

[10] R. Atta and M. Ghanbari, "Low-Complexity Joint Temporal-Quality Scalability Rate Control for H.264/SVC," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 28, no. 9, pp. 2331-2344, Sep. 2018. [Article (CrossRef Link)](#)

[11] Y. Zhang and C. Lu, "Efficient Algorithm Adaptations and Fully Parallel Hardware Architecture of H.265/HEVC Intra Encoder," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 29, no. 11, pp. 3415-3429, Nov. 2019. [Article (CrossRef Link)](#)

[12] Z. Pan, Y. Zhang, and S. Kwong, "Efficient Motion and Disparity Estimation Optimization for Low Complexity Multiview Video Coding," in *IEEE Transactions on Broadcasting*, vol. 61, no. 2, pp. 166-176, June 2015. [Article (CrossRef Link)](#)

[13] G. Sanchez, J. Silveira, L. V. Agostini, and C. Marcon, "Performance Analysis of Depth Intra-Coding in 3D-HEVC," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 29, no. 8, pp. 2509-2520, Aug. 2019. [Article (CrossRef Link)](#)

[14] J. E. Lim, J. Kim, K. N. Ngan, and K. H. Sohn, "Advanced rate control technologies for 3D-HDTV," *IEEE Transactions on Consumer Electronics*, vol. 49, no. 4, pp. 1498-1507, Nov. 2003. [Article (CrossRef Link)](#)

[15] B. B. Vizzotto, B. Zatt, M. Shafique, S. Bampi, and J. Henkel, "A Model Predictive Controller for Frame-Level Rate Control in Multiview Video Coding," in *Proc. of IEEE International Conference on Multimedia and Expo*, pp. 485-490, 2012. [Article (CrossRef Link)](#)

[16] S. Park and D. Sim, "An efficienct rate-control algorithm for multi-view video coding," in *Proc. of IEEE 13th International Symposium on Consumer Electronics*, pp. 115-118, 2009. [Article (CrossRef Link)](#)

[17] F. Shao, G. Jiang, W. Lin, M. Yu, and Q. Dai, "Joint Bit Allocation and Rate Control for Coding Multi-View Video Plus Depth Based 3D Video," *IEEE Transactions on Multimedia*, vol. 15, no. 8, pp. 1843-1854, Dec. 2013. [Article (CrossRef Link)](#)

[18] J. Xiao, M. M. Hannuksela, T. Tillo, M. Gabbouj, C. Zhu, and Y. Zhao, "Scalable Bit Allocation Between Texture and Depth Views for 3-D Video Streaming Over Heterogeneous Networks," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 25, no. 1, pp. 139-152, Jan. 2015. [Article (CrossRef Link)](#)

[19] P. Gao and W. Xiang, "Rate-Distortion Optimized Mode Switching for Error-Resilient Multi-View Video Plus Depth Based 3-D Video Coding," *IEEE Transactions on Multimedia,* vol. 16, no. 7, pp. 1797-1808, Nov. 2014. Article (CrossRef Link)

[20] L. Fang, N. Cheung, D. Tian, A. Vetro, H. Sun, and O. C. Au, "An Analytical Model for Synthesis Distortion Estimation in 3D Video," *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 185-199, Jan. 2014. Article (CrossRef Link)

[21] J. Xiao, M. M. Hannuksela, T. Tillo, M. Gabbouj, C. Zhu, and Y. Zhao, "Scalable Bit Allocation Between Texture and Depth Views for 3-D Video Streaming Over Heterogeneous Networks," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 25, no. 1, pp. 139-152, Jan. 2015. Article (CrossRef Link)

[22] H. Roodaki, Z. Iravani, M. R. Hashemi, and S. Shirmohammadi, "A View-Level Rate Distortion Model for Multi-View/3D Video," *IEEE Transactions on Multimedia*, vol. 18, no. 1, pp. 14-24, Jan. 2016. Article (CrossRef Link)

[23] X. Wang, S. Kwong, H. Yuan, Y. Zhang, and Z. Pan, "View synthesis distortion model based frame level rate control optimization for multiview depth video coding," *Signal Processing*, vol. 112, no. 1, pp. 189-198, Jan. 2015. Article (CrossRef Link)

[24] J. Lei, X. He, H. Yuan, F. Wu, N. Ling, and C. Hou, "Region Adaptive R- λ Model-Based Rate Control for Depth Maps Coding," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 28, no. 6, pp. 1390-1405, June 2018. Article (CrossRef Link)

[25] Y. Sun and I. Ahmad, "A robust and adaptive rate control algorithm for object-based video coding," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 14, no. 10, pp. 1167-1182, Oct. 2004. Article (CrossRef Link)

[26] P. Xu, S. Zou, Y. Li, and W. Xu, "A New Linear Rate Control Algorithm for H.264 Based on PID Controller," in *Proc. of the 2nd IEEE Conference on Industrial Electronics and Applications*, pp. 2058-2061, 2007. Article (CrossRef Link)

[27] Y. Zhou, Y. Sun, Z. Feng, and S. Sun, "PID-Based Bit Allocation Strategy for H.264/AVC Rate Control," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 58, no. 3, pp. 184-188, Mar. 2011. Article (CrossRef Link)

[28] L. Shen, Z. Liu, Z. Zhang, and X. Shi, "Rate control based on incremental Proportional-Integral-Differential algorithm," *Optical Engineering*, vol. 46, no. 7, pp. 0770021-0770027, July 2007. Article (CrossRef Link)

[29] D. Fani and M. Rezaei, "Novel PID-Fuzzy Video Rate Controller for High-Delay Applications of the HEVC Standard," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 28, no. 6, pp. 1379-1389, June 2018. Article (CrossRef Link)

[30] T. Yan, I. Ra, Q. Zhang, X. Hang, and L. Y. Huang, "A novel rate control algorithm based on ρ Model for Multiview High Efficiency Video Coding," *Electronics*, vol. 9, no. 1, p. 166, 2020. Article (CrossRef Link)

Tao Yan received the Ph.D. degrees in communication and information systems from Shanghai University, Shanghai, China, in 2010. He has been with the faculty of the School of Information Engineering, Putian University, where he is currently a Associate Professor. His major research interests include Multiview High Efficiency Video Coding, rate control, video codec optimization. He has authored or co-authored more than 20 refereed technical papers in international journals and conferences in the field of video coding and image processing. He currently presides National Natural Science Foundation project.

In-Ho Ra received the Ph.D. degrees in Computer Engineering from Chung-Ang University, Seoul, Korea, in 1995. He has been with the faculty of the School of Computer, Information and Communication, Kusan National University, where he is currently a Professor. His major research interests include block chain. He currently presides the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2016R1A2B4013002).

Deyang Liu received his BS degree in communication engineering from Anqing Normal University, Anqing, China, in 2011, and his MS degree and Ph.D. degree in signal and information processing from Shanghai University, Shanghai, China, in 2014 and 2017. He is currently an Associate Professor in School of Computer and Information, Anqing Normal University. From 2019 to 2020, he was a visiting scholar with University of Technology Sydney, Sydney, NSW, Australia. He is in charge of several projects including the National Natural Science Funds project. His research interests include 3-D video processing, video coding.

Qian Zhang received the Ph.D. degrees in communication and information systems from Shanghai University, Shanghai, China, in 2011. She has been with the faculty of the School of Information and Electromechanical Engineering, Shanghai Normal University, where she is currently a Associate Professor. His major research interests include Multiview High Efficiency Video Coding. She has authored or co-authored more than 30 refereed technical papers in international journals and conferences in the field of video coding and image processing.