ADVANCED RATE CONTROL TECHNOLOGIES FOR MVC

Tao Yan[1], In-Ho Ra[2], Na Song[1], Yu Youhao[1],
Zhicheng Wu[1] and Zhongyu Shi[1]

[1]School of Information Engineering, Putian University, Putian 351100, China
[2]School of Computer, Information and Communication Engineering,
Kunsan National University, Gunsan 54150, South Korea

ABSTRACT

After analyzing the research status and existing problems of multi-view video coding and bit rate control, we found that in addition to achieving higher coding efficiency, scalability characteristics, and quality consistency, reasonable bit rate control is urgent What needs further research. The paper proposes a multi-view video coding rate control algorithm based on the quadratic rate distortion (RD) model is presented. There are already many rate control algorithms. However, the testing work is very important, and different sequences need to be tested to effectively judge the effectiveness of the algorithm. Experimental simulation results show that the algorithm can effectively control the bit rate of multi-view video coding, while maintaining efficient coding efficiency, compared with the current MVC using JVT with fixed quantization parameters.

KEYWORDS

MVC(multi-view video coding), Rate control, Bit allocation, Basic unit layer.

1. INTRODUCTION

From the perspective of information theory, video compression is to remove the redundancy in video information, that is, to retain uncertain information and remove certain information. Video coding is an important part of digital video processing. Its main purpose is to represent video information with as few bits as possible while ensuring a certain reconstructed image quality. In video compression coding, how to effectively remove information redundancy is the main research issue. After decades of development, there have been many methods to eliminate information redundancy.

Although MVC is the extension of JM (Joint Model) of H.264, it has more picture types and estimation methods for disparity and motion vectors. Differing from the motion estimation, disparity estimation makes the H.264 rate control strategy incompatible. However, current most rate control algorithms concentrate on 2D video which are not suit for multi-view video coding(2-4). New MVC rate control strategy should allocate the appropriate bits temporally to avoid the overflow of buffer while it should also allocate bits between inter views to keep the quality of each view totally equal.

There are already many rate control algorithms(5-11). However, the testing work is still very heavy, and different sequences need to be tested to effectively judge the effectiveness of the algorithm. This reduces the change in lighting and ensures that in the left and right images, the same image block of the same object has the same gray value. However, in practical applications, the geometric characteristics of parallel optical axes cannot be accurately realized. In practical applications, since the left and right cameras cannot be exactly the same, two important deviations occur. First of all, it is difficult to ensure that the bottom edges of the imaging areas of the two cameras are collinear, or even accurate to within one pixel. Second, after digitization, the variance and mean of the pixel gray levels of the left and right images cannot be completely equal. Both of these points can be accomplished with a technique called camera calibration. Because of the fact that multi-view video has three or more views and more complicated bit allocation and rate control strategies are needed for it, it is not reasonable to continuedly adopt rate control model for stereoscopic video coding with two views to multi-view video. The MVC rate control algorithms are being preliminarily developed recently and most of the researchers have just studied the binocular stereo video. The stereo vision imaging system theoretically requires two cameras to have identical optical characteristics.

2. STRUCTURE OF RATE CONTROL

Multi-view video is a technique in which an object or a scene is recorded using several synchronous cameras from different positions, disparity-compensated prediction together with motion-compensated prediction are exploited to reduce all kinds of redundancy. Using the JVT MVC encoder, our rate control algorithm has 4 layers, GGOP (the group of group of pictures), GOP, Frame and Basic-Unit. GOP, Frame and Basic-Unit are conformable to JVT-G012. GOP_k is the kth view in a GGOP and there are N_{view} GOP in a GGOP if let N_{view} denote the number of views.

3. RATE CONTROL ALGORITHM FOR MVC

Our algorithm adopts the fluid-flow traffic model, HRD (hypothetical reference decoder) and linear prediction model of MAD (the mean absolute difference). The main procedures of the test algorithm are described as follows:

This section also uses the previous algorithm. In terms of frame level rate control of JVT-G012, the target bit rate for each frame is determined according to the target buffer level, the predefined frame rate, the available channel bandwidth, the actual buffer occupancy and the remaining bits, in which residual energy is not considered, as a result, the above scheme generally occurs skipped frame and quality degradation. Lei proposed an optimal bit allocation scheme based on residual energy for coding frame as follows,

$$T(j) = \frac{MAD_j}{MAD_a} \cdot \frac{(T - \sum_{m=1}^{M} C_m)}{M} + C_j \qquad (1)$$

where T denotes total bit budget for M frames, C_j and Cm denotes bit budget used for non-texture information of jth frame and mth frame, respectively, MAD_j denotes mean absolute

difference for jth frame, MAD_a denotes the average value of mean absolute difference for all frames. From formula (1), it is obvious that the more target bits should be assigned to those frames with larger MAD_j and C_j value. The proposed rate control scheme replaces M with $N(i)$, then T denotes bit number assigned for xth view, $T_{GOP}(n_{x,0})(x = 1, 2, \cdots, N_{view})$.

In multi-view video coding, in order to determine coded block pattern for macro-block, quantization level should be given in advance, however, quantization level is determined before mode decision when no residual coefficients are generated needed for current frame, and residual energy used to implement target bit allocation for frame layer could not be obtained. To settled the problem, Shen[12] proposed a method to predict current frame coding complexity on the basis of mean absolute difference for the previous frame, temporal activity for current frame and temporal activity for the previous frame. Frame difference between two adjacent frames is applied to define temporal activity for frames, temporal activity for jth frame is defined as follows:

$$FD(j) = \frac{1}{X \cdot Y} \cdot \sum_{x=0}^{X} \sum_{y=0}^{Y} \left| I_j(x, y) - I_{j-1}(x, y) \right| \qquad (2)$$

where X, Y denotes the number of luminance samples in the horizontal direction and vertical direction, respectively, x and y denote indexs of luminance samples in the horizontal direction and vertical direction, respectively, $I_j(x, y)$ and $I_{j-1}(x, y)$ denote the luminance values of pixels in position (x, y) in current frame and previous frame.

Based on target bit allocation technique illustrated in equation (2), target bit number allocated for $(j-1)^{th}$ frame is given by:

$$T'_r(j-1) = [\frac{MAD_{j-1}}{MAD_a} \cdot (\frac{T_{GOP}(n_{x,0})}{N(i)} - C_a) + C_{j-1}] \qquad (3)$$

where MAD_{j-1} is mean absolute difference for jth frame, MAD_a is the average value of mean absolute difference of all encoded frames for the current group of pictures, C_a is the average value of non-texture information of all coded frames for the current group of pictures.

It is confirmed that those frames with larger temporal activity need more bits, and vice versa. Thus, we improve the target bit allocation method as illustrated in equation (4) in terms of mean absolute difference and temporal activity for previous frame, temporal activity for current frame, to predict current frame coding complexity. The target bits for current frame is computed as follows.

$$T_r(j) = T_r'(j-1) \cdot \frac{\sum_{l=1}^{L} W(l) \cdot 2^n}{\sum_{l=1}^{L} \chi \cdot \frac{FD(l-1)^2}{\frac{1}{L-1} \cdot \sum_{k=1}^{L-1} FD(k)^2} \cdot W(l) + \sum_{l=1}^{L} \gamma \cdot W_B(l) \cdot (2^n - 1)} + T_j + \psi \qquad (4)$$

where $FD(j)$ and $FD(j-1)$ represent temporal activity for j[th] frame and (j-1)[th] frame, respectively. χ, γ and ψ are adjustment parameters, which are obtained through a large number of experiments.

After each frame been encoded, $T_{GGOP}(sn_{i,j})$ will be refreshed as follows:

$$T_{GGOP}(sn_{i,j}) = T_{GGOP}(sn_{i,j-1}) - A(sn_{i,j-1}) \qquad (5)$$

where, $A(sn_{i,j-1})$ denotes the real number of bits of the (j-1)[th] frame in the ith GGOP.

3. EXPERIMENT RESULT

This article basically uses existing algorithms for testing. There are already many rate control algorithms. However, the testing work is still very heavy, and different sequences need to be tested to effectively judge the effectiveness of the algorithm. We carried out experiments using several sequences with different image properties , which include Akko& kayo, Rena, Vassar1, flameon2. The Flameco2' sequence is obtained by resampling the Flameco2 sequence. The other sequences are similar.
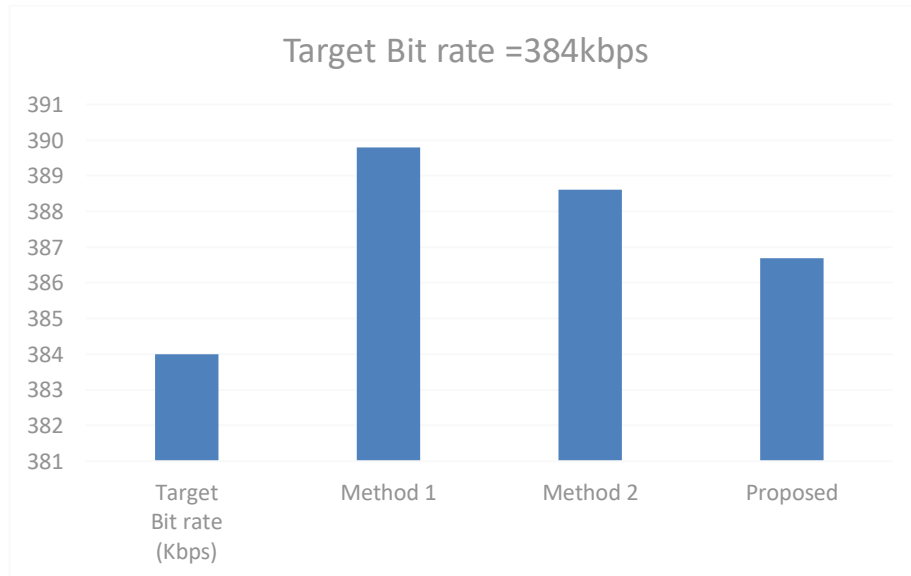

Figure 1 Experimental results

Table. 1 Simulation Results

Sequence	Target bits /(kbps)	Actual bits /(kbps)			Error/(%)		
		Method 1	Method 2	Algorithm used	Method 1	Method 2	Algorithm used
Flameco2'	256	259.07	258.07	258.02	1.20%	0.81%	0.79%
	392	396.78	395.72	395.33	1.22%	0.95%	0.85%
	512	521.48	520.12	519.05	1.85%	1.59%	1.38%
Vassar1'	256	258.15	258.82	257.79	0.84%	1.10%	0.70%
	392	400.9	399.51	395.21	2.27%	1.92%	0.82%
	512	523.96	518.66	515.19	2.34%	1.30%	0.62%
Akko&Kayo'	384	389.8	388.61	386.69	1.51%	1.20%	0.70%
	512	518.14	517.89	516.25	1.20%	1.15%	0.83%
	768	774.91	774.14	772.84	0.90%	0.80%	0.63%
Rena'	256	260.76	259.3	258.33	1.86%	1.29%	0.91%
	392	396.61	395.37	394.76	1.18%	0.86%	0.70%
	512	516.35	516.1	514.65	0.85%	0.80%	0.52%
Aveg.					1.43%	1.15%	0.79%

Table 1 illustrates the coding results of the test rate control scheme, which demonstrates that the test scheme can efficiently control the bit rate with an average rate control error of 0.79%. We takes note of that when the target bit rate is 512kbps, the Flameco2 sequence has larger rate control error, which is due to that the disparity is larger and movement is intense so that results in large error in bit allocation and MAD prediction. In our experiment, both modeled scene and real scene are used for evaluate validity of the test method.

4. CONCLUSION AND FUTURE WORK

There is still not much work on MVC RC. After analyzing the disadvantages and character of MVC, we propose a MVC-RC algorithm based on the quadratic RD model after analyzing the

character of MVC. It maintains the constant quality by distributing bit rate among different views rationally based on correlation analysis. Experimental results show that the test algorithm can control the bit rate accurately. When a scene changes, an effective video scene detection method needs to be further considered to ensure that the algorithm can quickly and accurately detect the video scene. Scene change detection is an issue to be considered in the future.

REFERENCES

[1] K. Muller, H. Schwarz, D. Marpe et al., "3D high-eiciency video coding for multi-view video and depth data," IEEE Transactions on Image Processing,vol.22,no.9,pp.3366–3378,2015.

[2] D. K. Kwon, M. Y. Shen, and C. J. Kuo. Rate control for H.264 video with enhanced rate and distortion models. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 5, 517-529.

[3] Y. Liu, Z. G. Li, and Y. C. Soh. Rate control of H.264/AVC scalable extension. IEEE Transactions on Circuits and Systems for Video Technology, 2004, 1, 116-121.

[4] L. Q. Shen, Z. Liu, Z. Y. Zhang. Frame-level bit allocation based on incremental PID algorithm and frame complexity estimation. Journal of Visual Communication and Image Representation, 2009, 1, 28-34.

[5] Joint Video Team of ITU-T VCEG and ISO/IEC MPEG. WD Reference software for MVC , San Diego, USA, 19–26, 2006.

[6] 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, 2015, 2, 667-673.

[7] J. E. Lim, J. Kim, K. N. Ngan, and K. H. Sohn. Advanced rate control technologies for 3D-HDTV. IEEE Transactions on Consumer Electronics, 2003, 4, 1498-1507.

[8] 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.

[9] Bruno Boessio Vizzotto, Bruno Zatt, Muhammad Shafique, Sergio Bampi.A model predictive controller for frame-level rate control in multiview coding[C]. 2012 IEEE International Conference on Multimedia and Expo, Melbourne, Australia, 2012, 485- 490 .

[10] Li Z G, Pan F, Lim K P, Feng G N. Adaptive basic unit layer rate control for JVT [C]// JVT-G012, the 7th Meeting, Pattaya, Thailand. 2003: 7–14.

[11] J. Y Xu, M. Xu, Y. A Wei, Z. L. Wang, and Z. Y. Guan, "Fast H.264 to HEVC Transcoding: A deep learning method," IEEE Transactions on Multimedia, Vol. 21, No. 7, pp. 1633-1645, July 2019.