

A FRAME LEVEL RATE CONTROL ALGORITHM FOR SCREEN CONTENT CODING

Shanshe Wang¹, Junru Li¹, Shiqi Wang²

¹ Institute of Digital Media, Peking University
Beijing, China

Siwei Ma¹, Wen Gao¹

²City University of Hong Kong
Hong Kong, China

Abstract—In this paper, we propose a frame level rate control scheme for screen content video (SCV) coding. Firstly, inter frame correlation (IFC) is computed to describe the similarity of successive frames for SCV. Based on the IFC, the frames in SCV are further classified into key frames (KFs) and non-key frames (NKF). Secondly, an efficient bit rate allocation scheme is proposed for KFs and NKFs based on the IFC and Hypothetical Reference Decoder (HRD), which can well ensure the R-D performance. Finally, rate quantization (R-Q) models for KFs and NKFs are established by considering the characteristics of the two types of frames, respectively. As such, accurate coding bits allocation at frame level can be achieved by the proposed rate control algorithm. The proposed rate control scheme introduces low coding complexity burden as well as low coding delay, enabling its application in real screen sharing scenarios. Compared with the recommended rate control scheme in HEVC, experimental results demonstrate that the proposed rate control scheme can achieve better performance in terms of both rate control accuracy and R-D performance. In particular, the average bit rate mismatch is within 1.4%, and the average R-D performance improvements are over 19%.

Key words: Rate control, screen content coding, bit allocation, R-Q model

I. INTRODUCTION

Rate control is always a hot research topic in the video coding [1][2][3][4][5]. Compared with the natural videos, SCVs have significantly different characteristics such as sharp content, repeating patterns, large and unnatural motion. Based on these distinct features, many advanced coding tools have been developed to improve the coding performance for the SCC extension of HEVC [6]. Due to the special characteristics of SCV and the newly adopted coding technologies [8][9][10][11][12], rate control for SCV needs to be specifically designed. In [13], the authors incorporated the R- λ model in HEVC [14] into the rate control of SCC based on a sliding window. High rate control accuracy and better R-D performance can be achieved. However, the proposed method requires complexity estimation for all frames in the sliding window, which introduces significant coding delay.

In this paper, we propose a similarity evaluation criterion for two successive frames in SCV, termed as inter frame

correlation (IFC). Based on IFC, the frames are classified into two categories: key frame (KF) and non-key frame (NKF). Then a frame level rate control algorithm is proposed. In particular, IFC based bit allocation is performed, and virtual buffer is incorporated to assist the bit allocation to avoid the bit rate overflow or underflow. Moreover, improved R-Q models are established by considering the properties of the KF and NKF. Experimental results show that compared to the R- λ model in HEVC, the proposed scheme can achieve better R-D performance and smaller bit rate mismatch.

The rest of the paper is organized as follows. Section II analyzes the distinguished characteristics of SCV. In Section III, the proposed rate control scheme is described in detail including bit allocation and improved R-Q model. Section IV presents the experimental results. Finally, we conclude the paper in section V.

II. INTER-FRAME CORRELATION AND RATE DISTORTION ANALYSIS FOR SCV

For SCV, distinguished properties have been observed in temporal domain. For successive frames, in contrast with the camera captured video (CCV), the content of SCV remains for a short time instead of gradually changing. In particular, some successive frames in a sequence are almost identical except for a small region. However, in other occasions, the two successive frames are almost completely different.

Based on the above observations, we define inter frame correlation (IFC) for SCV to evaluate the similarity between the two successive frames. The calculation of IFC is detailed as follows. Firstly, the two successive frames, i.e., the current coding frame and previous reconstructed frame, are divided into non-overlapping blocks with size $N \times N$. Here, N is set to be 16. Subsequently, for the i -th block of current coding frame, denoted as C_i , its inter correlation with previous frame is calculated according to the collocated block in the previous frame, denoted as P_i . Firstly, the Sum of Absolute Difference (SAD) of the two corresponding blocks, C_i and P_i , is calculated as

$$SAD_i(P_i, C_i) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} |P_i(j, k) - C_i(j, k)|, \quad (1)$$

where $P_i(j,k)$ and $C_i(j,k)$ are pixels of P_i and C_i , respectively. If the achieved SAD is smaller than a predefined threshold, then the two collocated blocks will be regarded as Similar Blocks (SB). In our paper, the threshold is set to be $2.5 \times N \times N$.

When completing the calculation for all the blocks, the number of SBs, denoted as N_{SB} , can be obtained and IFC of the current coding frame is defined as follows,

$$\delta = \frac{N_{SB}}{N_T}, \quad (2)$$

where δ is the IFC between the current coding frame and its previous reconstructed frame, and N_T is the total number of $N \times N$ blocks of a frame. Based on δ , the frames can be categorized into two types: key frame (KF) and non-key frame (NKF). In particular, if δ is smaller than a predefined threshold τ , it belongs to KF. Otherwise, it is regarded as NKF. In this paper, the threshold τ is set to be 0.99.

III. PROPOSED LOW DELAY RATE CONTROL FOR SCREEN CONTENT CODING

In this section, proposed rate control algorithm is detailed. It mainly includes the following aspects: IFC and hypothetical reference decoder (HRD) based bit allocation, improved R-Q model.

A. IFC and HRD Based Bit Allocation

Bit allocation is an essential part in rate control. A well designed bit allocation method can achieve high rate control accuracy and better R-D performance. Since content characteristics and the R-D characteristic of KF and NKF in SCV are remarkably different, it is more suitable to conduct bit allocation for them respectively. The proposed bit allocation methods mainly includes the following parts.

Firstly, if there is always no KF or NKF detected with the proposed method in Section II, then average bit allocation is directly utilized as follows,

$$R_t = \frac{BR_{remaining}}{F_{remaining}}, \quad (3)$$

where $BR_{remaining}$ denotes the remaining bits, $F_{remaining}$ is the number of remaining coding frame.

Otherwise, once KF and NKF are both detected, the bit allocation changes to the following scheme. For the KF, the target bits, R_{t_key} , is allocated as,

$$R_{t_key} = \frac{BR_{remaining}}{F_{remaining}} \times O_{key}, \quad (4)$$

where O_{key} is the overflow of the coded KFs which is presented as follows,

$$O_{key} = \frac{T_{actual_key}}{Target_{key}}, \quad (5)$$

where T_{actual_key} denotes the sum of actual consumed bits for the KFs. $Target_{key}$ is the corresponding total allocated target bits. The initial value of O_{key} is 1.

For the NKF, the target bits, $R_{t_non_key}$, are allocated as follows,

$$R_{t_non_key} = \frac{BR_{remaining}}{F_{remaining}} \times O_{non_key}. \quad (6)$$

where O_{non_key} is the overflow of the coded non-key frames which is presented as follows,

$$O_{non_key} = \frac{T_{actual_non_key}}{Target_{non_key}}, \quad (7)$$

where $T_{actual_non_key}$ denotes the actually consumed bits for the NKFs. $Target_{non_key}$ is the corresponding total allocated target bits. The initial value of O_{non_key} is 1.

Furtherly, for the coding frame with IFC equaling to 1, the target bits is further modified to half of the target bits based on Eq.(4) or Eq.(6).

For the above bit allocation, another important issue is to avoid the overflow and underflow. In this paper, the proposed rate control scheme takes HRD into consideration. Specifically, after achieving the target bit for the coding frame, R_t for the frame is further clipped to satisfy the HRD constraints, which is given by

$$R_t = MIN(T_U, MAX(R_t, T_L)), \quad (8)$$

where T_U and T_L are the upper and lower boundary of HRD requirements. They are initialized at each IDR frame and updated with the actual coding bits after encoding one frame. The initialization process is performed as,

$$T_U = W_{vb} \cdot \alpha, \quad (9)$$

$$T_L = BR / F_r, \quad (10)$$

where α is a constant which is set to be 0.8, F_r is the frame rate and W_{vb} is the virtual buffer size which is two times of the target bit rate. BR is the target bit rate.

The updating process is performed as follows:

$$T_U += BR / F_r - b, \quad (11)$$

$$T_L += BR / F_r - b, \quad (12)$$

where b is the actual coding bits of the currently encoded frame.

B. Improved R-Q Model for SCV

After the bit allocation, the subsequent procedure is to determine a suitable QP to meet the target bit rate. In our previous work [3], we proposed an adaptive rate control model for camera captured videos in HEVC. The model is a linear model which can be represented as follows

$$R_n = \frac{\theta \cdot X_n}{Q_n}, \quad (13)$$

where R_n is the target bit rate for the coding frame and Q_n indicates the quantization scale (QS), θ is the model parameter. X_n represents the complexity estimation of the coding frame which is given by,

$$X_n = \left(\frac{\omega_i \times X_{n-1} + SATD_n}{1 + \omega_i} \right)^{0.4} \times QP_{n-1} \times R_{n-1}, \quad (14)$$

where $SATD_n$ is the sum of the absolute transformed difference of the current coding frame. ω_i is a weighting factor which is set to be 0.5. QP_{n-1} and R_{n-1} are the actual QP and coding bits for the previous $(n-1)$ -th frame.

From Eq.(14), it can be concluded that the complexity estimation of the current coding frame has much dependency on the previous coded frames. For the camera captured videos, it may bring more accurate complexity estimation since the successive frames in natural video have similar content characteristic. However, for SCV, it may be not suitable due to the special content characteristic. As shown in Section II, R-D characteristics of KF and NKF differ greatly. Therefore, it is natural to design different rate control scheme for KF and NKF. In this paper, based on the analysis in the above sections, we propose an improved R-Q model for screen content rate control as follows,

$$R_t = \frac{\theta \cdot X_\Omega \cdot f_\Omega}{Q} (\Omega = key, non_key), \quad (15)$$

where Ω denotes the frame type including KF and NKF. f_Ω is considered to be an adjustment factor for different types of frames, which is related to the target allocated bits and the actual assumed bits for the previously encoded frames. For the KF, f_Ω is always set to be 1. For the NKF, f_Ω is determined as

$$f_\Omega = \frac{T_{total} / N_{coded}}{BR}, \quad (16)$$

where T_{total} is the total actual consumed bits and N_{coded} is the number of the coded frames.

The complexity estimations of KF and NKF also work in their own way. Intuitively, the key frame has less inter correlation with the previous coded frame. Thus it is more appropriate for the key frame that the updating of w_i should be less dependent with the previously coded frames. As such, the weighting factor for the KF in Eq.(14) is set to be 0.3. On the contrary, for NKFs, it can be concluded that the NKF has more inter correlations with the previous coded frame. In some occasions, it is almost the same with the previous frame. Thus, the updating weighting factor for the non-key frame in Eq.(14) is set to be 0.75.

Different from the calculation of SATD in [3], in the proposed scheme, the calculation of SATD is detailed as follows. Firstly, the current coding frame is divided into non-overlapped blocks with size of 32x32. Then $SATD$ of each block is calculated. Finally, $SATD$ of the whole frame is set to be the sum of $SATD$ of all the blocks.

In order to better meet the target bit rate, we propose to utilize the adjustment factor and IFC to further modify the QP for the current coding frame. For the frames with the IFC smaller than 0.5, the achieved QP does not change furtherly. For the frames with the IFC between 0.5 and 0.99, an additional QP offset is added to the achieved QP as follows,

$$QP_{offset} = \begin{cases} 3 & f_\Omega > 1.2 \\ 2 & 1.1 < f_\Omega \leq 1.2. \\ 0 & else \end{cases} \quad (17)$$

For the frames with the IFC higher than 0.99, the achieved QP is modified as follows,

$$QP = \begin{cases} \min(QP_{pre} - 2, QP) & f_\Omega < 0.97 \\ QP & else \end{cases} \quad (18)$$

where QP_{pre} is the QP of the previous coded frame.

IV. EXPERIMENTAL RESULTS

In this section, experiment results are presented to verify the effectiveness of proposed rate control scheme. Our proposed scheme is implemented on HM-SCM8.2. As the proposed scheme focuses on SCV and real time transmission, all the video sequences recommended by HEVC SCC common test condition are used as test sequences and all these video sequences are tested under low delay (LD) testing configuration. The target bitrates are achieved from the HM-SCM8.2 anchor under fixed QP. The performance of rate control algorithm is measured by the bit rate mismatch and BD-rate [15].

Bit rate mismatch, denoted as M , is calculated by

$$M = \frac{|R_{target} - R_{actual}|}{R_{target}} \times 100\%, \quad (19)$$

where R_{actual} denotes the actual coding bit rate of the encoded frames. R_{target} is the given target bit rate.

Table 1 shows the detailed performance for the classic SCVs. It can be seen that the proposed scheme can achieve smaller bit rate mismatch. The average bit rate mismatch error is only 1.38. However, the bit rate error of R- λ scheme is over 4.9%. In terms of rate distortion performance, it can be observed that the proposed scheme can also obtain significant rate distortion performance improvement. The average coding gain is over 19%.

V. CONCLUSIONS

In this paper, we analyze the distinguished properties of SCVs from the perspectives of temporal correlations and R-D behaviors. Inter frame correlation is subsequently introduced to characterize the similarity between successive frames. Based on the obtained IFC, a frame level rate control scheme is carefully designed. Specifically, efficient bit allocation is performed based on IFC and Hypothetical Reference Decoder constraint. Then improved R-Q models are further developed for KF and NKF respectively to derive appropriate QP. Experimental results show that the proposed scheme can achieve higher rate control accuracy and better R-D performance.

Table 1 Rate control performance comparisons between R- λ rate control and proposed rate control scheme for YUV444 sequences

Sequence	resolution	Target bit rate (kbps)	R- λ scheme in SCM8.2			Proposed			BD-rate
			Bit rate (kbps)	PSNR (dB)	Bit rate Mismatch	Bit rate (kbps)	PSNR (dB)	Bit rate Mismatch	
Basketball_Screen (YUV)	2560×1440	5258.57	5163.27	42.67	1.81%	5284.94	45.15	0.50%	-36.8%
		2481.32	2467.94	38.88	0.54%	2486.53	41.14	0.21%	
		1289.00	1288.67	34.82	0.03%	1290.81	37.77	0.14%	
		735.70	733.84	31.19	0.25%	739.36	35.86	0.50%	
MissionControlClip3 (YUV)	1920×1080	2068.80	2586.98	44.20	25.05%	2099.50	45.25	1.48%	-37.1%
		1131.74	1345.89	39.98	18.92%	1144.49	41.13	1.13%	
		632.23	716.23	35.86	13.29%	639.86	38.94	1.21%	
		375.49	393.89	31.82	4.90%	381.06	35.09	1.48%	
sc_desktop (YUV)	1920×1080	898.01	912.16	38.43	1.58%	914.87	47.59	1.88%	-18.2%
		844.56	854.90	37.00	1.22%	862.74	42.37	2.15%	
		780.20	785.11	34.89	0.63%	791.81	40.15	1.49%	
		704.16	705.27	32.88	0.16%	723.32	37.66	2.72%	
sc_console (YUV)	1920×1080	4510.68	4567.64	53.36	1.26%	4608.16	53.88	2.16%	-4.8%
		3625.24	3693.38	48.22	1.88%	3703.96	49.23	2.17%	
		2844.25	2852.27	42.24	0.28%	2884.60	43.84	1.42%	
		2161.09	2162.47	37.06	0.06%	2203.12	38.39	1.94%	
sc_programming (YUV)	1280×720	6013.97	6022.92	47.52	0.15%	6020.70	48.15	0.11%	-13.4%
		3334.12	3338.67	43.19	0.14%	3341.94	44.09	0.23%	
		1618.55	1622.20	38.78	0.23%	1621.94	39.66	0.21%	
		741.14	743.70	34.54	0.35%	743.29	35.59	0.29%	
sc_SlideShow (YUV)	1280×720	596.13	567.87	47.35	4.74%	599.54	52.28	0.57%	-27.0%
		341.86	334.86	44.68	2.05%	342.12	47.45	0.08%	
		199.42	197.97	40.97	0.73%	195.40	43.00	2.01%	
		123.73	123.90	37.57	0.14%	121.81	39.00	1.56%	
Map (YUV)	1280×720	1538.68	1314.67	44.14	14.56%	1524.34	47.11	0.93%	-14.4%
		990.33	870.41	40.36	12.11%	970.11	42.56	2.04%	
		601.00	547.49	36.63	8.90%	601.07	38.68	0.01%	
		356.88	350.21	33.43	1.87%	352.69	34.48	1.17%	
WebBrowsing (YUV)	1280×720	164.24	179.89	44.30	9.53%	166.10	46.13	1.14%	-21.7%
		131.05	151.12	41.55	15.32%	138.23	42.37	5.48%	
		101.04	118.12	37.80	16.90%	104.18	40.60	3.11%	
		77.79	90.62	33.95	16.49%	82.70	37.83	6.31%	
Robot (YUV)	1280×720	5491.34	5490.09	40.62	0.02%	5505.50	40.37	0.26%	-4.2%
		1796.78	1798.09	37.11	0.07%	1790.38	37.15	0.36%	
		616.87	615.79	34.32	0.17%	619.75	34.60	0.47%	
		244.96	244.21	31.99	0.30%	246.50	32.36	0.63%	
Average					4.91%		1.38%	-19.7%	

ACKNOWLEDGMENT

This work was supported in part by the General Financial Grant from the China Postdoctoral Science Foundation (2016M600866), National Natural Science Foundation of China (61571017), Top-Notch Young Talents Program of China and Shenzhen Peacock Plan, which are gratefully acknowledged.

REFERENCES

- [1] H. Choi, J. Nam, J. Yoo, D. Sim, I.V. Bajić, Rate control based on unified RQ model for HEVC, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG118th Meeting: San José, CA, USA, 1–10 February, 2012.
- [2] S. Wang, S. Ma, S. Wang, D. Zhao, and W. Gao, "Rate-GOP Based Rate Control for High Efficiency Video Coding," *IEEE Journal of Selected Topics in Signal Processing*, vol.7, no.6, pp.1101-1111, Dec.2013.
- [3] J. Si, S. Ma, W. Gao, "Adaptive rate control for HEVC," ISO/IEC JTC1/SC29/WG119, JCTVC-I0433, Geneva, CH, 27 April-7 May 2012.
- [4] Y. Tan, C. Ye, Z. Li, "Single-Pass Rate Control With Texture and Non-Texture Rate-Distortion Models," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 22, no. 8, pp. 1236-1245, 2012.
- [5] B. Li, Rate control by R-lambda model for HEVC, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T 1/SC 29/WG 11, 11th Meeting: Shanghai, CN, 10–19 Oct. 2012.
- [6] R. Joshi, and J. Xu, "High Efficiency Video Coding (HEVC) Screen Content Coding: Draft 2," *JCTVC-S1005*, Sapporo, JP, July 2014
- [7] D. Kwon and B. Madhukar, "Fast intra block copy (IntraBC) search for HEVC screen content coding." *Circuits and Systems (ISCAS)*, 2014 IEEE International Symposium on. IEEE, 2014.
- [8] T. Lin, S. Wang, P. Zhang and K. Zhou, "AHG8: P2M based dual coder extension of HEVC," *JCTVC-L0303*, Geneva, CH, Jan. 2013.
- [9] S. Yang, H.J. Shim, D. Lee, B. Jeon, "NonCE2: Transform skip signalling for intra block copy", *JCTVC-S0036*, Strasbourg, France, Oct. 2014.
- [10] Z. Ma, F. Xin and X. Meng, "Advanced screen content coding using color table and index map." *Multimedia and Expo Workshops (ICMEW)*, IEEE International Conference on. IEEE, 2014.
- [11] L. Guo, M. Karczewicz, and J. Sole, RCE3: Results of Test 3.1 on Palette Mode for Screen Content Coding, document *JCTVC-N0247*, Jul. 2013
- [12] H. Zhang et al. "HEVC-based adaptive quantization for screen content by detecting low contrast edge regions." *Circuits and Systems (ISCAS)*, 2013 IEEE International Symposium on. IEEE, 2013.
- [13] Yaoyao Guo, Bin Li, Songlin Sun, Jizheng Xu, "Rate Control for Screen Content Coding in HEVC," *ISCAS*, 2015.
- [14] B. Bross, W.-J. Han, G. J. Sullivan, J.-R. Ohm, T. Wiegand (Editors), "High Efficiency Video Coding (HEVC) text specification draft 6, " Document: JCTVC-H1003, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG118th Meeting: San José, CA, USA, 1–10 February, 2012.
- [15] G. Bjontegaard, "Calculation of average PSNR difference between RD curves," in Proc. ITU-T Q.6/SG16 VCEG 13th Meeting, Austin, TX, Apr. 2001, document VCEG-M33.