An Efficient Frame-Content Based Intra Frame Rate Control for High Efficiency Video Coding

Miaohui Wang, *Student Member, IEEE*, King Ngi Ngan, *Fellow, IEEE*, and Hongliang Li, *Senior Member, IEEE*

***Abstract*—Rate control plays an important role in the rapid development of high-fidelity video services. As the *High Efficiency Video Coding* (HEVC) standard has been finalized, many rate control algorithms are being developed to promote its commercial use. The HEVC encoder adopts a new R-lambda based rate control model to reduce the bit estimation error. However, the R-lambda model fails to consider the frame-content complexity that ultimately degrades the performance of the bit rate control. In this letter, a gradient based R-lambda (GRL) model is proposed for the intra frame rate control, where the gradient can effectively measure the frame-content complexity and enhance the performance of the traditional R-lambda method. In addition, a new coding tree unit (CTU) level bit allocation method is developed. The simulation results show that the proposed GRL method can reduce the bit estimation error and improve the video quality in HEVC all intra frame coding.**

***Index Terms*—Frame-content complexity, HEVC, lagrange multiplier, perceptual coding, rate control.**

I. Introduction

With the rapid development of high-fidelity video services in recent years, how to efficiently compress video data has been considered by many organizations and companies. To meet this demand, ITU-T Q.6/SG16 (VCEG) and ISO/IEC JTC1/SC29/WG11 (MPEG) have established the Joint Collaborative Team on Video Coding (JCT-VC) to develop a new video coding standard–*High Efficiency Video Coding* (HEVC) [1]. As the HEVC has adopted many advanced encoding tools, it can significantly improve the compression performance of high definition (HD) videos and save half of the bit rate compared to the H.264/Advanced Video Coding (AVC) [2] for the same perceptual video quality.

Rate control plays a significant role in transmitting high-quality video data via the communication channel. If the channel bandwidth is smaller than the output bit rate of the encoder, the encoded bits will accumulate in the encoder buffer. When the buffer size of these accumulated bits is too large, the encoder needs to skip some encoding frames to alleviate the buffer delay and avoid the buffer overflow. On the contrary, if the channel bandwidth is larger than the output bit rate of the encoder, this means that some channel bandwidth is wasted, and the encoder buffer may be underflow. Thus, the goal of rate control is to achieve a balance between the compressed video quality and the channel bandwidth to meet specific applications.

Intra frame rate control is more important in HEVC than in H.264/AVC. As the motion compensated prediction (MCP) is greatly improved in HEVC, the bit rate saving for MCP frames (i.e., B or P frames) is higher than for the intra-coded frames. Consequently, intra frame bits of the HEVC occupy more channel bandwidth than those of the H.264/AVC in a compressed video stream. In addition, initial quantization parameter (QP) of the first frame (i.e., intra-coded frame) greatly affects the channel bandwidth. On the one hand, a high-quality intra frame needs much more bandwidth than the MCP frames in a group of pictures (GOP), which may cause the buffer overflow and frame-skip on the encoder side. On the other hand, a low-quality intra frame that is used as a reference frame for the MCP frames degrades the video quality in the current GOP and may waste the channel bandwidth. Thus, intra frame rate control algorithms [3]–[6] have been considered in both the H.264/AVC and HEVC encoders.

In this letter, we propose a new gradient based R-lambda (GRL) model for the HEVC intra frame rate control, where the gradient is used to measure the frame-content complexity. In addition, a novel bit allocation method is developed for the coding tree unit (CTU) rate control. To the best of our knowledge, this is the first experimental study in which the relationship between rate-gradient and lambda is modeled and a new content complexity based bit allocation is developed for the HEVC intra frame rate control. Simulation results show that in HEVC, the proposed GRL method can significantly improve the performance of the intra frame rate control compared to state-of-the-art methods.

The rest of the letter is organized as follows. We briefly introduce the related works in Section II. The proposed GRL method is addressed in Section III. The simulation results are presented in Section IV while the concluding remarks are given in Section V.

II. Related Rate Control Models in HEVC

Recently, many rate control models are being considered in the HEVC encoder. In this section, we give a brief review of different rate models to facilitate the study of the rate control algorithms. In HEVC, rate models can be roughly classified into three categories: quadratic model, ρ-domain model and R-lambda model.

Manuscript received May 26, 2014; revised August 18, 2014; accepted November 25, 2014. Date of publication December 04, 2014; date of current version December 08, 2014. This work was supported in part by a grant from the Research Grants Council of the Hong Kong SAR, China, under Project CUHK 415712. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Marco Mattavelli.

M. Wang and K. N. Ngan are with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: mhwang@ee.cuhk.edu.hk; knngan@ee.cuhk.edu.hk).

H. Li is with School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 610051, China (e-mail: hlli@uestc.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2014.2377032

In [7], [8], the quadratic method has been extended to the early version of the HEVC reference software. However, two limitations of the quadratic model need to be further considered. Firstly, the quadratic model only determines the encoding bits of the transform coefficients but ignores the overhead bits. Secondly, the quadratic model contains a well-known "chicken and egg" dilemma [4] in the computation of quantization step.

In [9], the ρ-domain method has been studied in HEVC, where ρ is the percentage of zeros among quantized coefficients. The experiments in [9] show a little improvement in comparison with the quadratic method on the HEVC reference software. However, it should be noted that the ρ-domain method is developed for the codec with a fixed transform block size, which is not compatible with the quad-tree based coding structure in HEVC, especially when the transform block size can be varying from 4×4 to 32×32.

In modern video codecs (i.e., H.264/AVC and HEVC), the Lagrange multiplier (i.e., lambda) not only explicitly affects the bits of a block but also implicitly affects the determination of its encoding modes for rate distortion optimization (RDO) scheme. Thus, the Lagrange multiplier has been considered as an important factor for the rate control methods [10], [11] in H.264/AVC. Recently, Li *et al.* [12] developed a new R-lambda method for the HEVC codec.

Compared with the quadratic and ρ-domain models, the R-lambda model considers the overall bit rate including both the transform coefficient bits and the overhead bits. In addition, parameters in the R-lambda model can be easily computed. Due to its high efficiency and low complexity, the R-lambda model has been adopted as the rate control method for the HEVC codec by the JCT-VC. However, the model does not consider the frame complexity, which should enhance the performance of the R-lambda based rate control methods in HEVC. Thus, in this letter, the frame-content complexity is studied and incorporated into an improved R-lambda model.

III. Proposed Rate Control Method

In this section, we introduce the proposed intra frame rate control method for the HEVC encoder, including a new frame-content complexity based R-lambda model and a new CTU level bit allocation method.

A. Modeling the Relationship Between Rate-Gradient and Lambda for the HEVC Intra Frame Coding

In general, the output bits of the encoded frames in a video sequence have varying content from each other. The reason is that different frames have different encoding complexities. To reduce the bit estimation error, the frame-content complexity measure is incorporated into the proposed method for the HEVC intra frame coding.

The gradient based frame-content complexity measure has been considered in the H.264/AVC intra frame rate control [4], [5]. It is reported that the gradient per pixel (GPP) of the intra-coded frames has an approximated linear relationship with their encoded bits. The GPP is defined by

$$GPP = \frac{1}{H \times W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} \begin{pmatrix} |I_{i,j} - I_{i+1,j}| \\ + |I_{i,j} - I_{i,j+1}| \end{pmatrix}, \quad (1)$$

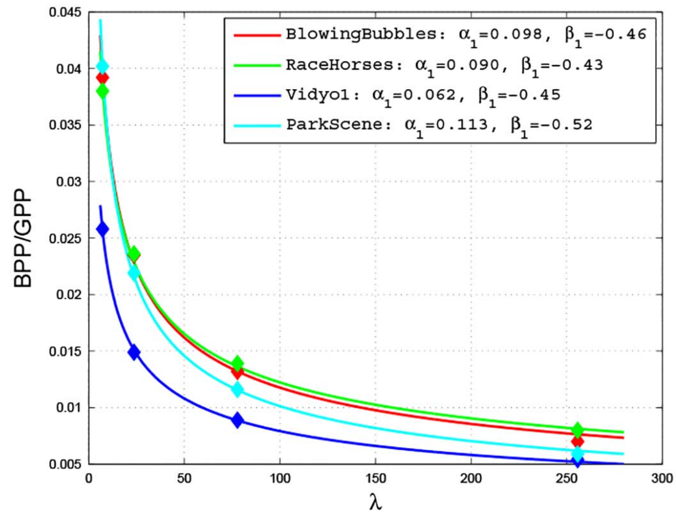


Fig. 1. Modeling the rate-gradient and lambda for intra frame coding. Four QPs (i.e., $\{22, 27, 32, 37\}$) are used in the experiments.

where H and W are the height and width of one CTU or one frame, and $I_{i,j}$ is the pixel value at position (i, j).

Since the HEVC utilizes similar encoding tools as the H.264/AVC in intra frame coding, it is believed that the GPP can also be used to measure content complexity of the HEVC intra-coded frames. However, the HEVC employs more flexible partitions, prediction directions and transform sizes than the H.264/AVC in intra frame coding. Thus, the relationship between rate-gradient and lambda should be modeled in HEVC intra frame coding.

Fig. 1 shows the relationship between rate-gradient and lambda in HEVC. In the experiment, the encoder is configured the same as in Section IV. There are a total of 120 frames encoded with the All-intra Main configuration. The average results of BPP (i.e., bits per pixel), GPP and λ of four QPs (i.e., $\{22, 27, 32, 37\}$) are used to model the relationship between rate-gradient and lambda. It is observed that in HEVC intra frame coding, a hyperbolic function can adequately fit the data.

$$\frac{BPP}{GPP} = \alpha_1 \lambda^{\beta_1}, \quad (2)$$

where α_1 and β_1 are the model parameters. Other video sequences also present similar characteristics. In the proposed GRL method, equation (2) is used to compute content based λ that is used to compute a corresponding QP.

As mentioned above, the initial QP greatly affects the performance of many rate control algorithms. Given the model parameters, the traditional R-lambda method computes a fixed initial QP for different video sequences. Obviously, the fixed initial QP inevitably degrades the performance of the R-lambda method, because different frames have varying content complexities as well as varying bits. On the contrary, our method can obtain a frame-content based λ and the initial QP for different video sequences based on their frame gradients. One can see that our method can obtain a more accurate initial QP than conventional R-lambda methods according to the experiments in Section IV.

B. Bit Allocation

A typical rate control algorithm usually consists of two steps: allocating the target bits and achieving the allocated target budget. As mentioned above, the target bit rate can be achieved by three rate models in HEVC. Based on the same rate model, different bit allocation methods can yield quite distinct results. Consequently, bit allocation is an important component of a rate control algorithm. The proposed method follows a classical three-level bit allocation structure: GOP level, frame level and CTU level.

GOP Level Bit Allocation: Given a target bit rate R_{tar} and the frame rate f, the target bits allocated to the current GOP is

$$T_{GOP} = \frac{R_{tar} \times (N_{coded} - sw) - R_{coded} \times f}{sw \times f} \times N_{GOP}, \quad (3)$$

where T_{GOP} is the target bits for the current GOP, R_{coded} is the bits of the number of pictures (i.e., N_{coded}) that have already been encoded, N_{GOP} is the number of pictures in the current GOP, and sw is the smooth window [13].

Frame Level Bit Allocation: Since in all intra frame coding, $N_{GOP} = 1$, the target bits allocated to the current picture is

$$T_{CurrPic} = T_{GOP}, \quad (4)$$

CTU Level Bit Allocation: When the frame level $T_{CurrPic}$ is allocated, the proposed CTU level bit allocation method is based on the GPP of each CTU in the current frame.

$$R_i = \frac{GPP_{CurrCTU}}{\sum_{k=1}^{N_{CTU}} GPP_k} T_{CurrPic}, \quad (5)$$

where GPP_k is the k th CTU gradient, R_i is the allocated bits for the i th CTU, and N_{CTU} is the number of CTUs in the current frame.

The CTU level bit allocation is updated as

$$T_{CurrCTU} = \left[R_{RemBits} + \frac{\sum_{j=1}^{i-1} (R_j - R_{act,j})}{sw} \right] w_{CurrCTU}, \quad (6)$$

where $T_{currCTU}$ is the allocated bits for the current (i.e., i th) CTU, $R_{RemBits}$ is the remaining bits used to encode the rest of CTUs in the current frame, $R_{act,j}$ is the actual bits of the j th CTU and $w_{CurrCTU}$ is the weight for the current CTU defined by

$$w_{CurrCTU} = \frac{GPP_{CurrCTU}}{\sum_{k=i}^{N_{CTU}} GPP_k}. \quad (7)$$

In traditional R-lambda methods, bit-budget of a CTU is allocated according to the mean absolute difference (MAD) of the CTU at the same position in the previous decoded frame. Obviously, MAD of the previous CTU may be quite different from the current one, and hence the MAD value used as a weight performs poorly in allocating the bit-budget. Instead of the MAD weight, we propose to use the gradient of the current CTU to allocate the bit-budget, because the gradient can be obtained directly from the current CTU, and it has an approximated linear relationship to the bit rate in intra frame coding. In the experiments (see Section IV), one can see that the gradient method can achieve a more accurate bit estimation than the MAD method.

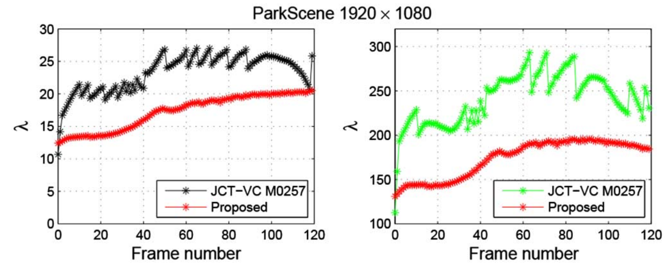


Fig. 2. Comparison of the updated λ between the JCT-VC M0257 [14] and the proposed method. In the experiments, the target bit rate is obtained by the anchor HEVC encoder at QP $= 27$ (left) and QP $= 37$ (right), respectively.

C. Model Parameter Update

After encoding one CTU or one frame, we use the actual encoded bits to update the values of λ_{new}, α_{new} and β_{new}.

$$\lambda_{new} = \alpha_{old} \cdot \left(\frac{BPP_{act}}{GPP}\right)^{\beta_{old}}$$
$$\alpha_{new} = \alpha_{old} \times \left(\frac{R_{act}}{R_{tar}}\right)^{c\beta_{old}}$$
$$\beta_{new} = \beta_{old} + c\left(\frac{R_{act} - R_{tar}}{R_{tar}}\right), \quad (8)$$

where R_{act} is the actual bits of an encoded CTU or frame, and c is a scaling factor that reduces the speed of β changes, and in our experiments it is set to 0.125. Note that the larger the scaling factor c is, the larger variation of λ is obtained.

Fig. 2 shows four results of the updated λ versus the frame number in both the proposed method and the JCT-VC M0257 [14] that is proposed for the intra frame rate control in HEVC. It can be seen that our method can attain a smaller λ fluctuation that indicates a smaller QP fluctuation and a more stable buffer status in comparison with the JCT-VC M0257.

D. Summary of the Proposed Algorithm

The proposed GRL algorithm is as follows.

1) Bit allocation according to equations (3)~(7);
2) λ_{ctu} computation for each CTU according to equation (2);
3) QP computation for each CTU by $QP_{ctu} = a \times ln(\lambda_{ctu}) + b$, where a and b are two constants;
4) Encoding and recording related encoding parameters;
5) Parameter update according to equation (8);
6) Going to Step 1 until the end of encoding.

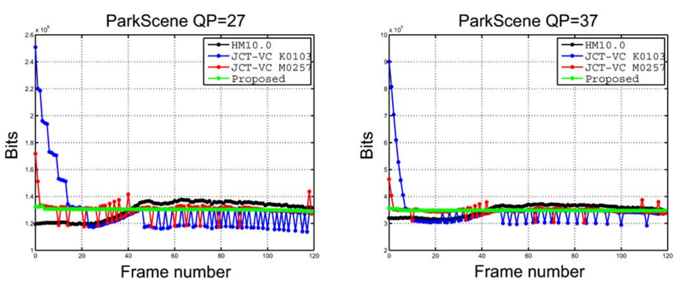
IV. Simulation Results

To evaluate performance of the proposed GRL method, we have implemented it on the HEVC reference software HM10.0 [15] configured with the common test conditions as suggested in [16]. The All-intra Main configuration is used in the experiments, where the RDQ and RDQTS are disabled. When the rate control is enabled, a target bit rate is assigned to each frame. The target bit-budget assigned to each frame is the average of the bit rates obtained by the anchor HEVC with four QPs=$\{22, 27, 32, 37\}$. In the simulation, we have performed four methods with the same configuration.

- HM10.0: the original HM10.0 without rate control.
- JCT-VC K0103: the original HM10.0 with the default rate control.

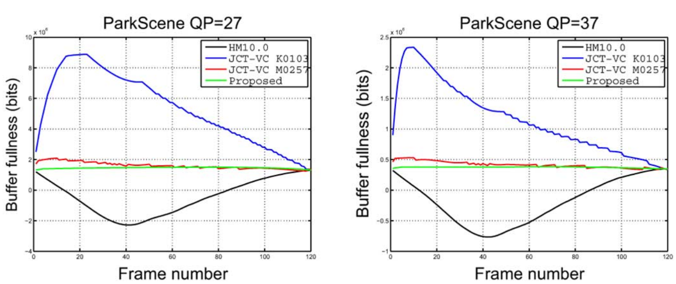
TABLE I
SIMULATION RESULTS FOR THE PERFORMANCE COMPARISON IN HEVC RATE CONTROL

Sequences		QP	HM10.0		JCT-VC K0103		JCT-VC M0257		GRL	
			PSNR variation (Std. Dev.)	Average bit rate (M bps)	PSNR variation (Std. Dev.)	Bit estimation (NRMSE %)	PSNR variation (Std. Dev.)	Bit estimation (NRMSE %)	PSNR variation (Std. Dev.)	Bit estimation (NRMSE %)
		22	0.01	10.5	0.46	7.98	0.23	3.89	0.04	0.25
		27	0.02	6.4	0.68	13.54	0.20	3.77	0.04	0.37
Class A	2500×1600	32	0.02	3.7	0.78	18.03	0.23	4.73	0.07	0.90
		37	0.02	2.1	0.83	21.34	0.24	4.52	0.12	1.48
		22	0.06	9.3	0.76	12.56	0.39	6.19	0.24	0.24
		27	0.13	6.5	0.78	15.79	0.40	4.94	0.35	0.53
Class B	1920×1080	32	0.21	2.4	0.83	17.64	0.52	4.67	0.48	0.68
		37	0.27	1.3	0.90	20.50	0.62	4.51	0.57	1.04
		22	0.13	3.0	1.51	17.49	0.77	3.35	0.72	0.63
		27	0.21	1.8	1.85	25.71	0.82	3.86	0.78	0.28
Class C	832×480	32	0.29	1.1	2.86	39.81	0.84	5.25	0.79	0.63
		37	0.37	0.6	1.93	46.62	0.80	5.81	0.74	1.12
		22	0.32	2.4	1.77	18.96	1.07	3.52	0.99	1.13
		27	0.38	1.4	2.77	31.77	1.18	4.26	0.91	0.76
Class D	416×240	32	0.41	0.8	3.32	45.28	1.16	5.34	0.84	1.34
		37	0.42	0.5	2.32	53.57	0.83	6.23	0.76	1.73
		22	0.17	0.8	0.33	5.64	0.33	3.99	0.26	0.40
		27	0.15	0.5	0.28	3.64	0.32	3.67	0.28	0.46
Class E	1280×720	32	0.13	0.3	0.40	4.88	0.37	4.01	0.31	0.60
		37	0.15	0.2	0.56	7.30	0.45	4.41	0.41	1.71
Encoding time			100%		105%		106%		106%	



(a) (b)

Fig. 3. Bits comparison versus frame number for *ParkScene*.



(a) (b)

Fig. 4. Buffer fullness comparison versus frame number for *ParkScene*.

- JCT-VC M0257: the original HM10.0 with the default intra frame rate control.
- Proposed method: the proposed GRL method.

In the experiments, the normalized root mean square error (NRMSE) in equation (9) is the measure for bit estimation accuracy of each method, and the standard deviation of PSNR is the measure of video quality variation.

$$NRMSE = \frac{1}{\overline{R}_{act}} \sqrt{\frac{\sum_{k=0}^{N-1} \left(R_{act}^{k} - R_{tar}^{k} \right)^2}{N}}, \tag{9}$$

where N is the number of encoded frames, $\overline{R}_{act}$ is the average of actual bits, R_{act}^{k} is the number of actual bits and R_{tar}^{k} is the target bits in the k th frame. Note that a smaller NRMSE means that a better bit estimation result is achieved.

Fig. 3 shows the bit estimation comparison of five methods for the sequence "*ParkScene*" (1920 × 1080). It is observed that at the frame level, our method can more accurately achieve the allocated target bit budget than other methods as shown by the smooth bit rate variation over the frames. Since our GRL method considers the frame-content complexity, a better initial QP is obtained in comparison with other methods, which guarantees that the encoder can quickly approach to a given channel bandwidth.

Fig. 4 shows the buffer fullness comparison of five algorithms for the sequence "*ParkScene*". We can see that the buffer fullness variation of our method is much smaller than other methods. In this case, we know that the encoder buffer is operating at a safer level (i.e., not underflow or overflow) when employing the proposed GRL method as compared to other methods.

More detailed numerical experiments are tabulated in Table I. From this table, we can observe that the proposed GRL method is able to achieve a more accurate bit-budget estimation. The NRMSE of the GRL method ranges from 0.24% to 1.73%. Compared with the JCT-VC M0257 [14] whose NRMSE ranges from 3.35% to 6.23%, our GRL method shows a significant improvement. The smaller bit-budget estimation errors result in a smoother buffer fullness variation and hence more consistent video quality over the video frames (see PSNR variation). This has increased the perceptual quality of the reconstructed frames. In addition, it is worth noting that both the GRL method, and the traditional methods do not increase the decoding complexity at all.

V. CONCLUSION

In this letter, a frame-content based rate control method is proposed for the HEVC intra frame coding. The frame-content complexity is measured by its gradient, which has been incorporated into an improved R-lambda model. In addition, a new bits-budget allocation method is developed for intra frame coding at the CTU level. Experimental results show that the proposed method can significantly reduce the bit estimation error for the HEVC intra frame rate control, and hence more consistent video quality over the video frames is obtained in comparison with conventional methods.

REFERENCES

[1] G. J. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, 2012.

[2] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the h. 264/avc video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, 2003.

[3] J. Lee, I. Shin, and H. Park, "Adaptive intra-frame assignment and bit-rate estimation for variable GOP length in H. 264," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 10, pp. 1271–1279, 2006.

[4] X. Jing, L.-P. Chau, and W.-C. Siu, "Frame complexity-based rate-quantization model for H. 264/AVC intraframe rate control," *IEEE Signal Process. Lett.*, vol. 15, pp. 373–376, 2008.

[5] B. Yan and M. Wang, "Adaptive distortion-based intra-rate estimation for H. 264/AVC rate control," *IEEE Signal Process. Lett.*, vol. 16, no. 3, pp. 145–148, 2009.

[6] Y. Zhou, L. Tian, and X. Ning, "Intra frame constant rate control scheme for high efficiency video coding," in *IEEE Int. Conf. Computing, Networking and Communications (ICNC)*, 2013, pp. 648–652.

[7] H. Choi, J. Nam, J. Yoo, D. Sim, and I. Bajić, "Rate control based on unified RQ model for HEVC," JCT-VC H0213, 2012.

[8] H. Choi, J. Yoo, J. Nam, D. Sim, and I. Bajic, "Pixel-wise unified rate-quantization model for multi-level rate control," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1112–1123, Dec. 2013.

[9] X. Liang, Q. Wang, Y. Zhou, B. Luo, and A. Men, "A novel rq model based rate control scheme in HEVC," *Vis. Commun. Image Process.*, pp. 1–6, 2013.

[10] M. Jiang and N. Ling, "On lagrange multiplier and quantizer adjustment for H.264 frame-layer video rate control," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 5, pp. 663–669, 2006.

[11] M. Wang and B. Yan, "Lagrangian multiplier based joint three-layer rate control for H. 264/AVC," *IEEE Signal Process. Lett.*, vol. 16, no. 8, pp. 679–682, 2009.

[12] B. Li, H. Li, L. Li, and J. Zhang, "Rate control by R-lambda model for HEVC," JCT-VC K0103, 2012.

[13] L. Xu, D. Zhao, X. Ji, L. Deng, S. Kwong, and W. Gao, "Window-level rate control for smooth picture quality and smooth buffer occupancy," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 723–734, Mar. 2011.

[14] X. Wang and M. Karczewicz, "Intra frame rate control based on SATD," *JCT-VC M0257, Incheon*, 2013.

[15] "HEVC reference software," [Online]. Available: http://hevc.kw.bbc.co.uk/trac/browser/jctvc-hm/tags 2012

[16] F. Bossen, "Common test conditions and software reference configurations," Joint Collaborative Team on Video Coding, JCT-VC G1200, Nov. 2011.