

Studying Rate Control Methods for UHD TV Delivery Using HEVC

Ivan Zupancic¹, Matteo Naccari², Marta Mrak², Ebroul Izquierdo¹

¹ School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK

² Research and Development Department, British Broadcasting Corporation, London, UK

Email: i.zupancic@qmul.ac.uk

Abstract—Since the early video coding standardisation efforts, rate control has been considered essential for almost any application. With the advent of improved video coding standards and the introduction of advanced flexible coding tools, previous Rate-Distortion (RD) models used for rate control have become obsolete. To address this issue, some rate control methods have been recently proposed specifically for the current state-of-the-art High Efficiency Video Coding (HEVC) standard which introduce many useful features, such as a robust correspondence between the rate and Lagrange multiplier λ . However, when applying these rate control methods on sequences in the new Ultra High Definition Television (UHD TV) format, degraded coding performance was observed. In this paper, an analysis of the state-of-the-art HEVC rate control method was performed and two directions for its improvement were evaluated. These improvements target frame-level bit-allocation and model parameter initialisation. When compared to the rate control method implemented in the HEVC reference software, these improvements result in reduced BD-rate losses of 3.1% and 2.1%, versus the 8.8% provided by the reference algorithm. Moreover, the proposed improvements increase the accuracy in hitting the target bit-rate.

Keywords—HEVC, rate control, UHD video

I. INTRODUCTION

Ultra High Definition Television (UHD TV) is the new format which is expected to deliver a greater impact, deliver more presence and immersion than the current High Definition Television (HDTV). Besides higher resolution, UHD TV has the potential to deliver wider colour gamut, high dynamic range and high frame rate [1]. Due to an enormous volume of data associated with UHD TV signals, improved compression techniques should be employed when delivering UHD TV services. To answer these needs, the new state-of-the-art H.265/High Efficiency Video Coding (HEVC) standard [2] has been developed recently. It can provide the same perceived video quality as its predecessor H.264/Advanced Video Coding (AVC) [3] for up to 60% bit-rate reduction for UHD TV content [4].

Even though improved compression technology is key in enabling the delivery of UHD TV content, it is also equally important to distribute the available bit-budget so that the quality of the decoded signal is maximised. A rate control method aims to optimise the visual quality given the limited

bandwidth constraints. Generally speaking, rate control can be divided into two main steps. The first one allocates the right amount of bits to each level of the coding process, i.e. Group Of Pictures (GOP), frame, macroblock or Coding Unit (CU) in HEVC. In the second step, the allocated rate is used to derive the amount of compression to be applied over a given part of the video sequence. Depending on the latency requirements, rate control can be performed in single- or multi-pass fashion.

In this paper, the state-of-the-art rate control method in HEVC reference software (HM) [5] was analysed and two possible approaches for improving its performance are then proposed. These approaches examine the impact of frame-level bit-allocation and model parameter initialisation.

II. RELATED WORK

Rate control is one of the crucial tools for any practical video codec. Throughout the years, various models were proposed to quantify the relationship between coding parameters and coding rate. Early attempts in modeling the relationship between coding rate and Quantisation Parameter (QP) date back to the MPEG-2 Video [6] and MPEG-4 Part 2 (Visual) [7] standards. The reference implementation of the AVC standard uses a rate control method based on a quadratic Rate-Quantisation ($R - Q$) relationship [8], where Laplacian distribution of the residual information is assumed [9]. A rate control method based on the quadratic $R - Q$ model [10] was used in early versions of the HEVC reference implementation [5]. However, due to the flexible quadtree partitioning used in HEVC, the $R - Q$ model is not sufficiently accurate to quantify the relationship between rate and quantisation step.

Another group of rate control methods tries to build a relationship between the rate and the percentage (ρ) of coefficients which are quantised to zero [11]. A quadratic ρ -domain rate model was proposed by Wang *et al.* [12] and used in a hierarchical bit-allocation scheme for rate control in an HEVC codec. Nonetheless, transform blocks in HEVC can vary in size, resulting in loosen relationship between ρ and rate.

The relation between Lagrange multiplier λ and coding rate was firstly analysed by Li *et al.* [13]. The proposed hyperbolic $R - \lambda$ model shows higher correlation when compared with the above mentioned models. The $R - \lambda$ model is utilised in the rate control method used in HM, where the bit-budget is allocated at three different levels of granularity. This rate control method was further improved for intra frames [14] using

Part of the work described in this paper has been conducted within the project COGNITUS. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687605.

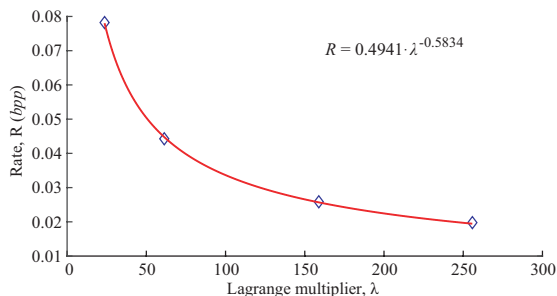


Fig. 1: $R - \lambda$ curve for *Manege* test sequence. The sequence was encoded using 4 QP values (27, 31, 35, and 37) and rates obtained are denoted as blue diamonds. The fitted $R - \lambda$ curve is depicted in red.

the Sum of Absolute Transformed Differences (SATD) as a complexity measure. Finally, Wang and Ngan [15] proposed a method which uses the distortion of collocated blocks in the previous frame to establish a linear relationship between distortion and λ used for different bit-allocation.

III. RATE CONTROL IN HM IMPLEMENTATION

The existing rate control method used in HM is based on the recently proposed $R - \lambda$ model [13]. It was shown that there exists a robust relation between the rate R (in bits per pixel) and the Lagrange multiplier λ which can be expressed with a hyperbolic function:

$$R = a \cdot \lambda^b, \quad (1)$$

where a and b are parameters related to the video source. An example of $R - \lambda$ relationship is shown in Fig. 1. Due to its improved accuracy and robustness, the rate control method based on the $R - \lambda$ model defined in (1) has been included in the HM reference implementation since Version 9.0. The algorithm can be divided in two parts: bit-allocation and actual coding using the $R - \lambda$ model. Bit-allocation is performed at three different levels, namely GOP, frame, and basic unit level. Basic unit in this context is represented by square areas of 64×64 image pixels associated with the luminance component and denoted as Coding Tree Unit (CTU) in the HEVC standard [2]. When allocating bits at frame level, the amount of bits a frame has assigned is weighted according its position in the GOP. Throughout this paper, when the QP assigned to each frame is selected according to a predefined pattern based on its temporal layer, the encoding will be denoted as Variable Bit-Rate (VBR) coding. At basic unit level, the weights to allocate the available bit-budget are calculated dynamically using the prediction error from a collocated basic unit in the previously coded frames of the same temporal layer.

Once the target rate is determined, it is straightforward to compute λ using the inverse of (1):

$$\lambda = \alpha \cdot R^\beta, \quad (2)$$

where α and β are the model parameters. However, the main problem here is how to determine the parameters α and β , which are generally content dependent. Also, in case of random access GOP structure, different temporal layers may have different model parameters, and hence multiple sets of parameters have to be used within the sequence. In the existing

approach, the corresponding α and β are continuously updated after encoding one basic unit or one frame. Finally, the QP value is determined as:

$$QP = c_1 \cdot \ln(\lambda) + c_2, \quad (3)$$

where c_1 and c_2 are set to 4.2005 and 13.7122, respectively. For practical use, QP is rounded to the nearest integer value. Finally, to keep the video quality consistent, both λ and QP should not vary significantly with time. Thus, the values for λ and QP are bounded to a range centred around the values used in previously encoded frames and basic units.

IV. HM RATE CONTROL PERFORMANCE ANALYSIS

This section presents the performance analysis for the rate control method implemented in HM. Coding conditions used for experimental evaluation are described first. The experimental evaluation of the existing rate control method in HM is presented next with two possible performance improvements.

A. Test Material and Coding Conditions

The test set used in this paper comprises 16 sequences with 8 bits per component, 4:2:0 chroma format, 3840×2160 spatial resolution, and frame rate of 50 and 60 fps. More information on the test sequences can be found elsewhere [16]. Each sequence is encoded with four QP values which were determined by visually inspecting the test set compressed with QP ranging from 22 to 45. All the sequences were encoded according to the JCT-VC Common Test Conditions (CTC) [17] using the selected QP values and the Random Access Main (RA-Main) configuration. Compression efficiency and rate accuracy are used as performance metrics in this paper. For compression efficiency, Bjøntegaard Delta-rate (BD-rate) [18] is used, computed between the anchor data and the data obtained after encoding with a tested encoder. In this context, negative BD-rate values correspond to compression efficiency gains. The rate accuracy is measured as an absolute percentage deviation from the target rate. Lower value corresponds to higher rate accuracy. Experimental analysis presented in this paper was performed using a fast HEVC encoder implementation based on HM version 12.0 [5], denoted as *HM-fast*, which was tailored for fast encoding of UHD content [16]. All the results presented here use the *HM-fast* codec run in VBR mode as anchor.

B. Analysis of the Rate Control Method in HM

Although the rate control method described in Section III shows improved coding performance compared to previous methods proposed for HM, it was noticed it is performing suboptimally near the beginning of the sequence. This results in degraded visual quality every time a scene change happens in a video sequence. In particular, very high QP values (e.g. up to 51) were used at the beginning of the sequence, as shown in Fig. 2. This can be explained by the fact that the initial α and β values for frames of all temporal layers are set to predetermined values of 3.2003 and -1.3670 , respectively. That is clearly suboptimal, as λ and the corresponding QP are not calculated using the right model parameters. With model

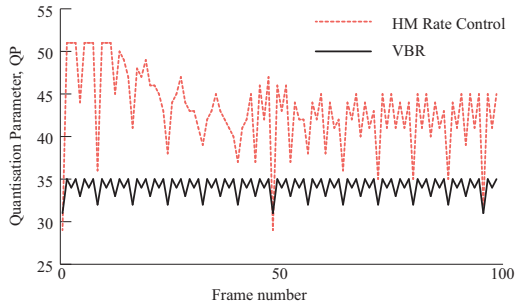


Fig. 2: QP values for the first 100 frames of the *Somersault* test sequence encoded with VBR mode with QP=31 (black), and HM rate control (red).

parameters α and β getting continuously updated, the model gradually becomes more accurate resulting in a better visual quality with time. However, in applications with frequent or continuous scene changes, such as broadcasting, this type of behavior is highly undesirable, as it results in high quality fluctuations of the decoded signal.

The coding performance of the existing rate control method in HM was then evaluated. The *HM-fast* encoder in VBR mode was used as an anchor in this experiment and bit-rate obtained from such encoding was set as a target rate for the rate control. It has to be noted that using a fast encoder, due to its limited number of available encoding tools, makes the rate control process even more challenging. The BD-rate and rate accuracy for the existing rate control method in HM are shown in Table I as *HM Rate Ctrl*. As may be noted, the existing rate control method in HM produces significant coding losses compared with VBR encoding. For instance, BD-rate losses higher than 20% are reported in some cases. However, the rate accuracy for all sequences is reasonably high.

After thorough sequence examination, it was inferred that the main source of such high losses was due to the suboptimal performance near the beginning of the sequence when the model was still inaccurate. This is supported by the fact that the existing rate control method in HM uses predetermined parameter values for $R - \lambda$ model at the beginning of the sequence, since it has no prior knowledge of the content currently being encoded. One possible solution for this issue would be to obtain the sequence information and calculate the model parameters prior to encoding with rate control. The following subsection will present the theoretical framework to improve the performance of the rate control method in HM.

C. Approaches for Improving the Rate Control Efficiency

Based on the previously presented findings, two possible approaches for improving the performance of HM rate control are described here. In both cases, the data after encoding a sequence with a VBR mode is assumed to be available prior to actual encoding with rate control. In the first approach, GOP- and frame-level bit-allocation in the HM rate control method were bypassed and the bit-budget was instead derived from the number of bits spent on each frame during VBR encoding. To handle the cases of bit underspending or overspending, a simple rate management scheme was added to redistribute

TABLE I: BD-RATE AND RATE CONTROL ACCURACY FOR THE THREE EXPERIMENTS DESCRIBED IN SECTION IV.

Sequence	HM Rate Ctrl		Approach 1		Approach 2	
	BD-R [%]	Acc. [%]	BD-R [%]	Acc. [%]	BD-R [%]	Acc. [%]
ParkAndBuildings	4.2	1.3	4.7	0.0	2.3	0.0
NingyoPompoms	6.5	0.0	3.5	0.0	3.5	0.0
ShowDrummer1	23.4	0.0	2.2	0.0	1.2	0.0
Sedof	3.9	0.0	5.0	0.0	5.0	0.0
Petitbato	8.8	0.1	2.1	0.1	1.6	0.0
Manege	1.4	0.0	2.6	0.0	2.0	0.0
ParkDancers	5.0	1.1	-0.1	0.0	-0.4	0.0
CandleSmoke	16.2	0.0	4.2	0.0	2.3	0.0
TableCar	8.2	1.8	2.0	0.0	1.3	0.0
TapeBlackRed	13.6	0.2	2.3	0.0	0.8	0.0
Hurdles	8.9	0.1	2.3	0.0	2.3	0.0
LongJump	5.0	0.0	4.0	0.0	3.9	0.0
Discus	4.1	0.0	3.1	0.0	3.8	0.0
Somersault	21.9	0.0	7.5	0.0	1.1	0.0
Boxing	6.5	0.0	1.5	0.0	2.3	0.0
Netball	4.1	0.0	2.4	0.0	1.2	0.0
Average	8.8	0.3	3.1	0.0	2.1	0.0

the difference bits to future frames based on their temporal layer. The aim of this experiment is twofold: on the one hand, the bit-allocation as designed in the rate control method of HM can be evaluated and its accuracy assessed. On the other hand, the accuracy of the $R - \lambda$ model can also be thoroughly investigated, since the errors that may be caused by the bit-allocation module are eliminated. The BD-rates for the luminance component and rate accuracy associated with this experiment are shown in Table I as *Approach 1*. It can be observed that replacing the existing GOP- and frame-level bit-allocation with the frame size obtained from VBR encoding mode improves the performance significantly. The worst case performance has been drastically improved. Moreover, the accuracy of achieving the target rate was further improved compared to the existing rate control method in HM. This experiment confirms that the existing bit-allocation scheme in HM can be improved using some prior sequence knowledge.

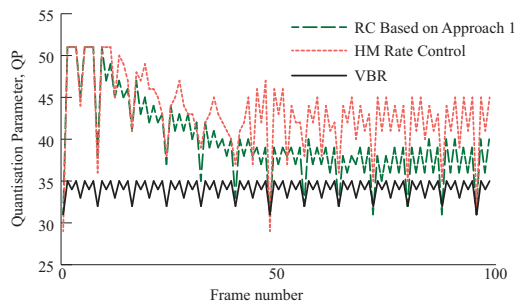
In the second approach, the initial model parameters were also updated with the values obtained from the model fitting with VBR data. The bit-rates and associated λ values, obtained after VBR encoding with four different QP values, were used to fit the $R - \lambda$ model from (2). The fitting is performed separately for each temporal layer and for each sequence. The cost minimised during the fitting is the sum of absolute differences between the QP value predicted by the model and the one used during encoding, with the relation between QP and λ from (3). This can be formally written as:

$$\arg \min_{\alpha, \beta} \sum_{i=1}^4 |(c_1 \cdot \ln(\alpha R_i^\beta) + c_2) - (c_1 \cdot \ln(\lambda_i) + c_2)|. \quad (4)$$

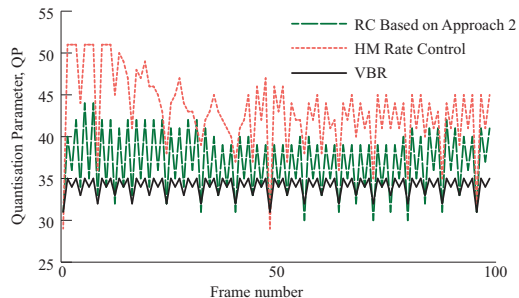
Since c_1 and c_2 are constants, this can be simplified to:

$$\arg \min_{\alpha, \beta} \sum_{i=1}^4 \left| \ln \left(\frac{\alpha \cdot R_i^\beta}{\lambda_i} \right) \right|, \quad R_i, \lambda_i > 0 \quad \forall i, \quad (5)$$

where R_i and λ_i are corresponding rate and λ value for each QP value, respectively. The reason for minimising the cost using the QP value is because a poor performance of the $R - \lambda$ model was observed when minimisation was applied to λ directly. In fact, small differences in λ value can translate



(a) QP values used by rate control based on Approach 1.



(b) QP values used by rate control based on Approach 2.

Fig. 3: QP values for the first 100 frames of the *Somersault* test sequence used when encoding with VBR mode (black), HM rate control (red dotted line), and two improved rate control approaches (green dashed line).

into large differences for QP, when λ values are small. This results in undesirable *weights* for different λ values. The α and β values obtained from the fitting process were used to initialise the corresponding parameters for frames of each temporal layer. As in the previous experiment, GOP- and frame-level bit-allocation were replaced with the coding rate obtained from VBR encoding. The results of this experiment are shown in Table I as *Approach 2*. It can be seen that the encoding performance of the modified rate control method has been further improved, with rate the accuracy achieving almost theoretical maximum, i.e. zero.

The performance near to the beginning of each sequence of the two above described approaches has been further evaluated. Fig. 3 shows the QP values used at the beginning of the sequence for the two approaches. It can be seen that QP values used by the two approaches are considerably lower than those used by the existing rate control method, and generally correlate more with QP values from the VBR encoding mode. This is especially true for *Approach 2*, where model parameters were initialised based on data obtained after model fitting.

These results suggest that the existing rate control method in HM can be improved by replacing the frame-level bit-allocation with the coding rate associated with VBR encoding and initialising the parameters based on fitting the actual rate in the $R - \lambda$ model. For most practical applications, this information may not be available prior to encoding and to obtain it, a full sequence needs to be encoded using at least 3 different QP values. This can result in a great computational overhead. However, the findings presented here are useful to guide the development of practical rate control solutions.

V. CONCLUSIONS

The rate control method currently used in the HEVC reference implementation is based on a recently proposed $R - \lambda$ model. Although it shows an improved performance compared to previous methods, it was noticed it performs suboptimal near the beginning of the sequence or after a scene change. In this paper, a study was performed to analyse the performance of the existing rate control method in HM and to identify the directions for its improvement. Specifically, frame-level bit allocation and model parameter initialisation were replaced with the information obtained after encoding video content in VBR mode. Albeit the amount of computational complexity associated with such approaches is significant and not suitable for most practical applications, these proposals serve as a proof of concept for the development of other faster methods. Future research will utilise these findings to develop a fast two-pass HEVC rate control method.

REFERENCES

- [1] ITU-R, "BT.2020: Parameter values for ultra-high definition television systems for production and international programme exchange," ITU-R, Tech. Rep., Aug 2012.
- [2] G. Sullivan, J. Ohm, W. 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.
- [3] T. Wiegand, G. 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, Jul 2003.
- [4] T. K. Tan, R. Weerakkody, M. Mrak, N. Ramzan, V. Baroncini, J. R. Ohm, and G. J. Sullivan, "Video quality evaluation methodology and verification testing of HEVC compression performance," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 4–19, Jan 2016.
- [5] HM Reference Software. Jan 2016. [Online]. Available: <https://hevc.hhi.fraunhofer.de/HM-doc/>
- [6] Test Model 5. Feb 2016. [Online]. Available: <http://www.mpeg.org/MPEG/MSSG/tm5/>
- [7] H.-J. Lee, T. Chiang, and Y.-Q. Zhang, "Scalable rate control for MPEG-4 video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 6, pp. 878–894, Sep 2000.
- [8] K.-P. Lim, G. Sullivan, and T. Wiegand, "Text description of joint model reference encoding methods and decoding concealment methods," JVT-N046, Tech. Rep., Jan 2005.
- [9] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 246–250, Feb 1997.
- [10] H. Choi, J. Nam, J. Yoo, D. Sim, and I. V. Bajic, "Rate control based on unified RQ model for HEVC," JCTVC-H0213, Tech. Rep., Feb 2012.
- [11] Z. He, Y. K. Kim, and S. Mitra, "Low-delay rate control for DCT video coding via rho-domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 928–940, Aug 2001.
- [12] S. Wang, S. Ma, S. Wang, D. Zhao, and W. Gao, "Quadratic rho-domain based rate control algorithm for HEVC," in *Acoustics, Speech and Signal Processing (ICASSP), IEEE Int. Conf. on*, May 2013, pp. 1695–1699.
- [13] B. Li, H. Li, L. Li, and J. Zhang, "Lambda domain rate control algorithm for high efficiency video coding," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 3841–3854, Sep 2014.
- [14] M. Karczewicz and X. Wang, "Intra frame rate control based on SATD," JCTVC-M0257, Tech. Rep., Apr 2013.
- [15] M. Wang and K. N. Ngan, "Optimal bit allocation in HEVC for real-time video communications," in *Image Processing (ICIP), 2015 IEEE Int. Conf. on*, Sep 2015, pp. 2665–2669.
- [16] M. Naccari, A. Gabriellini, M. Mrak, S. G. Blasi, I. Zupancic, and E. Izquierdo, "HEVC coding optimisation for ultra high definition television services," in *Picture Coding Symp.*, May 2015, pp. 20–24.
- [17] F. Bossen, "Common test conditions and software reference configurations," JCTVC-L1100, Tech. Rep., Oct 2012.
- [18] G. Bjontegaard, "Improvements of the BD-PSNR model," ITU-T SG16/Q6, 35th VCEG Meeting, Doc.VCEG-A111, Jul 2008.