# QUADRATIC $\rho$-DOMAIN BASED RATE CONTROL ALGORITHM FOR HEVC

*Shanshe Wang[1], Siwei Ma[2], Shiqi Wang[2], Debin Zhao[1], Wen Gao[2]*

[1] Department of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
[2]Institute of Digital Media, Peking University, Beijing, China

**ABSTRACT**

In this paper, a novel quadratic ρ-domain frame layer rate control algorithm is proposed for Low Delay (LD) configuration in High Efficiency Video Coding (HEVC). Firstly, we propose a quadratic ρ-domain rate quantization (R-Q) model to establish the relationship between bit rate and quantization parameter. Subsequently, inspired by the specialized QP determination and reference picture set (RPS) mechanism, we design a RPS based hierarchical partition structure for rate control. Based on the hierarchical partition, an efficient hierarchical bit allocation scheme is provided. Finally, the quantization parameter (QP) is calculated via the proposed R-Q model after the bit allocation to meet the target bits. Experimental results demonstrate the proposed algorithm can significantly improve the R-D performance compared to the state-of-the-art rate control scheme for HEVC. Moreover, our proposed algorithm also provides smaller mismatch between the target bit rate and the actual bit rate.

***Index Terms*—** HEVC, Rate Control, Rate-GOP

## 1. INTRODUCTION

High Efficiency Video Coding (HEVC) is the latest video coding standard developed by JCT-VC (Joint Collaborative Team on Video Coding) [1] which significantly improves the coding efficiency than the previous video coding standards, such as H.264/AVC [2]. In HEVC, many new technologies are adopted to improve the coding performance, especially the quad-tree coding structure in which three block concepts named coding unit (CU), prediction unit (PU) and transform unit (TU) [3] are introduced. The specialized structure is introduced to specify the coding, prediction and transform information. This highly well-suited coding structure achieves a leap in coding performance and provides the encoder great flexibility to improve the coding efficiency.

This work was supported in part by Major State Basic Research Development Program of China (973 Program, 2009CB320903), National Science Foundation (61121002, 61103088), National High-tech R&D Program of China (863 Program, SS2012AA010805) and National Sci-Tech Support Plan (2011BAH08B01).

Besides the quad-tree coding structure, the specialized reference frame set (RFS) [4] and QP determination mechanism are other two technologies which bring much coding gains.

Rate control plays a crucial role in any coding standard. In the previous video coding standard, many rate control algorithms and rate quantization (R-Q) models have been investigated for rate control, such as TM5 for MPEG-2[5], VM8 for MPEG-4[6] and TMN8 for H.263[7] etc.. According to the video applications, rate control can be classified to CBR (constant bit rate) rate control [2][3][4] and VBR (Variable Bit Rate) rate control[8]. According to the rate control implementation method, it can be classified into one pass [6], two pass or even multi-pass rate control [9][10]. In the previous coding standard H.264/AVC, the quadratic R-Q model in [6] shows great efficiency in rate control which is proposed under the assumption that the predicted residues follow a single Laplacian Distribution. Many modified algorithms based on quadratic R-Q model are proposed to improve the rate control performance [11][12][13][14]. A novel rate control algorithm with an accurate linear model in $\rho$-domain in proposed in [15], which is proved to be very effective in all the previous coding standards including wavelet based video coding. At present, rate control for HEVC has not been thoroughly studied yet. Based on the traditional quadratic R-Q model, [16] proposed a unified R-Q model called quadratic pixel-based unified rate-quantization (URQ) model for rate control of HEVC. The algorithm considered the new features that the size of PU varies so the bit allocation must be accordance with the number of pixels. However, the performance is much worse than HM anchor. In [17], based on the linear correlation between the QP and the parameter Lambda, a $R-\lambda$ model based rate control algorithm is proposed for HEVC which provides much gain over the URQ model.

Owing to the specialized coding structure and new coding technologies of HEVC, in this paper, we propose a quadratic ρ-domain frame layer rate control algorithm for HEVC. Firstly, based on the distribution of residual information, an R-Q model in $\rho$-domain is proposed. Then we design a hierarchical rate control structure based on the specialized reference frame set (RFS) of HEVC. Besides, an efficient hierarchical bit allocation method is proposed to keep the frames in RFS be of high video quality. Finally the proposed quadratic $\rho$-domain R-Q model is utilized to calculate an accurate quantization parameter (QP) to meet the target bits.

The rest of the paper is organized as follows. In section 2, a quadratic $\rho$-domain R-Q model is proposed based on the distribution of the residual information. Section 3 proposes a Rate-GOP based multi layer partition structure and an efficient hierarchical bit allocation scheme for rate control. Section 4 presents the experimental results. Finally, we conclude the paper in section 5.

## 2. PROPOSED QUADRATIC ρ-DOMAIN BASED RATE MODEL

For rate control, the most crucial step is to compute a suitable QP which heavily depends on the R-Q model to meet the target bits of the frame. In the previous work [4], under the assumption that the residual information belongs to a single Laplacian distribution, the source rate is modeled as a quadratic function of the quantization step [6]. The R-Q model shows great efficiency in H.264/AVC.

But it has been stated in [18] that a single Laplacian distribution is not appropriate to capture the distribution of transformed coefficients in HEVC due to the quad-tree structure, and thus a more accurate mixed Laplacian distribution is proposed to represent the distribution as follows.

$$f(x) = \sum_{i=0}^{3} f_i(x) = \sum_{i=0}^{3} \frac{1}{2}\omega_i \lambda_i e^{-\lambda_i |x|} \qquad (1)$$

where $f_i(x)$ denotes the distribution function of transform coefficients of CUs in *i-th* depth. In [15], it has been proved in $\rho$-domain that a linear relation between the texture bits and the number of non-zero transformed coefficients exists for a Laplacian distribution as follows.

$$R = \theta' \cdot (1-\rho) = \theta \cdot N_{non\_zero} \qquad (2)$$

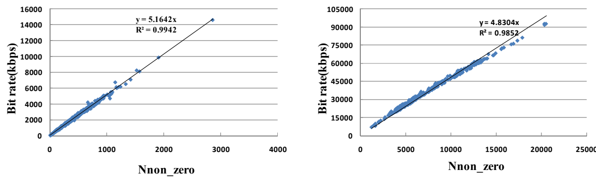
where ρ and $R$ denote the percentage of zero transformed coefficients and the bit rate of the coding frame. $N_{non\_zero}$ indicates the total number of non-zero transformed coefficients of the frame.

According to our experiments, the linear relation still holds firmly in HEVC. Fig. 3 shows the linear relation in $\rho$-domain for the sequence *BascketballPass* and *BQSquare*. It can be seen that the accurate linear relation still holds firmly with zero intercept. Similar observations can be drawn for other benchmark sequences.

Based on the RDOQ of HEVC, for a certain *i-th* level in (1), the distortion can be calculated via (3) and the percentage of zero coefficients, $\rho_i$, can be computed by (4).

$$D_i(q) = 2\int_0^{\delta q} f_i(x)dx + 2\sum_{j=1}^{\infty}\int_{(j-\delta)q}^{(j+\delta)q} f_i(x)\left|x - jq\right|dx \qquad (3)$$

$$\rho_i = \int_{-\delta q}^{\delta q} f_i(x)dx = \int_{-\delta q}^{\delta q} \frac{\lambda_i}{2} e^{-\lambda_i |x|}dx = 1 - e^{-\delta\lambda_i q} \qquad (4)$$




(a) BascketballPass (b) BQSquare

**Fig.1.** The linear relation between the number of non-zero transformed coefficients and the texture bits in $\rho$-domain

As illustrated in [15], for the *i-th* level with a Laplacian distribution, we still can obtain the linear relation as follows

$$R_i = \theta_i \cdot (1-\rho_i) \qquad (5)$$

where $\rho_i$ and $R_i$ denote the percentage of zero transformed coefficients and the bit rate of the *i-th* level. Then bit rate of texture information for the frame can be represented as

$$R = \sum_{i=0}^{3}\theta_i \cdot (1-\rho_i) = \sum_{i=0}^{3}\theta_i \cdot e^{-\delta\lambda_i q} \qquad (6)$$

Under a single Laplacian distribution, a mapping scheme between $\rho$ and QP is provided in [15] to determine a suitable QP to meet the target bits. It may be not suitable to utilize the scheme in HEVC due to the complicated distribution as (1). Consequently, it is necessary to find a suitable way to figure out the quantization parameter. Assume that $N_i$ denotes the number of pixels of *i-th* level and combine with (4), we can get

$$N_{i,non\_zero} = N_i \cdot (1-\rho_i) = N_i \cdot e^{-\delta\lambda_i q} \qquad (7)$$

where $N_{i,non\_zero}$ denotes the number of non-zero transformed coefficients of *i-th* level in a coding frame. Thus the total number of non-zero transformed coefficients of the frame can be obtained as

$$N_{non\_zero} = \sum_{i=0}^{3} N_i \cdot e^{-\delta\lambda_i q} = \sum_{i=0}^{3} N_i \cdot e^{-\delta\lambda_i q} \qquad (8)$$

A Taylor expansion of the above equation yields that the non-zero number of transformed coefficients can be represented by the quadratic function of quantization step.

$$N_{non\_zero} - N = a \cdot q^2 + b \cdot q \qquad (9)$$

where $N$ denotes the total number of pixel of the frame and the two parameters can be represented as

$$a = \sum_{i=0}^{3}(\delta\lambda_i)^2, b = -\sum_{i=0}^{3}\delta\lambda_i \qquad (10)$$

## 3. RATE-GOP BASED MULTI LAYER PARTITION AND HIERARCHICAL BIT ALLOCATION

### 3.1 Rate-GOP Based Multi Layer Partition

The specialized QP determination and RPS mechanism under Low_Delay configuration significantly improves the encoding efficiency. In HEVC Low_Delay configuration, four successive frames except I frame are considered as a coding group called as Rate-GOP in which the QP of each frame has fixed difference with QP of I frame, $QP_I$, as illustrated in (11).

$$QP = \begin{cases} QP_I + 1 & if(POC\%4 == 0) \\ QP_I + 2 & if(POC\%4 == 2) \\ QP_I + 3 & else \end{cases} \quad (11)$$

Different with RPS of H.264/AVC, RPS of HEVC is composed of one frame with the nearest temporal distance and three frames with lowest QP in Decoded Picture Buffer (DPB). For each frame, the *GopId* is firstly calculated as follows

$$GopId = (POC - 1)\%4 \quad (12)$$

Then the RPS for the current frame is determined by the corresponding Delta POC depending on *GopId*. Delta POC is difference between POC of the current frame and POC of previous coded frame. Table 1 presents the Delta POC of the reference frame for different *GopId*. Such specialized RPS takes more consideration of inter dependency into account and significantly improves the coding performance.

**Table 1.** Reference Picture Set (RPS) in HEVC

| GopId | Delta_POC | | | |
|---|---|---|---|---|
| 0 | -1 | -5 | -9 | -13 |
| 1 | -1 | -2 | -6 | -10 |
| 2 | -1 | -3 | -7 | -11 |
| 3 | -1 | -4 | -8 | -12 |

From the Table 1, we can conclude different frame in a Rate-GOP has different influence on the following frames. Generally, the frame with QP as $QP_I$+1 has relatively high reconstructed video quality and will be referred more than once which plays a crucial role for the final performance. Other frame in the same Rate-GOP is only referred once, so they are less important than the frame with $QP_I$+1 in terms of the reference by the following frames. The frame with QP as $QP_I$+1 is considered as a key frame in our paper.

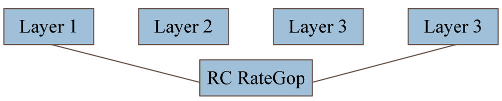



**Fig. 2.** Hierarchical Structure of Rate Control RateGop

Based on the specialized RPS and different influence on the following frames, we define a Rate Control RateGop (RCRG) which is composed of four frames including one key frame and its following three frames and then divide the frames into three layers. Since the first frame plays a crucial role which will be referred by the following frame more times and has great influence on the final coding performance, so the first layer includes the first frame in a RCRG. The second layer includes the second frame of the Rate-GOP, whose video quality mainly depends on the first layer because its nearest reference frame usually has high video quality and the other reference frames is much further. The remaining two frames whose two nearest reference frames locate in the same RCRG belongs to the third layer. Fig. 2 shows our proposed hierarchical structure of RCRG.

### 3.2 Hierarchical Bit Allocation

Bit allocation is a most crucial part during the rate control utilized in the calculation of QP and determines the smoothness of bit rate. Based on the multi layer partition as illustrated in section 3, we propose a hierarchical bit allocation scheme.

Firstly, the bits for a RCRG are allocated as following.

$$T_{RateGop} = N_{RateGop} \times R_{remaining} / F_{num} \quad (13)$$

where $T_{RateGop}$ denotes the target bits for a RCRG, $R_{remaining}$ is the remaining bits after encoding the previous frames, $N_{RateGop}$ indicates the number of the frame in a RCRG and $F_{num}$ denotes the remaining frames to be coded. Due to the temporal correlation between the adjacent RCRGs and in order to get a smooth bit rate, the allocated bits for a RCRG is further modified as,

$$T'_{RateGop} = \alpha \times T_{RateGop} + (1\text{-}\alpha) \times T_{RateGop\_pre} \quad (14)$$

where $T_{RateGop\text{-}pre}$ denotes the actual bits of the previous RCRG and $\alpha$ is set as 0.5 empirically.

Secondly is to allocate suitable target bits for each frame in a RCRG. The frame in upper layer should be allocated more bits due to its crucial role and fewer bits are allocated for the frames in lower layers. So based on the proposed hierarchical partition, the target bits allocated for the frames in different layers is implemented as follow,

$$T_{i,j} = \alpha_{i,j} \times T'_{RateGop} \qquad j = 1,2,3 \quad (15)$$

where $T_{i,j}$ and $\alpha_{i,j}$ denotes the target allocated bits and bits ratio of the frame in *j*-th layer of *i*-th RCRG respectively. The initial value of the parameters in (5) is set as 0.5, 0.3 and 0.2. Then the parameters will be updated adaptively as

$$\alpha_{i,j} = \frac{1}{3}(\alpha_{i\text{-}3,j} + \alpha_{i\text{-}2,j} + \alpha_{i\text{-}1,j}) \quad j=1,2,3 \quad (16)$$

The proposed bit allocation scheme ensures that the frames in layer 1 can be allocated more bits, so frames in the current RFS can be of high video quality and improved coding performance can be expected. After the determination of bits allocation, the proposed rate model (2) and (9) is utilized to calculate the QP to meet the target bits.

**Summary of the proposed algorithm**

**begin**
Allocate target bits for i-th RCRG based on (13)
**begin**
For each frame in the i-th RCRG
1) Allocate target bits for j-th frame as in (14)
2) Predict the value of θ as $\theta = \frac{1}{n}\sum_{i=1}^{n}\theta_i$
3) Calculate the $N_{non}$ based on (2)
4) Determine QP based on quadratic $\rho$-domain model (9)
5) Encoding with the QP
**end**

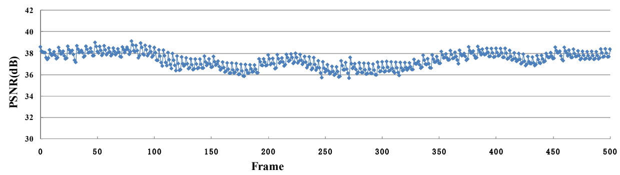
**begin**
Updating the parameters *a*, *b* and θ
**end**
**End**

## 4. EXPERIMENTAL RESULTS

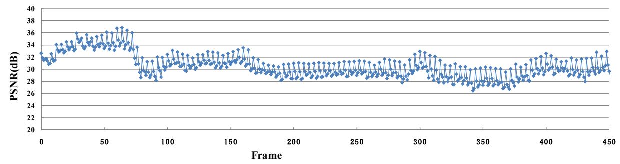
In order to verify the efficiency of our proposed algorithm, we first integrated the proposed algorithm scheme into the HM8.0 and compared with the HM anchor and state-of-art rate control algorithm [17] which is the recommended rate control algorithm in HM at present. The R-D comparisons is tested by the scheme in [19] and the rate estimation accuracy is measured by the frame layer mismatch ratio by

$$M\%=\frac{\left|R_{target}\text{-}R_{actual}\right|}{R_{target}}\times 100\%$$

where $R_{actual}$ and $R_{target}$ denote the actual bits and the target allocated bits of the video sequence respectively. Standard test sequences in different classes provided by HEVC are adopted to test the proposed algorithm.



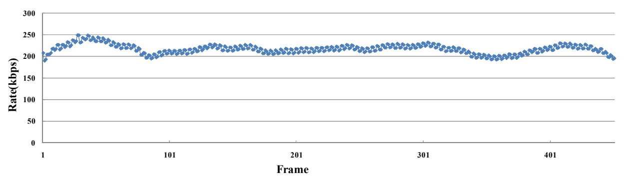

(a) FourPeople ( 1280x720)



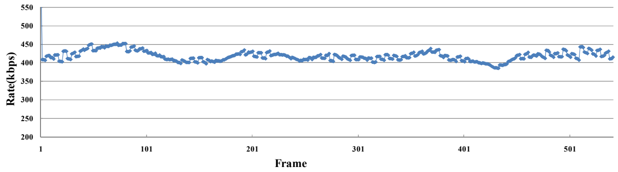

(b) BachetballPass (416x240)

**Fig. 3.** Quality fluctuation curves of the proposed algorithm

Table 2 states the performance comparisons for the proposed algorithm with HM anchor and the state-of-art rate control R-λ algorithm. It can be observed that the proposed algorithm works well for both high resolution and low resoution sequences. From the tables, we can conclude the proposed algorithm presents much better encoding performance. Compared with R-λ algorithm [17], the BD_Rate and BD_PSNR gain can be over 5% and 0.18dB on average. Table 3 shows the bit rate mismatch ratio comparisons. It is shown that the proposed algorithm generates smaller mismatch between target bit rate and actual bit rate. Fig. 3 and Fig. 4 shows the flunctuation curves of quality and bit rate of *FourPeople* and *BascketballPass*. It can be seen the proposed shceme also provides smooth video quality and bit rate.




(a) BascketballPass (416x240) Target bit rate = 216kbps




(b) FourPeople (1280X720) Target bit rate = 435kbps

**Fig. 4.** Bit rate fluctuation curves of the proposed algorithm

**Table 2.** Performance comparison for the proposed algorithm with HM anchor and R-λ[17] for LB-main and LP-Main coding configuration testing

| Sequences | LB-main | | | |
|---|---|---|---|---|
| | Gain over HM8.0 | | Gain over R-λ[17] | |
| | *BD-Rate* | *BD-PSNR(dB)* | *BD-Rate* | *BD-PSNR(dB)* |
| ClassB | 3.5% | -0.07 | -4.0% | 0.07 |
| ClassC | -1.4% | 0.09 | -3.9% | 0.20 |
| ClassD | -2.0% | 0.08 | -7.5% | 0.32 |
| ClassE | -3.5% | 0.11 | -5.2% | 0.16 |
| Avg | -0.9% | 0.05 | -5.2% | 0.19 |
| Sequences | LP-main | | | |
| | Gain over HM8.0 | | Gain over R-λ[17] | |
| | *BD-Rate* | *BD-PSNR(dB)* | *BD-Rate* | *BD-PSNR(dB)* |
| ClassB | 2.0% | -0.05 | -4.6% | 0.10 |
| ClassC | -2.4% | 0.09 | -4.8% | 0.18 |
| ClassD | -1.6% | 0.07 | -6.7% | 0.25 |
| ClassE | -3.1% | 0.11 | -4.6% | 0.14 |
| Avg | -2.3% | 0.06 | -5.2% | 0.17 |

**Table 3.** the rate mismatch of the proposed algorithm for LB-main and LP-main configuration

| Sequences | Rate Mismatch | |
|---|---|---|
| | LB-main | LP-main |
| ClassB | 0.31% | 0.22% |
| ClassC | 0.15% | 0.68% |
| ClassD | 0.30% | 0.28% |
| ClassE | 0.10% | 0.09% |
| Avg | 0.22% | 0.32% |

## 5. CONCLUSIONS AND FUTURE WORK

Based on the new coding technologies of HEVC, a novel quadratic ρ-domain rate model is proposed and a frame layer rate control algorithm is studied in our paper. The novelty of this paper lies in that we employ the ρ-domain rate model and hierarchical bits allocation scheme to get a better rate control performance. The efficiency of the proposed algorithm can be verified by the experimental results. Moreover, due to the simple rate model and parameters updating scheme, almost no extra computational complexity is added for the HM platform. We focused on the Low Delay configuration in this work and will continue to investigate rate control algorithms in Random Access setting in future.

## 6. REFERENCES


[1] 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.

[2] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification, ITU-T Rec. H.264 and ISO/IEC 14 496-10 AVC, Joint Video Team, Mar. 2003.

[3] I. K. Kim, J. Min, T. Lee, W. J. Han, and J. Park, "Block Partitioning Structure in the HEVC Standard", *IEEE Trans. Circuits Syst. Video Technol.*, in publication

[4] R. Sjöberg, J. Samuelsson, Absolute signaling of reference pictures, Joint Collaborative Team on Video Coding (JCT-VC)of ITU-T of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG116th Meeting: Torino, 2011.

[5] Test Model 5. [Online]. Available: http://www.mpeg.org /MPEG / MSSG / tm5.

[6] T. Chiang and Y. Zhang, "A new rate control scheme using quadraticrate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7,no. 2, pp. 287–311, Apr. 1997.

[7] J. Corbera and S. Lei, "Rate Control for Low-Delay Video Communications," ITU Study Group 16, Video Coding Experts Group, Portland, Documents Q15-A-20, 1997.

[8] M. Hamdi, J. W. Roberts and P. Rolin, "Rate control for VBR video coders in broad-band networks," *IEEE Journal on Selected Areas in Communications*, Vol. 15, No. 6 Aug. 1997.

[9] P. H. Westerink, R. Rajagopalan and C. A. Gonzales, "Two-pass MPEG-2 variable-bit-rate encoding," *IBM J.RES. DEVELOP*, Vol. 43 No.4 Jul. 1999.

[10] D. Zhang, K. N. Ngan, and Z. Chen, "A two-pass rate control algorithm for H.264/AVC high definition video coding," *Signal Processing: Image Communication*, vol. 24, no. 5, pp. 357--367, May. 2009.

[11] S. Ma, W. Gao and Y. Lu, "Rate-distortion analysis for H.264/AVC video coding and its application to rate control," IEEE Trans. Circuits Syst. Video Technol., vol. 15, no. 12, pp. 1533-1544, Dec. 2005.

[12] JVT-G012 Z. G. LI, F. Pan, K. P. Lim and G. N. Feng. Adaptive Basic Unit Layer Rate Control for JVT. Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-G012, 7th meeting, Pattaya, Thailand, 7-14(2003).

[13] M. Jiang and N. Ling, "On enhancing H.264/AVC video rate control by PSNR-based frame complexity estimation," *IEEE Trans. Consum. Electron.*, vol. 51, no. 1, pp. 281-286, Feb. 2005.

[14] D. Kwon, M. Shen and C. C. Jay Kuo, "Rate control for H.264 video with enhanced rate and distortion models," *IEEE Trans. Circuits Syst. Video Techn. ,17*(5): 517-529, 2007.

[15] Z. He and S. Mitra, "Optimum bit allocation and accurate rate control for video coding via -domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 10, pp. 840–849, Oct. 2002.

[16] 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.

[17] 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.

[18] Bumshik, M. Kim, "Modeling Rates and Distortion Based on a Mixture of Laplacian Distribution for Inter-Predicted Residuals in Quadtree Coding of HEVC," *IEEE signal processing letters*, vol. 18, no. 10, October 2011.

[19] G. Bjontegaard, "Calculation of average PSNR difference between RDcurves,"in *Proc. ITU-T Q.6/SG16 VCEG 13th Meeting*, Austin, TX,Apr. 2001, document VCEG-M33.