An Adaptive Frame Complexity Based Rate Quantization Model for Intra-Frame Rate Control of High Efficiency Video Coding (HEVC)

Lin SUN, Oscar C.Au, Wei Dai, Yuanfang Guo, and Ruobing Zou
Department of Electronic and Computer Engineering
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
Email: {lsunece, eeau, weidai, eeandyguo, zouruobing }@ust.hk

***Abstract*—An efficient and accurate R-Q model is greatly important for intra-frame rate control of the latest High Efficiency Video Coding (HEVC) standard. However, previous methods pay more attention to the gradient based rate quantization (R-Q) model for the intra-frame rate control. In this paper, we analyze the drawbacks of the gradient based frame complexity measure when applied different Quantization Parameters (QPs). Then we propose a novel edge based frame complexity measure using the Gaussian Gradient operator with properly selected parameters. In order to tackle the problems that the gradient based rate quantization model fails when using the different QPs, we propose an adaptive frame complexity based R-Q model for intra-frame rate control based on these two complexity measures. Simulations have been conducted based on HM6.2 which is the latest reference software of HEVC. Note that we may be the first to do this work, so we do not have the classical methods which have been implemented in HEVC to compare with. In this paper we implement the traditional gradient based and the Cauchy distribution based rate quantization model in HEVC and compare the performance with each other. Here we use bit rate mismatch ratio as the evaluation method. The simulation results show that by using our proposed scheme, up to 33.1% mismatch ratio reduction compared with the Cauchy distribution based model and up to 13% mismatch ratio reduction compared with the gradient based model for intra frames can be achieved.**

I. Introduction

The latest state-of-the-art video coding standard is H.264/AVC [1] developed by ITU-T and ISO/IEC/MPEG. Recently, ITU-T and ISO/IEC/MPEG have formed a Joint Collaborative Team on Video Coding (JCT-VC) on next generation video coding standard called High Efficiency Video Coding (HEVC) [2]. The HEVC test model (HM) still belongs to block-based hybrid video coding framework, except that HEVC introduces the new concept of CU, PU and TU. The overall coding structure is characterized by the various sizes of CU, PU and TU in a recursive manner, once the size of largest coding unit (LCU) and the hierarchical depth of CU are defined. HEVC can achieve higher coding efficiency than the state-of-the-art H.264, mainly due to the fact that the HEVC encoder employ more complicated and accurate approaches in the coding procedure. However, there is still no mature rate control algorithm for HEVC now, particularly for intra frames. Furthermore, HEVC is still facing the same problem as H.264/AVC that is the typical chicken and egg dilemma [3]. Li et al. in JVT-G012 [3] have proposed an adaptive rate control framework for H.264/AVC. In general, for P frames, a single-pass rate control method based on the classic quadratic rate quantization (R-Q) model is used, and a linear model for mean absolute difference (MAD) prediction is employed to solve the above dilemma. However, all the scheme focus on the inter frame rate control, no explicit R-Q model for an intra frame is discussed. In fact, the Quantization Parameter (QP) for an intra frame depends on the average QP of P-frames in the previous group of pictures (GOP). Since the bit rate of the different intra frames may change significantly when a fixed QP is used. Also usually the bits generated by intra coded frames are as many as more than 3 times than the bits generated by inter coded frames. It will cause the buffer overflow or underflow problem if not well controlled. Thus an accurate R-Q model is necessary for estimating intra bit rate and calculating the QP correspondingly. Although some R-Q models [4] [5] [6] [7] are proposed which can adaptively determinate QP for intra frames, such as the Cauchy distribution based R-Q model which was proposed by Kamaci et al. [5] by assuming the distribution of the coefficients is Cauchy distribution. However, these methods did not consider the complexity of the frame itself which is obviously highly related the final bit rate. To tackle this problem, Jing et al. [6] [7] proposed one rate control method with the consideration of the frame gradient. The proposed algorithm is based on the observation that the output bit rate of each intra frame is highly correlated with its gradient value. Actually, this phenomenon has already reported in [8] in which the authors dealt with the still images. According to our observation, the linear correlation between gradient-based frame complexity measure and the intra frame bit rate is high only when particular QP is used. On the other hand, intra prediction plays an important role for compressing intra frames and when it will generate more bits intra prediction is not so accurate. What is more, the intra prediction is not so accurate when the edge exists in the block [9]. So in this paper we propose to use the edge based complexity measure as an additional measure for the intra rate quantization model. Herein we use the Gaussian Gradient to detect the frame edge and calculate Edge Pixel Ratio (EPR), that is the number of edge pixels versus the total number pixels in whole frame. We also verified that it exists a strong linear relationship between the bit rate (evaluated by bit per pixel (bpp)) and EPR. More importantly, it can achieve better performance when the gradient based complexity fails. Based on this observation, we propose an adaptive complexity measure which combine both complexity measures. We apply proposed algorithm to the intra frames rate control in HEVC and obtain a better performance compared with the Gradient based and Cauchy-Distribution based method.

The rest of this paper is organized as follows: Section II gives a brief analysis of the two complexity measure methods-gradient based complexity measure and edge based complexity measure. In this part we also propose to use the Gaussian Gradient as the edge detection operator and give the details of the edge based complexity measure. Section III proposes our adaptive complexity measure. Simulation results of our proposed method are shown in Section IV. Section V contains our conclusions.

II. Frame Complexity Measures

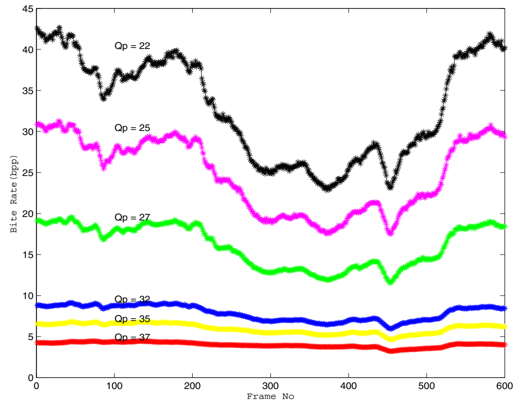
We exhaustively encode various test sequences only using the intra mode based on HM6.2 [10] using different QPs. Fig.1(a) shows the bit rate results of the different frames of BQMall sequence whose resolution is 832×480. We can get two major observations from this figure. Firstly, encoding the different frames will generate different bit rates even using the same QP. So the encoding complexity of different frames is different. Also that the shape of the generated curves is similar presents the high correlation between the bit rate and the frame property. Secondly, the larger the QP is, the less bit rate difference between the frames will be observed. Thus we achieve two conclusions from these observations. On one hand, it should exist some measures which can describe the characteristic of the frame complexity. On the other hand, using single one measure to describe the characteristic of the sequence under different QPs may not be proper. In the following part we will give two complexity measures and based on these two methods we propose an adaptive complexity measure for intra bit rate control.

A. Gradient Based Complexity Measure for Intra Frame

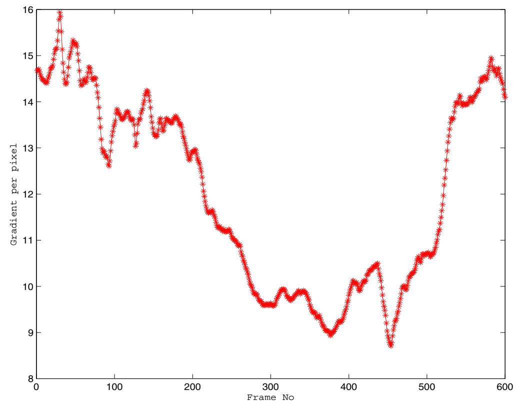
It is believed that gradient value per pixel of each frame has a great linear relationship with the bit rate of this intra coded frames. This method is currently used in the rate estimation problem for the intra frame coding in H.264/AVC. The calculation of gradient per pixel is defined by [8]

$$Grad = \frac{1}{N_1 \times N_2} \sum_{i}^{N_1-1} \sum_{j}^{N_2-1} (\sqrt{|I_{i,j} - I_{i+1,j}|} + \sqrt{|I_{i,j} - I_{i,j+1}|}) \tag{1}$$

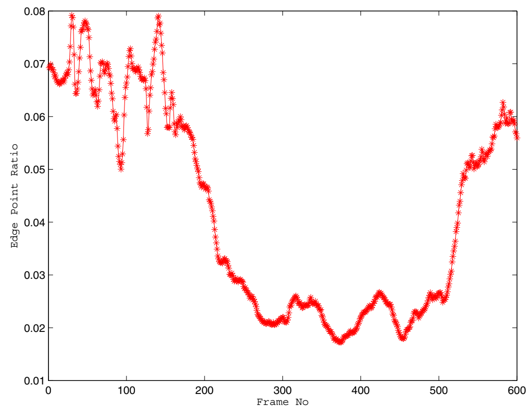
where $I_{i,j}$ presents the luminance value at the corresponding position (i, j). N_1 and N_2 are the horizontal and vertical dimensions of one frame, respectively.



(a)



(b)



(c)

Fig. 1: (a) Intra bit rate of the BQMall under different QPs (b) Gradient per pixel for different frames of BQMall (c) Edge pixel ratio for different frames of BQMall

In practice, in order not to increase the calculation complexity so much, the gradient based complexity measure is simplified as indicated in [6] [7]

$$Grad = \frac{1}{N_1 \times N_2} \sum_{i}^{N_1-1} \sum_{j}^{N_2-1} (|I_{i,j} - I_{i+1,j}| + |I_{i,j} - I_{i,j+1}|) \tag{2}$$

Fig.1(b) shows the gradient value of different frames of the BQMall sequence. From this figure we can see that the overall trend of the gradient is similar as the bit rate generated by intra coding.

B. Proposed Edge Based Complexity Measure for Intra Frame

According to previous statement, the intra coded bits should have a strong relationship with the EPR. In this work, first we should select an edge extraction operation to detect the edge. This operator normally return a binary image $B(i,j)$ of the same size as the original image, with $1'$s where the functions find edges and $0'$s elsewhere. Once the edges are detected, we use the ratio of the edge points as a complexity measure, defined by

$$EPR = \frac{1}{N_1 \times N_2} \sum_{i}^{N_1-1} \sum_{j}^{N_2-1} B(i,j) \times 100\% \tag{3}$$

Fig.1(c) shows the EPR value of the different frames of BQMall sequence. It is slightly different from the gradient value curve and importantly, it also shows the overall trend of the bit rate of the whole sequence.

C. Gaussian Gradient

Here we should apply the edge detection first, Gaussian Gradient [11] is used as our operator. One advantage of using Gaussian Gradient is that even the original frame exists some noise it can detect the real or nearly real edge and texture region. Considering a single row or column of one image signal, we plot the intensity of the signal as a function of the position.

Fig.2(a) shows 1-D signal *f(x)* with noise. The first order derivative of the signal $\frac{d}{dx}f(x)$ is used as the edge detection method, terribly getting Fig.2(b). Herein we can not point out where is edge for the edge is already corrupted by the noise. However, if we smooth the image first using the popular gaussian kernel *h(x)*, *f(x)*h(x)* is shown in Fig.2(c). Then we also apply the first order derivative of the smoothed signal $\frac{d}{dx}(f(x) * h(x))$, getting Fig.2(d). At this time we can see clearly that the edge is the highest magnitude of the output signal. The first order derivative of the smoothed signal is what we call Gaussian Gradient.

Here we use the Gaussian Gradient as our operator to detect the edge.

$$E(x,y) = \frac{\partial}{\partial x \partial y}(Gauss(x,y,\sigma) * I(x,y)) \tag{4}$$

where $E(x,y)$ is the edge image at the corresponding position (x,y) and $Gauss(x,y,\sigma)$ is gaussian kernel with the parameter σ. $I(x,y)$ is the luminance value at the position (x,y). In this step, a threshold TH should be used to determine which point can be considered as a edge point in $E(x,y)$. When $|E(x,y)| > TH$, we consider it as the edge point and vice versa. We select the best TH based on maximize the Pearson Correlation Coefficient between the bit rate and the EPR. Also in order to keep balance between restoring details and reacting to the noise, we set the gaussian kernel parameter σ to 0.5.

III. Proposed Adaptive Frame Complexity Measure Based R-Q Model

A. Adaptive Frame Complexity Measure

To evaluate the accuracy of the complexity measure, the quantity R^2 is used to measure the degree of the data variation from a given model

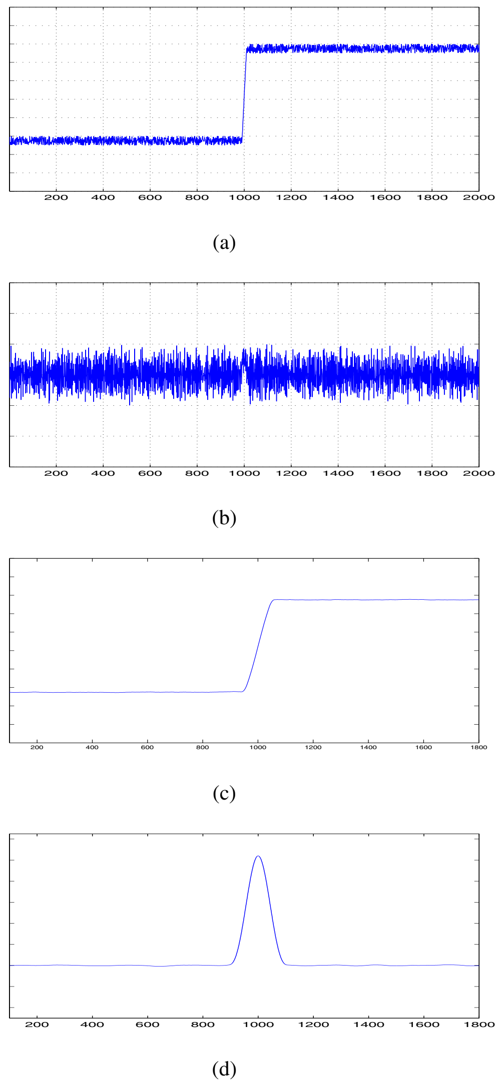
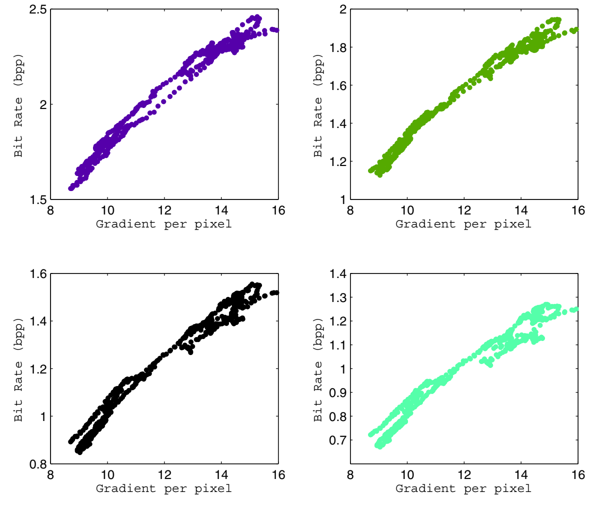


Fig. 2: The example of detection of 1-D noisy edge signal. (a) 1-D noisy edge signal (b) First order derivative of 1-D noisy edge signal (c) Gaussian filter filtered 1-D noisy edge signal (d) First order derivative of smoothed noisy edge signal

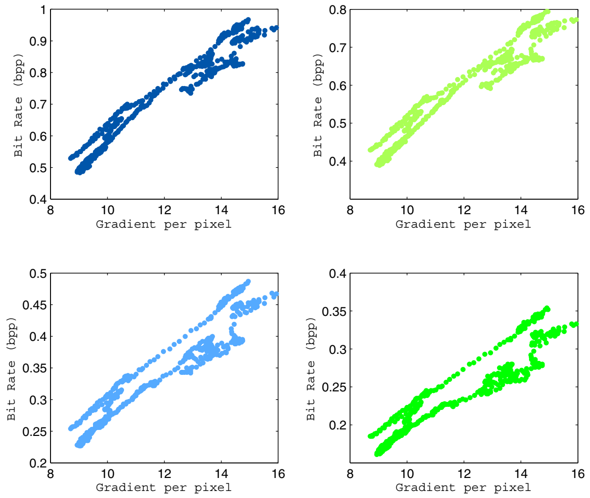
$$R^2 = 1 - \frac{\sum_i (X_i - \hat{X}_i)^2}{\sum_i (X_i - \bar{X}_i)^2} \quad (5)$$

where X_i and $\hat{X}_i$ are the real and estimated values of one data point i. $\bar{X}_i$ is the mean of all the data points. The closer the value of R^2 to 1, the more accurate is the model.

Here we first evaluate the accurate of the linear relationship between the bpp and the Gradient per pixel. Where bpp is defined as



(a)



(b)

Fig. 3: (a) Scatter plots of the bit rate (bpp) versus gradient per pixel value for BQMall sequence. Corresponding QPs (from top to bottom, left to right) are: 18, 20, 22, 25 (b) Scatter plots of the bit rate (bpp) versus gradient per pixel value for BQMall sequence. Corresponding QPs (from top to bottom, left to right) are: 27, 32, 35, 37

TABLE I: Performance Comparison of Different Complexity Models for Intra Bit Rate of BQMall sequence with resolution 832×480

Quantization Parameter (QP)	Gradient (R^2)	EPR(R^2)	Adaptive (R^2)
QP = 18	0.9759	0.9676	0.9860
QP = 20	0.9793	0.9555	0.9901
QP = 22	0.9701	0.9363	0.9857
QP = 25	0.9575	0.9601	0.9735
QP = 27	0.9450	0.9527	0.9701
QP = 32	0.9375	0.8928	0.9540
QP = 35	0.9066	0.8602	0.9201
QP = 37	0.8703	0.8291	0.9003

$$bpp = \frac{bit\_rate}{f \times N_{pic}} \quad (6)$$

where f is the frame rate and N_{pic} is the number of pixels in one frame.

As indicated in Table I, the accuracy of the linear relationship between bit rate and gradient fluctuates when the QP changes. This can also be seen clearly from the figures in Fig.3. From the figure we can see visually that when QP equals to 20, it shows the best linear relationship while in other QPs this linear relationship becomes weak. This phenomenon also exists in the relationship between bit rate and EPR. However, from Table I we can see that for some QPs the EPR shows better relationship while verse vice versa. Also the EPR signals the intra prediction accuracy, so how about combining these two complexity measures together? Then an adaptive complexity measure is proposed

$$Complexity = \alpha Grad + \beta EPR \quad (7)$$

where Grad is the gradient based complexity measure and EPR is edge based complexity measure as previously discussed. α and β is the weighted coefficients and $\alpha + \beta = 1$. And α and β is selected adaptively which will be discussed in the later section.

The R^2 value of this adaptive complexity measure also shown in Table I. We can see the proposed adaptive method presents the best linear relationship in all different QPs.

B. Adaptive Complexity Based Intra Frame R-Q Model

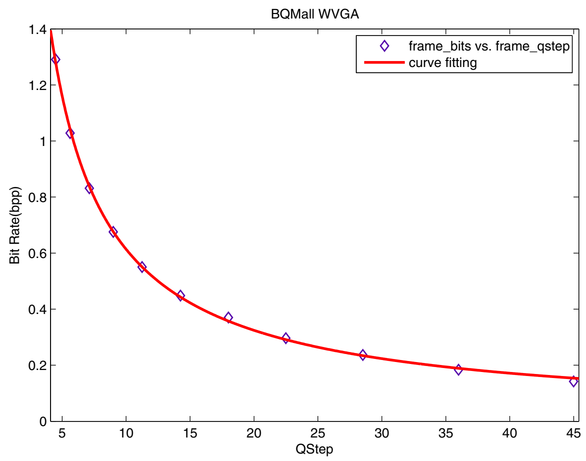
We encode one frame of the sequence using the QP frame 18 to 38 with the step 2 and plot the quantization step size (Q_{step}) vs. actual bit rate (bpp), as shown in Fig.4. We find that the bit rate and Q_{step} has the following relationship

$$R(Q_{step}) = a \cdot Q_{step}^b \quad (8)$$

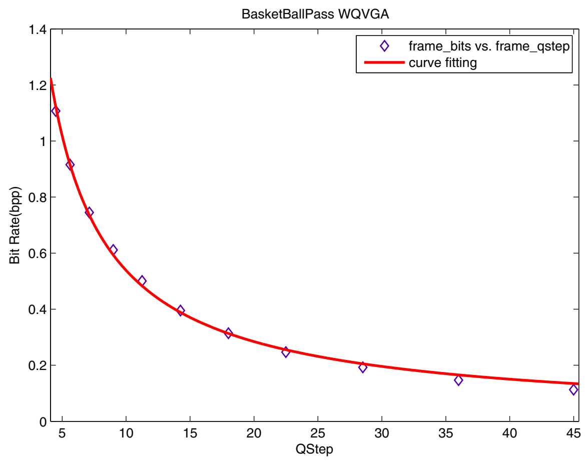
a and b is parameter that depend on the video content. According to our experimental results when b is in the range of (-0.88 -0.98), we can get good performance. So for simplicity we set b to -0.92 and change a from frame to frame for all the test sequences. Here we substitute a using our proposed adaptive complexity measure and propose a R-Q model as follows:

TABLE II: Comparison of Average Mismatch Ratio with Different Initial QP

Sequence	QP_0 = 18			QP_0 = 24			QP_0 =38		
BasketballPass WQVGA	Cauchy Distribution	Proposed	$\delta M(\%)$	Cauchy Distribution	Proposed	$\delta M(\%)$	Cauchy Distribution	Proposed	$\delta M(\%)$
	5.25	3.63	33.1	5.14	3.47	29.5	5.91	5.01	16.0
	Gradient Based	Proposed	$\delta M(\%)$	Gradient Based	Proposed	$\delta M(\%)$	Gradient Based	Proposed	$\delta M(\%)$
	4.04	3.63	10.1	3.99	3.47	13	5.61	5.01	10.9
BlowingBubbles WQVGA	Cauchy Distribution	Proposed	$\delta M(\%)$	Cauchy Distribution	Proposed	$\delta M(\%)$	Cauchy Distribution	Proposed	$\delta M(\%)$
	5.23	4.64	11.28	5.38	4.87	9.48	4.46	4.10	8.07
	Gradient Based	Proposed	$\delta M(\%)$	Gradient Based	Proposed	$\delta M(\%)$	Gradient Based	Proposed	$\delta M(\%)$
	5.04	4.64	7.94	5.34	4.87	8.80	4.35	4.10	5.75
BQMall WVGA	Cauchy Distribution	Proposed	$\delta M(\%)$	Cauchy Distribution	Proposed	$\delta M(\%)$	Cauchy Distribution	Proposed	$\delta M(\%)$
	4.10	3.89	5.12	4.40	3.95	10.23	3.93	3.75	4.58
	Gradient Based	Proposed	$\delta M(\%)$	Gradient Based	Proposed	$\delta M(\%)$	Gradient Based	Proposed	$\delta M(\%)$
	4.01	3.89	2.99	4.33	3.95	8.78	4.04	3.73	7.18
BasketballDrill WVGA	Cauchy Distribution	Proposed	$\delta M(\%)$	Cauchy Distribution	Proposed	$\delta M(\%)$	Cauchy Distribution	Proposed	$\delta M(\%)$
	4.44	3.91	11.94	4.73	4.40	6.98	3.93	3.76	4.33
	Gradient Based	Proposed	$\delta M(\%)$	Gradient Based	Proposed	$\delta M(\%)$	Gradient Based	Proposed	$\delta M(\%)$
	4.06	3.91	3.69	4.69	4.40	6.18	3.77	3.76	0.27



(a)



(b)

Fig. 4: (a)Scatter plot of the actual bit rate vs. QStep and the curve fitting using $a * QStep^b$, herein a = 5.10, b = -0.92 for BQMall (b) Scatter plot of the actual bit rate vs. QStep and the curve fitting using $a * QStep^b$, herein a = 4.49, b = -0.92 for BasketBallPass

$$R(Q_{step}) = Complexity \cdot Q_{step}^b \tag{9}$$
$$R(Q_{step}) = (\alpha Grad + \beta EPR) \cdot Q_{step}^b \tag{10}$$

where Complexity is the weighted average complexity value per pixel calculated by Eq.7, α and β is the parameters which is updated after encoding one frame, also they obey $\alpha + \beta = 1$. And we update the weights α using

$$\alpha_{k+1} = \gamma \cdot \alpha_k + (1-\gamma) \cdot \frac{\frac{R_k}{(Q_{step})_k^b} - EPR_k}{Grad_k - EPR_k} \tag{11}$$

where R_k, $Grad_k$, EPR_k and $(Q_{step})_k$ are the actual output bit rate (bpp), the gradient value per pixel, EPR and the quantization step size for the kth intra frame. α_k is the previous weighted value. After obtaining α, we can calculate $\beta_{k+1} = 1 - \alpha_{k+1}$. As the first QP is determined manually, α_0 and β_0 can be obtained as follows:

$$\alpha_0 = \frac{\frac{R_0}{(Q_{step})_0^b} - EPR_0}{Grad_0 - EPR_0} \tag{12}$$
$$\beta_0 = 1 - \alpha_0 \tag{13}$$

Also γ is a forgetting parameter with the typical value of 0.5.

IV. Experiment Results

To evaluate the performance of our proposed intra frame R-Q model, we have applied our proposed scheme, the gradient based method and the cauchy distribution based method to the HEVC reference software HM6.2 [10] and encoded a number of video sequences only using intra high efficiency coding mode. All the work satisfies the test condition [12]. In our experiments, the first frame of a sequence is coded with an initial QP_0 and the QP values of the following frames are calculated by our R-Q model. The encoded frame number is 100 and the QP is selected from small to large, such as 18, 24, 38. In order to measure the accuracy of the model, we define the frame bit rate mismatch ratio as

$$M = \frac{|R_{target} - R_{actual}|}{R_{target}} \cdot 100\% \tag{14}$$

where R_{target} and R_{actual} is the bit rate of the first frame and one following frame, respectively. And the average value of M is

$$\bar{M} = \frac{\sum_{i=0}^{N-1} M_i}{N} \tag{15}$$

$$\delta M = \frac{|\bar{M} - \bar{M}_{proposed}|}{\bar{M}} \cdot 100\% \tag{16}$$

where N is the number of the frames to be encoded and δM shows the mismatch reduction compared between the existing method and the proposed method. Table II shows the value of $\bar{M}$ for different sequence with the different initial QPs. From this table we can see that the bit rate estimation has been largely improved by our proposed scheme. Importantly, it is effective in both large OP and small QP value.

V. Conclusion

In this paper we propose an adaptive complexity measure based R-Q model for intra frames' bit rate control. Because the gradient based complexity measure is not so accurate when the QP changes, it will be not suitable for the bit rate estimation in this situation. On the other hand, the edge point ratio shows its better performance when the gradient based fails. When employing edge detection, we propose to use the Gaussian Gradient to balance the false detection and detail restoration. In implementation, we fix the gaussian parameter to balance the complexity and the smooth degree and select the threshold. Finally we propose the adaptive complexity measure based rate quantization model. The simulation results of the selected sequences based on the HM6.2 which is the latest reference software of HEVC show its effect in bit rate control of the intra frames. The proposed scheme can reduce the bit mismatch ratio by 33.1% compared with cauchy distribution based scheme and 13% compared with the gradient based method.

VI. Acknowledgement

This work has been supported in part by the Research Grants Council (RGC) of the Hong Kong Special Administrative Region, China. (GRF Project no. 610210)

References

[1] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the h.264/avc video coding standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 560–576, Jul 2003.

[2] I. K. K. McCann, W.J. Han, "Samsung's Response to the Call for the Proposals on Video Compression Technology," in *Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-A124*, Apr 2010.

[3] K. P. L. G. N. F. X. L. Z. G. Li, F. Pan and S. Rahardaj, "Adaptive Basic Unit Layer Rate Control for JVT," in *Joint Collaborative Team ISO/IEC JTC1/SC29/WG11 and ITU-T SG16/Q.6, JCT-G012*, 2003.

[4] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 7, no. 1, pp. 246–250, Feb 1997.

[5] N. Kamaci, S. Member, Y. Altunbasak, S. Member, and R. M. Mersereau, "Frame bit allocation for the h.264/avc video coder via cauchy-density-based rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol*, vol. 15, pp. 994–1006, 2005.

[6] X. Jing, L.-P. Chau, and W.-C. Siu, "Frame complexity-based rate-quantization model for h.264/avc intraframe rate control," *Signal Processing Letters, IEEE*, vol. 15, pp. 373–376, 2008.

[7] X. Jing and L.-P. Chau, "A novel intra-rate estimation method for h.264 rate control," in *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on*, May 2006.

[8] W. J. Kim, J. W. Yi, and S. D. Kim, "A bit allocation method based on picture activity for still image coding," *Image Processing, IEEE Transactions on*, vol. 8, no. 7, pp. 974–977, Jul 1999.

[9] F. Zou, O. Au, J. Dai, C. Pang, W. Yang, X. Wen, and Y. Liu, "Edge-based adaptive directional intra prediction," in *Picture Coding Symposium (PCS), 2010*, Dec. 2010, pp. 366–369.

[10] *Test Model 6.[Online].Available:http://www.mpeg.org/MPEG/MSSG/tm6.2*.

[11] L. Sun, O. Au, R. Zou, W. Dai, X. Wen, S. Li, and J. Li, "Adaptive bilateral filter considering local characteristics," in *Image and Graphics (ICIG), 2011 Sixth International Conference on*, Aug. 2011, pp. 187–192.

[12] F. Bossen, "Common test conditions and software reference configurations," in *Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-B300*, Jul 2010.