A STUDY ON THE RATE DISTORTION MODELING FOR HIGH EFFICIENCY VIDEO CODING

Siwei Ma[1],*Junjun Si*[1], *Shanshe Wang*[2]
swma@pku.edu.cn, {jjsi,sswang}@jdl.ac.cn
[1]Institute of Digital Media, Peking University
[2] Department of Computer Science, Harbin institute of Technology

ABSTRACT

This paper gives a study on the rate distortion modeling for High Efficiency Video Coding (HEVC) and new rate distortion models are proposed by considering the coding framework change of HEVC. Although HEVC is also characterized by the traditional prediction/transform hybrid coding framework with more improved coding tools, e.g. large size coding unit (CU), prediction unit (PU) and transform unit (TU), two new coding modules are added as sample adaptive offset(SAO) and adaptive loop filter(ALF). SAO and ALF are added after the motion compensation reconstruction, which would reduce the distortion of the reconstructed image. With considerations on these changes, new rate distortion models are proposed. For rate modeling, this paper proposes a separate model for header bits and coefficient bits respectively. For distortion modeling, the quantization distortion and the distortion reduction induced by SAO/ALF are modeled jointly. The experimental results show that the coding rate and distortion can be estimated by the proposed model accurately. These models can be used for rate distortion cost estimation for mode decision or rate control in the future.

Index Terms— HEVC, rate distortion model, Laplacian distribution

1. INTRODUCTION

High Efficiency Video Coding (HEVC) is emerging as a new video coding standard with significant coding improvement over the preceding standards, such as H.264/AVC. HEVC is also characterized by the traditional prediction/transform hybrid coding framework, as shown in Fig. 1[1]. In the framework, HEVC has adopted more improved coding tools, e.g. large size coding unit (CU), prediction unit (PU) and transform unit (TU) etc. CU, the basic coding unit similar to macroblock in H.264/AVC, can have various sizes and allows recursive quad-tree splitting. Given the size of Largest Coding Unit (LCU) and the maximum hierarchical depth, CU can be expressed in a recursive quad-tree representation adapted to the picture content. Once the splitting of CU hierarchical tree is finished, the leaf node CUs can be further split into PUs. PU is the basic unit for prediction and it allows multiple different shapes to encode irregular image patterns. PU size is limited to that of CU with square or rectangular shape. Besides CU and PU, TU is also defined to represent the basic unit for transform. The size of TU cannot exceed that of CU, and it is independent with PU size for inter coding.

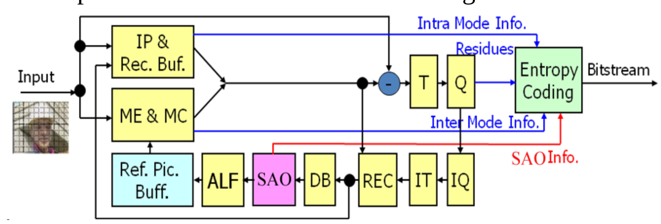


Fig. 1. HEVC Encoder Framework, IP: Intra prediction, T: Transform, Q: Quantization, IT: Inverse Transform, IQ: Inverse Quantization, REC: Reconstruction, DB: Deblock-filter, ME: Motion Estimation, MC: Motion Compensation

Besides these improved coding tools, two new coding modules are added to the coding framework. They are sample adaptive offset (SAO) and adaptive loop filter (ALF). SAO and ALF are added after the motion compensation reconstruction, which is used to reduce the compression distortion of reconstructed image by adding an offset or an adaptive filter processing on the reconstructed pixels. This paper gives a study on the distortion reduction brought by SAO and ALF. As de-block only works on the edge pixels, and the rate-distortion (R-D) model for H.264/AVC considers this part already, the effect of de-block filter is ignored in the paper.

R-D modeling is an important part in video optimization coding researches, such as rate control and fast mode decision etc., and many related researches have been done on it. The rate distortion models are usually derived theoretically or by experimental verification. The traditional rate distortion model include the models proposed in TM5 [2], TMN8 [3] and rho-domain model [4] etc. In TM5 a simple linear rate-distortion model, i.e. $R(QP)=X/QP$, is employed. In TMN8 and VM8, the more accurate R-D model is proposed, i.e. $R_i=A*(K\sigma^2/QP^2_i+\underline{C})$ in TMN8 [3]. In VM8, the rate model is derived as $R = X_1\times MAD_i/QP_i + X_2\times MAD^2_i/QP^2_i$ which is approximated from a Taloy expansion for R-D function of Gaussian distribution [4]. In

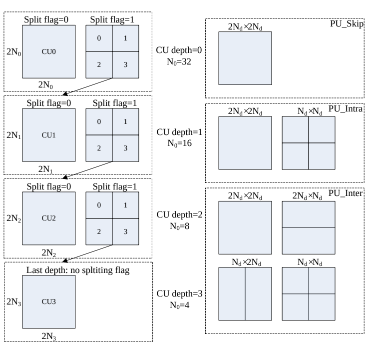
[5], the relation between rate and quantization parameter is indirectly represented with the relation between the rate and ρ, where ρ indicates the percentage of zero coefficients after quantization.

The rate distortion modeling problem for HEVC is also studied in [6]. In [6], the coefficient rate and quantization distortion model is derived based on the assumption of laplacian distribution for transform coefficients. In this paper the rate distortion model for HEVC is studied from the coding framework viewpoint by experimental methods. For rate modeling, the header bits and coefficient bits are modeled independently. For distortion modeling, with the joint consideration on quantization distortion, a new distortion model is proposed for HEVC. The experimental results show that the proposed distortion model can be used to estimate the coding distortion accurately.

The rest of the paper is organized as follows: Section 2 and Section 3 provide the derivation of rate and distortion model for HEVC respectively. Section 4 gives more experimental results for the derived distortion model and Section 5 concludes the paper.

2. RATE MODELING FOR HEVC

Considering the characteristics of HEVC, we model the rate in two parts: header bits for coding CU, PU and the coefficients bits for TU. As said in Section 1, a recursive quad-tree structure is used in HEVC, as shown in Fig. 2. To represent the coding structure, skip flag and prediction type are coded at CU header, and merge index, intra prediction mode, motion vector and reference index etc. are coded into PU header.

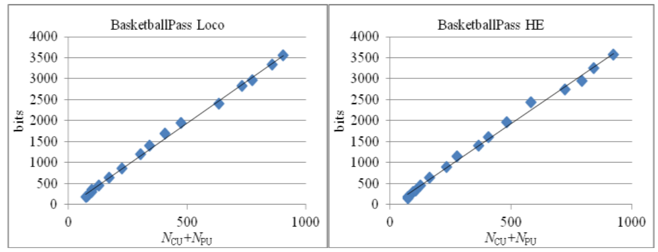


(a) (b)

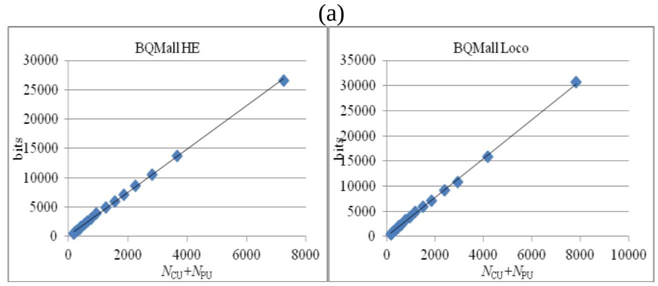
Fig. 2. (a) Maximum possible recursive CU structure in HM. (LCU size= 64, maximum hierarchical depth = 4), (b) Possible PU splitting for skip, intra and inter in HM.

In [7], the header bits for inter marcroblock are modeled with the number of nonzero motion vector elements and motion vectors. Inspired by [7], in our experiments it was found that the header bits have very good linear relationship with the number of CU and PU partitions, as shown in Fig. 3. The x-axis is the sum of CU and PU number. The y-axis is the total header bits. From the curve, it can be seen that the header bits R_h can be modeled as:

$$R_h = l*(N_{CU} + N_{PU})$$



(a)



(b)

Fig. 3. The relationsship between the heder bits R_h and the total CU number NCU and PU number NPU (a) basketballPass (b) BQMall

N_{CU}, N_{PU} is the number of coding unit and prediction unit respecitively. The linear model is also verified on more test sequences. The average goodness of fitting R^2 for HEVC test sequences at different resolution and frame rate is over 0.994.

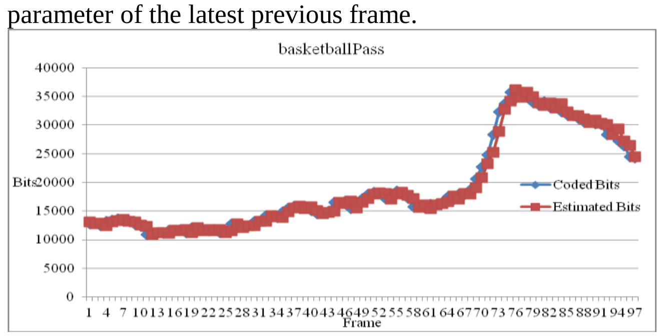
For HEVC coefficient modeling, [6] has given some analysis based on laplacian distribution assumption of transform coefficients, yet the model is complex for real rate estimation. Different from [6], we give a study on the traditional model used in TM5, shown as:

$$R_c = X / QP \quad (1)$$

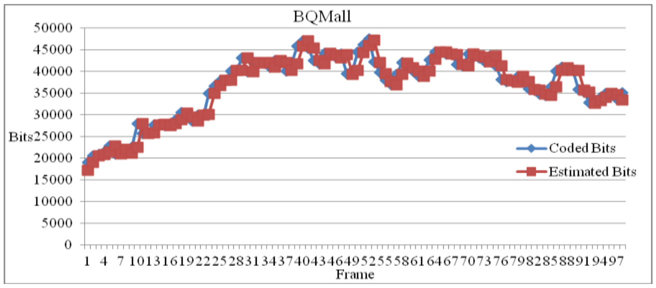
R_c is the coefficients bits. QP is quantization parameter. As the ABR rate control used in X264, X is computed as,

$$X = \left(\sum_{i=0}^{n} SATD_i / \sum_{i=0}^{n-1} SATD_i \right)^{1-qcom} \times R_{n-1} \times QP_{n-1} \quad (2)$$

i is the frame number and $SATD_i$ is the $SATD$ of the ith frame. n is the current frame number. $qcom$ is a constant, the recommended value is 0.6. R_{n-1} is the generated source bits of the latest previous frame. QP_{n-1} is the quantization parameter of the latest previous frame.



(a)



(b)

Fig. 4. The coded bits and the bits estimated by rate model, (a) basketballPass (b) BQMall

The rate model is also tested on various test sequences. As shown in Fig. 4, the estimated rate is very close to the coded bits. So the rate model is proved to work well for HEVC.

3. DISTORTION MODELING FOR HEVC CODING

As pointed out in Section 1, in HEVC coding framework, two new modules SAO and ALF are added and they would reduce the quantization distortion of the reconstructed image. So, for the distortion model, besides the quantization distortion, we need consider the distortion reduction brought by SAO and ALF. So let O denote the original image. Y denotes the final reconstructed image after SAO and ALF. The final reconstructed image Y can be expressed as:

$$Y = O + e_{total} = O + e_q - e_{ALF} - e_{SAO} \tag{3}$$

$$e_{total} = e_q - e_{ALF} - e_{SAO} \tag{4}$$

e_{total} is the total reconstruction error. e_q,e_{ALF},e_{SAO} is the quantization error, ALF error reduction and SAO error reduction respectively.

We first give a study on the quantization error e_q. It was found that the relationship between the quantization error e_q and the quantization parameter can be modeled with an exponential function, shown as

$$e_q = k_1 * QP^x$$

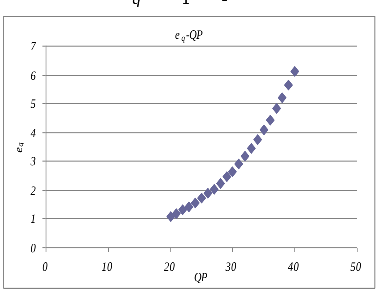


Fig. 5. Quantization error e_q and quantization parameter QP, BQSquare test sequence

For the ALF error reduction e_{ALF} and SAO error reduction e_{SAO}, it was found that both e_{ALF} and e_{SAO} have good linear correlation with the quantization parameter. That means the ALF and SAO error reduction are also related with the quantization. Assume e_{ALF} , and e_{SAO} can be modeled as:

$$e_{ALF} = k_2 * QP + b \tag{5}$$

$$e_{SAO} = k_3 * QP + c \tag{6}$$

So the total error and the total distortion D can be modeled as

$$e_{total} = k * QP^x + k' * QP + d \tag{7}$$

$$D = MSE = E\left(e^2{}_{total}\right) \tag{8}$$

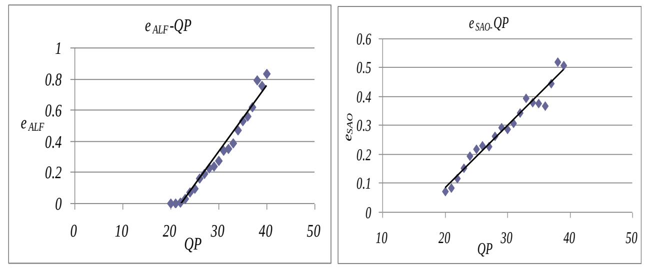


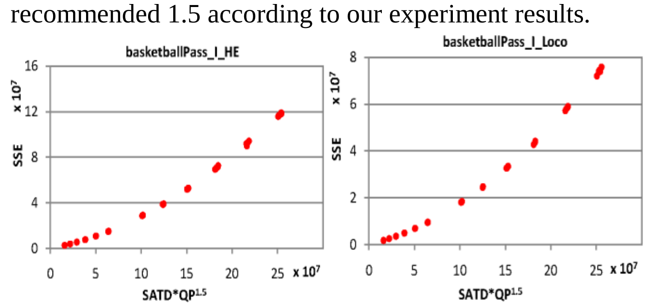
Fig. 6. The relationship between SAO/ALF error reduction and the quantization parameter QP, BQSquare test sequence

So we can see that the total distortion with ALF and SAO can be also modeled with quantization parameter, which is the traditional way for distortion modeling. As the distortion is also related with the content, we consider the relationship between the distortion and $SATD$ of the current picture for distortion modeling. The distortion of the coded slices, measured in Sum of Squared Error (SSE), versus the $SATD \times QP^{\gamma}$ for different slice types, both high efficiency and low complexity settings for HEVC coding are showed in Fig.7.

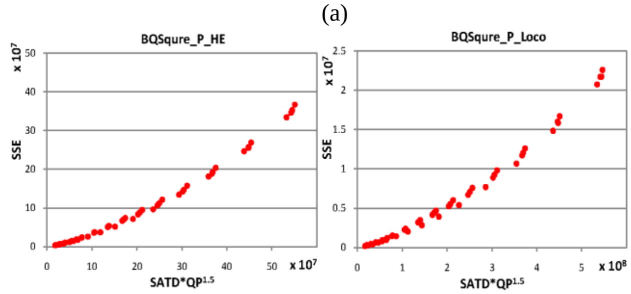
From the experiment results we can see that there is a strong exponential relationship between the slice distortion and $SATD \times QP^{\gamma}$. So far, as a variation of Equation (8), a new distortion model is proposed as:

$$D = \alpha \times (SATD \times QP^{\gamma})^{\beta} \tag{9}$$

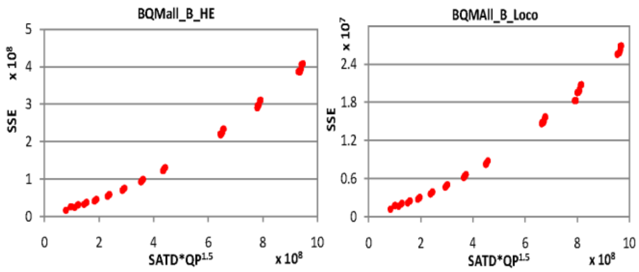
Where α and β are model parameters. And the value of γ is recommended 1.5 according to our experiment results.



(a)



(b)



(c)

Fig.7: The relation between distortion *SSE* and *SATD**$QP^{1.5}$. (a) I slice (b) P slice (c) B slice.

4. EXPERIMENTAL RESULTS

The rate model accuracy has been well proved in Section 2. This section will focus on the verification of the newly proposed distortion model. To verify the accuracy of the proposed distortion model, the relationship of the estimated distortion and the actual distortion of slices are plotted in Fig.8. Least square method is used to update the model parameters α and β of (9). It can be seen that the proposed distortion model can estimate the actual distortion accurately, especially for I slice.

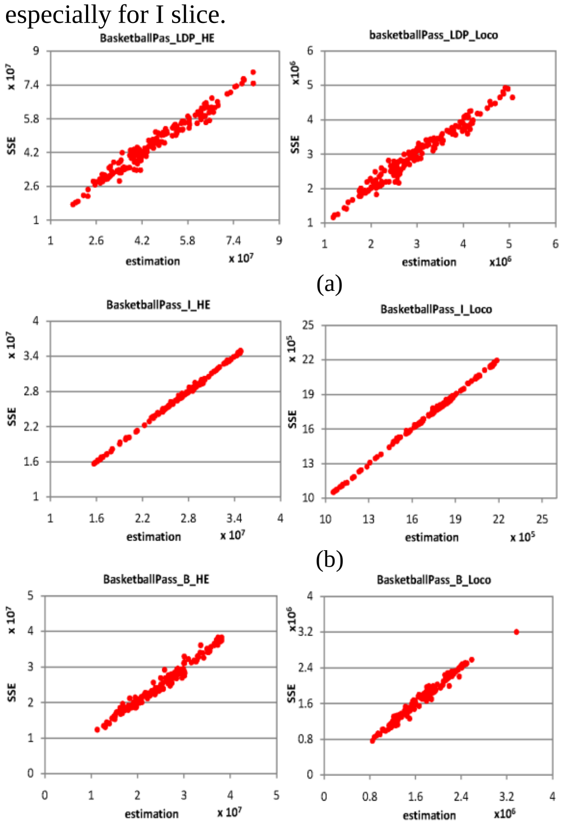


Fig.8: The relation between estimated distortion and the actual distortion. (a) P slice (b) I slice (c) B slice

Table 1 shows the average goodness of fitting R^2 values of several test sequences for I, P and B slice respectively. The R^2 can be up to 0.996 for I slice and 0.995 for P and B slice.

Table. 1: The distortion model goodness of fitting R-square on I, P and B slices

Sequence	I Slice	P Slice	B Slice
BasketballDrill	0.997	0.994	0.996
PartyScene	0.998	0.998	0.997
BQMall	0.993	0.993	0.989
RaceHorses	0.9973	0.9995	0.9996
Average	0.996	0.995	0.995

5. CONCLUSION

This paper gives a study on the rate distortion modeling for HEVC. And, new rate and distortion models are proposed by considering the new coding framework of HEVC. For rate modeling, a separate model for header bits and coefficient bits is proposed respectively. For distortion modeling, the quantization distortion and the distortion reduction induced by SAO/ALF are modeled jointly. The experimental results show that the proposed model can estimate the coding rate and distortion accurately. In the future the rate and distortion model will be further studied for using in mode decision and rate control.

6. ACKNOWLEDGMENT

This research is supported by the 973 program (2009CB320903) and the National Science Foundation of China (60833013, 61103088) of China, which are gratefully acknowledged.

7. REFERENCES

[1] G. J. Sullivan and J.-R. Ohm, “Recent developments in standardization of high efficiency video coding (HEVC),” SPIE Applications of Digital Image Processing XXXIII, Andrew G. Tescher (Editor), Proceedings of SPIE Volume 7798, Paper number 7798-30, August 2010.
[2] J. Katto and M. Ohta, “Mathematical analysis of MPEG compression capability and its Application to rate control,” IEEE International Conference on Image Processing, ICIP’95, vol. II, pp. 555-559, 1995.
[3] J. Corbera and S. Lei, “Rate control for low-delay video communications,” ITU Study Group 16, Video Coding Experts Group Documents Q15-A-20, Portland, June 1997.
[4] T. Chiang and Y. Zhang, “A new rate control scheme using quadratic distortion model,” IEEE Trans. Circuits Syst. Video Technol., vol. 7, no.1, pp. 246–250, Feb. 1997.
[5] 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.
[6] B. Lee and M. Kim, “Modeling rates and distortions based on a mixture of Laplacian distributions for inter-predicted residues in quadtree coding of HEVC, ” IEEE Signal Processing Letters, Vol. 18, No. 10, Oct. 2011.
[7] 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.