IMPROVING THE RATE-DISTORTION MODEL OF HEVC INTRA BY INTEGRATING THE MAXIMUM ABSOLUTE ERROR

Karina Jaskolka and André Kaup

Multimedia Communications and Signal Processing
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Cauerstr. 7, 91058 Erlangen, Germany

ABSTRACT

Normally, the mean squared error in conjunction with the rate is used to optimize the compression in hybrid video coding. However, in some areas, such as medical image coding, not only the average error but also the maximum error should be considered. Recently, it has been shown that incorporating the maximum absolute error into the calculation of the rate-distortion optimization (RDO) of the HEVC encoder can improve image quality, while preserving the average error and the rate. In this paper, we optimize the inclusion of the maximum absolute error in the RDO by considering the whole quantization parameter range as well as the different prediction unit sizes. Compared to the existing extended RDO, up to 1.5 % bitrate can be saved for the same maximum absolute error reduction.

***Index Terms*—** HEVC, rate-distortion optimization, intra prediction, medical volumes, high bit depth video.

1. INTRODUCTION

Hybrid video coders are the most successful class of video compression designs. This principle was introduced in the H.261 in 1991 [1]. In a hybrid video coder different combinations of mode decision are possible such as segmentation of the input picture, prediction mode decision, choice of quantization level, etc. The optimization goal of these coders is to find the best mode combination. Therefore, rate-distortion optimization (RDO) is included in such coders [2], [3]. With RDO a trade-off between the needed rate and the distortion occurring in a specific block is considered.

In the current coding standard HEVC [4] the distortion is calculated by the sum of squared differences in each block, so the average objective quality of a picture is involved in the RDO. In some application areas, especially for critical data like medical images, it is more important to reduce the maximum absolute error instead of the average error. Nevertheless, the average error should also be considered. In [5] a solution for this problem was introduced by extending the RDO of HEVC by the calculation of the maximum absolute error. However, in this method the introduced parameter are not optimized, since not the whole range of the Quantization Parameter (QP) is considered.

In this paper, an improved RDO$_{L_\infty}$ approach is proposed in order to decrease the maximum absolute error while preserving the rate compared to the existing RDO approach from [5]. This is achieved by considering the whole fidelity range of QP and all possible parameter combinations.

The paper is organized as follows: The next section introduces the state-of-the-art RDO in HEVC and the extended RDO from [5]. Based on this, the novel RDO$_{L_\infty}$ with an optimized calculation of the maximum absolute error is presented in Section 3. In Section 4 the simulation results are presented and compared. Section 5 concludes this work.

2. RATE-DISTORTION OPTIMIZATION IN HYBRID VIDEO CODING

In hybrid video coding the optimization task is to find the most similar representative under a given bit rate constraint. However, many options are possible to find this candidate. Since the various coding options show different behaviors on bit rate and different scene content, the goal should be to minimize the distortion D, subject to a constraint R_c, on the number of bits R used:

$$\min\{D\}, \text{ subject to } R < R_c. \tag{1}$$

2.1. Lagrangian Rate-Distortion Optimization

In [3] it was shown that (1) can be elegantly solved by Lagrangian optimization, where a rate term is weighted against a distortion term, the so called rate-distortion optimization:

$$J = D + \lambda \cdot R. \tag{2}$$

This equation is valid for all hybrid video coders, however, rate and distortion are determined in different ways.

In HEVC [4] the distortion D is calculated between the original block s and the predicted block p by the sum of squared differences (SSD), i.e., the squared L_2-norm:

$$D = L_2^2 = \sum_{(x,y)} |s[x,y] - p[x,y]|^2. \tag{3}$$

The Lagrange multiplier λ depends on a constant C, which depends on the prediction mode, and the Quantization Parameter (QP):

$$\lambda = C \cdot 2^{(\mathrm{QP}-12)/3}. \qquad (4)$$

The rate R is the number of needed coded bits of each prediction unit (PU). By minimizing (2) in the HEVC encoder the optimal solution will be found.

2.2. Lagrangian Rate-Distortion Optimization Considering the L_∞ Norm

The original RDO in HEVC considers only the average error of a picture, since the L_2-norm is calculated for minimizing J. In [5] it was proposed to use not only the average distortion D but also the maximum squared distortion $D_{\max}$ in calculation of the RDO. The maximum squared distortion is calculated by the squared L_∞-norm:

$$D_{\max} = L_\infty^2 = \max_{x,y}\{|s[x,y] - p[x,y]|^2\}. \qquad (5)$$

Here, s and p are defined as in the original RDO.

By adding the maximum squared distortion into the calculation of the RDO, J is calculated by:

$$J = \frac{(2-\alpha)\cdot D + \alpha \cdot D_{\max} \cdot \beta_{\mathrm{PU}}(\mathrm{QP})}{2} + \lambda \cdot R. \qquad (6)$$

With $\alpha \in [0,2]$ a weighting factor between D and $D_{\max}$ is introduced. The higher α is chosen, the more important is the maximum squared distortion in the calculation of the RDO. Depending on the value of α, different RDOs can be achieved:

$$J = \begin{cases} D + \lambda \cdot R, & \text{if } \alpha = 0, \\ \dfrac{D + D_{\max}\cdot \beta_{\mathrm{PU}}(\mathrm{QP})}{2} + \lambda \cdot R, & \text{if } \alpha = 1, \\ D_{\max} \cdot \beta_{\mathrm{PU}}(\mathrm{QP}) + \lambda \cdot R, & \text{if } \alpha = 2. \end{cases} \qquad (7)$$

$\beta_{\mathrm{PU}}(\mathrm{QP})$ is a weighting factor which shifts $D_{\max}$ into the same range as D, so both distortion values have the same scale and can be added. This weighting factor is necessary because in the HEVC encoder D and R are not normalized as their values depend on the PU size. Thus, for a bigger PU size the influence of $D_{\max}$ would be very small without $\beta_{\mathrm{PU}}(\mathrm{QP})$. In [5] $\beta_{\mathrm{PU}}(\mathrm{QP})$ is defined as a linear equation depending on the QP value for the four different PU sizes, which are optimized on a small range, i.e., four equations are defined for the different PU size which are optimized for the QP values 12, 17, 22, 27. This is a significant drawback as actually the whole fidelity range of QP should be take into account for the calculation of β.

3. PROPOSED EXTENSION

In this paper, a new solution for the weighting factor β is proposed in order to improve the RDO from [5]. It consists of two major parts: to consider the whole QP range of the HEVC coder from 1 to 51 and to combine the four separate equations for calculating β in [5] to one equation depending on all possible combinations of PU size and QP value.

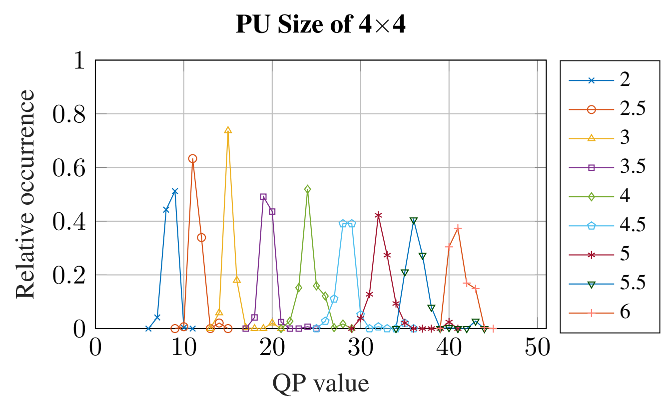

Fig. 1: Relative occurrences of the QPs for the different β-values shown in the legend for a PU size of 4×4. The tested input volumes are listed in Table 1. Only the even frames are used for training.

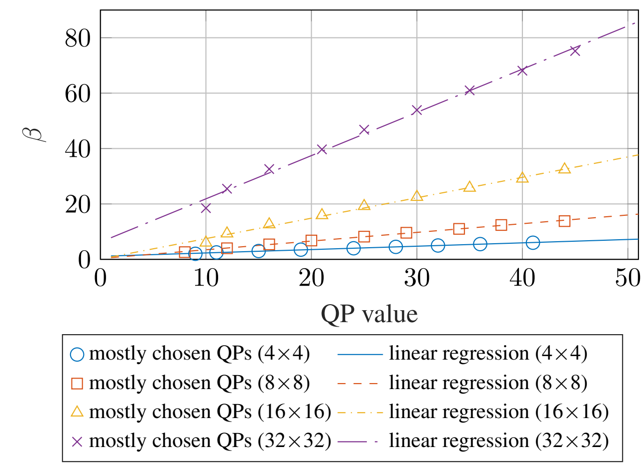

Fig. 2: Experimental relation between QP and β. The markers correspond to the maxima obtained by the method explained in Fig. 1, the curves are the proposed approximation.

3.1. Considering the QP Range

First, the new method for considering the whole QP range is introduced. Therefore, the obtained values from [5] for the different PU sizes and QP values are used as basis for the calculation of β. These calculated values for β are fixed for a specific PU size in the encoder, and also the PU size is constrained to this specific one, while the encoder is allowed to modify the QP value in a given range. Hence, the coder can choose different QP values for the different frames of a sequence for a specific PU size. After encoding different test sequences the relative occurrence of the different QP values can be analyzed.

In Fig. 1 the relative occurrence of the QPs for the different β-values for a PU size of 4×4 is depicted. The β-values are fixed to nine values in the range of [2,6] for a PU size of 4×4. The exact values can be seen in the legend. The QP values range from [5,45] in a step size of 5 and an offset of ± 5. So e.g., for a given β-value of 3.5, the QP value could vary from 15 to 25.

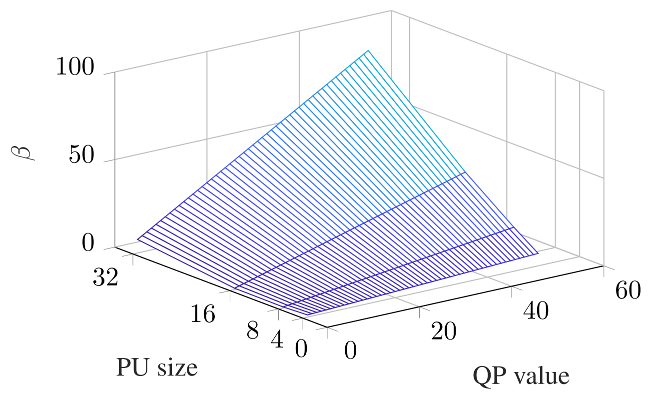

Fig. 3: Resulting surface from (8) as a function of PU size and QP value.

To calculate the relative occurrence for the different PU sizes separately, the PU size is fixed in the coder. This calculation is done for all PU sizes. As training set the even frames of the sequences from Table 1 are used.

We can observe that peaks occur for every value of β. In Fig. 2 the experimental relation between the dominant QPs and the fixed β-values as well as the approximated curves are depicted for all possible PU sizes. The circles are the obtained peaks from Fig. 1, and the straight line is the linear regression [6] to these points. This way we get four new equations for the β-QP-relation which can handle now the whole QP range.

3.2. Merging the β-QP-equations

To find an elegantly solution for β, a surface depending on both parameters is fitted to the curves obtained in Fig. 2. Therefore, a second order polynomial model is used. In Fig. 3, the resulting surface is depicted. The surface is linear in QP-β view and quadratic in PU-β view. The polynomial function depending on PU and QP is the result of curve fitting on the relative occurrences of the QPs:

$$\begin{aligned}\beta(\text{PU}, \text{QP}) = 2.59 - 0.091 \cdot \text{QP} - 0.426 \cdot \text{PU} \\ +0.0516 \cdot \text{QP} \cdot \text{PU} + (0.13 \cdot \text{PU})^2.\end{aligned} \quad (8)$$

The goodness of fit can be evaluated by the root-mean-square error (RMSE) and sum of squared error (SSE): RMSE is 0.13 and SSE is 3.21. Thus, a new equation for β can be found which can handle the whole QP range and the parameters PU and QP at the same time.

4. EXPERIMENTS

4.1. Experimental Settings

For a fair comparison with the results from [5] the HM reference software version 16.9 [7] is used for the simulation of HEVC encoder. The configuration file *intra_main_rext* [8] is selected for coding the sequences in intra mode. As test data set, different medical volumes with 12 bpp (*Med_12*) are used[1]. Since the size of the medical test set is quite limited, the even frames of these medical volumes are used for training and the odd frames are used for evaluation. So, two separate data sets are obtained. The properties of the volumes are summarized in Table 1. All medical test volumes have a spatial change in different directions as third dimension. However, *CT_heart* consists of a spatial and a temporal direction. Thus, we also evaluated a 3D+t CT hypervolume. *CT_heart* at position $t = 5$ defines a volume with spatial direction, *CT_heart* at position $z = 63$ defines a volume with temporal direction.

Table 1: Properties of the used medical test volumes *Med_12*.

	Sequence		Resolution	Frames	bpp
	CT_abdomen12		512×512	123	12
	CT_neck12		512×512	92	12
	MR_head		256×256	176	12
	MR_head_axial		448×512	23	12
	MR_abdomen		448×448	30	12
3D+t	*CT_heart*	$t = 5$	512×512	127	12
		$z = 63$		10	

Furthermore, three HEVC test sequences from *RangeExtension* (*RExt*) are evaluated as natural sequences with 10 bpp [9]: *BirdsInCage*, *EBURainFruits* and *Kimono*. The first 50 frames are used for the experiment. For a fair comparison the natural sequences are evaluated in 4:0:0 chroma format, because the medical test volumes are only available in luminance format.

The test sequences are encoded by the original RDO of HEVC, the extended RDO from [5] (Sec. 2.2) and the proposed RDO_{L_∞}. For evaluation, four QP values are chosen: 10, 20, 30, 40. To analyze the QP values, Bjøntegaard-Delta [10] is computed for rate (BD-Rate) and for maximum absolute error (BD-Max) for five different values of α. A detailed explanation about the calculation of BD-Rate and BD-Max can be found in [5].

4.2. Experimental Results

In Fig. 4, BD-Max is plotted over BD-Rate for the average values of *Med_12* (top) and *RExt* (bottom). Measurement points are referring to increasing values of α from left to right. The tested values for α are 0, 0.25, 0.5, 1, 1.5 and 2. It can be seen that for all tested values of α the proposed RDO_{L_∞} gives better results than the RDO from [5]. Not only less bitrate for a given α is needed, but also the maximum absolute error (Max) is further decreased. E.g., for a reduction of Max of around 56 in average for the medical volumes instead of 3.2 % more bitrate just 1.7 % more bitrate is needed compared to the original RDO (see yellow line in Fig. 4). Also for low values of α a reduction in the maximum absolute error can be observed. This behavior can also be seen for natural video sequences. Here, the reduction

[1] Some of the data set was kindly provided by Siemens Healthineers.

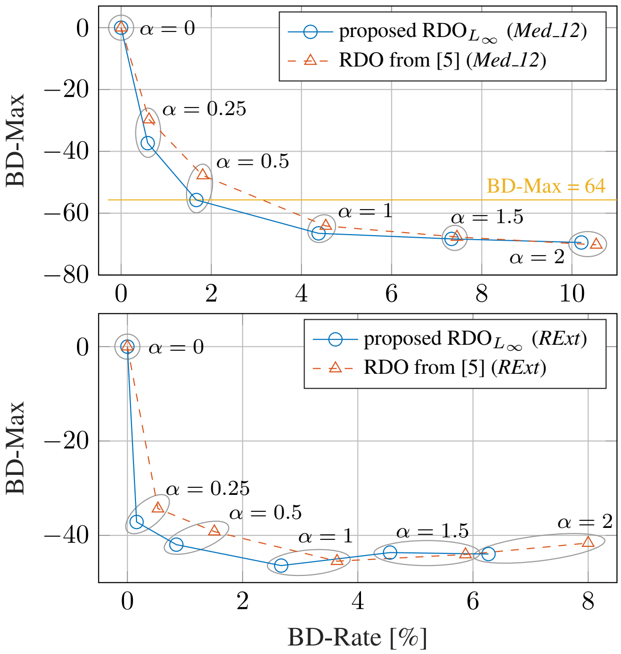

Fig. 4: BD-Max over BD-Rate of the proposed RDO$_{L_\infty}$ and the RDO from [5] for the average values for the classes *Med_12* (top) and *RExt* (bottom) for different values of α. The used QP values are 10, 20, 30, 40.

in rate for a specific value of α is getting larger, the higher the α-value is chosen. As the whole curve is shifted to the left, less bitrate is needed for coding natural sequences while reduction the maximum absolute error.

Two detail examples are shown in Fig. 5 for a visual comparison of the performance of the original RDO, the RDO from [5] and the proposed RDO$_{L_\infty}$. The used parameters are a QP value of 15 and 30 for *MR_head* and *Kimono*, respectively, and an α-value of 0.25. The shown rate, maximum error and PSNR values are calculated on the entire frame. In the upper row, an original frame of the test set can be seen. In the second and third row the two references, original RDO and RDO from [5], are depicted. In the lowest row, the proposed RDO$_{L_\infty}$ with $\alpha = 0.25$ is shown. For *MR_head* the details of the brain are better reconstructed. The complicated structure of the brain is better coded and looks more similar to the original. Additionally to the reduction of the maximum absolute error, the PSNR could be increased compared to the RDO from [5]. In *Kimono* details are better coded with the proposed RDO$_{L_\infty}$. The shapes appear more similar to the original. Furthermore, the maximum absolute error and the bitrate can be reduced compared to the RDO from [5]. Thus, the data fidelity can be further enhanced for natural and medical data.

5. CONCLUSION

In this paper, an improved RDO$_{L_\infty}$ approach including the L_∞ norm in HEVC has been proposed. This new method is able to reduce the maximum absolute error in natural and medical sequences, while rate and PSNR can be kept constant. A surface was included into the calculation of the RDO to consider the whole QP range and different PU sizes at the same time. For a reduction of the maximum absolute error of around 64 in medical volumes, 1.7 % instead of 3.2 % more rate has to be spend. Subjectively, the coded images are better reconstructed and less errors occur. In a next step, the principle of RDO including L_∞ norm will be applied to inter frame coding.

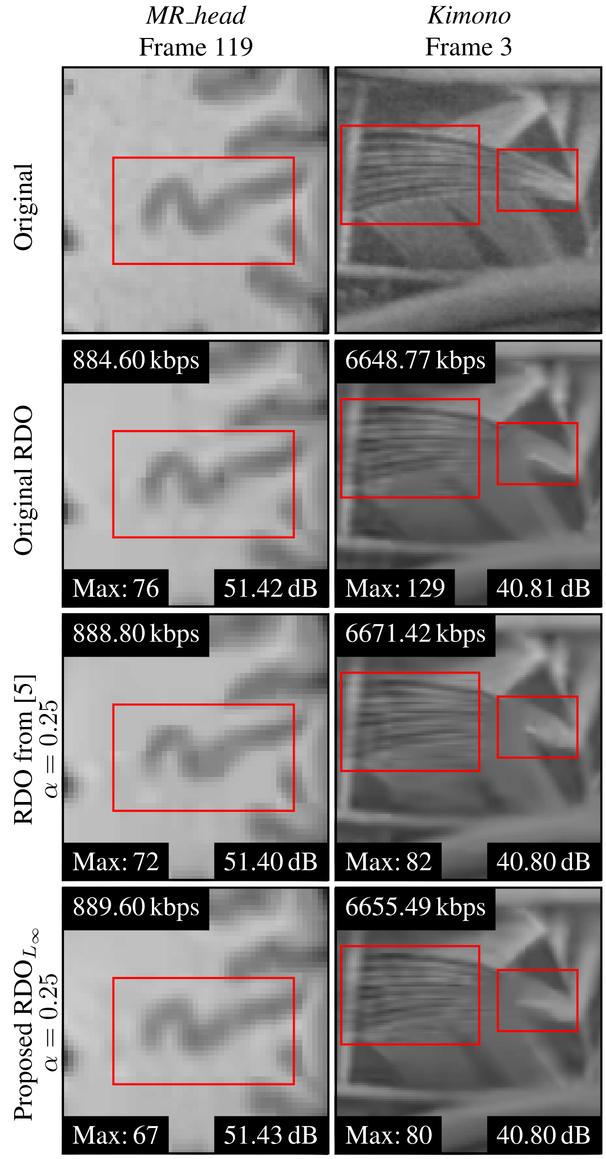

Fig. 5: Detail example of different frames for a visual comparison of the performance of the original RDO, the RDO from [5] and the proposed RDO$_{L_\infty}$ in HEVC using intra mode prediction and a QP value of 15 for *MR_head* and 30 for *Kimono*. Rate, PSNR and Max are measured on the entire frame. For the RDO from [5] and the proposed RDO$_{L_\infty}$ α is chosen to 0.25.

6. REFERENCES

[1] ITU-T (formerly CCITT), "Video Codec for Audiovisual Services at p $\times$ 64 kbit/s," ITU-T Recommendation H.261, version 2, Mar. 1993.

[2] A. Ortega and K. Ramchandran, "Rate-Distortion Methods for Image and Video Compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 23–50, Nov. 1998.

[3] G. J. Sullivan and T. Wiegand, "Rate-Distortion Optimization for Video Compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 74–90, Nov. 1998.

[4] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[5] K. Jaskolka and A. Kaup, "Joint Optimization of Rate, Distortion and Maximum Absolute Error for Compression of Medical Volumes Using HEVC Intra," in *Proc. Picture Coding Symposium (PCS)*, June 2018, pp. 1–5.

[6] D.C. Montgomery, E.A. Peck, and G.G. Vining, *Introduction to Linear Regression Analysis*, vol. 821, John Wiley & Sons, 2012.

[7] ITU/ISO/IEC, "HEVC Test Model HM," [Online]. Available: https://hevc.hhi.fraunhofer.de/.

[8] D. Flynn, D. Marpe, M. Naccari, T. Nguyen, C. Rosewarne, K. Sharman, J. Sole, and J. Xu, "Overview of the Range Extensions for the HEVC Standard: Tools, Profiles, and Performance," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 4–19, Jan. 2016.

[9] F. Bossen, "Common test conditions and software reference configurations," Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Jan. 2013.

[10] G. Bjøntegaard, "Calculation of Average PSNR Differences Between RD-Curves," ITU-T VCEG and ISO/IEC MPEG document VCEG-M33, Apr. 2001.