Optimizing the Lagrange Multiplier in Perceptually-Friendly High Efficiency Video Coding For Mobile Applications

Sima Valizadeh, Panos Nasiopoulos and Rabab Ward
Electrical and Computer Engineering, University of British Columbia

***Abstract*—In mobile applications, the amount of memory and bandwidth is restricted. Efficient compression of video is essential to provide better quality streaming on a mobile device. High efficiency video coding (HEVC) standard provide substantial bitrate savings compared with the previous standards. The compression efficiency of HEVC has been improved by integration of a perceptual video quality metric inside the encoder. In the rate distortion optimization process, PSNR-HVS has been used to measure distortion for the coding unit mode selection. In this paper, we find the optimal Lagrange multiplier based on the quantization parameter. Experimental results for various test sequences show the compression efficiency of the perceptual HEVC will be improved by finding the optimal Lagrange multiplier.**

***Index Terms*— Perceptual video coding, rate distortion optimization (RDO), Optimal Lagrange multiplier, human visual system (HVS), PSNR-HVS, high efficiency video coding (HEVC)**

I. INTRODUCTION

In video coding, the bitrate is minimized for a certain fixed distortion, or the distortion is minimized for a certain fixed rate. The minimization is formulated via a non-negative Lagrange multiplier in the rate distortion optimization (RDO) process [1]. In HEVC, a picture is partitioned into coding tree blocks (CTB) [2]. The encoder selects the size of the CTB based on the memory and delay constraints. CTBs are further partitioned into multiple coding blocks (CB) to form a quadtree structure. The minimum size of the CB is 8*8 and its maximum size is equal to the size of the CTB. The luma CB and the chroma CB, together with the associated syntax, form a coding unit (CU). In H.264, CUs were not present. Instead, a picture was partitioned into macroblocks (MB) with the fix size of 16*16.

There are spatial and temporal dependencies between blocks. However, these dependencies are ignored for practical applicability in the hybrid block-based video coding algorithms [3]. Alternatively, frame level distortion can be minimized instead of performing R-D cost optimization. However, computation complexity of minimizing the distortion over the entire frame becomes drastically high which makes it not practical due to the amount of complexity and delay.

This work was supported by the NPRP grant # NPRP 4-463-2-172 from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors

In HEVC, rate distortion optimization is done for each coding unit (CU) in four stages [4]: 1- coding unit mode (intra/inter) selection 2- intra prediction 3- motion estimation 4- quantization. In the rate distortion optimization process, distortion is measured by sum of square error (SSE), sum of absolute differences (SAD), or sum of absolute transform differences (SATD) in each of these four stages. Integration of a perceptual video quality metric in the rate distortion optimization process improves the compression efficiency of the encoder.

In the H.264 standard, the SSIM metric has been used in inter frame prediction and mode selection [5]-[9]. Also, SSIM has been integrated into the rate control of H.264 for intra frame coding [10]. In [11]-[14], a perceptual mode selection scheme based on SSIM is integrated into the rate distortion optimization of H.264. In a more recent study, SSIM has been used in the rate distortion optimization of the HEVC standard [15].

In the process of selecting coding unit (CU) mode (inter/intra), PSNR-HVS is integrated in the rate distortion optimization process [16]. PSNR-HVS is a perceptual quality metric that is based on the characteristics of human visual system (HVS) [17]. PSNR-HVS compared to PSNR, shows higher correlation with the subjective video quality evaluations.

In the rate distortion optimization process, the Lagrange multiplier acts as a knob that controls the trade-off be-tween rate and distortion. Finding the optimal Lagrange multiplier in the rate distortion optimization process is crucial for obtaining maximum compression efficiency. In our previous work [16], we have used a linear scaling relationship between the modified and the reference La-grange multipliers in HEVC. The compression efficiency of HEVC was improved by integrating a perceptual quality metric inside the rate distortion optimization process. However, optimal Lagrange multiplier should be deter-mined for the proposed approach. In this paper, we derive the optimal Lagrange multiplier for the perceptually modified HEVC.

The rest of the paper is organized as follows: Section 2 explains the proposed approach as well as the optimal

Lagrange multiplier. Section 3 includes the experimental results along with the optimal Lagrange multiplier formulation. Finally, section 5 concludes the paper and discusses the future directions of our work.

II. The Proposed Approach

For coding unit mode decision, distortion is measured by Sum of Square Error (SSE) in the rate distortion optimization process:

$$J(M) = D(M) + \lambda_{mode} R(M) \quad (1)$$

where D is the distortion, R is the required bitrate and M represents the mode for inter or intra coding. λ_{mode} is defined as [18]:

$$\lambda_{mode} = \alpha \times w_k \times 2^{((QP-12)/3.0} \quad (2)$$

where QP is the quantization parameter and α is defined as:

$$\alpha = \begin{cases} 1.0 - Clip3(0{,}0.5{,}0.05 \times number\_of\_B\_frames) \\ \qquad\qquad\qquad\qquad for\ referenced\ pictures \\ 1.0 \qquad\qquad for\ non\_referenced\ pictures \end{cases}$$

$$Clip3(x,y,z) = \begin{cases} x; z < x \\ y; z > y \\ z; otherwise \end{cases} \quad (3)$$

Coding unit mode decision is accompanied by intra prediction and inter-prediction of PU units. Motion estimation for each inter PU partition is broken into two parts: integer-sample precision and sub-sample precision. For integer-sample precision, distortion term is calculated by Sum of Absolute Differences (SAD) between original PU block and its motion compensated reconstructed block. For sub-sample precision, distortion is measured by Hadamard Transform to the block difference between reference and reconstructed block.

For intra PU prediction, minimization process is performed in parts. In the first part, minimizing the prediction cost function over 33 angular prediction directions, planar and DC mode gives a fixed number of candidate for intra prediction modes. In this part, distortion is measured by the Hadamard transform of the difference block. In the second part, the candidates list is augmented by three most probable modes. The best candidate is determined in the rate distortion optimization process.

At the last stage of the rate distortion optimization, the transform coefficient levels in the quantization process are adjusted. Distortion term is determined in transform domain. Finally, the Coding Unit mode decision is obtained recursively for the Coding Tree Unit (CTU).

Measuring the distortion by SSE and minimizing it in the rate distortion optimization process, leads to better PSNR in the coded video. However, the PSNR has been shown to have limited correlation with subjective tests. Various other video quality metrics have been developed to better represent how subjects evaluate the video quality.

PSNR-HVS [17] is a full reference video quality metric which takes into account the characteristics of the human visual system. One of the characteristics of the HVS is that its sensitivity decreases at high spatial frequencies. PSNR-HVS is defined as:

$$PSNR - HVS = 10 \log\left(\frac{255^2}{MSE_{HVS}}\right) \quad (4)$$

$$MSE_{HVS} = K \sum_{i=1}^{I-7} \sum_{j=1}^{J-7} \sum_{m=1}^{8} \sum_{n=1}^{8} ((X[m,n]_{ij} - X[m,n]_{ij}^{e}) T_C[m,n])^2$$

where K=1/[(I-7)(J-7)64]. I, J are the image width and height. X_{ij} is the DCT coefficient of an 8×8 image block with its upper left corner at (i,j). X_{ij}^{e} is the DCT coefficient of the corresponding block in the original image. T_c is a matrix adopted from the JPEG quantization table proposed in the JPEG [19].

Distortion measurements with a video quality metric changes this trade-off based on the range of output values. New optimal value of λ needs to be determined to have the best quality of video while minimizing the required bitrate. In our previous work [16], a constant scaling factor is used to denote the relationship between the new λ and the λ used in HEVC. However, the Lagrange multiplier should be optimized based on the quantization parameter (QP). In this paper, we find the optimal Lagrange multiplier in the rate distortion optimization process based on:

$$\lambda_{optimal}(\mathrm{QP}) = c(QP) \times \lambda_{mode}(QP) \quad (5)$$

In order to find the optimal Lagrange multiplier, contours of equal distortion are plotted in the plain of the two axes of the quantization parameter (QP) and coefficient c. On each contour, the point that leads to the least amount of bitrate is determined. By connecting these points and fitting a curve to them, the relationship between the optimal Lagrange multiplier and the quantization parameter is derived.

III. Experimental Results

For the test video sequences, we used MPEG standard videos as summarized in Table 1. These videos belong to the class of the mobile applications. All test video sequences are in the YCbCr 4:2:0 format. The length of each sequence is 10 seconds.

TABLE 1
TEST VIDEO SEQUENCES

Sequences	Resolution	Frame Rate	Total Frames
BQ Square	416× 240	60	600
BlowingBubbles	416 × 240	50	500
RaceHorses	416 × 240	30	300
PartyScene	832x480	50	500
BQMall	832x480	60	600
BasketballDrill	832x480	50	500

Contours of equal distortion are drawn for the BQSquare video sequence in Fig. 1. On each contour, the point associated with the least amount of bitrate is selected. These points lead to the minimum amount of bitrate resulting in the same amount of distortion. Fig. 2 shows the optimal coefficients based on the quantization parameter. These points are the extremum values in Fig. 1. By curve fitting, we can find the relationship between the optimal coefficient and the quantization parameter as:

$$c(QP) = 2.24e^{0.05\times QP} \quad (6)$$

By replacing $c(QP)$ from eq. 6 in eq.5, $\lambda_{optimal}$ can be found as:

$$\lambda_{optimal} = \alpha \times w_k \times 2^{((1.22QP-8.51)/3.0} \quad (7)$$

Using $\lambda_{optimal}$ in the rate distortion optimization process with the perceptual distortion measurement for the coding unit mode selection improves the compression efficiency of the high efficiency video encoder.

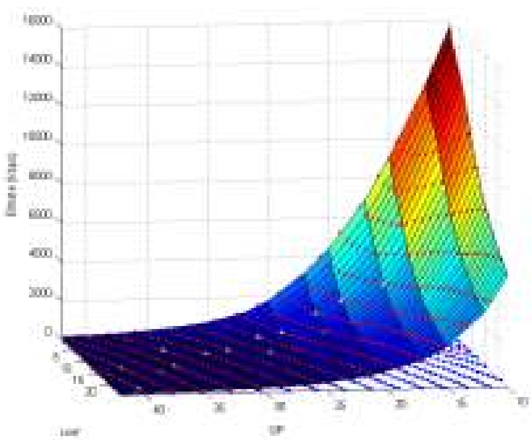

Fig. 1 Equal quality contours for the BQSquare video sequence. On each contour, the point associated with the least amount of bitrate is selected. These points lead to the minimum amount of bitrate resulting in the same amount of distortion

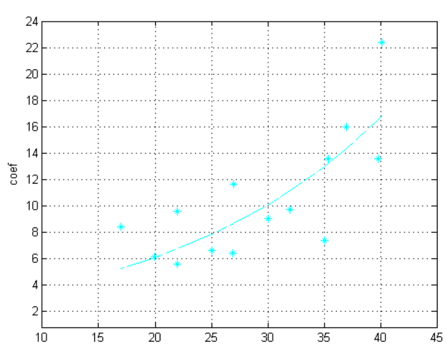

Fig. 2 Optimal coefficient based on the quantization parameter

Fig. 3 shows the Rate Distortion performance for video sequence BQSquare. The quality of the video is measured with Mean PSNR-HVS. PSNR-HVS has more correlation with subjective tests than PSNR [17]. Fig. 3 shows that our proposed method achieves higher quality at the same bitrate compared to HEVC reference software over the full range of Quantization Parameter (QP) values.

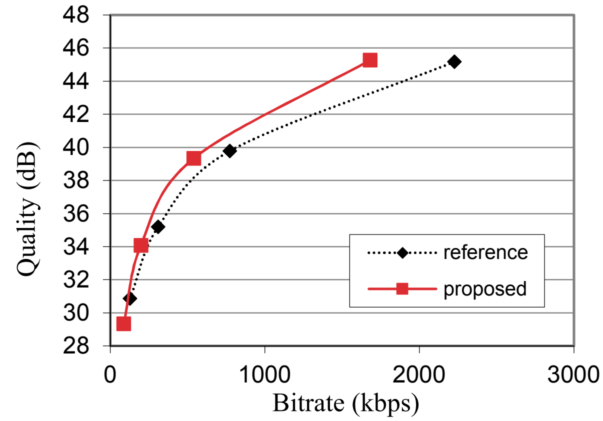

Fig. 3 Rate-Distortion Curves for BQSquare video. The quality is measured with Mean PSNR-HVS

The average bitrate difference between the proposed and reference rate-distortion curves is referred to Bjontegaard's Delta (BD) Rate [20]. A cubic polynomial approximation is derived using four data points (quality and bitrate points). The difference between the two curves is integrated in the horizontal direction for BD Rate. BD Rate measures the average bitrate savings by the proposed approach compared to the HEVC reference software. Bitrate savings of the proposed approach is summarized in Table 2.

TABLE 2
BITRATE SAVING OF THE PROPOSED APPROACH COMPARED TO THE REFERENCE HEVC SOFTWARE

Sequences	$\Delta rate$
BQ Square	-21.6%
BlowingBubbles	-10.1%
RaceHorses	-5.7%
PartyScene	-9.0%
BQMall	-6.5%
BasketballDrill	-3.8%
Average	-9.5%

5. CONCLUSION

In this paper, we found the optimal Lagrange multiplier based on the quantization parameter for the perceptually improved high efficiency video coding. In the perceptual HEVC, a video quality metric was integrated inside the rate distortion optimization procedure of the video coding. PSNR-HVS is a full reference quality metric based on the characteristics of the human visual system. PSNR-HVS was used as distortion measurement in the coding unit mode selection stage. The optimal Lagrange multiplier was derived based on the quantization parameter. The results show that our proposed scheme requires on average 9.5% less bitrate with the same perceived video quality compared to HEVC. In our future work, we will investigate the employment of the proposed approach in 3D-HEVC.

REFERENCES

[1] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *Signal Processing Magazine, IEEE,* vol. 15, pp. 74-90, 1998.

[2] V. Sze, M. Budagavi, G. Sullivan, *High Efficiency Video Coding (HEVC)*, Springer, 2014.

[3] J. Ohm, G. J. Sullivan, H. Schwarz, Thiow Keng Tan and T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC)," *Circuits and Systems for Video Technology, IEEE Transactions on,* vol. 22, pp. 1669-1684, 2012.

[4] G. J. Sullivan, J. Ohm, Woo-Jin Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *Circuits and Systems for Video Technology, IEEE Transactions on,* vol. 22, pp. 1649-1668, 2012

[5] Z. Mai, C. Yang, K. Kuang and L. Po, "A novel motion estimation method based on structural similarity for H.264 inter prediction," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing,* vol. 2, Feb. 2006, pp. 913-916, 2006.

[6] C. Yang, R.Leung, L. Po and Z. Mai, "An SSIM-optimal H.264/AVC inter frame encoder," in *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on,* 2009, pp. 291-295.

[7] C. Yang, H. Wang and L. Po, "Improved inter prediction based on structural similarity in H.264," in *Signal Processing and Communications, 2007. ICSPC 2007. IEEE International Conference on,* 2007, pp. 340-343.

[8] Z. Mai, C. Yang, L. Po, and S. Xie, “A new rate-distortion optimization using structural information in H.264 I-frame encoder,” *in Proc. ACIVS* 2005, pp. 435–441.

[9] Z. Mai, C. Yang and S. Xie, "Improved best prediction mode(s) selection methods based on structural similarity in H.264 I-frame encoder," in Proc. IEEE Int. Conf. *Sys. Man Cybern.,* May 2005, pp. 2673-2678 Vol. 3.

[10] B. H. K. Aswathappa and K. R. Rao, "Rate-distortion optimization using structural information in H.264 strictly intra-frame encoder," in *System Theory (SSST), 2010 42nd Southeastern Symposium on,* 2010, pp. 367-370.

[11] S. Wang, A. Rehman, Z. Wang, S. Ma and W. Gao, "Perceptual Video Coding Based on SSIM-Inspired Divisive Normalization," *Image Processing, IEEE Transactions on,* vol. 22, pp. 1418-1429, 2013.

[12] S. Wang, A. Rehman, Z. Wang, S. Ma and W. Gao, "SSIM-Motivated Rate-Distortion Optimization for Video Coding," *Circuits and Systems for Video Technology, IEEE Transactions on,* vol. 22, pp. 516-529, 2012.

[13] T. Ou, Y. Huang and H. H. Chen, "SSIM-Based Perceptual Rate Control for Video Coding," *Circuits and Systems for Video Technology, IEEE Transactions on,* vol. 21, pp. 682-691, 2011.

[14] C. Yeo, H. Li Tan and Y. H. Tan, "On Rate Distortion Optimization Using SSIM," *Circuits and Systems for Video Technology, IEEE Transactions on,* vol. 23, pp. 1170-1181, 2013.

[15] A. Rehman and Z. Wang, "SSIM-inspired perceptual video coding for HEVC," in *Multimedia and Expo (ICME), 2012 IEEE International Conference on,* 2012, pp. 497-502.

[16] S. Valizadeh, P. Nasiopoulos, R. Ward, “Perceptually-friendly rate distortion optimization in high efficiency video coding,” *European Signal Processing Conference (EUSIPCO-2015)*, August 2015.

[17] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, M. Carli, New full-reference quality metrics based on HVS, *Proc. of the Second International Workshop on Video Processing and Quality Metrics*, Scottsdale, USA, 2006, 4 p.

[18] K. McCann, B. Bross, W. Han, I. Kim, K. Sugimoto, G. Sullivan, High Efficiency Video Coding (HEVC) Test Model 13 (HM 13) Encoder Description, Joint Collaborative Team on Video Coding (JCT-VC), Document JCTVC-O1002, Geneva, Oct. 2013.

[19] G. Wallace, “The JPEG still picture compression standard,” *Comm. of the ACM*, vol. 34, No.4, 1991.

[20] G. Bjontegaard, “Calculation of average PSNR difference between RD curves,” *in Proc. 13th Meeting ITU-T Q.6/SG16 VCEG* , Austin, TX, Apr. 2001.