RATE-DISTORTION OPTIMIZED QUANTIZATION IN HEVC: PERFORMANCE LIMITATIONS

Jakub Stankowski, Cezary Korzeniewski, Marek Domański, Tomasz Grajek
Chair of Multimedia Telecommunications and Microelectronics,
Poznań University of Technology
60-965 Poznań, Polanka 3 Street, POLAND
{jstankowski, tgrajek}@multimedia.edu.pl

***Abstract*—Rate-Distortion Optimized Quantization (RDOQ) is an encoding optimization technique that may be modified with no implications to the bitstream compliance with the standard. In the paper, we propose an RDOQ variant with exact cost estimation resulting in higher compression ratios as compared to HEVC Test Model. This improvement is obtained at the cost of the complexity increase that may be scaled by replacing the individual steps of exact estimation by the simplified ones from HEVC Test Model. In that way, a family of RDOQ variants is created. The performance of the exact variant from this paper defines a kind of the performance limit for RDOQ.**

I. INTRODUCTION

The hybrid video coding scheme is the most common and the widest adopted approach to compressed video transmission. During the last two decades three generations of hybrid video coding technologies have been developed (Fig. 1) [1].

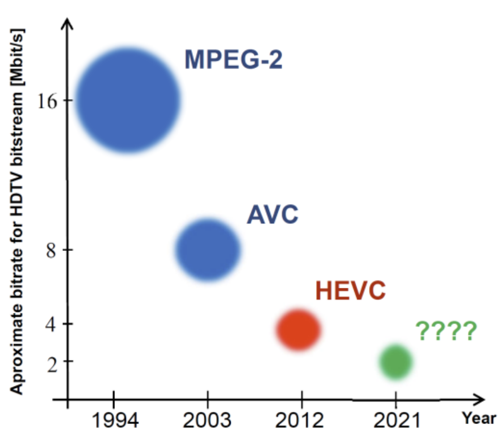

Figure 1. Generations of video coding technology. The reqired bitrate for every new generation is halved (with preserving the same quality).

In the most widespread video compression techniques like MPEG-2 Video/H.262 [2], AVC (MPEG-4 part 10 and H.264) [3,4] and novel HEVC (MPEG-H part 2 and H.265) [5,6] only bitstream syntax, bitstream semantics and decoder operations are standardized. Therefore, there is a wide margin of flexibility in encoder operation leading to the development of a number of encoder control algorithms. The development of the new control algorithms result in further improvement in coding efficiency within single standard (Fig. 2).

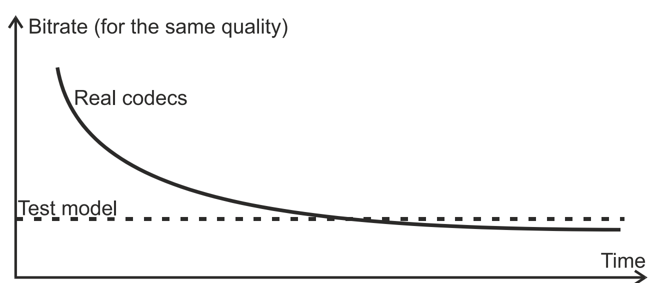

Figure 2. Bitrate for the same quality over time for singe compression standard.

In the paper the authors focus on calculation of quantized transform coefficients due to the fact that those coefficients constitute about 50-70% of compressed video bitstream [7]. Therefore, efficient calculations of quantized transform coefficients have significant influence on coding efficiency. However, the selection of quantization parameter (QP) value is out of the scope of this work.

II. QUANTIZED COEFFICIENTS CALCULATION

In the video encoder operation, the quantization stage is the only one responsible for lossy data reduction by both quantizing and reducing the number of transform coefficients. Therefore, the calculation of the quantized transform coefficients has a significant impact on the compression efficiency. The compression efficiency (in term of rate-distortion) offered by the encoder utilizing scalar quantization can be significantly improved by choosing the more sophisticated way of calculation of the quantized transform coefficients (it is not standardized and may be performed in any manner).

During the evolution of video compression techniques a number of approaches targeting the improvement of quantized coefficient calculation has been developed. For example:

- Dead zone quantizer aiming at the reduction of low magnitude coefficients (in particular magnitude equal to 1) by introducing the dead zone step in uniform quantizer [8,9].
- Coefficient thresholding "outliers" in AC coefficients of magnitude equal to 1 after quantization are discarded due to too high cost of transmission [10].
- Adaptive Rounding based on adaptive adjustment of a rounding offset used in coefficient quantization stage [11].
- Adaptive Quantization Matrix Selection (AQMS) allows for adjustment of quantization matrices to encoded signal also on the decoder side [12].
- Rate Distortion Optimized Quantization based on rough estimating the RD cost of modification or removal of selected transform coefficient or transform coefficients group [13].

Due to its efficiency the RDOQ has been selected as one of two possible quantization approaches in the HEVC test model development [14]. Therefore, in this paper the RDOQ has been selected as a most promising technique for future improvements. The performance limitations for improved RDOQ have been evaluated.

III. RATE-DISTORTION OPTIMIZED QUANTIZATION (RDOQ)

The improved quantizer can take into account both the quantization error and the number of bits required to transmit the transform coefficient and determine the optimal set of quantized transform coefficients. The optimization can be performed for every block of transform coefficients (i.e. Transform Unit (TU) in HEVC) and the optimal cost is calculated by minimizing the Lagrangian function [15].

The above introduced approach has been proposed in [16] for JPEG and MPEG, and then in [17,18] for MPEG-4 AVC and in [13] for KTA. The general idea of presented method is to find an optimal set of quantized transform coefficients which corresponds to the lowest RD cost. Theoretically the determination of optimal set of quantized coefficients requires exhaustive search by evaluating all possible combinations. The exhaustive approach is impractical due to extreme computational complexity of the encoder. Therefore, fast suboptimal approaches have been introduced. Additionally, in [13] the name "Rate-Distortion Optimized Quantization (RDOQ)" has been proposed, which is widely adopted

The purpose of the RDOQ is to find the optimal or suboptimal set of quantized transform coefficients representing a residual data in a encoded block. The RDOQ calculates the image distortion (introduced by quantization of transform coefficients) in encoded block and a number of bits needed to encode the corresponding quantized transform coefficient. Based on these two values, the encoder chooses better coefficient value, by calculating RD cost.

IV. THE RDOQ IN HEVC

The RDOQ has been included in the HEVC reference software (HM) and intensively used during HEVC development and performance. This section describes the RDOQ algorithm adapted to HEVC.

The adaptation of the RDOQ in HM is closely related to HEVC residual coding techniques. In the HEVC the Transform Unit size may vary from 4x4 to 32x32 pixels and only square units are allowed. After transformation and scanning, coefficients are divided into coefficient groups (CG) containing 16 transform coefficients (Fig. 3). The detailed description of HEVC transform coefficient coding could be found in [19].

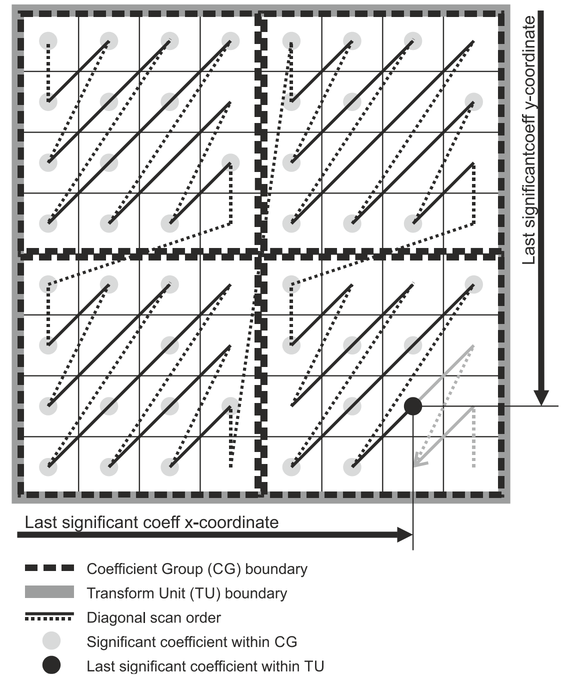

Figure 3. The exemplary 8x8 Transform Unit (TU) with division into four Coefficient Groups (CGs), diagonal scan order and last significant coeff marked.

The RDOQ operation in encoder can be divided into three stages: quantization of transform coefficients, elimination of coefficient groups (CG) and selection of the last non-zero coefficient.

A. Quantization of transform coefficients

In this stage the encoder performs calculation for each of transform coefficients separately. In the first step, the encoder calculates the value *Level* by quantizing the magnitude of transform coefficient by using the uniform quantizer without dead zone. In the next step, the encoder considers two additional magnitudes of the analyzed quantized coefficient: *Level-1* and *0*. For every of the mentioned coefficient magnitudes, the encoder calculates the RD cost of encoding the coefficient with the selected magnitude and chooses the one with the lowest RD cost.

It is worth to mention that, when compared to coefficient magnitude value equal to *Level* setting the magnitude to *0* value allows for more significant bitrate reduction than setting the lower magnitude value (*Level-1*). However, the elimination of selected transform coefficient by setting the magnitude to 0 may cause significant distortion.

B. Elimination of coefficient groups

In this stage the encoder performs the calculation for each of transform coefficient group (CG). The encoder calculates the RD cost of eliminating the whole CG. The elimination of the whole CG is performed by quantization of all coefficient in CG to magnitude zero. The encoder calculates the RD cost of elimination of the analyzed CG and, if the elimination allows for cost reduction, the selected CG is eliminated.

The elimination of the entire CG can lead to substantial bitrate reduction (there is no need to transmit *sig_coeff_flag* for every coefficient within CG) while introducing substantial distortion to the reconstructed image.

C. Selection of the last non-zero coefficient

The last stage of RDOQ is performed after steps A and B for all remaining CGs in TU. RDOQ algorithm analyzes coefficients to find the best (in term RD cost) last non-zero coefficient position. This step is included because the encoder has to encode the (x, y) coordinates of the last non-zero coefficient in the bitstream.

D. Calculation of RD Cost

During RDOQ operation the encoder has to calculate the cost of every considered set of transform coefficients or coefficient groups. This cost (*RD_cost*) may be calculated by taking into account the number of bits (B) required to encode selected coefficient, CG or TU, the introduced distortion (D) and weighting both values by Lagrange multiplier (λ) [13]:

$$RD\_cost = D + \lambda \cdot B \tag{1}$$

In different implementations, encoders may use the exact or estimated values of introduced distortion and a number of bits required to encode selected transform coefficient, coefficient group or transform unit (see Table I).

TABLE I. EXEMPLARY PERFORMANCE/COMPLEXITY TRADEOFS.

	Estimated number of bits calculation	**Exact number of bits calculation**
Estimated distortion calculation	RDOQ implementation in HEVC	Posible performance/complexity tradeof
Exact distortion calculation	Posible performance/complexity tradeof	RDOQ implementation evaluated in this paper

The usage of estimated values leads to some mistakes in the best coefficient set selection and causes some compression performance degradation. However the estimation of rate and distortion can speed up the encoder operation.

V. THE SIMPLIFIED RDOQ IN HEVC REFERENCE SOFTWARE

In the RDOQ implemented HEVC test model (HM16) [20] the encoder uses only estimated values of introduced distortion (represented by square quantization error) and a number of bits required to encode selected transform coefficient, coefficient group or transform unit.

For example, for every of the examined coefficient magnitude the encoder calculates the cost RD_cost*(L, c)* of encoding the coefficient c with the magnitude L according to (2) and chooses the case with the lowest RD cost.

$$RD\_cost\,(L, c) = est\_D(L, c) + \lambda \cdot est\_B(L, c), \tag{2}$$

where:
c – transform coefficient identifier,
L – value of quantized transform coefficient c,
RD_cost(L, c) – cost of quantization coefficient c to value L,
est_D(L, c) – square quantization error,
est_B(L, c) – estimated number of bits needed do encode coefficient c quantized to value L,
λ – Lagrange multiplier.

The detailed description of RDOQ implementation in HEVC can be found in [21].

VI. THE RDOQ WITH EXACT COST CALCULATION

The RDOQ implementation evaluated in this paper is called “Exact cost RDOQ”, because its main feature is the exact calculation of introduced distortion and a number of bits required to encode selected transform coefficient, coefficient group or transform unit. Moreover, the higher number of possible modifications of coefficient magnitude is considered – besides *L' = Level*, *L'=Level-1* and *L'=0*, the *L'=Level+1* is also examined.

The exact calculation of introduced distortion is performed in TU granularity by scaling the quantized transform coefficients, inverse transformation of scaled coefficients, calculation of reconstructed image block (corresponding to currently processed TU) and calculation of Sum of Square Differences (SSD) between original and reconstructed blocks. The exact calculation of a number of bits required is also performed for entire TU. The encoder uses CABAC engine to calculate a number of bits. The CABAC internal state is stored, the examined TU is encoded in order to calculate a number of required bits and the state of CABAC is restored (Fig. 4).

Both abovementioned operations are performed at TU level, therefore the examination of each modification of single transform coefficient requires reconstruction (SSD calculation) and encoding (number of bits calculation) of entire TU containing examined coefficient. For each of examined coefficient value the number of bits needed to encode TU and image distortion are calculated:

$$RD\_cost\,(L,\,c) = SSD(L,\,c) + \lambda \cdot B(L), \tag{3}$$

where:
c – transform coefficient identifier,
L – value of quantized transform coefficient c,
$RD\_cost(L, c)$ – cost of quantization coefficient c to value L,
$SSD(L, c)$ – sum of square differences between original and reconstructed image blocks corresponding to encoded TU,
$B(L)$ – exact number of bits needed do encode TU with coefficient c quantized to L value,
λ – Lagrange multiplier.

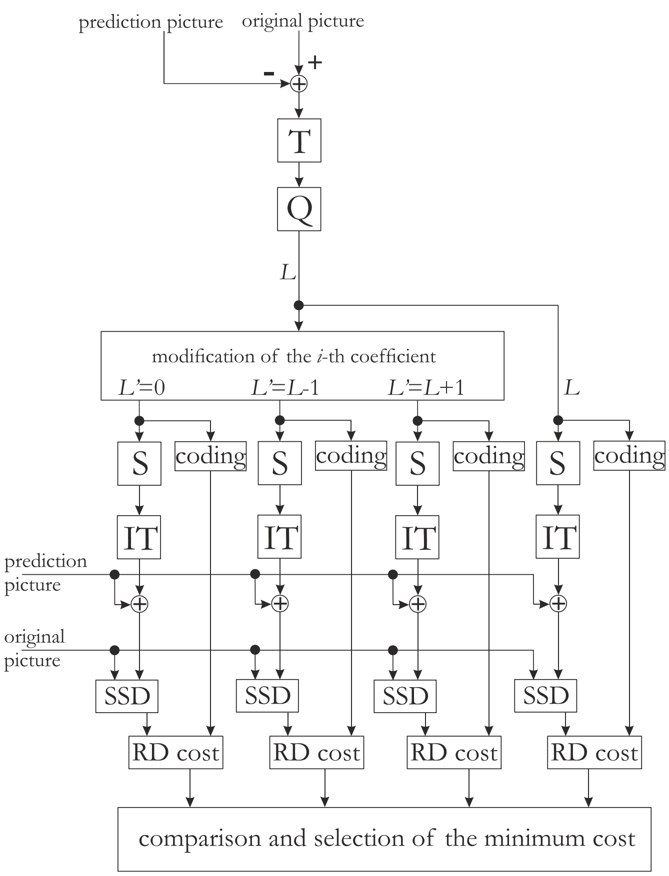

Figure 4. Flowchart proposed RDOQ improvement
(T – transform, Q – quantization, S – scaling, IT – inverse transform, SSD – Sum of squared differences)

VII. Experimental Evaluation

The experimental evaluation has been performed under the conditions described in the Common Test Conditions and Software Reference Configurations [14]. The test conditions described in [14] and previous versions of this document have been used by JCT-VC (Joint Collaborative Team on Video Coding of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11) to evaluate proposals during the development of HEVC technology.

The set of test sequences described in [14] consists of 6 classes differing in resolution and content, where: class A – natural content, 4k (cropped to 1600p) resolution, class B – natural content 1080p resolution, class C – natural content WVGA resolution, class D – natural content WQVGA resolution,. class E – teleconference 720p resolution and class F – synthetic and mixed content, different resolutions.

The [14] defines four coding scenarios depending on the used picture types and GOP structure: “All-Intra” – with intra coded pictures only, “Random access” – with hierarchical GOP, “Low-delay P” and “Low-delay B” – only with forward temporal prediction. In this paper only “All-Intra” and “Random access” scenarios have been evaluated.

VIII. Experimental Results

The experimental results of “All-Intra” test scenario have been shown in Table I. and results of “Random access” test scenario have been shown in Table II. The results are presented as BD-delta (Bjontegaard delta) [22]. However, due to interpolation issues and their influence on BD-delta results, the 3rd order polynomial interpolation has been replaced by Piece-wise Cubic Interpolation [23].

Table I. Experimental results for “All-Intra” test scenario.

	HM RDOQ performance [%]			**Exact cost RDOQ implementation performance [%]**			**HM RDOQ vs Exact cost RDOQ [%] (performance limit)**		
Class	***Y***	***Cb***	***Cr***	***Y***	***Cb***	***Cr***	***Y***	***Cb***	***Cr***
A	-4.1	1.0	-0.7	-5.1	0.7	-1.2	-1.0	0.0	-0.4
B	-5.4	0.1	-1.4	-6.5	-0.9	-1.1	-1.2	0.0	-0.2
C	-3.5	1.0	-0.3	-4.4	1.1	-0.2	-1.0	0.4	0.3
D	-3.3	1.2	0.0	-4.2	1.3	0.2	-1.0	0.3	0.4
E	-3.7	-0.9	-2.9	-4.5	-1.0	-3.2	-0.9	0.1	-0.2
F	-2.6	0.0	-0.9	-3.8	-0.3	-1.3	-1.2	-0.1	-0.2
All	**-3.8**	**0.4**	**-1.0**	**-4.8**	**0.1**	**-1.0**	**-1.1**	**0.1**	**0.0**

Table II. Experimental results for “Random access” test scenario.

	HM RDOQ performance [%]			**Exact cost RDOQ implementation performance [%]**			**HM RDOQ vs Exact cost RDOQ [%] (performance limit)**		
Class	***Y***	***Cb***	***Cr***	***Y***	***Cb***	***Cr***	***Y***	***Cb***	***Cr***
A	-3.3	-4.9	-6.4	-4.0	-3.0	-4.7	-0.9	-0.2	-0.3
B	-5.1	-5.9	-6.3	-5.6	-6.7	-6.4	-0.9	0.0	-0.4
C	-4.0	-5.9	-6.4	-4.8	-5.2	-5.6	-0.9	0.3	0.0
D	-3.9	-5.0	-6.1	-4.8	-4.9	-5.6	-0.9	0.1	0.4
E	-3.2	-2.6	-4.3	-4.0	-2.8	-4.9	-0.9	-0.3	-0.7
F	-3.7	-4.6	-5.1	-4.9	-4.4	-5.0	-1.3	-0.2	-0.2
All	**-4.0**	**-5.0**	**-5.8**	**-4.7**	**-4.5**	**-5.4**	**-1.0**	**-0.1**	**-0.1**

As shown in the Table I for All-Intra and Table II for Random access test scenario the HM RDOQ allows to achieve 3.8% and 4.0% average bitstream reduction for the same image quality respectively. The achieved performance improvement limit (for exact cost calculation algorithm) is 4.8% and 4.7% respectively. Therefore, in the extreme case the HM RDOQ performance could be improved at least by 1% in term of bitstream reduction.

The average encoding time for RDOQ with precise calculations has been 3 to 4 times longer than for HM RDOQ. Therefore the fast HM RDOQ implementation offers 1% lower performance but with significantly lower computational complexity.

IX. Conclusions

In this paper the performance limitation of RDOQ in HEVC has been evaluated. The increase of compression efficiency provided by the fast RDOQ implementation has been measured. Moreover the RDOQ implementation with exact calculation of distortion and number of bits has been introduced. The described approach to RDOQ algorithm allows to find the performance limit which could be reached by the exact calculation of bits number and distortion.

In this paper the upper limit of RDOQ introduced performance improvement has been determined. However it is possible to achieve another performance-encoding complexity tradeoff i.e. by using the exact calculation of distortion in combination with the estimation of the number of bits.

Acknowledgement

Research project was supported by National Science Centre, Poland, according to the decision DEC-2012/05/B/ST7/01279.

References

[1] M. Domański, T. Grajek, D. Karwowski, J. Konieczny, M. Kurc, A. Łuczak, R. Ratajczak, J. Siast, J. Stankowski, K. Wegner, "Coding of multiple video+depth using HEVC technology and reduced representations of side views and depth maps", 29th Picture Coding Symposium, PCS 2012, Kraków, Poland, May 2012.

[2] ISO/IEC 13818-2:2013, Information technology - Generic coding of moving pictures and associated audio information - Part 2: Video / Recommendation H.262 : Information technology - Generic coding of moving pictures and associated audio information: Video, International

[3] ISO/IEC 14496-10:2014, Information technology - Coding of audio-visual objects - Part 10: Advanced Video Coding / Recommendation H.264 : Advanced video coding for generic audiovisual services, International Telecommunication Union - Telecommunication Standardization Sector (ITU-T), Feb. 2014.

[4] T. Wiegand, G.J. Sullivan, G. Bjontegaard, A. Luthra, "Overview of the H.264/AVC video coding standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560-576, Jul. 2003.

[5] ISO/IEC 23008-2:2013, Information technology – Coding of audio-visual objects – Part 2: High efficiency video coding / Recommendation ITU-T H.265, High efficiency video coding for generic audiovisual services, 2013.

[6] 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. 1648–1667, Dec. 2012.

[7] J. Stankowski, D. Karwowski, T. Grajek, K. Wegner, J. Siast, K. Klimaszewski, O. Stankiewicz, M. Domański, "Bitrate Distribution of Syntax Elements in the HEVC Encoded Video", International Conference on Signals and Electronic Systems, ICSES 2014, Poznań, Poland, Sep. 2014.

[8] G.J. Sullivan, "On embedded scalar quantization", IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2004, vol.4, pp. 605-608, Montreal, Canada, May 2004.

[9] T. Wedi, S Wittmann, "Quantization offsets for video coding", IEEE International Symposium on Circuits and Systems, ISCAS 2005, pp. 324-327, Kobe, Japan, May 2005.

[10] Y.-K. Tu, J.-F. Yang, M.-T. Sun, "Rate-Distortion Modeling for Efficient H.264/AVC Encoding", IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 5, pp. 530-543, May 2007.

[11] G.J. Sullivan, "Adaptive quantization encoding technique using an equal expected-value rule", JVT meeting contribution JVT-N011, Hong Kong, Jan. 2005.

[12] A. Tanizawa, T. Chujoh, "Adaptive Quantization Matrix Selection", ITU-T Q.6/SG16 VCEG, document: D.266, 2006.

[13] M. Karczewicz, Y. Ye, I. Chong, "Rate Distortion Optimized Quantization", ITU-T SG 16/Q 6 VCEG, document: VCEG-AH21, Jan. 2008.

[14] 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, document: JCTVC-L1100, Geneva, Switzerland, Jan. 2013.

[15] T. Wiegand, M. Lightstone, D. Mukherjee, T. Campbell, S. Mitra "Rate-distortion optimized mode selection for very low bit-rate video coding and the emerging H.263 standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, no. 2, pp. 182 -190, Apr. 1996.

[16] K. Ramchandran, M. Vetterli, "Rate-distortion optimal fast thresh-olding with complete JPEG/MPEG decoder compatibility," IEEE Transactions on Image Processing , vol. 3, no. 5, pp. 700–704, Sep. 1994.

[17] E.-H. Yang, X. Yang, "Rate Distortion Optimization of H.264 with Main Profile Compatibility", IEEE International Symposium on Information Theory, ISIT 2006, pp. 282 – 286, Seatlle, USA, July 2006

[18] E.-H. Yang, X. Yu, "Rate distortion optimization for H.264 inter-frame coding: A general framework and algorithms," IEEE Transactions on Image Processing , vol. 16, no. 7, pp. 1774–1784, Jul. 2007.

[19] J. Sole, R. Joshi, N. Ngyuen, T. Ji, M. Karczewicz, G. Clare, F. Henry, A. Duenas, "Transform Coefficient Coding in HEVC", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1765-1777, Dec. 2012.

[20] K. McCann, C. Rosewarne, B. Bross, M. Naccari, K. Sharman, G. J. Sullivan (editors), "High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description", Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, document: JCTVC-R1002, Sapporo, Japan, Jun. 2014.

[21] J.-R. Ohm, G.J. Sullivan, H. Schwarz, T.K. Tan; T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC)," IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1669-1684, Dec. 2012.

[22] G. Bjontegaard, "Improvements of the BD-PSNR model," ITU-T SG16 Q.6 Document, VCEG-AI11, Berlin, Germany, Jul. 2008.

[23] J. Wang, X. Yu, D. He, "On BD-rate calculation", Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, document: JCTVC-F270, Torino, Italy, Jul. 2011.