

The impact of complexity in the rate-distortion optimization: A visualization tool

Ahmed Aldahdooh, Marcus Barkowsky, Patrick Le Callet

▶ **To cite this version:**

Ahmed Aldahdooh, Marcus Barkowsky, Patrick Le Callet. The impact of complexity in the rate-distortion optimization: A visualization tool. 22nd INTERNATIONAL CONFERENCE ON SYSTEMS, SIGNALS AND IMAGE PROCESSING (IWSSIP 2015), Sep 2015, Londres, United Kingdom. pp.45-48, 10.1109/IWSSIP.2015.7314173 . hal-01346405

HAL Id: hal-01346405
https://hal.science/hal-01346405

Submitted on 18 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Impact of Complexity in the Rate-Distortion Optimization: a Visualization Tool

Ahmed Aldahdooh, Marcus Barkowsky, and Patrick Le Callet
LUNAM University, University of Nantes, IRCCyN UMR CNRS 6597
Polytech Nantes, Rue Christian Pauc BP 50609 44306 Nantes Cedex 3, France
Email: *firstname.lastname@univ-nantes.fr*

***Abstract*—Recently, the new video coding standard, High Efficiency Video Coding (HEVC), was released. A 50% bitrate reduction at the same visual quality relative to the previous standard was achieved. These gains are partly related to tools that increase complexity such as coding decisions taken on smaller image areas or further improvements in motion compensated coding. This paper addresses the complexity issue, i.e. the execution time, and its impact when it is combined with the bitrate and distortion optimization of the video coding. Given a set of encoder configurations that restrict the encoding such that the complexity varies, an analysis including a visualization tool is proposed to help the user to select the best configuration for a specific amount of rate, distortion and complexity. A possible targeted applications are also introduced.**

***Keywords*—Video compression optimization; RDC optimization; Execution time; Complexity; HEVC; UHD**

I. Introduction

Video coding aims to reduce the required bitrate to send/store the video sequences while preserving a satisfying quality. HEVC [1], [2] came with new intra and inter prediction modes which adds new complexity since the rate-distortion optimization [3] needs to test all the modes. A number of research efforts were conducted to reduce the candidates of intra prediction modes for H.264, see for example [4]–[6] which can be adopted for HEVC. HEVC also came with new or improved tools such as quad-tree block structure in which a suitable partitioning is selected by the optimization process. Since this feature adds additional complexity, other research activities are also made to predict the coding tree unit depth, see for example [7]–[9]. Choosing encoder parameters values also trade-off the quality and the complexity. For instance, selecting a smaller motion search range value, accelerates the encoding process at the price of quality and a larger value may slow down the encoding process. Another factor that may increase the complexity is the sequence resolution and formats. Moreover, in the real world, different video classes like, natural scenes, cartoons, sports, news broadcasting and computer-generated videos exist and each class may be categorized into subclasses. Research showed that these content types have an impact in video coding. Pitrey et al. in [10] showed that the video content influences the video encoding. In more detail, one may encode a given video with several configurations (C_1, ..., and C_n), and may get a statistically equivalent Mean Opinion Score (MOS) for the output videos, where one configuration yields minimal computational complexity. This minimum may change from one content to another depending on video content characteristics. Designing parallel encoder is one possible solution to deal with the added complexity but such encoders are expensive. Therefore, tools to reduce the encoding time without compromising the coding efficiency and the perceived quality are important.

The challenge is to choose the best coding parameters to optimize the trade-off between compression ratio, quality, and complexity. This paper addresses this problem analytically by applying a linear optimization model and providing a visualization tool that helps to select the best configuration for a specific rate, distortion, and complexity point.

The rest of the paper is organized as follows: the experiment setup is explained in Section II. Section III describes the proposed visualization tool and the experimental results. A discussion and the drawn conclusion are introduced in Section IV and V respectively.

II. Experiment Setup

A. Video sequences

In this paper, a set of SJTU 4K Video Sequences [11] that are available for research purposes are investigated, snapshots are shown in Figure 1. Conditions and limitations mentioned in [12] and [13] are considered.

B. Encoder parameters configurations

To study the impact of the complexity, seven encoder configurations, for HM13.0 encoder, are prepared. In terms of complexity, 3 low, 2 medium, and 2 high complexity configurations are chosen as shown in Table I. They differ in the coding tree size and its depth and the motion search range. One can notice that the quantization parameter is fixed and hence, a certain quality of the overall quality range is selected for all configuration.

III. Analysis steps and results

In this section, the analysis steps and the build of visualization tool are described. In this paper, the "execution time" is used as a complexity measure and it is measured by encoding the sequences using recent computer. Firstly, when a video needs to be encoded, there are many encoder configurations that can be used and the one that yields the lowest execution time is needed to be chosen in order to



Figure 1: Snapshots of video sequences

compare it with others. To know which configuration yields the lowest execution time in general, the video sequences are encoded with configurations listed in Table I, and then the average of execution time for each configuration is calculated and the minimum-complexity configuration (C_{minC}) is selected. In order to distinguish between a subtle or obvious difference between two execution times or between two bitrate values, a relative measure is used to model the linear optimization formula. Hence, the corresponding bitrate (R_{minC}) and the distortion (D_{minC}) of the configuration of the lowest execution time are taken as anchor to compare with other configurations as shown in Tables II and III. Configuration #1 is considered to be of lowest complexity.

A comparison can be performed between configurations. For instance, to choose between configuration #6 and #7 that give approximately the same bitrate reduction, configuration #6 saves (2.27-1.88=0.39 (39%)) in complexity and the same can be done between configurations #4 and #5 and between configurations #1, #2 and #3. One can notice that the bitrate is mostly equivalent between some configurations although the complexity varies.

TABLE I: 7-DIFFRENT ENCODER CONFIGURATIONS

Parameters	Low			Medium		High	
	1	2	3	4	5	6	7
Coding Unit/Depth	16/1	16/2	32/2	32/2	32/3	64/3	64/4
Transform Unit min-max	2-2	2-3	2-3	2-3	2-4	2-4	2-5
Motion Search Range	32	32	32	64	64	Full	Full
IntraPeriod	8	8	8	16	16	32	32
- GOP is 8 (hierarchical B-Frames) with QP increased by one in each level - QP is 32 - Full: Full search mode							

TABLE II: THE AVERAGE TIME FOR EACH CONFIGURATION OVER 13 VIDEOS WITH THE BITRATE SAVINGS FOR ONE SEQUENCE

Config.	Average time (h:m:s)	Bitrate Mbps
1	3:52:4	4.9
2	5:55:11	5.1
3	5:22:22	4.9
4	5:51:13	2.6
5	8:5:54	2.5
6	7:26:42	1.5
7	9:1:4	1.5

TABLE III: AVERAGE-COMPLEXITY AND BITRATE SAVING FOR ONE SEQUENCE

Config.	Time saving factor	Bitrate saving factor
1	1.00	1.00
2	1.49	1.03
3	1.35	1.00
4	1.48	0.54
5	2.04	0.51
6	1.88	0.32
7	2.27	0.30

In order to judge the configuration against others, a linear combination of the three components, bitrate saving (R), distortion saving (D), and complexity saving (C), is established as an optimization criterion and expressed as in equation 1. This may be visualized in three dimensional space or using an equal side triangle, each angle represents a component, i.e. R, D, and C. Figure 2 illustrates the analysis space and the three coefficients (α, β, γ) need to be tuned to obtain the optimization criterion value. These coefficients are restricted to a sum of one. In other words, each point in the analysis space represents the contribution factor of each component of the optimization criterion, i.e. bitrate, distortion, and complexity and for the visualization purposes, these points represents the best configuration. Different real X-Y coordinates are generated and projected to the analysis space (equal side triangle) using the following matrix to obtain (α and β) and then $\gamma = 1 - \alpha - \beta$:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0.5 \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & -\frac{\sqrt{3}}{3} \\ 0 & \frac{2\sqrt{3}}{3} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$O = \alpha C + \beta R + \gamma D, \tag{1}$$

Where,
R: the ratio of bitrate to the C_{minC},
D: the ratio of distortion to the D_{minC},
C: the ratio of complexity to the R_{minC} , and
$\alpha + \beta + \gamma = 1$

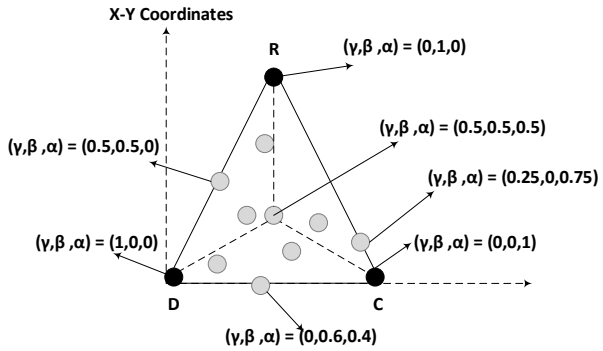


Figure 2: The analysis space

An example of applying the proposed visualization tool to the 13 video sequences using configurations mentioned in Table I is shown in Figure 3. Each point in the analysis space represents the best configuration. The selection of the best configuration is locally optimized within a limited bitrate and distortion range since the quantization parameter is fixed. It can be concluded from Figure 3 that:

- Configurations 2 and 3 should not be considered for sequence #1 as no gain is obtained for any rate-distortion-complexity operation point in comparison to the other configurations.
- Configuration 1 can be considered for less complexity, configuration 4 can be considered for balancing the three components, and configurations 6 and 7 can be considered for the best quality, for sequence #1.
- For sequence #13, configurations 6 and 7 cannot be used since better results are given with 4 and 5.
- There are sequences that behave alike which is important to note as it points to content properties similarities.

In general, the optimization criterion is critical as it acts as decision maker for which configuration should be selected. Changing this criterion may alter the selected configuration. Many changes can be done to this criterion. Firstly, the (log) function can be applied to the bitrate (R) and to the complexity (C) due to their logarithmical behavior. Secondly, another distortion measurement can be applied rather than the usual PSNR due to its limitation to human visual system (HVS).

IV. Discussion

It was shown in the previous section that the best configuration is obtained for specific values of rate, distortion, and complexity. Here, for instance, some targeted applications are mentioned.

- During encoding process, the encoder parameters can be altered to meet the application requirements. Bandwidth variations can be monitored and accordingly a suitable configuration can be selected. Specifically, in the encoder side, each sequence is attached with different quality levels, i.e. different visual diagrams each with different quantization parameter. Depending on the channel situation, the encoder has different possibilities. It may choose to stay on the current quality level, i.e current visual diagram, but with different bitrate, distortion, complexity level. Another choice is to change the quality level. The quality change can be done directly to the required level or can be gradually changed to the required level at the group of picture or intra period level. For instance, Figure 4 shows three different visual diagrams and the corresponding absolute values of rate, distortion, and quality ranges are shown in Table IV.
- The proposed visual diagram can be updated through encoding process. For instance, during the encoding of the first group of picture, it was notice that it takes much time. At that time, the encoder parameters can be changed for the next group of picture. This process can be applied as much as possible until the suitable configuration is selected.
- Content can be manually or automatically analyzed and characterized, and the best configuration for each class can be recommended or predicted. For instance, a small motion search range is recommended for sequences that labeled as low motion activity sequences. This application scenario is motivated by the fact shown in the previous section that some sequences have similar behavior and this behavior can be interpreted such that these sequences have similar content properties.

TABLE IV: Absolute value of visualization diagrams in Figure 4 for Sequence#10.

QP	R (Mbps)		D (dB)		C (H)	
	Min	Max	Min	Max	Min	Max
28	5	10.1	39.2	39.7	5	10.5
32	3	6.4	37.3	38	6	11.5
36	1.8	4.1	35.2	36.1	6.5	10

V. Conclusion

This work shows the impact of considering the complexity in the optimization process of video coding. The benefit is clear when two configurations yield similar bitrate and distortion, the proposed visualization tool indicates the one with lower complexity. More investigations are required to study the effect of changing the optimization criterion and using other distortion measurements. This tool is also valuable in real world applications since it chooses the best configuration for the encoding process.

Acknowledgment

This work is supported by the Marie Sktodowska-Curie under the PROVISION (PeRceptually Optimised VIdeo Com-

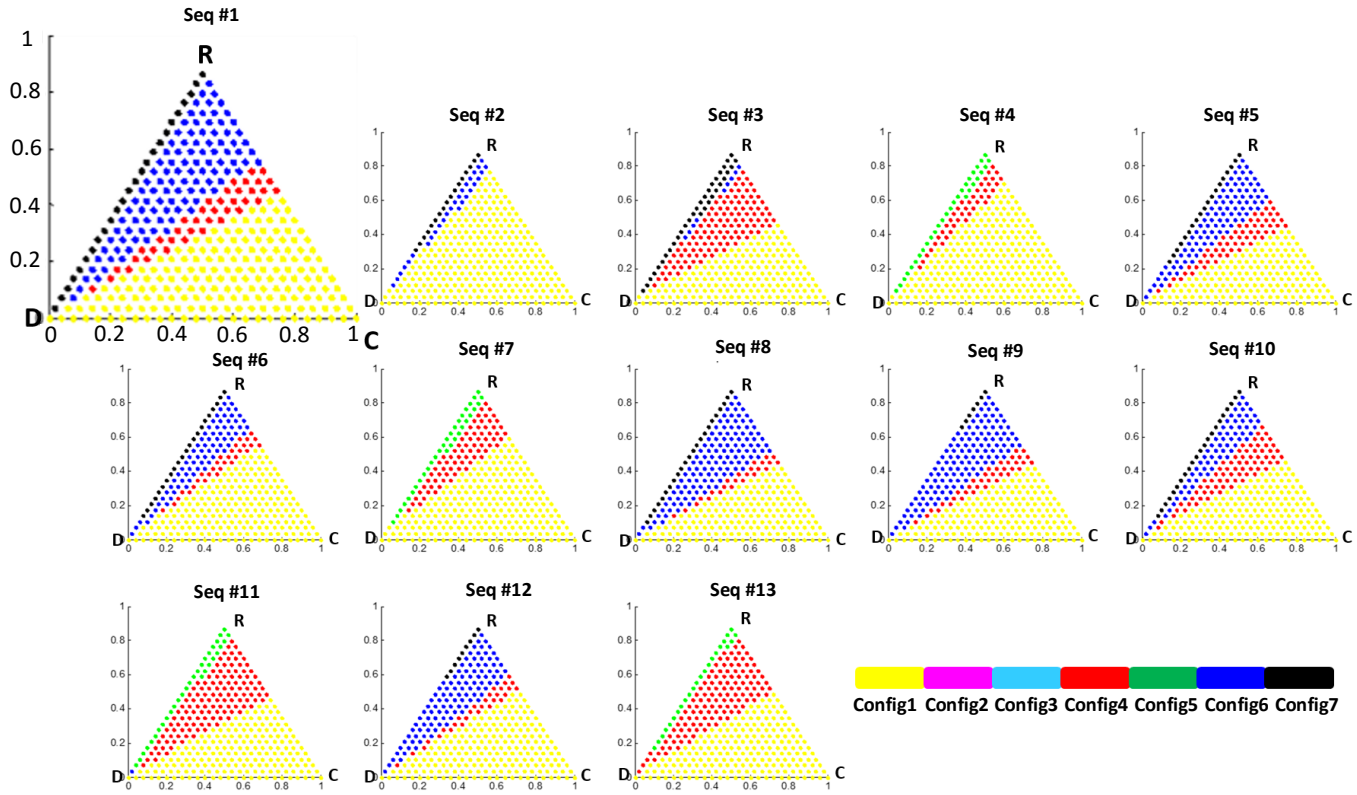


Figure 3: Visual analysis of optimization criterion with 7-encoding configurations for 13 sequences. The colors refer to the configurations as shown in Table I

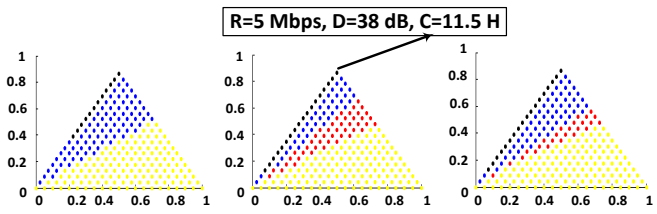


Figure 4: Optimization visualization for three different range of qualities for Sequence#10 "Traffic and building". Left: QP=28, middle: QP=32, and right:QP=36

presSION) project bearing Grant Number 608231 and Call Identifier: FP7-PEOPLE-2013-ITN.

REFERENCES

[1] G. J. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1649–1668, 2012.

[2] F. Bossen, B. Bross, K. Suhring, and D. Flynn, "HEVC complexity and implementation analysis," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1685–1696, 2012.

[3] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *Signal Processing Magazine, IEEE*, vol. 15, no. 6, pp. 74–90, 1998.

[4] B. Meng, O. C. Au, C.-W. Wong, and H.-K. Lam, "Efficient intra-prediction mode selection for 4× 4 blocks in H. 264," in *Multimedia and Expo, 2003. ICME'03. Proceedings. 2003 International Conference on*, vol. 3. IEEE, 2003, pp. III–521.

[5] B. La, M. Eom, and Y. Choe, "Fast mode decision for intra prediction in H. 264/AVC encoder," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 5. IEEE, 2007, pp. V–321.

[6] J.-H. Kim and B.-G. Kim, "Fast block mode decision algorithm in H. 264/AVC video coding," *Journal of Visual Communication and Image Representation*, vol. 19, no. 3, pp. 175–183, 2008.

[7] J. Leng, L. Sun, T. Ikenaga, and S. Sakaida, "Content based hierarchical fast coding unit decision algorithm for HEVC," in *Multimedia and Signal Processing (CMSP), 2011 International Conference on*, vol. 1. IEEE, 2011, pp. 56–59.

[8] X. Shen and L. Yu, "CU splitting early termination based on weighted SVM," *EURASIP Journal on Image and Video Processing*, vol. 2013, no. 1, pp. 1–11, 2013.

[9] L. Shen, Z. Zhang, and Z. Liu, "Effective CU size decision for HEVC intracoding," *Image Processing, IEEE Transactions on*, vol. 23, no. 10, pp. 4232–4241, 2014.

[10] Y. Pitrey, M. Barkowsky, R. Pépion, P. Le Callet, and H. Hlavacs, "Influence of the source content and encoding configuration on the perceived quality for scalable video coding," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012, pp. 82 911K–82 911K.

[11] L. Song, X. Tang, W. Zhang, X. Yang, and P. Xia, "The SJTU 4K video sequence dataset," in *QoMEX*, 2013, pp. 34–35.

[12] M. H. Pinson, M. Barkowsky, and P. Le Callet, "Selecting scenes for 2D and 3D subjective video quality tests," *EURASIP Journal on Image and Video Processing*, vol. 2013, no. 1, pp. 1–12, 2013.

[13] T. V. Q. E. Group, "The validation of objective models of multimedia quality assessment," PHASE I 2008.