A GLOBAL MODEL OF AVC/H.264 VIDEO ENCODERS

Tomasz Grajek and Marek Domański

Chair of Multimedia Telecommunication and Microelectronics
Poznań University of Technology, Poland, http:// multimedia.edu.pl

ABSTRACT

For AVC/H.264 video encoders, the paper describes a simple global model that defines a relationship between the bitrate and the quantization step Q_{step} for I-, P- and B-frames. Proposed is a general model for any video shot. As such a model must be dependent on video content, the model parameters have to be estimated from experimental data individually for each video shot, and individually for each frame type. The model is defined by formulas given individually for 3 intervals covering whole eligible range of Q_{step} values. Extensive experiments with numerous test video sequences have proved good accuracy of the model proposed.

Index Terms **—** compression, video coding, AVC, MPEG-4, H.264, video coder modeling.

1. INTRODUCTION

For video encoders, development of efficient control algorithms is still an open problem that permanently gains a lot of attention. The problem was already quite challenging for classic video encoders like MPEG-2 [2] but it became even more difficult for advanced video encoders like AVC/H.264 [1]. The main control parameter is quantization index *QP* that is directly related to quantization step Q_{step}. The latter parameter directly controls quantization, and therefore bitrate and quality of decoded video. Therefore adjusting the value of this parameter is used as a tool to control bitrate of the bitstream produced by the encoder. Unfortunately, there is no generic mathematical model that allows exact calculations of quantization step Q_{step} from given bitrate and video quality. Therefore, we are searching for a model that will define a bitrate B as a function of the quantization parameter Q_{step}. Of course, such a relation depends strongly on video content, so the model needs to take it into account.

Like other video coding standards, also AVC/H.264 suggests its own rate control algorithm [3] that obviously is not a normative part of the standard. It exploits computationally-intensive Rate-Distortion Optimization (RDO). In references, there are described many other approaches that exploit some modeling of video encoders. He and Mitra in [4] proposed different approach to rate control by introducing a linear ρ-domain source model, where ρ denotes percentage of zeros in quantized transform coefficients. It turned out to be very accurate in source content estimation hence several new rate control models have been developed based on their observations e.g. [5, 6]. For example, in [7], authors have proposed a rate-complexity-quantization model based on observations that coded bits have linear relationship with proposed frame complexity measure and exponential relationship with *QP* index. In [8], authors use three linear mathematical models to describe relationships between *QP* parameter, quality (PSNR used as a measurement metric) and bitrate.

In our proposal, an encoder is treated as a "black box" with one input (i.e. video sequence) and one output (bitstream). The encoder is controlled by only one parameter – quantization step Q_{step}. For this box we try to find appropriate mathematical relation between quantization step Q_{step} and bitrate B. Such a model would be very useful both in design of bitrate control algorithms as well as in teaching on advanced video compression.

2. PROPOSED MODEL

Our objective is to find a formula that defines number of bits per frame B as a function of quantization parameter Q_{step}

$$B = f(Q_{step}) \qquad (1)$$

The goal is to find a general type function $f(\cdot)$. This model $f(\cdot)$ will be established by analysis of experimental data. As such a model must be dependent on video content, model parameters are assumed to be estimated for a video shot individually for each frame type.

This approach is the same as the one already successfully used for MPEG-2 video encoders [9, 10]. In these papers, a simple global model of MPEG-2 bitstreams has been proposed. In that way an MPEG-2 bitstream is modeled with high accuracy using only one content-dependent parameter.

Assuming that similar considerations can be made with reference to AVC/H.264 standard, the authors applied the function fitting method [11, 12] to experimental data from AVC/H.264 reference encoder [13]. Because finding a good approximation of experimental data for the whole eligible range of Q_{step} values turned out to be much more difficult that for the MPEG-2, the entire eligible range of Q_{step} values has been divided into 3 intervals (Fig. 1), and the model has been defined individually in each interval. Practical motivation for such an approach is that the entire range of eligible values of Q_{step} is never used in a single application.

In the paper, we deal with 4CIF (704 × 576), 25 Hz video sequences. For such sequences, the eligible interval of Q_{step} <0.625, 224> has been divided into 3 intervals setting the borders on Q_{step} = 4.5 and Q_{step} = 64 (Fig. 1). The most interesting central interval ("2nd interval") covers Q_{step} values corresponding to bitrates from about 1 to 3 Mbit/s for test sequences. In practical applications, it is the most useful range of bitrates for 4CIF sequences.

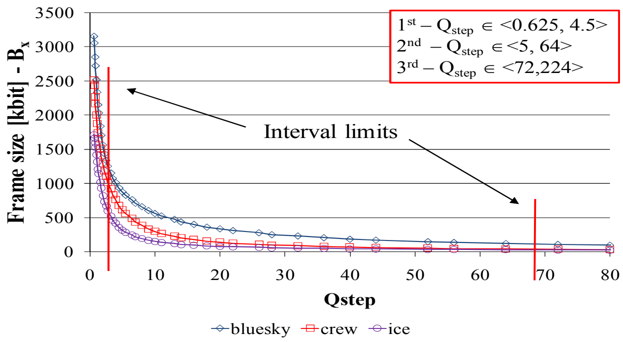

Fig. 1. The experimental curves for an I frame for 3 test sequences. For the sake of clarity, range of Q_{step} has been clipped to 80.

The research has been made on various 4CIF sequences with different motion characteristics. All sequences have been encoded with AVC/H.264 reference software version JM_13.2 [13] (main profile, CABAC and RDO enabled, GOP: IBBPBBPBBPBBP). Sequences *bluesky*, *pedestrian*, *riverbed*, *rushhour*, *station2*, *sunflower* and *tractor* have been cropped to 4CIF resolution from their original size - 720p (1080x720 pixels). For the sake of brevity, the results for only few sequences are reported in the paper but the research has been done using 21 test video sequences.

2.1. Model for the 2nd interval

Function fitting applied to the data from the 2nd interval resulted in hyperbolic model as follows:

$$B(Q_{step}, a, b, c) = \frac{a}{Q_{step}^{b} + c} \quad , \tag{2}$$

where a, b and c are model parameters that depend on sequence content and $B(Q_{step})$ is the average number of bits per frame for a given Q_{step} value. The parameters' values have been estimated by minimization of maximum approximation error over all 2nd interval of Q_{step}:

$$\varepsilon(Q_{step}, a, b, c) = \frac{\left|B_X(Q_{step}) - B(Q_{step}, a, b, c)\right|}{B_X(Q_{step})} * 100\% \quad , \tag{3}$$

$$\min_{a,b,c} \max_{Q_{step}} \varepsilon(Q_{step}, a, b, c) \, , \tag{4}$$

where $B_X(Q_{step})$ denotes measured number of bits per frame and $B(Q_{step},a,b,c)$ denotes the value calculated from the model (1).

Figs. 2,3 and 4 show experimental and approximated curves for 3 frame types for 3 exemplary sequences. Average relative error for most sequences is below 4%, 8% and 19% for I-, P- and B-frames, respectively.

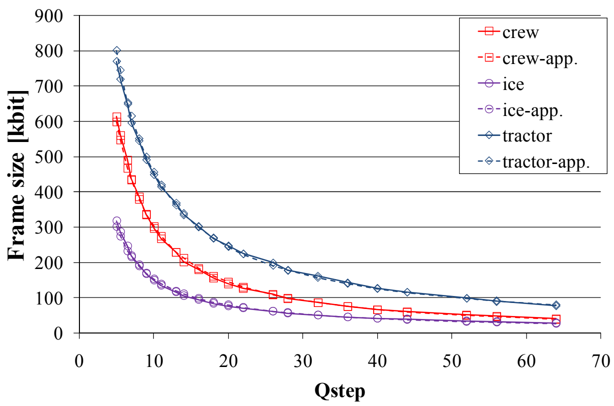

Fig. 2. Experimental and approximated curves for an I-frame for 3 exemplary test sequences (2nd interval).

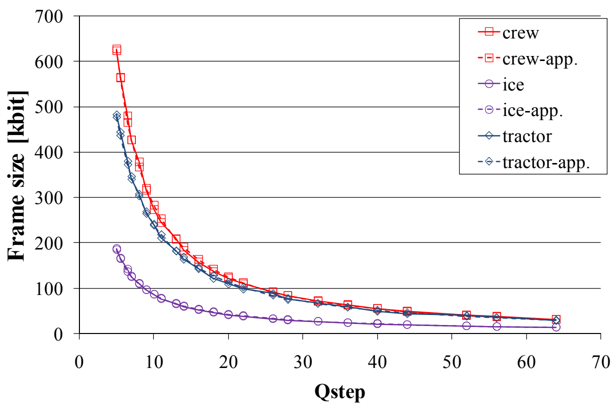

Fig. 3. Experimental and approximated curves for an P-frame for 3 exemplary test sequences (2nd interval).

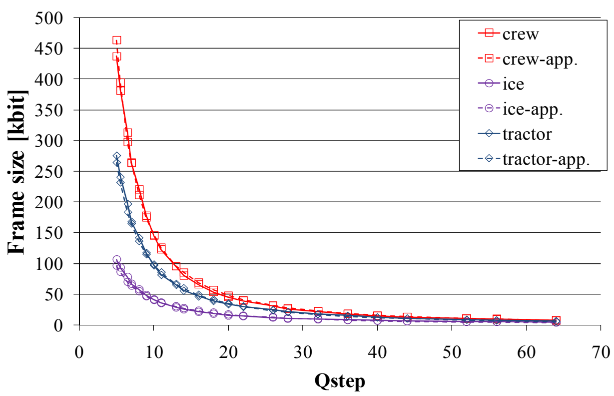

Fig. 4. Experimental and approximated curves for an B-frame for 3 exemplary test sequences (2nd interval).

2.2. Model for the 1st interval

Function fitting applied to the data from the 1st interval resulted in simpler cubic model with 4 parameters. Detailed analysis showed that there exists linear relationship between these parameters. Therefore, a model with only one free parameter d has been evaluated separately for each frame type (Eq. 5):

$$B(Q_{step}, d) = (a_1 d + a_2)Q_{step}^3 + (b_1 d + b_2)Q_{step}^2 + (c_1 d + c_2)Q_{step} + d \quad (5)$$

where a_1, a_2, b_1, b_2, c_1 and c_2 are model constants different for each frame type (see Table 1) and d is the model parameter.

Table 1. Values of model constants in 1st interval.

Const.	Frame type		
	I	P	B
a_1	-0.017	-0.0196	-0.0173
a_2	6385.8	9666.6	3133.3
b_1	0.1702	0.1786	0.16
b_2	-41606	-31199	25669
c_1	-0.5875	-0.5662	-0.5289
c_2	1451.1	-147954	-277773

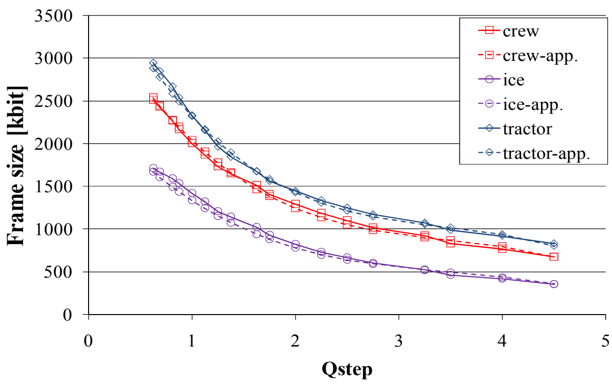

Fig. 5. Experimental and approximated curves for an I-frame for 3 exemplary test sequences (1st interval).

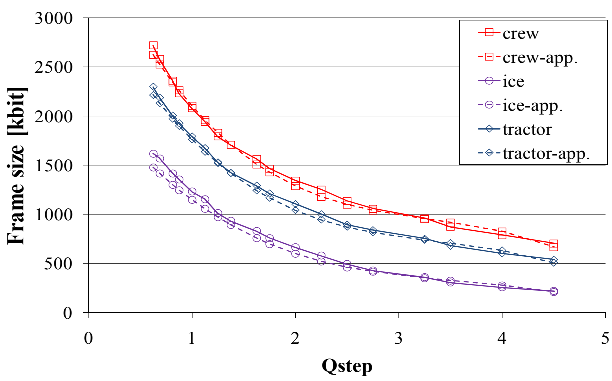

Fig. 6. Experimental and approximated curves for an P-frame for 3 exemplary test sequences (1st interval).

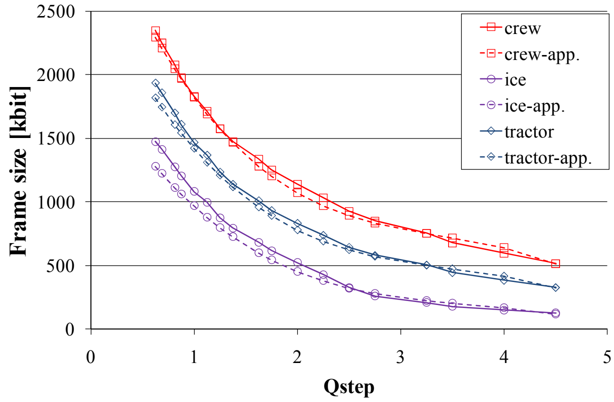

Fig. 7. Experimental and approximated curves for an B-frame for 3 exemplary test sequences (1st interval).

Fig. 5,6 and 7 show experimental and approximated curves for 3 frame types for 3 exemplary sequences. Average relative error for most sequences is below 2%, 4% and 7% for I-, P- and B-frames, respectively.

2.3. Model for the 3rd interval

Function fitting applied to experimental data from the 3rd interval resulted in quadratic model with 3 parameters.

Again, the model can also be simplified. A model with only one free parameter c and 4 universal constants has been found sufficient for good approximation (Eq. 6, Table 2). Average relative error for most sequences is below 5%, 7% and 20% for I-, P- and B-frames, respectively.

$$B(Q_{step}, c) = (a_1 c + a_2)Q_{step}^2 + (b_1 c + b_2)Q_{step} + c \quad (6)$$

Table 2. Values of model constants in 3rd interval.

Const.	Frame type		
	I	P	B
a_1	0.000017	0.000019	0.000026
a_2	0.10547	-0.02979	-0.05413
b_1	-0.00768	-0.00813	-0.01001
b_2	-13.545	13.794	14.695

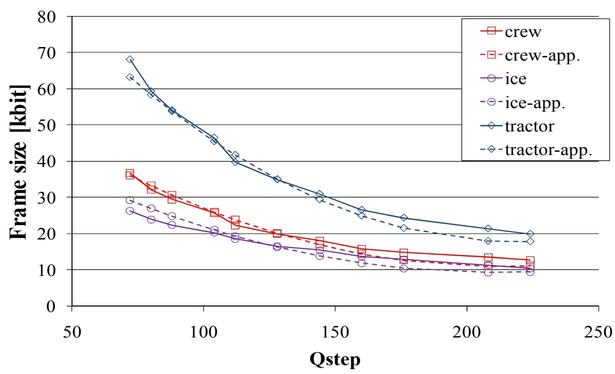

Fig. 8. Experimental and approximated curves for an I-frame for 3 exemplary test sequences (3rd interval).

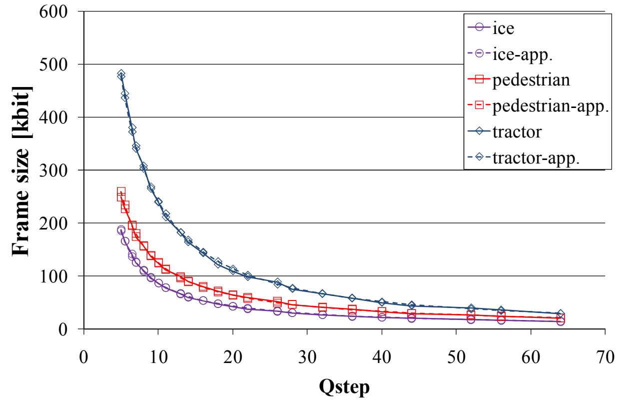

Fig. 9. Experimental and approximated curves for an P-frame for 3 exemplary test sequences (3rd interval).

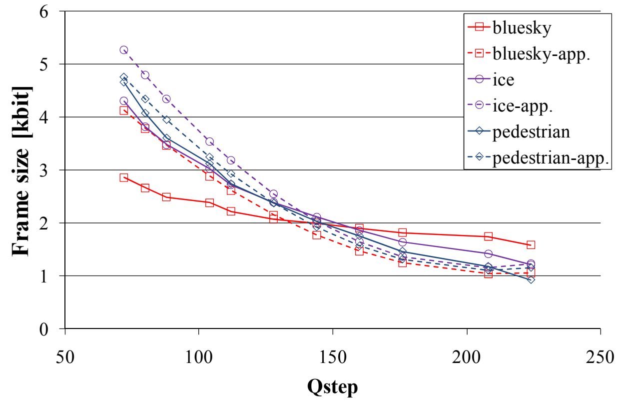

Fig. 10. Experimental and approximated curves for an B-frame for 3 exemplary test sequences (3rd interval).

3. CONCLUSIONS

A global model of AVC/H.264 coders has been described. The research started with careful analysis of experimental data gathered for 21 video test sequences with 200 frames each. These sequences were: *bluesky, city, crew, harbour, ice, pedestrian, riverbed, rushhour, soccer, station2, sunflower, tractor, basket, bus, cheer, flow, football, icon, stefan, universal* and *warner.* The analysis of this huge set of experimental data resulted in proposal of the function type that is able to fit well the experimental data in individual intervals of quantization step, and for different picture types. For two intervals, only one model parameter is needed that depends on sequence content. For these intervals, other parameters have been estimated as universal constants.

This model can be used to set a value of the quantization parameter *QP* for a given number of bits for an I- P- or B-frame. Tests showed that it fits experimental data very well in all intervals. For most sequences relative approximation error is lower than 5% for I-frames, 8% for P-frames and below 20% for B-frames. However sequences with motion characteristics like in *station2* and *city* have bigger errors, what may be caused by reduction of number of parameters in 1st and 3rd intervals.

The model estimates average number of bits per frame of a given type (i.e. I-, P- or B-frame). Therefore the content-dependent model parameters should be calculated individually for video shots limited by scene cuts.

Acknowledgements. This work was supported by the public funds as a research project in year 2009.

4. REFERENCES

[1] ISO/IEC FDIS 14496-10: *Information Technology – Coding of Audio-Visual Objects – Part 10: Advanced Video Coding*

[2] ISO/IEC 13818-2: *Information technology – Generic Coding of Moving Pictures and Associated Audio Information -- Part 2: Video*

[3] Wiegand, T., Schwarz, H., Joch, A., Kossentini, F., Sullivan, G.J.: "Rate-Constrained Coder Control and Comparison of Video Coding Standards", IEEE Trans. Circ. Sys. Video Techn., vol. 13, no. 7, pp. 688-703, (2003)

[4] He, Z., Mitra, S.K.: "Optimum Bit Allocation and Accurate Rate Control for Video Coding via ρ-Domain Source Modeling", IEEE Transactions on Circuits and Systems for Video Technology, vol. 12, no. 10, (2002)

[5] Lin, G., Zheng, S., Hu, J.: "A Two-Stage ρ-Domain Rate Control Scheme for H.264 Encoder", IEEE International Conference on Multimedia and Expo ICME'2008, pp. 713--716, (2008)

[6] Lim, S.C., Na, H.R., Lee, Y.L.: "Rate Control Based on Linear Regression for H.264/MPEG-4 AVC", Signal Processing: Image Communication, Volume 22, Issue 1, pp. 39--58, Elsevier (2007)

[7] Zhou, Y., Sun, Y., Feng, Z., Sun, S.: "New Rate-Complexity-Quantization Modeling and Efficient Rate Control for H.264/AVC", IEEE International Conference on Multimedia and Expo ICME'2008, (2008)

[8] Zhuo, L., Gao, X., Wang, Z., Feng, D.D., Shen, L.: "A Novel Rate-Quality Model based H.264/AVC Frame Layer Rate Control Method", 6th Int. Conference on Information, Communications & Signal Processing, Singapore (2007)

[9] Łuczak, A.: Digital *Video Bitrate Control Using Coder Models*, Doctoral Dissertation, Poznań (2003)

[10] Domański, M., Łuczak, A.: "Simple Global Model of an MPEG-2 Bitstream", Proceedings of IEEE International Conference on Multimedia and Expo ICME'2002, (2002)

[11] Bronshtein, I.N., Semendyayev, K.A., Musiol, G., Muehling, H.: *Handbook of Mathematics*, 4th edition, Springer – Verlag, New York (2004)

[12] Brandt, S.: *Data Analysis: Statistical and Computational Methods for Scientists and Engineers*, 3rd ed., Springer – Verlag, New York (1999)

[13] Fraunhofer Institute for Telecommunications Heinrich-Hertz-Institut Image Processing Department, http://iphome.hhi.de/suehring/tml/index.htm