A Simple Quantitative Model of AVC/H.264 Video Coders

Tomasz Grajek and Marek Domański

Poznań University of Technology
Chair of Multimedia Telecommunication and Microelectronics
ul. Polanka 3, 60-965 Poznań, Poland
{tgrajek, domanski}@et.put.poznan.pl

Abstract. The paper describes a simple quantitative model of AVC/H.264 coders. The model defines the relationship between the bitstream and the quantization step (*Qstep*) for I- and P-frames. The whole allowed range of *Qstep* values has been divided into 3 intervals. In 1st and 3rd interval, the proposed model has only one parameter that depends on sequence content, whereas in 2nd interval the proposed model has four parameters that depend on sequence content. The experiments have been conducted on 4CIF sequences and showed that proposed model fits experimental data very well in all intervals.

Keywords: compression, video coding, AVC, MPEG-4, H.264, video coder modeling.

1 Introduction

Despite of hundreds of millions of video coders working worldwide, designing efficient control algorithms is still an open problem that permanently gains a lot of attention. The problem was already quite difficult for classic video encoders like MPEG-2 [2], but it has become even more severe with emerging of the new generation of advanced video encoders. Among them, there are the state-of-the-art video encoders that are compliant with ITU-T H.264 and MPEG-4 AVC standards [1], called also briefly AVC/H.264 encoders. In particular, bitrate control is quite difficult for such video coders.

Recently, the bitrate control problem became even more crucial because of wide proliferation of video streaming in communication networks with rapidly varying throughput. It is related to numerous applications of wireless video transmission. In such applications, we need bitrate control techniques that are capable to cope with rapid variations of the available channel throughput. The techniques have to influence the Video Compression Layer in such a way that the number of produced bits will match the currently available channel throughput.

Similarly to prior video compression standards, AVC/H.264 [1] does not standardize any rate control algorithm (some suggestions can be found in informative part of the standard). Therefore, encoders' designers can freely optimize their rate control algorithms that have essential influence on encoding efficiency and performance.

The basic parameter that can be used to control an encoder is the quantization parameter Q_{step} that defines quantization step for transform coefficients. A typical goal of adjusting the parameter Q_{step} is to match the available channel bitrate. The quantization parameter Q_{step} may be adjusted on frame level, slice level as well as on macroblock level. The results of this paper are relevant to global bitrate control, i.e. bitrate control on the frame and slice levels only.

Unfortunately, there does not exist an universal quantitative mathematical model that allows for exact calculations of coder parameters from given bitrate and video quality. In particular, 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 would need to take it into account.

Similar problems have been already considered by several authors [3-11]. A very brief review of these solutions will be given in the next section. Nevertheless, none of those references proposes a simple statistical model of AVC/H.264 coder similar to that proposed in this paper.

Moreover, we assume that the relation $B(Q_{step})$ is stable in time, so the model of the currently encoded frame may be deduced from the previous frames. Such an assumption has been justified by many empirical observations [3, 6, 10].

2 Advanced Video Coder Modeling

In informative part of AVC/H.264 standard, mode selection algorithm is based on Rate-Distortion Optimization (RDO). The algorithm needs to know a value of quantization index QP (directly related to Q_{step}), but it can be determined based on the mean absolute difference (MAD) only after optimization. To cope with this problem a linear model for MAD prediction has been proposed [3]. Then quadratic rate-quantization model is used to calculate QP value [4]. Although, RDO results in a “chicken and egg” dilemma and makes rate control algorithm more complicated, many existing rate control schemes are based on these two models.

Different approach to rate control presented He and Mitra in [5] by proposing a linear ρ-domain source model, where ρ denotes percentage of zeros in quantized transform coefficients. Because it turned out to be very accurate in source content estimation, several new rate controls models have been developed based on their observations e.g. [6, 7].

Other proposals how to avoid a “chicken and egg” dilemma are described in [8] and [9]. In [8], 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 [9], authors use three linear mathematical models to describe relationships between QP parameter, quality (PSNR used as a measurement metric) and bitrate.

Our proposal is to treat a coder as a “black box” with one input (sequence) and one output (bitstream), controlled by only one parameter – Q_{step} and find appropriate mathematical description of its behavior, which allowed us to estimate a given bitrate with high accuracy.

3 Proposed Model

Our objective is to find the relationship between the number of bits B and quantization parameter Q_{step} for a given frame type in a given sequence. The relation will be established by analysis of experimental data. To collect these data, we encoded many test sequences with various, but constant Q_{step} values and calculated frame size as an average number of bits needed to encode frame over frames of the same type. Of course, each value of Q_{step} was set indirectly by properly setting the value of quantization index *QP*.

In [10, 11], a simple global model of MPEG-2 bitstream has been proposed. It was created by function fitting to experimental data [12, 13]. On the assumption that similar model may be obtained for AVC/H.264 standard, the authors applied the function fitting method to experimental data obtained from AVC/H.264 encoder. Because finding a good approximation of experimental data for the whole allowed range of Q_{step} values turned out to be much more difficult that for the MPEG-2 data, the authors decided to divide the experimental curves into 3 intervals (Fig. 1) and perform function fitting in each one separately. These intervals are:

- 1st – $Q_{step} \in$ <0.625, 2.25>;
- 2nd – $Q_{step} \in$ <2.5, 104>;
- 3rd – $Q_{step} \in$ <112,224>;

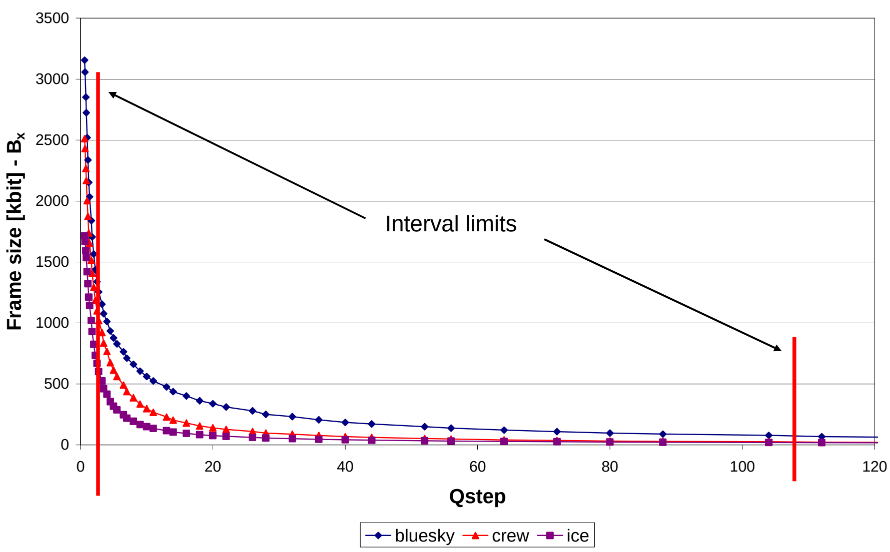

Fig. 1. The experimental curves for an I frame for 3 test sequences. For the sake of clarity, range of Q_{step} has been clipped to 120.

All research has been made on various 4CIF (704x576 pixels) sequences with different motion characteristics. All sequences have been encoded with AVC/H.264 reference software version JM_13.2 [14] (main profile, CABAC and RDO enabled). Sequences *bluesky, pedestrian, riverbed, rushhour, station2, sunflower* and *tractor* have been cropped to 4CIF resolution from their original size - 720p (1080x720 pixels).

3.1 Model for the 1st Interval

Function fitting applied to the data from the 1st interval resulted in quadratic model as follows:

$$B(Q_{step})=a*Q_{step}^{2}+b*Q_{step}+c, \tag{1}$$

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

$$\varepsilon(Q_{step},a,b,c)=\frac{\left|B_X(Q_{step})-B(Q_{step},a,b,c)\right|}{B_X(Q_{step})}*100\%, \tag{2}$$

$$\min_{a,b,c}\max_{Q_{step}}\varepsilon(Q_{step},a,b,c), \tag{3}$$

where $B_X(Q_{step})$ denotes measured value and $B(Q_{step}, a, b, c)$ denotes the approximated one.

Further, detailed analyses showed that there is linear relationship between those parameters. Therefore, model with only one free parameter has been evaluated separately for I- (Eq. 4) and P-frames (Eq. 5).

$$B(Q_{step})=[(0.1169*c)-26042.8]*Q_{step}^{2}+[(-0.549*c)+8228.2]*Q_{step}+c, \tag{4}$$

$$B(Q_{step})=[(0.1417*c)-41521.7]*Q_{step}^{2}+[(-0.58*c)-58830.1]*Q_{step}+c, \tag{5}$$

Fig. 2 shows experimental and approximated curves for an I-frame for 3 exemplary sequences. Parameter's values, maximum and average errors for all analyzed sequences have been shown in Table 1. Average relative error for most sequences is below 2% and 3 % for I- and P-frames, respectively.

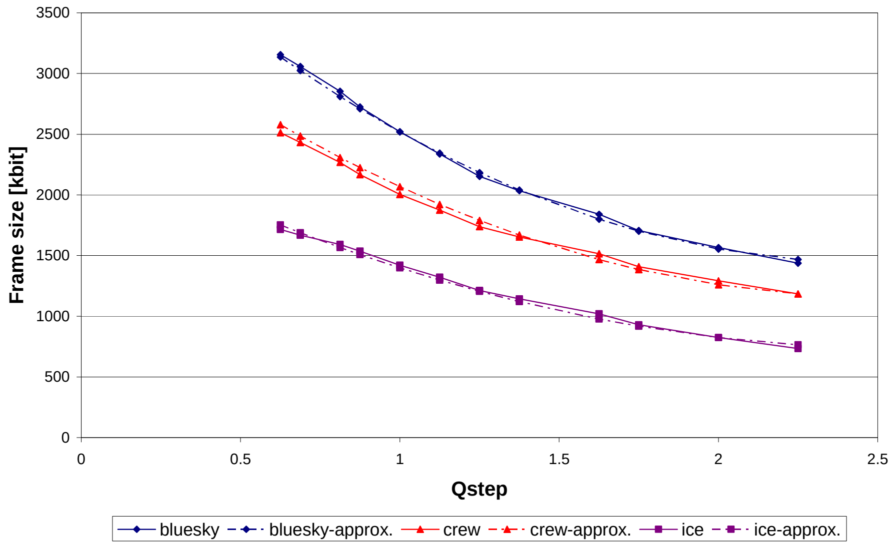

Fig. 2. Experimental and approximated curves for an I-frame for 3 exemplary test sequences.

Table 1. Estimated parameters, maximum and average error for I- and P-frames (1st interval).

Sequence	I frames			P-frames		
	c	max err	avg. err	*c*	max err	avg. err
basket	5 116 391	2.23	0.80	4 007 522	3.94	2.01
bluesky	4 477 530	2.23	0.94	2 633 699	4.29	2.06
bus	3 889 417	1.78	0.80	3 044 718	4.31	2.83
cheer	4 774 414	2.50	1.37	3 927 236	3.82	2.48
city	4 483 201	1.99	0.92	3 675 306	3.81	2.03
crew	3 682 216	3.15	2.17	3 975 767	3.43	1.95
flow	6 214 351	2.91	1.62	3 587 493	3.17	1.85
football	3 102 568	4.28	2.13	3 209 110	3.92	2.67
harbour	4 599 594	2.02	1.06	4 091 791	3.10	2.27
ice	2 507 134	4.20	1.86	2 316 404	5.14	3.13
icon	1 858 923	3.70	2.42	2 040 221	5.63	4.16
pedestrian	2 699 532	3.25	1.32	2 437 282	5.29	2.84
riverbed	4 127 412	2.37	1.15	4 485 542	3.89	1.88
rushhour	2 481 953	6.50	3.97	2 313 825	9.35	5.06
soccer	3 794 015	1.90	0.88	3 583 624	3.36	2.04
station2	3 542 499	2.79	1.40	2 203 927	12.07	6.71
stefan	3 876 281	1.97	0.85	3 144 886	4.14	2.56
sunflower	3 399 590	3.63	2.08	2 150 666	11.35	5.76
tractor	4 143 423	2.73	1.35	3 357 022	3.86	1.43
universal	1 660 066	3.38	1.69	1 927 193	8.61	6.35
warner	1 900 413	2.52	0.93	2 473 887	4.28	2.89

3.2 Model for the 2nd Interval

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

$$B(Q_{step}) = \frac{a}{c * Q_{step}{}^{b} + d}, \qquad \textbf{(6)}$$

where a, b, c and d are real constants that depend on sequence content and $B(Q_{step})$ is the number of bits per frame for a given Q_{step} value. The parameters' values can be estimated by minimization of maximum approximation error similarly as in section 3.1. Fig. 3 shows experimental and approximated curves for an I-frame for 3 exemplary sequences. The values of all four parameters, maximum and average errors for all analyzed sequences have been shown in Table 2 and Table 3. For I-frames average relative error for most sequences is lower than 4% and for P-frames is below 7%. However, sequences *city* and *station2* have bigger average relative errors for P-frames mainly due to very complicated content and motion characteristic.

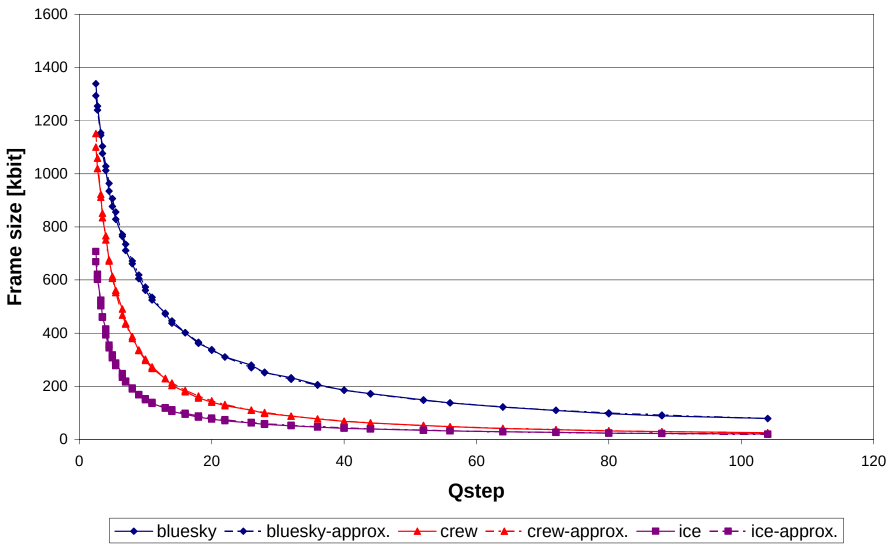

Fig. 3. Experimental and approximated curves for an I-frame for 3 exemplary test sequences.

Table 2. Estimated parameters, maximum and average error for I-frames (2nd interval).

Sequence	*a*	*b*	*c*	*d*	max err	avg. err
basket	425 309	0.9733	0.0553	0.1423	3.41	1.53
bluesky	1 045 098	0.9528	0.1541	0.4391	3.39	1.64
bus	431 254	1.0313	0.0710	0.1767	4.41	2.22
cheer	351 458	0.8916	0.0688	0.0776	2.81	1.34
city	540 958	1.1974	0.0479	0.2404	8.04	4.88
crew	836 141	1.0928	0.2116	0.1505	4.63	1.94
flow	379 723	0.9348	0.0413	0.1065	4.10	1.75
football	1 242 636	0.9924	0.5728	-0.0903	3.68	1.83
harbour	423 828	1.0265	0.0609	0.1443	3.92	2.01
ice	247 208	0.8350	0.2703	-0.2312	5.70	3.09
icon	388 630	1.2537	0.1155	0.4223	5.32	2.68
pedestrian	353 370	1.0105	0.1566	0.0869	3.19	1.52
riverbed	498 196	1.1560	0.0605	0.2557	6.11	3.27
rushhour	1 010 005	0.8252	0.8970	-0.4215	5.00	2.34
soccer	1 220 150	1.2884	0.1171	0.7268	7.01	3.64
station2	518 880	1.3068	0.0603	0.3216	6.94	3.79
stefan	1 060 521	0.8663	0.3098	0.2146	2.94	1.29
sunflower	388 893	0.9322	0.1117	0.1337	3.85	1.84
tractor	1 150 525	1.0652	0.1709	0.5024	3.77	1.96
universal	895 121	0.9454	0.8267	0.0146	3.69	1.87
warner	249 888	0.9509	0.2104	0.0138	10.23	5.55

Table 3. Estimated parameters, maximum and average error for P-frames (2nd interval).

Sequence	*a*	*b*	*c*	*d*	max err	avg. err
basket	1 064 274	1.2572	0.1369	0.5022	4.17	1.89
bluesky	824 428	1.2466	0.2969	0.3092	5.39	2.62
bus	958 129	1.2107	0.2708	0.4073	8.10	3.94
cheer	989 202	1.0059	0.2379	0.2493	3.11	1.33
city	272 749	1.8128	0.0136	0.3473	36.67	18.60
crew	200 163	1.1975	0.0425	0.0396	5.50	2.62
flow	934 342	1.5041	0.0800	0.7578	14.29	7.02
football	1 198 311	1.0383	0.5353	-0.0357	3.76	2.05
harbour	603 166	1.5378	0.0331	0.4212	9.88	4.82
ice	941 494	0.9022	1.5221	-1.6175	6.33	3.15
icon	225 684	1.0130	0.2146	-0.0171	13.12	6.85
pedestrian	998 135	0.8926	1.1675	-0.9283	6.86	2.76
riverbed	1 070 022	1.1669	0.1159	0.5051	5.01	2.28
rushhour	1 012 675	1.0798	0.8858	-0.3082	4.59	1.79
soccer	204 503	1.1473	0.0546	0.0345	7.79	3.87
station2	1 162 131	1.0569	2.7667	-4.6550	19.91	11.39
stefan	251 886	1.3048	0.0606	0.1099	9.10	5.37
sunflower	247 466	1.1594	0.2157	-0.0724	10.47	5.51
tractor	849 928	1.1578	0.2239	0.2882	5.11	2.18
universal	957 089	1.0087	0.8367	0.1828	5.96	3.28
warner	282 082	0.9732	0.1783	0.0045	3.84	1.89

3.3 Model for the 3rd Interval

Similarly to results achieved in the section 3.1, function fitting applied to data from the 3rd interval resulted in quadratic model (Eq. 1). Detailed analyses showed that the model can also be simplified. Therefore, model with only one free parameter has been evaluated separately for I-(Eq. 7) and P-frames (Eq. 8).

$$B(Q_{step})=[(0.000014*c)-0.00199]*Q_{step}^{2}+[(-0.00677*c)+11.173]*Q_{step}+c, \qquad (\mathbf{7})$$

$$B(Q_{step})=[(0.000013*c)-0.01656]*Q_{step}^{2}+[(-0.00671*c)+10.5]*Q_{step}+c, , \qquad (\mathbf{8})$$

Fig. 4 shows experimental and approximated curves for an I-frame for 3 exemplary sequences. Parameter's values and obtained maximum and average errors for all analyzed sequences have been shown in Table 4.

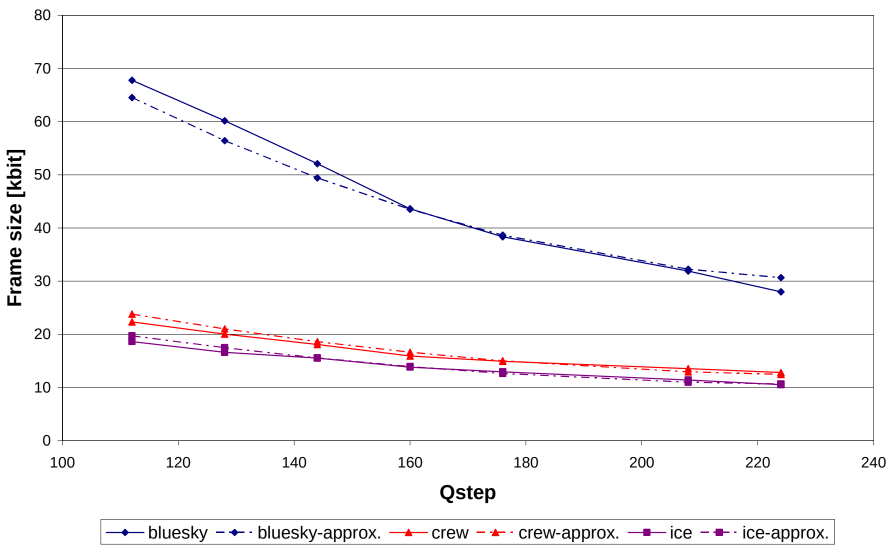

Fig. 4. Experimental and approximated curves for an I-frame for 3 exemplary test sequences.

For I- and P-frames average relative error for most sequences is lower than 5%. However, several sequences have bigger relative errors, what can be caused by reduced number of parameters.

Table 4. Estimated parameters, maximum and average error for I-and P-frames (3rd interval).

Sequence	I frames			P-frames		
	a	**max err**	**avg. err**	***a***	**max err**	**avg. err**
basket	164 627	2.64	1.26	48 465	5.38	3.36
bluesky	151 564	6.63	3.85	18 255	7.31	5.22
bus	97 923	3.00	1.77	28 755	4.22	2.49
cheer	183 918	9.21	5.39	85 862	7.82	3.84
city	62 512	18.44	9.79	12 569	5.80	3.76
crew	54 052	6.22	4.02	31 859	7.04	4.30
flow	239 994	3.28	1.61	24 500	4.12	3.00
football	47 987	6.53	4.11	40 081	7.30	4.42
harbour	105 654	7.37	3.95	20 642	20.98	11.24
ice	44 285	5.46	2.97	20 936	7.42	3.50
icon	19 213	2.47	1.22	23 482	6.31	4.16
pedestrian	39 074	1.64	1.11	29 271	7.35	4.34
riverbed	64 046	11.11	5.45	76 236	5.58	3.33
rushhour	49 008	2.44	1.56	11 068	15.74	9.12
soccer	46 043	5.30	2.70	42 461	8.14	3.95
station2	27 029	20.10	11.27	438	65.67	43.85
stefan	121 209	2.80	2.14	20 814	4.17	2.01
sunflower	101 826	9.65	5.68	9 701	5.71	3.39
tractor	94 891	3.88	2.29	30 638	17.88	11.81
universal	29 158	9.79	5.59	22 693	4.26	2.76
warner	31 949	5.04	2.56	33 976	3.92	2.50

4 Conclusions

A simple quantitative model of AVC/H.264 coders has been described. This model can be used to set a value of the quantization parameter Q_{step} for a given number of bits for an I- or P-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 and below 7% for P-frames. However sequences with specific motion characteristics like in *station2* and *city* sequences exhibit larger approximation errors. Some experiments show that these errors can be reduced by using additional parameters in the model for 1st and 3rd intervals.

The model may be used in global procedures for bitrate control. Previously encoded pictures may be used to identify the model parameters. Usually, the model parameters are nearly constant in time, mostly even in longer temporal intervals. In that way, the model may be used in order to explicitly calculate quantization parameter Q_{step} for the assumed bitrate. Such a technique may be used in bitrate control techniques that are appropriate for video streaming in the communication channels with rapid variations of the channel throughput, i.e. wireless channels.

Acknowledgements. This work was supported by the public funds as a research project in years 2007-2009.

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. Li, Z., Pan, F., et al.: Adaptive Basic Unit Layer Rate Control for JVT, JVT-G012, 7th meeting, Thailand (2003)
4. Chiang, T., Zhang, Y.-Q.: A New Rate Control Scheme Using Quadratic Rate Distortion Model, IEEE Transactions on Circuits and Systems for Video Technology, vol. 7, pp. 246–250, Feb. (1997)
5. 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)
6. 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, June 23-26, Hannover, Germany (2008)
7. 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)
8. 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, pp. 717--720, June 23-26, Hannover, Germany (2008)
9. 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 International Conference on Information, Communications & Signal Processing, Singapore (2007)
10. Łuczak, A.: Digital Video Bitrate Control Using Coder Models, Doctoral Dissertation, Poznań (2003)
11. Domański, M., Łuczak, A.: Simple Global Model of an MPEG-2 Bitstream, Proceedings of IEEE International Conference on Multimedia and Expo ICME'2002, August 26-29, Lausanne, Switzerland (2002)
12. Bronshtein, I.N., Semendyayev, K.A., Musiol, G., Muehling, H.: Handbook of Mathematics, 4th edition, Springer – Verlag, New York (2004), in Polish: Matematyka, poradnik encyklopedyczny, Wydawnictwo Naukowe PWN, wydanie 20, Warszawa (2004)
13. Brandt, S.: Data Analysis: Statistical and Computational Methods for Scientists and Engineers, 3rd ed., Springer – Verlag, New York (1999), in Polish: Analiza danych, metody statystyczne i obliczeniowe, Wydawnictwo Naukowe PWN, wydanie 2, Warszawa (1999)
14. Fraunhofer Institute for Telecommunications Heinrich-Hertz-Institut Image Processing Department, http://iphome.hhi.de/suehring/tml/index.htm