MODELING OF BITSTREAMS PRODUCED BY HYBRID VIDEO CODERS

Adam Łuczak, Marek Domański

Poznań University of Technology, Institute of Electronics and Telecommunications,
ul. Piotrowo 3a, 60-965 Poznań, Poland
E-mail: { aluczak, domanski } @ et.put.poznan.pl

ABSTRACT: The paper describes a simple global model of MPEG-2 video coders. This global model is defined as a relation between the bitrate and the quantization parameter for individual I-, P- and B-frames. The model has five parameters that depend on image content. Nevertheless it is shown that four of them can be fixed in order to obtain simply one-parameter model which can be used for encoder control. The experimental results are reported for some video test sequences in the BT.601/4CIF format encoded by an MPEG-2 MP@ML coder.

KEYWORDS: MPEG-2, rate control, video coder modeling.

1. INTRODUCTION

Many international standards for digital video data compression have been established hitherto. One of the most popular standards is MPEG-2 [1,2] that is widely used in digital media storage as well as for digital television transmission. Similarly as other video compression standards, the MPEG-2 standard is mostly a generic description of semantics and syntax of the bitstream and it does not describe an algorithm which could be used to set the encoder parameters in order to obtain required output bitstream or video quality.

The video coding part of MPEG-2 describes a buffer control algorithm which is only a part of the bitrate and distortion control mechanisms. Therefore each implemented encoder may use a different algorithm in order to obtain required bitstream or quality of sequence. In the encoder, there exist several variables which can be set in order to control the encoding process, i.e. also the bitrate and video quality. For example, macroblock type, maximum motion vector length or quantization factor Q are the important control parameters. The last parameter is extremely important as it directly influences the quantization of the DCT coefficients. The number of nonzero transform coefficients and their values define the number of bits needed for encoding the transform coefficients. These bits constitute the main component of the bitstream. Therefore, the quality of the control algorithm deeply influences the overall coding performance of the video coding system.

Efficient control of video coders is still an open problem that gains a lot of attention [3] because of its importance for industry and communications. In the references, various approaches have been already described [3-10]. Some methods use Lagrangian multiplier techniques [e.g. 4-6,9] which require large computational complexities. Other techniques exploit picture models [e.g. 8].

The goal of the control algorithm is to maximize the video quality by a given bitrate constrains or minimize the bitrate by assumed quality. The control unit estimates the control parameter, and first of all the quantization parameter Q. There exist two types of control:

- Global control, i.e. adjusting the Q value in a frame header,
- Local control, i.e. adjusting the Q value in the individual macroblock headers.

Global control algorithm estimates the average (global) value GQ of the parameter Q for whole frames. Local control modifies the values of Q in certain range around GQ in order to match the encoding process to local statistics within a frame. Moreover local control algorithm help use output buffer in optimal way.

The paper addresses the first problem i.e. the problem of global control. The goal is to establish a simple empirical model for a bitrate as function of Q set for individual frames. The model should be valid for typical conditions of operation of a coder. In the paper, the experimental results are given for I, P and B-frames. The model can be used for global control in the constant bitrate mode of operation. For the sake of simplicity, absence of adjustment of Q in the individual macroblocks is assumed.

The importance of global bitrate control can be highlighted by a simple example. Assume that the coder encodes a flat portion of an image that can be easily encoded by use of quit small number of bits. The value of Q is being slowly decreased in the flat area. After exceeding an edge of the textured region, the value of Q shows to be too small because of unacceptably high bitrate. Therefore, the control unit tries to increase the value of Q step by step (Fig. 1). During this process, some portion of the available bit budget was allocated in an improper way.

Global control algorithm estimates the average GQ (global Q) and helps to avoid this situation. Local modifications of the parameter Q around the average value GQ can be done within certain range only.

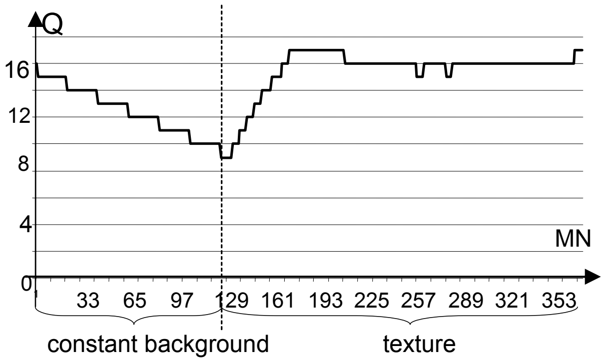

Fig. 1. Variations of the quantization parameter Q versus the macroblock number MN in a picture of the test sequence *Basket*.

2. BITSTREAM STRUCTURE

The total number of bits B allocated to an individual frame can be partitioned in the following way:

$$B = B_{\mathrm{CONST}} + B_{\mathrm{VAR}}(Q) \quad (1)$$

where B_{CONST} – the portion that does not depend on Q,

$B_{\mathrm{VAR}}(Q)$ – the portion that depends on Q.

In an I-frame, there is

$$B_{\mathrm{CONST}} = B_{\mathrm{CTR}} + B_{\mathrm{YDC}} + B_{\mathrm{CDC}} \quad (2)$$

where B_{CTR} denotes the number of bits for the frame, slice and macroblock headers and B_{YDC} and B_{CDC} denote the number of bits for intra-DC coefficients in the luminance and the chrominance component, respectively.

Similarly, in P- and B-frames, there is

$$B_{\mathrm{CONST}} = B_{\mathrm{CTR}} + B_{\mathrm{MV}}, \quad (3)$$

where B_{MV} denote the number of bits needed to encode motion vectors.

For a given sequence, all the components of B_{CONST} are relatively similar for various frames both for I, P- and B-frames (Figs. 2 and 3, Table 1).

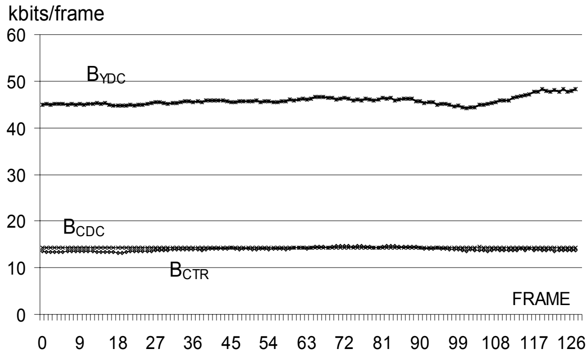

Fig. 2. The bitstream components that are independent from Q in I-frames of the test sequence *Basket*.

For various frames of three test sequences, the simulation results are summarized in Table 1 for the bitstream components that are independent from Q. The standard deviation of the bitrate calculated for 128 frames remains small.

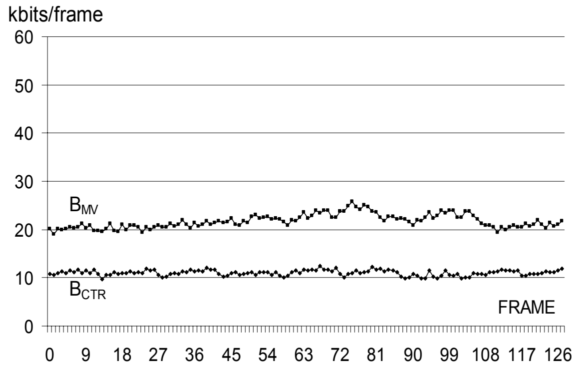

Fig. 3. The bitstream components that are independent from Q in P-frames of the test sequence *Basket*.

Table 1. The bitstream components for I- and P-frames that are independent from Q in various frames of three test sequences.

	bits/frame				
Frame	I-frames			P-frames	
Flower garden					
	B_{CTR}	B_{YDC}	B_{CDC}	B_{CTR}	B_{MV}
1	14356	43895	14193	11120	9847
16	14355	45007	14837	10808	9887
32	14394	45694	14940	9784	9250
1-128	14364	45521	15279	10075	9650
Std.dev.	13	527	37	632	296
Cheer					
1	14383	43849	17010	7594	8310
16	14394	44196	17363	4848	7317
32	14395	43823	16936	7145	8198
1-128	14383	43714	16892	5674	7724
Std.dev	12	361	263	832	389

Stefan					
1	14364	37783	12536	8870	10801
16	14372	37626	12370	8915	12122
32	14377	38173	12583	8879	11127
1-128	14382	38330	12532	8842	11324
Std.dev.	11	439	146	162	698

Table 2. The bitstream components for B-frames that are independent from Q in various frames of three test sequences.

	bits/frame					
	B-frames					
	Flower garden		Cheer		Stefan	
Fr.	B_{CTR}	B_{MV}	B_{CTR}	B_{MV}	B_{CTR}	B_{MV}
1	12044	19812	10898	19985	10477	18765
16	11557	21735	11566	21067	10432	17597
32	11950	20505	11483	19890	10688	16959
1-128	12121	22385	11522	19614	10682	18162
S.dev.	320	889	311	797	226	1136

Therefore, the control unit of the encoder is able to estimate B_{CONST} using the value of B_{CONST} calculated for a previous frame. In particular, the value of B_{CTR} is almost constant for I-frames.

3. GENERAL BITSTREAM MODEL

Let us consider the bitstream component that depends on the frame quantization factor Q.

$$B_{VAR}(Q) =, \; B_{YV}(Q) + B_{CV}(Q) \quad (4)$$

where $B_{YV}(Q)$ and $B_{CV}(Q)$ denote the bits needed for encoding of the DCT coefficients (except the intra-DC ones) in the luminance and the chrominance, respectively. In the MPEG-2 bitstream, the nonzero DCT coefficients are encoded using Huffman variable length codes. Nevertheless the average codeword length is almost constant when calculated over sufficiently large set of image blocks [8]. In the "codeword number per macroblock" and "bits per macroblock" coordinates, the two–dimensional histogram (Fig. 4) has its maxima along a line defined as

$$\textit{Number of bits} \approx 6.05 \cdot \textit{Number of codewords} \quad (5)$$

The above number of bits plus the number of bits needed for the EOB codes (codes for *End of block*) is equal to $B_{YV}(Q)$ or $B_{CV}(Q)$ for luminance and chrominance, respectively. The number of DCT-codewords is equal to the number of nonzero DCT coefficients. The latter one strongly depends on quantization factor Q. Therefore the values of $B_{YV}(Q)$ and $B_{CV}(Q)$ can be modeled statistically as functions of Q. The reason for that is that a 4CIF frame contains a sufficiently large set of blocks.

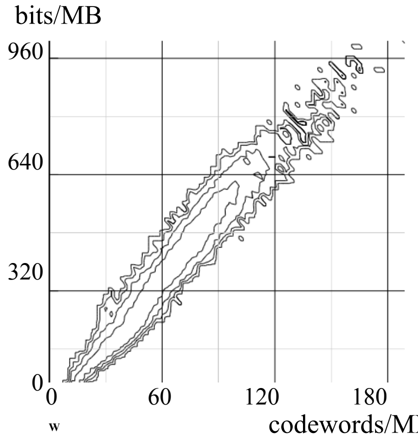

Fig. 4.Two-dimensional histogram "codewords – bits" for the test sequence *Basket*.

In order to match experimental data the following model has been chosen for $B_{YV}(Q)$ and $B_{CV}(Q)$

$$B_x(Q) = \frac{a}{c(Q^b + e) + d} \quad , \quad (6)$$

where *a, b, c, d* and *e* are real constants that depend on the frame content. The parameter values can be estimated by approximation error minimization over the interval <1,31> of the allowed values of Q. Precision of approximation of the experimental data B_e is measured by

$$\varepsilon = \frac{|B_e - B_x|}{B_e} \cdot 100\% \quad , \quad (7)$$

and ε_{max} means the maximum value of ε and ε_{ave} means the average value of ε.

The parameters *a, b, c, d* and *e* can be estimated for a given frame from a given test sequence. The results for I- and P-frames can be found in Tables 3 and 4. Similar results have been obtained for B-frames as well as for the chrominance components.

Table 3. Parameters for the luminance in I-frames.

Frame	Parameters					ε_{ma}	ε_{ave}
	$a/10^6$	*b*	*c*	*d*	*e*		
Flower garden							
0		0.91	2.06	0.11	0.87	7.9	2.9
32	6.4	0.91	2.05	0.11	0.88	6.8	2.9
64	6.5	0.91	2.05	0.11	0.89	7.1	2.6
96	3.3	0.86	2.28	0.10	0.86	6.2	2.3
Cheer							
0	1.4	0.95	0.61	0.06	0.71	7.6	2.8
32	1.3	0.93	0.63	0.05	0.68	8.2	2.8
64	1.4	0.95	0.64	0.05	0.67	8.0	2.9
96	1.3	0.94	0.63	0.05	0.67	8.3	2.9
Stefan							
0	1.9	0.89	0.59	0.05	0.38	7.0	2.3
32	1.1	0.88	0.66	0.05	0.44	5.8	2.0

64	1.1	0.86	0.67	0.05	0.38	6.2	2.2
96	1.1	0.85	0.67	0.05	0.37	5.5	1.9

Table 4. Parameters for the luminance in P-frames.

Frame	Parameters					ε_{max}	ε_{ave}
	$a/10^6$	b	c	d	e		
Flower garden							
0		1.07	0.90	0.53	0.56	13.8	4.9
32	2.1	1.31	0.47	0.77	0.49	15.6	8.2
64	1.2	1.20	0.31	0.39	0.03	13.7	8.5
96	2.2	1.18	0.62	0.62	0.10	14.7	8.7
Cheer							
0	2.3	1.24	0.56	0.56	0.47	7.3	2.8
32	2.1	1.14	0.63	0.66	0.41	15.8	4.2
64	2.0	1.26	0.58	0.70	0.42	9.9	4.1
96	2.0	1.13	0.61	0.64	0.41	12.6	4.6
Stefan							
0	1.7	1.11	0.59	0.67	0.42	4.2	2.2
32	1.9	1.09	0.64	0.76	0.41	10.6	5.6
64	1.9	1.22	0.73	0.68	0.41	13.8	5.6
96	2.1	1.23	0.64	0.92	0.42	8.6	4.3

If the parameters are estimated for individual frames, the model $B_x(Q)$ approximates the experimental data $B_e(Q)$ so well that often it is difficult to distinguish the experimental and the modeled curve (Fig. 4).

4. BITSTREAM MODEL WITH ONE FREE PARAMETER

For the bitrate control, it would be very practical to have a model with only one parameter representing the frame content. Therefore it was assumed that only the parameter c depends on frame content, while the other parameters assume general values that can be found by minimization of the approximation error $|B_e\text{-}B_x|$ over a set of frames from some training sequences. Parameter c was left as a free variable which depend on frame content. The minimization was performed in a window $Q_0 \pm dQ$ for various Q_0 and dQ (Fig. 5). Precision of approximation has been measured with different set of sequences (Tables 5-7).

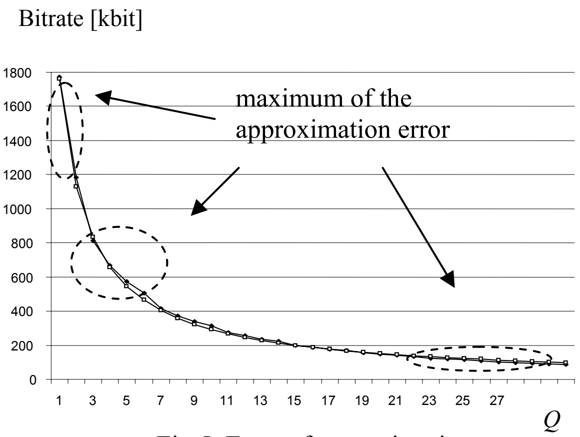

Fig.5. Error of approximation.

As we can see approximated curve fits well to experimental data. Maximum approximation error for is usually located Q near value 31. For large values of Q, the bitrates are very small as compared to the whole range of bitrates. Moreover, also in the middle of range of Q, the approximation error may be significant because many ESCAPE codes may be generated there.

Table 5. Precision of approximation for a general model with one free parameter for I-frames.

Seq.	ε_{max} [%]			ε_{ave} [%]		
.	dQ=2	dQ=4	dQ=8	dQ=2	dQ=4	dQ=8
Flower garden – luminance						
Q_0=10	0.9	1.8	5.9	0.4	0.8	2.0
Q_0=16	0.6	1.1	2.6	0.4	0.4	0.8
Q_0=22	0.6	1.0	1.6	0.3	0.4	0.5
Cheer– luminance						
Q_0=10	0.9	3.6	9.4	0.5	1.3	3.3
Q_0=16	0.9	1.7	5.3	0.5	0.6	1.6
Q_0=22	0.8	0.9	2.3	0.3	0.3	0.6
Stefan– luminance						
Q_0=10	1.0	3.3	4.9	0.6	1.0	1.8
Q_0=16	1.0	2.2	4.9	0.4	0.6	0.9
Q_0=22	1.2	2.6	5.0	0.7	1.1	1.4
Flower garden - chrominance						
Q_0=10	0.9	3.6	10.9	0.4	1.1	3.9
Q_0=16	2.0	2.6	3.4	0.9	1.0	1.8
Q_0=22	2.4	1.7	3.07	1.1	0.7	0.9
Cheer - chrominance						
Q_0=10	4.0	7.6	21.3	1.9	2.5	3.6
Q_0=16	2.8	5.6	8.2	1.2	.2.1	3.6
Q_0=22	4.9	3.7	7.7	2.3	1.7	3.5
Stefan - chrominance						
Q_0=10	6.0	10.2	22.9	2.3	2.5	5.2
Q_0=16	4.0	7.3	18.2	1.3	2.0	4.3
Q_0=22	2.8	5.0	10.6	1.0	1.8	2.2

Table 6. Precision of approximation for a general model with one free parameter for P-frames.

	ε_{max} [%]			ε_{ave} [%]		
Seq.	dQ=2	dQ=4	dQ=8	dQ=2	dQ=4	dQ=8
Flower garden - luminance						
Q_0=10	8.2	8.5	28.9	1.5	2.3	4.9
Q_0=16	6.9	8.1	16.9	1.3	3.1	4.9
Q_0=22	3.3	6.3	7.9	0.8	1.35	2.4
Cheer- luminance						
Q_0=10	7.5	19.3	29.9	1.7	2.9	5.5
Q_0=16	8.6	17.9	36.7	2.6	4.8	8.4
Q_0=22	8.5	13.6	19.9	2.4	3.9	5.9
Stefan- luminance						
Q_0=10	8.8	8.9	21.0	1.7	2.2	5.8
Q_0=16	10.9	9.3	16.6	2.5	2.8	4.0
Q_0=22	11.9	16.5	24.4	1.52	3.0	5.1
Flower garden - chrominance						
Q_0=10	14.4	29.5	63.6	4.9	5.5	8.8
Q_0=16	23.0	27.8	36.7	6.6	6.2	10.9
Q_0=22	15.6	24.2	40.9	3.6	8.1	11.8
Cheer- chrominance						
Q_0=10	.19.7	22.3	43.2	4.9	6.3	8.9
Q_0=16	20.5	25.4	34.8	3.7	6.5	8.7
Q_0=22	14.7	20.8	26.2	4.4	5.9	7.1
Stefan- chrominance						
Q_0=10	19.0	33.25	101.8	6.9	9.4	16.3
Q_0=16	39.8	33.7	43.2	7.1	8	12.9
Q_0=22	26.7	23.6	68.7	5.4	6.8	14.7

The dynamic range of bitrates is much larger for B-frames as compared to the cases of I- and P-frames. Therefore the errors of approximation have been defined for the B-frames in another way.

$$\varepsilon' = \frac{|B_e - B_x|}{B_{TOT}} \cdot 100\%$$

Where B_{TOT} means the whole bitstream value for encoded frame.

Table 7. Precision of approximation for a general model with one free parameter for B-frames.

Sequance	ε'$_{max}$ [%]			ε'$_{ave}$ [%]		
	dQ=2	dQ=4	dQ=8	dQ=2	dQ=4	dQ=8
Flower garden - luminance						
Q_0=10	0.00	0.05	0.48	0.00	0.00	0.04
Q_0=16	0.07	0.17	0.14	0.01	0.03	0.03
Q_0=22	0.07	0.14	0.18	0.01	0.04	0.03
Cheer- luminance						
Q_0=10	0.00	0.02	0.07	0.00	0.00	0.02
Q_0=16	0.01	0.02	0.03	0..00	0.00	0.00
Q_0=22	0.00	0.01	0.06	0.00	0.00	0.04
Stefan- luminance						
Q_0=10	0.01	0.05	0.21	0.00	0.00	0.03
Q_0=16	0.02	0.04	0.12	0.00	0.00	0.02
Q_0=22	0.02	0.04	0.04	0.00	0.00	0.04

5. IMPLEMENTATION

Proposed statistical model was implemented in MPEG-2 encoder for bitrate control. In experiment encoder had modified only global Q. Structure of GOP was set to sequence IBBPBBPBBPBB. Figure 6 shows changes of Q from frame to frame and changes of PSNR.

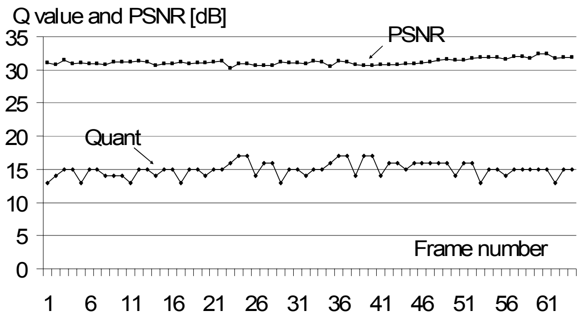

Fig.6. Example of performing global control algorithm with use of proposed model.

Test was performed for several progressive video sequences and for bitrate range <3Mbit, 12Mbit>. Without local control we obtain high accuracy of required bitrate with an error of bitstream rate less then 1%.

6. CONCLUSIONS

The paper presents a simple global model that well approximates the experimental values of bitrate as a function of the quantization factor Q. The model was tested for I-, P- and B-frames. It was shown that the version of this model with one free parameter can be efficiently used for bitrate control.

ACKNOWLEDGEMENT

The work was financed by the Committee for Scientific Research as the research project in the years 2002-2003.

REFERENCES

[1] ISO/IEC IS 13818, “Information Technology - Generic Coding of Moving Pictures and Associated Audio Information”.

[2] B. Haskell, A. Puri, A. N. Netravali, *Digital Video: An Introduction to MPEG-2*, New York: Chapman & Hall, 1997.

[3] G.J. Sullivan, T. Wiegand, “Rate-distortion optimization for video compression,” *IEEE Signal Processing Magazine*, vol. 15, pp. 74-90, Nov. 1998.

[4] J.H. Choi, D.C. Park, “A stable feedback control of the buffer state using Lagrange multiplier method” IEEE Trans. Image Processing, vol. 3, pp. 546-558, Sept. 1994.

[5] W.Y. Lee, J.B. Ra, “Fast algorithm for optimal bit allocation in rate-distortion sense, Electronics Letters, vol. 32, pp. 1871-1873, Sept. 1996.

[6] T. Wiegand, B. Girod, “Lagrange multiplier selection in hybrid video coder control,” *Proc. IEEE Int. Conf. Image Proc.*, pp. 542-545, Oct. 2001.

[7] W. Ding, B. Liu, rate control of MPEG-2 video coding and recording by rate-quantization modeling,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 12-20, June 1996.

[8] T. Kim, B. Roh, J. Kim, “An accurate bit-rate control for real-time MPEG video encoder,” *Signal Processing: Image Communication,* vol. 15, pp. 479-492, 2000.

[9] L. Wang “Rate control for MPEG video coding”, *Signal Processing: Image Communication* 15, pages 493-511, 2000.

[10] Y. Saw, M. Grant, J. Hannh, B. Mulgrew “Video rate control using a radial basis function estimator for constant bit-rate MPEG coders”, *Signal Processing: Image Communication* 13, pages 183-199, 1998.