RATE CONTROL FOR MULTI-SEQUENCE H.264/AVC COMPRESSION

Andrzej Pietrasiewicz and Grzegorz Pastuszak

Institute of Radioelectronics, Warsaw University of Technology, Nowowiejska 15/19, 00-665 Warsaw, Poland

Keywords: Video Compression, Rate Control, H.264/AVC.

Abstract: Multi-sequence video coding allows bit-budget to be distributed among sequences. This paper presents the method of selection of a common quantization parameter, which is applied concurrently to each sequence. The approach takes into account ρ-domain rate-distortion models kept independently for each video sequence and builds a common model. The output buffer is verified jointly for all the sequences and drives a joint bit allocation process. The method has been verified in simulation to demonstrate its usefulness in video encoding.

1 INTRODUCTION

Statistical multiplexing allows better utilization of available bandwidth for the transmission of several video sequences in the common channel. This feature is useful in such applications as broadcasting and video streaming over networks. The accurate control of the size of the output stream involves using sophisticated algorithms to perform this task in reasonable time. The bit allocation process in H.264/AVC encoder (ISO/IEC, 2003) is performed only by the selection of the quantization parameter Qp. One of the most important elements while controlling the process is an RD model. The model built in the domain of Mean Absolute Difference is non-linear and the accuracy is not high enough (Chiang, 1997). The RD model built in the domain of the parameter ρ denoting the percentage of zero quantised transform coefficients (He, 1996, Bobinski, 2004, Pietrowcew, 2005) provides much better results in terms of estimation accuracy, robustness, and complexity.

In this paper, the rate control based on ρ-domain is examined for single- and multi-sequence H.264/AVC encoding. The proposed rate control takes advantage of rate-distortion modelling based on ρ-domain and improves the methods for bit allocation and buffer verification inherited from the G012 rate-control (Li, 2003) used in the JM reference model. The usefulness of the multi-sequence approach is proved in simulations.

The rest of the paper is organized as follows. Section 2 reviews the rate-control algorithm based on the ρ-domain for coding a single frame. Section 3 describes the rate-control algorithm adopted to process several sequences concurrently. In particular, subsections analyze functional modules constituting the joint rate control. Section 4 presents simulation results, and the paper is concluded in Section 5.

2 RATE CONTROL BASED ON THE LINEAR MODEL

The purpose of the rate control is the adjustment of compression parameters in such a way, that bandwidth consumption is maximized but does not exceed a given limit. Also, the rate-control algorithm should react to achieve smooth quality changes and to prevent overflow and underflow of the output buffer. If the buffer fullness is high, it means that latest frames have utilized more bit-budget than assigned. Consequently, rate-control should allocate less bits for the following frames. In the opposite case, more bits can be assigned to the following frames. As statistics for I, P, and B frames differ, the bit allocation should take into account variable complexity weights computed separately for each frame type. Also, the RD model should have separate instances for each frame type to provide a reasonable prediction.
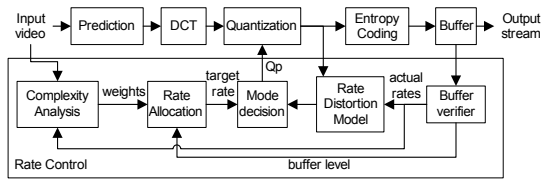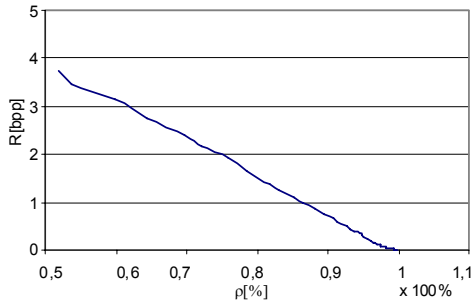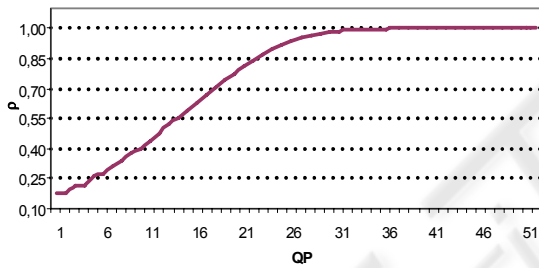
Figure 1: Rate control modules in the video encoder.

Figure 2: Dependence of bit-rate R on percentage of zero coefficients ρ.

Figure 3: Dependence of percentage of zero coefficients ρ on quantization parameter Qp.

The rate control is achieved by the modification of the value of the quantization parameter Qp, which trades bit-rate for quality. In Fig. 1, the modules of the rate control are shown with reference to main blocks of the video encoder.

The concept of rate-control based on the ρ-domain is shown in Fig. 2 and 3. The rate-distortion model counts the number of zero transform coefficients remaining after quantization and normalizes it to the total number of coefficient. It has been shown that in typical video coding systems the dependency between rate R and the percentage of zero coefficients ρ is linear, as can be seen in Fig. 2. This observation can be expressed as:

$$R(\rho) = \Theta(1 - \rho) \qquad (1)$$

The slope θ is modelled on the base of the previously encoded frame and is given by the formula:

$$\Theta = \frac{R_{prev}}{1 - \rho_{prev}} \qquad (2)$$

Parameters R_{prev} and ρ_{prev} denote the bit-rate and the zero fraction in the previous frame, respectively. The second dependency of the RD model keeps values of the parameter ρ calculated for all QP values (see Fig. 3). Thus, the selection of Qp for the next frame amounts to finding Qp for which percentage of zero coefficients ρ matches that calculated from the equation (1). To create the mapping between ρ and Qp used for the next frame encoding, the encoder has to apply all possible quantization parameters Qp to each block of transformed coefficients in the current frame. Note that this process repeats the forward quantization in the loop to count zero-valued coefficients.

3 MULTI-SEQUENCE RATE CONTROL

The purpose of multi-sequence rate control is to adjust compression parameters in such a way, that joint bandwidth consumption is maximized but does not exceed a given limit. Additionally, it is desirable to balance quality between sequences by removing limits on the bit rate assigned to each single sequence.

3.1 Joint Complexity Analysis

It is assumed that encoding for all sequences uses a common periodic pattern of frames and the same frame rate. Therefore, corresponding frames in each sequence make up a composite frame of the same type for the purpose of the rate control. The consistency of frame types allows the demonstration of the rate-control efficiency for the worst-case conditions.

Complexity weights W_X for the j-th sequence, where X corresponds to either I, P, or B, are computed based on the quantization parameter Qp used for the last coded frame of a given type and the actual number of utilized bits for that frame:

$$Wx_j = \frac{R_{j,i,X} * 2^{Qp_X/6}}{S_j} \qquad (3)$$

where i denotes the frame number and S_i is the frame area (width x height). Unlike in the G012 version, where weights are proportional to Qp. Instead, the weights depends exponentially on Qp normalized to six. This reflects the fact that doubling of the quantization step size is performed when Qp

is increased by six, which statistically leads to decreasing the actual bit rate by half. Average complexity weights used in the G012 version are not needed in the presented rate control.

Complexity weights for composite frames take into account area S_i of a single frame from the i-th sequence:

$$W_X = \frac{\sum_{j=0}^{NumberOfSequences-1} S_j * Wx_j}{\sum_{j=0}^{NumberOfSequences-1} S_j} \qquad (4)$$

3.2 Joint Buffer Verifier

The buffer verifier keeps track of the occupancy of the output buffer, which receives codestreams from several video encoders concurrently and releases joint stream (e.g., transport stream) at a given rate (e.g., channel bandwidth). Thus, after coding i-th frame, the buffer occupancy (level) BL_i is:

$$BL_i = BL_{i-1} + \sum_{j=0}^{NumberOfSequences-1} R_{j,i-1} - \frac{ChannelBitRate}{FrameRate} \qquad (5)$$

where $R_{i,j}$ denotes the number of bits utilized to code a given frame in the j-th sequence. The desired occupancy should be close to zero. Although the occupancy can assume negative value in that approach, the real implementation will have positive values by the introduction of a delay for removal of codestreams from the buffer.

For each P or I frame, the buffer occupancy is checked, and the target buffer level is updated. After coding of the first frame (I frame) in a GOP, the buffer occupancy may be considerably far from zero due to the inaccurate RD model (i.e., statistics for I-frames are updated relatively rarely). The deviation is distributed among the remaining frames in the GOP. Therefore, the target buffer level TBL_i is determined after coding of the i-th P frame and the following B frames as follows:

$$TBL_i = TBL_{i-1} - \frac{BL_0}{N_P} \qquad (6)$$

where N_P and BL_0 denotes the total number of P frames in the GOP and the buffer level after coding of the first frame in the GOP, respectively. Note that TBL_0 is equal to BL_0.

Due to changes in video content, the buffer occupancy deviates from the target buffer level. Thus, the rate control should compensate for this changes. Particularly, the deviation is taken into account to determine the target rate resulting from the buffer verifier:

$$T_{buffer} = \frac{ChannelBitRate}{FrameRate} + \gamma(TBL_i - BL_i) \qquad (7)$$

where γ is a constant that determines the strength of the buffer regulation. In the G012 version, the constant is equal 0.75 when there is no B frames and 0.25 otherwise.

3.3 Joint Rate Allocation

Allocation of bits for the multi-sequence coding is similar to that used in the G012 version. However, the proposed allocation refers to bit-rates and complexity weights computed for composite frames. Joint rate allocation is performed with reference to the hierarchy of frames. On the top level, there is a Group of Pictures (GOP), which is a contiguous block of frames from an I frame, inclusive, up to the next I frame, exclusive. On the second level, GOP consists of sections of pictures including one I or P frame and B frames following in the decoding order. The third level distinguishes single frames.

Before encoding each composite GOP, the bit budget for this GOP is estimated as a quotient of channel bit rate and frame rate. The quotient denotes an ideal number of bits per a composite frame, which, when multiplied by this GOP length, yields the bit budget under Constant Bit Rate (CBR) conditions:

$$R_i = N_{GOP} \frac{ChannelBitRate}{FrameRate} + R_{i-1} \qquad (8)$$

N_{GOP} and R_i denote the number of frames in GOP and the number of allocated bits for the i-th GOP in the sequence, respectively. The number of allocated bits is decreased by the actual number of utilized bits $R_{i,p}$ after coding of each (indexed by p) P frame and associated B frames:

$$R_i = R_i - R_{i,p} - \sum_b R_{i,p,b} \qquad (9)$$

After coding of the entire GOP, the remainder from the equation (9), which may be negative, is used to allocate bits for the next GOP (equation 8). During processing of a GOP, the number of remaining bits is allocated to P frames based on complexity weights of composite frames as follows:

$$T_P = \frac{W_P}{W_P N_P + W_B N_B} R_i \qquad (10)$$

where N_P and N_B denote the number of P and B frames remaining to code, respectively. Finally, the target rate for a given P frame is computed as a weighted average from the allocated bits and the target rate from the buffer verifier:

$$T_P = \beta * T_P + (1 - \beta) * T_{buffer} \qquad (11)$$

where β is a constant equal to 0.5 when there are no B frames and 0.9 otherwise. The number of bits for B frames is allocated in a similar way as for P frames:

$$T_B = \frac{W_B}{W_p(N_p-1)+W_B N_B}(R_i - T_P) \qquad (12)$$

3.4 Joint RD Model

Before encoding each I or P frame and associated B frames, a joint ρ-domain model is updated. The joint model is calculated with reference to models assigned to each sequence. This procedure takes into account the area of a single frame in a given sequence. Separate models are calculated for all three frame types. Thus, in the following equations, the index X is to be substituted by either I, P, or B. The mapping between the quantization parameter Qp and the fraction of zero-valued coefficients is calculated using the following formula:

$$\rho_X[Qp] = \frac{\sum_{j=0}^{NumberOfSequences-1} \rho_{j,X}[Qp] S_j}{\sum_{j=0}^{NumberOfSequences-1} S_j} \qquad (13)$$

for Qp in the range from 0 to 51. The fraction of zero-valued coefficients for the previous composite frame of a given type is calculated using the following formula:

$$\rho_X = \frac{\sum_{j=0}^{NumberOfSequences-1} \rho_{j,X} S_j}{\sum_{j=0}^{NumberOfSequences-1} S_j} \qquad (14)$$

The slope Θ for a given frame type is calculated using the following formula:

$$\theta_X = \frac{\sum_{j=0}^{NumberOfSequences-1} \theta_{j,X}(1-\rho_{j,X}) S_j}{(1-\rho_X)\sum_{j=0}^{NumberOfSequences-1} S_j} \qquad (15)$$

The RD model can be kept only for luma coefficients. In this case, the target rate assigned to a frame to be coded is scaled down according to the weight of the luma component. The target rate is used to find the final quantization parameter Qp applied to a composite frame (to all sequences). Qp is determined in the similar way as in the case of single-sequence coding.

Qp calculated from the RD model is verified with reference to previous frames. First, it is assumed that Intra frames have Qp not greater than that for the previous P frame. Second, it is assumed

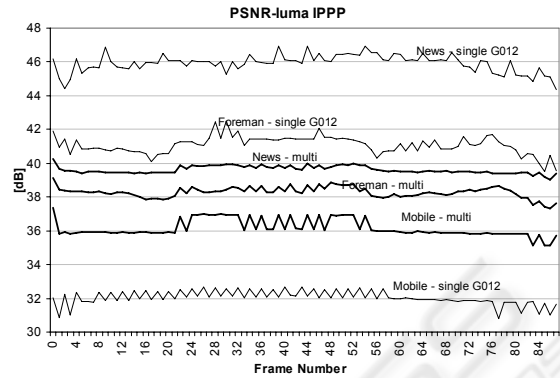that B frames have Qp not lower than that for the last I/P frame in the decoding order.

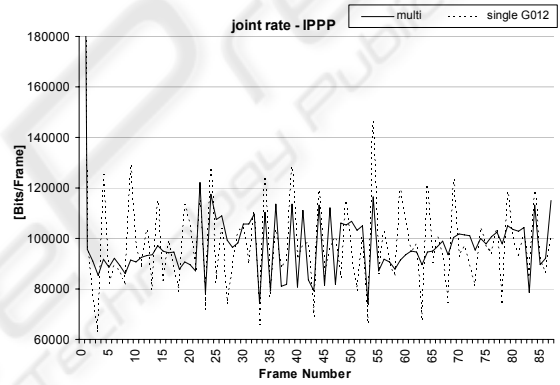Figure 4: Dependence of PSNR over frames using the IPPP pattern.

Figure 5: Dependence of joint bit-rate over frames using the IPPP pattern.

4 SIMULATION RESULTS

The rate control based on the ρ-domain is implemented in the H.264/AVC JM11 software reference model adapted to process several video sequences concurrently. The concurrency is achieved by switching between processed sequences when I or P frame and associated B frames are coded. In particular, the context of global and static variables is switched to keep data consistency. The multi-sequence rate control is verified in terms of the stability and compared with the single-sequence rate control. In particular, the original G012 version is used for the comparison. Results obtained for the single-sequence coding based on the ρ-domain are similar to those for the G012 proposition and are not shown for the clarity of plots.

459

Table 1: Comparison of qualities (PSNR) for the single- (G012) and the multi-sequence compression.

Joint Bit-rate [Mbps]	PSNR – luma [dB]					
	single-sequence			multi-sequence		
	News	Foreman	Mobile	News	Foreman	Mobile
0.75	37.57	34.40	24.75	33.12	32.28	27.80
1.5	41.90	37.68	28.57	36.87	35.37	31.50
3	45.55	40.85	31.83	40.24	38.33	35.56
6	48.93	43.73	35.65	43.34	41.46	40.22
12	53.00	47.08	40.05	47.11	45.76	45.38

Test are performed for following CIF sequences: Mobile, News, and Foreman. At 30 Hz frame rate, the bit rate is set on constant values equal to 1M bit/sec and 3M bit/sec for single- and multi-sequence coding, respectively. The encoder operates with Main Profile using Context Adaptive Binary Arithmetic Coding.
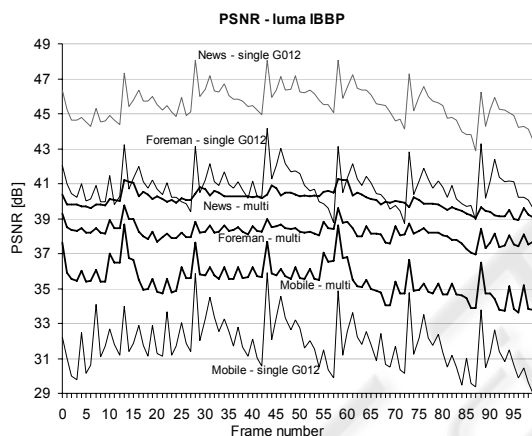

Figure 6: Dependence of PSNR over frames using the IBBP GOP pattern

Fig. 4 and 5 show simulation results for PSNR and the joint bit-rate, respectively. The results are obtained using the IPPP frame pattern. As can be seen, the multi-sequence rate control achieves the better stability. Moreover, quality is more balanced compared to independent encoding of each sequence. In Fig. 6, curves have periodic variations owing to Intra frames. These frames have higher PSNRs compared to Inter frames even though the same Qp is used. The variations for the multi-sequence compression are smaller since the Qp is selected to achieve the quality similar to the preceding P frame. By contrast, the G012 rate control analyzes the whole previous GOP and favours Intra frames. As original sequence at 30 Hz includes pairs of almost identical frames, the RD model fails to predict the accurate rate. This causes deviations in both the quality and the rate. Better

stability requires the use the finer rate control updated after coding some macroblocks not a whole frame (Li, 2003). This approach allows the RD model to predict rates more accurately for both the single- and the multi-sequence compression. Fig. 6 demonstrates the quality for each frame achieved when using the GOP structure for 30 Hz. Owing to the joint rate allocation and more accurate complexity weights, qualities are more balanced between sequences while keeping the target rate. This relation is valid for different bit rates as can be seen in Table 1. For high quality multi-sequence compression, differences in PSNR decrease. Selection of quantization parameter values with reference to content analysis would allow more similar quality between sequences with different complexity.

5 CONCLUSIONS

The rate control based on the ρ-domain allows better stability of encoded video compared to the G012 version. Thanks to exponential dependence of complexity weights on the quantization parameter, more accurate bit allocation for frames in a GOP is achieved. Moreover, the simpler buffer verifier proves its usefulness, i.e., the mismatches inherited from the G012 version are removed. The multi-sequence video compression allows a better quality balance between sequences. Future works will concentrate on balancing the quality based on the complexity analysis of the video content and updating the rate control on the macroblock level. Also, the use of various GOP patterns will be enabled to keep the total rate as constant as possible.

ACKNOWLEDGEMENTS

REFERENCES

ISO/IEC 14496-10:2003 | ITU-T Recommendation H.264, *Advanced Video Coding (AVC) for Generic Audiovisual Services/ MPEG-4 Part 10*, 2003.

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, February 1997.

He, Z., Mitra, S.K., *A unified rate-distortion analysis framework for transform coding*, IEEE Trans. On Circuits and Systems for Video Technology, vol. 11, no. 12, pp. 1221–1236, Dec. 2001.1996.

Li, Z. G., Pan, F., Lim, K. P., Feng, G. N., Lin, X., Rahardaj, S.: *Adaptive basic unit layer rate control for JVT*, doc. JVT-G012, 7th meeting, Pattaya, Thailand, March 2003.

Bobiński, P., Skarbek, W., *Analysis of RD models for coding efficiency in H.264 standard*, International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 2004, Lisboa, Portugal, Apr. 2004.

Pietrowcew, A., Buchowicz, A., Skarbek, W., *Bitrate control algorithm for ROI enabled video coding*, The 11th International Conference on Computer Analysis of Images and Patterns CAIP 2005, Rocquencourt, France, Sept. 2005.