BIT-RATE CONTROL FOR VIDEO CODING WITH ROI

Adam Pietrowcew, Andrzej Buchowicz, Władysław Skarbek

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

ABSTRACT

A novel bit-rate control algorithm for encoding of the video sequence with ROI has been presented in this paper. The algorithm is based on the linear R-D model in the ρ-domain. It ensures better quality of the reconstructed video in ROI and gradual quality degradation outside the ROI, while preserving the global constraint on the bit-rate of the encoded video sequence.

1. INTRODUCTION

The key issue in video compression is the optimal coder control, that is the selection of appropriate set of coding parameters, that will guarantee the demanded bit-rate with minimal lost in fidelity of the reconstructed video sequence. The quality of the whole frames of video sequence is usually taken into account. However in many applications, e.g. video monitoring and surveillance, telemedicine, some areas in the consecutive frames of the video sequence are more important than others. It is desirable to encode those areas, called region of interest (ROI) with smaller distortion than the rest of the sequence (background). Such functionality is available in the JPEG-2000 standard [1] for still image compression, but video coding standards do not provide it. This goal can be achieved by modifying the bit-rate control algorithm in such a way that more bits are allocated to the ROI and fewer bites to the background, while preserving the global constraint on the bit-rate of the encoded sequence.

This paper presents the bit-rate control algorithm for video sequence encoding with ROI. The algorithm is based on the linear rate-distortion (R-D) model in ρ domain [2]. It has been integrated and tested with the H.264/MPEG-4 AVC [3] reference software [4]. The algorithm outline and experimental results achieved with it will be presented in the following sections.

2. ROI BIT ALLOCATION ALGORITHM

In typical transform image/video coding both rate R and distortion D depend on the quantization parameter denoted here by q. The main task in designing a rate control algorithm is to find functions $R(q)$, $D(q)$, called R-D functions, for the particular coding scheme.

2.1. Linear RD model in ρ-domain

Methodology of our ROI bit budget allocation technique is based on ρ-domain algorithm [2, 5], which in turn is running on top of JVT-G012 proposal [6]. It has been shown [2] that for typical coding algorithms the R-D functions can be expressed by linear equations in the new domain of parameter ρ, which is the percentage of insignificant (quantized to zero) transform coefficients.

We assume that generated bit count values for intra and non-intra frames in encoded sequence satisfy the linear model in ρ-domain:

$$R(\rho) = \theta(1-\rho) .$$

Context in which ρ is computed depends on data source encoding options. It can be defined for frames or basic units of type I, P and B in one dimension or in color components Y, C_b, C_r correspondingly. In presented method such context was defined only for P frames and Y color component.

The slope θ is modeled on the base of previous context and given by the formula:

$$\theta = \frac{r_{prev}}{1-\rho_{prev}} .$$

Parameters r_{prev} and ρ_{prev} denotes the bit-rate and zero fraction in the previous context, accordingly. The experimental results justifying the above formulas were presented in [5, 7].

The mapping from ρ to quantization index q is straightforward, having lookup table build on basis of histogram for zero quantized coefficients obtained for all possible indices q. While model parameter θ depends strongly on local data, the relationship between ρ and q is less connected to data predication errors.

2.2. ROI bit-rate control

ROI in encoded sequence can be defined as one compact area or a set of them [7]. It's composed of macroblocks for which at least one pixel intersects with ROI (Fig. 1.). Let's denote all such macroblocks by L_1. Then we can define level sets L_i (for $i > 1$) as a sets of macroblocks which are 4-neigbours of macroblocks in L_{i-1} and are not included in L_j for all $j < i$ (Fig. 2).

All such level sets defined on frame based on its ROI structure can be further described by $N_i = |L_i|$ - count of macroblocks on given level set, ρ_i - zero fraction on L_i and r_i – bit-rate for L_i.

In order to distribute bit budget, allocated for frame by ρ-domain algorithm, we decided to increase the fraction of zeros in consecutive level sets L_i by $\Delta\rho$ such that:

$$\rho_i = \rho_1 + (i-1)\Delta\rho .$$

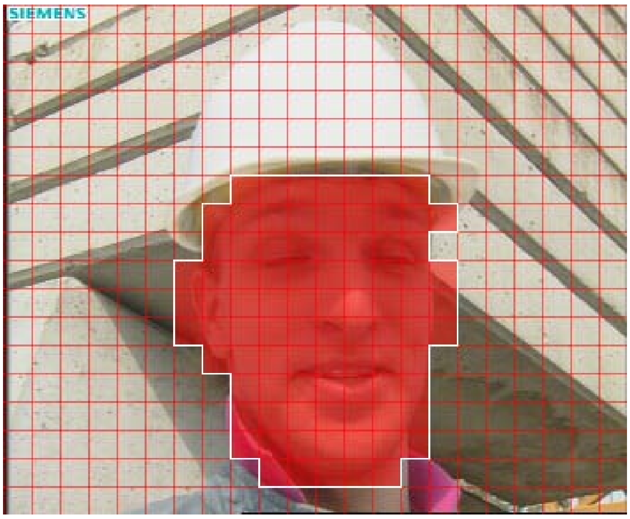

Figure 1.ROI definition for the first frame in Foreman sequence.

4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
4	4	4	4	4	4	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4
4	4	4	4	4	3	3	2	2	2	2	2	2	2	2	2	3	3	4	4	4	4
4	4	4	4	4	3	2	2	1	1	1	1	1	1	1	2	2	3	4	4	4	4
4	4	4	4	3	3	2	1	1	1	1	1	1	1	1	1	2	3	4	4	4	4
4	4	4	4	3	2	2	1	1	1	1	1	1	1	1	2	2	3	4	4	4	4
4	4	4	4	3	2	1	1	1	1	1	1	1	1	1	1	2	3	4	4	4	4
4	4	4	4	3	2	1	1	1	1	1	1	1	1	1	1	2	3	4	4	4	4
4	4	4	4	3	2	1	1	1	1	1	1	1	1	1	1	2	3	4	4	4	4
4	4	4	4	3	2	2	1	1	1	1	1	1	1	1	2	2	3	4	4	4	4
4	4	4	4	3	3	2	2	1	1	1	1	1	1	1	2	3	3	4	4	4	4
4	4	4	4	4	3	3	2	1	1	1	1	1	1	1	2	3	4	4	4	4	4
4	4	4	4	4	4	3	2	1	1	1	1	1	1	1	2	3	4	4	4	4	4
4	4	4	4	4	4	3	2	2	1	1	1	1	1	2	2	3	4	4	4	4	4
4	4	4	4	4	4	3	3	2	2	2	2	2	2	2	3	3	4	4	4	4	4

Figure 2.Level sets surrounding ROI in the first frame of Foreman sequence.

Assuming that global frame bit-rate is distributed to level sets L_i proportionally to their sizes we have got:

$$r = \sum_i w_i r_i \text{ , where } w_i = \frac{N_i}{N}$$

$$\rho_1 = 1 - \frac{r}{\theta} - \Delta\rho \sum_i w_i (i-1) .$$

The N and r in above formulas denote the count of all macroblocks in a frame and bit-rate (in bits per pixel) for frame, accordingly.

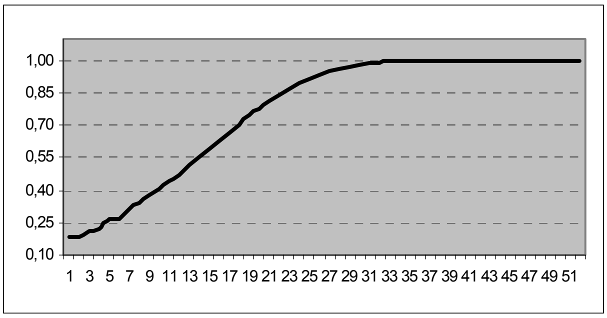

Figure 3. Dependency between ρ and q indices for P frames on Foreman sequence (average over P frames in range from 1 to 60).

Hence for the fixed $\Delta\rho$ we compute ρ_1 by the formula given above and next ρ_i for all $i > 1$. Having ρ_i we can use further steps of ρ-domain algorithm, i.e. getting the quantization index q_i from look-up table $\rho[q_i]$, establishing the encoding mode for each macroblock in L_i and model updating. In order to limit the maximum quantization index q_i the value of ρ_i is clipped to 0.999.

Proposed ROI bit allocation scheme works well on P frames. For I frames similar technique as in JVT-G012 proposal was applied. On each level set L_i we calculate sum of q_i indices for all P fames in a GOP - S^i_{pq}. Then quantization indices for ROI level sets on I frame in given GOP were calculated using formula:

$$q_i = \frac{S^i_{pq}}{N_p} - \min\left(2, \frac{N_{gop}}{15}\right),$$

where N_p denotes number of P frames in previous GOP and N_{gop} is the size of GOP.

Described previously ROI bit budged allocation scheme based on consecutive increasing zero fraction by constant $\Delta\rho$ is one of possibilities. Better results can be achieved if ρ_i on level set L_i will also be depended on its complexity. This complexity can be measured as frame pixels variance after motion compensation.

3. PRACTICAL APPLICATIONS

The proposed in previous section algorithm was evaluated on several standard test sequences. The ROI definitions was loaded to the encoder from external files. These files were created manually by test sequence browsing. The algorithm worked correctly on all sequences recorded with CIF and Q-CIF resolutions and at different bit-rates. The required bit-rate was preserved for all GOPs in test videos while at the same time quality of ROI region was better in relation to the original JVT-G012 implementation.

Figures 4 and 5 presents PSNR measures calculated on defined ROI level sets for proposed bit allocation algorithm and original JVT implementation respectively. From them, we can observe, that PSNR measure variance is lower for bit allocation method working in ρ-domain. Also quality of ROI region is better for almost all frames for more than 1dB.

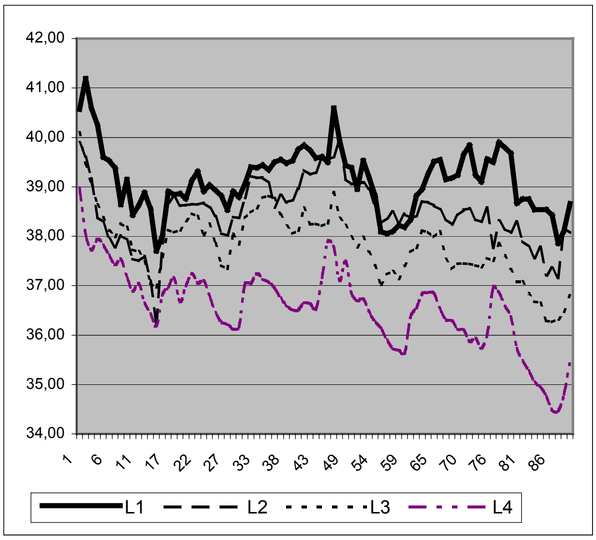

Figure 4. PSNR on ROI level sets for Foreman sequence, bit-rate 270Kb/s, $\Delta\rho = 0.0125$.

Quality on level sets L_1 and L_2 is comparable in both implementations. Only on the last level set quality in JVT implementation is better, as in ROI ρ-domain algorithm more bits were allocated to the area of interest.

In described ROI bit allocations algorithm we proposed splitting of all frame macroblocks into level sets in order to avoid strong quality change on the border between interest region and rest of the image. Such solution allows gradually decreasing quality on consecutive level sets lying farther from ROI.

Similar results (Fig. 6 and 7) were obtained for News video sequence, but here PSNR at ROI level set was slightly lower then on level set L_2. This was result of video sequence content, which presents scene with very dark background, two speakers at the first plane and dancing pair in second plane. As interest region was defined on speakers' heads and dancing pair this looks obvious.

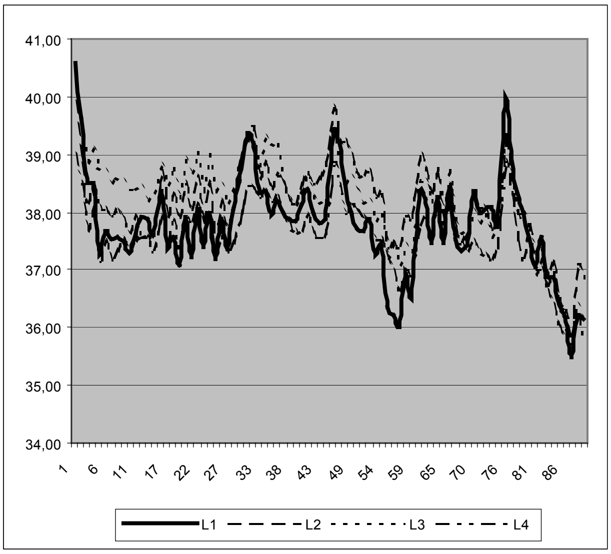

Figure 5. PSNR on level sets for Foreman sequence in original JVT-G012 implementation, bit-rate 270Kb/s.

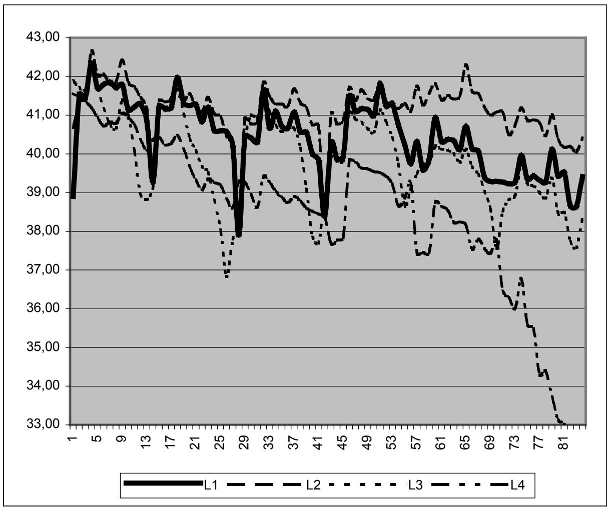

Figure 6. PSNR on ROI level sets for News sequence, bit-rate 230Kb/s, $\Delta\rho = 0.025$.

In comparison to sequence encoded by JVT-G012 algorithm quality on ROI was improved on average by 3dB. Also the quality in ROI not decreases such rapidly at the end of sequence. It's generally more stable. Average bit-rate for all GOPs is closer to desired value in

ROI ρ-domain method (230,71Kb) than in original implementation (232,49Kb), too.

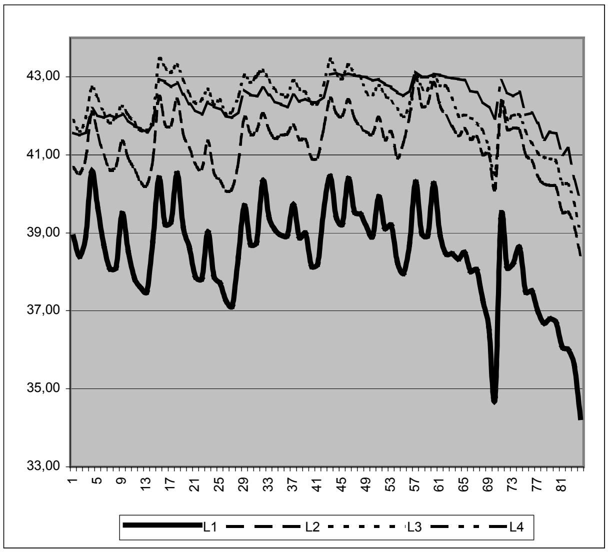

Figure 7. PSNR on level sets for News sequence in original JVT-G012 implementation, bit-rate 230Kb/s.

The PSNR average values for Foreman and News sequences were presented in Table 1.

Table 1. Sequence quality after decoding: Foreman 270 Kb/s, News 230 Kb/s

Sequence	no ROI (PSNRdB)			ROI (PSNR dB)		
	Y	U	V	Y	U	V
Foreman (CIF)	37,90	41,68	44,40	35,61	41,02	44,08
News (CIF)	41,33	43,59	44,78	39,19	42,46	43,65

4. CONCLUSSIONS

The bitrate control algorithm allowing encoding of the video sequence with ROI has been presented in this paper. It is based on the linear R-D model in ρ-domain. The algorithm guarantees higher quality of the reconstructed video in the ROI and gradual video quality degradation outside the ROI. The available bit budget was distributed among level sets in the consecutive frames by increasing the fraction of zero quantized coefficients by a constant value $\Delta\rho$. The value of the $\Delta\rho$ parameter was experimentally determined for each sequence and was constant for the entire sequence. It seems better results can be achieved if the value of the $\Delta\rho$ parameter is modified accordingly to the variation of the scene complexity. The presented algorithm was used only for P frames encoding. Its extension for B frames is currently being developed. It is also desirable to combine the proposed algorithm with object detection and tracking algorithm to simplify creation of the ROI definitions.

Acknowledgment: The work presented was developed within VISNET, a European Network of Excellence (http://www.visnet-noe.org), funded under the European Commission IST FP6 programme.

9. REFERENCES

[1] D.S. Taubman, and M.W. Marcellin, *"JPEG2000 Image Compression Fundamentals, Standards and Practice",* Kluwer Academi Publisher, Boston,2002

[2] Z. He, and S.K. Mitra, "A Unified Rate-Distortion Analysis Framework for Transform Coding" *IEEE Transaction on Circuits and Systems for Video Technology*, pp. 1221-1236, Vol. 11, No. 12, December 2001.

[3] ISO/IEC, *Generic Coding of Audiovisual Objects – Part 10: Advanced Video Coding (MPEG-4 AVC)*, ISO/IEC 14496-10, October 2003

[4] *H.264/AVC Software Coordination, http://iphome.hhi.de/ suehring/tml/*

[5] P. Bobiński, and W. Skarbek, "Analysis of RD Models for Coding Efficiency in H.264 Standard", *5th International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 2004*, Lisboa, Portugal, April 2001.

[6] Z. Li., F. Pan, K.P. Lim, G. Feng, X. Lin, and S. Rahardja, "Adaptive Basic Unit Layer Rate Control for JVT", *JVT-G012-r1*, Pattaya, Thailand, March 2003.

[7] W. Skarbek, A.Buchowicz, A.Pietrowcew, and F.Pereira, "Bit-rate Control for Compression of Video with ROI", *International Conference on Computer Vision and Graphics ICCVG 2004*, Warszawa, Poland, September 2001.