Journal of ELECTRICAL ENGINEERING, VOL. 63, NO. 2, 2012, 129-132

ANALYSIS OF CODING GAIN AND OPTIMAL BIT ALLOCATION IN MOTION–COMPENSATED VIDEO COMPRESSION

Andreja Samčović *

This paper describes mathematical frameworks on temporal and spatial predictive processing in the motion-compensated video compression. Firstly, the coding gain over intra coding is derived, regarding the bit allocation algorithm and Lagrange multiplier method. The optimal ordering of three different picture types (I, P and B pictures) is clarified according to image source characteristics. Secondly, a novel framework with the block-based multihypothesis motion-compensated optimal coding gain and bit allocation are derived in a closed-form expression.

K e y w o r d s: coding gain, bit allocation, video, compression, Lagrangian method, pictures

1 INTRODUCTION

Several video standards have been introduced which mainly address the compression of video data for digital storage and communication services. All of these standards apply block-based motion-compensated prediction (MCP). Motion-compensated schemes achieve data compression by exploiting the similarities among successive frames of a video signal. Most of the work for the design and optimization of video codecs is carried out experimentally, [1]. A theoretical treatment of motion-compensated video coding requires many assumptions and simplifications for the analysis of a complicated system processing real signals. Nevertheless, even an approximate theory can provide useful insights in the mechanisms and give guidance for the design of state-of-the-art video coders. The standard algorithms employ intra, inter and bidirectional picture coding types. Bidirectional prediction is one of the factors characterizing the MPEG (*Motion Picture Expert Group*) video compression standards. Coding efficiency is highly improved for some image sequences when the bidirectional prediction is utilized in an adequate manner. However, the picture ordering has been empirically defined without sufficient theoretical backgrounds [2].

This paper presents the mathematical frameworks for these problems. In Section 2, a coding gain and the optimal bit allocation are provided, based on Gaussian dis-

* University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11040 Belgrade, Serbia, andrej@sf.bg.ac.rs

DOI: 10.2478/v10187-012-0020-z, ISSN 1335-3632

tribution of an input signal and the Lagrange multiplier method. A parameter reflecting persistent tendency of interframe correlations is introduced, and the optimum picture ordering is indicated. Also, the improved algorithm with the multihypothesis is proposed. Finally, in Section 3 we come to some conclusions.

2 MATHEMATICAL ANALYSIS OF CODING GAIN AND OPTIMAL BIT ALLOCATION

The MPEG video standards have three picture types from the point of view of their temporal processing: intra-coded pictures with no temporal prediction (I-pictures), predictive-coded from earlier pictures (P-pictures) and bidirectionally predictive-coded pictures (B-pictures), that is the pictures have a prediction from earlier and later pictures in a sequence [3]. One example of such defined group of pictures (GOP) is shown in Fig. 1.

We can represent this GOP structure by

$$B^{M-1}I(B^{M-1}P)^{L-1}$$

where M is a distance between "core pictures" (I and/or P-pictures), L is the number of core pictures in the GOP, while the number of $L\times M$ is equal to the total number of pictures in the GOP, *ie* an interval between I-pictures.

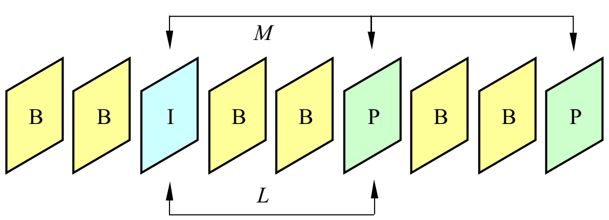

Fig. 1. GOP structure in the MPEG standard

By derivation a coding gain of this structure in comparison with the intra coding, we assume following relationships

$$\sigma_q^2 = \varepsilon^2 2^{-2R}\sigma_s^2 \tag{1}$$

where σ_s^2 is the variance of an input video signal s, σ_q^2 is a quantization error variance, R is the overall bitrate, while ε^2 is a quantization performance factor. The closed-loop prediction means that $\sigma_{r,i}^2 = \sigma_{q,i}^2$ for every i =I, P, B-picture, taking into account that σ_r^2 represents a reconstruction error variance.

Prediction error variances of the considered pictures are given by

$$\begin{aligned} \sigma_{r,I}^2 &= \sigma_{q,I}^2 = \varepsilon^2 2^{-2R_I}\sigma_s^2\,,\\ \sigma_{r,P}^2 &= \sigma_{q,P}^2 = \varepsilon^2 2^{-R_P}\sigma_s^2 P(M)\,,\\ \sigma_{r,B}^2 &= \sigma_{q,B}^2 = \varepsilon^2 2^{-R_B}\sigma_s^2 B(M)\,. \end{aligned} \tag{2}$$

We made an assumption that the quantization performance factor is the same in all pictures. R_I, R_P and R_B are the numbers of the allocated bits per picture element — pel, in the I, P and B-pictures, respectively. We assume also that the bits needed for pictures of the same type are equal for every picture of that type. $P(M)$ and $B(M)$ are the following terms

$$\begin{aligned} P(M) &= 2\left[1-\mathrm{cor}(M)\rho_s^{\Delta_r}\right],\\ B(M) &= \frac{3}{2}+\frac{1}{2}\mathrm{cor}(M)\rho_s^{\Delta_h} - \\ &\qquad \mathrm{cor}(j)\rho_s^{\Delta_r} - \mathrm{cor}(M-j)\rho_s^{\Delta_r} \end{aligned} \tag{3}$$

where j is the distance from the past core picture, Δ_h is the distance between two blocks, Δ_r is the radial displacement of each block, ρ_s is the spatial correlation coefficient, while $\mathrm{cor}(k)$ represents a correlation between k-frame apart pictures.

We introduce an input source model in which $\mathrm{cor}(k)$ is given by [4]

$$\mathrm{cor}(k) = \rho^{(k^w)} \tag{4}$$

where ρ is a correlation coefficient between neighboring pictures and w is a weighting factor. When $w=1$, it is the autoregresive AR(1) model. The objective of w is to reflect persistent tendency of interframe correlations, which is suggested by simulation results using real image sequences. Typical values of w are around 0.25–0.5. In image sources with highly continuous correlations, for instance in the sequence *Mobile* & *Calendar*, w takes the small value of 0.125.

Based on these equations, the optimum bit allocation problem is formulated as follows: minimizing the average reconstruction error variance [5]

$$\min\sigma_r^2 = \frac{1}{LM}\Big[\sigma_{r,I}^2 + (L-1)\sigma_{r,P}^2 + L\sum_{j=1}^{M-1}\sigma_{r,B}^2(j)\Big]^2, \tag{5}$$

taking into account the constant bitrate constraint

$$R_I + (L-1)R_P + L(M-1)R_B = LMR = \text{const.} \tag{6}$$

By applying the Lagrange multiplier method [6], we should minimize

$$\begin{aligned} \min J = \min\Big\{&\frac{1}{LM}\Big[\sigma_{r,I}^2 + (L-1)\sigma_{r,P}^2 + L\sum_{j=1}^{M-1}\sigma_{r,B}^2(j)\Big] + \\ &\lambda[R_I + (L-1)R_P + L(M-1)R_B]\Big\} \end{aligned} \tag{7}$$

where J is the Langrangian rate-distortion functional, and λ is the Lagrangian multiplier. We obtained the following optimal bit allocations

$$\begin{aligned} R_I &= R - \frac{L-1}{2LM}\log_2 P(M) - \frac{M-1}{2M}\log_2 B(M)\,,\\ R_P &= R + \frac{L(M-1)+1}{2LM}\log_2 P(M)\\ &\qquad - \frac{M-1}{2M}\log_2 B(M)\,,\\ R_B &= R - \frac{L-1}{2LM}\log_2 P(M) + \frac{1}{2M}\log_2 B(M)\,. \end{aligned} \tag{8}$$

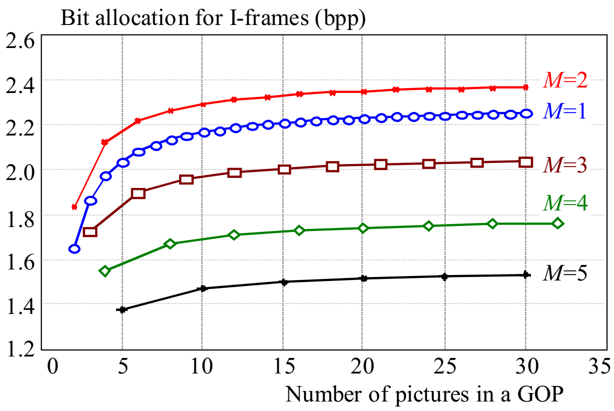

Fig. 2. Bit allocation for I-frames

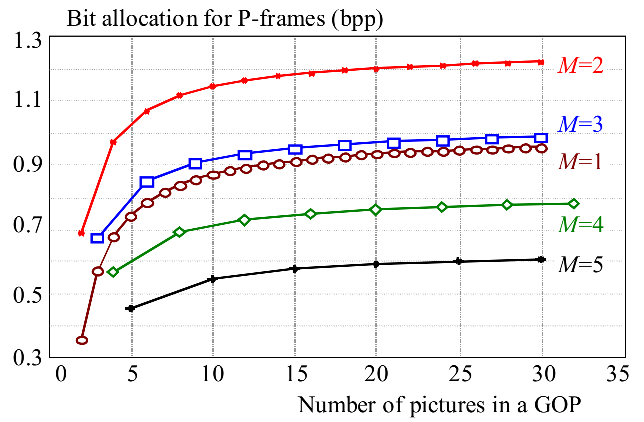

Fig. 3. Bit allocation for P-frames

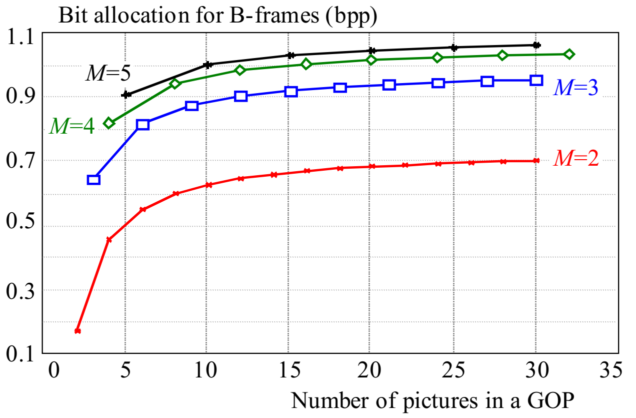

Fig. 4. Bit allocation for B-frames

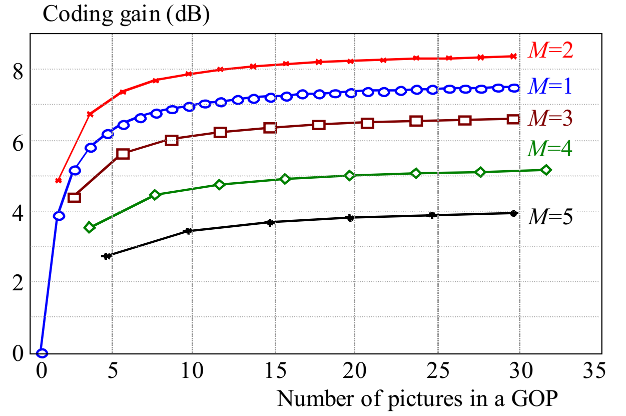

Fig. 5. Coding gain for *Mobile & Calendar*, $\rho = 0.95$ and $w = 0.125$

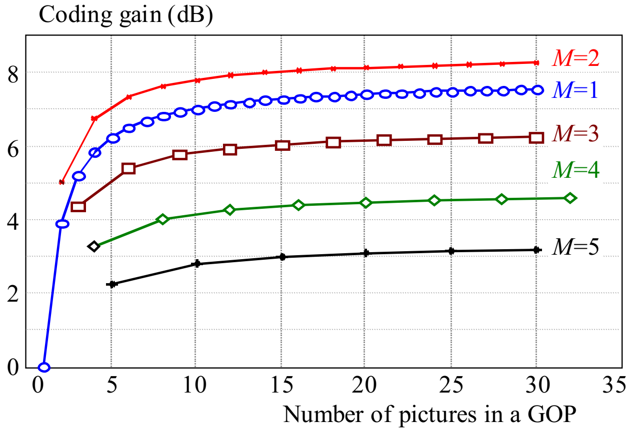

Fig. 6. Coding gain for pictures with moderate details, $\rho = 0.95$ and $w = 0.5$

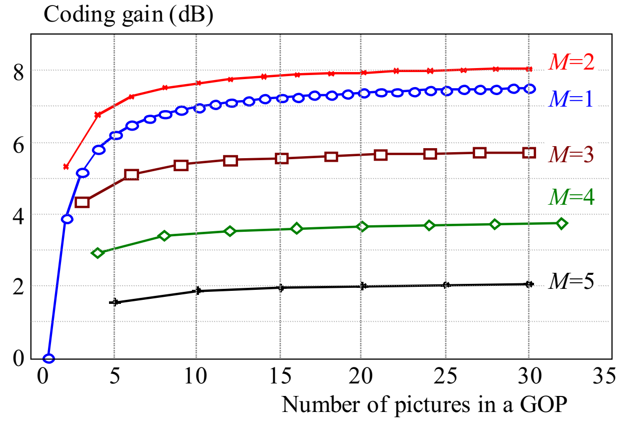

Fig. 7. Coding gain for pictures with AR(1), $\rho = 0.95$ and $w = 1.0$

Figures 2, 3 and 4 represent the bit allocations for I, P and B-pictures, respectively, for an example of $\rho = 0.95$; $w = 0.5$; $\Delta_r = 0.5$ and $\rho_s = 0.93$. This corresponds to the case of ordinary image sequences. When the number of pictures increases, one should spend more bits per pixel, but the saturation is exceeded with 20 pictures in a group. For I-pictures, R_I takes values from 1.35 (when $M = 5$) to 2.35 ($M = 2$) bits per pixel (bpp). P-pictures require significantly more bits per pixel than B-pictures.

Taking into account the optimal values, the coding gain of the GOP structure is defined and calculated by

$$G = \frac{\sigma_q^2}{\sum_{t=I,P,B} \sigma_t^2} = \frac{1}{P(M)^{(L-1)/LM} B(M)^{(M-1)/M}} = 10 \log \frac{1}{P(M)^{(L-1)/LM} B(M)^{(M-1)/M}} \text{ (dB)} . \quad (9)$$

Figure 5 shows the example of the sequence with the still high correlation between distant pictures, such as the case of *Mobile* & *Calendar*, where $\rho = 0.95$; $w = 0.125$; $\Delta_r = 0.5$ and $\rho_s = 0.93$. The coding gain increases when the number of pictures in a GOP decreases. Adequate insertion of B-pictures brings signal-to-noise ratio (SNR) gains comparing with $M = 1$ (no B-picture). Use of a single B-picture between the I and P-frames yields the best SNR gain. Increasing the number of B-pictures results in decreased compression performance. In Fig. 6 we can see that for the typical images with moderate details ($\rho = 0.95$; $w = 0.5$; $\Delta_r = 0.5$ and $\rho_s = 0.93$), the coding gain is a little bit more than 8 dB for the best case $M = 2$, and is generally smaller in comparison with the highly correlated sequences. Figure 7 represents the AR(1) process, when $\rho = 0.95$; $w = 1$; $\Delta_r = 0.5$ and $\rho_s = 0.93$. The difference between $M = 2$ and $M = 1$ is ~ 0.5 dB, while the gain between the best ($M = 2$) and the worst case ($M = 5$) is greater than for all other values of the weighting factor, and takes value of about 6 dB.

3 CONCLUSION

Adequate insertion of B-pictures brings SNR gains. Use of a single B-picture between the I and P-frames yields the best SNR gain. Increasing the number of B-pictures results in decreased compression performance. For the typical images with moderate details, the coding gain is smaller in comparison with the highly correlated sequences. The coding gain increases when the number of pictures in a GOP increases, but the saturation is obtained still with 20 pictures in a group.

References

[1] BANIC, M.—LUCENTEFORTE, M.—MARCOZZI, E.: MPEG-2 Optimized Video Coder, International Picture Coding Symposium PCS'94, Sacramento, Cal, USA, 21-23 September 1994, pp. 79–82.

[2] WIEGAND, T.—ZHANG, X.—GIROD, B.: Long-Term Memory Motion-Compensated Prediction, IEEE Trans. on Circuits and Systems for Video Technology **9** No. 1 (Febr 1999), 70–84.

[3] SCHWARZ, H.—MARPE, D.—WIEGAND, T.: Analysis of Hierarchical B Pictures and MCTF, 2006 IEEE International Conference on Multimedia and Expo, Toronto, Canada, 9–12 July 2006.

[4] KATTO, J.—OHTA, M.: Mathematical Analyis of MPEG Compression Capability and its Application to Rate Control, International Conference on Image Processing ICIP'95, vol. II, Washington, USA, 23-26 October 1995, pp. 555–558.

[5] RAMCHANDRAN, K.—ORTEGA, A.—VETTERLI, M.: Bit Allocation for Dependent Quantization with Applications to Multiresolution and MPEG Video Coders, IEEE Trans. on Image Processing **3** No. 5 (Sept 1994), 533–545.

[6] LI, X.—OERTEL, N.—HUTTER, A.—KAUP, A.: Laplace Distribution based Langrangian Rate Distortion Optimization for Hybrid Video Coding, IEEE Trans. on Circuits and Systems for Video Technology **19** No. 2 (Febr 2009), 193–205.

Received 15 March 2011

Andreja Samčović was born in 1963, Belgrade, Yugoslavia. From 1983/1989 he attended the Faculty for Electrical Engineering, University of Belgrade. Since 1991, he has been a teaching assistant at the Faculty of Transport and Traffic Engineering, Belgrade. He spent 9 months in 1991/92 as a scholar of the OEAD at the TU Vienna, Austria. He attended the postgraduate studies at the University of Belgrade. In 1995 he wrote the magister thesis Some techniques of subband still image coding. He spent the October 1995 in studying the German language at the Goethe- Institute in Prien, Germany. Also, he spent 6 months in 1999/2000 at the Fridrich-Alexander University Erlangen-Nuremberg, in Erlangen, Germany, as a scholar of DAAD. He published in 2000 the monograph Subband Image Coding, in English and Serbian, published by Andrejević Endowment, Belgrade. In 2003, he served as a lecturer at the Uni-Adrion summer school in Ammudia-Preveza, Greece. In 2003-2004, he participated at the research project with the Virtual University of the Adriatic-Ionian basin, as well as in 2004-2005 at the bilateral project with Technical University Kosice, Slovakia. In 2005, he made his PhD thesis at the University of Belgrade, in the field of Digital Image Compression.