

Spatiotemporal adaptive quantization for an efficient video rate control

Si-Woong Lee

Hanbat National University
Division of Information Comm. & Computer Eng.
16-1, Dukmyung-dong, Yuseong-gu
Daejeon, 305-719, Korea

Wook-Joong Kim

Kyuheon Kim
Electronics and Telecommunications
Research Institute
Broadcasting Media Research Department
161, Gajeong-dong, Yuseong-gu
Daejeon, 305-350, Korea

Abstract. A new algorithm for the rate control of videos considering the sensitivity of the human visual system (HVS) is presented. The method adopts the three-step structure of MPEG-2 Test Model 5 (TM5) rate control, while a new measure for the macroblock (MB) activity based on spatiotemporal sensitivity is introduced. Experimental results show that the proposed activity measure outperforms the spatial activity of TM5 in picture quality. © 2005 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.1920068]

Subject terms: video rate-control; MPEG-2 Test Model 5(TM5); adaptive quantization; human perception.

Paper L040977 received Dec. 20, 2004; revised manuscript received Feb. 20, 2005; accepted for publication Mar. 10, 2005; appeared online Mar. 14, 2005; published online Jun. 10, 2005.

1 Introduction

The conventional rate control algorithms can be classified into two classes according to whether they use the rate-quantizer (R-Q) model or not. The model based approaches, such as H.263 TMN Ver. 10 (TMN10)¹ and MPEG-4 Verification Model Ver. 8 (VM8)² have been considered as rate control for maximizing the video quality at a given bandwidth. However, they are inappropriate for use in a wide variety of bit rates. Among model-free approaches, the MPEG-2 TM5 rate control³ is the most representative. It consists of three steps to adapt the MB quantization parameter for controlling the bit rates.^{3,4} Among them, the adaptive quantization step modulates the quantization parameter adaptively according to the spatial activity of the MB, considering the characteristic of the human perception. The human perception, however, depends not only on the spatial activity but also on the temporal activity of the MB. Based on this observation, we propose a new spatiotemporal activity measure for adaptive quantization.

2 Overview on TM5 Rate Control

The first step of the TM5 rate control is the target bit allocation. Based on several factors including the picture type, buffer fullness, and picture complexity, the number of bits

available for coding the current picture is estimated. The second step is the rate control step. In this step, the number of bits for each MB is determined by the rate control algorithm within a picture, and then a quantization parameter is derived from the number of bits available for the MB to be coded. Adaptive quantization is the last step of the TM5 rate control. It is noted that for active areas or busy areas, the human eyes are not so sensitive to the quantization noise, while the smooth areas are more sensitive to the quantization noise according to the frequency masking effect of the HVS. Based on this property, the adaptive quantization modulates the quantization parameter obtained from the previous step in such a way to increase it for active areas and reduce it for smooth areas.

3 Adaptive Quantization Using Spatiotemporal Activity Measure

The adaptive quantization in the TM5 considers only the spatial characteristic of the human visual perception. Thus, the quantization parameter is modulated according to the spatial activity measure of the current MB. There still, in fact, exists another feature in the HVS called temporal masking, which can be suitably used to enhance the performance of the adaptive quantization. Temporal masking states that it takes a while for the HVS to adapt itself to the scene when the scene changes abruptly. During this transition the HVS is not sensitive to details. Considering both spatial and temporal masking properties of the HVS, the perceptual performance improvements can be obtained by introducing an adaptive quantization based on the spatiotemporal MB activity.

In the proposed method, the modulated quantization parameter for the *j*th MB is obtained by

$$mquant_j = Q_j \cdot N_act_j \quad (1)$$

where Q_j is the quantization parameter obtained in the second step of TM5 rate control, and N_act_j is the MB activity factor, which is calculated as

$$N_act_j = \alpha \cdot N_act_j^s + (1 - \alpha) \cdot N_act_j^t \quad (2)$$

The above equation indicates that, differently from TM5, the activity factor of the current MB is calculated as the weighted sum of the spatial activity factor $N_act_j^s$ and the temporal activity factor $N_act_j^t$, which are calculated as

$$N_act_j^s = \frac{2 \times act_j^s + avg_act^s}{act_j^s + 2 \times avg_act^s} \quad (3)$$

$$N_act_j^t = \frac{2 \times act_j^t + avg_act^t}{act_j^t + 2 \times avg_act^t}$$

Table 1 Performance comparison.

Input sequence	TM5			Proposed		
	PSNR	Rate	QP	PSNR	Rate	QP
Carphone	37.26	127.92	13.85	37.99	128.02	14.00
Coastguard	30.23	127.87	20.02	31.11	128.13	19.52
Container	39.15	127.98	8.48	40.19	128.17	8.71
Foreman	35.18	128.00	16.39	35.90	127.88	16.16

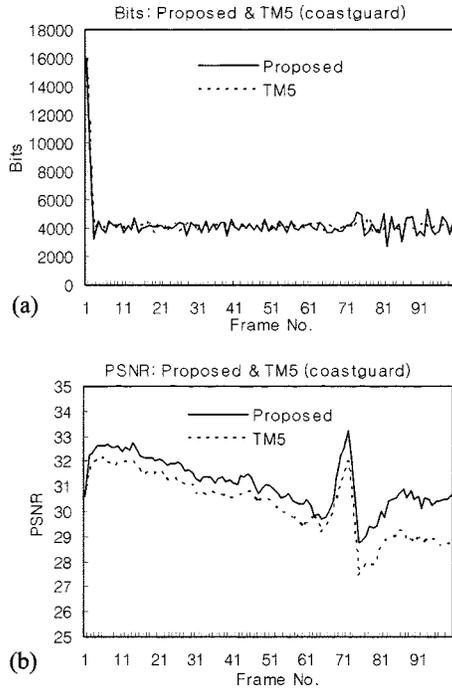


Fig. 1 Results of the “Coastguard” sequence: (a) generated bits; (b) PSNR.

where avg_act^s and avg_act^t are the average values of the spatial activity measure act_j^s and the temporal activity measure act_j^t of the last picture to be encoded, respectively; act_j^s and act_j^t for the j th MB are calculated as

$$act_j^s = 1 + \underset{blk=1,8}{\text{Min}} (\text{var_blk})$$

$$act_j^t = 1 + |Mvp(j)| \tag{4}$$

where var_blk is the variance of each spatial 8×8 block, and $Mvp(j)$ is the predicted motion vector of the j th MB. For the case where the motion vector for the current MB is not available in rate control process, such as the H.264 video encoder, we use the predicted motion vector instead of the motion vector in Eq. (4).

The parameter α in Eq. (2) is the weight of the spatial activity in the MB activity factor. For intrapictures, $\alpha=1.0$ is used, and for interpictures $\alpha=0.5$ is used. Note that $\alpha=1.0$ means the TM5 rate control algorithm.

4 Experimental Results

To evaluate the performance of the proposed algorithm, four standard test sequences of “Carphone,” “Container,” “Coastguard,” and “Foreman” are used in the experiment, and the results are compared with those of the TM5. The rate control algorithms are implemented on the H.264/AVC JM2.0 video codec.⁵ The frame rate is 30 frames/s, and the input sequences of 4:2:0 QCIF format are coded at the channel rate of 128 kbps. Except for the first frame, which is intracoded, all the subsequent frames are predictive-coded.

Table 1 shows the average performance over 100 frames of both methods. It is noted that both methods provide ex-

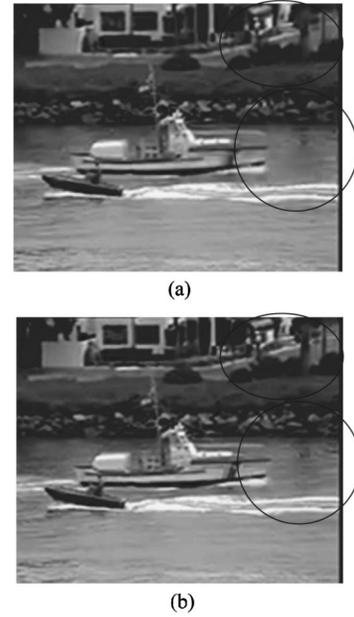


Fig. 2 Reconstructed frames: (a) TM5; (b) proposed.

act rate control results. In the respect of average PSNR, the proposed method shows improvements about 0.7–1.1 dB over the TM5. Figure 1 shows the number of generated bits and PSNR values of the “Coastguard” sequence. Note that there is a little bit larger interframe variation in bit rates in the proposed method. This indicates that the proposed method modulates the number of bits so that more bits are allocated to stationary frames. The results of Table 1 and Fig. 1 show that the spatiotemporal activity measure for the adaptive quantization is superior to the spatial activity measure in TM5 in objective picture quality. Figure 2 shows the reconstructed frames of the 87th frame of the “Coastguard” sequence. It can be seen that, in the background and the right sea of the white coast guard ship (circled areas), the proposed method shows better picture quality. This demonstrates that the proposed method assigns more bits to still areas than the area of the fast-moving ship.

5 Conclusions

This letter proposes a TM5-like video rate control algorithm using a new adaptive quantization scheme. In our method, considering the temporal characteristic of the human perception, a spatiotemporal MB activity measure for the adaptive quantization is proposed. By evaluating the MB temporal activity according to the magnitude of the motion vector, the quantization parameter can be modulated so that more bits are assigned to the more perceptible still areas.

References

1. ITU-T/SG16, “Video codec test model near-term version 10” (Apr. 1998).
2. ISO/IEC JTC1/SC29/WG11, “Text of ISO/IEC 14496-2 MPEG video VM-version 8.0” (July 1997).
3. ISO/IEC JTC/SC29/WG11, “Test Model 5,” Draft (Apr. 1993).
4. Y. Q. Shi and H. Sun, *Image and Video Compression for Multimedia Engineering*, CRC Press, Boca Raton, FL (2000).
5. ISO/IEC JTC1/SC29/WG11, “Working draft of reference software for advanced video coding” (Mar. 2003).