

# A New Quantization for Rate Control with Frame Variation Consideration

Seonki Kim  
Digital Media Laboratory  
LG Electronics, Inc.  
Seoul, Korea  
Email: seonki@lge.com

Seong-Jae Lim  
Dept. of Information and Communications  
Gwangju Institute of Science and Technology  
1 Oryong-dong, Buk-gu, Gwangju, 500-712, Korea  
Email: sjlim@gist.ac.kr

**Abstract**—The main role of rate control in video coding is to handle a trade-off between a bandwidth limit and a video quality. In the low-delay video communication systems, its responsibility is more important. In this work, we propose a new rate control algorithm. In this paper, we have two approaches: one is a consideration of variations among the frames at the frame-level bit allocation, and the other is a model-based adaptive quantization at the macroblock-level. The proposed scheme is done with a low computational complexity. For simulations, we implement the proposed algorithm to H.263 video codec since it is widely used as a video coding tool in real-time mobile AV systems. The proposed scheme shows more improved picture quality than H.263 TMN8 and generates coding bits close to the target bits. The results also show more regular PSNR fluctuation than H.263 TMN8.

## I. INTRODUCTION

The video codec plays an important role in the mobile video communication systems. The block-based hybrid coding schemes, such as MPEG and H.26x series [1]–[3], are widely used to the video communication systems. Together with the coding efficiency of video codec, the quality control is the most important factor, which should be considered. The performance of general coding tools, such as an accuracy level of motion estimation/compensation and variable length coding, is determined by the capability of the systems. However, the general video coding is a lossy coding that the loss of information is done by the quantization process and it is controlled by the rate control.

The main role of rate control is to control the trade-off between coding efficiency and video quality. It regulates the coded bit streams to meet the channel rate while keeping a good picture quality. The other objective of the rate control is a buffer control to avoid an overflow and underflow. Especially, the real-time video communication requires a very small end-to-end delay to provide better quality service.

The rate control for low-delay system is done at two layers: frame-level and macroblock-level. The macroblock-level rate control does a regularization of coding bits close to the target bits. In the recent research, its approaches are based on the R-D model [5], [11]. The rate control solutions introduced in MPEG-4 and H.263 TMN8 are designed based on the R-D model [2]–[4].

The rate control algorithm proposed in this paper is also

based on R-D model. The previous R-D model-based solutions defined the source characteristics as a single model. However, it can cause an exhaustive approximation problem. In order to solve that, we design a R-D model, which is done at macroblock-level to determine the quantization parameter, based on the Generalized Gaussian Distribution (GGD) [6]–[9]. We also consider the variation among frames at frame-level rate control. It can make more adaptive bit-allocation to the frames and reduce the weakness of the predictive coding, such as a scene change.

In the following two sections, we describe the proposed rate control algorithm in two-level: frame and macroblock levels. In section IV, we demonstrate the simulation results for the proposed algorithm compared with H.263 TMN8 rate control algorithm. Finally, we conclude the meaning of this paper.

## II. FRAME-LEVEL RATE CONTROL

In real-time video, such as video conferencing, the end-to-end delay is to be very small. Therefore, the frame-level optimization is an inadequate to the above system since it requires more elaborate source statistics. In this point, we make an optimization process in a quantization parameter decision that it is done at the macroblock level.

Our bit-allocation scheme is based on the algorithm shown in H.263 TMN8. The bit allocation used in H.263 TMN8 just considers buffer level to adjust target bits to allocate the current frame. It leads a weakness for the case of scene change. Especially, real-time video, such as video telephony, the change of scene has a high possibility being occurred. In order to compensate this limit, we reflect the difference/variation ratio among frames to give more adaptive bit allocation.

The total coding bits, that are spent to encode frames for a second, should be similar to the target bits, which is the pre-determined channel capacity. It is expressed in Eq. (1).

$$R = \sum_{i=1}^F R_i \quad (1)$$

$R$  and  $R_i$  indicate the target bits per second and the number of coding bits used to encode the  $i$ -th frame, respectively.  $F$  is the number of frames to be encoded.

The target bits for a frame,  $R_i$ , is simply obtained as shown in Eq.(2).

$$R_i = K_i \cdot \frac{R}{F} - \Delta, \text{ where } K_i = \frac{M_{(i,i-1)}}{\frac{1}{5} \sum_{j=1}^5 M_{(i-j,i-j-1)}} \quad (2)$$

The  $K_i$  denotes the frame variation ratio, and  $M$  indicates the MAD(Mean Absolute Difference). In order to see an overall variation among previous coded frames, we consider five reconstructed reference frames. The value of  $K_i$  can be increased on a sudden. In this case, we may have a trouble by a shortness of bits required to encode the remaining frames. Therefore, we restrict the range of  $K_i$  between 0.8 and 1.2.

The  $\Delta$  is obtained as a same method used in H.263 TMN8.

$$\Delta = \begin{cases} \frac{W}{F}, & W > Z \cdot M \\ W - Z \cdot M, & \text{otherwise} \end{cases} \quad (3)$$

in which by default  $Z$  is set by 0.1.  $\Delta$  is a small value that provides feedback from the buffer fullness  $W$ . If  $W$  is larger than 10 percentage of  $M$ , the frame target  $B$  is slightly decreased. Otherwise,  $B$  is slightly increased. Since  $B' \approx B$ , the buffer fullness is  $W \approx \max(W_{prev} - \Delta, 0)$  and, as a result, the  $\Delta$  correction will help maintain a small number of bits in the buffer without underflow.

### III. MACROBLOCK-LEVEL RATE CONTROL

In this work, we are concerned about the R-D function of GGD:

$$R = \frac{1}{\gamma} \log_2 \left( \frac{\sigma^\beta}{D} \right) \quad (4)$$

where  $\beta$  is the shape factor, which describes a decay rate of the distribution, and  $\gamma$  is the coding parameter that depends on the shape factor and the standard deviation of the source. As a change of the shape factor, the distribution form is determined. The selection of the proper shape parameter is important since the quantization parameter is determined as the source characteristics. Together with the importance of shape parameter, the distortion, which is caused by the quantization parameter, is also key factor should be considered to determine a quantization parameter. We define a distortion model that is designed as a function of the quantization parameter.

In general, the distortion by the quantization is defined as follows,

$$D = c \cdot Q_i^2 \quad (5)$$

where  $c$  is the distortion parameter that represents a relationship between distortion and quantization parameter. The value of distortion parameter gives an inspiration to make a region-based rate control scheme. We don't handle that issue in this paper. The decision of distortion parameter will be described in the next section.

From Eq. (4) and Eq. (5), we can derive a rate-quantization model as follows:

$$Q_i = \sqrt{\frac{\sigma_i^\beta \cdot 2^{-\gamma_i R_i}}{c_i}} \quad (6)$$

In Eq. (6), the shape factor  $\beta$  and  $\gamma$  take charge of a key role to model the source distribution. We will explain how we get those parameters in following sections.

#### A. Shape factor

The shape factor is an important parameter to model a form of the source distribution. However, the exact estimation of the shape parameter for the source is not easy. It also requires a high computational complexity. In order to avoid this problem, some researches have been studied to find the optimal shape factor with small complexity [7], [8]. In this work, we simply decide the shape factor by a similarity between two adjacent images.

$$\beta = \begin{cases} 1, & \omega > 0.9 \\ 2, & \omega < 0.6 \\ 2 - 3.3 \cdot (\omega - 0.6), & a \leq \omega \leq b \end{cases} \quad (7)$$

$$\omega = \frac{1}{256} \sum_{i=1}^{256} v(i), v(i) = \begin{cases} 1, & |x - \hat{x}| < 2 \\ 0, & \text{o.w.} \end{cases} \quad (8)$$

where  $x$  and  $\hat{x}$  denotes the samples, which are placed on the corresponding point, in the current and the reference frames, respectively. In this work, we assume that the source distribution is modeled as one of forms existing between Laplacian, that is defined by the value of 1, and Gaussian, which is defined by the value of 2, distributions.

#### B. Model and Distortion parameters decision

The model parameter  $\gamma$  is important to determine a quantization parameter. In Eq.(9), the model parameter is derived by  $\beta$ ,  $\sigma$ , and  $\Gamma$ -function.

$$p(x) = \frac{\beta \cdot \gamma}{2\Gamma(1/\beta)} e^{-(\gamma|x|)^\beta}, \gamma = \frac{1}{\sigma} \sqrt{\frac{\Gamma(3/\beta)}{\Gamma(1/\beta)}}. \quad (9)$$

As mentioned in the previous subsection, we limit the range of the shape factor,  $\beta$ , between 1 and 2. In this case, the range of  $\sqrt{\frac{\Gamma(3/\beta)}{\Gamma(1/\beta)}}$  is placed between 0.7071 and 1.4142. Therefore, we can derive the equation to obtain a model parameter,  $\gamma$ , as follows:

$$\gamma = \delta \cdot \frac{\beta}{\sigma} \times 0.7071. \quad (10)$$

To compensate the difference between the coding bits and target bits, we adjust a model parameter by the adjustment value,  $\delta$ . The update and computation of delta is explained in the coming subsection.

In Eq.(5), we design a distortion model by a distortion parameter,  $c$ , and a quantization parameter,  $Q_i$ . In order to decide the distortion parameter for each macroblock, we consider the typical distortion measure of quantization [3].

$$D(i) = \alpha_i \frac{Q_s^i}{12} \quad (11)$$

where  $Q_s^i$  denotes the quantization step size, and  $\alpha$  is a distortion weight for the  $i$ -th macroblock. In H.263 video coding standard,  $Q_s^i = 2 \times Q_i^2$ , where  $Q_i$  is the quantization parameter.

From Eq.(5) and Eq.(11), we can derive the equation to compute a distortion parameter,  $c$ , as follows:

$$c_i = \frac{\alpha_i}{3} \quad (12)$$

$$\alpha_i = \begin{cases} \frac{B_T}{\sigma_i \times (256N)}, & \frac{B_T}{256 \times N} < \frac{1}{2} \\ 1, & o.w. \end{cases} \quad (13)$$

where  $B_T$  denotes the target bits for a frame,  $\sigma_i$  is the standard deviation for the  $i$ -th macroblock.  $N$  is the number of macroblocks in a frame.

### C. QP Computation and Update Parameters

The quantization parameter is computed by the designed rate-quantization model shown in Eq.(6). The parameters, that are required in Eq.(6), are obtained in pre-processing stages before determining a quantization parameter.

After encoding a macroblock, we update information to be used for the next coding. In parameters being updated, the adjustment value for model parameter should be taken into account to compensate the difference between the estimating result and the actual result. Let  $R_T$  be the target bits per pixel and  $R_A$  be the coding bits. Theoretically,  $R_A$  and  $R_A$  follow the rate-quantization model since it was based on the rate-distortion function. Thus, we can express  $R_A$  and  $R_A$  as follows:

$$R_T = \frac{1}{\gamma_T} \log_2 \frac{\sigma^\beta}{c \cdot Q^2} \quad (14)$$

$$R_A = \frac{1}{\gamma_A} \log_2 \frac{\sigma^\beta}{c \cdot Q^2} \quad (15)$$

where  $\gamma_T$  and  $\gamma_A$  denote the estimated and the actual model parameters. From the relationship between Eq.(14) and Eq.(15), we can compute the adjustment value,  $\delta$ .

$$R_T - R_A = \left( \frac{1}{\gamma_T} - \frac{1}{\gamma_A} \right) \cdot A \quad (16)$$

$$\Leftrightarrow \frac{1}{\gamma_T} - \frac{1}{\gamma_A} = \frac{R_T - R_A}{A} \quad (17)$$

Therefore,

$$\delta = \frac{\gamma_A}{\gamma_T} = \frac{A}{A - (R_T - R_A) \cdot \gamma_T} \quad (18)$$

where

$$A = \log_2 \left( \frac{\sigma^\beta}{c \cdot Q^2} \right) \quad (19)$$

The value of A is small.

In here, we should keep in mind the meaning of  $\delta$ . In Eq.(10),  $\delta$  is a value to adjust the current model parameter, and it is based on the ratio between the estimated model parameter and the real model parameter.

We also update the counter. When each maroblock is coded, the counter is increased by 1.

## IV. EXPERIMENTAL RESULTS

### A. Experimental Conditions

In order to estimate the performance of the proposed algorithm, we implement the proposed algorithm into the H.263 video codec. Table I. shows the experimental conditions for our simulation.

TABLE I  
SIMULATION CONDITIONS

RDO	On
GOP Structure	IPPP
QP for I-frame	13
Target bit-rate	36, 48, 64
Frame rate	10, 15
Reference Frames	1

The proposed algorithm is applied for the P-frames because B-frame coding is not used in the low bit-rate real-time communication system. The encoding of the frame type of B does a bi-directional predictive coding, and it requires much complexity and time consumption. Therefore, we use IPPP GOP structure for simulation. For the I-frames, we use the fixed quantization parameter value, which is 13. Test video sequences used for simulation are "Foreman," "News," and "Carphone" sequences. The format of sequences is QCIF with 4:2:0. We use three target bits, which are 36, 48 and 64 kbps. These target bits are generally used as a target bits in video conference video systems, such as 3G handset. Frame rates used in experiment are 10 fps and 15 fps.

### B. Performances

In order to evaluate the performance of the proposed rate control algorithm, we compare the results with the results of H.263 TMN8 [3]. The main purpose of the rate control is to provide an enhanced video quality within given constraints. Therefore, we can pass a judgment by considering two factors, PSNR and coding bits.

One of measures to evaluate the performance of the rate control method is the coding bits that should be generated close to target bits. When the rate control algorithm fully uses a target bits, we can get more improved results. For experiments, we use three low target bits, 36, 48, and 64 kbps. As shown in Table II, the proposed algorithm generates the number of bits close to the target bits.

Table III describes the comparison of the average PSNR values for the proposed rate control algorithm and H.263 TMN8. The proposed scheme provides improved coding efficiency rather than that of the other. For the overall sequences, the proposed scheme have PNSR gains over 0.1dB.

Fig. 1, Fig. 2, and Fig. 3 show comparisons of the R-D curve during encoding "Carphone," "News," "Foreman" sequences. The solid line indicates the result by the proposed scheme and the dashed line is the result by H.263 TMN8 method. The proposed algorithm provides less PSNR fluctuation than H.263 TMN8. It means that our scheme can provide better quality stabilization.

TABLE II  
COMPARISON FOR TARGET BITS VS. CODING BITS, UNIT: KBITS

Test sequence	Target bits	Proposed algorithm	TMN8	FR(fps)
Carphone	36	35.73	36.00	10
Carphone	48	47.38	47.91	10
News	48	47.57	46.35	10
News	64	62.91	34.52	15
Foreman	64	63.56	64.01	15

TABLE III  
AVERAGE PSNR COMPARISON, UNIT: DB

Test sequence	Target bits	Proposed algorithm	H.263 TMN8
Carphone	36	31.61	31.53
Carphone	48	32.97	32.68
News	48	33.82	33.13
News	64	34.52	33.58
Foreman	64	30.78	30.50

## V. CONCLUSION

In this paper, we proposed an adaptive model-based rate control scheme considering frame variations. The frame variation is to allocate the target bits to the frame to be coded, and it does an role to reduce the exhaustive quality fluctuation. The quantization parameter decision at macroblock-level rate control is done by the rate-quantization model, which is designed based on the rate-distortion function of GGD. The experimental results describe the proposed rate control algorithm provide more regular PSNR curve than H.263 TMN8. We can also show that the proposed rate control scheme generates coding bits close to the target bits and provides improved coding efficiency at low bit rates.

## REFERENCES

- [1] MPEG-2, MPEG-2 Test Model5 (TM5) Doc. ISO/ IEC/ JTC1/SC29/WG11/N0400, Test Model Editing Committee, April 1993.
- [2] MPEG-4, MPEG-4 Video Verification Model version 16 (VM16) Doc. ISO/IEC JTC1/SC29/WG11/ N3312, March 2000.
- [3] "Video Codec Test Model, Near-Term, Version 8 (TMN8)," ITU-T VCEG(ITU-T SG16 Q.6), Q15B46.doc, September 1997.
- [4] J. Ribas-Corbera and S. Lei, "Rate Control in DCT Video Coding for Low-Delay Communications," *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 9, No. 1, pp.172-184, February 1999.
- [5] B. Tao, B.W. Dickinson, and H.A. Peterson, "Adaptive Model-Driven Bit Allocation for MPEG Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 10, No. 1, pp. 147-157, Feb. 2000.
- [6] Toby Berger, "Rate Distortion Theory," Englewood Cliffs, NJ:Prentice Hall, 1971, pp. 85-102
- [7] E.Y. Lam and J.W. Goodman, "A Mathematical Analysis of the DCT coefficient Distribution for Images," *IEEE Transactions on Image Processing*, Vol. 9, No. 10, pp.1661-1666, Oct. 2000.
- [8] G.S. Yovanof and S. Liu, "Statistical Analysis of the DCT Coefficients and Their Quantization Error," *The thirtieth Asilomar Conference on Signals, Systems and Computers*, Vol. 1, pp. 601-605, Nov. 1996.
- [9] S. Kim, S.-H. Kim, and Y.-S. Ho, "Adaptive Model-Based Quantization for H.264 Video Coding," *TENCON2004*, Vol. 1, pp. 331-334, Nov. 2004.
- [10] J.-C. Tsai and C.-H. Hsieh, "Modified TMN8 Rate Control for Low-Delay Video Communications," *IEEE Trans. on CSVT*, Vol. 14, No.6, pp. 864-868, June 2004.
- [11] H. Song and C.-C. Jay Kuo, "A Region-Based H.263+ Codec and Its Rate Control for Low VBR Video," *IEEE Trans. on Multimedia*, Vol. 6, No.3, pp. 489-500, June 2004.

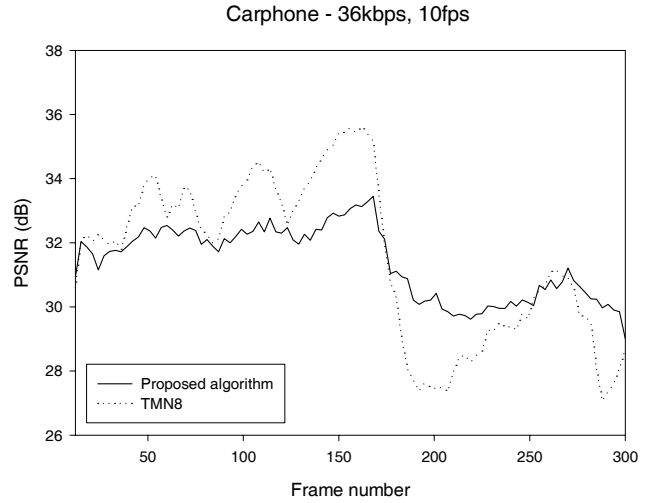


Fig. 1. R-D Curve: Carphone sequence, 10fps and 36kbps

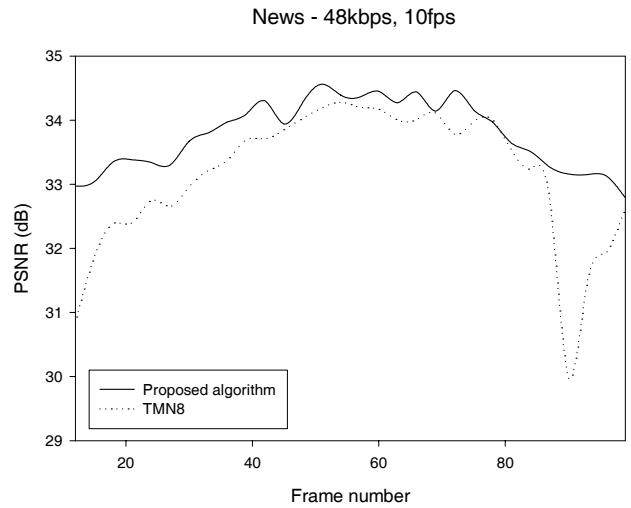


Fig. 2. R-D Curve: News sequence, 10fps and 48kbps

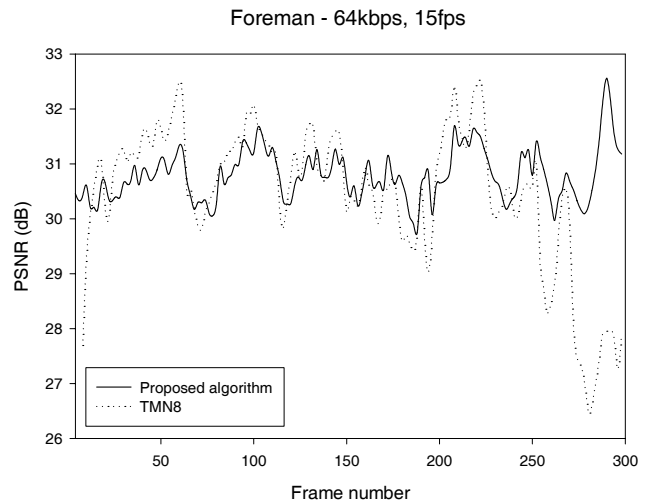


Fig. 3. R-D Curve: Foreman sequence, 15fps and 64kbps