ADAPTIVE MODEL-BASED QUNATIZATION FOR H.264 VIDEO RATE CONTROL

Seonki Kim, Seung-Hwan Kim and Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST)
1 Oryong-dong Buk-gu, Gwangju, 500-712, Korea
Email: {sunkee, kshkim, hoyo}@gist.ac.kr

ABSTRACT

In general, the source data to be quantized cannot be defined by a single distribution because it has a problem caused by the exhaustive approximation of the actual source distribution. In order to avoid this problem, we use the generalized Gaussian distribution (GGD) for modeling the distribution of the source. GGD is a parametric family of distributions, including the uniform, Laplacian, and Gaussian distributions as special cases. In this paper, we propose a quantization algorithm for H.264 video rate control based on the rate-quantization model, which is derived from the rate-distortion function of GGD. We implement our rate control scheme for low-delay inter-frame coding, and compare its performance with those of TM5 and TMN8 for various video sequences.

1. INTRODUCTION

The H.264 video coding standard is jointly developed by ISO/IEC and ITU-T [1]. By utilizing various predictive and entropy coding schemes, H.264 has substantially outperformed existing video coding standards. H.264 also adopts the rate-distortion optimization (RDO) using the Lagrangian method. However, H.264 does not include the rate control algorithm in the normative part of the video coding standard [1].

The main objective of developing H.264 is to enhance coding efficiency. The goal of the rate control scheme is to handle a trade-off between image quality and channel capacity. In order to deal with the trade-off, we need to select a proper quantization parameter that does not cause buffer underflow or overflow problems. We also consider the rate control scheme to improve visual quality under given coding conditions, such as frame rates and bit rates.

In this paper, we propose a new quantization algorithm for the H.264 video rate control. The proposed rate control algorithm, applied for inter-frames, is based on the rate-quantization model. In order to represent a proper source model, we need a shape factor that represents the difference between macroblocks in adjacent frames. Experimental results demonstrate that the proposed algorithm provides improved coding efficiency, compared to MPEG-2 TM5 and H.263 TMN8 rate control algorithms.

0-7803-8560-8/04/$20.00©2004IEEE

2. GENERALIZED GAUSSIAN DISTRIBUTION

In order to model the distribution of the source data in video coding, various ideas have been investigated. Design of quantization and coding parts relies on statistical modeling of the source data. Several probabilistic distributions have been used to model the source data for the rate control [2-3].

In general, the distribution of differential data to be quantized is modeled by the Laplacian distribution. Although the Laplacian distribution provides the most general statistical model for residuals, the source data cannot be properly modeled by a single distribution. In order to model the source distribution adaptively, the generalized Gaussian distribution (GGD) has been studied. It is a parametric family of distributions that include the uniform, Laplacian, and Gaussian distributions as special cases [4].

The approximation of the best probability density function for the source data can be adaptively achieved by GGD, defined by

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

where β is the shape parameter and σ is the standard deviation of the source.

When $\beta = 2$ and $\gamma = \sqrt{2}$, it becomes a standard Gaussian distribution. As $\beta \to 0$, $p(x)$ becomes an impulse function having tails and nonzero variance. As $\beta \to \infty$, it approaches the uniform distribution having variance σ^2. The Laplacian distribution is obtained by setting $\beta = 1$ and $\gamma = 1/\lambda$. The shape parameter β takes charge of the exponential rate of decay: if β increases, the distribution becomes flat; otherwise, the distribution is more peaked.

3. ADAPTIVE RATE CONTROL ALGORITHM

3.1. Rate-Quantization Model

In this work, we begin with the R-D function of GGD.

$$R = \frac{1}{\gamma} log_2 \left(\frac{\sigma^{\beta}}{D} \right) \quad (2)$$

where β is the shape factor, which describes a decay rate of the distribution, and γ is the coding parameter that depends on the shape factor and the standard deviation of the source.

As the shape factor changes, the distribution form is altered. Selection of the proper shape factor is important since the quantization parameter is determined by the source characteristics. The distortion is caused by the quantization parameter used for coding each macroblock. We define a distortion model that is definded as a function of the quantization parameter.

$$D = c \cdot Q^2 \tag{3}$$

where *c* is the distortion factor that represents the relationship between distortion and quantization parameter. If we use a large distortion factor in the important region, we can increase a quality of that region in the image. However, we do not consider the perceptual rate control for the special object in the image. We have just focused on the design of the quantization algorithm for the whole image.

By combining Eq. (2) and Eq. (3), we can derive a rate-quantization model as follows:

$$Q = \sqrt{\frac{\sigma^{\beta} \cdot 2^{-\gamma R}}{c}} \tag{4}$$

where the shape factor β and the coding parameter γ play important roles in modeling the source distribution.

3.2. Decision of Shape Factor

The shape factor is an important parameter to model the source distribution. However, the exact estimation of the shape factor for the source is not easy. It also requires a high computational complexity. In order to avoid this problem, we need to find an optimal shape factor with small complexity [5]. In this work, we simply decide the shape factor using similarity between two adjacent images. The shape parameter β is determined by

$$\beta = \begin{cases} 1, & ratio > a \\ 2, & ratio < b \\ 2 - \dfrac{ratio - a}{|b-a|}, & a \le ratio \le b \end{cases} \tag{5}$$

and

$$ratio = \frac{\textit{Number of pixels for } |P_c - P_p| < 2}{M} \tag{6}$$

where M denotes the number of pixels in the macroblock. P_c is a pixel value in the current image, and P_p is a pixel value in the previous image. In this work, we assume that the source distribution is modeled by one of forms existing between Laplacian and Gaussian distributions. If many pixels of the macroblock in the current image are similar with pixels of the macroblock in the previous image, the distribution of the source data is defined by the Laplacian distribution and the shape factor is one.

3.3. Quantization Parameter Decision

We should select a proper quantization parameter for each macroblock to obtain the best quality within the given target bits.

First of all, we need to predict the standard deviation for the current macroblock to determine the quantization parameter by the rate-quantization model. In H.264, the use of RDO creates a chicken-and-egg problem. In order to determine the quantization parameter, we need to compute the Lagrangian multiplier before starting the encoding process; however, the standard deviation of the residual data is obtained after motion compensation. This problem occurs in all the macroblock modes except the Inter16×16 mode. In order to solve this problem, we predict the standard deviation σ for the source data by

$$\sigma = \begin{cases} 1.1 \times \sigma_P, & \dfrac{\sigma_P^L + \sigma_P^T}{2} > 1.1 \times \sigma_P \\ 0.9 \times \sigma_P, & \dfrac{\sigma_P^L + \sigma_P^T}{2} < 0.9 \times \sigma_P \\ \dfrac{6 \times \sigma_P + 2 \times \{\sigma_P^L + \sigma_P^T\}}{10}, & \text{otherwise} \end{cases} \tag{7}$$

where σ is the predicted standard deviation for the current macroblock and σ_P is the actual standard deviation of the macroblock in the previous frame. σ_P^T and σ_P^B are standard deviations of the left and top macroblocks for σ_P in the previous frame, respectively.

The number of average target bits per pixel in the current macroblock is calculated by $R = B_R/(P_T - P_C)$, where P_T is the number of pixels in the frame, and P_C is the number of coded pixels for the frame. R is used to compute the quantization parameter by Eq. (4). B_R is the remaining bits to be used for encoding the remaining macroblocks in the current frame.

In the second step, we compute the quantization parameter for the macroblock. The calculated quantization parameter is rounded to the nearest integer value. The quantization parameter is then adjusted within ± 2 to reduce side effects by the difference of quantization parameters between the current and the previous macroblocks.

$$QP^* = \begin{cases} QP - 2, & \text{if } QP - QP_{prev} < -2 \\ QP + 2, & \text{if } QP - QP_{prev} > 2 \\ QP, & \text{otherwise} \end{cases} \tag{8}$$

where QP^* is the adjusted quantization parameter. After adjusting the quantization parameter, we set the quantization parameter value in the range of 0 to 51, as specified in the H.264 video coding standard [1].

In the third step, we update the counter and the model parameter. We update the number of remaining bits for coding macroblocks in the current frame and the coded macroblock counter. After each macroblock is coded, we increase the number of coded macroblocks by one. We also update the model parameter for the next macroblock to be coded.

3.4. Update Parameters at Macroblock Level

After encoding each macroblock, we update the coding information. The number of remaining bits in the current frame is updated by

$$B_R = B_R - B_{MB} \tag{9}$$

where B_R is the number of remaining bits for macroblocks in the frame, and B_{MB} is the number of coding bits used to encode the macroblock. B_R is decreased by B_{MB}.

We update the model parameter after encoding each macroblock. Let R_T be the target bits and R_A be coding bits. R_T can theoretically be determined by the rate-quantization model that is based on the rate-distortion function. Thus, we can define R_T and R_A as follows:

$$R_T = \frac{K}{\gamma(j-1)} \log_2 \frac{\sigma^\beta}{c \cdot Q^2} \tag{10}$$

$$R_A = \frac{K}{\gamma(j)} \log_2 \frac{\sigma^\beta}{c \cdot Q^2} \tag{11}$$

where K is the number of samples in the macroblock, $\gamma(j-1)$ is the model parameter used for the current macroblock, and $\gamma(j)$ is the model parameter that will be used for the next macroblock.

$$R_T - R_A = \left(\frac{1}{\gamma(j-1)} - \frac{1}{\gamma(j)} \right) \cdot K \cdot A \tag{12}$$

$$\Leftrightarrow \frac{1}{\gamma(j)} = \frac{1}{\gamma(j-1)} - \frac{1}{A} \left(\frac{R_T - R_A}{K} \right) = T \tag{13}$$

$$\gamma(j) = \frac{1}{T} \tag{14}$$

where

$$A = \log_2 \left(\frac{\sigma^\beta}{c \cdot Q^2} \right) \tag{15}$$

which is small.

3.5. Update Parameters at Frame Level

Before starting an encoding process, we allocate the number of bits for the GOP by considering frame rates and target bits. If the encoding for the last macroblock in the frame is finished, we update the number of remaining bits at GOP level.

$$B_{GOP} = B_{GOP} - T_{frame} \tag{16}$$

where B_{GOP} denotes the number of remaining bits in GOP and T_{frame} denotes the number of bits used to encode the frame.

We also update the model parameter for the frame as follows:

$$\gamma = \frac{1}{MB_{cnt}} \sum_{j=1}^{MB_{cnt}} \gamma(j) \tag{17}$$

where j is the macroblock number and MB_{cnt} is the number of macroblocks in the frame. γ is the model parameter for the first macroblock in the next frame.

4. EXPERIMENTAL RESULTS

4.1. Simulation Conditions

In this work, we set simulation conditions based on the baseline profile specified in the H.264 standard. Table 1 shows our simulation conditions.

Table 1. Simulation Conditions

RDO	On
GOP Structure	IPPP
Symbol Mode	CAVLC
MV search range	32
Reference Frames	1

The proposed algorithm is applied only to P-frames because B-frame coding is not included in the baseline profile. Therefore, we use the GOP structure of IPPP. For I-frames, we use the fixed quantization parameter value. We test two video sequences: ”Foreman” and ”News.” The format of both sequences is QCIF with 4:2:0. We set two target bits, 48 and 64 kbps, and the frame rate is 10 fps.

4.2. Performance Evaluation

In order to evaluate the performance of the proposed algorithm, we implement MPEG-2 TM5 and H.263 TMN8 rate control algorithms in the H.264 video codec. We also compare results of the proposed algorithm with Siwei’s rate control algorithm [6].

In Section 3.3, we explain the reason why the standard deviation should be predicted. We propose the standard deviation prediction method for the macroblock to solve that problem. Fig. 1 shows the comparison of the standard deviation curve for actual and predictive values. While the dashed-line indicates the curve for predictive values, the solid-line is for the original values.

One performance measure for the rate control algorithm is to check whether the number of coding bit is close to that of the target bits. As shown in Table 2, the proposed algorithm generates the number of coding bits close to the number of the target bits within 1% difference.

Table 3 compares average PSNR values for the proposed algorithm and other algorithms, including MPEG-2 TM5, H.263 TMN8, and Siwei’s algorithms. The proposed scheme improves coding efficiency over other schemes. For ”Foreman” sequence, we improved the average PSNR value by more than 1 dB. ”News” sequence also proves increased coding efficiency. H.263 TMN8 provides the lowest results.

Fig. 2 and Fig. 3 compares PSNR variations during encoding ”Foreman” and ”News” sequences, respectively. In Fig. 3, the solid line indicates the result by the proposed algorithm, and the dashed line is the result by the TM5 method. The proposed algorithm generally provides better coding efficiency than TM5.

5. CONCLUSIONS

In this paper, we proposed an adaptive model-based quantization scheme for effective coding of inter-frames. The proposed rate-quantization model is designed from the rate-distortion function of GGD, which can adaptively represent various distributions by changing the shape factor. The shape factor is simply obtained from the comparison of differences between macroblocks. The quantization parameter is calculated by the proposed rate-quantization model. Simulation results demonstrate that the proposed rate control scheme generates the coding bits close to the target bits and provides improved coding efficiency at low bit rates.

6. ACKNOWLEDGEMENTS

This work was supported in part by Gwangju Institute of Science and Technology (GIST), in part by the Ministry of Information and Communication (MIC) through the Realistic Broadcasting Research Center (RBRC) at GIST, and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.

7. REFERENCES

[1] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, "Study of Final Committee Draft of Joint Video Specification," Doc. JVT-F100, Dec. 2002.

[2] L.-J. Lin and A. Ortega, "Bit-Rate Control Using Peicewise Approximated Rate-Distortion Characteristics," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 8, No. 4, pp. 446-459, Aug. 1998.

[3] Z. He and S.K. Mitra, "A Linear Source Model and a Unified Rate Control Algorithm for DCT Video Coding," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 12, No. 11, pp. 970-982, Nov. 2002.

[4] E.Y. Lam and J.W. Goodman, "A Mathematical Analysis of the DCT coefficient Distribution for Images," *IEEE Trans. on Image Processing*, Vol. 9, No. 10, pp.1661-1666, Oct. 2000.

[5] B. Aiazzi, L. Alparone, and S. Baronti, "Estimation Based on Entropy Matching for Generalized Gaussian PDF Modeling," *IEEE Signal Processing Letters*, Vol. 6, No. 6, pp.138-140, June 1999.

[6] S. Ma, W. Gao, P. Gao, and Y. Lu, "Rate Control for Advanced Video Coding Standard," *IEEE International Symposium on Circuits and Systems*, Vol. 2, pp.892-895, May 2003.

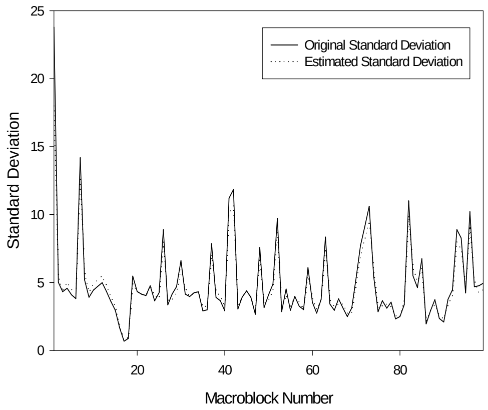

Fig. 1. Standard Deviations

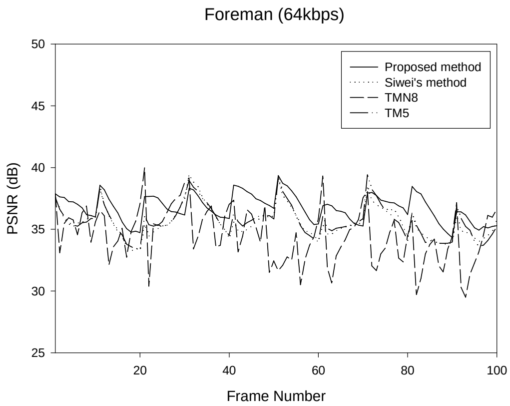

Fig. 2. PSNR Variations, "Foreman", 64 kbps

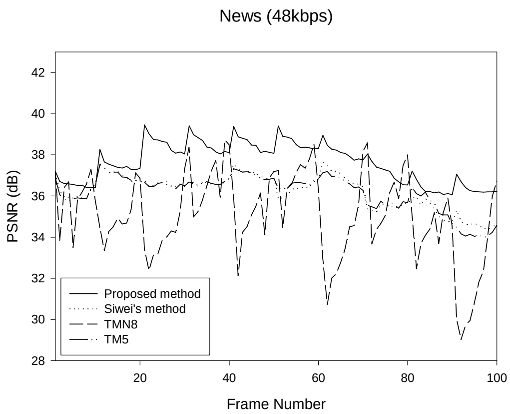

Fig. 3. PSNR Variations, "News", 48 kbps

Table 2. Target Bits vs. Coding Bits

Test Sequence	Target Bits (kbps)	Coding Bits (kbps)	Difference (kbps)
FM	48	48.05	0.05
FM	64	63.83	-0.17
News	48	47.93	-0.07
News	64	64.10	0.10

Table 3. Comparison of Average PSNRs

Test Sequence	TM5 (dB)	TMN8 (dB)	Siwei's (dB)	Proposed (dB)
FM(48)	34.13	33.28	34.15	35.19
FM(64)	35.67	34.52	35.60	36.73
NEWS(48)	36.20	35.02	36.23	37.63
NEWS(64)	38.09	37.37	38.22	39.78