저 전송률 부호화기를 위한 프레임 특성에 근간한 균등 비트 할당 기법

Equal Bit Rate Control for Low Bit-rate Coder based on Frame Statistics

서 동 완*, 최 윤 식*

Dong-Wan Seo*, Yoon-Sik Choe*

요 약

본 논문은 프레임간의 특성 변화를 이용한 균등 비트 할당 기법을 제안한다. 모델 기반 비트율 조정 알고리즘에 대한 기존의 연구는 양자화 파라미터와 비트율 및 왜곡 모델간의 관계를 찾는 것에 초점을 맞추어 진행해 왔다. 본 논문에서는 비트율-왜곡 모델을 새로이 설정하는 것 대신에 모델 기반 접근 알고리즘 내에서 프레임 특성 변화에 따라 모델 파라미터를 조절하여 모델 기반의 비트율 조정 알고리즘의 성능을 향상시킨다. 모델 기반 접근의 성능은 비트율-왜곡 모델 자체의 정확성과 프레임 간의 비트율-왜곡 특성 차이를 보상하는 모델 파라미터의 정확한 설정이 알고리즘의 성능을 좌우한다. 제안 알고리즘은 프레임의 특성 변화를 모델 파라미터에 의해 파악하고 모델 파라미터를 현재 프레임에 맞도록 조정한다. 제안 알고리즘은 MPEG-4 비디오 부호기를 이용하여 설계하고, 성능은 모델 기반 접근의 대표적인 방법인 TMN8과 비교하여 0.6dB 정도의 성능 향상을 나타낸다.

Abstract

This paper presents an equal bit rate control algorithm utilizing the statistical change between the previous frame and the current frame. The previous studies on the model-based rate control have focused on the models of bit rate and distortion in types of coders, in terms of the quantization parameter. The proposed algorithm improves the typical model-based rate control by updating a model parameter instead of modeling a better model of the rate and distortion. The proposed algorithm updates this model parameter by recognizing the change in statistics between the previous frame and the current frame. We implement the proposed algorithm in MPEG-4 coders and verify its performance while comparing it to the TMN8's approach (up to 0.6dB of improvement).

Key words : MPEG-4, Video Coding, Rate Control

I. Introduction

Real time video transmission has gained importance in recent years, thus generating the need for a stable and fine video compression algorithm. This provides the motivation to investigate rate control algorithm. The objective of rate control is to select the quantization parameters so that source distortion is minimized for a given coding bit budget [1][2].

In earlier studies of rate control, the buffer control algorithm has only been considered. The buffer control schemes typically have the objective of preventing overflow and underflow. After that time, improving image quality as well as controlling the buffer are considered by such techniques as Lagrangian

*연세대학교 전기전자공학부
접수 일자 : 2005. 10. 14   수정 완료: 2005. 10. 26
논문 번호 : 2005-4-2

optimization and Dynamic programming [3]. In these approaches, all of rates and distortions are computed in order to design the optimal quantizers. Due to high computational complexity, these schemes are not suitable for real time transmission of video data.

For "real-time" applications, a classical framework for rate control techniques is to use a model-based rate-distortion theory, which deals with the minimization of the model when measured for source distortion [4-6]. Most model-based rate control algorithms create models for measuring the encoding bit rate and distortion not from the current frame but from the previous frame. This modeling technique creates a problem in that bits are allocated unsuitably without regard to the statistics of the current frame. For instance, if too many bits are allocated to the earlier macroblocks of a frame due to complexity of the previous frame, few bits are allocated to the later macroblocks.

In this paper, we introduce a model-based bit-rate control algorithm by using the statistics more suitable to the current frame with low computational complexity. In order to resolve the unsuitable bit allocation problem, we must recognize the change in frame statistics. Because the existing methods set up initial model parameters calculated from the statistics of the previous frame, the identified quantizer for the current frame is affected by the previous frame. If we find a large amount of variance in the statistics between frames, we must encode the current frame by using the initial model parameter, which is obtained from the statistics of the current frame.

This paper is organized as follows. In Section Ⅱ, we introduce models of bit rate and distortion in the TMN8 [6]. The problem of existing model-based rate control is defined in Section Ⅲ. In Section Ⅳ, we show how to allocate bits suitably by using the statistics of the current frame. Finally, Sections Ⅴ and Ⅵ present our experimental results and conclusions, respectively.

Ⅱ. Model of Rate and Distortion

In a typical block-based video coder such as the H.263 and MPEG-4, bit-rate and distortion of macroblock(MB) are controlled by using a quantizer parameter. Therefore, TMN8[6] defines the model used for determining bit-rate $R(Q)$ and distortion $D(Q)$ by using a quantizer step size of a given MB explained as follows:

$$R(Q)=\sum_{i=1}^{N}A\left(K\frac{\sigma_i^2}{Q_i^2}+C\right) \tag{1}$$

$$D(Q)=\frac{1}{N}\sum_{i=1}^{N}\alpha_i^2\frac{Q_i^2}{12} \tag{2}$$

where A is the number of pixels in the MB, C is the length of the header, Q_i is the quantizer of the i-th MB, σ_i^2 is the variance of i-th MB, and α_i^2 is the distortion weight of i-th MB, and N is the number of macroblocks in one frame.

In order to find the optimal solution for the bit allocation, the rate control algorithm must obtain the set of optimal quantization step size, $Q_1^*, Q_2^*, \cdots, Q_N^*$, which minimizes the distortion subject to the constraint that the total number of bits, $\sum_{i=1}^{N}B_i$, must be equal to a given rate B. This bit allocation problem is defined as equation (3) and equation (3) is changed into the unconstrained problem as equation (4) by using a Lagrange multiplier (λ) [7].

$$Q_1^*, Q_2^*, \cdots, Q_N^* = argmin_{Q_1,Q_2,\cdots,Q_N}\sum_{i=1}^{N}D(Q_i) \quad s.t. \sum_{i=1}^{N}B_i \leq B \tag{3}$$

$$Q_1^*, Q_2^*, \cdots, Q_N^* = argmin_{Q_1,Q_2,\cdots,Q_N}\sum_{i=1}^{N}D(Q_i)+\lambda\sum_{i=1}^{N}B_i \tag{4}$$

The TMN8 solves the bit allocation problem with a differential equation in terms of the Lagrage multiplier and quantization step size of each MB. As a result, equation (5) is obtained as referred in [6]. As shown equation (5), it is useful to determine the quantizer by using the statistics related to the variance and the number of bits of i-th MB. The value of K in equation (5) stands for the statistics of i-th MB. The performance of this model-based rate control is up to whether the model parameter, K may catch up with the statistics of the current frame. So, we improve the overall gain by using more suitable model parameter to the current frame.

$$Q_i=\sqrt{\frac{AK}{B-ANC}\frac{\alpha_i}{\sigma_i}\sum_{i=1}^{N}\alpha_i\sigma_i} \tag{5}$$

III. Mismatched Model Statistics

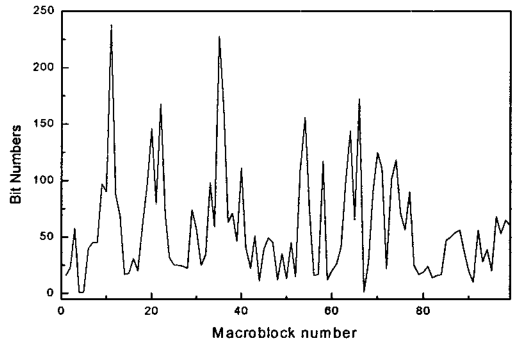
Fig. 1 illustrates the typical features of "Foreman" sequences, which in the earlier part of frames are filled with white blocks and the later parts of frames are composed of complicated blocks such as object's face, jacket collar, and hand. Here, we can easily guess that the earlier parts of the frames need only few bits, because these can be encoded by using only motion estimation. On the other hand, the complicated parts, such as the face and hand contain more information, so they must be allocated more bits than the earlier parts of the frames. In the TMN8, however, bits are unsuitably allocated, without regard to the information of the real current frame.

(a)

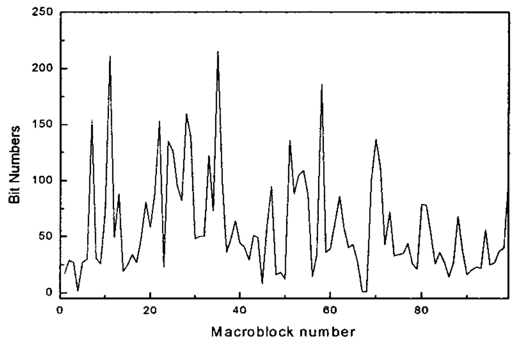
(b)

Fig. 1. "Foreman" sequence.
(a) 68-th frame, (b) 69-th frame.

As you can see in Fig. 2, the TMN8 rate control allocates many bits in the earlier parts of a frame, which consists of a white wall and a white helmet. On the other hand, fewer bits are allocated in the later parts of a frame, because too many bits are used in the earlier parts. This problem indicates that bits are allocated unequally, since the initial model parameter for the current frame is calculated from the statistics of the previous frame. If the statistics between the current frame and the previous frame change extremely, the problem of unsuitable bit allocation occurs. In this case, the model parameter cannot catch up with the statistics of the current frame. For example, if the previous frame is complicated, the earlier parts of the current frame use too many bits.



(a)



(b)

Fig. 2. Bit ditribution in the TMN8 (48kbps).
(a) 68-th frame, (b) 69-th frame.

IV. Model Parameter Update Rate Control based on frame statistics

In equation (4), parameters such as α_i and σ_i can be obtained independently at each MB. Then, these parameters have no relation to other MBs. However, the model parameter, K has dependency with the other MBs, which are presented as follows.

It is assumed that the current frame has similar statistics of the previous frame. The TMN8 encodes the first MB of the current frame using the model parameter calculated at the previous frame. It encodes the rest MB of the current frame using the weighted average value of the model parameter at the previous frame. In the TMN8, the model parameter obtained from previously encoded MB of the current frame like as equation (6).

$$K = \tilde{K}\frac{i}{N} + \hat{K}\frac{(N-i)}{N} \tag{6}$$

where $\tilde{K}$ is the average value of the model prameter obtained from previously encoded MB and $\hat{K}$ is the average value of the model parameter at the previous frame.

The TMN8 finds the quantization step size using equation (5) with the model parameter, K in equation (6). Then, the model parameter of the coded macroblock is calculated by using the real rate and variance of the coded MB as equation (7). The value of $\tilde{K}$ is the average value of the K_i.

$$K_i = \frac{B_{LC,i}(2QP)^2}{16^2\sigma_i^2} \tag{7}$$

where, $B_{LC,i}$ is the real bits of the i-th coded MB and QP is the quantization parameter of the coded MB.

We do not know the model parameter of the current MB because the current frame is not coded yet. So, The TMN8 guesses that the model parameter of the current MB and the remaining MB is estimated by the model parameter of the previous frame. The $\hat{K}$ substitudes the model parameter of the uncoded MB in the current frame.

The role of the model parameter is the controller of bit rate and distortion in a given MB. We can consider it as the statistics of a given frame. When there is a great difference in statistics between the current frame and the previous frame, there is a great difference between the model parameter of the current frame and the model parameter of the previous frame. After all, we can recognize a great change in statistics by comparing the current value of K with the $\hat{K}$ value obtained after encoding all MB's. We recognize the statistical change when the value of K does not satisfy the condition of $|\hat{K} - K| \leqq Th \cdot \hat{K}$. The threshold value (Th) is selected as 0.1 experimently. Consequently, if a change of statistics is found, we must encode again by using K obtained after encoding all MB's. In this re-encoding process, we only need quantization, inverse quantization and inverse DCT in terms of a new selected quantization parameter by using the updated K value. There is no doubt that when there is a small change, we need not encode the current frame again.

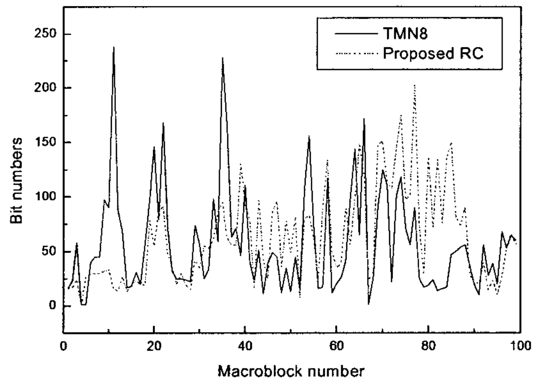


Fig. 3. The bit distribution of each macroblock comparing TMN8 and the proposed algorithm ("Foreman" sequence, 48kbps, 68-th frame)

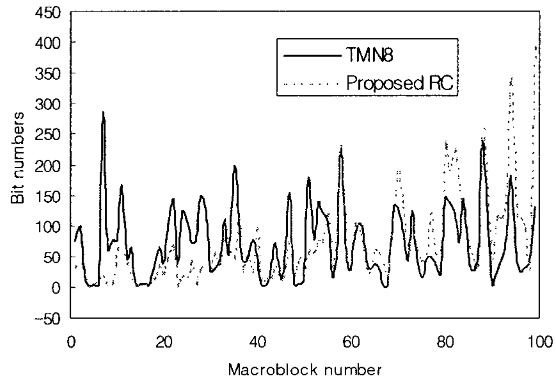
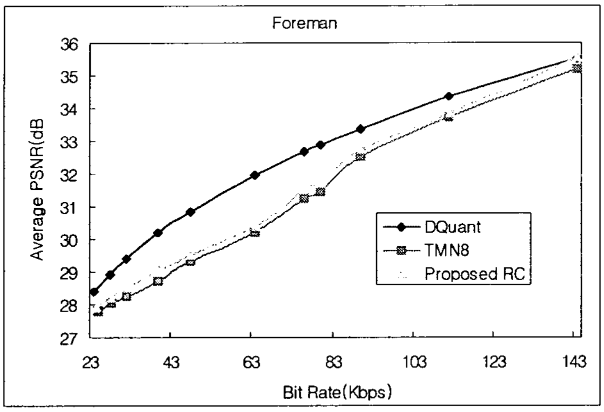


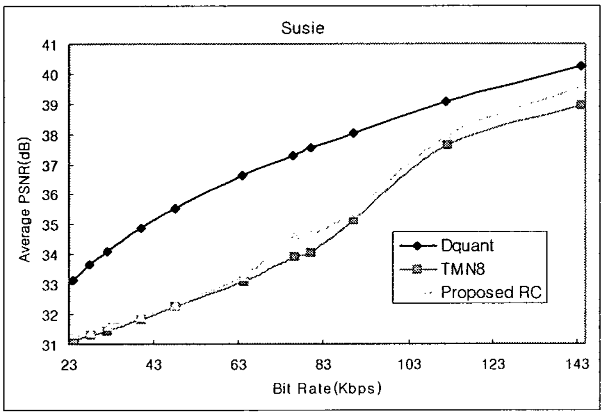
Fig. 4 The bit distribution of each macroblock comparing TMN8 and the proposed algorithm ("Foreman" sequence, 64kbps, 69-th frame)

V. Experimental Results

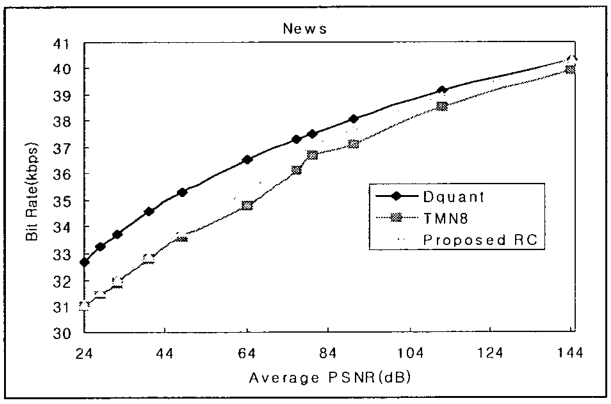
We implement our algorithm in MPEG-4 coders and compare its performance using R-D characteristics with the TMN8 and the optimal rate control algorithm which is our previous work [8]. The proposed algorithm evauated on the "Foreman" and "Susie" sequences which are in the standard QCIF format. We encoded these sequences based on experimental conditions that the target frame rate was 10 and Intra update did not exist for various bit rates between



(a)



(b)



(c)

Fig. 5 Rate-Distortion curves comparing, the TMN8 the optimal approach, and the proposed algorithms. (a) "Foreman" sequence, (b) "Susie" sequence, (c) "News" sequence.

24-140 kbps. We used $QP=5$ for the first I frame, interceded the rest as P frames using rate control.

Figure 3 and 4 show a bit distribution for each MB, which concerns the TMN8 and the proposed rate control agorithm. In the figure, the vertical axis denotes the number of bits used for each MB and the horizontal axis indicates the MB's index number for each frame in which there are 99 blocks in a frame. We can verify that the proposed rate control algorithm resolves the unequal bit allocation problem in the TMN8. The proposed algorithm allocates bits in each MB in accordance with the real complexity in a given frame. In Fig. 3, the proposed algorithm allocates 273 bits on the first 33 macroblocks While TMN8 allocates 1211 bits on the same macroblocks. In Fig. 4, the proposed algorithm allocates 751 bits on the first 33 macroblocks while the TMN8 allocates 2147 bits on the same macroblocks. As you can see in Fig. 3 and 4, the proposed algorithm allocates a substantial amount of bits especially in the complicated blocks such as the hands and the jacket of the worker.

Figure 5 represents the R-D plot of the TMN8, the proposed algorithm, and the optimal rate control method. We could identify about 20 frames of 94 frames that there is a great change in statistics. As we know from Fig. 4, our proposed algorithm has a more efficient performance average of 0.2 dB and maximally of 0.7 dB. Comparing frames which have the change of the statistics, our proposed algorithm has 2 dB gain maximally. Table 1 represents R-D gain after converting the same PSNR for comparing. It shows the rate gain at the same PSNR.

Table 1. R-D gain comparing.

	PSNR(dB)	TMN8 (kbps)	Proposed RC(kbps)	R-D gain (%)
Foreman	28	28.20	25.64	9.08
	32	85.09	82.85	2.63
	35	139.61	131.41	5.87
Susie	32	43.16	41.60	3.61
	35	88.68	84.60	4.60
	38	120.79	113.58	5.96
News	32	32.79	31.94	2.59
	35	64.90	60.49	6.79
	39	122.70	112.91	7.97

VI. Conclusions

To reduce the huge amount of computations required for offline rate control, significant efforts have been undertaken to speed up the process of video encoders. Those efforts have been mainly focused on the models used for bit rate and distortion in terms of the quantization parameters. Because previous research on the rate control schemes were related to the model-based rate control, which are based on the statistics of the previous frame, the previous model-based approaches occur the problem such that bits are allocated unequally without regard to current frame statistics. The problem needs to be addressed. In this paper, we propose an equal bit allocation scheme using the current frame statistics and solve the above problem through updating the initial model parameter. The experimental results show that the proposed method can achieve significant coding gains over TMN8's approach (up to 0.6 dB in PSNR).

References

[1] A.R. Reibman and B.G. Haskell, "Constraints on variable bit-rate video for ATM networks," *IEEE Trans. CSVT*, no. 2, pp. 361-372, 1992.

[2] C.Y. Hsu, A. Ortega, and A. Reibman, "Joint selection of source and channel rate for VBR transmission under ATM policing constraints," *IEEE J. Select Areas Commun.*, vol. 15, pp. 1016-1028, Aug. 1997.

[3] K. Ramchandran, A. Ortega and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders," *IEEE Trans. on IP*, vol. 3, no. 5, pp 553-545, 1994.

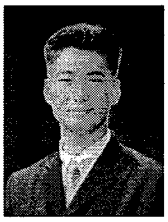
[4] Tihao Chiang and Ya-Qin Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. on CSVT*, vol. 7, no. 1, p. 246-250, 1997.

[5] Junfeng Bai, Qingmin Liao, Xinggang Lin, "Hybrid models of the rate-distortion characteristics for MPEG video coding," *WCC-ICCT 2000*, vol. 1, , pp 363-366. 2000.

[6] Ribas-Corbera, J. and Shwmin Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. on CSVT.*, vol. 9, no. 1, pp. 172-185, 1999.

[7] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizer," *IEEE Trans. on Acoustics, Speech Signal Process.*, vol. 36, no. 9, pp. 1445-1453, Sept. 1988.

[8] D. Seo, S. Han, Y. Kim, and Y. Choe, "Optimal quantization parameter set for MPEG-4 bit-rate control," *IEICE Trans. On Commun.*, vol. E87-B, no. 11, Nov. 2004.



Dong-Wan Seo received his B.S. degree in electrical engineering from the Yonsei University in 1999. He received M.S. degree in electrical and electronic engineering from the Yonsei University in 2001. He is working toward the Ph.D. degree in electrical and electronic engineering at the same university. His current research interests include multimedia communication system, transcoding technique, and bit-rate control.

Yoon-sik Choe received the B.S degree in electrical engineering from the Yonsei. University, Seoul Korea in 1979, and the M.S.E.E. in systems engineering, M.S. and Ph.D. degrees all in electrical engineering from the Case Western Reserve University, Cleveland, OH, the Pennsylvania State University, University Park, PA and the Purdue University, West Lafayette, IN, in 1984, 1987, and 1990, respectively. From 1990 to 1993 he was a Principal Engineering at the industrial electronics research center in the Hyundai Electronics Industries, Co. Ltd. Since 1993 he has been with the Department of Electrical and Electronic Engineering at the Yonsei University, Seoul Korea, where he is Associate Professor. His research interests include video coding, video communication, statistical signal processing and digital image processing.