토큰 버킷을 이용한 낮은 비트율 비디오의 실시간 비트율 제어

박상현* · 오원근*

Real-Time Rate Control with Token Bucket for Low Bit Rate Video

Sang-Hyun Park* · Won-Geun Oh*

이 논문은 순천대학교 공과대학 학술재단 학술 연구비 지원으로 수행되었음

요 약

낮은 대역폭을 가지는 네트웍을 통해 비디오를 전송하기 위해 개발된 H.263과 같은 낮은 비트율 비디오는 대역폭에 맞게 출력 데이터를 조절하여야 한다. 본 논문에서는 토큰 버킷 알고리즘을 사용하여 입력 트래픽을 관리하는 네트웍에 적합한 비트율 제어 알고리즘을 제안한다. 제안하는 알고리즘은 계산의 복잡도를 줄이기 위해 반복 연산을 수행하지 않는 최적화 방법을 사용하였고, 전체 시퀀스의 평균 왜곡을 최소화하면서도 인접 프레임 사이의 왜곡 차이를 최소화 하도록 알고리즘을 설계하였다. 비디오 품질의 변동을 감소시키기 위하여 제안된 방법은 슬라이딩 윈도우 개념을 도입하였으며, 이 방법은 선분석(pre-analysis) 처리를 요구하지 않는 특징이 있다. 따라서, 제안된 알고리즘은 비디오를 압축할 때 추가적인 지연을 필요로 하지 않게 되어 실시간 처리와 낮은 복잡도를 요구하는 비디오 압축기에 적용될 수 있다. 실험 결과에서 제안된 방법은 기존의 비트율 제어 알고리즘에 비해 영상의 품질이 우수한 것을 보여준다.

ABSTRACT

A real-time frame-layer rate control algorithm with a token bucket traffic shaper is proposed for low bit rate video coding. The proposed rate control method uses a non-iterative optimization method for low computational complexity, and performs bit allocation at the frame level to minimize the average distortion over an entire sequence as well as variations in distortion between frames. In order to reduce the quality fluctuation, we use a sliding window scheme which does not require the pre-analysis process. Therefore, the proposed algorithm does not produce time delay from encoding, and is suitable for real-time low-complexity video encoder. Experimental results indicate that the proposed control method provides better visual and PSNR performances than the existing rate control method.

키워드

frame-layer rate control, token bucket traffic shaper, video traffic management

Ⅰ. Introduction

Video communication over the Internet has become popular due to the fast development of networking and compression technologies. The digital video compression technique plays an important role in development of an

* 순천대학교 정보통신공학부

접수일자 : 2006. 9. 6

audiovisual communication system. A near-term enhancement of H.263 known as H.263+ [1] is suitable for low bit rate visual communications such as video over the Internet. To transmit compressed video efficiently over the Internet, we should consider both the underlying video content and channel conditions, and develop an effective rate control scheme accordingly.

Current Internet Protocol (IP) provides a service that can be characterized as a best-effort service. The best-effort model is sufficient for non-real-time video or non-real-time data communication applications like web browsing, file transfer, remote login, and electronic mail. In order to provide QoS guarantee to real-time applications, the IETF has defined the integrated services (IntServ) and the differentiated services (DiffServ). The IETF has specified a token bucket algorithm as a traffic policing mechanism in the IntServ [2]. In the case of the DiffServ, the token bucket algorithm is also one of the most commonly used methods in edge routers. In the DiffServ network, a service subscriber first sets up a service profile with an Internet service provider regarding the desired type of service. At the ingress of the DiffServ network, edge routers classify all packets passing through them into several predefined service classes, and mark the packets with different drop precedences according to the subscriber's service profile using the token bucket algorithm [3, 4]. Therefore, this work focuses on an effective rate control for low bit rate video over the Internet where video traffic is policed by the token bucket shaper.

In this paper, we propose a real-time frame-layer rate control method with a token bucket policing for low bit rate video compression standard, H.263+. In order to achieve accurate rate control, a new frame-layer rate-distortion (R-D) model is derived. Furthermore, we use a non-iterative method with low computational complexity for real-time rate control. The proposed sliding window scheme can reduce the quality fluctuation without the pre-analysis process. It is seen that the proposed rate control algorithm does not produce time delay from encoding.

The paper is organized as follows. In the next section, the conventional TMN8 rate control algorithm, which was designed for low-delay video communications, is briefly introduced. The proposed frame-layer rate control scheme is presented in Section 3. Section 4 presents and discusses the experimental results. Finally, our conclusions are given in Section 5.

Ⅱ. Conventional TMN8 Rate Control

In H.263+, the current video frame to be encoded is decomposed into macroblocks of 16×16 pixels per block, and the pixel values for each of the four 8×8 blocks in a macroblock are transformed into a set of coefficients using the DCT. These coefficients are then quantized and encoded with some type of variable-length coding. The number of bits and distortion for a given macroblock depend on the macroblock's quantization parameter used for quantizing the transformed coefficients. The TMN8 rate control uses a frame-layer rate control to select a target number of bits for the current frame and a macroblock-layer rate control to select the values of the quantization step-sizes for the macroblocks. In the following discussions, the following definitions are used:

- B : target number of bits for a frame;
- R : channel rate in bits per second;
- F : frame rate in frames per second;
- W : number of bits in the encoder buffer;
- M : maximum value indicating buffer fullness;
- W_{prev} : previous number of bits in the buffer;
- B' : actual number of bits used of encoding the previous frame.

In the frame-layer rate control, a target number of bits for the current frame is determined by

$$B = R/W - \Delta, \tag{1}$$

$$\Delta = \begin{cases} W/F, & if\ W > Z \cdot M, \\ W - Z \cdot M, & others, \end{cases} \tag{2}$$

$$W = max\,(W_{prev} + B' - R/F, 0), \tag{3}$$

where $Z = 0.1$ by default. The frame target varies depending on the nature of the video frame, the buffer fullness, and the channel throughput. To achieve low delay, the algorithm tries to maintain the buffer fullness at about 10 % of the maximum M. If W is larger than 10 % of the maximum M, the frame target B is slightly decreased. Otherwise, B is slightly increased.

Ⅲ. Proposed Frame-Layer Rate Control

For frame-layer rate control, we first estimate the rate and the distortion of the current frame as a function of the average quantization parameter (QP). Then, the target number of bits for the current frame is determined according to the parameters of the token bucket.

3.1 Estimation of Rate and Distortion

To estimate the rate and the distortion of each frame, we employ an empirical data-based frame-layer R-D model using the quadratic rate model and the affine distortion model [5] with respect to the average QP in a frame, which is given by

$$\hat{R}(\bar{q}_i) = (a \cdot \bar{q}_i^{-1} + b \cdot \bar{q}_i^{-2}) \cdot MAD(\hat{f}_{ref}, f_{cur}) \quad (4)$$

$$\hat{D}(\bar{q}_i) = a' \cdot \bar{q}_i + b', \quad (5)$$

where a, b, a', and b' are the model coefficients, $\bar{q}_i$ is the average QP of all macroblocks in the ith frame, and $\hat{R}(\bar{q}_i)$ and $\hat{D}(\bar{q}_i)$ are the estimated rate and the estimated distortion of the ith frame, respectively. Here, $\hat{f}_{ref}$ is the reconstructed reference frame at the previous time instant, f_{cur} is the uncompressed image at the current time instant, and $MAD(\cdot)$ is the mean of absolute difference between two frames.

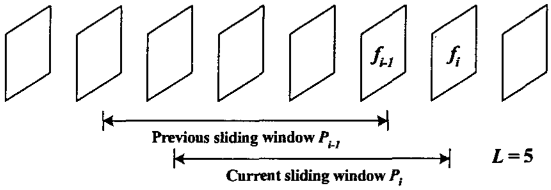


Fig. 1. Concept of the proposed sliding window method
그림 1. 제안된 슬라이딩 윈도우 방법의 개념

3.2 Determination of Target Number of Bits

The previous bit allocation approaches used a jumping window method for the analysis of scene characteristics [6]. However, the jumping window scheme requires a pre-analysis process for analyzing a group of frames before encoding, and thus causes an additional delay. We introduce a sliding window method that can analyze scene characteristics without time delay. Fig. 1 shows the sliding window method where the window moves one frame at a time to determine the target number of bits for each frame. P_i is the current sliding window consisting of frames $\{f_{i-L+1}, f_{i-L+2}, \cdots, f_{i-1}, f_i\}$, where L is the number of frames within the sliding window.

At the encoding time of the ith frame, the target bit rate R_i^T is determined using the sliding window P_i. In the proposed scheme, rate allocation is performed accordingly to achieve the constant video quality since token bucket policing allows the output rate to vary depending on the state of the token bucket.

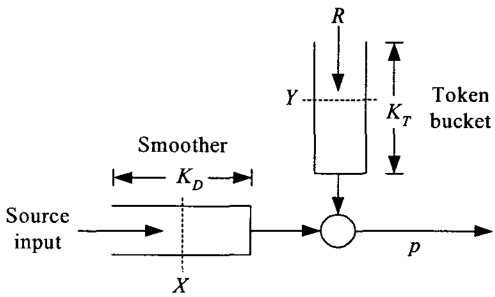


Fig. 2. Token bucket traffic shaper
그림 2. 토큰 버킷 트래픽 세이퍼

To manage the token bucket, we use a virtual buffer random variable method. Fig. 2 shows the token bucket shaper where K_D is the smoothing buffer size, K_T is the token bucket size, p is the peak rate, and R is the token generation rate which is equivalent to the channel rate of the CBR network. Two random variables X and Y represent the occupancy of the smoothing buffer and the token bucket, respectively. Note that the token bucket shaper has following properties: The smoothing buffer can be occupied only if the token bucket is empty. Conversely, the token bucket can be occupied only if the smoothing buffer is empty. We now define the following single virtual buffer random variable.

$$\widetilde{W} \equiv X - Y + K_T. \tag{6}$$

To preserve smooth video quality, we define the target distortion D_T at the encoding time of the ith frame as

$$D_i^T = \begin{cases} D_{i-1}, & \text{if } \; W \geq \alpha K or \; \widetilde{W} \leq \beta K \\ D_{i-1}^T, & others, \end{cases} \tag{7}$$

where $K = K_D + K_T$ and α and β are parameters used for the safety margin to avoid the buffer overflow and underflow, respectively. That is, D_i^T is changed to the distortion of the previous frame only if $\widetilde{W}$ is within the safety margin. Otherwise, D_i^T preserves the same value of D_{i-1}^T. In our experiment, α and β are set to 0.9 and 0.1, respectively.

Next, we find the target bit rate which causes the same distortion as D_i^T. With the estimated R-D curves and D_i^T, an adaptive rate allocation is easily obtained by

$$q_i^T = (D_i^T - b')/a', \tag{8}$$

$$R_i^T = \{a(q_i^T)^{-1} + b(q_i^T)^{-2}\} MAD(\hat{f}_{ref}, f_{cur}). \tag{9}$$

Since the bit allocation method depends on the MAD value and the estimated R-D model, the buffer overflow may occur. To avoid this problem, the target bit rate is recalculated using the previous virtual buffer occupancy $\widetilde{W}_{prev}$ as

$$\hat{R}_i^T = min(R_i^T, \alpha \cdot K + R/F - \widetilde{W}prev). \tag{10}$$

Once the target bit budget $\hat{R}_i^T$ is allocated to the current frame using the aforementioned frame-layer rate control, the TMN8 macroblock layer rate control algorithm allocates the bit budget to each macroblock.

Ⅳ. Experimental Results and Discussion

Extensive experimental testing and comparison were performed on two sequences with different characteristics: Foreman and Carphone. These sequences are in QCIF format (176×144) and the frame rate F is 30 fps. The token generation rate R, the size of the sliding window L, and the size of the smoothing buffer K_D are set to 64kbps, 12, and R/F, respectively.

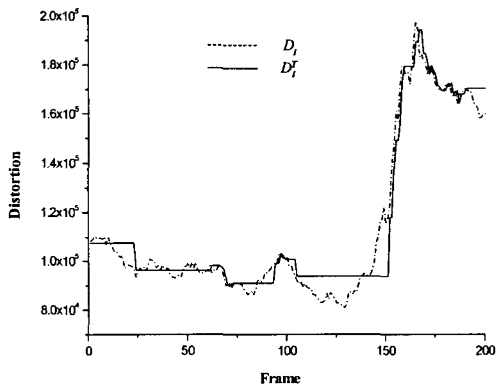


Fig. 3. Comparison between the actual distortion and the target distortion
그림 3. 실제 distortion과 목표 distortion 비교.

We first present simulation results to validate the proposed method to determine the target distortion D_i^T in (7). Fig. 3

shows the actual distortion and the target distortion for each frame of the Carphone sequence. As shown in this figure, the proposed method works reasonably well. Note that D_i^T preserves the same value for some period since D_i^T has the same value of D_{i-1}^T while $\widetilde{W}$ is less than $\alpha \cdot K$ and greater than $\beta \cdot K$. It is also seen that D_i is properly regulated by D_i^T.

Table 1. Performance comparison of the proposed algorithm with TMN8
표 1. 토큰 버킷의 크기에 따른 제안된 알고리즘의 성능과 TMN8의 성능 비교

Test sequence	Rate control method	Average PSNR	σ of PSNR
Foreman	TMN8	29.52	0.825
	Proposed(R/F)	29.70	0.823
	Proposed($3R/F$)	29.84	0.724
	Proposed($6R/F$)	29.97	0.557
Carphone	TMN8	30.99	2.464
	Proposed(R/F)	31.16	2.371
	Proposed($3R/F$)	31.24	2.357
	Proposed($6R/F$)	31.33	2.276

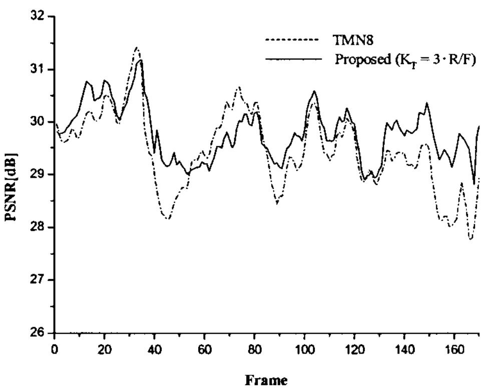


Fig. 4. PSNR comparison with a token generation rate at 64kbps
그림 4. 64kbps의 토큰 생성율에서의 PNSN 비교. QCIF Foreman sequence, $K_T = 3 \cdot R/F$.

Performance was mainly evaluated by peak-signal-to-noise ratio (PSNR) which is an objective measurement of the distance between an original image $f(x,y)$ and its reconstructed image $g(x,y)$. For an $m \times n$ image with [0, 255] gray-level range, PSNR can be defined as

$$PSNR = 10 log_{10} \frac{m \times n \times 255^2}{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \{(f(x,y) - g(x,y))\}^2}. \quad (11)$$

The performance of the proposed frame-layer rate control sheme is compared with that of TMN8. For the performance comparison for two test sequences, we show the average PSNR value and the standard deviation (σ) of PSNR in Table 1. Note that when compared with TMN8, the proposed frame rate control algorithm can not only improve the average PSNR value, but also reduce the variation of PSNR. It is also seen that the quality degradation and the quality variation decrease as the size of the token bucket increases. The PSNR plot associated with the Foreman sequences as a function of the frame number is shown in Fig. 4. Note that the proposed frame rate control scheme can reduce the quality degradation and the quality variation better than TMN8.

V. Conclusion

In this paper, we presented a new real-time frame-layer rate control for H.263+ over the networks where flows are regulated by a token bucket. In order to reduce the quality fluctuation, the sliding window method has been utilized. In addition, we introduced a frame-layer rate control scheme to minimize the average distortion over an entire sequences as well as variations in distortion between frames with the target distortion method. Since the proposed technique uses fast convergence method and does not require pre-analysis, it is suitable to real-time low-complexity video coding. The proposed algorithm has been tested on several sequences, and found to provide better visual and PSNR performances than the existing TMN8 rate control algorithm.

References

[1] ITU-T, Video coding for low bit rate communication, ITU-T Recommendation H.263 version 2, Jan. 1998.

[2] C Metz, "IP QoS: Traveling in First Class on the Internet," *IEEE Internet Computing*, vol. 3, pp. 84-88, Mar. 1999.

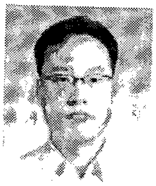
[3] H.Su and M. Atiquzzaman, "Comprehensive Performance Model of Differentiated Service with Token Bucket Marker," *IEE Proceedings-Communications*, vol. 150, pp. 347-353, Oct. 2003.

[4] E. C. Park and C. H. Choi, "Adaptive Token Bucket Algorithm for Fair Bandwidth Allocation in DiffServ Networks," in *Proc. IEEE Globecom'03*, pp. 3176-3180, 2003.

[5] T. Chiang and Y. Zhang, "A New Rate Control Scheme using Quadratic Rate Distortion Model," *IEEE Trans. Circuit Syst. Video Tech*, pp. 246-250, Feb. 1997.

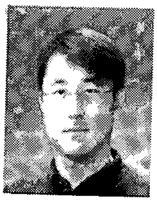
[6] Y. Yokoyama and S. Nogaki, "A Rate Control Method with Pre-analysis for Real-time MPEG-2 video coding," in *Proc. of Int. Conf. Image Processing*, pp. 514-517, Oct. 2001.

저자소개



박 상 현(Sang-Hyun Park)

1995년 2월 고려대학교 전자공학과 학사
1997년 2월 고려대학교 전자공학과 석사
2002년 2월 고려대학교 전자공학과 박사
2004년 2월 ~ 현재 순천대학교 정보통신공학부 조교수
※관심분야 : 영상처리, 영상압축, 멀티미디어통신



오 원 근(Won-Geun Oh)

1989년 2월 한양대학교 전자통신공학과 학사
1991년 2월 한양대학교 대학원 전자통신공학과 석사
1997월 2월 한양대학교 대학원 전자통신공학과 박사
1997년 3월 ~ 현재 순천대학교 정보통신공학부 부교수
※관심분야 : RFID, 음향신호처리, 지능시스템