

특집논문-08-13-6-03

실시간 H.264/AVC를 위한 적응적인 프레임 단위 비트율 제어 기법

김명진^{a)}, 김경환^{a)}, 홍민철^{a)‡}

Adaptive Rate Control in Frame-level for Real-time H.264/AVC

Myoung-Jin Kim^{a)}, Kyoung-Hwan Kim^{a)}, and Min-Cheol Hong^{a)‡}

요약

본 논문에서는 실시간 H.264/AVC를 위한 적응적인 프레임 단위 비트율 제어 기법에 대해 제안한다. QP값에 의한 프레임당 발생 비트량은 영상의 특성에 따라 다양하게 발생하고 있으며, 부호화하고자 하는 프레임은 인접한 프레임들과의 일정한 상관관계를 유지하는 특성이 존재한다. 제안 방식은 이러한 통계적 특성을 이용하여, 영상의 특성별 QP값에 따른 발생 비트량을 적용시키기 위하여 예측된 영상의 복잡도에 따라 QP값에 대한 발생 비트량의 변화를 구하여 부호화하고자 하는 프레임의 예측 비트량에 적용하고 이전 프레임들의 통계 정보인 QP와 발생비트량과의 가중치를 이용하여 현재 프레임의 복잡도를 예측한다. 실험 결과를 통해 동일 비트율에서 PSNR 성능 저하없이 기존 방식 대비 99% 이상 계산량 감소가 있음을 확인할 수 있었다.

Abstract

In this paper, we propose an adaptive rate control in frame-level for real-time H.264/AVC. For given QP, bits according to video characteristics, and current frame is close correlation between the adjacent frames. Using the statistical characteristic, we obtain change of occurrence bit about QP to apply the bit amount by QP from the video characteristic and applied in the estimated bit amount of the current frame. In addition, we use weight with QP and occurrence bit amount that is statistical information of encoded previous frames. Simulation results show that the proposed rate control scheme achieves time saving of more than 99% over JM 12.1 rate control algorithm. Nevertheless, PSNR and bit rate were almost same as the performances of JM.

Keywords : 비트율 제어, H.264/AVC, 통계적 특성, 영상 복잡도 예측, QP

1. 서론

H.264/AVC는 ITU와 ISO/IEC가 함께 진행하여 향상된 비디오 부호화 표준을 제공하기 위해서 개발된 국제 압축 표준으로서, 다양한 예측 부호화 기법과 엔트로피 부호화 방법들을 통해, H.264/AVC 비디오 부호화는 기존의 비디

오 부호화 표준들과 비교하여 최대 50%까지의 비트율을 절감할 수 있어 보다 향상된 부호화 효율을 제공한다^[1-8].

영상정보를 실시간으로 전송시, 영상의 특징과 종류에 따라 생성되는 비트량은 시간적으로 변화가 심할 수 있게 되는데, 이는 전송되는 채널의 특성상 매우 제한적이며, 고정된 대역폭을 가지고 있기 때문이다. 따라서 영상의 품질과 채널 용량 간의 균형을 조절하여 가장 좋은 부호화 성능을 제공하기 위해서는 부호화 변수들을 조절하는 것이 필수적이라 할 수 있다. 이와 같은 부호화 성능을 조절하기 위하여 H.264의 비트율 제어 알고리즘은 비트율-왜곡 최적

a) 숭실대학교 정보통신공학과

School of Electronic Engineering, Soongsil University

‡ 교신저자 : 홍민철 (mhong@ssu.ac.kr)

* 본 논문은 학술진흥재단 중점연구소 지원사업 (KRF-2006-005-J03801) 및 숭실대학교 교내 연구비 지원 사업에 의해 수행되었음.

화(RDO: Rate Distortion Optimization) 방법을 기반으로 하여 발생 비트와 양자화 왜곡의 비율을 효율적으로 조절함으로써 부호화 전에 각 프레임을 위한 적절한 QP (Quantization Parameter)를 결정하고, 각 부호화 모드의 비용을 측정하여 최소 비용의 모드를 최종 모드로 결정한다^[9-13]. 그러므로 H.264/AVC에서의 비트율 제어 방식은 모드 결정 및 움직임 벡터 결정을 하기 이전에 최적의 QP를 선택해야 하는 문제점이 발생한다. 이 문제를 해결하기 위해 해당 매크로 블록의 왜곡 값을 MAD(Mean Absolute Different)를 통해 이전 프레임의 매크로 블록으로부터 추정되는 기법이 사용된다. M. Jiang은 이런 MAD 기법을 개선하여, GOP 구간에서의 부호화된 P 프레임들의 평균 MAD값에 대한 현재 프레임의 예측된 MAD 값을 이용하여 프레임 복잡도를 예측하는 모델^[13]과 이 MAD 율과 함께 부호화된 P 프레임들의 평균 PSNR 오차율을 이용하여 현재 프레임의 복잡도를 예측하여 개선된 성능을 보여 주고 있다^[14]. 그러나 이와 같은 MAD 기법들은 많은 계산량을 필요로 하기 때문에 동영상의 응용에 많은 제약은 갖게 된다.

본 논문에서는 이와 같은 문제점을 해결하기 위하여 QP와 부호화 이후에 발생하는 실제 비트량과의 선형적 모델을 기반으로 하여 기존의 방식보다 복잡도를 줄이는 방식을 사용함으로써, 적은 계산량으로도 프레임의 발생 비트량을 정확히 예측하여 기존의 방식보다 비트율 제어 알고리즘의 전체적인 계산량을 줄이는 기법을 제안한다. H.264/AVC는 VBR(가변 비트율 부호화 방법)과 CBR(고정 비트율 부호화 방법)을 동시에 지원하지만, 이 논문에서는 프레임 단위를 기반으로 한 CBR 비트율 조절 방법만을 논문의 주제로 고려하겠다.

본 논문의 구성은 다음과 같다. II장에서는 기존 H.264/AVC의 비트율 제어 알고리즘에 대하여 소개하고, III장에서는 QP와 발생 비트량과의 선형적 모델을 기반으로, 프레임의 복잡도를 예측하여 적은 계산량으로도 효율적인 비트량을 할당하는 알고리즘을 제안한다. IV장에서는 기존 비트율 제어 알고리즘과 제안한 비트율 제어 알고리즘의 시뮬레이션 결과를 비교 분석한다. 마지막으로 V장에서는 본 논문의 결론을 맺는다.

II. H.264/AVC의 비트율 제어 알고리즘

H.264/AVC를 효율적인 비디오 부호화 표준으로 만드는 많은 기술들 중 하나로, 율-왜곡 최적화(Rate-Distortion Optimization) 방법을 들 수 있는데, 움직임 추정에 쓰이는 가변블록 모드와 공간 예측 모드(Intra16x16, Intra4x4), Skip 모드 등 모든 가능한 부호화 모드 중에서 가장 최적의 모드를 선택하는 방법이다. 이 방법은 효율화의 장점과 함께 연산량 증가에 의한 부호화기 복잡도의 증가라는 단점이 있다^[16]. 또한 율-왜곡 최적화 방법은 비트율 조절의 측면에서도 매우 다루기 까다로운 방법으로, 율-왜곡 최적화를 위해서는 모든 부호화 모드를 재귀적으로 검사를 하면서 최적의 모드를 선택하는데, 최적화가 끝난 후에 QP의 선택을 위해 사용되는 여러 변수들이 결정된다. 그림 1은 H.264/AVC 참조 모델의 비트율 제어 방법을 나타낸 것으로 목표 비트율이 할당되면 프레임 단위로 양자화 계수를 결정하여 부호화한다. 그런 다음 비트율-왜곡 최적화를 수행한 후, 화질과 비트율을 모두 만족하는 매크로블록 모드를 부호화한다.

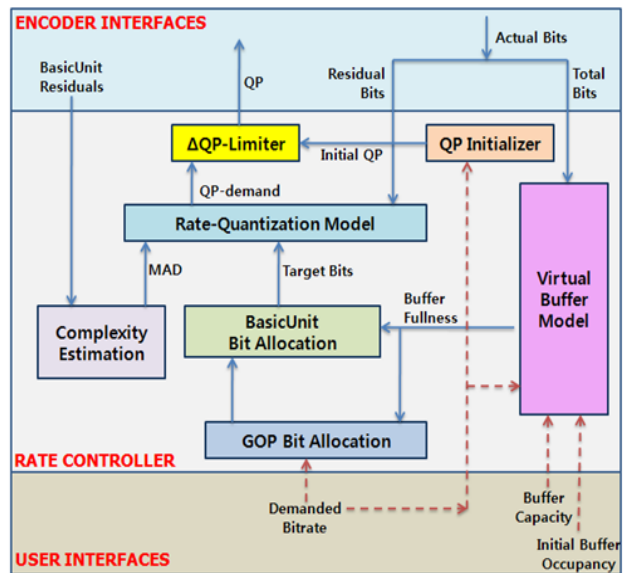


그림 1. H.264/AVC 참조 모델의 비트율 제어 방법
Fig. 1. The rate control of H.264/AVC reference model

1. 비트율-왜곡 최적화

H.264/AVC의 비트율-왜곡 최적화(RDO:Rate Distortion Optimization)의 최종 목적은 부호화측에서는 최소의 비트율을 부호화하여 손실없이 최종단말기에 전송하는데 목적이 있고, 복호화측에서는 최대의 화질을 제공받으자 왜곡 값이 최소인 매크로블록을 최적의 매크로블록 모드로 결정한다. H.264/AVC의 테스트 모델은 라그랑지 최적화 모드와 움직임 추정 방법^[6]에 대한 방법을 채택하고 있다. 이 방법에서 라그랑지 계수(λ)는 다음 식 (1)과 같이 각각의 매크로블록에 대하여 양자화 계수(QP)를 사용하여 라그랑지 값을 결정한다.

$$\lambda = 0.85 \times 2^{\frac{QP-12}{3}} \quad (1)$$

라그랑지 최적화 방법^[17,18]에서 라그랑지 계수가 작을수록 높은 비트율과 좋은 화질을 보인 반면에 라그랑지 계수가 높을수록 낮은 비트율과 나쁜 화질을 초래한다는 결과와 같이 라그랑지 계수는 화질과 비트율에 영향을 끼치는 것을 알 수 있다.

2. 프레임에 대한 초기화

비디오 시퀀스는 GOP(Group of Picture)로 구성되어 있고, GOP의 첫 번째 프레임은 인트라(Intra) 모드로 부호화하고, 그 이후의 프레임은 인터(Inter) 모드로 부호화하는 것이 일반적이다. 프레임 당 비트율을 제어하기 위해서는 먼저 첫 번째 프레임에 대한 가상버퍼 점유량(Virtual Buffer Occupancy)을 정해야 하는데, 이 첫 번째 가상 버퍼 점유량이 $B_c(1)$ 이면, 이때 초기 가상버퍼 점유량은 버퍼 사이즈(B_s)를 다음과 같이 1/8로 설정한다.

$$B_c(1) = \frac{B_s}{8} \quad (2)$$

H.264/AVC 참조 모델은 현재 버퍼 상태에 적합하도록

다음 식 (3)과 같이 초기 목표 버퍼레벨(TBL: Target Buffer Level)을 설정한다.

$$TBL(2) = B_c(2) \quad (3)$$

그리고 버퍼의 넘침(Overflow)과 고갈(Underflow)을 방지하기 위하여 버퍼의 상한($U(1)$) 및 하한($L(1)$)의 초기값을 식 (4)과 같이 설정한다. 이때 u 은 채널 대역폭으로 압축된 영상 시퀀스를 일정하게 전송하는 비트율이고, F_r 은 초당 부호화된 프레임 율이다.

$$\begin{aligned} L(1) &= \frac{u}{F_r} \\ U(1) &= \alpha \times B_s, (\alpha = 0.8) \end{aligned} \quad (4)$$

3. 프레임에 대한 목표 비트량 계산

n 번째 프레임을 부호화할 때마다 가상버퍼 점유량($B_c(n)$) 및 목표 버퍼레벨($TBL(n)$)은 각각 식 (4)와 식 (5)와 같이 유동적으로 변하는 유동 트래픽 모델(Fluid traffic model) 식을 따르고 있다.

$$B_c(n) = \min(\max(0, B_c(n-1) + A(n-1) - \frac{u}{F_r}), B_s), \quad 3 \leq n \leq N \quad (5)$$

$$TBL(n) = TBL(n-1) - \frac{TBL(2) - B_s/8}{N_p - 1} \quad (6)$$

식 (5)에서 $A(n-1)$ 는 $(n-1)$ 번째 프레임에 대한 부호화된 비트량이고, 식 (6)에서 N_p 는 하나의 GOP 내에 있는 P 프레임의 개수이다. n 번째 프레임을 부호화하기 이전에, 프레임 당 할당할 수 있는 목표 비트량을 예측함에 있어서 버퍼량의 예측이 필요하다. 이때 버퍼량 예측은 가상버퍼 점유량 및 목표 버퍼레벨을 이용하여 식 (7)과 같이 예측할 수 있다.

$$T'_{buf}(n) = \frac{u}{F_r} + r(TBL(n) - B_c(n)), \quad r = 0.7 \quad (7)$$

이렇게 프레임 당 버퍼량이 예측되면 버퍼 상태를 고려하여 식 (8)과 같이 버퍼의 상한과 하한 범위까지 고려하여 최종적으로 버퍼량($T_{buf}(n)$) 값을 결정한다.

$$T_{buf}(n) = \min(U(n), \max(L(n), T_{buf}(n))) \quad (8)$$

n번째 프레임에 대한 목표 비트량 $T(n)$ 는 식 (8)에서 예측한 버퍼량과 식 (9)에서 예측한 부호 비트량을 고려하여 식 (10)과 같이 최종적으로 결정된다.

$$T_r = \frac{R}{N-n} \quad (9)$$

$$T(n) = \beta \times T_r + (1-\beta) \times T_{buf}, (\beta = 0.5) \quad (10)$$

식 (9)에서 T_r 은 프레임당 부호화할 수 있는 비트량으로 현재 남아 있는 비트량(R)과 앞으로 부호화할 프레임 개수 즉, 남은 프레임 개수($N-n$)를 이용하여 계산한다.

4. MAD 예측 및 양자화 계수 예측

비트율 제어를 위해 먼저 현재 프레임과 참조 프레임간의 차이를 나타내는 MAD 를 이용하며, 현재 프레임의 MAD_c 를 예측하기 위해 식 (11)과 같이 계산한다.

$$MAD_c = C_1 \times MAD_p + C_2 \quad (11)$$

$$C_1 = \frac{1}{20} \sum_{i=1}^{20} \frac{MAD_i}{MAD_{i-1}}$$

식 (11)에서 MAD_p 는 이전 프레임의 MAD 를 나타내며 C_2 는 예측된 MAD 값들의 예측 오차에 대한 보상 값이다. C_1 과 C_2 에 대한 초기값은 각각 1과 0이다. 현재 H.264/AVC의 비트율을 제어하기 위한 프레임 단위에 대한 양자화 계수는 식 (12)과 같이 2차 방정식 비트율-왜곡 모델을 사용한다.

$$T(n) = \frac{x_1 \times MAD_n}{QP} + \frac{x_2 \times MAD_n}{QP^2} \quad (12)$$

식 (12)에서 $T(n)$ 는 n번째 프레임에서의 목표 비트량, MAD_n 는 n번째 프레임에서의 MAD , QP 는 양자화 계수, 그리고 x_1 와 x_2 는 각 프레임에서 얻어지는 2차원 비트율-왜곡값을 나타내는 변수이다.

5. 부호화 이후의 버퍼 상태 갱신

예측한 목표 비트량을 이용하여 부호화 수행 후, 현재 버퍼의 상한 및 하한 값을 식 (13)과 같이 갱신하여, 다음 프레임에서 버퍼가 넘치거나 고갈되지 않도록 한다. $\varpi (= 0.8)$ 은 버퍼의 넘침을 방지하기 위한 상수이다.

$$L(n) = \min(\max(0, L(n-1) + \frac{u(n-1)}{F_r} - A(n-1)), B_s) \quad (13)$$

$$U(n) = \max(\min(U(n-1) + (\frac{u(n-1)}{F_r} - A(n-1)) \times \varpi), B_s), 0)$$

III. 제안 방식

H.264/AVC에서 사용되고 있는 비트율 제어 기법은 실제 부호화에 의한 최적의 부호화 파라미터를 찾는 것으로, 라그랑지안 최적화 기법을 이용하여 가능한 부호화 파라미터 조합 중 주어진 비트율 조건을 만족시키고 왜곡을 최소화 하는 것을 예측하는 방식을 사용하고 있다. H.264/AVC에서 사용되는 비트율 조절 방법^[13]은 비트와 왜곡이 양자화 스템의 함수로서 표현되는 Quadratic R-D 모델에 기반을 두고 있다^[10]. 이 방법은 이미지에 대한 분석의 간단함 및 양자화에 대한 직접적인 조절 방법 등의 특성에도 불구하고, 비디오 부호화의 다양한 특성 때문에, 경우에 따라서는 비효율적인 비트율 조절 결과를 나타내는 경우가 있다. 이러한 방식은 가장 이상적인 방법이지만 모든 양자화 파

라미터에 따른 실제 비트율과 왜곡 값을 구해야 하므로 많은 계산량의 문제점으로 인해 실시간 멀티미디어 전송시에 부적합하다. 영상 정보의 실시간 전송은 대상 응용에 맞는 비트량을 예측하는 과정에서 불필요한 계산량을 줄이고 영상의 왜곡을 최소화하는 것을 목적으로 한다.

본 방식에서는 비트량을 예측하는 과정에서의 계산량을 줄이기 위하여 통계적 모델에 기반한 비트율 제어 기법을 이용하여 대상 응용에 적합하게 비트량을 산출하고 영상의 왜곡을 최소화하여 영상 정보를 실시간으로 부호화하는 기법에 대해 기술한다.

그림 2는 제안 방식에서의 전체적인 계층 구조를 나타낸 것으로서, 4개의 주요 처리 단계로 이루어져 있다.

- 1) 부호화를 위한 주요 파라미터 초기화
- 2) 프레임에 대한 목표 비트량 계산
- 3) 목표 비트량에 대한 QP값 결정
- 4) 프레임에 대한 복잡도 갱신

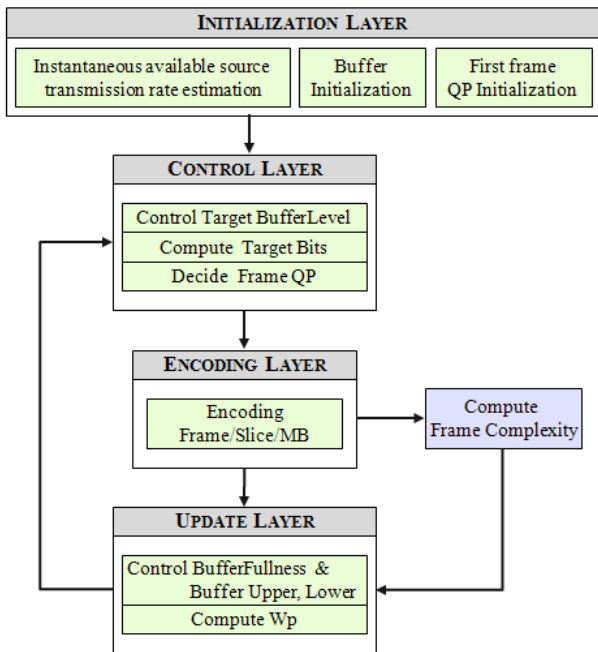


그림 2. 제안된 방식의 계층적인 구조
Fig. 2. The hierarchical diagram of the proposed rate control algorithm

단계별 내용에 대한 구체적인 설명은 다음과 같다.

1. 부호화를 위한 파라미터 초기화

단계 1에서는 부호화를 위한 프레임 정보 및 버퍼의 초기화, 그리고 영상의 초기 QP 값 등의 파라미터들을 초기화한다. I 프레임에 할당되는 비트량은 전체 영상의 화질에 큰 영향을 미치게 된다^[11]. 따라서 I 프레임의 QP를 결정하는 것은, I 프레임을 뒤따르는 프레임들의 화질에 영향을 미치므로, I 프레임의 QP를 결정하는 것은 매우 중요하다.

본 논문에서의 I 프레임에 대한 파라미터 초기화 과정은 H.264/AVC의 참조 모델 방식^[13]을 사용한다. 이 방법은 각 픽셀에 할당할 수 있는 비트수(bpp; bit per pixel)를 고려하여 QP값을 결정하는 방법으로 식 (14)와 같이 구한다.

$$bpp = \frac{u}{F_r \times (width \times height)} \tag{14}$$

식 (14)에서 width와 height은 영상의 너비와 높이를 나타내고, F_r 은 초당 부호화된 프레임 율이다. 초기 가상 버퍼의 점유량은 식 (2)와 같이 설정하고, 버퍼의 넘침과 고갈을 방지하기 위하여 버퍼의 상한($U(1)$) 및 하한($L(1)$)의 초기값을 식 (15)와 같이 설정한다.

$$\begin{aligned} L(1) &= \frac{u}{F_r} \\ U(1) &= u \times 2.048 \end{aligned} \tag{15}$$

2. 프레임에 대한 목표 비트량 계산

단계 2에서는 부호화 이후의 프레임에 대한 복잡도 갱신이 이루어진 정보들을 이용하여 현재 부호화 하고자 하는 프레임의 목표 비트량을 예측하는 단계로서, 부호화 하고자 하는 프레임에 대한 복잡도 가중치, 가상버퍼 점유율 그리고 목표 버퍼레벨을 고려하여 목표 비트량을 예측한다.

본 논문에서의 가상버퍼 점유량($B_c(n)$) 및 목표 버퍼레벨($TBL(n)$)은 각각 식 (16)과 식 (17)을 이용하여 구한다.

$$B_c(n) = A(n-1) - \frac{u}{F_r} + 0.5 \tag{16}$$

$$\begin{aligned}
 & \text{if}(codedPframe > 1) \\
 & \quad TBL(n) = TBL(n-1) - \frac{B_c}{N_p-1} \\
 & \text{else} \\
 & \quad TBL(n) = B_c - \frac{B_c}{N_p-1}
 \end{aligned} \tag{17}$$

식 (16)에서 $A(n-1)$ 는 $(n-1)$ 번째 프레임에 대한 부호화된 비트량이고, 식 (17)에서 $codedPframe$ 는 GOP 구간에서의 부호화된 P 프레임 개수이다. 식 (17)과 같이 목표 버퍼레벨은 I 프레임이 부호화 된 이후의 첫 번째 P 프레임의 경우는 버퍼 점유율에서 목표 버퍼레벨을 줄이고, 첫 번째 P 프레임이 부호화 된 이후에는 이전 목표 버퍼레벨에서 값을 줄인다. 현재 프레임을 부호화함에 있어서, 프레임에 할당할 수 있는 목표 비트량을 예측하기 위해서는 다음 식 (18)과 같이 계산된 버퍼 점유량 및 목표 버퍼레벨을 이용하여 버퍼량의 예측을 수행한다.

$$T_{buf} = \frac{u}{F_r} + \gamma \times (TBL(n) - B_c(n)), (\gamma = 0.8) \tag{18}$$

프레임 당 버퍼량이 예측되면 버퍼 상태를 고려하여 식 (8)과 같이 버퍼의 상한과 하한 범위까지 고려하여 최종적으로 버퍼량($T_{buf}(n)$) 값을 결정한다. n번째 프레임에 대한 최종 목표 비트량은 예측한 버퍼량과 식 (19)에서 예측한 부호 비트량을 고려하여 식 (20)과 같이 최종적으로 결정된다.

$$T_{r1} = \frac{W_p \times R}{N_p \times W_p} + 0.5 \tag{19}$$

$$T_{r2} = \max(0, \frac{u}{F_r} - \gamma \times (B_c(n) - TBL(n)) + 0.5)$$

$$T(n) = \beta \times (T_{r1} - T_{r2}) + T_{r2} + 0.5 \tag{20}$$

식 (19)에서 W_p 는 현재 P 프레임의 복잡도 가중치를 예측한 값으로서, 본 논문에서는 이전에 부호화된 프레임들의 발생 비트량 가중치와 QP값으로 계산하였다. R 은 현재 남아 있는 비트량이다. T_{r1} 과 T_{r2} 는 P 프레임에 대한 복잡

도 가중치를 이용한 프레임당 부호화 할 수 있는 비트량과 버퍼 점유량과 목표 버퍼레벨을 고려한 프레임당 부호화 할 수 있는 비트량이다. 식 (20)에서 β 는 버퍼 점유량과 목표 버퍼레벨에 대한 의존도를 의미하는 것으로 실험에 의해 0.5로 설정하였다.

3. 목표 비트량에 대한 QP값 결정

부호화 하고자 하는 프레임의 QP값을 결정하는 단계로서, 본 논문에서는 기존 H.264/AVC 방식의 식 (12)와 같이 QP변환 알고리즘을 사용하지 않고, 그림 2와 같이 다양한 영상의 실험에 의한 통계 정보를 이용하여 결정하도록 하였다. 그림 3은 각 영상의 특성에 따라 QP값에 따른 P 프레임의 발생 비트량 특성을 나타내고, 표 1은 각 프레임의 비트량을 예측하기 위하여 산출된 실험 데이터로서 각 영상의 P 프레임들의 평균 비트량을 나타낸 것이다.

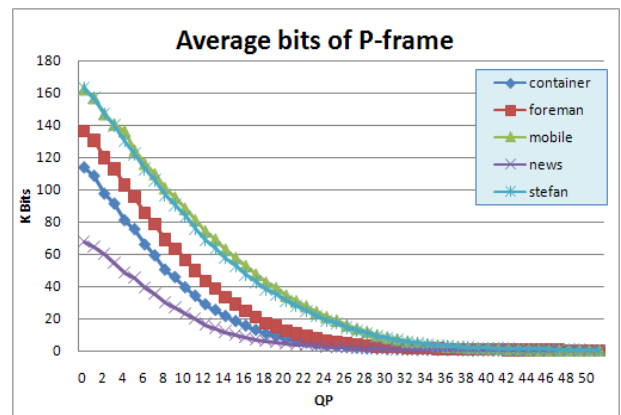


그림 3. 각 영상의 특성에 따른 발생 비트량
Fig. 3. The generated average bits of P-frames

다양한 유형의 영상 통계 정보를 이용하기 위하여 실험에 사용된 영상들은 빠르고 세밀한 영상의 "Mobile"와 "Stefan", 그리고 느리고 평탄한 영상의 "News"와 "Container"가 사용되었으며, 보통의 빠르기와 평탄함 및 세밀함이 골고루 있는 영상 "Foreman"이 사용되었다.

실험(표 1)을 통하여 P 프레임들의 평균 비트량이 QP값과 밀접한 관계를 가지고 있음을 확인할 수 있으며, 그림

표 1. 각 QP에 따른 P 프레임의 평균 비트량

Table 1. The average bits of P-frames by each QP value

QP	container	foreman	mobile	news	stefan	QR
...						
17	12,849	20,692	47,388	7,136	42,849	24,559
18	10,621	17,164	42,426	6,103	38,352	21,819
19	9,091	14,943	38,998	5,426	35,062	19,301
20	7,212	12,277	34,350	4,645	31,002	16,823
21	5,285	10,522	30,895	4,108	27,926	14,893
22	5,007	9,018	27,715	3,636	24,809	13,095
23	4,006	7,575	24,176	3,153	21,849	11,313
24	3,224	6,358	21,084	2,749	18,947	9,905
25	2,659	5,538	18,972	2,458	17,057	8,593
26	2,093	4,572	16,009	2,100	14,467	7,317
27	1,704	3,921	13,845	1,856	12,594	6,303
...						

3과 같이 QP값에 따라 할당되는 비트량의 크기가 선형적으로 변하는 것을 보여주고 있다. QP값과 발생 비트량과의 선형 관계는 다음과 같이 유도될 수 있다.

$$QP_{bits,n} = \alpha \times e^{(\beta \times (QP_{n-1} + 1))}, \quad (1 \leq n \leq 51) \quad (21)$$

식 (21)의 n 는 QP값 인덱스이며, $QP_{bits,n}$ 는 QP값 인덱스 n 에 대한 예측된 발생 비트량을 의미한다. α 와 β 는 그림 2와 표 1의 다양한 영상의 통계 특성에 의해 결정된 상수 값이다. 표 1에서의 QR(QP Range)은 현재 부호화하고자 하는 프레임에 할당 가능한 비트량 구간을 나타낸 것으로, 양자화 인덱스 함수로서의 비트량이 가우시안 분포를 가질 수 있기 때문에 식 (21)의 QP 인덱스에 대한 예측 가능한 발생 비트량 $QP_{bits,n}$ 는 식 (22)와 같이 재계산 될 수 있다.

$$QP_{bits,n} = QP_{bits,(n+1)} + \frac{|QP_{bits,n} - QP_{bits,(n+1)}|}{2}, \quad (22)$$

$$(0 \leq n < 51)$$

이렇게 QP 인덱스에 대한 예측된 발생 비트량 $QP_{bits,(0 \sim 51)}$ 은 초기화 단계에서 한 번만 계산되며 1차원 테이블에 값을 저장하여 사용한다. 이후에는 식 (20)에서 계산된 목표 비트량을 이용하여 식 (23)과 같이 현재 프레임에 대한 QP_c

값을 결정한다.

$$QP_c = \text{choice } QP(QP_{bits,n(1 \sim 51)}, T), \quad (1 \leq n \leq 51) \quad (23)$$

식 (23)에서 $\text{choice } QP()$ 는 QP 인덱스에 의한 예측 비트량 $QP_{bits,n(1 \sim 51)}$ 구간에서 해당 목표 비트량 T 를 찾는 함수로써, 반복 루프를 통하여 구해진다. 그리고 P 프레임의 QP_c 값이 결정될 때는 프레임들 간의 화질 열화를 줄이기 위하여 식 (24)와 같이 ΔQP 값으로 제한한다.

$$QP_c = \min(QP_p + \Delta QP, \max(QP_p - \Delta QP, QP_c)) \quad (24)$$

식 (24)에서 QP_p 는 이전 프레임에서의 QP값을 나타내며 프레임간의 QP값을 제한하는 ΔQP 값은 ± 2 로 설정하였다.

4. 프레임에 대한 복잡도 갱신

단계 4에서는 이전 프레임을 부호화 한 후에 발생한 실제 비트량을 이용하여 현재 프레임의 복잡도를 계산하는 단계로서, 가상 버퍼의 점유량 및 버퍼의 넘침과 고갈을 예방하기 위한 과정 그리고 이전에 부호화된 P 프레임들의 발생된 비트량과 현재 프레임의 발생 비트량을 이용하여 이후에 부호화시 사용될 P 프레임의 복잡도 가중치를 계산한다.

가상 버퍼의 점유량 B_c 및 버퍼의 넘침과 고갈을 예방하기 위한 버퍼의 상한 $U(n)$ 및 하한 $L(n)$ 값은 다음 식 (25)~(27)을 통해서 구해진다.

$$B_c = B_c + (A(n-1) - \frac{u}{F_r} + 0.5) \quad (25)$$

$$U(n) = U(n-1) - (A(n-1) - \frac{u}{F_r} + 0.5) \quad (26)$$

$$L(n) = L(n-1) - (A(n-1) - \frac{u}{F_r} + 0.5) \quad (27)$$

다음에 부호화 할 P 프레임의 목표 비트량을 예측하기 위한 P 프레임의 복잡도 가중치 W_p 는 식 (28)과 같이 구해진다.

$$W_P = (\lambda \times A(n-1) + (1-\lambda) \times S_{bits}) \times QP_p + 0.5 \quad (28)$$

본 논문에서는 P 프레임에 대한 복잡도 가중치를 예측하기 위한 부호화된 P 프레임들의 최대 개수를 20개로 제한하였으며, 식 (28)에서 S_{bits} 는 이전 P 프레임의 발생 비트량 $A(n-2) \sim A(n-20)$ 구간에서 QP_p 와 동일한 QP 값을 갖는 P 프레임들의 평균 비트량이다. 이는 프레임에 할당할 수 있는 비트량을 통하여 QP 값을 결정하기 위해서는 각 QP 값에 따라 발생하는 프레임 비트량을 알아야 하기 때문이다. 또한 부호화 하고자하는 P 프레임은 인접한 프레임들과의 밀접한 상관관계를 가지고 있다. 따라서 목표 비트량과 실제 발생된 비트량과의 오차를 줄이기 위한 $A(n-1)$ 와 S_{bits} 의 가중치 λ 값을 0.67로 설정하였다. 이 λ 값은 각각의 고정 QP 값과 움직임 추정을 위한 참조 프레임을 20으로 설정하고 시뮬레이션하여 얻어낸 결과 값이다.

IV. 실험 및 결과

본 논문에서는 H.264/AVC 부호기를 기반으로 한 효율적인 비트율 제어 방법을 제시하였다. 제안한 알고리즘의 성능을 검증하기 위하여 H.264/AVC 참조 소프트웨어 12.1 부호기를 사용하였다^[13]. 실험 환경은 표 2와 같고 실험에 사용된 영상은 부호기에서의 성능평가 테스트로 광범위하게 사용되는 QCIF급 CONTAINER, NEWS, FOREMAN, COASTGUARD, PARIS, SIGN IRENE, TABLE 영상을

표 2. 실험환경
Table 2. Experimental environment

Profile	Baseline 3.0
MV Resolution	1/4 Pel
Hardamard Transform	ON
RDO	ON
Search Range	±16
Reference Frames	1
Symbol Mode	CAVLC
GOP size	30
Encoding Frame	300
Test Channel Bandwidth	32K ~ 512K

사용하였다.

영상의 GOP 구조는 초기 I 프레임과 그 이후의 모든 프레임들은 P 프레임으로 구성하였고 B 프레임을 사용하지 않았다. 그리고 모든 영상들은 프레임 스킵 없이 QP 값을 프레임별로 적용하여 실험하였다. 기존의 비트율 제어 알고리즘인 JM12.1과 참고 논문^[14] 그리고 본 논문에서 제안한 비트율 제어 알고리즘은 영상 화질의 객관적 평가를 위해 PSNR(Peak Signal to Noise Ratio)을 사용하였으며, $M \times N$ 크기의 8bits 영상에 대하여 원 영상 대비 복호화된 영상의 잡음비율은 다음과 같이 정의된다.

$$PSNR = 10 \log \frac{MN \times 255^2}{\|f - \hat{f}\|^2} \quad (29)$$

식 (29) 에서 $\|\cdot\|$ 는 유클리드 노름(Euclidean norm)을 나타내고, f 는 원 영상, \hat{f} 은 복호화 된 영상을 나타낸다. 그리고 각 알고리즘의 계산량 비교를 위하여 고밀도 측정 타이머를 사용하였으며, 계산량 측정 범위는 각 알고리즘의 데이터 및 버퍼의 초기화 부분을 제외한 비트율을 제어하는 부분에서만 측정하였다. 표 3에서 표 5는 채널 전송대역폭이 각각 32, 64, 128Kbps이고 프레임율이 30fps일 때 기존 비트율 제어 방법과 제안한 비트율 제어 알고리즘에 대한 부호화된 비트량, PSNR 및 계산량 비교이다.

실험 결과표에서 알 수 있듯이 참고 논문(Ref.14)이 JM 12.1보다는 평균적으로 0.07Kbps의 적은 비트량을 절약하고도 PSNR 측면에서는 평균적으로 0.02~0.03dB 정도 높게 나와 더 좋은 성능을 보이고 있다. 그러나 Ref14의 방법은 PSNR 기반의 비트율을 제어하기 때문에 이에 따른 계산량 증가로 JM 12.1의 방법보다는 약 13%의 계산량이 더 필요한 것으로 나타났다. 그리고 JM 12.1과 Ref14의 방법은 채널 전송환경에 부적합하게 평균적으로 0.1Kbps 비트량을 초과 발생하였다. 반면, 제안한 방법은 PSNR 측면에서 JM 12.1보다는 평균적으로 0.02dB 정도 낮았고, Ref.14보다는 평균적으로 0.04dB 정도 낮았지만, 채널 전송 환경에서의 발생 비트량은 JM 12.1 방법보다 0.5Kbps 정도 절약하였고, Ref.14 방법보다는 0.12Kbps 정도 절약하였다. 또한 제안한 방법이 전체적으로 채널 전송 환경에 적합하

표 3. 32 kbps 채널 전송 환경에서 각 영상들의 부호화된 비트량, PSNR 및 계산량 비교

Table 3. Performance comparisons of the proposed rate control scheme with the existing schemes at 32 Kbps

Sequence	JM 12.1				Ref.14				Proposed		
	Bit rate	PSNR	계산량 (μs)	Mad %	Bit rate	PSNR	계산량 (μs)	Mad %	Bit rate	PSNR	계산량 (μs)
container	31.99	34.74	1,148.783	14%	32.02	34.74	1,318.194	12%	31.67	34.54	1.277
news	32.09	31.94	1,137.806	14%	32.04	31.80	1,313.095	12%	32.01	31.68	1.289
foreman	32.08	28.39	1,138.822	13%	32.00	28.58	1,310.744	12%	31.95	28.52	1.306
coastguard	32.05	27.46	1,135.405	13%	32.06	27.44	1,309.780	12%	32.03	27.48	1.294
paris	32.18	26.57	1,157.744	14%	32.10	26.53	1,328.955	12%	32.01	26.56	1.283
signirene	32.01	31.23	1,120.029	13%	32.04	31.18	1,294.267	12%	32.00	31.15	1.314
table	32.03	29.27	1,145.141	14%	32.04	29.51	1,319.145	12%	31.94	29.21	1.311

표 4. 64 kbps 채널 전송 환경에서 각 영상들의 부호화된 비트량, PSNR 및 계산량 비교

Table 4. Performance comparisons of the proposed rate control scheme with the existing schemes at 64 Kbps

Sequence	JM 12.1				Ref.14				Proposed		
	Bit rate	PSNR	계산량 (μs)	Mad %	Bit rate	PSNR	계산량 (μs)	Mad %	Bit rate	PSNR	계산량 (μs)
container	64.03	37.42	1,175.506	14%	64.04	37.42	1,346.126	12%	63.94	37.27	1.317
news	64.04	35.82	1,168.595	14%	63.97	35.87	1,342.050	12%	63.89	35.51	1.304
foreman	64.03	32.15	1,183.141	14%	63.99	32.17	1,353.496	12%	63.79	32.10	1.302
coastguard	64.02	29.76	1,173.391	14%	64.00	29.76	1,346.029	12%	64.01	29.75	1.312
paris	64.34	29.05	1,184.514	14%	64.12	29.05	1,356.702	12%	64.06	29.03	1.277
signirene	64.04	33.85	1,148.641	14%	63.97	33.93	1,323.441	12%	64.03	33.74	1.311
table	64.03	33.11	1,185.203	14%	64.02	33.16	1,357.471	12%	63.94	33.33	1.314

표 5. 128 kbps 채널 전송 환경에서 각 영상들의 부호화된 비트량, PSNR 및 계산량 비교

Table 5. Performance comparisons of the proposed rate control scheme with the existing schemes at 128 Kbps

Sequence	JM 12.1				Ref.14				Proposed		
	Bit rate	PSNR	계산량 (μs)	Mad %	Bit rate	PSNR	계산량 (μs)	Mad %	Bit rate	PSNR	계산량 (μs)
container	128.04	40.24	1,201.869	14%	127.95	40.16	1,372.095	13%	127.20	40.19	1.305
news	128.09	40.15	1,193.926	14%	128.07	40.17	1,368.743	12%	127.66	40.08	1.314
foreman	128.01	35.75	1,225.928	14%	128.02	35.78	1,399.074	12%	127.83	35.64	1.313
coastguard	128.14	32.17	1,212.770	14%	127.96	32.21	1,387.609	12%	128.06	31.97	1.314
paris	128.21	33.65	1,204.631	14%	128.01	33.69	1,374.881	12%	128.13	34.00	1.296
signirene	128.15	37.65	1,184.677	14%	127.91	37.64	1,358.640	12%	128.06	37.55	1.272
table	128.00	36.56	1,215.251	14%	128.04	36.63	1,386.727	12%	127.71	36.65	1.307

계 비트량을 잘 조절하고 있음을 알 수 있다. 계산량에 있어서는 JM 12.1 방법과 Ref.14 방법 모두 복잡한 MAD 계산과 이에 따른 비트량 조절을 위한 연산량이 제안한 방법보다 월등히 많음을 알 수 있다. 특히 현재 프레임에 대한 복잡도를 예측하는데 사용되는 MAD 계산량(MAD %)이 전체 비트율 제어 계산량의 평균적으로 약 12% ~ 14% 정도 차지하고 있음을 실험결과 확인할 수 있었다. 실험 결과에서 알 수 있듯이 제안한 방법이 PSNR 측면에서 기존 JM 12.1 방법과 Ref.14 방법과는 평균적으로 거의 비슷하였지만, 계산량에서는 평균적으로 월등히 적은 0.1%만의 계산량으로도 효율적인 비트율을 제어할 수 있음을 확인하였다. 제안한 방법의 많은 계산량 절감은 기존 방법과는 달리 복

잡한 MAD 계산과 이에 따른 계산 과정을 따르지 않고, 기존 영상들의 통계정보를 이용하고 실시간적으로 발생하는 발생 비트량을 다시 통계정보로 활용함으로써 영상의 특성에 적응적으로 비트율 제어가 이루어졌기 때문이다.

그림 4와 그림 5는 TABLE 영상을 128Kbps 채널로 전송할 경우, 제안한 방법과 기존 방법들에 대하여 프레임별 부호화된 비트량과 PSNR 비교이다. 제안한 방법이 기존 방법들에 비해 좀 더 안정적으로 발생 비트량을 발생시키면서도 높은 PSNR 성능을 보여주고 있다.

그림 6, 7, 8은 각각 CONTAINER, NEWS, TABLE 영상을 다양한 채널로 전송할 경우, 제안한 방법과 기존 방법들에 대한 PSNR 성능 비교이다. 대부분의 전송 채널에 있어

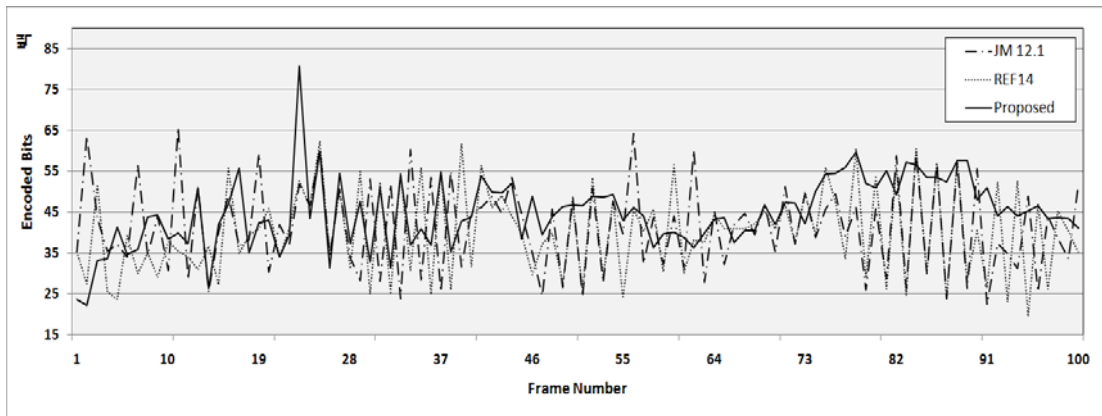


그림 4. TABLE 영상의 제안 및 기존 방법들의 프레임별 부호화된 비트량 비교(128 Kbps)
 Fig. 4. Performance comparison of the proposed rate control scheme with the encoded bits per frame(128 Kbps)

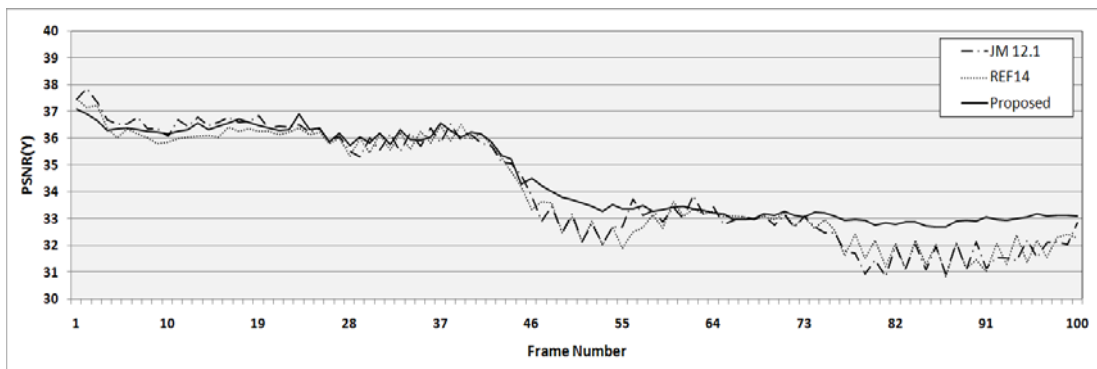


그림 5. TABLE 영상의 제안 및 기존 방법들의 프레임별 PSNR 비교(128Kbps)
 Fig. 5. Performance comparison of the proposed rate control scheme with the PSNR value per frame(128 Kbps)

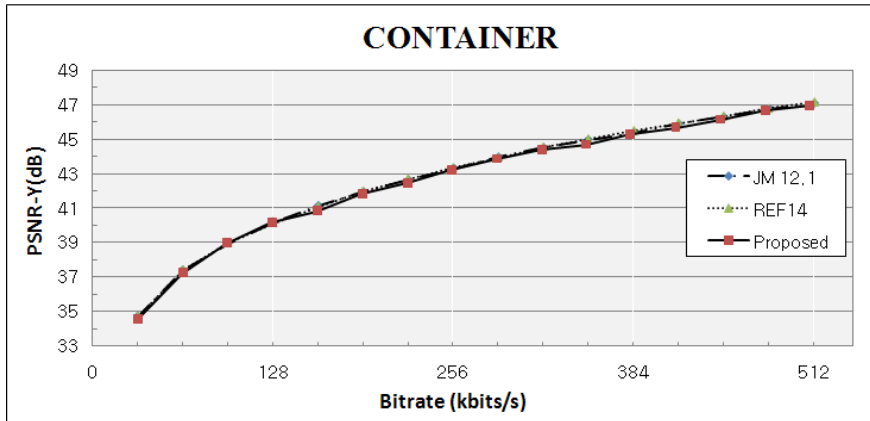


그림 6. CONTAINER 영상의 평균 PSNR Y(dB) (@30Hz)
 Fig. 6. Comparison of the PSNR performance (Container sequence, @30Hz)

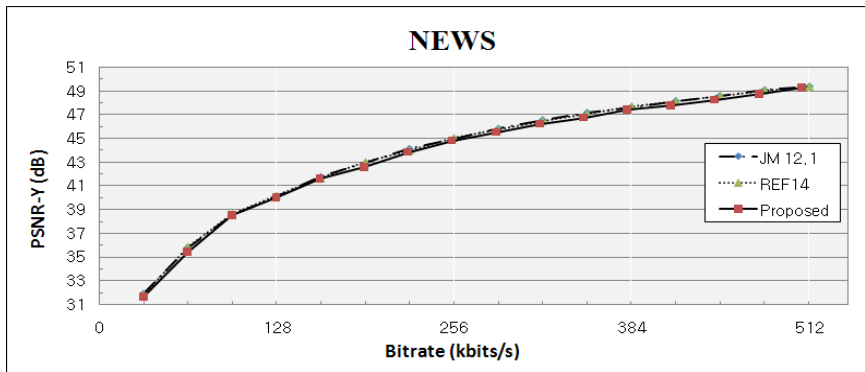


그림 7. NEWS 영상의 평균 PSNR Y(dB) (@30Hz)
 Fig. 7. Comparison of the PSNR performance (News sequence, @30Hz)

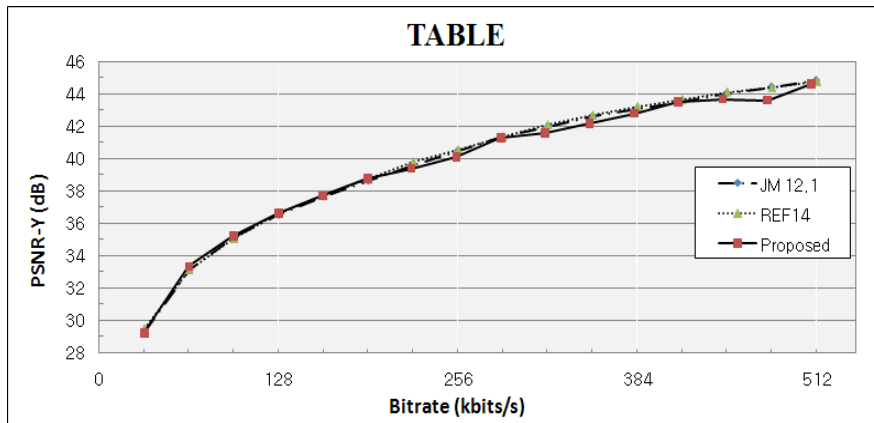


그림 8. TABLE 영상의 평균 PSNR Y(dB) (@30Hz)
 Fig. 8. Comparison of the PSNR performance (Table sequence, @30Hz)

서 기존 방법과 유사한 PSNR 성능을 보여주고 있음을 알 수 있다.

V. 결론

본 논문에서는 H.264/AVC 부호기에서 부호화된 비트스트림을 고정된 채널로 전송할 경우, 기존의 비트율 제어 방법을 소개하고 문제점을 제시하였다. 기존의 비트율 제어 방법의 큰 문제점은 현재 프레임의 부호화하는데 있어서 이전 프레임의 차분부호 정보량을 이용하기 때문에 많은 계산량을 필요로 한다는 것이다. 이러한 문제점들을 해결하기 위하여 본 논문에서는 이전 프레임의 차분부호 정보량이 아니라 다양한 영상의 통계정보와 이전에 발생된 비트량을 통계정보로 이용하여 영상 특성에 적응적으로 비트율을 제어하는 방법을 제안하였다.

본 논문에서 제안한 비트율 제어 알고리즘의 성능을 측정하기 위하여 기존의 비트율 제어 알고리즘과 비교 실험을 하였다. 실험 결과에서 알 수 있듯이 기존 비트율 제어 방법의 PSNR 성능과 거의 비슷하였고 다양한 채널 전송 환경에서도 기존 비트율 제어 방법보다 적합하게 효율적으로 부호화되는 것으로 나타났다. 특히 계산량 절감에 있어서는 제안한 방법이 기존 방법보다 99% 이상의 계산량이 감소하여 실시간 처리에 적합한 것으로 나타났다. 이는 영상과 관련된 응용 분야에 많은 활용이 있을 것으로 사료된다.

참 고 문 헌

- [1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264|ISO/IEC 14496-10 AVC) JVT-G050, May 2003.
- [2] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthar, "Overview of the H.264/AVC Video Coding Standard," IEEE Trans. Circuit and Systems for Video Technology, vol. 13, pp.560-576, July 2003.
- [3] Iain E.G. Richardson, H.264 and MPEG-4 Video Compression, Wiley, 2003.
- [4] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: tools, performance, and complexity," IEEE Circuits and Systems Magazine, vol. 4, pp. 7-28, 2004.
- [5] T. Wiegand, X. Zhang, and B. Girod, "Motion-compensating long-term memory prediction," Proc. IEEE Int. Conf. Image Processing, Oct. 1997.
- [6] T. Wiegand, X. Zhang, and B. Girod, "Block-based hybrid video coding using motion-compensated long-term memory prediction," Proc. Picture Coding Symp., Sep. 1997.
- [7] T. Wiegand, X. Zhang, and B. Girod, "Long-term memory motion-compensated prediction," IEEE Trans. Circuit and Systems for Video Technology, vol. 9, pp. 70-84, Feb. 1999.
- [8] T. Wiegand and B. Girod, Multi-Frame Motion-Compensated Prediction for Video Transmission. Norwell, MA: Kluwer, 2001.
- [9] Proposed Draft of Adaptive Rate Control, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISOWC JTCIISC29McGill and ITU-T SG16 4.6), May 2003
- [10] T. Chiang, Y.-Q. Zhang, "A new rate-control scheme using quadratic rate distortion model," IEEE Trans. Circuits Syst. Video Technoogy., vol. 7, no. 1, pp.246-250, Feb. 1997.
- [11] Feng Pan, Z. Li, K. Lim and G. Feng, "A Study of MPEG-4 Rate Control Scheme and Its Improvements," IEEE Trans. Circuits Syst. Video Technology, vol. 13, no. 5, pp. 440-446, May 2003.
- [12] Lee H J, Chiang T H, Zhang Y Q. "Scalable rate control for MPEG-4 video," IEEE Trans. Circuits System Video Technology, vol. 10, no. 6, pp. 878-894, Oct. 2000.
- [13] JM12.1, H.264/AVC reference software <http://iphome.hhi.de/suehring/tml/>
- [14] M. Jiang and N.Ling, "On Enhancing H.264/AVC Video Rate Control by PSNR-Based Frame Complexity Estimation," IEEE Trans. Consumer. Electronics. vol. 51, no. 1, pp. 281-286, Feb. 2005.
- [15] M. Jiang X. Yi, and N.Ling, "Improved frame-layer rate control for H.264 using MAD ratio," IEEE International Symposium on Circuits and Systems, vol. III, pp. 813-816, Vancouver, Canada, May 2004.
- [16] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini and G. Sullivan, "Rate-Constrained Coder Control and Comparison of Video Coding Standards," IEEE Trans. Circuit and Systems for Video Technology, vol. 13, No. 7, July 2003
- [17] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," IEEE Signal Processing Magazine, pp. 74-90, Nov. 1998.
- [18] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," IEEE Signal Processing Magazine, pp. 23-50, Nov. 1998.
- [19] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," IEEE Trans. Circuit and Systems for Video Technology, vol. 7, pp.246-250, Feb. 1997.

 저 자 소 개



김 명 진

- 2002년 2월 : 한국방송대학교 컴퓨터학과 졸업(학사)
- 2005년 2월 : 한국방송대학교 정보과학과 졸업(석사)
- 2005년 3월 ~ 현재 : 숭실대학교 정보통신공학과 박사과정
- 주관심분야 : 동영상 부호화 및 영상통신, Rate Control, 영상 복원 및 enhancement



김 경 환

- 2006년 2월 : 한성대 멀티미디어 정보처리과 졸업(학사)
- 2008년 2월 : 숭실대학교 정보통신공학과 졸업(석사)
- 2008년 2월 ~ 현재 : (주)셀런 연구원
- 주관심분야 : 동영상 부호화, Rate Control



홍 민 철

- 1988년 2월 : 연세대학교 전자공학과 졸업 (학사)
- 1990년 8월 : 연세대학교 전자공학과 졸업 (석사)
- 1990년 7월 ~ 1991년 8월 : LG정보통신 연구원
- 1997년 9월 : Northwestern University, 전기 및 컴퓨터공학과 졸업 (박사)
- 1997년 9월 ~ 1998년 8월 : Northwestern University, Research Fellow
- 1998년 9월 ~ 2000년 2월 : LG전자, 선임연구원
- 2000년 3월 ~ 현재 : 숭실대학교, 부교수
- 주관심분야 : 영상 복원 및 enhancement, 정지 및 동영상 필터링, 동영상 부호화, 비선형 필터링, Signal Blind Deconvolution