

논문 2006-43SP-1-4

선형 관계를 이용한 H.264/MPEG-4 AVC 비트율 제어 방법

(Rate Control based on linear relation for H.264/MPEG-4 AVC)

나 형 율*, 임 성 창*, 이 영 렬**

(Hyeong-Youl Na, Sung-Chang Lim, and Yung-Lyul Lee)

요 약

제한된 채널 대역폭이나, 저장 공간의 한계가 있는 경우 일정한 화질의 영상을 얻기 위해서는 영상의 움직임, 화면전환, 버퍼 용량, 채널 대역폭등의 변화에 순간, 순간 적용할 수 있는 비트율 제어 방법이 필요하다. 각 프레임마다 목표한 비트수를 얼마나 정확하게 예측할 수 있는지, 또는 버퍼에 의한 영상 지연 시간을 얼마나 짧게 할 수 있는지 등이, 효율적인 부호화기를 구성하는데 필요한 기술들이다. 따라서 본 논문에서는 비디오 부호화기에서 찾을 수 있는 여러 가지 선형 관계를 이용하여 위의 요구 조건을 만족하는 비트율 제어 방법을 제안한다. 제안된 방법에서는 3가지의 선형 관계에 대해서 설명을 한다. 첫 번째로, 비트수(R)와 양자화 된 변환 계수 중 Zero의 비율(ρ)과의 관계, 두 번째, PSNR과 양자화변수(QP) 사이의 관계, 그리고 마지막으로 QP와 ρ 에서의 선형적 특성을 찾을 수 있었다. 제안된 비트율 제어 방법과 H.264/MPEG-4 AVC JM9.3의 비트율 제어 방법을 비교 실험하여 본 결과, 제안된 방법에서 정확한 비트수 예측, 낮은 버퍼 충만도, 높은 PSNR을 관찰 할 수 있었다.

Abstract

The main purpose of rate control is to achieve the highest video quality when bandwidth or storage capacity is limited. For this purpose, we need a rate control algorithm which is adaptively controlled by the motion information of sequences, scene change, buffer capacity and time-varying bandwidth channels. A rate-control method in the encoder requires the accurate estimation of target bit for each frame and the low end-to-end delay for transmitting video data by intelligent selection of encoding parameters. In this paper, we suggest three kinds of linear relation in the encoder to satisfy the characteristics of rate control. The first relation is that between the percentage of zero quantized coefficients (ρ) and coded bits. Second relation is that between the PSNR of encoded frame and its Quantization parameter(QP). Finally, we can find out a linear approximation between QP and ρ . According to the experimental analysis, the proposed method results in an efficient rate control in terms of the bit estimation, the buffer capacity, and PSNR compared with the existing rate control in the H.264 JM 9.3.

Keywords : 비트율 제어, 선형 관계, PSNR, H.264, MPEG-4 AVC

I. 서 론

H.264/MPEG-4 AVC는 ISO/IEC와 ITU-T의 공동 작업에 의해 보다 향상된 비디오 부호화 표준을 제공하기 위해서 개발되었다^[1]. 다양한 예측 부호화 기법과 엔트로피 부호화 방법들을 통해, H.264/MPEG-4 AVC 비디오 부호화는 기존의 비디오 부호화 표준들보다 향상된 부호화 효율을 제공한다.

데이터를 전송할 경우, 전송 되는 채널은 특성상 매우 제한적이며, 고정된 대역폭을 가지고 있다. 따라서 비트율 제어의 목적은 채널의 시간적 다양성에 따라서, 부호화 변수들을 조절함으로써, 영상의 품질과 채널의 용량간의 균형을 조절하여 가장 좋은 부호화 성능을 제공하는 것이다.

이와 같은 부호화 성능을 조절하기 위한 방법으로 비트율 조절에서 많이 쓰이고 있는, Quadratic R-D 모델 [2-6]과 달리, 논문 [7-9]에서는 새로운 방법으로 비트율을 조절하는 효과적인 모델이 구현되었고, [8][9]에서는 이 모델이 여러 가지 비디오 부호화 표준에서 구현

* 학생회원, ** 정회원, 세종대학교 인터넷공학과
(Dep. of Internet Engineering, Sejong University)
접수일자: 2005년7월13일 수정완료일: 2005년11월11일

되는 예를 보여주고 있다. 이 모델에서는 낮은 비트 할당 오류와 부호화 오류를 보정하지 않는 특징 때문에, 다양한 대역폭의 조건에 적합함을 알 수 있다. 본 논문에서는 이 모델의 장점을 수용하고, 비디오 부호화에서 찾을 수 있는 선형 관계를 이용한 방법을 첨가하여, H.264/MPEG-4 AVC에 적합한 비트율 제어 방법을 구현한 방법을 설명하였다. H.264/MPEG-4 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.

다음 II 절에서는 영상의 비트수는 DCT(Discrete Cosine Transform) 영역에서 양자화 된 변환 계수 중 Zero의 비율([7]에서 ρ 라 지칭)에 선형 관계를 갖고 있다는 것을 기술한다. III 절에서는 본 논문에서 제안하는 향상된 비트율 제어 모델을 기술한다. 새로운 선형 관계인 ρ -QP(양자화변수) 모델을 제시하고, I 프레임의 QP를 결정하는 방법으로 PSNR-QP 선형 관계 모델을 제시한다. 이렇게 제안된 모델을 기반으로 하는 비트율 제어 방법을 IV절에 기술 한다. 이 절에서는 QP를, 기존의 방법에서처럼 DCT 계수 히스토그램이나 확률분포함수등을 고려하지 않고, ρ - E_p - E_{QP} -QP 관계에 의해서 직접적으로 선택 할 수 있음을 보인다.

H.264/MPEG-4 AVC의 Quadratic R-D 모델 기반의 비트율 조절 방법^[3]과 제안한 방법을 기반으로 한 비트율 조절 방법의 비교 테스트 결과를 V절에 기술하였다. 이 결과에서 우리는 제안된 방법이 H.264/MPEG-4 AVC에서 제공하는 비트율 제어 방법보다 화질적인 측면과 비트 제어의 관점에서 더 좋은 효율을 가지고 있음을 알 수 있다. 마지막으로, VI절에 본 논문에서 제안한 비트율 제어 방법에 대한 결론을 기술하였다.

II. ρ -도메인 기반의 비트율 조절

H.264/MPEG-4 AVC에서 사용되는 비트율 조절 방법^{[2][3]}은 비트와 왜곡이 양자화 스텝의 함수로서 표현되는 Quadratic R-D 모델에 기반을 두고 있다^[5]. 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.

비트율 조절의 기준으로 ρ -도메인을 사용하는 방법 [7]은 보다 정확한 이미지의 특성을 예측할 수 있는 새

로운 방법으로 제안되었으며, [8]에서 기존의 표준 부호화 방법(H.263, MPEG-4)에 적용될 수 있고, 그 결과가 매우 정확하며, 효율적인 예측을 할 수 있음을 보여 주었다.

ρ 는 양자화 된 변환 계수 중 Zero 값의 비율을 나타내며, 이 ρ 와 실제 생성되는 비트(R) 사이에는 선형 관계가 존재함을 이용하여, 목표 비트수를 결정한 후, 선형 관계식을 사용하여, 목표 비트수에 해당하는 ρ 값을 결정 할 수 있음을 알 수 있다. 이때 사용되는 선형관계식은 다음과 같은 수학적 R- ρ 모델로 표현된다^[7].

$$R(\rho) = \theta(1-\rho) \tag{1}$$

이때, θ 는 상수로서, 기울기를 나타낸다.

이 관계식의 예로, Foreman 영상에서 I 프레임 및 P 프레임의 관계 그래프를 그림 1에서 표현하였다. 본 논문에서는 계산의 편의를 위해서, ρ 값은 영상의 휘도 성분만 고려하였다. 그림 1에서 I 는 I (Intra) 프레임을 나타내고, P1,...,P9 는 각각 첫 번째 P (Predictive) 프레임부터 아홉 번째 P 프레임을 나타낸다.

그림 1에서 목표 비트수에 해당하는 ρ 를 구한 후, ρ 에 해당하는 QP를 구해야 한다. [7]에서는 ρ 와 QP 사이에 일대일 관계가 있음을 나타내고 있으며, 이에 해당하는 관계식은 다음과 같다.

$$\rho(\Delta) = \sum_{|a| < \Delta} p_x(a) \tag{2}$$

여기에서, Δ 은 각 프레임에 대한 양자화 스텝이고, $P_x(a)$ 는 변환 계수(Transformed Coefficient) x의 확률 분포이다. H.264/MPEG-4 AVC 표준에서는 Δ 값은 고정되어 있고, [0,51] 사이의 정수인 양자화 변수 (QP)값으

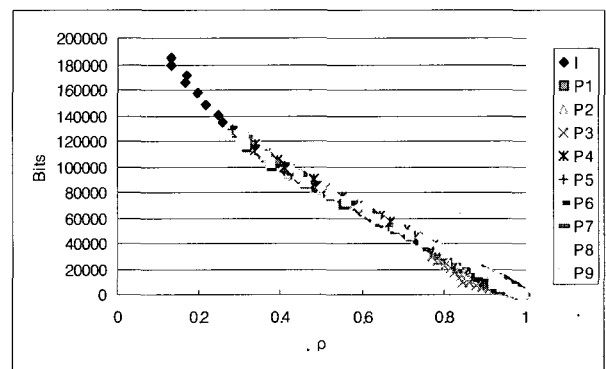


그림 1. ρ -비트 관계 그래프 : QP값을 0~51로 변화 시키면서 나오는 각 프레임의 ρ 와 비트를 표현함
Fig. 1. Plots of bit rate vs ρ , as QP varies between 0 and 51.

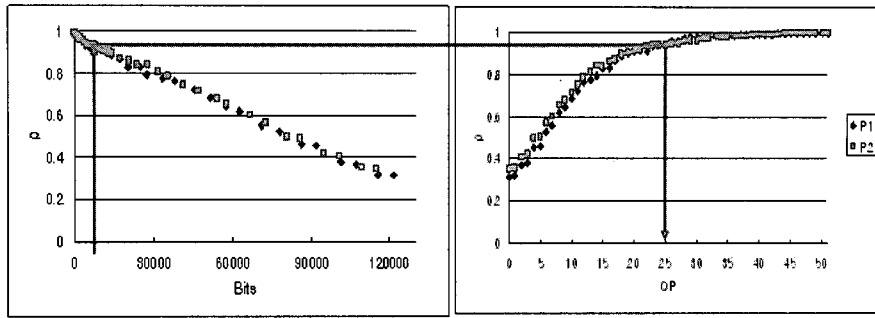


그림 2. R-p-QP 관계 그래프
Fig. 2. R-p-QP relation plots.

로 표현된다.

그림 2는 이 R-p-QP 사이의 상관관계를 개념적으로 표현한 그래프이다. 이 그래프를 보면, (1)식에 의해서 R를 통해 ρ 를 얻고, (2)식에 의해서 ρ 를 통해 QP를 결정함을 알 수 있다.

H.264/MPEG-4 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 공간예측 모드 (Intra16x16, Intra4x4), SKIP모드 등 모든 가능한 부호화 모드 중에서 가장 최적의 모드를 선택하는 방법이다. 이 방법은 효율화의 장점과 함께 연산량 증가에 의한 부호화 복잡도의 증가라는 단점이 있다^[10]. 또한 율-왜곡 최적화 방법은 비트율 조절의 측면에서도 매우 다루기 까다로운 방법이다. 율-왜곡 최적화를 위해서 모든 부호화 모드를 재귀적으로 검사를 하면서 최적의 모드를 선택하는데, 율-왜곡 최적화가 끝난 후에 QP의 선택을 위해 사용되는 여러 변수들이 결정되지만, 율-왜곡 최적화 도중에 이미 QP를 사용해야 하는, 닭이 먼저인지, 달걀이 먼저인지를 고민하는 것과 같은 딜레마가 생긴다^{[2][9]}.

이 딜레마를 풀기 위해서, [9][11]과 같은 논문에서는 이중 수행 처리 방법을 제안했고, [2]에서는 부호화된 이전 영상의 특성을 이용해서 현재 프레임의 특성을 예측하는 방법을 제안했다.

그림 1은 I 프레임과 각기 서로 다른 9개의 P 프레임의 비트수와 ρ 를 관찰한 결과이다. 이 결과에서 보면, 한 영상내의 각 프레임들의 특성 그래프는 거의 비슷한 특성을 갖고 있음을 알 수 있다. 따라서 본 논문에서는 이런 특성을 감안하여, 이전에 부호화된 프레임들의 통계적 특성을 현재 프레임 부호화에 이용하는 비트율 제어 방법을 제안했다.

III. 비디오 부호화내의 선형 관계

1. ρ -QP사이의 선형 관계

본 논문에서는 ρ 와 QP사이의 관계를 더욱 정확하게 나타내기 위해서, [7]에서 제안한 식(2)와 같은 계산식을 사용하지 않고, ρ 와 QP사이의 선형 관계를 유도하여 사용하였다.

그림 3은 ρ 와 QP사이에 지수함수를 사용하면, 선형적 특성을 찾을 수 있음을 나타낸다. 이때, 사용하는 지수함수는 식(4), 식(5)에 나타내었다. H.264/MPEG-4 AVC에서 사용하는 QP의 범위는 0~51의 범위이지만, 실제로 비디오 부호화에서 주로 사용하는 QP범위(대략 10~45) 안에서는, 그림 3에서 보이는 것처럼 거의 직선적인 관계가 있음을 알 수 있다. 이 부분에 대해서는 그림 4에 다시 한번 반투명 영역으로 표현하였다.

ρ 와 QP사이의 선형 모델은 선형 회귀 분석(Linear regression analysis)의 최소제곱법(Least Squares Method)에 의해서 구해 질 수 있는데, 다음과 같은 수학적식으로 표현가능하다.

$$E_{QP,i} = \beta_0 + \beta_1 E_{\rho,i} \tag{3}$$

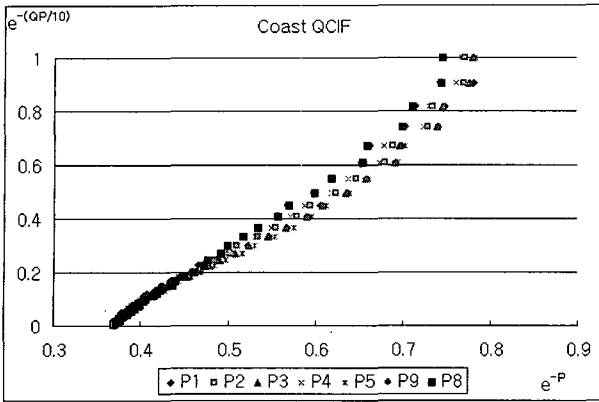
이때,

$$E_{\rho,i} = e^{-\rho,i} \tag{4}$$

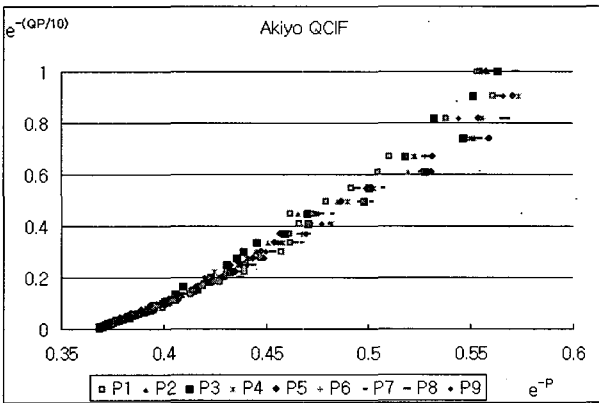
$$E_{QP,i} = e^{\left(\frac{QP,i}{10}\right)} \tag{5}$$

여기서 다음과 같이 최소제곱추정량(least squares estimator)인 β_0, β_1 을 구할 수 있다. β_0, β_1 을 구하기 위

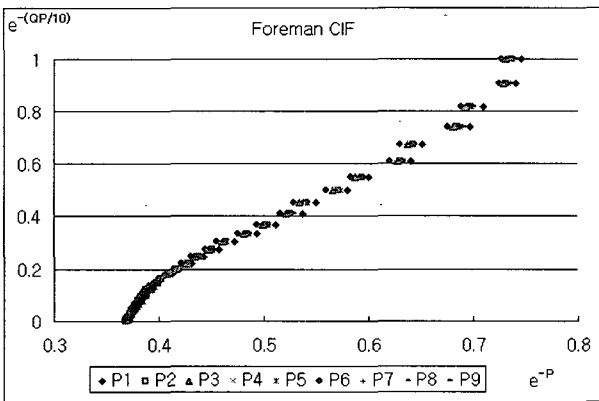
해서 왜곡값 $D = \sum_{i=1}^k (E_{QP,i} - (\beta_0 + \beta_1 E_{\rho,i}))^2$ 를 각



(a) Coast 영상 (QCIF)



(b) Akiyo 영상 (QCIF)



(c) Foreman 영상 (CIF)

그림 3. e^{-P} 와 $e^{-(QP/10)}$ 관계 그래프, 점선은 선형적 특성을 표현함

Fig. 3. e^{-P} vs. $e^{-(QP/10)}$ relation plot, The dotted line means the linear characteristic of the plots

각 β_0, β_1 에 대하여 편미분하여 0이 되도록 하는 답을 다음과 같이 구할 수 있다.

$$\beta_0 = \frac{\sum_{i=1}^k E_{QP,i} - \beta_1 \sum_{i=1}^k E_{\rho,i}}{k}$$

$$\beta_1 = \frac{k \sum_{i=1}^k E_{\rho,i} E_{QP,i} - \left(\sum_{i=1}^k E_{\rho,i} \right) \left(\sum_{i=1}^k E_{QP,i} \right)}{k \sum_{i=1}^k E_{\rho,i}^2 - \left(\sum_{i=1}^k E_{\rho,i} \right)^2} \quad (6)$$

k: 분석을 위해 값을 유지하고 있는 프레임 수 (=20)

그림 4에는 본 논문에서 제안하는 p-QP의 선형 관계를 이용한 비트율 조절 방법의 전체적인 개념을 나타내었다.

현재 프레임에 위한 목표 비트수(R)를 결정하면, 식 (1)에 의해서 ρ 값을 얻어 낼 수 있고(그림 4, (a)참고), 식 (4)에 의해서 E_p 를 얻어 낼 수 있다(그림 4, (b)참고). 그리고 식 (3)의 관계식에 의해 E_p 에 대응하는 E_{QP} 값을 얻어 낼 수 있고(그림 4, (c)참고), 최종적으로 E_{QP} 로부터 QP값을 결정하게 된다.

2. I 프레임의 QP 결정.

I 프레임에 할당되는 비트량은 전체 영상의 화질에 큰 영향을 미치게 된다^[4]. 따라서 I 프레임의 QP를 결정하는 것은, I 프레임을 뒤따르는 프레임들의 화질에 영향을 미치므로, I 프레임의 QP를 결정하는 효율적인 방법이 필요하다.

H.264/MPEG-4 AVC JM9.3[3]의 비트율 제어에서는 이전 GOP내에 포함된 P 프레임의 QP값의 통계적 정보를 사용하여 I 프레임의 QP를 결정하게 된다. 이 방법에서 이전 GOP내의 P 프레임들의 평균 QP값이 높을 경우, 현재 GOP의 I 프레임의 QP값도 높게 된다. 높은 QP값은 낮은 PSNR을 의미하므로, 이 I 프레임을 참조하는 현 GOP내의 P 프레임들은 연쇄적으로 낮은 PSNR을 갖게 된다. 결국 전체적으로 보면 I 프레임의 PSNR을 어떻게 결정하느냐에 따라서 전체 영상의 화질이 달라질 수 있다.

따라서 본 논문에서는 I 프레임의 QP를 결정하는데, 화질에 큰 영향을 미치는 요소인 PSNR에 따른 QP결정 방법을 제안한다.

한 영상에서 각 프레임에 대해서 QP값을 0에서 51까지 변화시키면서, 각 프레임의 PSNR을 계산하였을 때, QP와 PSNR과의 관계 그래프를 그려보면, 그림 5와 같

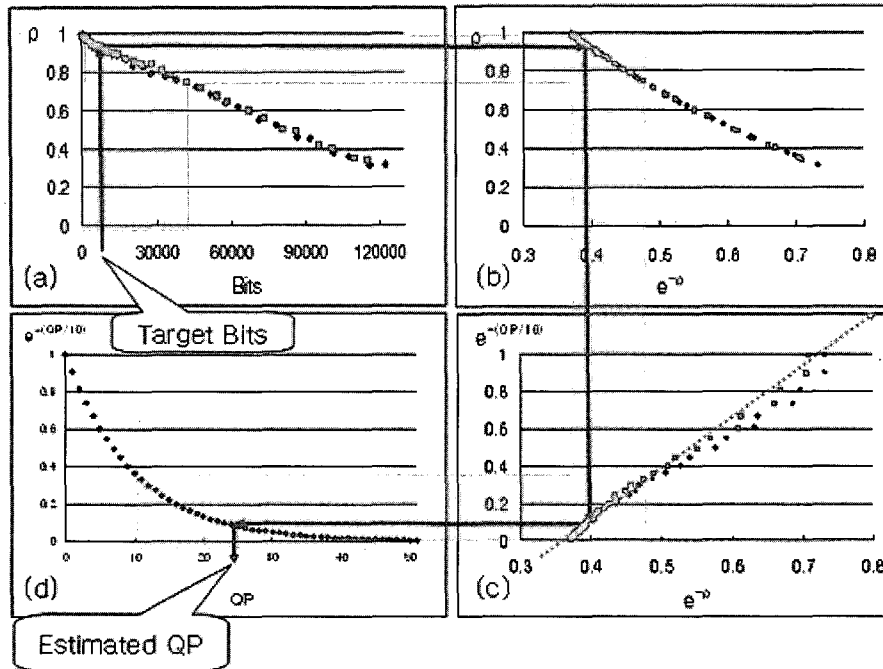


그림 4. R-p-Ep-EQP-QP 관계 그래프 :
 (a) 비트수- ρ 그래프 (b) ρ -Ep 그래프 (c) Ep-EQP 그래프 (d) EQP-QP 그래프
 반투명 영역 : 비디오 부호화에서 주로 많이 사용하는 QP 영역 (10~45)
 Fig. 4. R-p-Ep-EQP-QP relation plot :
 (a) Bit vs ρ plot (b) ρ vs Ep plot (c) Ep vs EQP plot (d) EQP vs QP plot
 Translucent area : QP range which is the most popularly used for video coding (10~45).

이 QP와 PSNR사이에는 선형 관계가 존재함을 알 수 있다. 그리고 그 선형 관계는 프레임들 사이에 거의 비슷한 관계를 가지고 있음을 역시 알 수 있다.

따라서 전체 영상의 화질을 유지하기 위해서 필요한 I 프레임의 PSNR을 결정하면, 그림 5에서 나타나는 선형 관계에 의해서 QP를 결정할 수 있다. 이때 현재 I 프레임의 PSNR은 이전 GOP내에 포함된 P 프레임의 PSNR의 정보를 이용하여 다음과 같이 결정한다.

$$PSNR_{0,n}^I = \left(\frac{1}{k} \sum_{i=1}^k PSNR_{i,n-1}^P \right) + \delta \quad (7)$$

식 (7)에서 $PSNR_{0,n}^I$ 은 n번째 GOP내의 I 프레임의 PSNR을 나타내고, $PSNR_{i,n-1}^P$ 은 (n-1)번째 GOP의 i 번째 P 프레임의 PSNR을 나타내고, k는 한 GOP내의 P 프레임수이고, δ 는 실험에 의해서 결정된 상수 (= 3)이다. 이때 δ 값은 P프레임들에 의해서 참조되는 I 프레임의 PSNR을 이전 GOP내의 P프레임의 평균값 보다

높여 줌으로써 현재 GOP내의 P프레임들의 화질을 높여주는 역할을 해주며, δ 값에 0,1,2,3,4... 등의 여러 가지 값들을 넣어서, 여러 동영상에서 공통적으로 높은 PSNR이 나오는 값을 선택하였다.

PSNR과 QP의 선형 관계식은 선형회귀분석에 의해서 다음과 같이 정의 될 수 있다.

$$PSNR_{0,n}^I = \alpha_0 + \alpha_1 QP_{0,n}^I \quad (8)$$

이때, α_0, α_1 은 다음과 같이 계산된다.

$$\alpha_0 = \frac{\sum_{i=1}^k PSNR_{i,n-1}^P - \alpha_1 \sum_{i=1}^k QP_{i,n-1}^P}{k}$$

$$\alpha_1 = \frac{k \sum_{i=1}^k QP_{i,n-1}^P - PSNR_{i,n-1}^P}{k \sum_{i=1}^k QP_{i,n-1}^P QP_{i,n-1}^P - \left(\sum_{i=1}^k QP_{i,n-1}^P \right)^2} \quad (9)$$

$$\frac{\left(\sum_{i=1}^k QP_{i,n-1}^P \right) \left(\sum_{i=1}^k PSNR_{i,n-1}^P \right)}{k \sum_{i=1}^k QP_{i,n-1}^P QP_{i,n-1}^P - \left(\sum_{i=1}^k QP_{i,n-1}^P \right)^2}$$

k : 분석을 위해 값을 유지하고 있는 프레임수 (= 20)

제안 방법의 실효성을 검증하기 위해서, I 프레임의 QP결정법을 H.264/MPEG-4 AVC JM9.3의 비트율 제어에서 사용한 방법과, 본 논문에서 제안한 방법을 나머지 조건을 전부 다 같이 설정한 상황에서 비교해 보면, 표 1과 같이 전체적으로 본 논문에서 제안한 방법에서 높은 화질 (PSNR)을 얻을 수 있음을 알 수 있다.

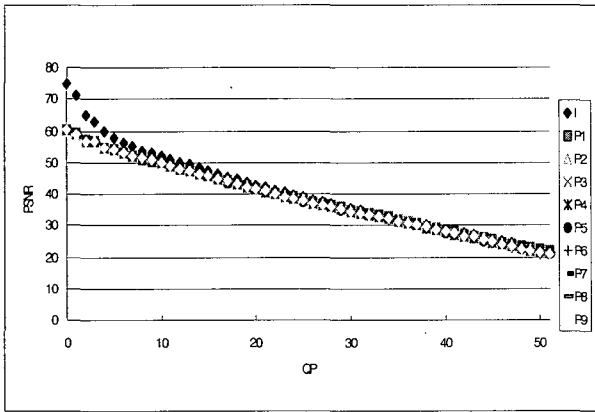


그림 5. PSNR와 QP사이의 관계 그래프, I 프레임 1개와 9개의 P 프레임에서 QP=0~51까지의 결과를 나타냄

Fig. 5. PSNR vs QP relation plot. QP varies from 0 to 51 for one I-picture and nine P-pictures.

표 1. I 프레임의 QP 결정법의 비교 실험
Table 1. Comparative test of QP estimation for I Frame

영상	종류	JM9.3	제안방법	Gain
Foreman	Qcif(128K)	35.37	35.43	+0.06
Carphone	Qcif(64K)	33.26	33.34	+0.08
paris	Cif(256K)	31.94	32.35	+0.41
news	Cif(128K)	34.99	35.18	+0.19
irene	Cif(256K)	33.69	33.77	+0.08

IV. H.264/MPEG-4 AVC를 위해 제안된 비트율 제어 방법

제안된 비트율 제어 방법은 다음과 같이 정리할 수 있다.

제 1단계 : 현재 프레임을 위한 목표 비트수 계산

목표 비트수를 계산하는 방법은 [3]에서 사용하는 방법과 같지만, 정확한 추정을 위해서, 상수 BETAP의 값을 실험에 의해서 다음과 같이 수정하였다.

BETAP : 0.25 (B 프레임이 없는 경우)

0.9 (B 프레임이 있는 경우)

H.264/MPEG-4 AVC JM9.3에서는 BETAP값이 원래 B 프레임이 없는 경우는 0.5, B 프레임이 있는 경우는 0.9 였으나, B 프레임이 없는 경우, BETAP값을 0.25로 수정함으로써 목표 비트수를 결정할 때 버퍼 상태에 대한 의존도를 더 높임으로써, 부호화를 해나가면서 생기는 비트율의 변화를 더욱 민감하게 조절할 수 있게 하였다.

제 2 단계 : 현재 상태에 따른 QP 갱신단계 선택 조건

- 현재 프레임이 P 프레임인 경우, 제 3 단계(QP 갱신 단계)는 다음의 조건에 의해서 선택되어 진다.

R_T 는 목표 비트수이고, B_C 는 현재 버퍼 충전 상태를 나타낸다. B_C 와 비교 기준으로, 실제 버퍼 상태가 아닌, 계산에 의해서 구해진 버퍼 충전 상태 B_I 는 다음과 같이 예측할 수 있다.

$$B_{I,n} = B_{I,n-1} + \frac{R_{GOP} - R_{Ipicture}}{N_{GOP} - 1} - \frac{R_{GOP}}{N_{GOP}} \quad (10)$$

이때,

$B_{I,n}$: n번째 프레임에서의 계산적 버퍼 충전 상태.

R_{GOP} : 현재 GOP에 할당된 총 비트수.

$R_{Ipicture}$: I 프레임에 할당된 비트수.

N_{GOP} : 현재 GOP 내의 프레임 수.

QP 갱신단계 선택 조건은 다음과 같다.

if($B_C < B_{I,n} * BLEVEL$)

 Estimated QP = Previous QP - DeltaQP

else if($R_T < 0$)

 Estimated QP = Previous QP + DeltaQP

else

제 3 단계(QP 갱신단계) 실행.

위 조건에서 BLEVEL 은 0.5 이고, DeltaQP 는 2 이다. 'Previous QP'는 이전 프레임의 QP값을 나타내고, 'Estimated QP'는 현재 프레임을 부호화하기 위해서 사용될 QP값을 나타낸다.

제 4 단계 : QP 갱신단계

1) I 프레임인 경우 :

- 첫 번째 GOP의 I 프레임의 경우, 부호화기에서 설정된 기본값을 사용하는데, 이 방법은 [3]에서 사용하는 방법과 같다.
- 위의 경우 이외의 I 프레임인 경우, 식 (7)에 의해서 선택된 목표 PSNR을 사용하여, 식 (8)에 의해서 갱신한다.

2) P 프레임인 경우 :

- 목표 비트수 R_T 를 식(1)에 사용하여, ρ 값을 선택한다.
- 식 (4)에 의해서 E_p 을 계산한다.
- 식 (3)에 의해서 E_p 에 대응하는 E_{QP} 를 결정한다.
- E_{QP} 값은 식 (5)에 의해서 구할 수 있다.
- QP는 $QP - E_{QP}$ 일대일 관계에서 의해서 구할 수 있다.

제 5 단계 : 부호화 후, 각종 변수 갱신 단계

- 식 (1)의 기울기 θ 는 다음의 식을 사용하여, 각 프레임을 부호화 할 마다 갱신 하여 준다. 이는 ρ 와 R 사이의 관계를 정확하게 구하기 위해서 필요한 작업이다.

$$\theta_{n+1} = \frac{1}{n} \sum_{k=1}^n \frac{R_k}{(1-\rho_k)} \quad (11)$$

이때,

- θ_{n+1} : 다음 프레임에서 사용할 기울기 θ .
- R_k : k번째 P프레임을 부호화 한 후 생성된 비트수
- ρ_k : k번째 P프레임의 ρ 값
- n 현재 GOP에서 부호화된 프레임의 수
- $E_{QP} - E_p$ 관계식 (식 (3))의 계수(β_0, β_1)을 식 (6)에 의해 갱신한다.
- PSNR-QP 관계식 (식 (8))의 계수(α_0, α_1)를 식 (9)에 의해 갱신한다.

V. 실험 결과

본 논문에서 제안한 알고리즘의 성능을 평가하기 위해서 H.264/MPEG-4 AVC JM9.3[3] 코드를 기반으로

베이스라인 프로파일(Baseline Profile)상에서 구현하였고, 성능 평가는 제안된 알고리즘으로 수정된 비트율 제어 루틴과, H.264/MPEG-4 AVC JM9.3에 기본적으로 구현되어 있는 비트율 제어 루틴을 비교하였다. GOP내의 첫번째 프레임은 I 프레임이고, 나머지 프레임들은 전부 P 프레임으로 설정하였다. I 프레임은 매 30 프레임마다 하나씩 나오도록 설정하였다. 실험환경으로, UVLC 모드를 선택하였고, ME 검색 범위(Search Range)를 16으로 설정했다. 참조 프레임수(Reference Frames)값은 5로 설정하고, 울-왜곡 최적화(RDO)값은 활성화 했다. 그 외 비트율 제어 방법의 실험에 영향을 미치지 않는 부호화 매개변수들은 공정한 성능 비교를 위해서 동일하게 설정하였다.

두 부호화기에서 초기 상태를 똑 같이 할 수 있도록, 'Initial QP' 값을 '35'로 설정하고, 'BasicUnit'은 QCIF영상에서는 '99', CIF영상에서는 '396'으로 설정하여 프레임 레이어 비트율 제어를 할 수 있도록 하였다.

실험에 쓰인 영상들은 QCIF와 CIF 크기의 영상을 사용했고, 또, 이 영상들을 서로 연결시켜서 생성된 영상들도 사용하였다. 이렇게 영상을 선택함으로써, 느린 움직임, 빠른 움직임, 그리고 화면 전환(Scene change)와 같은 특성들을 고려해서 실험을 할 수 있도록 하였다.

1. 상대적 비트율 제어 오류 비교

비트율 제어기의 성능을 측정하기 위해서, 다음과 같은 '상대적 비트율 제어 오류' (Relative rate control error)를 사용하였다.

$$E_{rc} = \frac{R - R_T}{R_T} \times 100\% \quad (12)$$

이때, R과 R_T 는 각 프레임의 실제 비트수와 목표 비트수를 의미한다.

표2은 각 영상들의 프레임별 '상대적 비트율 제어 오류'의 표준편차를 나타낸다. 표준편차는 비트율 예측이 얼마나 정확하게 행하여졌는지를 판단 할 수 있는 척도이다. 표2을 확인하여 보면, 제안된 비트율 제어기가 JM9.3의 비트율 제어기보다 보다 정확하게 비트수 예측을 하고 있음을 알 수 있다.

그림 6과 그림 7은 'Foreman'과 'News-Stephan' 영상에서의 '상대적 비트율 제어 오류'가 프레임 별로 어떻게 변화하는지를 나타내는데, 'Foreman'은 빠른 움직임에서의 특성을 볼 수 있고, 'News-Stephan'은 163번째 프레임에서 'News'영상과 'Stephan'영상을 연결한

표 2. '상대적 비트율 제어 오류'의 표준 편차
Table 2. Standard deviation of 'Relative rate control error'.

¹ News-Stephen, ² Akiyo-Coast-Container, ³ M&D-News-Silent-Stephen-Carphone, ⁴ Mobile-Paris

영상이름	종류	비트수	Frame수	표준편차	
				JM9.3	제안방법
M&D	qcif	24k	300	4016	88
foreman	qcif	48k	300	2657	67
NS ¹	qcif	64k	300	1081	304
ACC ²	qcif	100k	300	1460	663
LNSSC ³	qcif	64k	500	2113	703
tempete	cif	250k	240	579	26
mobile	cif	250k	300	4919	23
MP ⁴	cif	250k	300	4894	53

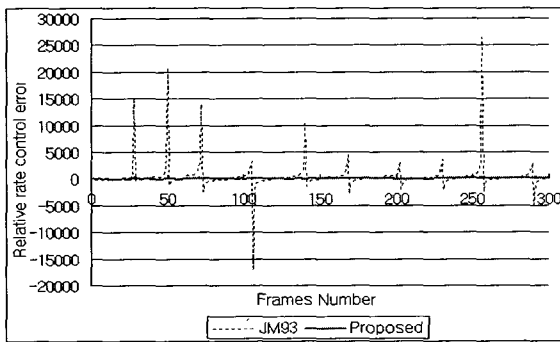


그림 6. 프레임별 상대적 비트율 제어 오류, Foreman 영상(목표비트수 : 48Kbps), 실선은 제안된 방법, 점선은 JM9.3

Fig. 6. Relative rate control error comparison for the sequence Foreman (Target bit : 48Kbps), solid line is for the proposed method, dotted line is for the JM9.3 method.

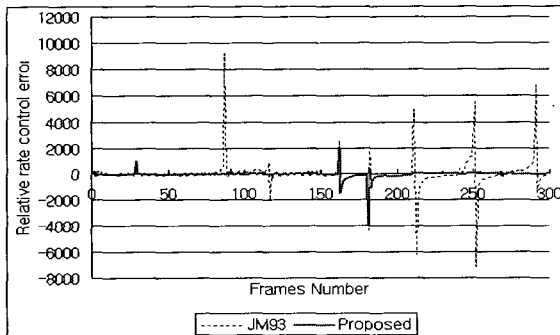


그림 7. 프레임별 상대적 비트율 제어 오류, News-Stephan 영상 (목표비트수 : 64Kbps), 실선은 제안된 방법, 점선은 JM9.3 방법

Fig. 7. Relative rate control error comparison for the sequence News-Stephan (Target bit : 64Kbps), solid line is for the proposed method, dotted line is for the JM9.3 method.

것으로 장면 전환(scene change)의 특성을 볼 수 있다. 이때 사용된 'News'영상은 느린 움직임 특성을 가지고 있으므로, 각 특성상에서의 '상대 비트율 제어 오류'를 확인 해 볼 수 있다.

표 2에서 빠른 움직임 특성을 가지고 'foreman'의 경우, JM9.3의 비트율 제어 오류 표준편차는 2657이고, 제안된 방법에서는 67의 상대적으로 매우 낮은 값을 갖는다. 줌아웃(Zoom out) 특성을 갖는 tempete의 경우, JM9.3은 579 이고, 제안된 방법은 26의 값을 갖는다. 패닝(Panning) 특성을 갖는 Mobile의 경우, JM9.3은 4919이고, 제안된 방법의 경우, 23의 매우 낮은 값을 갖는다. 그리고, 여러 영상들을 연결해서 장면 전환을 특성을 갖는 영상들에서도 제안된 방법이 JM9.3의 결과 값에 비해서 매우 낮은 결과를 나타낸다. 따라서, 제안된 방법은 여러 가지 영상 특성 하에서 비트수 예측을 JM9.3보다 매우 정확하게 할 수 있음을 알 수 있다.

2. 버퍼 충만 상태(Buffer Fullness) 비교

버퍼 충만 상태는 낮은 전송 지연 시간을 요하는 환경에서는 그 의미가 매우 중요하다. 예를 들어, 화상 전화나, 화상 회의와 같은 응용 분야에서는 영상을 부호화 하여, 복호화기에 전송되는데 걸리는 시간이 짧아야 실시간성을 제공할 수 있다. 따라서 버퍼에 누적되는 영상의 정보가 프레임을 부호화할 때의 상태에 따라서 크게 변화하지 않는 것이 중요하다.

그림 8과 그림 9는 각각 Foreman영상과, 여러 실험 영상을 매 100장마다 연결하여 생성한 영상의 버퍼 충만 상태 그래프를 나타낸다. 이 두 예제는 각각 빠른 움

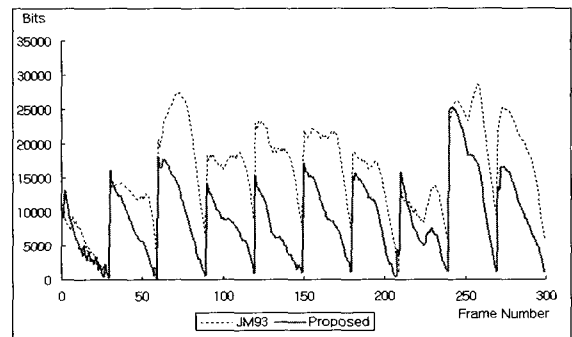


그림 8. Foreman영상에서의 버퍼 충만 상태 그래프 (목표 비트수 : 48Kbps), 실선은 제안된 방법, 점선은 JM9.3

Fig. 8. Buffer fullness comparison for the sequence Foreman (Target bit : 48Kbps), solid line is for the proposed method, dotted line is for the JM9.3 method.

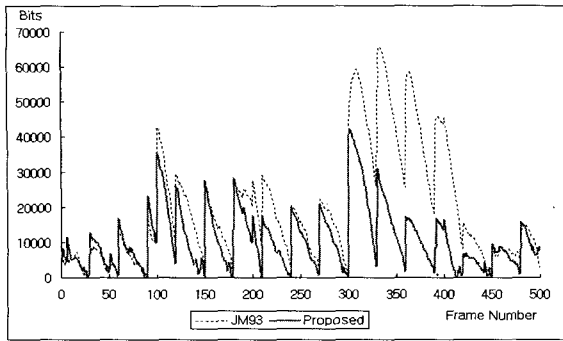


그림 9. M&D-News-Silent-Stephan-Carphone영상을 매 100프레임에 연결한 영상에서의 버퍼 총만 상태 그래프 (목표 비트수 : 48Kbps). 실선은 제안된 방법, 점선은 JM9.3

Fig. 9. Buffer fullness comparison for the sequence M&D-News-Silent-Stephan-Carphone which are concatenated at every 100th frame each other. (Target bit : 48Kbps), solid line is for the proposed method, dotted line is for the JM9.3 method.

직접 특성과 장면 전환 특성에서의 버퍼 총만 상태를 관찰 한 것이다.

이 결과를 보면, 제안된 비트율 제어 방법이 JM9.3의 비트율 제어 방법보다 더 낮은 버퍼 총만 상태를 유지함을 알 수 있다. 특히 그림 9에서 처럼, 장면 전환이 있는 경우는 비트율 제어 방법의 특성상 버퍼에 누적되는 비트의 값이 매우 크게 되는데, 제안된 방법의 경우, 버퍼의 누적 정도가 JM9.3에 의해 누적된 정도보다 상대적으로 낮음을 확인 할 수 있다.

3. PSNR 비교

지금까지, 제안된 방법이 JM9.3에 비해서 비트율 제어 면에서 향상된 기능을 제공할 수 있다. 성능을 비교하기 위해서는 추가로 실제 비디오 부호화의 가장 중요한 요소인 화질의 향상을 고려해야 한다. 표 3에는 여러 영상에 대한, 제안된 방법과 JM9.3의 평균 PSNR 과 그 표준편차를 나타내었다. 여기에서 우리는 JM9.3 보다 제안된 방법이 보다 향상된 화질을 보여줄 수 있음을 알 수 있다. 프레임에 따른 PSNR의 변화를 나타내기 위해서, 그림 10, 11에 단일 영상인 'Foreman', 'M&D'의 PSNR변화 그래프와 그림 12, 13에 연결 영상인 'Akiyo-Coast-Container'와 'Mobile-Paris'의 PSNR 변화 그래프를 나타내었다. 표 3에서 PSNR의 표준편차를 보면, 'Foreman'의 경우, JM9.3에서 2.76 이지만, 제안된 방법은 1.6이다. 'M&D'의 경우, JM9.3에서 2.55이지만, 제안된 방법은 2.16이다. PSNR의 표준편차가 작

다는 의미는 영상 화질의 변화가 크지 않고, 안정적인 화질을 나타냄을 의미한다. 따라서 표준 편차와 그림 10, 그림 11을 확인 해보면, 제안된 방법이 보다 JM9.3 보다 안정적이고, 향상된 화질을 보여줄 수 있다. 반면에 화면 전환 특성을 테스트하기 위해서 여러 영상을 연결한 'Akiyo-Coast-Container'와 'Mobile-Paris'의 경우, 그 특성상 PSNR의 표준편차가 JM9.3에 비해서 크게 감소하지는 않았지만, 그림 12, 그림 13을 보면, 장면 전환이 있는 후의 PSNR변화에서, JM9.3보다 제안된 방법이 더 안정적인 PSNR변화를 보여준다. 따라서 다양한 영상 특징에 대해서, 제안된 비트율 제어 방법이 JM9.3에 비해서 더 효과적인 화질 관리 능력을 보여줄 수 있다.

표 3. 평균 PSNR과 PSNR 표준편차

Table 3. Average PSNR and standard deviation of PSNR.

¹ News-Stephen, ² Akiyo-Coast-Container,

³ M&D-News-Silent-Stephan-Carphone, ⁴ Mobile-Paris

영상	종류	비트 수	frame 수	평균 PSNR		Gain	PSNR 표준편차	
				JM9.3	제안		JM9.3	제안
M&D	qcif	24k	300	31.55	31.94	0.39	2.55	2.16
foreman	qcif	48k	300	30.15	30.53	0.38	2.76	1.6
NS ¹	qcif	64k	300	30.01	30.21	0.2	5.69	5.56
ACC ²	qcif	100k	300	36.1	37.07	0.97	7.46	6.41
LNSSC ³	qcif	48k	500	31.95	32.28	0.33	4.83	4.32
tempeste	cif	250k	240	28.54	28.72	0.18	1.03	1.06
mobile	cif	250k	300	25.43	25.8	0.37	1.41	1.23
MP ⁴	cif	250k	300	27.44	27.89	0.45	2.81	3.03

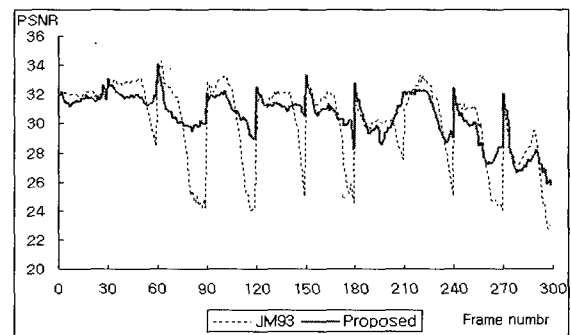


그림 10. PSNR 변화 그래프, Foreman 영상 (목표 비트수 : 48Kbps). 실선 : 제안된 방법, 점선 : JM9.3

Fig. 10. PSNR comparison plot for the sequence Foreman (Target bit : 48Kbps), solid line is for the proposed method, dotted line is for the JM9.3 method.

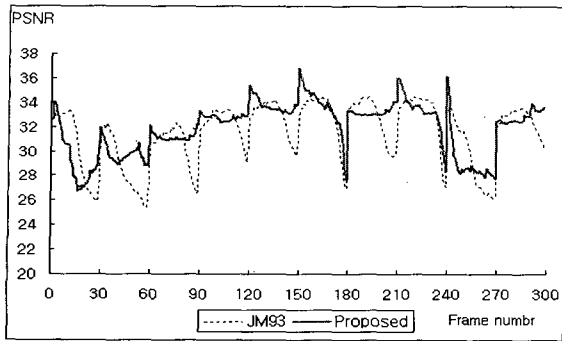


그림 11. PSNR 변화 그래프, M&D영상 (목표 비트수 : 24Kbps). 실선 : 제안된 방법, 점선 : JM9.3
 Fig. 11. PSNR comparison plot for the sequence M&D (Target bit : 24Kbps), solid line is for the proposed method, dotted line is for the JM9.3 method.

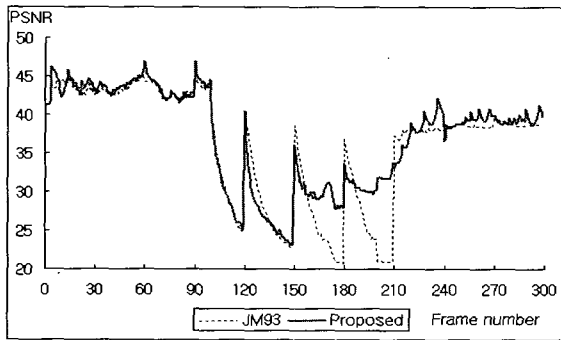


그림 12. PSNR 변화 그래프, Akiyo-Coast-Container의 영상을 매 100 프레임마다 연결한 영상 (목표 비트수 : 100Kbps). 실선 : 제안된 방법, 점선 : JM9.3
 Fig. 12. PSNR comparison plot for the sequence Akiyo-Coast-Container which are concatenated at every 100th frames each other (Target bit : 100Kbps), solid line is for the proposed method, dotted line is for the JM9.3 method.

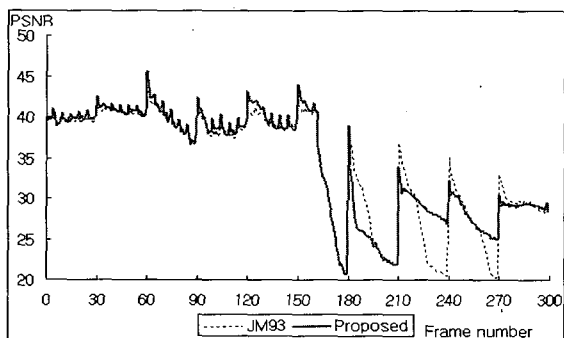


그림 13. PSNR 변화 그래프, Foreman과 Stephan영상을 163번째 프레임에서 연결한 영상 (목표 비트수 : 64Kbps). 실선 : 제안된 방법, 점선 : JM9.3
 Fig. 13. PSNR comparison plot for the sequence Foreman-Stephan which are concatenated at 163th frame (Target bit : 64Kbps), solid line is for the proposed method, dotted line is for the JM9.3 method.

VI. 결 론

본 논문에서는 비디오 부호화에서 구할 수 있는 여러 가지 선형 관계를 이용하여, 빠른 움직임이나, 장면 전환의 특성이 있는 영상에서 보다 향상된 효과를 나타내는 H.264/MPEG-4 AVC의 비트율 제어 방법에 대해서 제안하였다. I 프레임에 대한 QP값은 PSNR-QP의 선형 관계를 이용하여 결정할 수 있고, P 프레임의 경우, 지수 함수를 이용하여 ρ 와 QP의 선형 관계를 형성하고, 그곳에서 ρ 에 대한 QP값을 선택할 수 있다. 여러 가지 특성이 있는 영상을 대상으로 실험을 하여 본 결과, 제안된 방법이 JM9.3의 비트율 제어 방법보다, 비트수 예측이나 PSNR의 측면에서 보다 향상된 성능을 나타냄을 확인하였다. 특히, 제안된 방법의 상대적으로 낮은 버퍼 충전 상태는, 제안된 방법이 화상 전환이나 화상회의와 같은 실시간성이 필요한 영역에도 적합함을 알 수 있다.

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" in Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVTG050, 2003.
- [2] Proposed Draft of Adaptive Rate Control, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISOWC JTCIISC29MrGil and ITU-T SG16 4.6) 8~ Meeting: Geneva, May 20-26, 2003 Document: JVT-H017 Filename: JVT-H017.doc
- [3] JM9.3, H.264/AVC reference software <http://iphome.hhi.de/suehring/tml/>
- [4] 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 Technol., Vol. 13, pp.440-446, May. 2003.
- [5] T. Chiang, Y.-Q. Zhang, A new rate-control scheme using quadratic rate distortion model, IEEE Trans. Circuits Syst. Video Technol. 7 (February 1997) 246-250.
- [6] Lee H J, Chiang T H, Zhang Y Q. Scalable rate control for MPEG-4 video[J]. IEEE Trans Circuits Syst Video Technol, 2000, 10(6): 878-894.
- [7] Z. He, Y. K. Kim, Sanjit K. Mitra "Low-Delay Rate Control for DCT Video Coding via ρ -domain Source Modeling", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 11, No. 8, pag.928-940, Aug. 2001.

- [8] Z.He, S. K. Mitra "A Unified Rate-Distortion Analysis Framework for Transform coding", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 11, No. 12, Pag. 1221-1236, Dec 2001.
- [9] Z. He and T. Chen, "Linear rate control for JVT video coding,"International Conference on Information Technology: Research and Education, Newark NJ 2003.
- [10] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini and G. Sullivan, "Rate-Constrained Coder Control and Comparison of Video Coding Standards," IEEE Transactions on Circuits & Systems for Video Technology, 13, #7, July 2003
- [11] I.H. Shin, Y.L. Lee, and H.W. Park, "Rate control using linear rate- ρ model for H.264," Signal Processing: Image Communications, Vol. 19, pp. 341-352, April. 2004.

저 자 소 개



나 형 울 (학생회원)
 1997년 전남대학교 전산학과
 학사
 2000년 3월 Lotus Development
 Korea 연구원
 2000년 4월~현재 (주) 셀런
 책임 연구원

2003년 9월~현재 세종대학교 인터넷공학과
 석사 과정
 <주관심분야 : 영상 처리, 멀티미디어 시스템, 임
 베디드 시스템>



이 영 렬 (정회원)
 1985년 서강대학교 전자공학과
 학사
 1987년 서강대학교 전자공학과
 석사
 2000년 학국과학기술원 전기 및
 전자공학과 박사

2001년 8월 삼성전자 중앙연구소 DM Lab.
 수석 연구원
 2001년 9월~현재 세종대학교 인터넷공학과
 부교수
 <주관심분야 : 영상 처리, 멀티미디어 시스템,
 비디오 트랜스코딩, 3차원 비디오 코딩,
 RGB 코딩>



임 성 창 (학생회원)
 2006년 세종대학교 컴퓨터공학과
 학사 예정
 <주관심분야 : 영상 처리, 영상
 압축>