

논문 2007-44SD-6-10

저 비트율 영상에 대한 향상된 H.264 프레임 단위 데이터율 제어 알고리즘

(Enhanced Frame-Layer Rate Control for High-motion
Low-bit rate Video in H.264)

이 창 현*, 이 원 재*, 김 재 석**

(Changhyun Lee, Wonjae Lee, and Jaeseok Kim)

요 약

기존 H.264 프레임 단위 데이터율 제어는 버퍼 상태와 영상의 복잡도를 효율적으로 사용하지 못하여 움직임이 많은 영상에 있어 영상 품질의 변화가 심한 단점이 있다. 특히 인코딩시 설정된 비트율이 낮고 (19.2 kbps 이하) 프레임율이 높아 (30 fps), 영상 프레임당 사용할 수 있는 비트가 작아지면 영상 품질의 변화는 더욱 심해진다. 본 논문에서는 안정된 영상품질과 버퍼 상태를 얻기 위해 향상된 H.264 프레임 단위 데이터율 제어 알고리즘을 제안한다. 실험 결과, 제안한 알고리즘은 기존 알고리즘에 비해 향상된 결과를 보였다.

Abstract

In the existing frame-layer rate control for H.264, buffer status and content complexity are used improperly, causing quality fluctuations of high-motion video at low bit rates and high frame rates (under 19.2 kbps at 30 fps). We propose an enhanced H.264 frame-layer rate control scheme to obtain steady video quality and stable buffer management. Experimental results showed that the proposed scheme performed better than existing schemes.

Keywords : Video content complexity, H.264, rate control

I. 서 론

데이터율이 제한되어 있는 비디오 코딩에 있어서 데이터율 제어(rate control)는 중요하다. 좋은 데이터율 제어의 결과는 비디오 화질이 좋고, 비디오 화질의 변화폭이 낮으며, 버퍼의 상태가 안정적이고, 목표로 정한 비트율과 인코딩 된 비트율의 차이가 낮아야 한다. H.264/AVC는 현존하는 최고의 압축 효율을 보이는 비디오 코딩 표준이고^{[1],[6]-[7]} 기존의 데이터율 제어는

JVT(joint video team)를 기준으로 한다.^[2] 데이터율 제어는 다음의 3가지 과정을 거친 후 최종적으로 양자화 파라미터(quantization parameter)를 출력한다.^{[2]-[3]}

- 목표로 하는 비트수 예측
- 양자화 파라미터 계산
- 양자화 파라미터 조정

이러한 JVT의 데이터율 제어는 목표 비트수 예측시, 한 시퀀스(sequence) 내에서 콘텐츠 복잡도(content complexity)를 고려치 않기 때문에 움직임이 많은 영상에서 최적의 결과를 얻지 못하는 단점을 가진다. M. Jiang은 이러한 단점을 해결하고자 목표 비트수 예측시, 한 시퀀스 내에서 상대적으로 콘텐츠 복잡도(relative content complexity)를 적용하여 움직임이 많은 영상에서 비디오 화질의 변화정도를 감소시켜 성능을 향상시켰다.^[4] 상대적 콘텐츠 복잡도는 현재 프레임에서 예측된

* 학생회원, ** 정회원, 연세대학교 전기전자공학과
(Department of Electrical and Electronic
Engineering Yonsei University)

※ 본 연구는 정보통신부 및 정보통신연구진흥원 대학 IT연구센터 육성·지원사업 및 삼성전자 정보통신총괄에 의해 지원되었음.

접수일자: 2007년4월12일, 수정완료일: 2007년5월22일

평균 절대차(mean absolute difference; MAD)와 이전에 인코딩된 모든 P(forward predicted) 프레임에서 계산된 절대 평균차값들의 평균에 대한 비이다. 그러나 이러한 방법도 여전히 목표 비트율과 인코딩된 비트율의 차이가 크고 버퍼 관리가 불안정한 단점을 갖고 있다. 이는 2가지 이유에서 기인한다. 첫째, 상대적 콘텐츠 복잡도가 높을 때, 버퍼 상태를 고려치 않고 목표 비트수를 증가시키기 때문이고, 둘째, 상대적 콘텐츠 복잡도를 목표 비트수 예측에 사용하여 얻은 양자화 파라미터로 버퍼 점유율(buffer occupation ratio)을 안정하게 유지 못할 수도 있기 때문이다. 따라서 버퍼 상태와 상대적 콘텐츠 복잡도를 함께 고려해야 한다. 또한 이 두가지 변수를 목표 비트수를 예측할 때 보다는 양자화 파라미터를 조정할 때 사용하는게 더 합리적이다.

프레임당 주어진 비트수(허용 비트수)가 낮고, 예를 들어 낮은 목표 비트율과 높은 목표 프레임율(frame rate) 이고, 상대적 콘텐츠 복잡도가 연속적으로 높으면 목표 비트수가 영 이하인 음수로 떨어질 확률이 매우 높다. 이러한 경우에, 현재 프레임에 대한 양자화 파라미터는 이전 프레임의 양자화 파라미터 보다 2 만큼 증가된 값이 된다.^{[2]-[4]} 이는 비디오 화질을 열화 시킨다. 또한 목표 비트수가 자주 음수값을 가질수록 비디오 화질은 크게 출렁거릴 것이다. 따라서, 이러한 비디오 화질의 변화폭을 줄이기 위해 목표 비트수가 음수값을 갖지 못하도록 해야 한다.

앞서 살펴본 바에 의해, 본 논문에서는 허용 비트수가 낮은 응용분야와 움직임이 많은 영상에 대해, 목표 비트수를 최대한 양수값으로 유지시킬 목적으로 양자화 파라미터를 버퍼 상태와 상대적 콘텐츠 복잡도를 활용하여 조정한다.

II. 제안 방식

일반화를 저해하지 않는 범위내에서, GOP(group of picture) 구조는 IPPP,...,P 로 가정한다. 여기서 I는 인트라(intra) 코딩된 프레임을 뜻하고 P는 이전 프레임을 참조하여 예측된(forward-predicted) 프레임을 뜻한다. 데이터율 제어의 궁극적인 목적은 주어진 데이터율 내에서 최고의 성능을 낼수 있는 양자화 파라미터를 출력하는 것이다. 매 프레임에 대한 데이터율 제어는 3가지 과정으로 이루어져 있다. 즉, 프레임당 목표 비트수 예측, 양자화 파라미터 계산, 그리고 양자화 파라미터 조정이다.

프레임 목표 비트수를 예측하기 위해 현재 남아 있는 비트수를 파악해야 하고, 버퍼의 고갈과 넘침을 예방하기 위해 현재 버퍼 충만도(current buffer fullness)와 목표 버퍼 레벨(target buffer level)을 활용한다. 현재 프레임을 인코딩하기 전에 목표 비트수, $T_{b,i}$ 은 수식(1)과 같이 구할 수 있다.^{[2]-[3]} 수식(1)에서 $R_{b,i}$ 는 i 번째 프레임에서 인코딩 하지 않은 모든 P 프레임에 사용될, 남아 있는 비트수이고, $N_{Pr,i}$ 는 i 번째 프레임에서 인코딩하지 않은 P 프레임의 수를 나타낸다. 또한 b_r 은 목표 비트율을, f_r 은 목표 프레임율을, CBF_i 는 현재 버퍼 충만도를, TBL_i 는 목표 버퍼 레벨을 나타낸다. 마지막으로 b , G 는 상수로서 일반적으로 각각 0.5와 0.75 값을 갖는다.

$$T_{b,i} = \left[\beta \frac{R_{b,i}}{N_{Pr,i}} \right] + \left[(1-\beta) \left\{ \frac{b_r}{f_r} - \Gamma(CBF_{i-1} - TBL_i) \right\} \right] \quad (1)$$

1. 양수 목표 비트수의 경우

목표 비트수가 양수이면, 양자화 파라미터는 2차원율-왜곡 모델(quadratic rate-distortion model)^[5]에 의해 다음과 같이 계산 된다.

$$\frac{T_{tb,i}}{PMAD_i} = \frac{x_1}{Q_{c,i}} + \frac{x_2}{Q_{c,i}^2} \quad (2)$$

수식(2)에서 $T_{tb,i}$ 는 i 번째 프레임의 텍스처(texture) 비트수이다. 이는 목표 비트수와 이전에 인코딩된 헤더 비트수와의 차로 얻어진다. $PMAD_i$ 는 i 번째 프레임에서 예측된 평균 절대차를 나타내고, $Q_{c,i}$ 는 i 번째 프레임에서 계산된 양자화 파라미터이다. 또한 x_1 , x_2 는 각각 1차, 2차 계수값이다. 허용 비트수가 낮은 응용영역에서, 이전에 인코딩된 헤더 비트수가 현재 목표 비트수와 대등한 수준이 되면, 텍스처 비트수는 음수로 떨어질 수 있다. 이러한 경우에 텍스처 비트수는 1로 제한한다. 또한 콘텐츠 복잡도가 빠르게 변화하는 시퀀스에 있어서, 계산된 양자화 파라미터는 눈에 띄게 출렁거릴 수 있으므로 양자화 파라미터의 변화는 수식(3)처럼 프레임간 ± 2 단위로 제한한다.

$$Q_{lm,i} = \max \{ Q_{i-1} - 2, \min \{ Q_{i-1} + 2, Q_{c,i} \} \} \quad (3)$$

수식 (3)에서 Q_i 는 i 번째 프레임 인코딩 시 사용할 최종 양자화 파라미터이다. 수식(3)에 의해 제한된 양자화 파라미터를 버퍼 상태와 상대적 콘텐츠 복잡도를 활용하여 최종적으로 수식(4)와 같이 조정한다. CM_i 는 i 번째 프레

$$Q_i = \begin{cases} Q_{lm,i} - 1; [(Q_{i-1} - Q_{lm,i}) < 2] \text{ and } [CM_i > 1.09] \text{ and } [(CBF_{i-1} - TBL_i) < b_r / (f_r \times \Gamma)] \\ Q_{lm,i} + 1; [CM_i < 0.99] \text{ and } [(CBF_{i-1} - TBL_i) > b_r / (f_r \times \Gamma)] \end{cases} \quad (4)$$

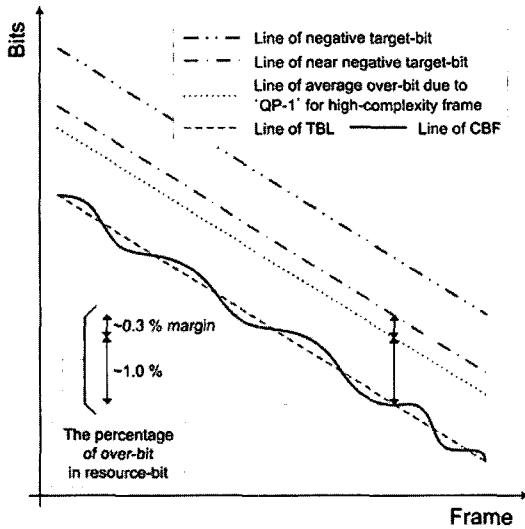


그림 1. 음수 목표 비트수 측면에서 양자화 파라미터 1 유닛 감소가 미치는 영향

Fig. 1. The effect of the number of over bits with respect to the negative number of target bits.

임의 상대적 콘텐츠 복잡도를 나타낸다. 기존 방식에서는 CM_i 로서 평균 절대차 비율(MAD ratio)을 사용했다.^[4] 본 논문에서도 상대적 콘텐츠 복잡도로 평균 절대차 비율을 사용했으며 수식(5)와 같다. 여기서 $CMAD_j$ 는 j 번째 프레임에서 계산된 평균 절대차를 나타낸다.

$$CM_i = \frac{PMAD_i}{\frac{1}{i-1} \sum_{j=1}^{i-1} CMAD_j} \quad (5)$$

수식(4)의 위부분에서, 목표 비트수를 양수로 유지할 만큼 버퍼 상태가 충분히 낮고, 상대적 콘텐츠 복잡도가 높으면, 프레임간 비디오 품질의 변화를 줄이기 위해 양자화 파라미터를 1 유닛 줄였다. 실험으로 양자화 파라미터의 1 유닛 감소에 따른 증가 하는 비트수를 알아보았다. 모든 프레임을 동일한 양자화 파라미터로 인코딩하고, 콘텐츠 복잡도가 높은 단 하나의 프레임에서 양자화 파라미터를 1 유닛 감소시켜 증가하는 비트수를 알아보았다. 전 프레임을 인코딩 하는데 필요한 비트수 (resource-bit) 대비 1 유닛 감소된 양자화 파라미터에 의해 증가한 비트수(over-bit)의 비율은 1% 미만이었다. 이 수치는 수식(1)의 오른쪽 부분을 음수로 만들 정도는 못된다. 그림 1은 음수 목표 비트수 측면에서 증가된 비트수의 영향을 나타낸 것이다. 그림1에서 'near negative target-bit' 라

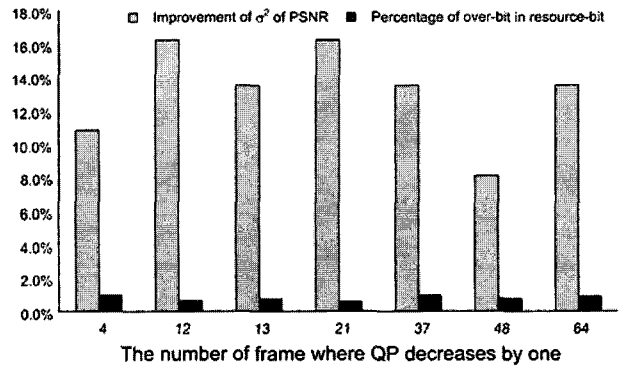


그림 2. 전체 사용된 비트수 대비 양자화 파라미터 1 유닛 감소에 따른 비트수 증가의 백분율 비교와 그에 따른 프레임간 비디오 품질 변화 감소율

Fig. 2. The percentage of over-bit in resource-bit and the improvement of fluctuation in video quality where the constant QP is 41 over 100 frames except for only one high-complexity frame that is quantized by 40.

인은 수식(1)의 두 번째 대괄호 부분이 음수임을 의미하고 'negative target-bit' 라인은 수식(1), 즉 목표 비트수가 음수임을 뜻한다. 그림1에서 보듯이 양자화 파라미터를 1 유닛 감소해도 목표 비트수가 양수를 유지할 만큼 여유가 있음을 알 수 있다. 그림2는 전 프레임을 인코딩 하는데 필요한 비트수 대비 복잡도가 높은 한 프레임의 양자화 파라미터를 1 유닛 감소시킴에 따른 비트 증가 비율을 나타낸다. 또한 양자화 파라미터의 1 유닛 감소에 따른 프레임간 비디오 품질 변화 감소율을 나타낸다. 그림2에서 1% 미만의 비트 증가율로 10% 이상의 프레임간 비디오 품질 변화 감소율을 얻을 수 있음을 알 수 있다.

실험조건을 간단히 언급하면, 시퀀스는 'foreman'을 사용했고, 프레임 개수는 100개이고, 양자화 파라미터는 콘텐츠 복잡도가 높은, 한 프레임을 빼고 모두 41을 사용했다. 나머지 한 프레임은 40을 사용했다. 수식(4)의 아래 부분에서, 버퍼 상태가 높고 상대적 콘텐츠 복잡도가 낮으면, 버퍼 충만도 레벨을 낮춰 다음 프레임의 목표 비트수가 음수화 되는 것을 예방하기 위해 양자화 파라미터를 1 유닛 증가시킨다.

2. 음수 목표 비트수의 경우

수식(1)에서 보듯이 음수 목표 비트수는 현재 버퍼 충만도가 목표 버퍼 레벨 보다 어느 정도 클 때 발생한

다. 일반적으로 현재 버퍼 충전도는 복잡도가 높은 영상에서 크고, 목표 버퍼 레벨은 낮은 허용 비트를 갖는 응용 분야에서 낮다. 이러한 조건에서 목표 비트수가 음수로 될 확률이 높다. 일단 목표 비트수가 음수이면, 현재 버퍼 충전도를 낮춰 가급적 다음 목표 비트수를 양수로 만들어야 한다. 이를 위해, 현재 양자화 파라미터를 수식(6)처럼 이전 양자화 파라미터 대비 증가시켜야 한다.

$$Q_i = \begin{cases} Q_{i-1} + 2; & CM_i > 1.09 \\ Q_{i-1} + 3; & otherwise \end{cases} \quad (6)$$

상대적 콘텐츠 복잡도가 낮으면, 양자화 파라미터는 이전 양자화 파라미터 대비, 더 크게 해서 버퍼 충전도 레벨이 더 빨리 감소되도록 했다. 이렇게 함으로서 설정 다음 프레임의 상대적 콘텐츠 복잡도가 높다 해도 그 다음번 프레임의 목표 비트수가 음수가 되지 않도록 버퍼의 공간을 잡아두게 된다. 수식 (4)와 수식(6)에 의해 얻어진 최종 양자화 파라미터로 현재 프레임을 인코딩 한 이후, 다음 프레임을 위해 선형 회귀 방법(linear regression method)^[5]을 사용하여 평균 절대차를 예측하는데 필요한 변수들과 2차원 윌-왜곡 모델에 필요한 두 가지 변수들(x1, x2)을 갱신한다. 마지막으로 제안한 데이터율 제어 알고리즘을 요약하면 그림3과 같다.

III. 실험 결과

실험은 움직임이 많은 'foreman'과 'carphone', 이 두 가지 시퀀스에서, 낮은 허용 비트 조건으로 이뤄졌다. 즉, 목표 비트율은 9.6 kbps 와 19.2 kbps 이고 목표 프레임율은 30 fps 이었다. 사용된 시퀀스는 QCIF 크기의 4:2:0 포맷을 가졌다. 테스트 플랫폼은 H.264 표준 소프트웨어 버전 JM6.1 이었고, 제안한 방식을 기존의 H.264 데이터율 제어 방식들과^{[3]-[4]} 비교했다. 모든 테스트는 다음과 같은 조건하에서 이루어 졌다.

- RDO(rate distortion optimization) 활성화
- 움직임 예측 조사 영역(search range for motion estimation)은 16
- 참조 프레임 수는 1개, 3개, 5개를 사용 (계산 복잡도와 성능 측면에서 고르게 실험해 보기 위함)
- 하다마드 변환 (Hadamard transform) 사용
- 엔트로피 코딩 방법은 CABAC (context-based adaptive binary arithmetic code)

실험에 필요한 다른 변수들은 3가지 방식이 모두 동일

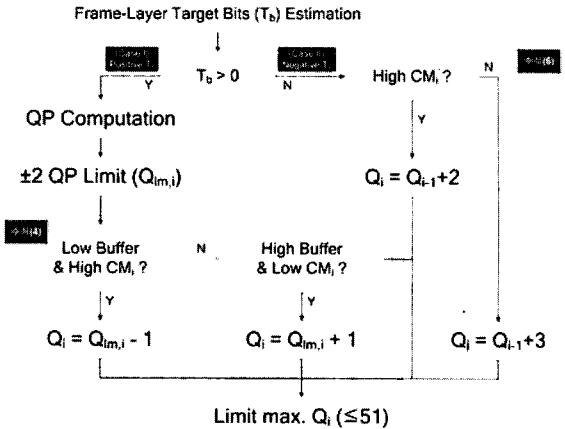
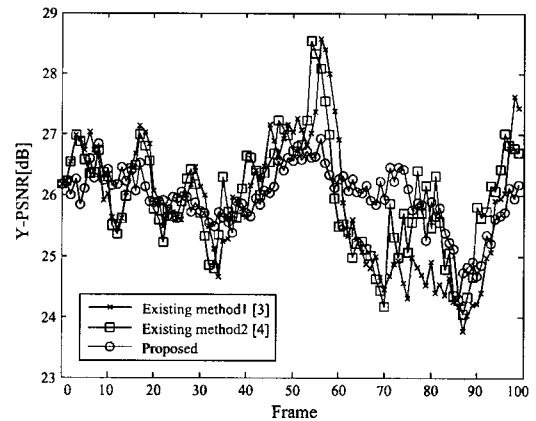
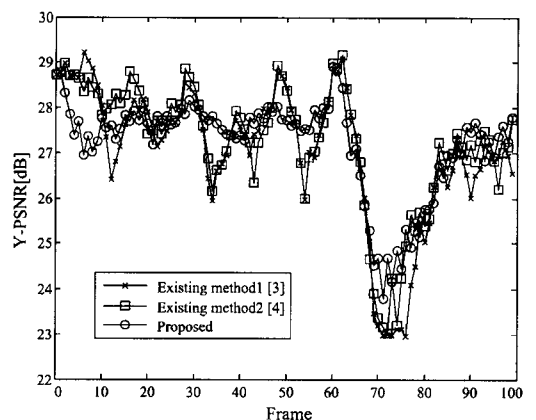


그림 3. 제안한 H.264 데이터율 제어 알고리즘의 흐름도
Fig. 3. The flow chart of the proposed H.264 rate control scheme.



(a)



(b)

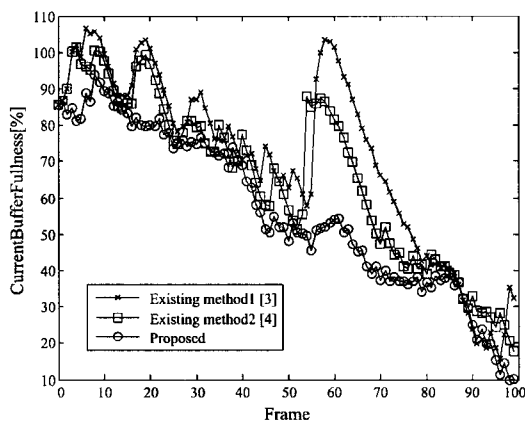
그림 4. 두 시퀀스에 대한 PSNR (a) 'carphone', 9.6 kbps/30 fps/1 참조 프레임 (b) 'foreman', 19.2 kbps/30 fps/1 참조 프레임

Fig. 4. PSNR curves for sequences (a) 'carphone' at 9.6 kbps/30 fps/1 reference frame, and (b) 'foreman' at 19.2 kbps/30 fps/1 reference frame.

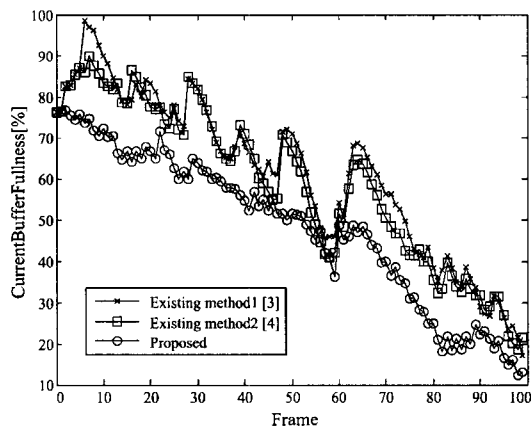
표 1. 기존 두가지 방식과 제안한 방식의 비교 (목표 프레임율은 30 fps)

Table 1. Performance comparisons of the proposed scheme with Reference 3 and 4 at 30 fps.

시퀀스	목표 비트율	참조 프레임수	평균 Y-PSNR (dB)			Y-PSNR의 표준 편차			인코딩 된 비트율 (kbps)		
			방식 ₁ ^[3]	방식 ₂ ^[4]	제안방식	방식 ₁ ^[3]	방식 ₂ ^[4]	제안방식	방식 ₁ ^[3]	방식 ₂ ^[4]	제안방식
carphone	9.6 kbps	1	25.85	25.98	25.95	1.05	0.87	0.49	9.89	9.68	9.57
		3	25.90	26.02	26.02	0.85	0.68	0.55	9.71	9.70	9.66
foreman	19.2 kbps	1	27.04	27.24	27.20	1.54	1.39	1.07	19.33	19.46	19.21
		5	26.99	27.29	27.13	1.70	1.28	0.97	19.32	19.35	19.25



(a)



(b)

그림 5. 두 시퀀스에 대한 현재 버퍼 충만도 (a) 'carphone', 9.6 kbps/30 fps/1 참고 프레임 (b) 'foreman', 19.2 kbps/30 fps/1 참고 프레임

Fig. 5. Current buffer fullness curves for sequences (a) 'carphone' at 9.6 kbps/30 fps/1 reference frame, and (b) 'foreman' at 19.2 kbps/30 fps/1 reference frame.

하도록 선택되어졌다.

그림 4는 두 시퀀스에 대한 비디오 품질을 PSNR (peak signal to noise ratio) 로 알아본 것이다. 그림4에

서 제안한 방식은 기존의 두 가지 방식보다 전체적으로 높은 PSNR을 유지하면서 그 변동폭이 작음을 알 수 있다. 자세한 결과는 표 1에 정리되어 있다. 제안한 방식은 PSNR 변동과 비트율 불일치 측면에서 크게 개선되었음을 알 수 있다. 즉, 제안한 방식의 PSNR 변동은 기존 두가지 방식과^{[3]-[4]} 비교하여, 'carphone' 영상에 있어서, 각각 53.3%, 43.7% 개선되었고, 'foreman' 영상에 있어서, 각각 42.9%, 24.2% 개선되었다. 버퍼 안정도 측면에서 살펴보면 그림 5와 같다. 실험에 사용된 버퍼의 크기는 목표 비트율의 반으로 설정했다. 그림5에서 제안한 방식은 기존의 두 가지 방식에 비해 버퍼 상태가 상당히 안정되어 있음을 알 수 있다.

IV. 결 론

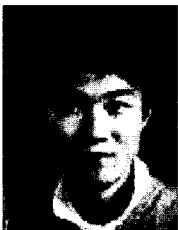
낮은 비트율과 높은 프레임율을 갖는 응용 영역에서 움직임이 많은 영상에 대해 향상된 H.264 프레임 레벨 데이터를 제어 방식을 제안했다. 제안한 방식은 버퍼 상태와 콘텐츠 복잡도를 동시에 고려하여, 목표 비트수를 가능한 양수이도록 양자화 파라미터를 조정하였다. 이러한 방식으로 기존 방식들 대비, 비디오 품질의 변화폭을 줄이고, 비트율의 불일치를 줄였으며, 보다 안정된 버퍼 충만도 레벨을 유지할 수 있었다.

참 고 문 헌

[1] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," IEEE Trans. Circuits Syst. Video Technol. Vol. 13, no. 7, pp. 688 - 703, 2003.
 [2] Z. G. Li, F. Pan, K. P. Lim, G. Feng, X. Lin, and S. Rahardja, "Adaptive basic unit layer rate

- control for JVT,” in 7th Meeting, Pattaya II, JVT-G012-r1, 2003.
- [3] http://ftp3.itu.ch/av-arch/jvt-site/2003_03_Pattaya/JVTG012r1_software.zip.
- [4] M. Jiang, X. Yi, and N. Ling, “Improved frame-layer rate control for H.264 using MAD ratio,” in IEEE Intl. Symp. Circuits Syst., Vol. III, pp. 813 - 816, 2004.
- [5] L. Hung-Ju, C. Tihao, and Z. Ya-Qin, “Scalable rate control for MPEG-4 video,” IEEE Trans. Circuits Syst. Video Technol., Vol. 10, no. 6, pp. 878 - 894, 2000.
- [6] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification, May 2003. Joint Video Team.
- [7] A. Joch, F. Kossentini, H. Schwarz, T. Wiegand, and G.J. Sullivan, “Performance comparison of video coding standards using Lagrangian coder control,” IEEE Intl. Conf. on Image Processing, Vol. II, pp. 501-504, Sept. 2002.

 저 자 소 개



이 창 현(학생회원)

1997년 경희대학교 전자공학과
학사 졸업.

1999년 연세대학교 전파공학과
석사 졸업.

2004년~현재 연세대 전기전자
공학과 박사 과정.

1999년~2004년 삼성전자 통신연구소 연구원
<주관심분야 : H.264, Rate control, 영상통신>



이 원 재(학생회원)

2001년 연세대 전기전자공학과
학사 졸업.

2003년 연세대 전기전자공학과
석사 졸업.

2003년~현재 연세대 전기전자
공학과 박사 과정.

<주관심분야 : 영상처리, SoC 설계>



김 재 석(정회원)

1977년 연세대 전자공학과
학사 졸업.

1979년 KAIST 전기전자공학과
석사 졸업.

1988년 Rensselaer Polytechnic
Institute, NY, 박사 졸업.

1993년~1995년 한국전자통신연구원 책임연구원
1996년~현재 연세대학교 전기전자공학과 교수
<주관심분야 : 통신 및 영상 시스템, VLSI 신호
처리, 임베디드 S/W 및 SoC 구현>