

# 히스토그램 기반의 비트율-왜곡 특성을 이용한 비트율 제어

정희원 홍성훈\*, 유상조\*, 박수열\*, 김성대\*

## Bit-Rate Control Using Histogram Based Rate-Distortion Characteristics

Sung-Hoon Hong\*, Sang-Jo Yoo, Su-Yeal Park\*, Seong-Dea Kim\* *Regular Members*

### 요 약

본 논문에서는 히스토그램 기반의 비트율-왜곡 (R-D) 추정 결과를 이용하여 이웃한 영상들간에 일관된 화질을 제공하는 비트율 제어방식을 제안한다. 히스토그램 기반의 R-D 추정 방식은 부호화에 사용되는 양자화 파라미터 (QP)에 따라 한 영상에서 발생하는 비트량과 왜곡을 예측하거나, 역으로 발생하는 비트량이나 왜곡에 대한 QP 값을 예측할 수 있는 수학적 모델을 제공한다. 이 추정 방식의 가장 큰 장점은 추정을 위한 주 연산이 양자화기에 입력되는 DCT 계수에 대한 히스토그램 또는 가중 히스토그램을 구하는 것이므로 계산량이 적은 것이다. 또 실제 비디오 부호화에 적용할 수 있을 만큼 정확하다. 따라서 이 추정 모델을 이용하는 제안된 비트율 제어 방식은 저 지연과 저 복잡도를 요구하는 응용 분야에 적합하고, 정확한 제어를 수행한다. 제안된 비트율 제어방식은 비디오 버퍼 제한 조건을 만족시키므로써 버퍼의 넘침이나 고갈이 발생하지 않도록 하고, 추가적으로 왜곡 제한에 의하여 이웃한 영상들간에 화질차이가 일정 범위를 벗어나지 않도록 한다. 그리고 하나의 영상에 대하여 하나의 QP를 사용하여 영상 내에서도 일관된 화질을 유지하며, 기준 영상과 비 기준 영상에 대한 버퍼 제한 조건을 차별화함으로써 기준 영상의 화질 열화에 의한 오류 전파를 감소시킨다. 제안된 방식과 MPEG-2의 TM5 (Test Model 5)에서 제안한 비트율 제어 방식의 성능을 비교한 실험으로부터 제안된 방식이 평균 PSNR을 0.521.84 dB 정도 향상시키고, 영상간 그리고 영상 내에서도 일관된 화질을 유지함을 확인하였다.

### ABSTRACT

In this paper, we propose a rate control scheme, using histogram based rate-distortion (R-D) estimation, which produces a consistent picture quality between consecutive frames. The histogram based R-D estimation used in our rate control scheme offers a closed-form mathematical model that enable us to predict the bits and the distortion generated from an encoded frame at a given quantization parameter (QP) and vice versa. The most attractive feature of the R-D estimation is low complexity of computing the R-D data because its major operation is just to obtain a histogram or weighted histogram of DCT coefficients from an input picture. Furthermore, it is accurate enough to be applied to the practical video coding. Therefore, the proposed rate control scheme using this R-D estimation model is appropriate for the applications requiring low delay and low complexity, and controls the output bit-rate and quality accurately. Our rate control scheme ensures that the video buffer do not underflow and overflow by satisfying the buffer constraint and, additionally, prevents quality difference between

\* 한국과학기술원 전기 및 전자공학과 (hsh@sdivision.kaist.ac.kr)  
논문번호 : 98555-1228, 접수일자 : 1998년 12월 23일

consecutive frames from exceeding certain level by adopting the distortion constraint. In addition, a consistent considering the maximum tolerance BER of the voice service. Also in Rician fading channel of  $K=6$  and  $K=10$ , considering  $CLP = 10^{-3}$  as a criterion, it is observed that the performance improvement of about 3.5 dB and 1.5 dB is obtained, respectively, in terms of  $E_b/N_0$ , by employing the concatenated FEC code with pilot symbols.

## I. 서론

MPEG-2<sup>[1]</sup>와 같은 표준 비디오 부호화 방식들은 MC-DCT(motion compensation discrete cosine transform) 방식에 의해서 공간 그리고 시간 방향 상관성을 제거하고 VLC(variable length coding)에 의해서 통계적 상관성을 제거한다. 따라서 이 표준 부호화 방식들은 필연적으로 가변 비트율(variable bit rate: VBR)의 비트열을 발생시킨다. 가변 비트율을 갖는 비디오 비트열을 대역이 제한된 채널을 통하여 전송하는 응용에 있어서는 비트열을 평활화하는 비디오 버퍼와 이 버퍼의 넘침(overflow)이나 고갈 (underflow)을 방지하기 위한 적합한 비트율 제어 방식이 제공되어야 한다.

비록 비트율 제어 그 자체는 비디오 부호화 방식에 관한 표준으로는 정의되어 있지 않으나, 비트율 제어의 성능은 비디오 부호화의 성능에 큰 영향을 미친다. 실제적으로 우수한 비트율 제어는 주어진 대역폭 하에서 복호화된 영상의 화질을 향상시킬 수 있다. 비트율 제어 방식을 설계하는데 있어서 고려해야 할 사항은 첫째, 필수적으로 부호기와 복호기 버퍼의 넘침이나 고갈을 방지해야 한다. 만일 이 조건을 만족하지 못하면 비디오 정보의 손실이 발생하고, 결국 복호화된 영상이 크게 훼손된다. 둘째, 주어진 대역폭 하에서 보다 높은 화질뿐만 아니라 일관된 화질을 제공해야 한다. 셋째, 실시간 비디오 서비스에 적용할 경우 비트율 제어로 인한 지연과 복잡도가 낮아야 한다.

주어진 대역폭 하에서 가장 높은 화질을 얻기 위한 비트율 제어 방식으로 'Lagrangian' 최적화 또는 'dynamic programming'에 의한 최적화 방식들이 제시되었다<sup>[2]-[4]</sup>. 그러나 이 방식들은 비트 할당을 위해 향후 부호화될 영상들에 대한 비트율-왜곡(R-D) 특성을 측정해야 하고 비트 할당 과정에서 반복적인 연산이 요구되므로 복잡도가 매우 높다. 따라서 이 방식들은 양방향 비디오 통신과 같이 저 지연과 저 복잡도를 요구하는 응용 분야에서는 적합하지 않고, 다만 성능 평가를 위한 기준으로 사용

된다.

영상전환나 영상회의와 같이 저 지연과 저 복잡도를 요구하는 응용에는 MPEG-2 TM5<sup>[5]</sup>와 같이 이미 부호화된 영상 데이터로부터 얻어진 정보를 이용한 제한 비트율 제어(feedback rate control) 방식들<sup>[6][7]</sup>이 적합한 방식으로 고려되고 있다. 이 방식들은 현재 관측 가능한 버퍼 충만도(buffer fullness), 이전 영상으로부터 발생한 비트량, 또는 국소적인 영상의 활성화(scene activity) 등을 이용하여 영상 부호화에 이용될 QP 값을 결정한다. 그러나 이 방식들은 과거 부호화 정보에 의존하여 제어를 수행하므로 장면 전환과 같이 급격한 변화를 갖는 영상 영역에서는 화질열화를 피하기 어렵다.

제한 비트율 제어 방식들의 성능을 향상시키기 위해서 현재 프레임 또는 향후 부호화될 프레임에 대한 활성화도 측정 또는 발생 비트량에 대한 모델을 이용하는 순방향 비트율 제어(forward rate control) 방식들<sup>[8]-[11]</sup>이 제시됐다. 그러나 이 방식들은 주로 비트 발생량만을 고려하므로 만족할 만한 화질 향상을 제공하지 못하거나 비트율 제어를 위한 지연과 복잡도가 크게 증가하는 문제가 있다.

이러한 문제점들을 해결하기 위하여 본 논문에서는 히스토그램 기반의 R-D 추정<sup>[12]</sup> 결과를 이용하여 이웃한 프레임들간에 일관된 화질을 유지하면서도 향상된 화질을 얻을 수 있는 순방향 비트율 제어 방식을 제안한다. 제안된 비트율 제어는 R-D 추정 모델을 이용하여 양방향 비디오 통신과 같이 저 지연과 저 복잡도를 요구하는 응용 분야에 적용이 가능하고 정확한 제어를 수행하도록 설계되어졌다. 또한 제안 방식은 비디오 버퍼 제한 조건을 만족시킴으로써 버퍼의 넘침이나 고갈이 발생하지 않도록 하고, 추가적으로 왜곡 제한에 의하여 이웃한 프레임들간에 화질차이가 일정 범위를 벗어나지 않도록 한다. 그리고 하나의 프레임에 대하여 하나의 QP 값을 사용하므로 프레임 내에서도 일관된 화질을 유지하며, 기준 프레임(anchor frame)과 비 기준 프레임에 대한 버퍼 제한 조건을 차별화함으로써 기준 프레임의 화질열화에 의한 오류전파를 감소시킨

다.

본 논문의 구성은 II장에서 본 논문에서 부호화 방식으로 사용되는 MPEG-2 비디오 부호화와 MPEG-2 TM5 비트율 제어 방식의 문제점에 대하여 간단히 기술하고, III장에서 히스토그램 기반의 R-D 추정방식에 대한 내용과 성능평가 결과를 보인다. IV장에서는 히스토그램 기반의 R-D 추정 결과를 이용한 MPEG-2 비디오 부호화에 대한 비트율 제어 방식을 제안한다. V장에서 실험을 통하여 제안된 비트율 제어 방식의 성능을 평가하고, VI장에서 결론을 맺는다.

## II. MPEG-2 비디오 부호화 특성

현재 MPEG-2와 같은 비디오 압축 부호화 표준안들이 제정되어 있다. 이들 표준안들 중 MPEG-2는 비디오 응용 분야에 가장 많이 사용되고 있고, 발생 비트량의 변화가 가장 크다. 따라서 본 논문에서는 제안된 비트율 제어의 평가를 위한 부호화 방식으로 MPEG-2 부호화 방식을 사용한다.

MPEG-2 비디오 부호화에서는 임의 접근(random access)을 위하여 입력 비디오 시퀀스를  $N_{GOP}$  개의 프레임들을 포함하는 GOP(Group Of Pictures) 단위로 나누고 각 프레임을 I-, P- 그리고 B-프레임 중 하나의 형태로 부호화한다. 그리고  $M_B$ 개의 B-프레임이 이웃한 기준 프레임들(I- 또는 P-프레임) 사이에 존재한다. 부호화 효율을 고려하여  $N_{GOP}$ 는 보통 12 또는 15사이의 값을 사용하고,  $M_B$ 는 2 또는 3을 사용한다<sup>[13]</sup>. 그림 1은  $N_{GOP}=12$ ,  $M_B=2$ 일 때의 GOP 구조와 각 프레임의 부호화 형식을 보여준다.

I-프레임은 이웃한 프레임들과 무관하게 현재 프레임 내에서 공간 상관성만을 제거한다. 따라서 I-프레임은 GOP의 첫번째 프레임으로 사용되며 전송 오류를 감소시키는 주기적인 재동기 기능을 제공한다. P-프레임은 움직임 추정과 보상에 의해 이전 I- 또는 P-프레임과의 시간적 상관성을 제거하므로 I-프레임에 비하여 부호화 효율이 좋다. B-프레임은 움직임 추정과 보상을 위해 이웃한 이전 그리고 이후의 I- 또는 P-프레임을 참조하므로 가장 부호화 효율이 좋다. 그러나 B-프레임은 양방향 예측을 수행하므로 지연과 복잡도가 증가하게 한다. 한편 B-프레임은 다른 프레임들의 부호화에 참조되지 않으므로 B-프레임에서의 화질열화나 부호화 오류는 다른 프레임으로 전파되지 않는다. 그러나 I- 또는 P-프레임에서의 화질열화는 다른 프레임에 전파되므로

이 기준 프레임들에 대해서는 어느 정도의 화질이 유지되어야 한다.

비트율 제어 방식으로는 MPEG-2 TM-5 방식이 구현이나 비교 기준으로 가장 많이 사용되고 있으며, 좋은 성능을 갖는 것으로 알려져 있다. 이 방식은 매 프레임마다 이전 프레임들의 복잡도(complexity) 정보를 이용하여 현재 부호화될 프레임의 목표 비트량을 할당하고, 가상 버퍼의 충전도 정보에 따라 MB (MacroBlock) 단위로 QP를 조절하여 현재 프레임에서 발생하는 비트량이 프레임 목표 비트량에 가깝도록 제어한다. 그리고 인간의 시각특성을 고려한 적응 양자화를 수행한다.

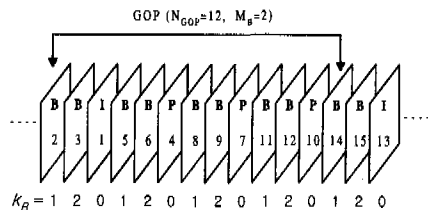


그림 1. GOP 구조와 영상 부호화 방식.

그러나 이 방식은 다음과 같은 문제가 있다.

첫째, 목표 비트 할당에 사용되는 TM5의 복잡도는 이전 프레임의 비트 발생량과 평균 QP 값의 곱으로 구해지므로 화면 전환이나 움직임이 큰 영상에 대해서는 현재 영상의 상태를 정확하게 나타내지 못한다. 따라서 현재 프레임에 대해서 적합하지 못한 목표 비트량이 할당될 수 있고, 이로 인해 이웃한 프레임간에 큰 화질을 변화가 나타날 수 있다<sup>[5]</sup>. 만일 이러한 화질 변화가 자주 발생할 경우 시청자들에게는 깜박이는 듯한 영상이 제공되게 된다. 특히 기준 프레임에 목표 비트량이 너무 적게 할당될 경우 화질열화 현상이 오래 지속될 수도 있다. 둘째, 각 MB에 기준 QP 값을 할당하는 비트율 제어 단계가 한 프레임에 있는 신호 성분들의 특성을 고려하지 않고 단지 가상 버퍼의 충전도만을 고려하므로 가상 버퍼의 상태에 따라 동일한 신호 성분들이 서로 다른 화질을 갖게 된다. 이로 인해서 한 프레임 내에서도 일관된 화질을 유지하지 못한다<sup>[11][14]</sup>. 셋째, 이 방식은 비디오 부호기와 복호기에 실제로 사용되는 비디오 버퍼 크기에 대한 고려와 이 버퍼들의 넘침이나 고갈에 대한 고려를 하지 않는다. 따라서 실제 사용되는 물리적인 버퍼의 넘침과 고갈이 발생할 수 있다<sup>[5]</sup>.

### III. 히스토그램 기반의 R-D 추정

MPEG-2와 같은 표준 비디오 부호화 방식들은 일반적으로 양자화 과정에 적용되는 양자화 스텝 크기 (Quantization Step Size: QS)를 조절하여 발생 비트량 또는 화질을 제어한다. 따라서 영상 부호화에 적용되는 QS에 대한 발생 비트량과 왜곡의 정확한 추정은 효율적이고도 정확한 비디오 부호화 제어에 도움을 준다. 여기서 QS는 양자화 파라미터 QP의 함수로써 n 번째 양자화 파라미터를  $QP_n$  (보통  $QP_n = n$ )이라 하고 이에 대응하는 QS를  $QS_n$ 이라 하면 QS와 QP의 관계는 보통  $QS_n = 2QP_n$ 이고  $QP_n < QP_{n+1}$ 이다. 본 장에서는 제안되는 비트율 제어에 사용되는 히스토그램 기반의 R-D 추정방식<sup>[12]</sup>을 소개하고 실험 결과를 보인다.

#### 1. 비트율 추정

MPEG-2 비디오 부호기의 출력 정보는 움직임 벡터, 각종 헤더 등과 같은 부가 정보와 DCT 계수로부터 발생하는 신호 정보로 구성된다. 여기서 QP의 변화에 영향을 받는 신호 성분은 intra-DC를 제외한 DCT 계수들이고, 높은 비트율로 부호화할 경우 부호기의 출력 비트량은 대부분 DCT 계수로부터 발생된다.

DCT 계수에 대한 부호화는 다음과 같다. 첫째, 양자화 처리가 수행된다. MPEG-2에서 사용되는 양자화와 역 양자화 과정은 MB (MacroBlock)의 부호화 방식에 따라 intra와 non-intra 양자화 과정으로 분류되는데, TMS에서 제시한 양자화기는 다음과 같다<sup>[5]</sup>.

$$\overline{ac}_q(i, j) = (16 \times ac(i, j)) // W_q(i, j) \quad (1)$$

$$QAC_q(i, j) = \left\{ \overline{ac}_q(i, j) + \delta_n(i, j) \right\} // (2 \times QP_n) \quad (2)$$

여기서

$$\delta_n(i, j) = \begin{cases} \text{sign}(\overline{ac}_q(i, j)) \times ((3 \times QP_n) // 4) & \text{for intra} \\ 0 & \text{for non-intra} \end{cases}$$

이고,  $ac(i, j)$ 는 intra-DC를 제외한 DCT 계수이며  $W_q(i, j)$ 는 양자화 방식  $qt \in \{I, N\}$ 에 대한 양자화 매트릭스(quantization matrix)의  $(i, j)$  번째 성분이다.

'//'는 반올림을 하는 나눗셈 연산자이고, '/'는 내림을 하는 나눗셈 연산자이다.

$ac(i, j)$ 에 대한 양자화 과정은 우선 식 (1)에 나타난 바와 같이 양자화 매트릭스 각 성분에 의한 개별적인 양자화를 수행한다. 그리고 양자화 매트릭스 처리된 DCT 계수 (즉,  $\overline{ac}_q$ )는 식 (2)에서와 같이 QP에 의해서 양자화 처리된다. 한편 MPEG-2에서 사용되는 역 양자화는 식 (3)으로 표현되는 역 양자화기를 적용하여 수행된다<sup>[1]</sup>.

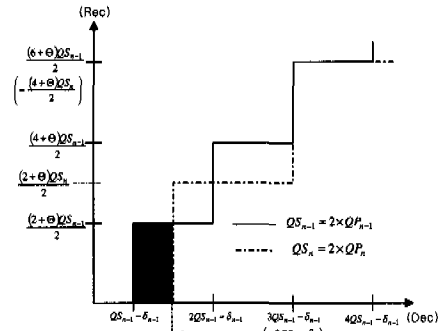


그림 2. MPEG-2 양자화 특성

$$REC_q(i, j) = \left\{ (2 \times QAC_q(i, j) + \Theta(i, j)) \times W_q(i, j) \times QS_n \right\} // 32 \quad (3)$$

여기서

$$\Theta(i, j) = \begin{cases} 0 & \text{for } qt = \text{intra} \\ \text{sign}(QAC_q(i, j)) & \text{for } qt = \text{non-intra} \end{cases}$$

양자화 매트릭스 영향을 무시할 경우, 그림 2는 양자화에 적용되는 QP가  $QP_{n-1}$ 과  $QP_n$  일 때 식 (2)와 (3)으로부터 얻어지는 양자화의 입출력 특성을 나타낸 것이다.

양자화 처리된 DCT 계수들은 run-length 부호화에 의하여 (zero-run, level)인 codeword로 변환된 후 VLC에 의하여 최종적인 비트열로 만들어진다. 여기서 (zero-run, level)로 표현되는 codeword의 수는 양자화 처리된 DCT 계수 중 영이 아닌 성분 (Non-Zero Coefficient: NZC)의 수와 일치한다. 따라서 직관적으로 양자화된 DCT 계수로부터 발생하는 비트량은 NZC의 수에 비례할 것이라고 생각할 수 있다. 그림 3은 I, P 그리고 B 프레임에 대하여 적용되는 QP를 1에서 31로 증가시킬 때 intra-DC를 제외한 DCT 계수로부터 발생하는 비트량과 NZC

수의 관계를 나타낸 것이다. 이 결과로부터 적용되는 QP에 따라 발생하는 비트량과 NZC 개수의 관계는 비록 VLC가 사용되더라도 불구하고 영상의 내용이나 부호화 방식에 거의 무관하게 비교적 선형적으로 비례함을 보여준다. 따라서 현재 영상에 적용되는 QP에 따라 한 DCT 계수들로부터 발생하는 비트량  $Sbit(QP)$ 을 NZC의 수를 추정함으로써 식 (4)와 같이 근사적으로 추정할 수 있다.

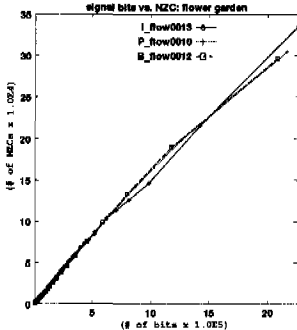


그림 3. 프레임 당 발생하는 비트수와 NZC 수의 관계.

$$Sbit(QP_n) = \sum_{qt \in \{I, P\}} (\alpha_{qt} \times NZC_{qt}(QP_n)) \quad (4)$$

여기서  $NZC_{qt}(QP_n)$ 는 현재 영상에 적용되는 QP가  $QP_n$  일 때  $qt$  방식의 양자화 처리된 DCT 계수들로부터 발생한 NZC의 수이다. 그리고 파라미터  $\alpha_{qt}$ 는 이전 영상들의 DCT 계수로부터 발생된 비트수와 NZC의 수의 비로부터 얻어지는데, 선형 회귀(linear regression)를 이용하여 아래 식과 같이 구한다.

$$\alpha_{qt} = \frac{\sum_{i=1}^P (Sbit_{qt}^i \times NZC_{qt}^i)}{\sum_{i=1}^P (NZC_{qt}^i)^2} \quad (5)$$

여기서  $P$ 는 관찰되어진 이전 프레임의 수이고,  $w$ 는 각각 이전  $i$  번째 프레임에서  $qt$  방식의 양자화 처리된 DCT 계수들로부터 발생한 신호 비트수와 NZC의 수이다.

적용되는 QP에 대응하는 NZC의 수는 다음 절차에 의해서 쉽게 구해지는데, 이 절차는 NZC와 양자화 특성 사이의 관계에 대한 관찰을 통하여 유도되었다. 그림 2에서 빗금 친 영역에 존재하는 DCT 계수들의 양자화 결과는 적용되는 QP가  $QP_{n-1}$  일 때는 '1'이 되고,  $QP_n$  일 때는 '0'이 된다. 이는 적

용되는 QP가  $QP_{n-1}$ 에서  $QP_n$ 으로 한단계 증가함으로써 감소하는 NZC 수와 회색 영역에 있는 DCT 계수의 수가 일치하고, 회색 영역에 있는 DCT 계수의 수는  $QP_{n-1}$ 을 적용하여 양자화된 결과가 '1'인 DCT 계수들의 수와 일치함을 의미한다.

1 단계)의 히스토그램을 다음과 같이 구한다.

$$H_{qt}[UQP_{qt}(i,j)]^{++} \quad (6)$$

여기서  $clipping_{qt}[]$ 는 양자화 방식  $qt$ 가 intra인 경우는 입력을 반올림한 후 0에서 31로 제한하는 연산자이고,  $qt$ 가 non-intra인 경우는 입력을 내림 처리한 후 0에서 31로 제한하는 연산자이다. 또 식 (2)로부터 유도되는 상수로 intra 양자화인 경우는 =0.8이고 non-intra 양자화인 경우는 =0.5이다. 즉  $UQP_{qt}(i,j)$ 는 양자화된 DCT 계수의  $(i,j)$  번째 성분들의 절대치를 1로 만드는 QP 값을 의미한다. 결국  $qt$ 방식의 양자화에 적용되는 QP가  $QP_n$ 에서  $QP_{n+1}$ 로 한단계 증가함으로써 감소되는 NZC의 수가  $H_{qt}[QP_n]$ 에 저장된다. 예를 들어 그림 2에서 회색 영역에 존재하는 입력성분이  $H_{qt}[QP_{n-1}]$ 에 저장되는 것이다.

2 단계)  $H_{qt}[QP]$ 를 다음과 같이 누적 시킨다.

$$NZC_{qt}[QP_n] = \sum_{QP=n}^{31} H_{qt}[QP], \quad 1 \leq n \leq 31 \quad (7)$$

식 (7)로부터  $NZC_{qt}[QP_n]$ 에는  $QP=QP_n$  일 때  $qt$  방식의 양자화 처리된 DCT 계수들로부터 발생하는 NZC의 수가 저장된다. 결국 식 (7)로부터 QP 값에 대응하는 NZC의 수를 쉽게 구할 수 있고, DCT 계수들로부터 발생하는 신호 비트수는  $NZC_{qt}[QP]$ 를 식 (4)에 대입함으로써 구할 수 있다. 또한 적용되는 QP 값이  $QP_n$  일 때 현재 영상으로부터 발생하는 전체 비트량의 예측치는 식 (8)과 같다.

$$\hat{E}(QP_n) = Sbit(QP_n) + Obits \quad (8)$$

여기서  $Obits$ 는 QP 값의 변화에 직접적으로 영향을 받지 않는 정보로부터 발생하는 비트수로 한번의 비트 카운트(bit count)에 의해서 구할 수 있다.

## 2. 왜곡 추정

왜곡 추정은 다음 사실에 근거한다. 첫째, MSE(mean squared error) 관점에서의 왜곡은 양자화에

적용되는 QS의 제곱에 비례한다. 만일 입력 신호가 균등 분포 (uniformly distributed)를 갖고 양자화가 이상적인 균등 양자화이면  $QS=QS_n$ 에 대한 MSE는 가 된다<sup>[15][16]</sup>. 한편 MPEG-2 양자화에서는 인간의 시각 특성을 고려하여 양자화 매트릭스를 사용한다. 이 양자화 매트릭스 영향을 고려하면 QS가  $QS_{n-1}$ 에서  $QS_n$ 으로 한단계 증가할 때 왜곡의 증가량은 에 비례한다. 둘째, QS가  $QS_{n-1}$ 에서  $QS_n$ 으로 한단계 증가할 경우  $QS=QS_{n-1}$ 에서 이미 '0'인 DCT 계수는 QS 값의 증가에 따른 왜곡 증가에 영향을 미치지 못하고 오직 NZC만이 왜곡 증가에 영향을 미친다. 이 사실들에 근거해서 적용되는 QP 값에 대응하는 MSE 관점에서의 왜곡을 추정한다.

양자화 매트릭스 영향을 고려하기 위해서 제안된 왜곡 추정은 다음 절차에 의해서 구해지는 가중 히스토그램과 이의 누적치를 이용한다.

1 단계)  $UQP_{qt}(i,j)$ 의 가중 히스토그램을 구한다.

$$w\_H_{qt}[UQP_{qt}(i,j)] = W_{qt}^2(i,j) \text{ if } ac(i,j) \neq 0. \quad (9)$$

여기서  $UQP_{qt}(i,j)$ 는 (6)에서 정의되었고,  $W_{qt}^2(i,j)$ 는 qt 형식의 양자화 매트릭스의 (i,j) 번째 성분의 제곱이다.

2 단계)  $w\_H_{qt}[QP]$ 를 다음과 같이 누적화 한다.

$$w\_NZC_{qt}[QP_n] = \sum_{QP=n}^{31} w\_H_{qt}[QP], \quad 1 \leq n \leq 31 \quad (10)$$

결국  $w\_NZC_{qt}[QP_n]$ 에는 qt 방식의 양자화에 적용되는 QP가  $QP_n$  일 때 양자화된 DCT 계수 중 영이 아닌 NZC에 대응하는 양자화 매트릭스 성분의 제곱의 합이 저장된다.

본 절에서 소개하는 왜곡 추정 모델은 NZC들이 이웃한 결정 준위(decision level) 사이에서 균일 분포를 갖는다는 가정 하에서 유도된다. 양자화 매트릭스의 영향을 무시할 경우 적용되는 QS가  $QS_n$  ( $n > 1$ ) 일 때  $[kQS_n, (k+1)QS_n]$  구간에 존재하는 입력신호는  $(2k + 1)QS_n/2$ 으로 복원된다. 따라서 MPEG-2 양자화기에 입력되는 입력신호가 균일 분포를 갖는다면  $QS = QS_n$ 인 양자화에 의해 발생하는 왜곡 (MSE)은

$$UD_{qt}(QP_n) = \frac{1}{QS_n} \int_{kQS_n-\delta_n}^{(k+1)QS_n-\delta_n} \left( x - \frac{(2k+\Theta)QS_n}{2} \right)^2 dx \quad (11)$$

이 된다 (그림 2 참조). 따라서 QS 가  $QS_{n-1}$ 에서  $QS_n$ 으로 한단계 증가할 때 ( $QS_n$ )보다 큰 DCT 계수에 대한 왜곡의 증가량은 근사적으로

$$\Delta E_{qt}(QP_n) = \{UD_{qt}(QP_n) - UD_{qt}(QP_{n-1})\} \times \frac{w\_NZC_{qt}[QP_n]}{16^2} \quad (12)$$

으로 추정할 수 있다. 그리고  $[QS_{n-1}, QS_n]$  구간에 존재하는 균일 분포를 갖는 신호에 대한 왜곡 증가량의 합은

$$\Delta UR_{qt}(QP_n) = \frac{\int_{QS_{n-1}-\delta_{n-1}}^{QS_n-\delta_n} x^2 dx}{(QS_n-\delta_n) - (QS_{n-1}-\delta_{n-1})} - UD_{qt}(QP_{n-1}) \quad (13)$$

으로 표현된다. 따라서  $[QS_{n-1}, QS_n]$  구간에 존재하는 DCT 계수에 대한 왜곡 증가량의 합은

$$\Delta RE_{qt}(QP_n) = \Delta UR_{qt}(QP_n) \times \frac{w\_H_{qt}[QP_{n-1}]}{16^2} \quad (14)$$

으로 추정할 수 있다. 결국 QS 가  $QS_{n-1}$ 에서  $QS_n$ 으로 한단계 증가할 때 qt 방식의 양자화에 의한 왜곡 증가량의 합은

$$\Delta D_{qt}(QP_n) = \Delta E_{qt}(QP_n) + \Delta RE_{qt}(QP_n) \quad (15)$$

으로 추정할 수 있고  $QS = QS_n$  일 때 부호화된 영상으로부터 발생하는 왜곡은

$$\hat{D}(QP_n) = \hat{D}(QP_{n-1}) + \sum_{qt \in \{I,N\}} \{d_{qt} \times \Delta D_{qt}(QP_n)\} / num \quad (16)$$

으로 추정되어진다. 여기서 num은 한 프레임에 있는 화소의 수이고,  $d_{qt}$ 는 부호화된 이전 영상들에 대해서 실제로 측정된 왜곡과 추정된 왜곡의 비로부터 얻어지는 비례계수로 선형 회귀를 이용하여 아래 식과 같이 구한다.

$$d_{qt} = \frac{\sum_{i=1}^P (D_{qt}^i \times \hat{D}_{qt}^i)}{\sum_{i=1}^P (\hat{D}_{qt}^i)^2} \quad (17)$$

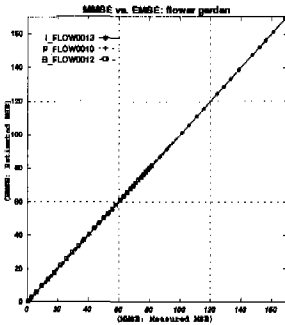


그림 4. "flower garden" 영상에 대한 측정왜곡(MMSE)과 식 (16)으로 구한 추정왜곡(EMSE)의 관계.

여기서  $P$ 는 관찰되어진 이전 프레임의 수이고,  $w$ 는 각각 이전  $i$  번째 프레임에서  $qt$  방식의 양자화 처리된 DCT 계수들에 대한 측정 왜곡과 추정 왜곡이다. 그림 4는 실험적으로 측정된 왜곡과 식 (16)에 의해 추정된 왜곡의 관계를 보여준다. 이 결과는 추정 왜곡들이 측정 왜곡에 대하여 선형적으로 비례하며 이들의 기울기가 영상의 부호화 방식에 무관하게 거의 '1'이 됨을 보여준다.

### 3. R-D 추정 모델의 적합성 검사

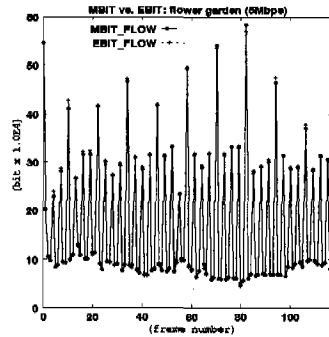
히스토그램 기반의 R-D 추정 방식의 성능을 평가하기 위해서 CCIR601 규격을 갖는 4종류의 표준 MPEG 비디오들 ("flower garden", "football", "mobile", 그리고 "popple") 을 이용하여 모의실험을 수행하고, 실제로 측정된 결과와 추정 결과를 비교하였다.

본 실험에서는 'IBBPBBPBBPBB'인 고정된 GOP구조를 적용하고 각 비디오 시퀀스를 4, 5, 그리고 6 Mbps의 전송률로 부호화하였다.

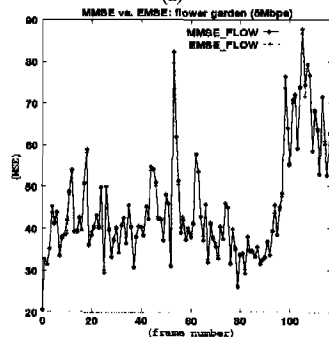
그림 5는 "flower garden" 시퀀스를 5Mbps로 부호화할 때 추정 결과와 측정 결과를 비교한 것으로, 이 추정 결과가 영상의 부호화 방식이나 내용에 무관하게 실제로 측정된 값과 거의 동일함을 보여준다. 표 1은 4 종류 of 비디오 시퀀스들을 각각 4, 5, 그리고 6 Mbps의 전송률로 부호화했을 때 비트율과 왜곡 추정에 대한 평균 상대적 오차를 보여준다. 상대적 오차  $R\_error(i)$ 는 다음과 같다.

$$R\_error(i) = \frac{|Measured(i) - Estimated(i)|}{Measured(i)} \times 100 \quad (\%)$$

여기서  $i$ 는 프레임 번호를 의미한다. 이 결과로부터 비디오의 특성이나 부호화에 적용되는 전송률에 무관하게 큰 추정오차가 발생하지 않음을 알 수 있다 (비트율에 대한 평균  $R\_error$ 는 2.5%를 넘지 않으며, 왜곡에 대한 평균  $R\_error$ 는 1.5%를 넘지 않는다).



(a)



(b)

그림 5. 5Mbps에서 부호화된 "flower garden" 시퀀스에 대한 (a) 비트와 (b) 왜곡의 추정 결과.

## IV. 제안된 비트율 제어

본 장에서는 III. 장에서 소개한 R-D 추정 결과를 이용하여 이웃한 영상들간에 일관된 화질을 유지하면서도 향상된 화질을 얻을 수 있는 비트율 제어 방식을 제안한다. 제안된 비트율 제어방식은 비디오 버퍼 제한 조건을 만족시킴으로써 버퍼의 넘침이나 고갈이 발생하지 않도록 하고, 추가적으로 왜곡 제한에 의하여 이웃한 영상들간에 화질차이가 일정 범위를 벗어나지 않도록 한다. 그리고 하나의 영상에 대하여 하나의 QP를 사용하여 영상 내에서도 일관된 화질을 유지한다. 또한 기존 영상의 화질열화에 의한 오류 전파 문제를 감소시키기 위해서 기존 영상들과 B-프레임과의 제어 절차를 차별화한다.

	Average $R_{error}$ for Bit Estimation (%)			Average $R_{error}$ for Distortion Estimation (%)		
	4Mbps	5Mbps	6Mbps	4Mbps	5Mbps	6Mbps
flower	1.683	1.471	1.711	0.790	0.944	1.199
football	0.811	0.902	1.043	0.603	0.905	0.920
mobile	2.023	2.203	2.347	0.556	0.572	0.925
popple	1.057	1.253	1.637	1.330	1.194	1.311

표 1. 비트율-왜곡 추정 결과

1. 비디오 버퍼와 왜곡 제한 조건

1.1 비디오 버퍼 제한 조건

부호기에서의 비트율 제어는 필수적으로 부호기와 복호기의 버퍼가 넘침이나 고갈되지 않도록 수행되어야 한다. 그런데 CBR 전송에 있어서 부호기와 복호기의 비디오 버퍼 크기가 동일한 경우, 부호기 버퍼가 넘치거나 고갈되지 않으면 복호기 버퍼도 넘침과 고갈이 발생하지 않는다<sup>[17]</sup>. 따라서 CBR 전송의 경우 부호기에서의 비트율 제어는 부호기 버퍼의 넘침과 고갈을 방지하면 된다. 부호화된 비트열들이 부호기 버퍼에 저장된 후 CBR 채널을 통하여 전송될 경우 부호기 버퍼의 넘침이나 고갈이 발생하지 않을 조건은 다음과 같다.

$$0 \leq B^e(i) \leq B_{max}^e \quad \forall i$$

$$B^e(i) = B^e(i-1) + E(i) - R_f \quad (18)$$

여기서  $B^e(i)$ 는  $i$ 번째 프레임이 부호화된 순간 부호기 버퍼 상태이고  $E(i)$ 는  $i$ 번째 프레임에서 발생하는 비트량이며  $R_f$ 는 한 프레임 기간동안 전송되는 평균 비트량이다. 식 (18)로부터 부호기 버퍼의 넘침이나 고갈이 발생하지 않을  $E(i)$ 의 범위는 다음과 같이 구할 수 있다.

$$E(i) \in [E^{min}(i), E^{max}(i)]$$

$$= [\text{MAX}(0, R_f - B^e(i-1)), B_{max}^e + R_f - B^e(i-1)] \quad (19)$$

따라서  $i$ 번째 프레임에 할당되는 목표 비트량  $T(i)$ 의 허용 범위는 다음과 같이 설정한다

$$T(i) \in [T^{min}(i), T^{max}(i)]$$

$$= [\text{MAX}(B_v^e, E^{min}(i)), E^{max}(i) - B_o^e] \quad (20)$$

여기서  $v$ 는 각각 부호기 버퍼 상태가 '0'과  $e$

지나치게 근접하는 것을 막기 위한 값으로 본 논문에서는 다음과 같이 설정 했다.

$$B_v^e = 0.1 \times B_{max}^e$$

$$B_o^e = 0.1 \times B_{max}^e + \frac{k \times R_f}{M_B} \quad (21)$$

여기서  $k_B$ 는 그림 1에 나타난 바와 같이 기준 프레임에 대해서는 '0'이고, 기준 프레임 이후에 부호화되는 B-프레임에 대해서는 값이 '1'씩 증가한다. 따라서 기준 프레임들은 이전 B-프레임에 비해서  $T_f$  만큼 큰 버퍼를 사용하게 되고, 부호기 버퍼 넘침을 막기 위해 적은 비트를 기준 프레임에 할당함으로써 발생하는 기준 프레임의 화질열화를 완화할 수 있다.

1.2 왜곡 제한 조건

이 제한은 이웃한 기준 영상들간에 화질차이가 일정 범위를 벗어나지 않도록 하는 조건을 부여한다. 부호화된 이전 기준 영상의 왜곡이  $D_p$ 이고 PSNR관점에서 화질변화의 허용 범위를 (dB)로 설정하면 현재 기준 영상에 대한 왜곡  $D$ 의 제한 범위는 식 (22)으로 나타낼 수 있다.

$$D \in [D^{min}, D^{max}]$$

$$= [D_p \times 10^{-0.1\lambda\Delta}, D_p \times 10^{0.1\lambda\Delta}] \quad (22)$$

B-프레임은 다른 영상이 참조하지 않고, 동일한 화질인 경우 기준 영상에 비하여 적은 비트수를 발생시킨다. 따라서 이 왜곡 제한은 B-프레임에 대해서는 적용하지 않고 기준 영상에 대해서만 적용한다.

1.3 결합 제한 조건

식 (20)에 나타낸 목표 비트량의 상한과 하한 대응하는 QP 값들을 식 (8)의 비트율 추정 결과를 이용하여 구하면 다음과 같다.



$$\begin{aligned}
 QP_{\min}^E &= \arg \min_{1 \leq QP \leq 31} |\hat{E}(QP) - T^{\max}(i)| \\
 QP_{\max}^E &= \arg \min_{1 \leq QP \leq 31} |\hat{E}(QP) - T^{\min}(i)|
 \end{aligned} \tag{23}$$

따라서 식 (20)의 목표 비트량의 허용 범위에 대응하는 QP의 제한 조건은 식 (24)로 표현된다.

$$C_{QP}^E = [QP_{\min}^E, QP_{\max}^E] \tag{24}$$

비록 식 (24)가 식 (20)과 동일한 조건을 의미하지는 않지만 식 (20)에서 부호기 버퍼에 대한 여유를 두므로 (와 를 이용)는 식 (24)는 실제 버퍼에 대한 제한 조건 (18)를 만족시킬 수 있다.

마찬가지로 식 (22)의 왜곡 제한 조건의 상한과 하한에 대응하는 QP 값들을 식 (16)의 왜곡 추정 결과를 이용하여 구하면 다음과 같다.

$$\begin{aligned}
 QP_{\min}^D &= \arg \min_{1 \leq QP \leq 31} |\hat{D}(QP) - D^{\min}| \\
 QP_{\max}^D &= \arg \min_{1 \leq QP \leq 31} |\hat{D}(QP) - D^{\max}|
 \end{aligned} \tag{25}$$

따라서 식 (22)의 왜곡 제한 조건에 대응하는 QP의 허용 범위는 다음과 같다.

$$C_{QP}^D = [QP_{\min}^D, QP_{\max}^D] \tag{26}$$

결국 식 (24)와 식 (26)으로부터 버퍼 제한과 왜곡 제한을 동시에 만족시키기 위한 QP의 허용 범위는 다음과 같이 구할 수 있다.

$$\begin{aligned}
 C_{QP}^C &= [QP_{\min}^C, QP_{\max}^C] \\
 &= C_{QP}^E \cap C_{QP}^D
 \end{aligned} \tag{27}$$

이 된다. 만일 와 의 교차 영역이 없다면 비디오 버퍼에 대한 제한은 반드시 지켜져야 하므로 는 로 설정한다.

## 2. 비트율 제어 절차

일반적으로 GOP 기간 정도의 짧은 구간 내에 있는 연속된 프레임들간의 영상 변화는 비교적 작은 경향이 있다. 따라서 아직 부호화되지 않은 향후 프레임들의 R-D 특성은 동일한 부호화 방식으로 가장 최근에 부호화된 프레임과 유사하다고 가정할 수 있다. 이러한 관점에서 제안된 비트율 제어는 복잡도와 지연 증가를 피하기 위하여 향후 GOP 기간 내에 부호화될 프레임들은 동일한 부호화 방식으로

가장 최근에 부호화된 프레임과 동일한 R-D 특성을 갖는다고 가정한다. 따라서 향후 GOP 기간 내에 있는 프레임들의 R-D 특성을 가장 최근에 동일한 부호화 방식으로 부호화된 프레임의 R-D 특성을 이용하여 예측한다. 이를 위하여 각 부호화 방식  $pct$  ( $I, P, B$ )에 대한 R-D 추정 결과들(와 을 저장하는데, 매 프레임 마다 식 (8)과 식 (16)으로 현재 프레임에 대한 R-D 추정을 수행하고, 이 결과로 현재 프레임과 동일한 부호화 방식에 대한 R-D 추정 결과를 갱신한다.

제안된 방식은 기존 프레임에 대한 처리와 B-프레임에 대한 처리 과정이 다르다. 수행 절차는 기존 프레임의 경우 이전 기준 프레임의 왜곡, 현재 부호기 버퍼 상태, 그리고 R-D 추정 값들을 이용하여 부호화에 사용될 QP 값을 구하고, 이 QP 값이 왜곡 제한과 버퍼 제한 조건을 만족하도록 수정한다. 그리고 이와 같이 구해진 QP 값을 이용하여 부호화한다. 한편 B-프레임의 경우는 부호화된 이전 그리고 이후 기준 프레임의 왜곡 값들을 이용하여 목표 왜곡을 구하고, 이 목표 왜곡에 가장 근접한 왜곡 추정 값에 대응하는 QP 값을 구한다. 그리고 이 QP 값이 버퍼 제한 조건을 만족하도록 수정하며, 이와 같이 구해진 QP 값을 이용하여 부호화한다.

### 2.1 기준 프레임에 대한 처리

GOP 주기인  $N_{GOP}$  프레임 기간에 속한 프레임들 중 부호화 방식이  $I, P, B$ 인 프레임의 수를 각각  $N_I, N_P, N_B$ 라 하면,  $N_{GOP}$  프레임 기간동안 전송되는 비트량은

$$\begin{aligned}
 T_{GOP} &= (N_I + N_P + N_B) \times R_f \\
 &= N_{GOP} \times T_{CBR}
 \end{aligned} \tag{28}$$

이다. 여기서  $R_f$ 는 한 프레임 기간동안 CBR 채널을 통해 전송되는 비트량이다.

현재 부호화되는 프레임이 부호화 방식  $X$  ( $I, P$ )인 기준 프레임이고 이 프레임에 적용되는 QP 값이  $Q^X$ 라 하면, 향후 ( $N_{GOP} - 1$ ) 프레임 기간동안 각 프레임으로부터 발생하는 왜곡이 현재 프레임으로부터 발생하는 왜곡과 동일한 값을 유지하기 위해 필요한 전체 비트량은 다음과 같이 예측할 수 있다.

$$\begin{aligned}
 \hat{E}_{GOP}(QP_X) &= N_X \times \hat{E}_X(QP_X) + \sum_{pct \neq X} N_{pct} \times \hat{E}_{pct}(QP_{pct}) \\
 QP_{pct} &= EQ[\hat{D}_{pct}(QP), \hat{D}_X(QP_X)]
 \end{aligned}$$

(29)

여기서  $\hat{E}_{pct}(QP)$ 와  $\hat{D}_{pct}(QP)$ 는 부호화 방식이 pct X인 프레임에 대한 비트 발생량과 왜곡의 추정 결과이고,  $\hat{E}_x(Q^x)$ 와  $\hat{D}_x(Q^x)$ 는 적용되는 QP가  $Q^x$  일 때 부호화 방식이 X인 현재 프레임에 대한 비트 발생량과 왜곡의 추정값이다. 따라서  $\hat{E}_{GOP}(Q^x)$ 는 현재 기준 프레임에 적용되는 QP가  $Q^x$ 일 때 이에 대응하는 왜곡  $\hat{D}_x(Q^x)$ 가 향후 ( $N_{GOP} - 1$ ) 프레임 기간동안 유지되기 위해 필요한 전체 비트량을 의미한다. 이와 같이 구해진  $\hat{E}_{GOP}(QP)$ 중  $T_{GOP}$ 에 가장 가까운 비트량에 대응하는 QP는

$$Q^x = \arg \min_{1 \leq QP \leq 31} |\hat{E}_{GOP}(QP) - T_{GOP}| \quad (30)$$

가 된다. 만일  $Q^x$ 가 식 (27)의 결합 제한 조건을 만족하면 이 QP를 이용하여 현재 프레임을 부호화 하고, 그렇지 않으면 결합 제한 조건의 경계치를 사용한다. 결국 부호화 방식이 X인 현재 기준 프레임의 부호화에 사용되는 QP 값은 다음과 같다.

$$Q^x = \begin{cases} Q^x, & \text{if } QP_{min}^C \leq Q^x \leq QP_{max}^C \\ QP_{min}^C, & \text{if } Q^x \leq QP_{min}^C \\ QP_{max}^C, & \text{if } QP_{max}^C \leq Q^x \end{cases} \quad (31)$$

### 2.2 B-프레임에 대한 처리

B-프레임은 양방향 예측을 사용하므로 이전 그리고 이후의 기준 프레임들이 부호화된 후에 부호화 된다. 따라서 이전 그리고 이후의 기준 프레임들에 대한 부호화 결과를 미리 알 수 있다. 만일 현재 부호화될 프레임이 그림 1에 나타난 바와 같이 기준 프레임 이후  $k_B$ 번째 B-프레임이라면 부호화 절차는 다음과 같다. 우선 현재 B-프레임에 대한 목표 왜곡  $D^T$ 를 이전 그리고 이후 기준 영상에 대한 추정 왜곡들을 가중 평균하여 다음과 같이 구한다.

$$D^T = \frac{(M_B + 1 - k_B) \times \hat{D}_{prev} + k_B \times \hat{D}_{next}}{M_B + 1} \quad (32)$$

여기서  $D_{prev}$ 와  $D_{next}$ 는 각각 이전 그리고 이후 기준 영상으로부터 발생된 왜곡의 추정값이고  $M_B$ 는

기준 영상 사이에 있는 B-프레임의 수이다 (그림1 참조). 그리고 현재 B-프레임의 왜곡 추정 결과를 이용하여 목표 왜곡  $D^T$ 와 가장 근접한 왜곡을 발생시키는 QP 값을 다음과 같이 산출한다.

$$Q^B = \arg \min_{1 \leq QP \leq 31} |\hat{D}_B(QP) - D^T| \quad (33)$$

마지막으로 이와 같이 구해진  $Q^B$ 가 식 (24)의 버퍼 제한 조건을 만족하는가를 평가한다. 결국 현재 B-프레임의 부호화에 사용되는 QP 값은 다음과 같다.

$$Q^B = \begin{cases} Q^B, & \text{if } QP_{min}^B \leq Q^B \leq QP_{max}^B \\ QP_{min}^B, & \text{if } Q^B \leq QP_{min}^B \\ QP_{max}^B, & \text{if } QP_{max}^B \leq Q^B \end{cases} \quad (34)$$

## V. 실험

제한된 비트율 제어의 성능을 평가하기 위하여 CCIR601 규격을 갖는 4종류의 표준 MPEG 비디오 들 ("flower garden", "football", "mobile", 그리고 "popple") 을 이용하여 모의실험을 수행하고, MPEG-2 TM5 방식과 비교하였다. 본 실험에서는 각 비디오 시퀀스에 대해서  $N_{GOP}=12$ 와  $M_B=2$  인 고정된 GOP구조를 적용하였고, 채널 전송률은 4, 5, 그리고 6 Mbps의 전송률로 부호화하였다. 그리고 양방향 비디오 통신 서비스를 고려하여 전체 지연을 200ms로 설정하고, 버퍼의 크기는 로 정하였다. 식 (22)의 왜곡 제한에서의 허용 범위는 1 (dB) 로 설정했다.

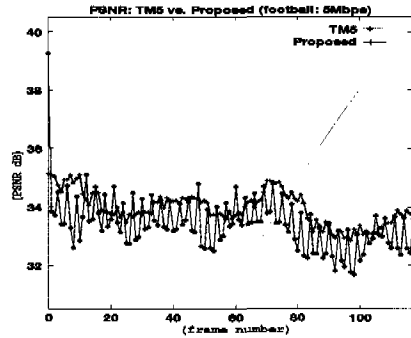
그림 6은 각각 "football" 비디오 시퀀스를 TM5 방식과 제안된 비트율 제어방식으로 부호화한 결과를 보여준다. (a)는 PSNR관점에서의 성능을 나타낸 것으로 제안된 방식이 향상된 화질을 유지할 뿐만 아니라 화질이 매우 일정하게 유지됨을 알 수 있다. (b)는 식 (35)의 판단 기준에 따른 이웃한 영상간에 PSNR 차를 나타낸 것으로 제안 방식이 이웃한 영상간에 화질차이를 크게 감소시킴을 알 수 있다. (c)는 버퍼 상태를 나타낸 것으로 두 가지 방식이 모두 안정된 영역에서 변화되고있다. (d)는 각 영상에서 MB의 MSE에 대한 표준 편차를 나타낸 것으로, 제안된 방식이 TM5 보다 표준 편차를 크게 감소시킴을 보여준다. 따라서 한 영상 내에서도 제안 방식이 일정한 화질을 유지함을 알 수 있다.

$$d\_psnr = PSNR(current\ frame) - PSNR(previous\ frame) \quad (35)$$

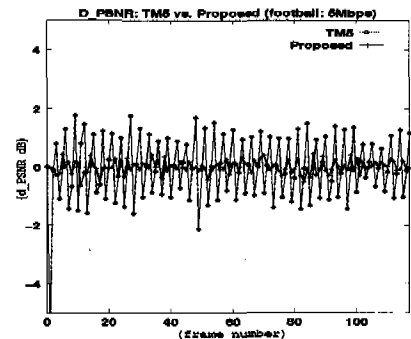
표 2는 4 종류의 비디오 시퀀스들을 각각 4, 5, 그리고 6 Mbps의 전송률에 대해서 TM5 방식과 제안된 비트율 제어방식을 사용하여 부호화했을 때 평균 부호화 성능을 비교한 것이다. 부호화에 사용된 평균 QP 값과 평균 버퍼 충만도 (Buf. St.)는 두 방식이 거의 같다. 그러나 평균 PSNR은 제안된 방식이 TM5보다 0.521.84 dB 정도 향상된다. 이는 제안된 방식의 경우 왜곡 제한에 의해서 영상들간의 화질차이를 감소시키기 때문이다. 즉 복잡한 영상에서의 비트율 증가에 따른 왜곡 감소 효과가 단순한 영상에서의 비트율 감소에 따른 왜곡 증가 효과보다 크므로, 단순한 영상에 할당되는 비트량의 일부를 복잡한 영상에 할당하여 균일한 화질을 유지하면 전체 화질이 향상되게 된다. 또 식 (35)에 정의된  $d\_psnr$ 의 제곱의 평균값 ( $d\_psnr^2$ )으로부터 제안된 방식이 이웃한 영상간에 보다 균일한 화질을 제공할 뿐만 아니라 두드러진 화질 변화가 발생하지 않음을 알 수 있다. 각 영상의 MB들의 MSE에 대한 표준 편차의 평균 ( $MSE(MB)$ )으로부터 제안된 방식이 한 영상 내에 있는 MB들간의 화질도 보다 균일하게 유지함을 알 수 있다. 이는 제안된 비트율 제어 방식은 한 영상에 대하여 동일한 양자화 파라미터를 적용하므로 복잡한 영상영역에서도 화질의 저하가 적은 반면 제한 비트율 제어방식을 사용하는 TM5의 경우는 복잡한 영상영역에서 버퍼 충만도가 높아지므로 이 영역에 대해서는 큰 QP 값을 사용하게 되고 따라서 이 영역에서의 화질의 저하가 크기 때문이다. 이러한 이유에서 제안된 방식은 평균 화질의 향상과 이웃한 영상간에 일관된 화질뿐만 아니라 한 영상 내에서도 일관된 화질이 유지된다. 또한 제안 방식의 성능 향상은 적용되는 전송률에 거의 무관하게 유지됨을 알 수 있다. 따라서 제안된 방식은 적용되는 환경에 큰 영향을 받지 않는 안정된 제어 방식이라 할 수 있다.

### VI. 결론

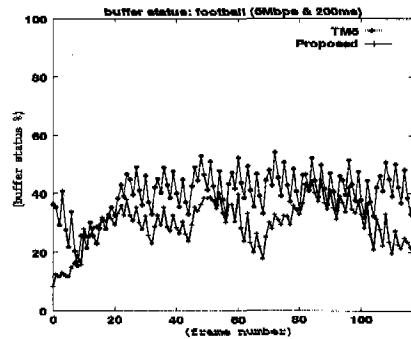
본 논문에서는 히스토그램 기반의 R-D 추정 결과를 이용하여 이웃한 영상들간에 일관된 화질을 유지하면서도 향상된 화질을 얻을 수 있는 비트율 제어 방식을 제안하고 실험을 통하여 성능을 평가했다. 제안된 비트율 제어방식은 비디오 버퍼 제한



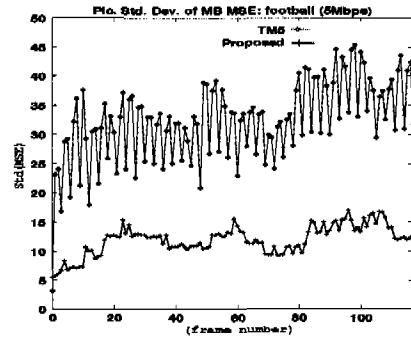
(a)



(b)



(c)



(d)

그림 6. "football" 영상에 대한 제안된 방식과 TM5 방식의 비트율 제어 결과 (5Mbps): (a) PSNR, (b) 이웃한 영상간의 PSNR차, (c) 버퍼 상태, (d) 픽처 당 MB의 MSE의 표준 편차.

(a) Input seq.	TM5 (4Mbps, 200ms)					Proposed (4Mbps, 200ms)				
	PSNR	$d_{psnr2}$	$\sigma_{MSE(MB)}$	QP	Buf. St.	PSNR	$d_{psnr2}$	$\sigma_{MSE(MB)}$	QP	Buf. St.
flower	29.53	0.270	63.73	16.32	47.75	30.37	0.160	34.86	16.43	45.17
football	32.29	1.041	42.35	13.10	40.80	32.81	0.029	17.29	13.08	30.79
mobile	28.69	0.212	86.34	16.37	52.05	29.39	0.073	40.87	16.14	52.50
popple	33.14	0.731	46.04	10.14	38.34	34.59	0.068	19.10	10.61	37.44
(b) Input seq.	TM5 (5Mbps, 200ms)					Proposed (5Mbps, 200ms)				
	PSNR	$d_{psnr2}$	$\sigma_{MSE(MB)}$	QP	Buf. St.	PSNR	$d_{psnr2}$	$\sigma_{MSE(MB)}$	QP	Buf. St.
flower	30.71	0.432	49.47	13.22	46.45	31.51	0.151	25.66	13.14	43.28
football	33.41	1.175	32.24	10.49	40.03	33.95	0.044	11.89	10.50	29.68
mobile	29.67	0.292	70.86	13.51	49.53	30.42	0.089	22.68	13.31	41.99
popple	33.89	0.895	37.36	8.51	37.68	35.56	0.073	12.93	8.52	33.46
(c) Input seq.	TM5 (6Mbps, 200ms)					Proposed (6Mbps, 200ms)				
	PSNR	$d_{psnr2}$	$\sigma_{MSE(MB)}$	QP	Buf. St.	PSNR	$d_{psnr2}$	$\sigma_{MSE(MB)}$	QP	Buf. St.
flower	31.69	0.578	39.78	11.15	45.01	32.50	0.158	19.92	11.03	44.37
football	34.33	1.334	25.34	8.81	39.08	34.93	0.081	8.60	8.75	39.73
mobile	30.53	0.382	58.64	11.52	47.88	31.40	0.082	22.68	11.25	37.21
popple	34.45	0.915	31.86	7.46	35.55	36.25	0.059	9.47	7.26	29.34

표 2. TM5와 제안된 방식의 평균 성능 비교 (a) 4Mbps, (b) 5Mbps, (c) 6Mbps.

조건을 만족시킴으로써 버퍼의 넘침이나 고갈이 발생하지 않도록 하고, 추가적으로 왜곡 제한에 의하여 이웃한 영상들간에 화질차이가 일정 범위를 벗어나지 않도록 한다. 그리고 하나의 영상에 대하여 하나의 QP를 사용하여 영상 내에서도 일관된 화질을 유지한다. 또한 기준 영상과 B-프레임에 대한 버퍼 제한 조건을 차별화함으로써 기준 영상의 화질 열화에 의한 오류 전파를 감소시킨다. 제안된 비트율 제어 방식과 TM5에서 제시한 제한된 비트율 제어 방식을 비교한 실험을 통하여 제안된 방식이 TM5에 비하여 평균 PSNR이 0.521.84 dB 정도 향상시키고, 영상간 그리고 영상 내에서도 일관된 화질을 유지함을 확인하였다. 마지막으로 제안된 방식은 전송률이나 비디오의 특성과 같이 적용 환경에 큰 영향을 받지 않고 항상 향상된 결과를 제공하므로 보다 안정된 제어 방식이라 할 수 있다.

참고 문헌

[1] ISO-IEC/JTC1/SC29/WG11, "Generic coding of moving pictures and associated audio information: Video," ISO-IEC 13818-2, Nov. 1994.  
 [2] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoustic, Speech, and Signal*

*Processing*, vol. 36, no. 9, pp. 1445-1453, Sep. 1988  
 [3] S.-W. Wu and A. Gersho, "Rate-constrained optimal block-adaptive coding for digital tape recording of HDTV," *IEEE Trans. Image Processing*, vol. 3, no. 1, pp. 100-112, Mar. 1991.  
 [4] A. Ortega, K. Ramchandran, and M. Vetterli, "Optimal trellis-based buffered compression and fast approximations," *IEEE Trans. Image Processing*, vol. 3, no 1, pp. 26-40, Jan. 1994.  
 [5] Document ISO-IEC/JTC/SC29/WG11, "Test Model 5," Draft, Apr. 1993.  
 [6] L. Wang, "Bit rate control for hybrid DPCM/DCT video codec," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 5, pp. 509-517, Oct. 1994.  
 [7] C.-T. Chen and A. Wong, "A self-governing rate buffer control strategy for pseudoconstant bit rate video coding," *IEEE Trans. Image Processing*, vol. 2, no. 1, pp. 50-59, Jan. 1993.  
 [8] K.-W. Chow and B. Liu, "Complexity based rate control for MPEG encoder," in *Proc. IEEE Int'l. Conf. Image Processing'94*, pp. 263-267.  
 [9] W. Ding and B. Lui, "Rate control of MPEG

video coding and recording by rate-quantization modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 1, pp. 12-20, Feb. 1996.

- [10] K.-C. Fan and K.-S. Kan, "An active scene analysis-based approach for pseudoconstant bit-rate video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 2, pp. 159-170, Apr. 1998.
- [11] S. K. Peak, J. S. Kang, and Y. S. Seo, "Rate control strategy on human visual sensitivity for MPEG video coder," in *Proc. SPIE Visual communications Image Processing'94*, pp. 322-330.
- [12] S.-H. Hong and S.-D. Kim, "Practical Rate-Distortion Estimation of MPEG-2 Video for Real-Time Processing," *Journal of Electrical Engineering and Information Science*, vol. 4, no. 1, pp. 57-64, Feb. 1999.
- [13] J. Katto and M. Ohta, "Mathematical analysis of MPEG compression capability and its application to rate control," in *Proc. IEEE Int'l. Conf. Image Processing'95*, pp. 555-558.
- [14] L. Wang, "Rate control for MPEG video coding," in *Proc. SPIE Visual Commun. and Image Processing '95*, pp. 53-64.
- [15] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice Hall, 1971.
- [16] A. N. Netravali and B. G. Haskell, *Digital Pictures: Representation and Compression*. New York, NY: Plenum, 1988.
- [17] A. R. Reibman and B. G. Haskell, "Constraints on variable bit rate video for ATM networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 2, no. 4, pp. 3612-372, Dec. 1992.

홍성훈(Sung-Hoon Hong) 정회원



1988년 2월: 영남대학교 전자공학과 졸업 (공학사)  
 1991년 2월: 한국과학기술원 전기 및 전자공학과 졸업 (공학석사)  
 1999년 8월: 한국과학기술원 전기 및 전자공학과

졸업 (공학박사)

1991년 2월 ~ 현재: LG 전자 DTV 연구소  
 <주관심 분야> 비디오 부호화, 멀티미디어 통신 등

유상조(Sang-Jo Yoo) 정회원

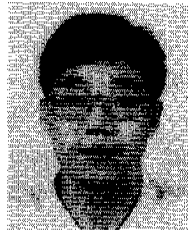


1988년 2월: 한양대학교 전자통신공학과 졸업 (공학사)  
 1990년 2월: 한국과학기술원 전기 및 전자공학과 졸업 (공학석사)  
 1990년 3월~현재:

한국통신 연구개발본부  
 1997년 3월~현재: 한국과학기술원 전기 및 전자공학과 박사과정

<주관심 분야> 멀티미디어 통신, 트래픽 모델링 및 큐잉해석, 멀티미디어 QoS 제어

박수열(Soo Yeal Park) 정회원



1987년 2월: 고려대학교 전자공학과 졸업 (공학사)  
 1990년 2월: 한국과학기술원 전기 및 전자공학과 졸업 (공학석사)  
 1990년 3월~1994년 9월:  
 한국통신 통신망연구소

1994년 10월~1995년 8월:

Hankuk Microtec Research

1995년 9월~현재: 한국과학기술원 전기 및 전자공학과 박사과정

<주관심 분야> 멀티미디어 통신, 영상압축 및 전송, Networking, RealTime OS

김성대(Seong-Dae Kim) 정회원

1977년: 서울대학교 전자공학과 졸업(공학사)  
 1979년: 한국과학기술원 전기 및 전자공학과 졸업 (공학석사)  
 1983년: 프랑스 INPT ENSEEIHT 졸업 (공학박사)  
 1984년~현재: 한국과학기술원 전기 및 전자공학과 교수

<주관심 분야> 영상통신, 컴퓨터 비전, VLSI구현 등