

왜곡-비트율 추정에 근거한 MPEG-2 비디오의 순방향 비트율 제어

정회원 홍성훈*, 김성대*, 최재각**, 홍성용***

Forward Rate Control of MPEG-2 Video Based on Distortion-Rate Estimation

Sung-Hoon Hong*, Seong-Dae Kim*, Jae Gark Choi**,
Sung Yong Hong*** *Regular Members*

요약

비디오 부호화에 있어서 평균화질의 향상뿐만 아니라 이웃한 영상들간에 일관된 화질을 유지하는 것은 매우 중요하다. 본 논문에서는 MPEG-2 비디오에 대한 왜곡-비트율 추정방식과 이 추정결과를 이용하여 일정 비트율(CBR)로 부호화되는 MPEG-2 비디오가 일관된 화질을 유지하면서도 향상된 평균화질을 얻도록 하는 순방향 비트율 제어방식을 제안한다. 본 논문에서 제안하는 왜곡-비트율 추정은 부호화에 사용되는 양자화 스텝 크기에 따라 한 영상에서 발생하는 비트량과 왜곡을 예측하거나, 역으로 발생하는 비트량이나 왜곡에 대한 양자화 스텝 크기를 예측할 수 있다. 제안된 추정 방식의 장점은 실제의 비디오 부호화 응용에 적용할 수 있을 만큼 계산량이 적고 정확하다는 것이다. 한편 제안된 비트율 제어는 프레임 단위로 왜곡-비트율 추정, 목표 비트 할당, 왜곡 제한 그리고 VBV(Video Buffer Verification) 제한 절차에 의해서 부호화에 적용할 양자화 파라미터를 결정하고, 이 양자화 파라미터를 적용하여 영상을 부호화 함으로써 일관되고도 향상된 화질을 유지한다. 또한 화면전환이나 기준영상의 화질열화 등에 의하여 화질저하가 오래 지속되는 문제를 B-picture 제거와 기준영상에 대하여 최소 비트량 할당을 보장함으로써 해결한다. 그리고 제안된 순방향 비트율 제어방식과 MPEG-2의 TMS (Test Model 5)에서 제안한 비트율 제어를 비교한 실험 결과들을 통하여 제안된 방식이 영상들 간에 일관된 화질을 유지할 뿐만 아니라 평균화질도 향상됨을 확인한다.

ABSTRACT

In video coding, it is important to improve the average picture quality as well as to maintain consistent picture

* 한국과학기술원 전기 및 전자공학과

** 경일대학교 제어계측공학과

*** 대구과학대학 전자계산과

論文番號 : 97378-1016

接受日字 : 1997年 10月 16日

quality between consecutive pictures. In this paper, we propose a distortion-rate estimation method for MPEG-2 video and a forward rate control method, using the proposed estimation result, to be able to obtain the improved and consistent picture quality of CBR (Constant Bit Rate) encoded MPEG-2 video. The proposed distortion-rate estimation enable us to predict the distortion and the bits generated from an encoded picture at a given quantization step size and vice versa. The most attractive features of proposed distortion-rate estimation are its accuracy and low computational complexity enough to be applied to the practical video coding. In addition, the proposed rate control first determines a quantization parameter per frame by following procedure: distortion-rate estimation, target bit allocation, distortion constraint and VBV(Video Buffer Verification) constraint. And then this quantization parameter is applied to the encoding so that improved and consistent picture quality can be obtained. Furthermore the proposed rate control method can solve the error propagation problem caused by scene change or anchor picture degradation by using the B-picture skipping and the guarantee of the minimum bit allocation for the anchor picture. Experimental results, comparing the proposed forward rate control method with TM5 method, show that the proposed method makes more improved and consistent picture quality than TM5.

I. 서 론

비디오 압축 부호화에 있어서 고려해야 할 사항은 주어진 대역폭 하에서 화질의 향상 뿐만 아니라 일관된 화질을 유지하는 것이다. MPEG-2⁽¹⁾와 같은 현재의 표준 비디오 부호화 방식들^{(2),(4)}은 가변길이 부호화(Variable Length Coding: VLC) 방식을 사용하므로 필연적으로 가변 비트율의 비트열을 발생시킨다. 만일 가변 비트율을 갖는 비트열을 CBR 채널을 통하여 전송할 경우 이 비트열을 평활화하는 부호기 버퍼와 이 버퍼의 넘침이나 고갈(overflow/underflow)을 방지하기 위한 적합한 비트율 제어방식이 제공되어야 한다.

일반적으로 비트율 제어는 버퍼의 충만도(buffer fullness)에 대한 정보를 이용하여 양자화 파라미터(quantization parameter)와 같은 부호화 변수들을 조절하는 폐환 비트율 제어(feedback rate control) 방식을 사용하는데, 대표적인 방식으로는 MPEG-2의 TM5⁽³⁾에서 제시한 방식이 있다. 이 방식은 매 프레임 주기마다 이전 영상들의 복잡도(complexity) 정보를 이용하여 목표 비트량을 할당하고, 가상 버퍼의 충만도에 따라 폐환 비트율 제어를 수행하며 인간의 시각특성을 고려한 적응 양자화를 수행한다. 그러나 이 방식은 버퍼 상태만을 고려하므로 동일한 영상 영역들에 대해서도 버퍼 상태에 따라 서로 다른 화질이 얻어지고, 비트 발생량과 평균 양자화 파라미터의 곱으로 구해지는 TM5의 복잡도는 실험적으로 구해진 근사화 결과이므로 동일한 영상에 대해서도 적용되는 양자화

파라미터에 따라 서로 다른 결과를 나타낸다. 따라서 일관된 화질을 보장하기 어렵다. 또한 화면 전환 등에 의한 화질열화 현상이 오래 지속되며, 버퍼의 넘침이나 고갈을 방지하지 못할 수도 있다. 이러한 문제점들을 해결하기 위하여 본 논문에서는 적용되는 양자화 파라미터에 따른 왜곡-비트율(distortion-rate) 추정결과를 이용하여 복원된 영상들간에 일관된 화질을 유지하면서도 향상된 화질을 얻을 수 있는 MPEG-2 비디오에 대한 순방향 비트율 제어(forward rate control) 방식을 제안한다.

제안된 순방향 비트율 제어는 현재 부호화할 영상에 대한 왜곡-비트율 추정결과를 이용하여 정확한 비트율 제어를 수행한다. 그리고 VBV(Video Buffer Verification) 제한에 의하여 버퍼의 넘침이나 고갈이 발생하지 않도록 하고, 추가적으로 왜곡 제한에 의하여 이웃한 영상들간에 화질차이가 일정 범위를 벗어나지 않도록 한다. 또한 한 화면에 대해서 동일한 양자화 파라미터를 사용하므로 영상 내에서도 일관된 화질이 유지된다. 그리고 화면전환이나 I 또는 P 영상과 같이 다른 영상에 영향을 주는 기준영상의 화질열화 등에 의하여 화질저하가 오래 지속되는 문제를 B-picture 제기와 기준영상에 대하여 최소 비트량 할당을 보장함으로써 해결한다.

본 논문의 구성은 II장에서 MPEG-2 비디오에 대한 왜곡-비트율 추정방식을 제시하고, III장에서는 II장에서 제시된 왜곡-비트율 추정 결과를 이용한 순방향 비트율 제어 방식을 제안한다. IV장에서는 실험을

통하여 제안된 비트율 제어 방식의 성능을 평가하고 V장에서 결론을 맺는다.

II. MPEG-2 비디오의 왜곡-비트율 추정

MPEG-2 비디오 부호화에서는 일반적으로 입력 비디오 시퀀스가 GOP (Group Of Pictures)들로 분리되고, 각 GOP들은 적어도 하나의 I (Intra coded) 프레임과 다수의 P (Predictive coded) 프레임들 그리고 선택적으로 I 또는 P 프레임들 사이에 B (Bi-directionally coded) 프레임을 포함한다. 그리고 각 프레임은 MB (MacroBlock)들로 구성되는데, I 프레임에 있는 모든 MB의 부호화 방식은 intra 방식이지만 P 또는 B 프레임의 경우는 intra와 non-intra (P 또는 B 형식) 부호화 처리되는 MB들이 공존한다. 한편 양자화는 부호화되는 각 MB의 부호화 방식에 따라 intra 양자화 또는 non-intra 양자화가 적용된다.

MPEG-2⁽¹⁾와 같이 MC-DCT 방식의 영상 부호화를 수행하는 표준 비디오 부호화 방식들⁽²⁻⁴⁾은 일반적으로 양자화 과정에 적용되는 양자화 스텝 크기(Quantization Step Size) QS를 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 양자화 스텝 크기 QS는 양자화 파라미터 QP의 함수로써 n 번째 양자화 파라미터를 QP_n (보통 QP_n = n)이라 하고 이에 대응하는 QS를 QS_n이라 하면 QS와 QP의 관계는 보통 QS_n = 2 * QP_n이고 QP_n < QP_{n+1}이다. 또한 MPEG-2 부호화에서 사용 가능한 양자화 파라미터의 개수는 31개이다.

이미 왜곡-비트율 추정에 관한 여러 연구들⁽⁶⁾⁻⁽⁹⁾이 발표되었으나 현행 표준 비디오 부호화 방식들에 대한 정확한 모델을 제시하지 못하고 여전히 많은 계산량이 요구된다. 따라서 본 장에서는 MPEG-2를 기본 부호화 방식으로 사용하는 비디오 부호화에 적용하기 위한 왜곡-비트율 추정 방식을 제시하고 실험 결과를 보인다. 실험에 사용되는 부호화기는 MPEG-2의 TM5를 적용하였다. 제안된 왜곡-비트율 추정은 추정을 위한 주된 연산이 입력 신호에 대하여 구간별로 히스토그램 또는 가중 히스토그램을 구하는 것이므로 계산량이 적고 최종 결과식이 간단하다. 또한 제안된 추정 과정은 H.261⁽²⁾, H.263⁽⁴⁾, MPEG-1⁽³⁾ 등과 같은 다른 표

준 비디오 부호화에도 동일한 방식으로 적용 가능하다.

1. 적용되는 양자화 파라미터에 대응하는 왜곡추정 제안된 왜곡추정은 다음 두 가지 사실에 근거한다. 첫째, 평균자승오차(Mean Squared Error:MSE) 관점에서 왜곡은 양자화에 적용되는 양자화 스텝 크기 QS의 제곱에 비례한다. 실제적으로 표준 비디오 부호화 방식들에 사용되는 양자화기는 균일 양자화기(uniform quantizer)인데, 만일 양자화 과정이 이상적인 균일 양자화이고 양자화기의 입력신호가 균일분포(uniform distribution)를 갖는다면 적용되는 QS에 대한 왜곡(MSE)은 QS²/12이다. 둘째, 양자화 스텝 크기가 QS_n에서 QS_{n+1}로 한단계 증가할 때 왜곡의 증가에 영향을 미치는 입력신호 성분들은 QS_n으로 양자화 처리했을 때 영이 되지 않는 성분(Non-Zero Coefficients:NZC)들이다. 본 절에서는 이와 같은 사실을 기반으로 하여 MPEG-2 부호화에 있어서 양자화 과정에 적용되는 양자화 파라미터에 대한 MSE관점에서의 왜곡을 추정한다. MPEG-2에서의 양자화와 역 양자화 과정은 다음과 같다.

- Intra quantizer in TM5

$$\overline{ac}_I(i, j) = (16 * ac(i, j)) // W_I(i, j) \quad (1)$$

$$QAC_I(i, j) = \frac{[\overline{ac}_I(i, j) + \text{sign}(\overline{ac}_I(i, j)) * (3 * QP) // 4]}{2 * QP} \quad (2)$$

- Non-intra quantizer in TM5

$$\overline{ac}_N(i, j) = (16 * ac(i, j)) // W_N(i, j) \quad (3)$$

$$QAC_N(i, j) = \frac{\overline{ac}_N(i, j)}{2 * QP} \quad (4)$$

- Intra dequantizer in MPEG-2

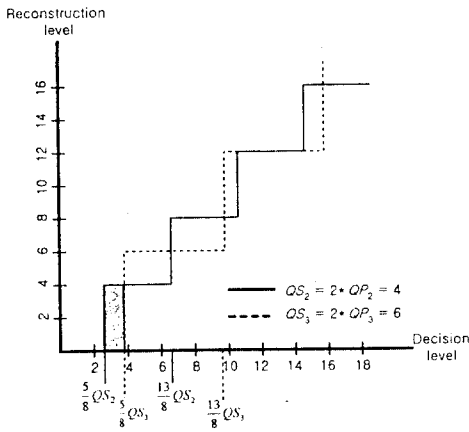
$$REC_I(i, j) = (QAC_I(i, j) * W_I(i, j) * QS) / 16 \quad (5)$$

- Non-intra dequantizer in MPEG-2

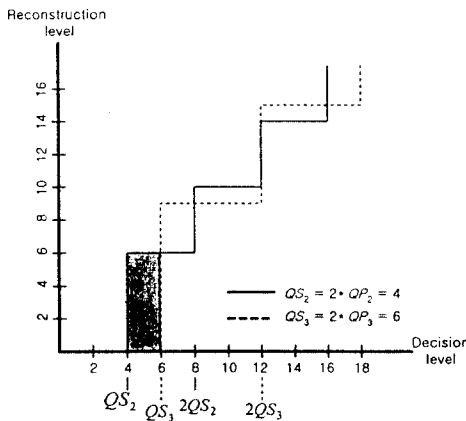
$$REC_N(i, j) = (2 * QAC_N(i, j) + \text{sign}(QAC_N(i, j)) * W_N(i, j) * QS) / 32 \quad (6)$$

여기서 $ac(i, j)$ 는 intra-DC를 제외한 입력 DCT 계수이고, $W_I(i, j)$ 와 $W_N(i, j)$ 는 양자화 매트릭스(quantiz-

ation matrix)의 (i, j) 성분이다. 한편 그림 1은 양자화 파라미터 QP 가 2, 3일 때 식 (2)와 (5)로부터 얻어지는 intra 양자화와 식 (4)와 (6)으로부터 얻어지는 non-intra 양자화의 입출력 특성을 나타낸 것으로 빗금친 영역에 존재하는 입력 신호(즉 $\overline{ac}_I(i, j)$ 또는 $\overline{ac}_N(i, j)$)들에 대한 양자화 처리 결과는 양자화에 적용되는 양자화 파라미터 QP 가 2일 때는 영이 아니고 3일 때는 영이 된다.



(a) Intra 양자화



(b) Non-intra 양자화

그림 1. MPEG-2 양자화 특성
Fig. 1 MPEG-2 quantization characteristics

양자화 과정에 적용되는 양자화 스텝 크기가 QS_n 에서 QS_{n+1} 으로 증가할 때 양자화 매트릭스의 영향을 고려한 왜곡의 증가량 $\Delta D(QP_n)$ 은 $(QS_n^2 - QS_{n-1}^2)$ 에 비례한다. 그러나 QS_n 을 적용했을 때 이미 영인 신호 성분은 왜곡의 증가에 영향을 미치지 못하고 다만 QS_n 을 적용했을 때 존재하는 영이 아닌 신호 성분들만이 왜곡의 증가에 영향을 미친다. 이러한 사실로부터 제안된 왜곡 추정은 다음 절차를 따른다.

① 양자화 매트릭스를 고려한 $AC(i, j) = clipping[\beta * abs(\overline{ac}(i, j))]$ 의 가중 히스토그램을 구한다.

$$QM_pdf[AC(i, j)] += W^2(i, j) \quad (7)$$

여기서 β (는 식 (2)또는 (4)로부터 유도되는 상수로서 intra 양자화기의 경우는 0.8이고 non-intra 양자화기의 경우는 0.5이다. $clipping[\cdot]$ 은 입력을 0에서 31로 제한된 정수값으로 변환하는 연산자이고 $W^2(i, j)$ 는 양자화 매트릭스의 (i, j) 번째 성분을 제공한 값으로 미리 계산된 값을 사용하면 되므로 실제적인 연산은 덧셈 처리이다. 따라서 식 (7)로부터 얻어지는 $QM_pdf[QP_n]$ 에는 양자화 처리된 DCT 계수 중 적용되는 QP 가 QP_n 인 경우는 영이 아니고 QP_{n+1} 인 경우 영으로 되는 DCT 계수에 대응하는 양자화 매트릭스 성분을 제공한 값들의 합이 저장된다.

② 가중 히스토그램의 누적값을 구한다.

$$QM_cdf[QP_M] = \sum_{QP=1}^{QP_M} QM_pdf[QP], 1 \leq M \leq 31 \quad (8)$$

식 (8)로부터 얻어지는 $QM_cdf[QP_M]$ 에는 적용되는 양자화 파라미터가 QP_M 일 때 양자화된 DCT 계수들 중 NZC 에 대응되는 양자화 매트릭스의 제공값들의 합이 저장된다. 그림 2는 MPEG-2의 기본 양자화 매트릭스를 사용했을 때 식 (7)로부터 얻어진 가중 히스토그램과 식 (8)로부터 얻어진 가중 히스토그램의 누적값의 예를 보여준다.

③ 양자화 방식에 따른 왜곡추정

본 절에서 제안하는 왜곡추정은 위에서 언급한 MPEG-2의 양자화 특성에서 결정준위(decision level) 사이에 존재하는 NZC 들의 분포가 균일하다는 가정하

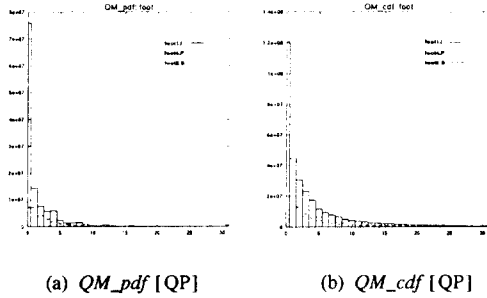


그림 2. $QM_pdf [QP]$ 와 $QM_cdf [QP]$ 의 예
Fig. 2 An example of $QM_pdf [QP]$ and $QM_cdf [QP]$

에서 유도된다. 또한 MPEG-2의 양자화 처리는 부호화 방식에 따라 intra와 non-intra 양자화 처리로 분류되고, 각 양자화 처리의 특성이 서로 다르므로 제안된 왜곡 추정방식은 적용되는 양자화 방식에 따라 다른 결과식을 갖는다.

i) Intra 양자화에 대한 왜곡 추정

MPEG-2의 intra 양자화 특성에서 결정준위 사이에 존재하는 입력신호가 균일분포를 갖는다면, 적용되는 양자화 스텝 크기가 QS_n ($n > 1$)일 때 발생하는 왜곡 (MSE)은

$$\frac{1}{QS_n} \int_{-\frac{5}{8}QS_n}^{\frac{13}{8}QS_n} (x - QS_n)^2 dx$$

이다(그림 1의 (a) 참조). 따라서 양자화 스텝이 QS_{n-1} 에서 QS_n 로 한단계 증가할 때 $\frac{5}{8}QS_n$ 보다 큰 값을 갖는 입력신호 성분에 대한 증가된 왜곡의 합은 (9)와 같다.

$$\Delta E_I(QP_n) = \frac{19 * (QS_n^2 - QS_{n-1}^2)}{192} * \frac{QM_cdf[QP_n]}{16^2} \quad (9)$$

한편 $\frac{5}{8}QS_{n-1}$ 와 $\frac{5}{8}QS_n$ 사이에 존재하는 균일분포를 갖는 입력신호에 대한 왜곡의 증가량은

$$\frac{1}{\frac{5}{8}(QS_n - QS_{n-1})} \int_{\frac{5}{8}QS_{n-1}}^{\frac{5}{8}QS_n} x^2 dx - \frac{19}{192} QS_{n-1}^2$$

이므로 이 구간에 존재하는 NZC들에 의해 증가되는 왜곡의 합은 식 (10)으로 표현할 수 있다.

$$\Delta RE_I(QP_n) = \frac{(6QS_{n-1}^2 + 25QS_{n-1}QS_n + 25QS_n^2)}{192} * \frac{QM_pdf[QP_{n-1}]}{16^2} \quad (10)$$

결국 양자화 스텝 크기가 QS_{n-1} 에서 QS_n 으로 한 단계 증가할 때 intra 양자화에 의해 증가되는 왜곡은 식 (11)으로 추정되고 왜곡은 식 (12)로 추정할 수 있다.

$$\Delta D_I(QP_n) = \Delta E_I(QP_n) + \Delta RE_I(QP_n) \quad (11)$$

$$D_I(QP_n) = D_I(QP_{n-1}) + \Delta D_I(QP_n) / I_pix_num \quad (12)$$

여기서 I_pix_num 은 intra 방식으로 부호화되는 부호화 단위 내에 있는 화소의 개수이다.

한편 intra 양자화에 대한 초기왜곡 $D_I(QP_1)$ 은 0에서 $\frac{5}{8}QS_1$ 사이에 존재하는 균일분포를 갖는 입력신

호에 대한 왜곡이 $\frac{1}{QS_1} \int_0^{\frac{5}{8}QS_1} x^2 dx$ 이므로 NZC에 대한 초기왜곡은 식 (13)으로 추정할 수 있다.

$$D_I(QP_1) = \frac{DC_mse + (25QS_1^2 * QM_pdf[0] + 19QS_1^2 * QM_cdf[1]) / (192 * 16^2)}{I_pix_num} \quad (13)$$

여기서 DC_mse는 DCT계수 중 DC 성분으로부터 발생하는 왜곡의 합이다.

ii) Non-intra 양자화에 대한 왜곡추정

Non-intra 양자화에 대한 왜곡추정도 intra 양자화에 대한 왜곡추정과 동일한 유도과정을 따른다. Non-intra 양자화 특성에서 결정준위 사이에 존재하는 입력신호가 균일분포를 갖는다면, 적용되는 QS가 QS_n ($n > 1$)일 때 왜곡은

$$\frac{1}{QS_n} \int_{QS_n}^{2QS_n} (x - \frac{3}{2}QS_n)^2 dx$$

이 된다. 따라서 QS가 QS_{n-1} 에서 QS_n 로 한단계 증가

할 때 QS_n 보다 큰 값을 갖는 입력 신호에 대한 증가된 왜곡의 합은 식 (14)와 같다.

$$\Delta E_N(QP_n) = \frac{(QS_n^2 - QS_{n-1}^2)}{12} * \frac{QM_cdf[QP_n]}{16^2} \quad (14)$$

한편 QS_{n-1} 에서 QS_n 사이에 존재하는 균일분포를 갖는 입력신호에 의한 왜곡의 증가는

$$\frac{1}{(QS_n - QS_{n-1})} \int_{QS_{n-1}}^{QS_n} x^2 dx - \frac{1}{12} QS_{n-1}^2$$

이다. 따라서 QS_{n-1} 에서 QS_n 사이에 존재 NZC들에 대해 증가되는 왜곡의 합은 식 (15)과 같다.

$$\Delta RE_N(QP_n) = \frac{(QS_n^2 + QS_n QS_{n-1} + \frac{3}{4} QS_{n-1}^2)}{3} * \frac{QM_pdf[QP_{n-1}]}{16^2} \quad (15)$$

결국 양자화 스텝 크기가 QS_{n-1} 에서 QS_n 으로 한 단계 증가할 때 non-intra 양자화에 의해 증가되는 왜곡의 합은 식 (16)으로 추정되고 왜곡은 식 (17)로 추정할 수 있다.

$$\Delta D_N(QP_n) = \Delta E_N(QP_n) + \Delta RE_N(QP_n) \quad (16)$$

$$D_N(QP_n) = D_N(QP_{n-1}) + \Delta D_N(QP_n) / N_pix_num \quad (17)$$

여기서 N_pix_num 은 non-intra 방식으로 부호화되는 부호화 단위 내에 있는 화소의 개수이다.

한편 non-intra 양자화에 대한 초기 왜곡 $D_N(QP_1)$ 은 0에서 QS_1 사이에 존재하는 균일분포를 갖는 입력 신호 성분에 대한 왜곡이 $\frac{1}{QS_1} \int_0^{QS_1} x^2 dx$ 이므로 NZC에 의한 초기왜곡은 식 (18)과 같다.

$$D_N(QP_1) = \frac{QS_1^2 * (4 * QM_pdf[0] + QM_cdf[1])}{12 * 16^2 * N_pix_num} \quad (18)$$

만일 왜곡추정을 한 프레임 단위로 수행할 때, I 프레임에는 intra 양자화 처리만이 사용되므로 식 (12)로 왜곡을 추정하면 되지만 P 또는 B 프레임의 경우는 각 MB의 부호화 방식에 따라 intra와 non-intra 양

자화 처리가 사용된다. 따라서 한 프레임 단위로 추정되는 왜곡은 식 (19)와 같다.

$$D(QP_n) = D(QP_{n-1}) + \{\Delta D_I(QP_n) + \Delta D_N(QP_n)\} / pixel_num \quad (19)$$

여기서 $pixel_num$ 은 한 프레임 내에 존재하는 화소의 개수이다. 한편 왜곡 증가량의 합인 $\Delta D_I(QP_n)$ 과 $\Delta D_N(QP_n)$ 의 산출에 사용되는 가중 히스토그램과 누적값은 부호화 방식에 따라 각각 구해진 값들을 사용해야 한다. 즉 $\Delta D_I(QP_n)$ 의 계산에는 intra 양자화 처리되는 DCT 계수들로부터 얻어진 가중 히스토그램과 누적값이 사용되고, $\Delta D_N(QP_n)$ 의 계산에는 non-intra 양자화 처리되는 DCT 계수들로부터 얻어진 가중 히스토그램과 누적값이 사용된다.

그림 3은 'football' 영상에 대하여 식 (19)에 의해 추정된 왜곡과 실제로 양자화 파라미터를 변화시키면서 측정된 왜곡의 관계를 보여준다. 이 결과는 제안된 왜곡 추정의 결과가 부호화 방식에 무관하게 실제로 측정된 왜곡과 선형적으로 비례함을 보여준다.

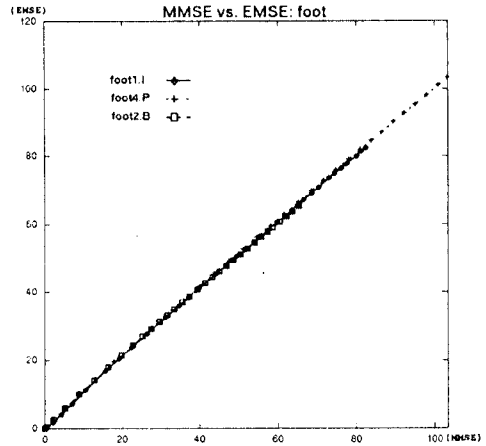


그림 3. 실험 영상에 대한 측정왜곡(MMSE)과 추정왜곡(EMSE)의 관계

Fig. 3 Relationship between measured distortion and estimated distortion for test image

2. 적용되는 양자화 파라미터에 대응하는 비트율 추정 MPEG-2 비디오 부호기의 출력 정보는 움직임 벡

터, 각종 헤더 등과 같은 부가정보와 DCT 계수로부터 발생하는 신호정보로 구성된다. 여기서 양자화 파라미터 QP의 변화에 영향을 받는 신호 성분은 intra-DC를 제외한 DCT 계수들이고, 높은 비트율로 부호화할 경우 부호화기의 출력 비트량은 대부분 DCT 계수로부터 발생된다. 양자화 처리된 DCT 계수들은 run-length 부호화에 의하여 (zero-run, level)인 codeword로 변환된 후 가변 길이 부호화(VLC)하여 최종적인 비트열로 만들어진다. 여기서 (zero-run, level)로 표현되는 codeword의 수는 양자화 처리된 DCT 계수 중 영이 아닌 성분(NZC)의 수와 일치한다. 한편 부가 정보로부터 발생하는 비트량은 동일한 부호화 방식으로 부호화된 이전 영상에서 발생하는 비트량과 비슷한 양을 갖는다. 또한 왜곡-비트율 이론⁽⁶⁾⁽⁷⁾으로부터 엔트로피 부호화를 수행하는 균일 양자화기에 입력되는 신호에 대한 비트와 왜곡의 관계는

$$Sbit(D(QP_n)) \propto \log\left(\frac{\epsilon^2 CDOT \sigma^2}{D(QP_n)}\right)$$

와 같다. 여기서 σ^2 는 입력신호의 분산이고 ϵ 는 입력신호와 관련된 매개 변수인데, 본 논문에서는 실험을 통하여 ϵ 를 intra 부호화의 경우는 5/8로, non-intra 부호화의 경우는 1로 설정했다. 그리고 $D(QP_n)$ 는 입력 신호를 QP_n 인 양자화 파라미터로 양자화했을 때 발생하는 왜곡으로 식 (19)로부터 구해진다. 따라서 적용되는 양자화 파라미터가 QP_n 일 때 한 부호화 단위(프레임 또는 MB)에서 발생하는 비트량은

$$\begin{aligned} Sbit(QP_n) &= \sum_{j \in A} \log\left(\frac{\epsilon^2 \cdot \sigma_j^2}{D_j(QP_n)}\right) \\ &= \log\left[\prod_{j \in A} \frac{\epsilon^2 \cdot \sigma_j^2}{D_j(QP_n)}\right] \end{aligned} \quad (20)$$

으로 표현할 수 있다. 여기서 영역 A에 속하는 입력 신호 j 는 QP_n 으로 양자화한 결과가 영이 아닌 신호 성분들(NZC)이다. 이와 같은 결과로부터 다음 절차에 따라 적용되는 QP에 대응하는 비트 발생량을 추정한다.

① $AC(i, j)$ 의 히스토그램을 구한다.

$$NSC_pdf[AC(i, j)] ++ \quad (21)$$

여기서 $AC(i, j)$ 는 식 (7)에서 정의된 값이다. 따라서 $NZC_pdf[QP_n]$ 에는 QP가 QP_n 에서 QP_{n+1} 로 증가할 때 이에 대응하는 NZC의 감소량이 저장된다. 즉 QP_n 을 적용하여 양자화했을 때는 영이 아니고 한단계 증가된 QP_{n+1} 를 적용하여 양자화하면 영으로 바뀌는 DCT 계수의 개수가 $NZC_pdf[QP_n]$ 에 저장되는 것이다. 그림 4는 식 (21)로부터 얻어진 $NZC_pdf[QP]$ 의 예를 보여준다.

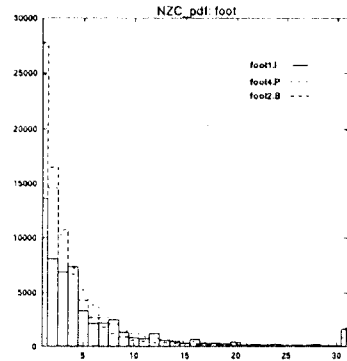


그림 4. $NZC_pdf[QP]$
Fig. 4 An example of $NZC_pdf[QP]$

② $NZC_pdf[QP]$ 와 식 (7)로부터 구해진 $QM_pdf[QP]$ 를 이용하여 식 (20)을 식 (22)로 근사화 함으로써 적용되는 QP에 대응하는 발생 비트량을 추정하는데, 근사화 과정은 양자화기에 입력되는 신호 성분들이 양자화 과정의 결정 준위사이에서 균일한 분포와 분산값을 갖는다는 가정에 근거한다.

$$Sbit(QP_n) = \sum_{k=n}^{31} NZC_pdf[QP_k] * \log\left(\frac{\epsilon^2 \cdot \sigma^2(QP_k)}{D(QP_n)}\right) \quad (22)$$

여기서 $\sigma^2(QP_k)$ 는 양자화 처리에 적용되는 양자화 파라미터가 QP_k 이면 양자화한 결과가 영이 아니고 QP_{k+1} 이면 영이 되는 신호 성분들에 대한 대표 분산값으로 다음 연산에 의해서 구해진다.

$$\sigma^2(QP_k) = \{(QS_k + 1)^2 * QM_pdf[QP_k]\} / (16 * NZC_pdf[QP_k])$$

여기서 QS_k 는 QP_k 에 대응되는 양자화 스텝 크기이고

초기 분산값 $\sigma^2(QP_{31})$ 은 다음과 같이 구한다.

$$\sigma^2(QP_{31}) = \{2 * ap_{31} + 3\}^2 * QM_pdf[QP_{31}] / (16 * NZC_pdf[QP_{31}])$$

$$ap_{31} = 4 * NZC_pdf[QP_{31}] / \sum_{k=27}^{30} NZC_pdf[QP_k] + QP_{31}$$

그림 5는 'football' 영상을 I, P 그리고 B 방식으로 부호화 했을 때 식 (22)에 의해 추정된 비트 발생량과 실제로 양자화 파라미터를 변화 시키면서 측정된 신호 성분(intra DC를 제외한 DCT 계수)으로부터 발생한 비트량의 관계를 보여준다. 이 결과들은 제안된 비트율 추정 결과가 부호화 방식에 무관하게 실제로 측정된 발생 비트량과 선형적으로 비례함을 보여주고, 실제 부호화 처리에 적용할 수 있을 만큼 정확함을 알 수 있다.

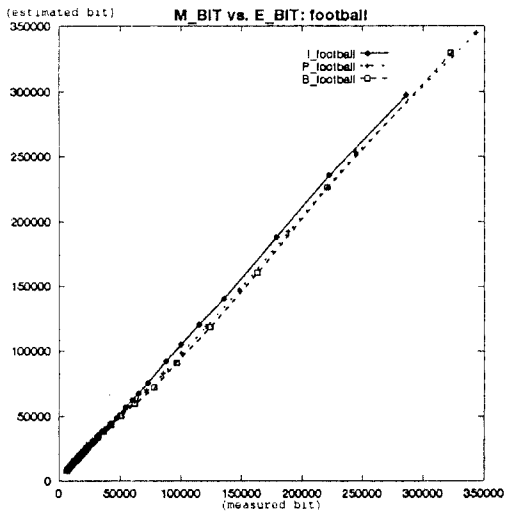


그림 5. 'football' 영상에 대한 비트율 추정 결과
Fig. 5 Rate estimation results for 'football' image

한편 한 영상에 대하여 적용되는 양자화 파라미터에 따라 발생하는 전체 비트량의 추정은 식 (23)과 같이 현재 영상의 신호성분으로부터 발생되는 비트량의 추정값 $Sbit(QP_n)$ 에 부호화 방식이 같은 이전 영상의 부가 정보로부터 발생한 비트량 $Obit$ 를 더함으로써 추

정할 수 있다.

$$Ebit(QP_n) = Sbit(QP_n) + Obit \quad (23)$$

본 장에서는 적용되는 양자화 파라미터에 따른 왜곡과 비트율의 추정 방식을 제안하였는데, 제안된 방식의 주된 연산은 입력 신호에 대하여 구간별로 히스토그램 또는 가중 히스토그램을 구하는 것이므로 계산량이 적고 비교적 정확한 결과를 갖으므로 실시간 부호화에 적용이 가능하다. 따라서 본 논문에서 제안하는 순방향 비트율 제어에는 제안된 왜곡-비트율 추정 결과를 사용한다.

III. 제안된 순방향 비트율 제어

본 장에서는 II 장에서 제시한 왜곡-비트율 추정 결과를 이용하여 복원된 영상들간에 일관된 화질을 유지하면서도 향상된 화질을 얻을 수 있는 순방향 비트율 제어 방식을 제안한다.

제안된 비트율 제어방식은 VBV 제한에 의하여 버퍼의 넘침이나 고갈이 발생하지 않도록 하고, 추가적으로 왜곡 제한에 의하여 이웃한 영상들간에 화질차이가 일정 범위를 벗어나지 않도록 한다. 또한 화면 전환이나 기준영상의 화질열화 등에 의하여 화질저하가 오래 지속되는 문제를 B-picture 제거와 기준영상에 대하여 최소 비트량 할당을 보장함으로써 해결한다.

수행 절차는 매 프레임마다 이전 영상의 목표 왜곡, 현재 버퍼 상태 그리고 현재 프레임에서 발생할 비트량의 추정 값으로부터 목표 비트량을 할당하고 이 목표 비트량이 왜곡 제한과 VBV 제한 범위 내에 있도록 수정한다. 그리고 비트율 추정 결과로부터 목표 비트량에 가장 근접한 비트를 발생시키는 양자화 파라미터를 선택하여 부호화한다. 만일 B-picture를 부호화한 후 버퍼의 상태가 일정 범위를 벗어날 것으로 판단되면 이 B-picture를 부호화하지 않는다. 제안된 비트율 제어의 초기화 과정과 세부 비트율 제어 절차는 다음과 같다.

1. 초기화 과정

이 과정에서는 초기 parameter들의 설정과 부호화 되는 첫번째 영상에 대한 처리를 수행한다.

i) Parameter 설정

$N_{k \in \{I, P, B\}}$: 한 GOP 구간에 존재하는 부호화 방식이 k 인 프레임의 수.

$BIT_{k \in \{I, P, B\}}$: 부호화 방식이 k 인 이전 영상에서 발생한 비트량.

N_{GOP} : 한 GOP 기간에 해당되는 프레임의 수.

$R_{GOP} = bit_rate * N_{GOP} / picture_rate$: 한 GOP 기간 동안 전송되는 비트량

$R_F = bit_rate / picture_rate$: 한 프레임 기간 동안 전송되는 비트량

ii) 첫번째 영상 (I 프레임) 부호화

① 식 (19)와 식 (23)을 이용하여 첫번째 영상(I 프레임)에 대한 왜곡의 추정치 $D(QP)$ 와 발생 비트량의 추정치 $Ebit(QP)$ 를 구한다. 여기서 $Ebit(QP)$ 산출에 사용되는 $Obit$ 는 $0.1 \times Sbit(QP)$ 로 대치한다.

② 부호화할 영상에 할당할 비트량 $Tbit$ 를 결정한다.

여기서 W_P 와 W_B 는 I 영상과 P또는 B 영상에 대한 발생 비트량의 비로서 본 논문에서는 실험적으로 $W_P = 0.7$, $W_B = 0.5$ 로 설정하였다.

③ 양자화 파라미터 결정 및 부호화

추정된 발생 비트량 $Ebit(QP)$ 로부터 ②에서 구한 $Tbit$ 에 가장 근접한 비트량의 추정치에 해당하는 양자화 파라미터 qp 를 구하고 이 양자화 파라미터로 부호화한다.

$$qp = EQ(Tbit, Ebit(QP))$$

$$= |QP| \text{abs}(Ebit(QP) - Tbit) < \text{abs}(Ebit(QP^*) - Tbit)$$

여기서 $EQ(A, E(QP))$ 는 QP 에 대응하는 값(비트량 또는 왜곡의 추정치)이 저장된 $E(QP)$ 중 A 에 가장 근접한 값에 해당되는 변수 QP 를 구하는 연산이다.

④ 초기값 설정

부호화 방식에 따라 한 영상에서 발생하는 비트량과 목표 왜곡 TD 의 초기값을 설정한다.

$$BIT_I = bit(qp)$$

$$BIT_P = W_P * bit(qp)$$

$$BIT_B = W_B * bit(qp)$$

$$TD = D(qp)$$

여기서 qp 는 초기화 절차 ③에서 구해진 양자화 파라미터이고 $bit(qp)$ 는 현재 영상을 양자화 파라미터 qp 로 부호화 했을 때 실제로 발생한 비트량이며 $D(qp)$ 는 qp 에 해당하는 왜곡의 추정값이다.

2. 비트율 제어 과정

부호화된 첫번째 영상이후의 영상들에 대해서는 다음 절차에 따라 순방향 비트율 제어가 수행된다.

i) 왜곡-비트율 추정

식 (19)와 식 (23)을 이용하여 현재 부호화할 영상에 대한 왜곡의 추정치 $D(QP)$ 와 발생 비트량의 추정치 $Ebit(QP)$ 를 구한다.

ii) 목표 비트량 결정

이 절차는 현재 부호화될 영상에 할당할 비트량을 결정하는 단계로 다음 절차를 따른다.

① 이전 영상의 목표왜곡 TD 에 가장 근접한 왜곡을 발생시키는 양자화 파라미터 qp 를 현재 영상에 대한 왜곡의 추정 결과인 $D(QP)$ 로부터 구한다.

$$qp = EQ(TD, D(QP))$$

② 향후 N_{GOP} 기간 동안 목표왜곡 TD 를 유지할 때 발생하는 비트량 e_bit_{GOP} 를 추정한다.

$$Ebit(qp) + N_P * BIT_P + N_B * BIT_B : I - picture$$

$$BIT_I + N_P * Ebit(qp) + N_B * BIT_B : P - picture$$

$$BIT_I + N_P * BIT_P + N_B * Ebit(qp) : B - picture$$

여기서 $Ebit(qp)$ 는 ① 항에서 구한 qp 에 대응하는 추정된 비트량이고, $BIT_{k \in \{I, P, B\}}$ 는 부호화 방식이 k 인 이전 영상에서 발생한 비트량이며 $N_{k \in \{I, P, B\}}$ 는 한 GOP 구간에 존재하는 부호화 방식이 k 인 프레임의 수이다.

③ e_bit_{GOP} 와 현재 버퍼 충만도 B^o 로부터 목표 비트량 $Tbit$ 를 결정하는데, 현재의 버퍼 상태가 버퍼크기의 1/2 이상인 경우만 버퍼 상태에 따른 가중치를 부여한다. 그리고 비트율 추정 결과인 $Ebit(QP)$ 로부터 $Tbit$ 에 근접한 비트량을 발생시키는 양자화 파라미터 qp 를 구한다. 또

한 기준 영상인 I와 P 영상에 대해서는 최소의 화질을 보장하기 위해서 R_F 이상의 비트를 할당한다. 이 과정을 C-code로 표현하면 다음과 같다.

```
ratio = R_GOP / e_bit_GOP;
B_factor = (1.5 - B_i^e / B_max^e) / *B_i^e:버퍼상태, B_max^e:버퍼크기*/
if (B_factor < 1.) {
    if (ratio < 1.) Tbit = Ebit(qp);
    else Tbit = ratio * Ebit(qp);
}
else Tbit = ratio * B_factor * Ebit(qp);
if (anchor picture) Tbit = MAX(Tbit, R_F);
qp = EQ(Tbit, Ebit(QP));
```

iii) 왜곡 제한 평가

이 과정은 이웃한 영상들간에 PSNR관점에서의 화질차이가 일정 범위를 벗어나지 않도록 하는 과정으로, 현재 영상에 대한 목표왜곡이 TD 이고 화질 변화의 허용 범위를 $\pm\Delta$ (dB)로 설정하면 왜곡의 제한 범위는 식 (24)와 같고, 이에 대응하는 양자화 파라미터는 식 (25)와 (26)에 의하여 구한다.

$$D_{\min} = TD * 10^{-0.1 * \Delta}, \quad D_{\max} = TD * 10^{0.1 * \Delta} \quad (24)$$

$$\begin{aligned} qp_D_{\max} &= UQ(D_{\max}, D(QP)) \\ &= \underset{1 < QP \leq 31}{MIN} \{QP \mid D(QP) \geq D_{\max}\} \end{aligned} \quad (25)$$

$$\begin{aligned} qp_D_{\min} &= LQ(D_{\min}, D(QP)) \\ &= \underset{1 < QP \leq 31}{MAX} \{QP \mid D(QP) \geq D_{\min}\} \end{aligned} \quad (26)$$

여기서 $UQ(A, E(QP))$ 는 변수 QP 에 따른 값(비트량 또는 왜곡)이 저장된 $E(QP)$ 에 대해서 A 보다 큰 값들중 최소값에 해당되는 변수 QP 를 구하는 연산이고, $LQ(A, E(QP))$ 는 A 보다 작은 값들중 최대값에 해당되는 변수 QP 를 구하는 연산이다. 따라서 이웃한 영상간에 화질차이가 $\pm\Delta$ (dB) 이하이기 위해서는 목표 비트량 결정과정에서 구해진 양자화 파라미터 qp 를 다음과 같이 제한해야 한다.

$$qp = \max(\min(qp, qp_D_{\max}), qp_D_{\min}) \quad (27)$$

iv) VBV (video Buffer Verification) 제한 평가

비트율 추정의 부정확성으로 인하여 추정된 목표 비트량과 실제로 부호화하여 발생하는 비트량이 일치하지 않으므로 목표 비트량이 지나치게 높거나 낮으면 부호기 버퍼의 넘침이나 고갈이 발생할 수 있다. 따라서 목표 비트량은 부호기 버퍼 상태를 고려하여 재 조절할 필요가 있다. 특히 버퍼의 넘침은 부호화된 영상으로부터 발생한 비트열의 훼손을 의미하므로 버퍼 넘침은 반드시 방지해야 한다.

부호화된 비트열들이 부호기 버퍼에 저장된 후 CBR 채널을 통하여 전송될 경우 부호기 버퍼의 넘침이나 고갈이 발생하지 않을 조건은 다음과 같다.

$$0 \leq B_i^e \leq B_{\max}^e \quad \forall i \quad (28)$$

$$B_i^e = B_{i-1}^e + E_i - R_F \quad (29)$$

여기서 B_i^e 는 i 번째 프레임이 부호화된 순간 부호기 버퍼의 총만도이며 E_i 는 i 번째 프레임에서 발생하는 비트량이다. 식 (28)과 (29)로부터 부호기 버퍼의 넘침이나 고갈이 발생하지 않을 E_i 의 범위는 다음과 같다.

$$\frac{MAX(0, R_F - B_{i-1}^e)}{(E_i^{\min})} \leq E_i \leq \frac{B_{\max}^e + R_F - B_{i-1}^e}{(E_i^{\max})}$$

따라서 부호기 버퍼의 넘침이나 고갈이 발생하지 않을 목표 비트량의 허용 범위는 다음과 같이 설정한다.

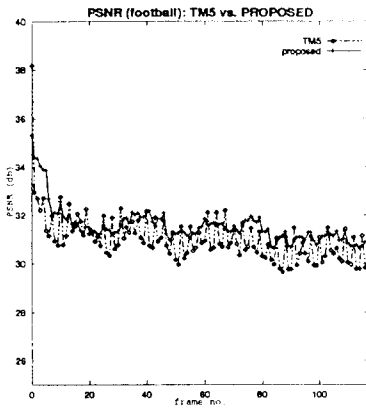
$$\frac{MAX(0.1 * B_{\max}^e, E_i^{\min})}{(Tbit_{\min})} \leq Tbit \leq \frac{\lambda * E_i^{\max}}{(Tbit_{\max})} \quad (30)$$

여기서 λ 는 부호기 버퍼 상태가 B_{\max}^e 에 지나치게 근접하는 것을 막기 위한 상수로 본 논문에서는 0.8를 적용하였다. 한편 발생 비트량과 적용되는 양자화 파라미터는 반비례하므로 식 (30)에 대응하는 양자화 파라미터의 허용 범위는 다음과 같이 설정된다.

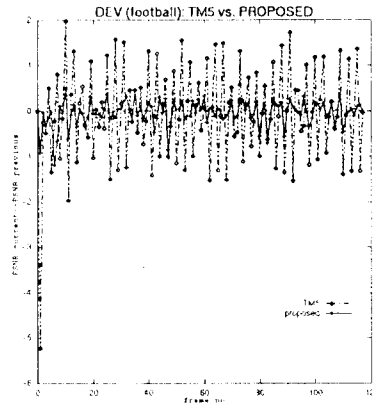
$$\begin{aligned} qp_B_{\min} &\leq qp \leq qp_B_{\max} \\ qp_B_{\min} &= UQ(Tbit_{\max}, Ebit(QP)) \\ qp_B_{\max} &= LQ(Tbit_{\min}, Ebit(QP)) \end{aligned}$$

여기서 qp_B_{min} 는 현재 영상에 대한 비트율 추정결과가 $Tbit_{max}$ 보다 작은 값들 중 최대값에 대응되는 양자화 파라미터이고, qp_B_{max} 는 현재 영상에 대한 비트율

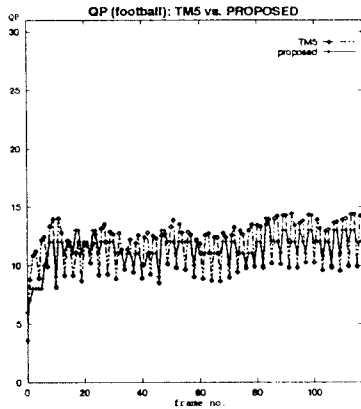
추정결과가 $Tbit_{min}$ 보다 큰 값들 중 최소값에 대응되는 양자화 파라미터이다. 따라서 식 (27)로부터 얻어진 왜곡 제한 처리된 양자화 파라미터 qp 는 다음과 같이



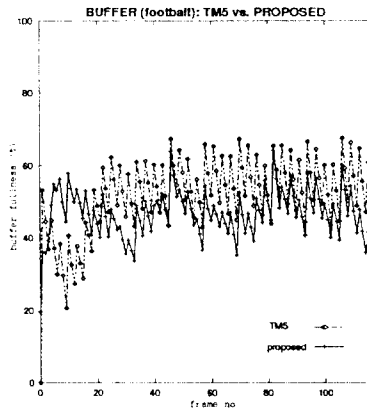
(a)



(b)



(c)



(d)

그림 6. Football 영상에 대한 제안된 방식과 TM5 방식의 비트율 제어 결과

- (a) PSNR
- (b) 이웃한 영상간의 화질차
- (c) 평균 양자화 계수
- (d) 버퍼 상태

Fig. 6 Rate control results of proposed method and TM5 method for football video sequence

- (a) PSNR
- (b) (c) picture quality difference between adjacent pictures
- (c) average QP
- (d) buffer status

VBV 제한된다.

$$qp^* = \max(\min(qp, qp_B_{max}), qp_B_{min}) \quad (31)$$

v) B-picture 제거

만일 B-picture를 부호화한 후 버퍼의 충만도가 버퍼 크기의 80% 이상일 것으로 추정되면 이 프레임은 부호화하지 않는다. 이 처리에 의하여 화면 전환 등에 의하여 저하된 화질이 오래 지속되는 문제를 해결한다.

$$\text{if}((\text{current_picture_type} == B) \&\& ((B_{i-1}^e + Ebit(qp) - R_F) / B_{max}^e > 0.8)) \quad \text{SKIP_B_picture} \quad (32)$$

vi) 목표 왜곡 및 발생 비트량 갱신

만일 현재 영상에 대한 부호화가 VBV 제한이나 B-picture 제거 없이 정상적인 부호화가 수행된 경우 다음 영상에 대한 비트율 제어에 사용되는 정보인 목표 왜곡과 부호화 방식에 따른 발생 비트량을 갱신한다. 즉 식 (27)에 의하여 구해진 양자화 파라미터 qp 와 식 (31)에서 구한 양자화 파라미터 qp^* 가 일치하고 식 (32)의 조건에 해당되지 않는다면 다음과 같이 목표 왜곡과 부호화 방식에 따른 발생 비트량을 갱신한다.

$$\text{if}(qp == qp^*) \&\& !\text{SKIP_B_picture} \{ \\ \quad TD = D(qp^*); \\ \quad BIT_{k \in \{I, P, B\}} = bit(qp^*); \\ \}$$

IV. 실험

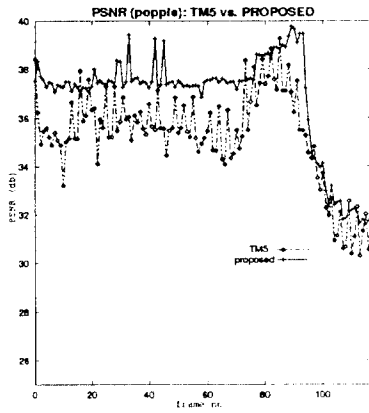
본 장에서는 제안된 순방향 비트율 제어의 성능을 평가하기 위하여 MPEG 표준 영상을 적용하여 실험을 수행하고, MPEG-2 TM5 방식과 비교하였다. 적용된 입력 비디오는 SIF 형식(352×240 spatial resolution, 30 frames/sec, 4:2:0 format)의 비디오 시퀀스이고, GOP 구조는 IBBPBBPBBPBB형식의 고정된 구조를 사용하였다. 채널 전송률 R_{CBR} 은 1.675 Mbps로 설정하였고, 통신 서비스의 경우 전체 지연을 보통 150ms로 설정하므로 버퍼의 크기는 $B_{max}^e = R_{CBR} \times 150ms$ 로 정하였다.

그림 6은 “football” 비디오 시퀀스를 제안된 비트율 제어방식과 TM5 방식으로 부호화한 결과를 보여 준다. (a)는 PSNR관점에서의 성능을 나타낸 것으로 제안된 방식이 향상된 화질을 유지할 뿐만 아니라 화질이 매우 일정하게 유지됨을 알 수 있다. (b)는 식 (33)의 판단 기준에 따른 이웃한 영상간에 PSNR 차를 나타낸 것으로 제안 방식이 이웃한 영상간에 화질 차이를 크게 감소시킴을 알 수 있다. (c)는 영상에 적용된 평균 양자화 파라미터를 나타낸 것으로 제안된 방식이 TM5 보다 일정한 양자화 파라미터 값이 사용됨을 보여준다. (d)는 버퍼 상태를 나타낸 것으로 두 가지 방식이 모두 안정된 영역에서 변화되고 있다.

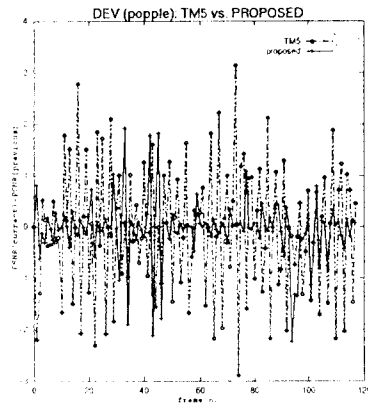
$$d_PSNR = PSNR(\text{currentframe}) - PSNR(\text{previousframe}) \quad (33)$$

그림 7은 “popple” 비디오 시퀀스를 제안된 비트율 제어방식과 TM5 방식으로 부호화한 결과를 나타낸 것으로 “football” 시퀀스와 유사한 결과를 얻는다. 그러나 한가지 특이한 결과는 (c)에 나타난 바와 같이 제안된 방식이 TM5에 비하여 평균적으로 큰 양자화 파라미터를 사용함에도 불구하고 (a)에 나타난 바와 같이 평균 PSNR이 향상된다. 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 양자화 파라미터 값을 사용하게 되고 따라서 이 영역에서의 화질의 저하가 크기 때문이다. 이러한 이유에서 제안된 방식은 평균 화질의 향상과 이웃한 영상간에 일관된 화질 뿐만 아니라 한 영상 내에서도 일관된 화질이 유지된다. 또한 제안된 방식은 한 영상 내에서 동일한 양자화 파라미터를 사용하므로 각 MB 마다 적용되는 양자화 파라미터를 표시하기 위한 부가 정보가 필요하지 않고, 이 부가 정보에 사용되는 비트량을 영상 부호화에 사용할 수 있다.

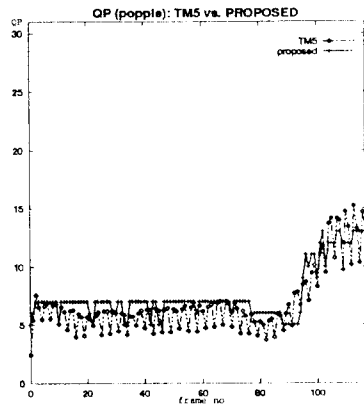
표 1은 제안된 방식과 TM5 방식으로 부호화된 영상 시퀀스의 평균 화질, 평균 발생 비트량, 그리고 평균 양자화 파라미터 관점에서 비교한 것으로 제안된 방식이 평균 화질(\overline{PSNR} 또는 \overline{MSE})의 향상과 이웃한 영상간에 일관된 화질($|\overline{d_PSNR}|$)을 유지함을 알 수 있다.



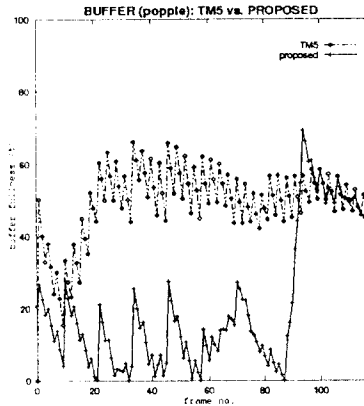
(a)



(b)



(c)



(d)

그림 7. Popple 영상에 대한 제안된 방식과 TM5 방식의 비트율 제어 결과

(a) PSNR (b) 이웃한 영상간의 화질차

(c) 평균 양자화 계수 (d) 버퍼 상태

Fig. 7 Rate control results of proposed method and TM5 method for popple video sequence

(a) PSNR (b) picture quality difference between adjacent pictures

(c) average QP (d) buffer status

표 1. TM5와 제안된 방식의 성능 비교(평균 값)

Table 1. Performance comparison between TM5 and proposed rate control method (average value)

입력 영상	TM5					제안 방식				
	\overline{PSNR}	\overline{MSE}	\overline{QP}	\overline{bits}	$ \overline{d_psnr} $	\overline{PSNR}	\overline{MSE}	\overline{QP}	\overline{bits}	$ \overline{d_psnr} $
Flower garden	28.66	91.25	14.02	55561.56	0.61	29.00	83.74	14.37	55920.16	0.36
football	30.87	54.05	11.86	56095.60	0.74	31.36	47.69	11.65	55388.74	0.17
popple	35.22	22.13	6.94	56240.02	1.03	36.73	16.18	7.65	56271.06	0.36

표 2는 제안된 방식과 TM5 방식으로 부호화된 영상 시퀀스의 화질, 발생 비트량 그리고 양자화 파라미터의 표준 편차 관점에서 비교한 것으로 왜곡의 표준 편차(σ_{MSE})와 d_PSNR 의 표준 편차(σ_{d_PSNR})로부터 제안된 방식이 각 영상 시퀀스 전체와 이웃한 영상간에 일관된 화질이 유지됨을 알 수 있다. 그리고 양자화 파라미터의 표준 편차(σ_{QP})로부터 제안된 방식이 TM5에 비하여 영상 부호화에 적용되는 양자화 파라미터가 균일함을 알 수 있다. 또한 각 영상으로부터 발생하는 비트량의 변화(σ_{bit})가 적으므로 보다 작은 버퍼를 사용할 수 있고, 따라서 버퍼에서의 지연을 감축할 수 있다.

화된 한 영상에서 발생하는 왜곡-비트율을 비교적 정확하게 추정하며, 특히 추정을 위한 주된 연산이 양자화에 입력되는 신호의 구간별 히스토그램 또는 가중 히스토그램을 구하는 것이므로 계산량이 매우 적다.

한편 제안된 비트율 제어방식은 VBV 제한에 의하여 버퍼의 넘침이나 고갈이 발생하지 않도록 하고, 추가적으로 왜곡 제한에 의하여 이웃한 영상들간에 화질차이가 일정 범위를 벗어나지 않도록 했다. 또한 한 화면에 대해서 동일한 양자화 파라미터를 사용하므로 영상 내에서도 일관된 화질이 유지되고 화면 전환이나 기준 영상에서의 화질저하 등에 의하여 저하된 화질이 오래 지속되는 문제를 B-picture 제거와 기준 영

표 2. TM5와 제안된 방식의 성능 비교(표준 편차)

Table 2. Performance comparison between TM5 and proposed rate control method (standard deviation)

입력 영상	TM5				제안 방식			
	σ_{MSE}	σ_{QP}	σ_{bit}	σ_{d_PSNR}	σ_{MSE}	σ_{QP}	σ_{bit}	σ_{d_PSNR}
Flower garden	21.61	2.61	40309.41	0.78	18.66	1.71	29522.64	0.52
football	8.38	1.83	27499.00	0.90	4.17	0.95	18035.08	0.23
popple	12.46	2.84	21470.20	1.27	10.86	2.34	15182.96	0.62

V. 결론

본 논문에서는 MPEG-2 비디오에 대한 왜곡-비트율 추정방식을 제안하였고, 이 추정 결과를 이용하여 복원된 영상들간에 일관된 화질을 유지하면서도 향상된 화질을 얻는 순방향 비트율 제어방식을 제시하고 실험을 통하여 성능을 평가했다.

제안된 왜곡-비트율 추정 방식은 MPEG-2로 부호

상에 최소 비트량 할당을 보장 함으로써 해결했다.

제안된 순방향 비트율 제어 방식과 TM5에서 제시한 변환 비트율 제어 방식을 비교한 실험을 통하여 제안된 방식이 부호화된 영상의 화질 향상뿐만 아니라 영상간 그리고 영상 내에서 일관된 화질을 유지함을 확인하였다. 그리고 부호화된 각 영상으로부터 발생하는 비트량의 변화가 작으므로 보다 작은 버퍼를 사용할 수 있고, 따라서 버퍼에서의 지연을 감축할

수 있음을 확인했다.

참 고 문 헌

1. ISO-IEC/JTC1/SC29/WG11, "Generic coding of moving pictures and associated audio information: Video," ISO-IEC 13818-2, Nov. 1994.
2. CCITT, Working Party XV/1, "Draft of recommendation H.261: video codec for audiovisual services at $P \times 64$ kbit/s," July 1990.
3. ISO-IEC/JTC1/SC29/WG11, "Coding of moving pictures and associated audio," CD 11172-2, Nov. 1991.
4. ITU-T Study Group 15, "Draft ITU-T recommendation H.263: video coding for low bitrate communication," May 1996.
5. Document ISO-IEC/JTC1/SC29/WG11, "Test Model 5," Draft, Apr. 1993.
6. T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice Hall, 1971.
7. H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. IT-14, PP. 676-683, Sept. 1968.
8. N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice Hall, 1984.
9. A. N. Netravali and B. G. Haskell, *Digital Pictures: Representation and Compression*. New York, NY: Plenum, 1988.

홍 성 용(Sung Yong Hong)

정회원

현재: 대구과학대학 전자계산과 조교수
한국통신학회 논문지 제22권 9호 참조

홍 성 훈(Sung-Hoon Hong)

정회원

현재: 한국과학기술원 전기 및 전자공학과 박사과정
한국통신학회 논문지 제 23 권 5 호 참조

김 성 대(Seong-Dae Kim)

정회원

현재: 한국과학기술원 전기 및 전자공학과 교수
한국통신학회 논문지 제 22 권 5 호 참조

최 재 각(Jae Gark Choi)

정회원

현재: 경일대학교 제어계측공학과 전임강사
한국통신학회 논문지 제 22 권 6 호 참조