

확률 분포를 고려한 저 전송률 비디오 부호기의 균등 비트 할당 기법 연구

한 성 옥(韓聖燮), 서 동 완(徐東完), 최 윤 식(崔潤植)

연세대학교 전기전자공학과

전화 : (02) 2123-2774 / 팩스 : (02) 362-6444

Equal Bit Rate Control for Low Bit-Rate Coder by Using Frame Statistics

Seong-wook Han, Dong-wan Seo, Yoon-sik Choe

Department of Electronics and Electrical Engineering Yonsei University

E-mail : suphan@image3.yonsei.ac.kr, md@image3.yonsei.ac.kr, yschoe@yonsei.ac.kr

Abstract

In typical block-based video coding, the objective of RC(Rate Control) is to select the quantization parameters so that the encoder produces bits at the rate of the channel and the overall distortion is minimized. To reduce the huge amount of computations required for offline RC, there have been significant efforts to speed up the process of video encoders. Those efforts have been mainly focused on the modes for bit rate and distortion in types of coders, in terms of the quantization parameters. Because previous works related to model based online RC are based on statistics of previous frame, it occurs the problem such that allocates bits unequally without regard to current frame statistics. In this thesis, an equal bit allocation scheme using current frame statistics is proposed.

I. 서론

비트율 제어 기법은 일정한 양의 비트를 사용하면서 보다 좋은 화질로 비디오를 압축하는 방법에 관한 연구이다. H.263+나 MPEG-4와 같은 DCT기반의 압축 표준들은 양자화기의 크기를 조절하여 비트량과 복원 영상의 화질을 결정하기 때문에 많은 비트율 제어 기법들이 프레임 비트량을 채널 용량에 맞추면서 부호화 오차를 최소로 만들기 위해 각각의 블록에 적합한 양자화기의 크기를 결정한다. 초기의 비트율 제어 기법들은 채널 용량이 제한되어 있기 때문에 주어진 용량 내에서 비트를 할당하는 기법들만을 연구하였다. [1][2] 그 이후 버퍼 조절과 함께 화질 개선을 동시에 하고자 하는 방법들이 제안되었는데 라그랑제 방법과 다이나

믹 프로그래밍을 이용한 연구들이 대표적인 것이다. 두 가지 방법에 대한 연구는 각각의 프레임이 독립적이라고 가정했을 때 뿐 만이 아니라 서로 상호 의존적인 관계에 있을 때의 최적의 해를 구했다. [3] 그러나, 현재 프레임에 대한 R-D의 특성을 미리 계산하여 양자화기를 추정하기 때문에 부호화 시간이 오래 걸리는 단점이 있다.

실시간 비디오 전송을 위해서는 복부호기 양단에서 비디오 정보를 전송하는 데 필요한 지연이 매우 작아야 한다. 이를 위해서는 두 가지 측면이 고려되어야 하는데, 첫째, 각각의 프레임을 부호화 했을 때 생기는 비트량이 채널 용량을 넘어서 생기는 버퍼 지연의 문제와 둘째, 부호화시 요구되는 계산상의 복잡성 문제이다. 버퍼 지연 문제는 일반적인 비트율 할당의 기본 전제인 비트 할당을 통해서 해결할 수 있지만 두 번째 문제는 현재 프레임의 R-D 특성을 미리 예측하는 모델링 과정 [4][5][6]을 통해 해결해야 한다.

양자화 오차를 R-D 특성에 모델링하는 TMN8[6]이 제안되었는데 TMN8은 각 프레임마다 할당된 비트 수를 비교적 정확히 맞추기는 했지만 프레임 내의 블록들에 비트를 불균등하게 할당하는 또 다른 문제가 발견되었다. 이전 프레임의 확률 분포를 이용하여 현재 프레임의 양자화 매개 변수 할당에 이용하였기 때문이었다. 특히, 프레임간에 변화가 심한 경우 전체 프레임의 앞부분 또는 뒷부분에만 너무 많은 비트를 할당하는 현상이 일어났다.

본 논문에서는 현재 프레임의 확률 분포를 고려하지 않아 발생하는 불균등 비트 할당 문제의 해결 방법에 관한 연구를 하였다. 이런 문제를 해결하기 위해서는 무엇보다도 프레임간 확률 분포의 변화를 인식하여야 한다. 기존의 모델 기반 비트율 제어 기법들은 프

레이를 부호화하기 시작할 때, 초기 매개 변수를 이전 프레임으로부터 계산된 값을 그대로 사용하기 때문에 현재 프레임이 이전 프레임으로부터 영향을 받는다. 그래서, 본 연구에서는 양자화기의 크기를 결정할 때 사용된 매개 변수의 변화를 조사하여 프레임간의 확률 변화를 찾고자 한다. 즉, 부호화전의 매개변수 초기 값과 부호화 후의 매개변수 값을 비교하여 확률 분포의 변화를 감지하였다. 이런 방법으로 확률 변화를 찾은 경우에만 새로 구한 매개변수를 이용하여 다시 부호화해서 앞에서 언급한 불균등 비트 할당 문제를 해결하였다.

II. TMN8의 비트율과 왜곡 모델링

H.263과 MPEG과 같은 전형적인 블록기반 비디오 부호화 과정에서 매크로 블록의 비트량과 왜곡은 변환 계수들을 양자화하는 양자화기의 크기에 의해 영향을 받는다. 따라서, TMN8에서는 비트량과 왜곡에 대한 특성을 매크로 블록의 양자화 스텝의 크기를 이용해 모델링하였다. 비트율 모델($R(Q)$)과 왜곡 모델($D(Q)$)은 아래와 같다.

$$R(Q) = \sum_{i=1}^N A(K \frac{\sigma_i^2}{Q_i^2} + C) \quad (1)$$

$$D(Q) = \frac{1}{N} \sum_{i=1}^N \alpha_i^2 \frac{Q_i^2}{12} \quad (2)$$

A : 매크로 블록의 픽셀 수

K : 모델 파라미터

C : 헤더 길이 (bpp)

Q_i : i 번째 매크로 블록의 양자화 크기

σ_i^2 : i 번째 매크로 블록의 분산

α_i^2 : i 번째 매크로 블록의 가중치

위에서 설정한 비트율과 왜곡에 대한 모델을 이용하여, 비트율 제어에 관한 최적의 해를 구할 것이다.

즉, 매크로 블록의 비트량, $\sum_i B_i$ 이 목표 비트량, B , 과 같으면서, 식(3.5)의 왜곡을 최소로 만드는 양자화 매개 변수 $Q_1^*, Q_2^*, \dots, Q_N^*$ 를 찾는 것이다.

$$Q_1^*, Q_2^*, \dots, Q_N^* = \arg \min_{Q_1, \dots, Q_N} \frac{1}{N} \sum_{i=1}^N \alpha_i^2 \frac{Q_i^2}{12} \quad (3)$$

$$\sum_{i=1}^N B_i = B$$

라그랑제 상수를 이용하여 식(3)의 해를 구하면 다음

과 같다. 프레임의 확률 분포와 상수만으로 양자화 변수를 구하는 식은 다음과 같이 구할 수 있다.

$$Q_i = \sqrt{\frac{AK}{B-ANC} \frac{\alpha_i}{\sigma_i} \sum_{i=1}^N \alpha_i \sigma_i} \quad (4)$$

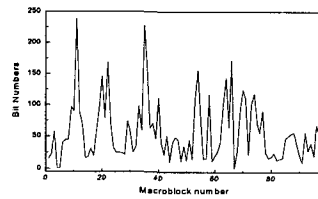
III. 불균등 비트 할당 문제

그림2는 Foreman 영상의 특징은 전체 프레임의 절반에 해당하는 앞부분 매크로 블록들은 거의 하얀 블록만으로 구성되어 있고, 프레임의 중간에 해당하는 얼굴부분부터는 복잡한 형태를 이루는 점이다. 따라서, 프레임의 앞부분에는 움직임 벡터만을 이용하여도 부호화가 가능하기 때문에, 매우 적은 량의 비트가 필요할 것이다. 하지만, 상대적으로 복잡한 얼굴 부분이나 프레임의 마지막에 해당하는 손 부분은 많은 정보를 담고 있으므로 많은 량의 비트를 할당해야한다는 것을 쉽게 생각할 수 있다. 하지만 실제로 TMN8에서는 실제 영상의 정보와는 관계없이 비트를 할당하는 불균등 비트 할당 문제가 발생한다.

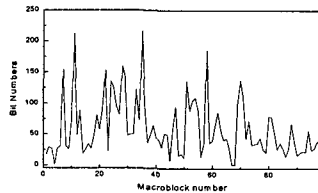


(a) 68번째 프레임 (b) 69번째 프레임

그림1 Foreman 실험 영상



(a) 48 kbps - 68th frame



(b) 48 kbps - 69th frame

그림 2 TMN8 기법의 비트 할당 실제 분포도

그림 2는 건물의 하얀 벽과 인부의 헬멧에 해당하는 하얀 블록이 많은 프레임의 앞부분에 많은 비트가 할당되었다. 반면에 상대적으로 복잡도가 매우 심한 프레임의 뒷부분에 해당하는 매크로 블록들에는 오히려 앞부분보다 적은 양의 비트를 사용하였다. 프레임의 앞부분에서 너무 많은 비트를 할당하였다.

실제 영상 정보와 관계없는 불균등 비트 할당 문제가 생기는 이유는 이전 프레임의 확률 분포를 이용하여 현재 프레임에 적용하였기 때문이다. 프레임간에 확률 분포 변화가 심한 경우에는 현재 프레임의 정보와 상관없이 복잡한 곳에 비트를 적게 할당하고 단순한 곳에 비트를 많이 할당하게 된다.

VI. 확률 분포를 고려한 비트 할당 기법

식(4)에서 α_i, σ_i 등의 매개변수는 각각의 매크로 블록마다 독립적으로 구할 수 있어서, 다른 매크로 블록들과 관계가 없다는 것을 알 수 있다. 식(4)에서는 K 만이 다른 매크로 블록들과 관계가 있는데, 아래와 같이 설정한다.

$$K = \frac{B'_{LC,i}(2QP)^2}{16^2\sigma_i^2} \quad (5)$$

식(5)에서 분산을 이용하여 i 번째 매크로 블록의 K 값을 계산하고,

$$K_i = K_{i-1}(i-1)/i + K/i \quad (6)$$

을 통해 프레임내의 i 번째 매크로 블록 이전까지 계산된 모든 K 들의 평균을 구한다. K 를 구하는 이유는 K 값의 평균을 구하여 현재 매크로 블록의 분산이 전체 프레임의 분산의 평균값과 얼마나 차이가 있는지를 알아보기 위해서이다. 만약 현재 프레임의 분산이 분산의 평균값과 차이가 많이 나면, 많은 비트를 할당하고, 차이가 적으면 작은 양의 비트를 할당하기 위해서이다. 매크로 블록의 분산의 변화를 통해 양자화 스텝 크기를 구하는데 사용되는 K 를 설정하는 식은 아래와 같다.

$$K = K_j(i/N) + K_1(N-i)/N \quad (7)$$

위의 식에서 K_1 값은 이전프레임에서 구한 K 값을 대입한다. 현재 매크로 블록의 분산이 커지면 K 의 값

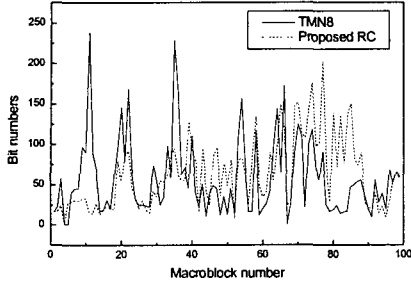
은 작아지고 K 의 값도 작아진다. 또한 식(7)을 통하여, 이전 프레임에서 구한 K_1 보다 K 값 작게 만들어, Q_i^* 의 양자화 스텝크기를 작게 한다. 다시 말해, K 매개변수의 역할은 분산이 커지는 경우, 양자화 크기를 작게 설정하여 비트를 보다 많이 할당하는 역할을 하는 것이다. 매크로 블록의 분산이 프레임의 평균 분산값과 차이가 많이 나면, 전체 프레임의 평균 양자화 스텝 크기보다 적은 값을 구한다. 분산의 차이가 적으면, 큰 양자화 스텝 크기를 구하여 비트 할당을 한다.

하지만 식(7)을 통해 알 수 있듯이 양자화 스텝 크기를 직접 구하는데 사용되는 K 값을 이전 프레임에서 구한 K 값의 평균인 K_1 으로부터 변화시켜 나갔다. 아직 부호화하지 않은 프레임의 K 값들을 구할 수 없기 때문에, 이전프레임의 평균 K 값을 K_1 이라 놓고, K_1 으로부터 현재 매크로 블록의 K 값들을 변화시켜 나갔다. 따라서 이전 프레임과 현재 프레임의 확률 분포가 많이 변하는 경우에는 현재 프레임을 부호화하고 구한 K 값들의 평균과 K_1 과 차이가 많이 나게 된다. 이점을 이용하여, 프레임간의 확률 분포 변화를 K_1 과 실제 부호화 하여 구한 K 값의 평균을 비교하여 인식하였다.

이와 같이 K 값의 변화를 통하여 확률 분포 변화를 인식한 경우, 실제 K 값들의 평균을 이용하여 현재 프레임을 다시 부호화 한다면 위에서 언급한 불균등 비트 할당 문제를 해결할 수 있다. 물론 K_1 과 부호화 후의 K 값의 평균과 차이가 별로 없는 경우에는 K_1 값을 그대로 사용하여 부호화 하면 복잡도를 줄일 수 있다.

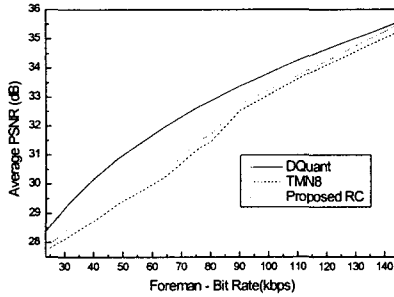
V. 실험 결과

아래 실험 결과를 통해 본 논문에서 제안한 기법이 3장에서 언급한 TMN8의 프레임의 확률 분포를 고려하지 않아서 생기는 비트의 불균등 할당 문제를 해결하였음을 확인할 수 있다. 그림 3은 Foreman 영상을 48kbps로 부호화했을 때, 69번째 프레임의 매크로 블록마다 사용한 비트 수를 비교한 것이다. 아래의 변화를 통해서 제안 기법이 실제 복잡도와 매우 근사하게 비트를 할당하는 것을 확인할 수 있다.

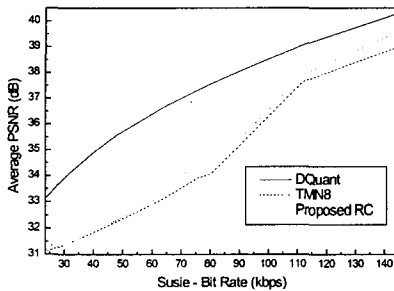


Foreman - 48kbps - 68th Frame

그림 3 매크로 블록마다 사용한 비트량 비교



(a) Foreman sequence



(b) Susie sequence

그림 4 균등 비트 할당 기법의 R-D 그래프

전체 94개의 P 프레임 중에서 약 35개, 높은 비트율일 경우에는 약 20개의 프레임에서 확률 분포의 변화를 찾을 수 있었다. Susie 실험 영상의 경우에는 총 31개의 P 프레임에서 6개의 변화를 찾았다. 평균 PSNR은 평균적으로는 0.2dB, 최대 0.7dB까지 성능이 향상됐음을 확인할 수 있다.

VI. 결론

실시간 부호화를 위한 비트율 제어 기법은 비트량과 왜곡간의 관계를 직접 계산하면 시간이 너무 오래 걸리기 때문에, 모델링을 통하여 최적의 해를 추정하고, 현재 프레임의 확률 분포를 예측할 수 없기 때문에, 이전 프레임의 확률 분포를 이용하여 QP를 구하였다. 그래서, 카메라의 이동이 많거나, 물체의 움직임이 많은 영상들과 같이 프레임간에 확률 분포가 많이 변하는 경우에는, 비트를 실제 영상 정보와는 관계없이 불균등하게 할당하는 문제점이 생겼다. 이에 본 논문에서는 확률 분포 변화를 감지하고, 변화가 인식된 경우, 현재 프레임의 확률 분포를 이용하여 부호화하였다. Foreman의 경우에는 전체 94개의 프레임중 약 30개 정도에서 변화가 인식되었고, Susie 영상에서는 30개중 6개의 프레임에서 확률분포 변화를 찾을 수 있었다. PSNR도 기존 TMN8에 비해 0.2dB에서 0.7dB 정도 증가하였다.

참고문헌

- [1] A.R. Reibman and B.G. Haskell, "Constraints on variable bit-rate video for ATM networks." EMI Trans. CSVT, vol. 2, pp. 361-372, 1992
- [2] C.Y. Hsu, A. Ortega, and A. Reibman, "Joint selection of source and channel rate for VBR transmission under ATM policing constraints." IEEE J. Select Areas Commun. vol. 15, pp. 1016-1028, Aug. 1997
- [3] K. Ramchandran, A. Ortega and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders", IEEE Trans. IP, vol. 3, no. 5, pp. 553-545, 1994
- [4] Tihao Chiang and Ya-Qin Zhang, "A new rate control scheme using quadratic rate distortion model", IEEE Trans. CSVT, vol. 7, no. 1, p. 246-250, 1997
- [5] Junfeng Bai, Qingmin Liao, Xinggang Lin, "Hybrid models of the rate-distortion characteristics for MPEG video coding", WCC-ICCT 2000, vol. 1, , pp 363-366. 2000
- [6] Ribas-Corbera, J. and Shwmin Lei, "Rate control in DCT video coding for low-delay communications", IEEE Trans. CSVT. vol. 9, no. 1, pp. 172-185, 1999