

DMB 응용을 위한 슬라이딩 윈도우 구조의 가변 비트율 제어 알고리즘

유영일, 정경훈, 강동욱, 김기두
국민대학교 전자정보통신공학부
young@kookmin.ac.kr

A Variable Bit Control Algorithm Based on the Sliding Window for DMB Application

Youngil Yoo, Kyeong Hoon Jung, Dong Wook Kang, Ki-Doo Kim
School of Electrical Engineering, Kookmin University,

요약

비트율 제어 알고리즘의 목적은 '비트율 제한조건을 만족하면서 화질을 일정하게 유지하는 것이다. 그러나 통상적인 비트율 제어 알고리즘에서는 화질을 일정하게 유지시키는 것 보다는 버퍼의 오버플로우 및 언더플로우가 발생하지 않도록 비트를 분배하는 것이 주된 목적으로서 시퀀스가 진행함에 따라 불가피하게 시각적으로 바람직하지 않은 화질의 등락이 발생한다. 본 논문에서는 비트율 제어의 본질적인 목적에 부합하도록 버퍼 제한조건을 만족시키는 범위에서 화질을 가능한 한 일정하게 유지하는 방법을 제시한다. 제안방법에서는 먼저 일반적인 CBR(constant bit rate) 알고리즘으로 부호화를 먼저 시작한 후, 일정 크기의 슬라이딩 윈도우 내에서 발생한 시퀀스를 분석하여 다음 단계에서 사용할 양자화 파라미터의 기준값을 결정한다. 마지막 단계에서는 다시 시퀀스의 처음으로 돌아가 얻어진 기준 양자화 파라미터를 사용하여 VBR(variable bit rate)의 부호화를 수행한다. 제안 방법은 실시간의 제한요소가 엄격하지 않은 응용분야에서 유용할 것으로 기대된다.

1. 서론

다양한 멀티미디어 동영상 서비스가 도입되면서 디지털 동영상 처리는 디지털 방송, 통신, 및 저장을 위한 핵심적인 기술이 되었다. 특히 MPEG-2, MPEG-4 그리고 H.264를 비롯한 부호화 표준들은 DTV(Digital Television) 및 DMB(Digital Multimedia Broadcasting) 서비스를 위한 필수적인 요소로서 활발한 연구가 이루어지고 있다. 동영상 부호화를 위한 요소기술 가운데 비트율 제어기법은 부호화기의 성능 향상을 위해 표준에서 규정하고 있지는 않지만 주어진 비트율의 제한을 만족시키면서 부호화된 동영상의 화질을 최적화하기 위해서 반드시 필요한 기술이다. 시각적으로 만족스러운 동영상의 품질을 얻기 위해서는 개개의 프레임이 최적의 화질로 부호되어야 할 뿐만 아니라 프레임 사이의 변화를 가능한 한 적게 유지하여야 flicker를 비롯한 부자연스러운 움직임을 피할 수 있다 [1].

일반적으로 부호화기의 출력 비트스트림은 CBR(Constant Bit Rate) 또는 VBR(Variable Bit Rate)로 부호화될 수 있다. 기존의 MPEG2의 TM5, MPEG4의 VM 그리고 H.264의 JM 등

다양한 비트율 제어 알고리즘이 제안되었는데 이들은 기본적으로 CBR에 근거한 방식이라고 할 수 있다. 주어진 동영상 시퀀스의 통계적 성질이나 채널의 상황을 미리 알 수 없다는 측면에서 실시간 또는 온라인 응용을 위해서는 출력 비트율을 가능한 한 일정하게 유지시키는 것이 합리적이라고 생각할 수 있다. 하지만 실제 동영상의 공간적 그리고 시간적 특성은 프레임마다 달라지는 것이 일반적이기 때문에 출력 비트율이 일정한 부호화 방법으로는 화질을 일정하게 유지시키는 것은 불가능하다. 한편 실시간이라는 제한요소가 없는 응용분야를 고려한다면에서는 비트율의 변화를 허용하면서 화질을 일정하게 유지시키는 VBR 기반의 비트율 제어가 바람직할 것이다. 예를 들어, 디지털 방송이나 DVD 등에서는 동영상 시퀀스가 반드시 실시간 일 필요가 없기 때문에 VBR 부호화를 통해 보다 안정적인 화질을 얻을 수 있다.

본 논문에서는 슬라이딩 윈도우 구조의 비트율 제어 알고리즘을 제안한다. 제안 방법에서는 기존의 통상적인 CBR 비트율 제어 알고리즘을 이용하여 동영상 시퀀스를 부호화 한 후 여기서 출력된 비트스트림과 통계값들을 분석하여 다음 단계 양자화 파라미터의 기준 값을 계산하고 부호화 단위의 공간적 그리고 시간적 특성을 살펴본다. 마지막 단계에서는 동영상의

본 논문은 2005년도 국민대학교 UCRC 연구센터 연구지원금으로 수행된 연구임.

시공간적 특성과 버퍼의 상태에 따라 양자화 파라미터를 얻는다. 이 과정은 단계별로 하나를 마친 후 다음 단계를 적용할 수 있겠지만, 부호화 시간을 줄인다는 측면에서 슬라이딩 윈도우의 개념을 사용하여 최종적인 VBR의 비트율 제어를 하기 위해 일정 프레임의 갯수만큼 선행된 CBR 알고리즘의 수행 결과만을 이용하는 것이 가능하다.

본 논문은 다음과 같이 구성된다. 우선 2절에서는 기존의 CBR 방식에 근거한 비트율 제어알고리즘에 대해 간단히 살펴본다. 3절에서는 제안하는 알고리즘을 설명하고 4절에서는 제안 알고리즘에 의한 실험결과를 보이며, 마지막으로 5절에서 결론을 맺는다.

2 CBR에 기반한 비트율 제어 알고리즘

가. 비트율 제어 및 버퍼

그림1은 전형적인 부호기의 출력 버퍼 및 복호기의 입력 버퍼를 나타낸다. 비트율 제어와 관련되어 이 버퍼들의 오버플로우 및 언더플로우를 주목해야 한다.

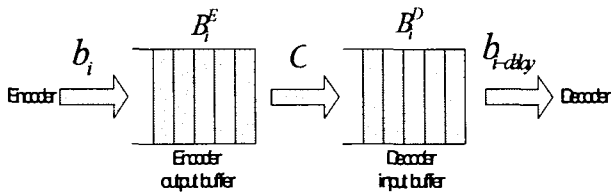


그림 1. 부호기 출력버퍼 및 복호기 입력 버퍼.

부호기의 버퍼가 i -번째 프레임에서 채워져 있는 상태를 B_i^E 로 나타낼 때, i -번째 프레임에서 발생하는 비트수를 b_i , 채널의 데이터율을 C 라고 하면 다음 관계가 성립한다.

$$B_i^E = B_{i-1}^E + b_i - C \quad (1)$$

부호기 버퍼의 크기를 B_{\max}^E 라 하면 부호기에서 오버플로우 및 언더플로우가 발생하는 것은 각각 다음의 식 (2) 및 식 (3)을 만족하는 경우이다.

$$B_i^E > B_{\max}^E \quad (2)$$

$$B_i^E < 0 \quad (3)$$

마찬가지로 복호기 버퍼의 채워져 있는 상태 및 버퍼의 크기를 각각 B_i^D 및 B_{\max}^D 로 나타내면, 복호기 버퍼의 상태와 오버플로우 및 언더플로우가 발생하는 경우를 다음의 식 (4), (5), (6)으로 생각 할 수 있다.

$$B_i^D = B_{i-1}^D + C - b_{i-delay} \quad (4)$$

$$B_i^D > B_{\max}^D \quad (5)$$

$$B_i^D < 0 \quad (6)$$

부호기에서 언더플로우가 발생하게 되면 전송할 데이터가 없어서 부호기에서는 원하지 않게 전송이 중단된다. 이의 원인으로서는 부호기에서 정해진 시간 내에 출력 비트스트림을 만들

지 못하는 경우도 생각할 수 있지만 이는 비트율 제어의 관점에서는 고려사항이 아니므로 우리가 대상으로 삼는 것은 부호기로부터 입력되는 데이터율이 채널의 데이터율보다 적은 경우이다. 부호기 발생하는 언더플로우를 피하는 통상적인 방법은 세밀한 양자화 파라미터를 사용하여 발생하는 비트수를 늘리는 것이다. 한편 부호화 버퍼의 오버플로우는 데이터의 발생율이 채널의 데이터율보다 빠른 프레임이 많이 발생하여 버퍼의 최대 크기를 넘어서는 경우에 발생한다. 이러한 경우에는 화질의 손해를 보더라도 양자화 파라미터의 값을 키워서 발생하는 데이터의 양을 줄여주어야 한다.

나. CBR 비트율 제어 알고리즘의 문제점

MPEG2의 TM5, MPEG4의 JM, H.264의 HRD 등 다양한 비트율 제어 알고리즘은 채널 또는 버퍼의 제한요소를 만족시키기 위해 사용된다 [2]-[4]. 이들 알고리즘의 과정을 다음과 같이 요약할 수 있다.

1. 각 프레임의 목표 비트율을 계산한다. 이 과정은 GOP(Group Of Picture)so 프레임의 수, 아직 부호화되지 않고 남아있는 GOP에 사용할 비트수, 버퍼의 최대크기, 그리고 각 프레임의 추정된 복잡도에 따라서 달라진다. 복잡도를 고려하는 이유는 이전 프레임이 복잡하면 다음 프레임도 복잡할 것으로 예상하여 상대적으로 많은 비트를 할당하기 위함이다.
2. 전체 프레임에 적용할 양자화 파라미터를 계산한다. 이 단계에서 출력 비트율과 양자화 파라미터 사이의 관계에 대한 모델링이 필요하다.
3. 프레임을 부호화한다.
4. 실제 발생한 비트수를 이용하여 모델링 과정에서의 파라미터를 조정한다.

이들 알고리즘을 사용하면, 버퍼의 오버플로우 및 언더플로우를 발생시키지 않고 상대적으로 일정한 율로 출력되는 비트 스트림을 얻을 수 있지만 이처럼 CBR에 근거한 방식에는 몇가지 문제점이 있다. 먼저 가장 근본적인 문제는 화질을 일정하게 유지시킨다는 관점에서 만족스럽지 못하다는 것이다. CBR 비트율 제어 알고리즘에서는 출력 화질이 우선적인 고려사항이 아니기 때문에 시퀀스가 시간적으로 진행하면서 시각적으로 거슬리는 현상이 발생할 가능성이 있다.

또 다른 문제점은 초기 양자화 파라미터의 영향에 관한 것이다. 시퀀스를 시작할 때는 그 시간적, 공간적 특성의 전혀 알 수 없기 때문에 양자화 파라미터의 초기값을 결정하는 것은 쉬운 작업이 아니다. 그림 2에서는 foreman 시퀀스를 부호화함에 있어서 양자화 파라미터의 영향을 타내내었다. 그림 2를 통해서 초기값이 잘못 선택 되었을 때 상당수의 프레임을 부호화하기 까지 그 영향이 지속됨을 볼 수 있다. 또한 양자화 파라미터의 초기값이 너무 작으면(QP=20) 안정화 이후의 PSNR값이 상대적으로 낮으며, 반대로 너무 크면 (QP=32) 이 값이 상대적으로 높은 것을 확인할 수 있다. 즉 부호화된 시퀀스의 화질 유지를 위해서 양자화 파라미터의 초기값은 신중하게 선택될 필요가 있다.

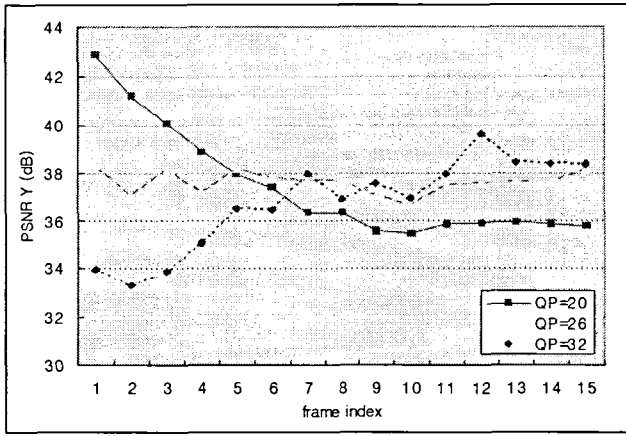


그림 2. 양자화 파라미터 초기값 선택의 영향.

다음으로 생각할 수 있는 문제점은 CBR의 경우에 양자화 파라미터가 프레임 단위로 요동하는 현상이 발생할 수 있다는 것이다. 다음의 그림3에 보인 바와 같이 연속되는 프레임 사이에서 양자화 파라미터의 값이 주기적으로 오르내리는 것을 확인할 수 있다. 이는 양자화 파라미터의 요동은 다시 화질의 요동으로 이어져 결국 시각적으로 거슬리는 결과를 나타내게 된다.

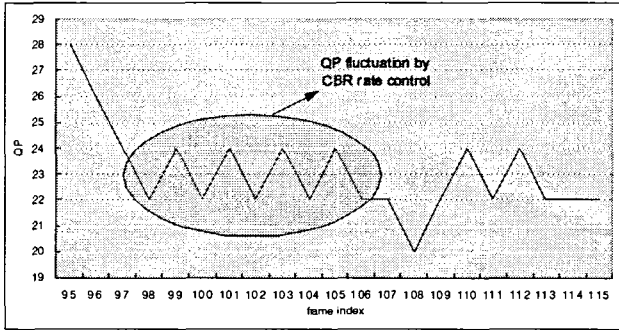


Fig. 3. CBR 알고리즘에서 프레임간 QP의 요동 현상.

3. 제안 비트율 제어 알고리즘

제안하는 비트율 제어 알고리즘의 개념도를 그림 4에서는 나타내었다. 제안하는 부호기는 일단 CBR 비트율 제어 알고리즘을 통해 얻어진 결과를 분석하여 VBR 제어를 하는 개념이다. 이 부호기에서 CBR의 수행결과를 분석하는 단계 이후를 생략하면 실시간 응용을 위한 통상적인 부호기와 동일하다.

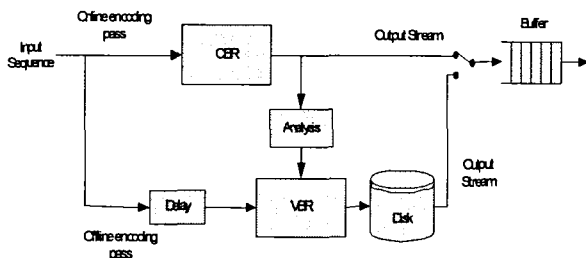


그림 4. 제안 알고리즘.

먼저 CBR의 비트율 제어 단계에서 양자화 파라미터는 채널의 데이터율을 만족하도록 조정된다. 채널의 대역폭을 최대한으로 활용하기 위해서는 비트율을 가능한 한 크게 하는 편이 유리하다. 실험에서는 JM 알고리즘을 사용하였다 [5]. 이 방식은 비트율 제어의 기본 단위로 프레임, 슬라이스, 또는 매크로블록을 사용하며, 현재 프레임의 현재 기본 단위의 MAD(Mean Absolute Difference)를 예측하기 위해 이전 프레임의 동일 위치에서의 해당값을 사용하는 선형 모델을 사용한다. 이 알고리즘을 사용하여 버퍼의 오버플로우 및 언더플로우가 발생함 없이 부호화된 비트스트림을 얻을 수 있다.

두 개의 부호화 단계 사이에는 분석 단계가 있다. 이 단계에서는 우선 VBR 단계에서 사용될 양자화 파라미터의 기준값을 결정한다. 기준값으로는 슬라이딩 윈도우내의 프레임들의 양자화 파라미터의 이동 평균을 사용한다. 이를 다음의 식 (7)에 나타내었다. 여기서 N 은 슬라이딩 윈도우의 크기이고 α_k 는 가중치를 의미한다.

$$QP_2(i) = \frac{1}{N} \sum_{k=i}^{i+N} \alpha_k \cdot QP_1(k) \quad (7)$$

식 (7)을 통하여 양자화 파라미터 값의 변화 정도가 줄어들기 때문에 초기값 설정에 따른 문제를 피할 수 있을 뿐 만 아니라 요동 현상이 발생하는 정도를 줄일 수 있다.

한편 이 단계에서 인간의 시각 특성에 따른 비트율 제어를 위해서 시퀀스 부호화 단위 마다 공간적인 활동도 및 시간적인 복잡도도 함께 추정한다. 이 특성들은 두 번째 VBR의 부호화 단계에서 양자화 파라미터 변화에 따른 부호화 단위의 민감한 정도에 따라 세밀한 조정을 위해 필요하다. 실험에서는 공간적인 활동도를 추정하기 위해 저주파 성분과 고주파 성분의 비율을 사용하였으며 시간적 활동도를 추정하기 위해 해당 부호화 단위의 MV(motion vector)의 합을 사용하였다.

마지막 단계에서는 분석단계에서 얻어진 기준 QP값을 바탕으로 버퍼의 상태를 관찰하면서 버퍼의 오버플로우 및 언더플로우가 발생하지 않는 범위내에서 최적의 화질을 얻도록 부호화를 수행한다. 실험에서 양자화 파라미터를 조정하는 가이드라인은 다음과 같다.

1. 분석단계에서 얻어진 기준 양자화 파라미터의 값의 변화는 시각적 화질을 가능한 한 일정하게 유지하기 위해 ± 1 로 제한한다.
2. 현재 버퍼의 채워져있는 상태를 B_i^D 라 할 때, i -번째 프레임의 양자화 파라미터의 차이값 Δ_i 는 다음과 같다.

$$\Delta_i = \begin{cases} +1, & B_i^D > T_1 \\ 0, & T_1 \leq B_i^D < T_2 \\ -1, & B_i^D \leq T_2 \end{cases} \quad (8)$$

식 (8)에서의 임계값 T_1 과 T_2 는 기본 부호화 단위의 공간적 시간적 활동도에 따라서 적응적으로 결정된다. 즉 일반적으로 말하면 부호화 단위가 복잡할수록 공간적 그리고 시간적 활동도가 높은 값을 가지므로 임계값도 높게 설정하는 것이 바람직하다.

4. 실험결과

제안하는 알고리즘을 DMB 환경에 적용하기 위해 H.264/AVC baseline profile과 JM9.4 [5]를 사용하였으며, foreman(CIF), CF(QVGA), Akiyo(CIF) 시퀀스에 대해 실험하였다. 표 1에서 실험에서 사용한 파라미터들을 정리하여 나타내었다.

표 1. 부호화 파라미터.

Items	Contents
Profile	H.264/AVC Baseline
Intra refresh interval	15 frames
Channel Bitrate	384 Kbps
Initial QP for CBR	32

그림 5에서 그림 7까지는 실험에서 사용한 각 시퀀스에 대해서 프레임 레벨에서의 PSNR을 나타내었다. 그림들을 통해 CBR을 사용한 경우의 PSNR의 변화가 VBR 단계를 거치면서 상당히 줄어든 것을 확인할 수 있고, 전반적인 PSNR도 향상되었음을 알 수 있다. 표 2에서는 각 시퀀스마다 PSNR의 평균값과 분산값을 정리하였다. 이를 통해 PSNR의 분산, 즉 변화 정도가 상당히 줄어든다는 사실과 더불어 평균적으로 0.75 dB의 부호화 이득이 발생하는 결과를 확인할 수 있다.

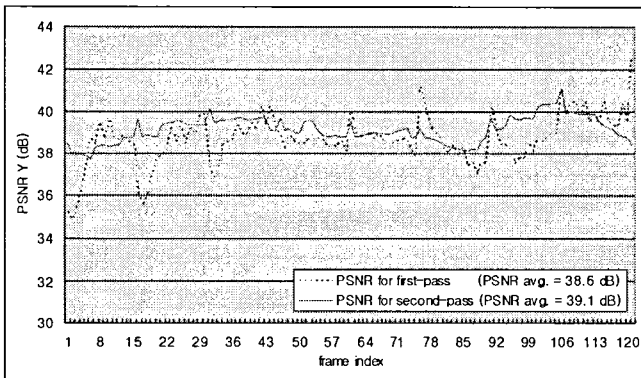


그림 5. PSNR 성능 (foreman 시퀀스의 경우).

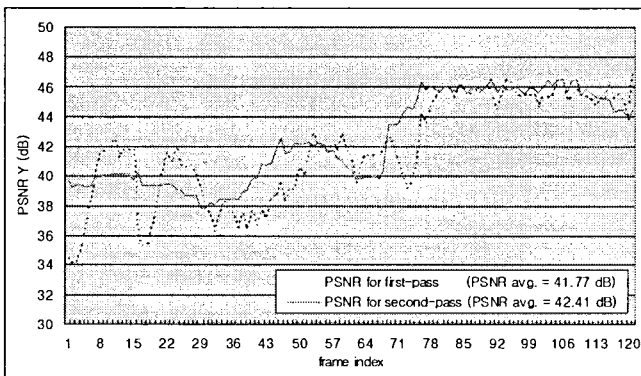


그림 6. PSNR 성능 (CF 시퀀스의 경우).

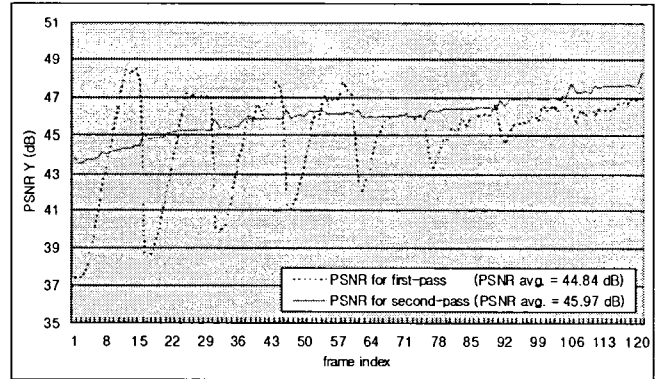


그림 7. PSNR 성능 (Akiyo 시퀀스의 경우).

표 2. CBR 및 VBR 알고리즘 성능 비교.

Sequence	Mean (dB)		Variance	
	First-pass	Second-pass	First-pass	Second-pass
Foreman (CIF)	38.64	39.12	1.41	0.39
CF (QVGA)	41.77	42.41	11.90	8.54
Akiyo (CIF)	44.84	45.97	6.94	1.09

다음의 그림 8에서는 주관적인 화질을 비교하기 위해 CF 영상의 299번째 프레임을 나타내었다. 그림에서 보듯이 특수한 상황에서 CBR알고리즘은 심각한 화질 저하를 가져올 수 있으나 제안 방법에서는 화질 저하를 주변 프레임과 나누어 감당하기 때문에 시각적으로 유리한 방법임을 볼 수 있다.



(a)



(b)



(c)

그림 8. CF 영상의 299번째 프레임. (a) 원영상, (b) CBR 알고리즘의 결과영상(PSNR = 25.77 dB), (c) 제안 알고리즘의 결과영상(PSNR = 33.75 dB).

5. 결론적 제언

본 논문에서는 VBR 비트율 제어 알고리즘을 제안하고 모의실험을 통해 그 성능을 살펴보았다. 제안 방법은 CBR 기반의 비트율 제어 알고리즘에서 일반적으로 발생하는 시각적으로 거슬리는 화질의 변화를 상당히 줄일 수 있을 뿐 만 아니라 초기 양자화 파라미터의 선택 문제 및 프레임간 양자화 파라미터의 요동 현상도 피할 수 있는 장점이 있다. 부호화 성능의 측면

에서도 실험을 통해 평균적으로 0.75dB의 이득이 있음을 보였다. 한편 제안방법에서는 CBR의 결과를 분석하고 VBR의 비트스트림을 얻기 위해 슬라이딩 윈도우의 개념을 사용함으로써 완전한 실시간은 아니지만 단지 슬라이딩 윈도우의 크기 만큼의 지연만이 발생한다. 따라서 제안하는 방법은 실시간의 제한요건이 중요하지 않은 방송이나 DVD 등의 응용분야에서 활용할 수 있을 것으로 기대된다.

References

1. Z. He, C. Chen, and J. Cai., "Lowpass filtering of rate-distortion functions for quality smoothing for real-time video recording and streaming" Proceedings of ICIP 2004, vol.2, pp.1101 - 1104. Oct. 2004.
2. Z. Li, W. Gao, and F. Pan, "Adaptive rate control with HRD consideration", JVT-H014. doc, Joint Video Team(JVT), 8-th meeting, Geneva, May, 2003.
3. ISO/IEC 14496-2, Coding of audio-visual objects - Part 2: Visual, 2001, Annex L.
4. Iain E.G. Richardson, H.264 and MPEG-4 Video Compression, John Wiley & Sons, 2003
5. JVT Reference Software version 9.4, <http://bs.hhi.de/~suehring/ttml/download>