

영상의 복잡도를 이용한 장면 전환 추정과 H.264/AVC 비트율 제어

김종구, 서재원
충북대학교

jgKim@cbnu.ac.kr, sjwon@cbnu.ac.kr

H.264/AVC Rate Control for Scene Change with Complexity of Sequence

Kim, Jong Gu, Suh, Jae Won
Chungbuk National University

요약

본 논문에서는 프레임 안에서의 인접픽셀간의 복잡도를 이용하여 할당비트 조절과 장면 전환을 추정하는 H.264/AVC 비트율 제어 기법을 제안한다. Quadratic 율-왜곡 모델은 프레임 할당 비트와 선형 모델을 통해 예측된 MAD정보를 필요로 한다. 이때 장면전환과 같은 급격한 화질 변화는 예측된 MAD정보만으로는 적합한 양자화파라미터를 결정하지 못하고 이후 비트율 제어를 어렵게 한다. 제안하는 방법은 프레임 안에서 인접픽셀간의 차분 값들의 절대 합을 구하여 이전프레임과 현재프레임의 비로 나타낸 후 이 값의 일정 범위를 벗어나면 장면전환이 이루어진 것으로 간주하여 이에 알맞은 양자화 파라미터 결정기법을 적용하도록 하였다.

1. 서론

미디어기술이 발전함에 따라 문화 사회 전반에서 그 기술을 필요로 하는 분야가 늘어났다. 미디어는 소리, 문자, 영상 등 다양한 방법으로 정보를 전달하고 있으며, 특히 전달효과가 가장 강한 영상을 통한 미디어 전송방식이 각광을 받게 된다. 이때 영상이 가지고 있는 정보는 그 양이 너무 크기 때문에 저장매체나 실시간 전송을 통해 전송할 때 효율적이지 못하다. 그래서 압축을 통해 실시간 전송이나 정보의 저장을 효율적으로 하기위해 동영상 압축 기술에 대한 표준안을 채택하여 영상을 효율적으로 전송하고 재생하는 기술인 디지털 비디오 코덱 표준을 만들었다.

디지털 비디오 코덱 표준의 한 종류인 H.264/AVC는 ITU-T(Video Coding Experts Group)과 ISO/IEC(Moving Picture Experts Group)이 공동으로 연구한 표준안이다. H.264/AVC에 사용된 압축 기술로는 시·공간적으로 중복된 정보를 제거하기 위해 복수 참조 영상을 이용한 가변 블록 크기의 움직임 추정, 16×16 및 4×4 크기의 화소를 이용한 공간적 예측 모드 등을 사용하여 좀 더 정밀한 예측을 가능케 하였다. 더욱이 비트율-왜곡 최적화 기법(RDO, Rate Distortion Optimization)을 기반으로 부호화 모드의 결정을 최적화 하여 예측 부호화의 정확도를 높여 부호화 효율을 크게 향상 시켰다. 또 DCT(Discrete Cosine

Transform)를 통해 정보를 변환 압축하며 최종적으로 가변길이 부호화 방식을 이용하여 압축정보를 전송하는 기술이 사용되고 있다.

또한 이러한 압축 기술을 이용해 실시간 영상 전송 효율을 높이게 되었다. 기존보다 높은 부호화 효율로 인해 보다 좋은 화질의 영상 전송을 가능하게 하였다. 물론 채널 대역이 충분히 커서 압축하지 않은 본래의 영상정보를 그대로 전송할 수 있다면 이러한 기술이 그다지 효율적이지 못하겠지만 현재로서는 전송 채널의 제한으로 인해 충분히 큰 정보를 한 번에 전송하는 것은 불가능하다. 따라서 높은 부호화 효율을 가진 압축 기술로 정보를 압축하여 전송하면 보다 좋은 화질의 정보를 실시간으로 받아 볼 수 있다.

하지만 보통 채널의 용량은 일정한 크기의 정보를 전송할 수 있게 되어 있으나 영상 정보는 한 화면마다 각기 다른 정보량을 가지게 된다. 여기서 우리는 일정한 채널 용량에 맞춰서 각각의 화면 정보를 전송할 필요가 있다. 이것이 비트율 제어(Rate Control)알고리즘이 필요한 이유이다. 비트율 제어는 기본적으로 채널 용량이 허용하는 비트 량과 영상이 가지고 있는 비트량과 화질을 고려하여 최적의 부호화 성능을 가지도록 하는데 목적을 둔다. 영상의 화질과 압축률을 조절하는 가장 효율적인 방법은 양자화 파라미터를 이용하는 것이다. 이것을 조절하여 영상의 비트량을 목표 비트에 가깝게 맞추면서 화질을 개선하는 것이

비트율 제어의 전체 흐름이다.

2. H.264/AVC 비트율 제어

비트율 제어는 양자화과라미터 결정→RDO→MAD산출→부호화 단계를 거치며 실제 발생 비트량과 목표 비트량의 간격을 좁히는 것이 목표이다. 여기서 양자화 파라미터는 비트율 제어를 통해서만 결정이 되는 것을 알 수 있다. 또한 MAD값은 양자화 파라미터가 결정되어져야만 RDO과정을 통해 산출되어진다. 하지만 앞에서 언급하였듯이 비트율 제어를 위해서는 영상의 복잡도 정보를 알아야만 한다. 이때 영상의 복잡도 정보인 MAD는 양자화 파라미터가 결정된 이후에야 구해질 수 있다. 이것이 비트율 제어의 딜레마이다.

가. 프레임단위 비트할당

프레임에 할당된 비트를 구하기 위해서는 두 가지가 고려되어진다. 전체프레임에 대한 한 프레임의 할당비트와 버퍼 점유율을 고려한 버퍼 할당을 가중치를 두어 합하여 프레임 할당비트가 결정된다.

$$\hat{T}_f(n_i) = \beta T'(n_i) + (1 - \beta) T''(n_i) \quad (2.3)$$

n_i 프레임 넘버를 의미하여, β 는 가중치 값으로 보통 0.5의 값을 갖는다. T', T'' 은 각각 가용비트를 이용한 프레임 할당비트와 버퍼 점유율 고려한 할당 비트를 의미하며 식(2.3)과 식(2.4)로 표현된다. 여기서 가용비트는 전체 프레임에 할당된 총 비트를 구한 다음 부호화로 사용된 비트를 빼줌으로써 사용가능한 비트를 나타낸다.

$$T'(n_i) = \frac{T_r(n_{i-1}) - b(n_{i-1})}{N_{p,r}} \quad (2.4)$$

이때 $T_r(n_i)$ 가 가용비트가 되며 $b(n_i)$ 는 부호화시 생성비트를 뜻한다. $N_{p,r}$ 은 아직 부호화하지 않은 P-프레임의 수를 의미한다.

버퍼는 실제 부호화 후 전송되기 전에 부족분 혹은 넘치는 비트를 저장하기 위한 공간으로 사용된다. I-프레임을 부호화 할 경우 P-프레임의 경우 보다 많은 비트가 생성되기 때문에 대부분은 I-프레임을 부호화하고 생성된 비트는 정해진 비트 율을 넘게 된다. 이렇게 넘겨진 비트들은 버퍼에 저장 되고 다음 프레임들을 부호화하면서 버퍼를 비워나가도록 설계된다.

$$T''(n_i) = \frac{u(n_i)}{F_r} + \gamma \times (tbl(n_i) - B_c(n_{i-1})) \quad (2.5)$$

$u(n_i), F_r$ 은 각각 비트율과 프레임 율을 의미하고, $B_c(n_i)$ 은 실제 버퍼 점유율을 뜻하며, $tbl(n_i)$ 는 목표 버퍼 레벨을 나타낸다. 목표 버퍼 레벨은 버퍼에 저장된 비트를 남아있는 P프레임 수로 나누어 그 값을 구해준다. 즉, 버퍼에 저장된 정보를 다음 전송을 통해 보내면서 점차적으로 버퍼를 비워주는 개념으로 이해하는 것이 좋겠다.

나. MAD예측을 위한 선형 모델

앞서 언급한 바와 같이 현재의 정확한 MAD정보를 산출하는 것은 불가능하다. 따라서 현재 매크로블록의 MAD를 이전 프레임과 동일한

위치에서의 MAD값을 이용해 현재 MAD를 예측하는 선형 모델이 사용된다.

$$MAD_{cb} = a_1 \times MAD_{pb} + a_2 \quad (2.2)$$

a_1, a_2 는 선형 모델의 모델 파라미터로 각 매크로블록을 부호화 해나감에 따라 갱신되며, 초기 값은 각각 1과 0이다. MAD_{cb}, MAD_{pb} 는 각각 현재 매크로블록과 이전 매크로블록을 의미한다.

다. 이차원 비트율-왜곡(Quadratic Rate-Distortion) 모델

H.264/AVC에서는 이차원 비트 율-왜곡 모델을 통해 양자화 파라미터를 조절한다. DCT계수 분포가 라플라스 분포(Laplace Distribution)를 가진다고 가정을 통해 Distortion과 Rate의 관계를 이용하여 양자화 파라미터와 목표 비트와의 관계를 다음과 같이 정의한다.[6]

$$R = MAD \times \left(\frac{X_1}{Qstep} + \frac{X_2}{Qstep^2} \right) \quad (2.1)$$

여기서 R은 부호화에 필요한 목표 비트 즉, $\hat{T}_f(n_i) - m_{hdr}(n_{i-1})$ 으로 구해진다. 앞에서 구한 프레임할당비트에서 헤더비트를 뺀 값이 실제 목표 비트가 된다. X_1, X_2 는 각각 이차원 비트 율-왜곡 모델의 1차, 2차 계수 값이다.

라. 양자화 파라미터 결정

기본적인 양자화 파라미터는 이차원 모델을 통해 결정되지만 이차원 모델은 확률을 기반으로 하기 때문에 오차를 동반하고 있다. 이러한 오차를 수정하기 위해 H.264/AVC에서는 몇 가지 사항을 추가하여 최종적인 양자화파라미터를 산출한다.

여기서 이루어지는 양자화 과정은 크게 두 가지 방법이 있다. 먼저, 프레임 할당비트가 0보다 작을 경우이다. 이 경우는 특별한 알고리즘 없이 양자화 수치를 크게 하여 목표 비트에 접근하는 방식을 사용하고 있다. 다른 한 가지는 프레임 할당비트가 0이상인 경우로 이 경우는 앞서 설명된 이차원 비트율-왜곡 모델을 통해 양자화 파라미터를 결정하게 된다. 이때 양자화 파라미터 변화를 통해 지나친 화질 변화를 막고자 최대 ± 4 로 양자화 파라미터 변화를 제한한다.

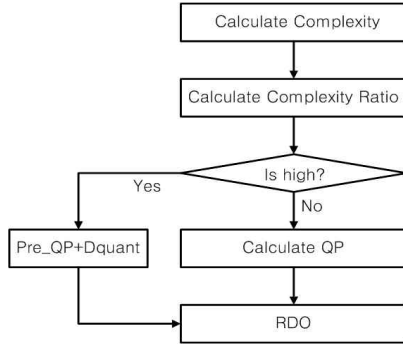
3. 장면전환 검출

H.264/AVC 비트율 제어 과정에서 사용된 모델들은 선형 혹은 확률을 기반으로 적용하고 있기 때문에 장면 전환과 같은 급격한 변화를 갖는 장면을 제어함에 있어 효율이 떨어지게 된다. 여기서는 장면전환이 이루어짐을 검출하고 이때 효율적인 비트율 제어가 이루어지도록 하는 알고리즘을 제안하겠다.

가. 장면전환 검출 알고리즘

장면전환을 검출하기 위해서는 현재 영상과 이전영상의 차이점을 나타낼 수 있는 복잡도 계산이 필요하다. 이때 장면이 가진 복잡도가 이

전과 일정범위 이상의 차이를 보이게 되면 장면전환이 이루어진 것으로 간주될 수 있다. 다음은 장면전환 검출 과정에 대한 블록도이다.



<그림 3.1> 장면전환 검출 블록도

이때 화면내의 복잡도를 계산하기 위해 인근 픽셀간의 차분치의 절대값을 이용한다. 영상은 인근 픽셀의 상관도를 이용해 많은 정보를 알 수 있다. 픽셀간의 화질 변화가 크면 화면은 많은 정보를 담고 있는 것으로 간주되어 복잡한 화면으로 구성되었음을 알 수 있고, 부호화된 코드 길이 또한 길어지는 것은 자명한 일이다. 또 영상의 인근 픽셀이 유사한 크기를 가지고 있다면 중복 정보가 많아져서 부호화 코드도 적어져 압축율이 증가된다. 이렇게 인근 픽셀 간의 정보를 이용하여 화면의 복잡도를 얻을 수 있다.

$$G_{MB}(i, j, k) = \sum_{n=0}^7 \sum_{m=0}^7 (|I_{x,y,k} - I_{x+1,y,k}| + |I_{x,y,k} - I_{x,y+1,k}|), \quad (3.1)$$

$$(x = 16i + 2m, y = 16j + 2n, 0 \leq i < \frac{M}{16}, 0 \leq j < \frac{N}{16})$$

여기서 i, j 는 프레임 안에서 매크로블록의 가로, 세로 위치를 뜻하며, k 는 부호화되는 프레임 넘버를 의미한다. n, m 은 매크로블록 안에서 픽셀의 가로, 세로 위치를 뜻한다. $I_{x,y,k}$ 는 픽셀의 크기를 의미한다. 이때 각각의 픽셀은 0부터 255의 크기를 가지고 있다. 식(3.1)에서와 같이 기준 픽셀을 중심으로 좌측 픽셀과의 차분치의 절대값, 아래쪽 픽셀과의 차분치의 절대값들을 더한 후 좌측으로 두 번째 픽셀을 기준으로 다시 같은 과정을 반복하여 매크로블록의 인접픽셀의 차분치의 절대치의 합을 모두 더해 각 매크로블록의 복잡도를 구해준다. 다음으로 영상의 가로 열에 대한 각 매크로블록의 복잡도를 더하여 $G_{MB,R}(j, k)$ 을 구하고, 끝으로 모든 매크로블록의 복잡도를 합하여 $G_F(k)$ 을 구해준다.

$$G_{MB,R} = \sum_{i=0}^{M/16-1} G_{MB}(i, j, k) \quad (3.2)$$

$$G_F(k) = \sum_{j=0}^{N/16-1} G_{MB,R}(i, k) \quad (3.3)$$

이렇게 구해진 각 프레임의 복잡도를 이전 프레임과 현재 프레임의 비를 적용하여 장면전환 여부를 판별한다.

$$G_{F,ratio}(k) = \frac{G_F(k)}{G_F(k-1)} \quad (3.4)$$

이때 장면전환이 비슷한 복잡도를 가진 화면으로 이루어지는 것을 검출하기 위한 또 다른 비례식이 필요하다.

$$G_{MB,R,diff}(k) = \frac{1}{M \times N} \sum_{j=0}^{N/16-1} |G_{MB,R}(j, k) - G_{MB,R}(j, k-1)| \quad (3.5)$$

식(3.5)은 앞에서 구한 $G_{MB,R}(j, k)$ 을 이용하여 구해진다. 이전 프레임과 현재 프레임의 차분치를 절대 가합한 후 가중치로 나누어 구하게 된다. 이때 장면전환을 만족하는 조건은 다음과 같다.

$$G_{F,ratio}(k) > 1.2 \text{ or } [G_{F,ratio}(k) < 0.97 \text{ and } G_{MB,R,diff}(k) > 1] \quad (3.6)$$

나. 목표 비트 할당

$G_{F,ratio}(k)$ 는 현재 영상과 이전 영상의 복잡도를 계산하여 비율로 나타낸 것으로 그 값이 커지면 현재 영상이 이전 영상보다 좀 더 복잡한 화면으로 구성되고 부호화 시 더 많은 비트가 소모될 것으로 생각해 볼 수 있다. 따라서 가용비트에 복잡도 비를 곱한 값을 이용해 좀 더 정확한 프레임할당 비트를 구하도록 설정한다.

$$\overline{T}_f(k) = G_{F,ratio}(k) \times \hat{T}_f(k) \quad (3.7)$$

장면전환이 이루어질 경우 생성될 비트는 두 가지로 나타난다. 장면전환이 이루어졌다고는 하나 이전프레임보다 더 낮은 복잡도를 가진 화면으로 장면전환이 이루어지면 생성될 비트는 큰 변화가 없게 나타난다. 하지만 이전 프레임보다 높은 복잡도를 가진 화면으로 장면전환이 이루어지면 프레임할당비트를 통해 구해진 $\hat{T}_f(k)$ 보다 많은 비트를 필요로 하게 된다. 따라서 프레임 할당 비트는 후자의 경우만 적용하여 다음과 같은 비트를 할당하였다.

$$\overline{T}_f(k) = \alpha \times \hat{T}_f(k) \quad (3.8)$$

여기서 α 는 실험치 값으로 2가 사용되었다.

다. 양자화 파라미터 결정

양수의 목표 비트가 할당되었을 때에는 먼저 이차원 모델을 통해 QP를 구해준다. 이때 이전 프레임의 목표 비트와 실제 생성 비트의 비를 통해 QP 값을 조절한다.

$$\text{if } \left(\frac{\hat{T}_f(k)}{b(k)} < 1.0 \text{ and } G_{F,ratio}(k) > a_1 \right) \quad (3.8)$$

$$\text{then } (QP(k) = \overline{QP}(k) + 1)$$

$$\text{if } \left(\frac{\hat{T}_f(k)}{b(k)} > 2.0 \text{ and } G_{F,ratio}(k) < a_2 \right) \quad (3.9)$$

$$\text{then } (QP(k) = \overline{QP}(k) - 1)$$

여기서 $\overline{QP}(k)$ 는 이차원 R-D모델을 통해 구해진 양자화 파라미터를 의미한다. a_1, a_2 는 실험치 값으로 0.996과 1.004로 설정되었다.

다음은 장면전환 시에 양자화 파라미터 결정 기법으로 앞서 언급한 바로 같이 복잡도가 증가하면서 장면전환이 이루어지는 경우에만 적용된다.

$$\text{if } G_{F,ratio}(k) > 1.2 \quad \text{then } QP(k) = QP(k-1) + \alpha_1 \quad (3.10)$$

α_1 은 실험치 값으로 8이고 이것은 장면전환 시 이후에 비트율 제어를 효과적으로 하기 위하여 강제적으로 큰 QP로 부호화하여 생성비트를 줄이도록 설계하였다. 또 장면전환이 이루어진 바로 다음프레임도 강제적 QP를 적용할 필요가 있는데 이는 이차원 R-D모델에서 사용될 MAD가 선형적으로 이뤄지도록 되어 있기에 이차원 R-D모델을 통해 효율적인 QP를 적용시키기 어렵기 때문이다.

$$\text{if } G_{F,ratio}(k-1) > 1.2 \quad \text{then } QP(k) = QP(k-1) + \alpha_2 \quad (3.11)$$

4. 실험 결과

본 논문에서는 제안한 비트율 제어 방법의 성능을 검증하기 위해서 H.264/AVC 표준 소프트웨어 JM-14.0 환경에서 실험을 수행하였다. 실험 조건은 <표 4.1>과 같다.

GOP structure	IPPP
RDO	On
Size	QCIF 4:2:0
Frame Rate	15fps
Initial QP	35
basic Unit	1 (frame layer)
Symbol Mode	CABAC

<표 4.1> 실험 조건

장면전환은 임의의 영상을 2가지로 구성되며 50번째 프레임에서 이루어지도록 설정하였다. 각각의 표는 4가지 영상에 대한 PSNR 평균과 비트 율을 나타내고 있다. 실험에서 적용한 비트율은 19.2kbps와 48kbps이다.

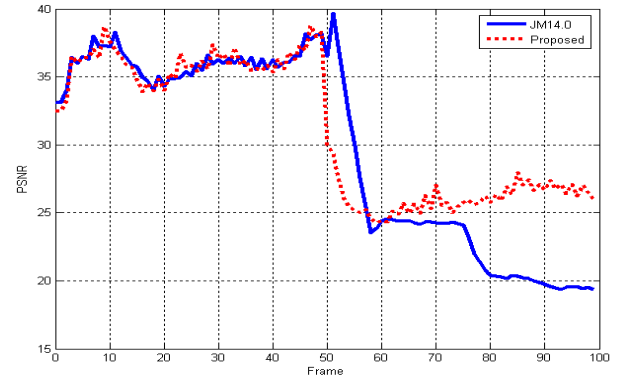
영상 시퀀스	비트율	JM14(dB)	제안(dB)	이득
(1)silent-stefan	19.2k	27.33	27.84	+0.51
	48k	29.97	31.87	+0.90
(2)silent-carphone	19.2k	32.39	32.32	-0.08
	48k	36.31	36.52	+0.21
(3)stefan-foreman	19.2k	26.62	26.39	-0.23
	48k	31.13	31.26	+0.13
(4)mother&dough-bus	19.2k	27.03	27.54	+0.51
	48k	29.85	31.04	+0.19

<표 4.2> 임의의 장면전환을 설정한 영상을 통해서 일정한 비트율 제어를 수행 시 제안하는 방법과 JM14.0의 평균 PSNR 비교

영상 시퀀스	비트율	JM14(bps)	제안(bps)	이득
(1)silent-stefan	19.2k	20.32k	19.47k	-0.85
	48k	49.15k	47.90k	-1.25
(2)silent-carphone	19.2k	19.45k	19.22k	-0.23
	48k	47.95k	47.91k	-0.04
(3)stefan-foreman	19.2k	20.21k	19.22k	-0.99
	48k	48.01k	48.21k	+0.20
(4)mother&dough-bus	19.2k	21.39k	19.80k	-0.59
	48k	50.14k	48.01k	-2.13

<표 4.3> 임의의 장면전환을 설정한 영상을 통해서 일정한 비트율 제어를 수행 시 제안하는 방법과 JM14.0의 실제 발생 비트율 비교 위 실험을 통해 (1),(4)영상은 복잡도가 증가하는 방향으로 장면전환이 이루어졌음을 확인할 수 있다. (2),(3)영상의 경우는 48k의 비트 율

로 부호화시 목표 비트에 근접한 결과를 얻을 수 있었으나 (1),(4)영상은 목표비트에 근접하지 못하는 비트율로 부호화 되었다. 복잡도가 증가하는 방향으로 장면전환이 이루어진 경우 이차원 R-D모델을 적용하여 양자화 파라미터를 결정하면 장면전환 이후 프레임의 비트율 제어가 제대로 이루어지지 않는다. 다음 그림은 (4)mother&doughther-bus 영상을 부호화하고 난 후 각 프레임에서의 PSNR 변화를 나타내었다.



<그림 4.1> (4)영상에서 각 프레임의 PSNR 변화

그림은 장면전환 이후 강제적인 QP 변화를 통해 이후 프레임의 비트율 제어 효율을 높이는 것을 보여준다.

5. 결론

실험에서 보여주듯 장면전환 등과 같은 급격한 화질 변화에 대한 즉각적인 QP 변화가 실현되지 않았을 경우 이후 프레임들의 비트율 제어를 어렵게 만들고 결국 목표 비트에 도달하는 것이 어렵게 된다. 따라서 이러한 예외적인 상황에 대한 비트율 제어 알고리즘이 필요하며 본 논문은 실험에서 나타난 바와 같이 즉각적인 QP조절을 통해 장면전환 이후의 비트 고갈 문제를 해결할 수 있다. 결과적으로 본 논문의 알고리즘이 급격한 화질 변화에 능동적으로 대처할 수 있도록 설계되었다는 것을 확인할 수 있다.

참 조

- [1] Tian Lan, Xuemai Gu, "H.264 Frame Layer Rate Control Based on Block Histogram Difference," ICC 2008 workshop proceeding
- [2] Xuan jing, Lap-pui Chau, "A Novel Intra-Rate Estimation Method for H.264 Rate Control", ISCAS 2006
- [3] Yun-Gu Lee, Byung Cheol Song, "AN INTRA-FRAME RATE CONTROL ALGORITHM FOR ULTRA LOW DELAY H.264/AVC CODING", ICASSP 2008
- [4] JM reference software version 14.0, <http://iphome.hhi.de/suehring/tml/>
- [5] Z.G. Li, W. Gao, F.Pan, S. Ma, K.P. Lim(2003), G. Feng, X. Lin, S. Rahardja, H. Lu and Y.Lu, "Adaptive Rate Control with HRD Consideration," JVT-H014, 8th Meeting, Geneva.
- [6] T. Chiang and Y. Zhang(1997), "A new rate control scheme using quadratic rate distortion model," IEEE Trans. on Circuit and Systems for Video Tecchnology, Vol.7, No.1, pp.246-250.