

MPEG-4 부호화기에서의 인트라 프레임 양자화 레벨 선정

김정우^{*}, 조성환^{**}

요 약

본 논문에서는 MPEG-4 부호화기에서 인트라 프레임의 양자화 레벨을 구하는 방법을 제안한다. 인트라 프레임은 GOP 내에서 참조 프레임으로써 다른 프레임에 계속 전파되므로 이 프레임의 성능에 의해 전체 GOP의 성능이 크게 영향을 받을 만큼 중요한 부분이다. 본 논문에서는 인트라 프레임의 양자화 레벨을 효과적으로 구할 수 있도록 할당된 비트 수와 입력되는 영상의 복잡도, 그리고 GOP의 구조를 이용하는 방법을 제안한다. 실험 결과에서는 기존의 방법들이 고정된 값을 사용하여 효율이 저하되거나 영상의 특성에 따라 상이한 성능을 보이는데 비하여 본 논문의 방법은 다양한 영상에 대하여 안정된 결과를 보였다. 제안한 방법을 MPEG-4 검증모델에서 선택한 Q2 알고리즘과 비교할 때 영상의 특성에 따라 조금씩의 차이는 있지만 높은 비트율에서 최대 3.49dB 정도의 높은 이득을 얻을 수 있었다.

Quantization Level Selection of Intra-Frame for MPEG-4 Video Encoder

Jeong-Woo Kim^{*}, Seong-Hwan Cho^{**}

ABSTRACT

This paper presents the method of calculating the quantization level of the intra-frame in MPEG-4 video encoder. The intra-frame is an essential part in that the quality of the whole GOP is affected by the quality of this frame since the intra-frame, which works as a reference frame within GOP, continuously propagates through other frames. This work proposes how to use bits assigned for gaining the quantization level of the intra-frame, complexity of input images, and GOP structures. The result shows that while existing approaches have the decline in efficiency by using fixed values or show different qualities depending on the characteristics of the images, the current approach shows the steady results in various images. Comparing with Q2 algorithm obtained in MPEG-4 VM, the approach suggested in this paper gains the benefit of maximum 3.49dB with some variations depending on the characteristics of the images.

Key words: MPEG-4 Video Encoder(MPEG-4 비디오 부호화기), Rate Control(비트율 제어), Quantization(양자화), Q2 Algorithm(Q2 알고리즘), Intra-Frame(인트라 프레임)

※ 교신저자(Corresponding Author): 김정우, 주소: 서울시 관악구 봉천7동 산4-1호(151-818), 전화: 02)2102-0275, FAX: 02)2102-0190, E-mail: myimage@chol.com
접수일: 2004년 8월 23일, 완료일: 2004년 11월 22일
^{*} 준회원, LG전자 DTV연구소 책임연구원
^{**} 정회원, 금강대학교 컴퓨터전공 부교수
(E-mail: shcho@ggu.ac.kr)

1. 서 론

근래에는 다양하고 복잡한 정보들이 동일한 전달 미디어를 통하여 종합적으로 저장 및 전송되어지는데 이를 멀티미디어(Multimedia)라 정의한다[1]. 멀티미디어의 대부분을 차지하는 것은 영상과 음향/

음성이며 이들 데이터의 양은 단순한 방법에 의해 전송되기에는 전달 미디어에 큰 부담이 된다. 이러한 이유로 오랜 기간동안 데이터를 줄이기 위한 효율적이고 실현 가능한 기술들[2-6]이 개발되어 왔다. 압축을 해야하는 정도는 주어진 채널의 주파수 넓이(Bandwidth) 혹은 저장 매체(Storage Media)의 용량에 의존하며 비트 율 제어(RC ; Rate Control) 방법에 의해 원하는 용량으로 조절된다. 압축 표준안들(IS; International Standards)에는 부호화기를 구현하는 업체의 고유한 기술과 융통성을 위하여 비트 율 제어 방법에 관한 구체적인 명시는 하지 않는다. 다만, ISO/IEC의 MPEG과 같은 표준안은 표준안에 적용할 기술의 증명과 부호화기와의 정합을 위하여 참조 소프트웨어(Reference Software, Part 5)와 검증모델(VM ; Verification Model)을 제시하고 있으며 이 내용 속에 비트 율 제어에 관한 방법이 포함되어 있다.

MPEG-4의 검증 모델8.0[7]에서는 비트 율 제어 기술을 위하여 A. Vetro가 제안한 Q2 알고리즘[8,9]을 사용한다. Q2 알고리즘의 기본 개념은 부호화기 내에 버퍼를 두고 부호화를 위해 발생한 비트가 버퍼의 50% 수준을 계속 유지하도록 조절하는 것이다. 이를 위하여 부호화 전에 프레임에 할당할 전체 비트 수를 미리 계산하고 이전 프레임의 양자화 레벨(Quantization Level), 버퍼의 현재 수준 등을 전체적으로 고려하여 현재 프레임을 위한 양자화 레벨을 조절하고 적용한다. 그러나 Q2 알고리즘에서는 P 프레임(Inter-frame)을 위한 양자화 레벨만을 계산하며 I 프레임(Intra-frame)을 위한 양자화 레벨은 미리 지정한 고정된 값을 사용한다. I 프레임은 GOP(Group Of Pictures) 내의 첫 번째 프레임으로써 다음 P 프레임의 참조 프레임(Reference Frame)이 되어 GOP 내의 모든 프레임이 부호화될 때까지 영향을 미친다. 그러므로 I 프레임이 얼마나 잘 부호화되었는가의 여부에 따라 GOP 내의 전체 프레임의 성능을 좌우한다. 앞서 설명한 바와 같이 Q2 알고리즘의 고정된 양자화 레벨은 다양한 영상의 특성을 반영하지 못하여 효율성이 떨어지게 된다.

본 논문에서는 이러한 I 프레임의 양자화를 위한 레벨을 구하는 방법을 제안한다. 가장 큰 영향을 미치는 비트 율의 변화에 따른 레벨의 설정, 현재 입력되는 영상의 특성을 이용하기 위하여 복잡도를 구하

고 레벨의 선정에 적용하는 방법, GOP의 구조에 따른 변화를 반영하기 위하여 GOP 내 P 프레임의 수를 이용하는 방법 등을 제시한다. 논문에서는 각각의 방법들을 조절하기 위하여 가중치를 이용한 하나의 단일 수식으로 통합하고 적용함으로써 단순화 하였다. 이 방법을 다양한 특성을 가진 여러 영상에 적용하여 기존의 방법과 비교함으로써 제안한 방법의 적합성과 성능을 증명하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 MPEG-4 비트 율 제어에 사용된 Q2 알고리즘과 기존의 방법에 대하여 설명하며 3장에서는 양자화 레벨 선정을 위한 새로운 방법을 제시하고 이의 적합성에 대하여 알아본다. 4장에서는 실험을 통하여 기존의 방법과 성능을 비교하며 마지막 5장에서 결론을 맺는다.

2. 기존의 비트 율 제어 방법

2.1 MPEG-4 Q2 알고리즘

MPEG-4 Q2 알고리즘은 A. Vetro에 의해 제안된 방법으로써 R-D 함수(Rate-Distortion function)가 2차식으로 표현되며 MPEG 위원회에 의해 검증모델 8.0부터 비트 율 제어 방법으로 적용되었다. 이 알고리즘의 목적은 부호화기에 버퍼를 두고 각 프레임의 부호화가 끝나고 난 후 버퍼의 점유율이 50%를 유지하도록 양자화기(Quantizer)를 제어하는 것이다. Q2 알고리즘은 크게 4가지의 단계로 이루어지며 여기서는 검증모델19.0[10]을 기준으로 설명한다.

Step 1. 초기화(Initialization)

첫 번째 프레임(I 프레임)의 부호화가 끝나고 난 후 비트 율 제어에 필요한 값들을 초기화하는 역할을 한다. 남은 비트 율과 남은 프레임 수, 버퍼의 크기와 점유율, 양자화 레벨 값, R-D 함수의 계수 값 등이 대표적인 변수들이다.

Step 2. 목표 비트 수 계산(Target Bit Allocation)

이 단계에서는 현재 부호화 할 프레임에 할당하고자 하는 비트 수를 다음의 식을 이용하여 계산한다.

$$T_1 = \max\left(\frac{R_s}{30}, \frac{R_r}{N_r} \times 0.95 + S \times 0.05\right) \quad (1)$$

R_s 는 세그먼트 혹은 GOP에 할당된 비트 율, R_r 은 부호화 후 남은 비트 수, N_r 은 남은 프레임 수를 의미

하며 S 는 이전 프레임의 부호화 시 발생한 비트 수이다. 이 식에서는 R_r 와 N_r 를 이용하여 각 프레임에 할당할 평균 비트 수를 구하여 95%를 적용하고 나머지 5%는 이전 프레임의 부호화시에 발생한 비트 수를 더한다. 최소한의 품질을 유지하기 위하여 계산된 값과 GOP의 평균 비트율 중 큰 값을 선택한다. 목표 비트 수는 버퍼 크기와 현재 버퍼의 점유도를 이용하여 수정한다. 사용된 식은 다음과 같다.

$$T_2 = T_1 \times \frac{2B_s - B}{B_s + B} \quad (2)$$

B 는 현재 버퍼의 점유도, B_s 는 버퍼의 크기이며 검증모델에서는 비트율의 절반 정도를 설정한다. 버퍼의 점유도가 절반 이하이면 비트를 좀더 많이 할당하고 절반 이상이면 비트 수를 조금 줄이는 역할을 한다. 현 단계의 마지막은 오버플로(Overflow)가 발생하지 않도록 비트 수를 조절하게 되며 버퍼 크기의 10%와 90%를 기준으로 이루어진다.

Step 3. 양자화 레벨 계산(Quantization Level Calculation)

현재 프레임의 양자화 레벨을 계산하기 위하여 다음의 2차 함수를 사용한다.

$$T_2 = x_1 \frac{Ec}{Qc} + x_2 \frac{Ec}{Qc^2} \quad (3)$$

단, Ec 는 움직임 보상 후의 MAD(Mean Absolute Difference) 값이며, Qc 는 현재 프레임을 위한 양자화 레벨, x_1 과 x_2 는 계수 값이다. 여기에서 계산된 양자화 레벨은 다음 식 (4)와 같은 클립 함수에 의하여 한정된다.

$$Qc = \min[\text{ceil}(Q_i \times 1.25), Qc, 31] \quad (4-1)$$

$$Qc = \max[\text{ceil}(Q_i \times 0.75), Qc, 1] \quad (4-2)$$

Q_i 은 이전 양자화 레벨이며 현재 양자화 레벨은 이전 양자화 레벨의 25% 범위 내에서 변하게 된다.

Step 4. R-D 함수 갱신

Step 3에서 계산된 Qc 를 이용하여 프레임을 부호화 하고난 후 필요한 변수들을 갱신한다. 장면 전환과 같은 급격한 화면의 변화에 의한 영향을 최소화하기 위하여 적용되는 이전 프레임의 수를 슬라이딩 윈도우(Sliding-Window)를 이용하여 조절한다. 또한, 통계적으로 이상점에 해당하는 값을 제거하기 위

한 단계를 거친다. 마지막으로 현재 버퍼의 점유도를 보고 버퍼 크기의 80% 이상이 되면 다음 프레임의 부호화를 건너뛰게 된다.

2.2 인트라 프레임의 양자화 레벨 설정 : F. Pan의 방법[11]

앞서 설명된 Q2 알고리즘에는 I 프레임에 위한 양자화 레벨의 설정 방법이 없다. I 프레임이 오면 Step 1을 이용하여 초기화 하므로 고정된 양자화 레벨이 항상 선택된다. 일반적으로 최대 양자화 레벨의 중간값($Qc=16$)을 사용하므로 비트율이 아주 높아서 매우 낮은 양자화 레벨이 필요하거나 반대의 비트율에서 아주 높은 양자화 레벨이 필요한 경우에 적절하게 대응할 수 없다. 잘못 부호화된 I 프레임은 다음에 이어지는 P 프레임의 참조 프레임으로 사용되므로 화질의 열화는 계속 전파되어 다음 I 프레임까지 계속된다. 그러므로 I 프레임에 적절한 양자화 레벨을 설정하는 것은 전체 부호화의 효율 측면에서 아주 중요한 역할을 하게 된다.

F. Pan의 논문에서는 I 프레임의 양자화 레벨을 결정하는 방법으로써 다음과 같은 수식을 제안하였다.

$$Q_i = \frac{16.34}{B^{2.05}} \times MAV_{DCT}^{1.0+0.29 \times \ln(B)} \quad (5)$$

단, Q_i 는 I 프레임의 양자화 레벨, B 는 비트율이며 MAV_{DCT} 는 다음의 식에 의해 계산된다.

$$MAV_{DCT} = \frac{1}{N \times M} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} \text{abs}[F(u, v)] \quad (6)$$

$F(u, v)$ 는 8×8 블록의 DCT(Discrete Cosine Transform) 계수 값이며, N 과 M 은 각각 영상의 수평, 수직 픽셀 수이다. 계산된 값은 다음의 식을 이용하여 [5,25] 사이의 정수 값으로 한정하는 클립 함수를 적용한다.

$$Q_i = \max[5, \min(Q_i, 25)] \quad (7)$$

Pan의 방법은 몇 가지 문제점을 가지고 있다. 먼저, Pan은 현재 영상의 복잡도를 계산하기 위하여 MAV_{DCT} 라는 값을 정의하였는데 이 값은 DCT 계수를 이용하여 계산된다. 8×8 블록의 영상을 DCT를 이용하여 주파수 영역으로 변환하면 다음 그림 1과 같은 주파수 특성을 가진다.

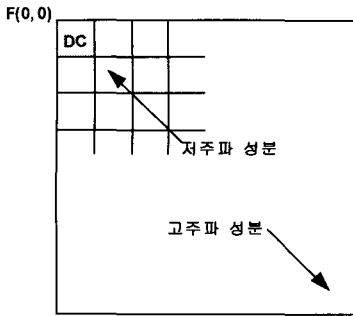


그림 1. DCT 계수 값의 분포

Lam[12]의 논문에 의하면 DCT 계수의 분포 중 $F(0,0)$ 은 DC 값으로써 8×8 블록의 평균이며 나머지 값들은 $F(0,0)$ 에 가까울수록 저주파에, 멀수록 높은 고주파에 해당하며 근사적으로 분산에 해당된다. 만약 영상이 플래시와 같은 밝은 조명이 주어진다면 DC의 값은 크게 올라갈 것이며 같은 형태라도 조명이 없다면 DC의 값은 크게 줄어든 것이다. 이러한 이유로 영상이 복잡한 지의 여부를 DC와 저주파 성분을 모두 포함한 값을 이용하여 계산한다면 상당한 오류를 내포하고 있을 가능성이 높다.

두 번째의 문제점은 식 (5)의 구성에 있다. 식에서 첫 번째 항을 그림 2에 그래프로 표현하였다.

x축은 비트율이며 5Kbps에서부터 1Mbps까지 값을 갖는다. 그래프는 5Kbps에서부터 비트율의 증가에 따라 급격히 줄어들고 약 25~30Kbps 정도에서 변곡점을 보이며 대략 60Kbps를 지나면서 포화 상태에 이른다. 이는 첫 번째 항이 5~60Kbps 정도의 낮은 비트율에서만 적합하도록 구성되어 있는 것으로 간주할 수 있다.

이상에서 살펴본 바와 같이 Pan의 방법은 낮은 비

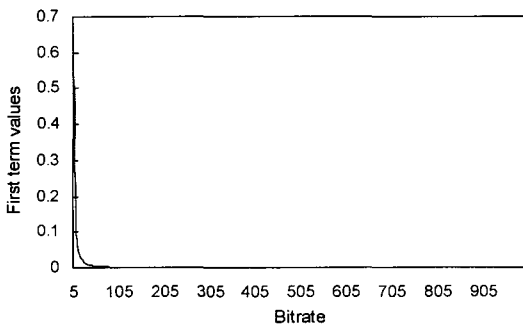


그림 2. F. Pan의 제안 수식 중 비트율 항의 분포

트율에서만 정상적인 동작을 보일 것으로 예상되며 입력되는 영상의 복잡도를 적용하는 점에서도 문제점이 있다. 그러므로 영상과 부호화기의 특성을 잘 표현할 수 있는 값들을 이용하여 양자화 레벨을 구성한다면 부호화기의 효율을 더욱 향상시킬 수 있을 것이다. 이를 위하여 본 논문에서는 다음 장에서 설명될 몇 가지 특성을 이용하여 양자화 레벨을 구하고 실험을 통하여 그 결과를 입증하고자 한다.

3. 제안한 알고리즘

3.1 비트율과 양자화 레벨의 관계

비트율(Bit-rate)은 부호화기에 할당된 초당 비트 수이며 한정된 채널을 통과하기 위하여 반드시 수렴해야 하는 값이다. 일반적으로 비트율이 높을수록 각 프레임에 할당되는 비트 수는 많아지게 되고 이로 인해 낮은 양자화 레벨이 주어져서 좋은 품질의 영상을 얻을 수 있다. 반대로, 비트율이 낮을수록 양자화 레벨은 높아지고 품질은 저하되며 남은 프레임 수를 고려할 때 주어진 비트율에 수렴할 수 없다면 프레임 스킵이 발생하게 된다. 이러한 관점을 고려하면 비트율과 양자화 레벨은 반비례의 관계를 갖는다. 본 논문에서는 이들의 관계를 정립하기 위하여 다음의 수식을 제안한다.

$$Q \propto \frac{1}{1 + ae^{(0.01B-5)\beta}} \quad (8)$$

단, B 는 비트율(Kbps), α 와 β 는 각각 그래프의 형태 정합을 위한 파라미터이다. MPEG에서는 양자화 레벨이 1~31의 정수 값만을 가질 수 있도록 클립 함수를 적용한다. 낮은 양자화 레벨은 양자화 과정에서 적은 에러를 발생시키도록 프레임에 많은 비트를 할당하며 반대로 높은 양자화 레벨에서는 비트율을 맞추기 위하여 최소한의 비트만을 할당하게 된다. 이러한 이유로 낮은 양자화 레벨에서의 레벨 변화는 급격한 비트의 증가 혹은 감소를 가져오며 높은 양자화 레벨에서의 변화는 비트의 변화가 상대적으로 적다. 아주 낮은 비트율에서는 영상의 형태에 상관없이 할당되는 비트 수가 적으므로 비트율의 변화에 둔감(데드존 ; Dead zone)해야 한다. 또한, 아주 높은 비트율에서는 단순한 영상의 경우에는 많은 비트를 할당하더라도 다 소비하지 못하는 경우가 많으므로

역시 비트 율의 변화에 둔감해야 하며 적절한 비트 율에서의 변화에만 양자화 레벨을 변경함으로써 한정된 비트를 효과적으로 배분하여야 한다. 그림 3을 살펴보면 수식 (8)의 그래프가 이러한 상황을 적절하게 표현하고 있음을 알 수 있다.

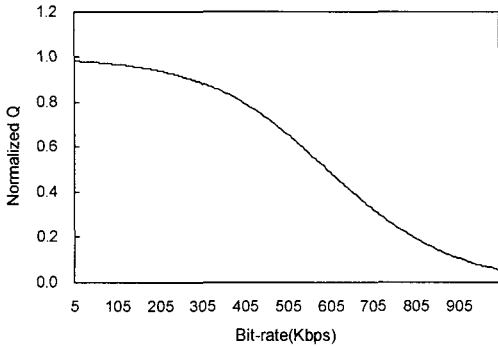


그림 3. 수식 (8)의 그래프(단, $\alpha=0.5, \beta=0.7$)

3.1.1 파라미터 α 의 역할

그림 4에는 파라미터 α 의 변화에 따른 그래프를 나타내었다.

그림에서 보는바와 같이 파라미터 α 의 값이 변할 때마다 그래프의 곡선은 기울기의 변화 없이 x 축을 따라 우측 혹은 좌측으로 움직인다. 이는 양자화 레벨을 적용하고자 하는 비트 율 구간의 변화 없이 단지 비트 율 값만을 변화시키고자 할 때 사용한다. 예를 들면, 그림 4에서 α 값이 1.0인 경우는 약 300Kbps ~ 650Kbps 구간에서 양자화 레벨을 급하게 변화시키며, α 값이 0.5인 경우는 400Kbps ~ 750Kbps의 구간에서 양자화 레벨을 급격하게 변화시키게 된다. 이러한 비트 율 값의 변경은 양자화 레벨에서의 데드존 구간에도 영향을 미치게 된다.

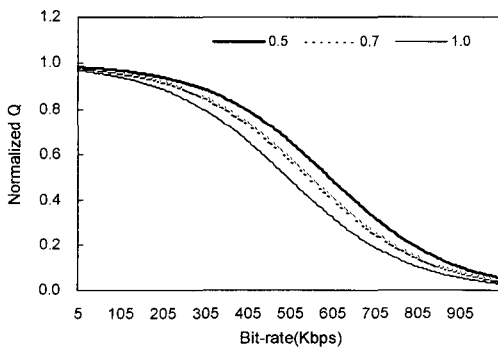


그림 4. 파라미터 α 에 따른 그래프의 변화(단, $\beta=0.7$)

3.1.2 파라미터 β 의 역할

그림 5에는 파라미터 β 에 따른 그래프를 나타내었다. β 의 값이 증가함에 따라 양자화 레벨의 변화 구간 폭은 줄어들고 기울기는 증가하며 데드존은 넓어진다. 반대로 낮은 β 값은 기울기가 낮아지며 데드존이 줄어들고 양자화 레벨이 점차 선형에 가까워진다.

결론적으로 앞서 설명된 바와 같이 α 와 β 의 값은 각 비트 율에서의 양자화 레벨을 적절하게 선택하기 위하여 중점적인 비트 율 구간과 그 구간에서의 기울기, 그리고 데드존을 설정할 수 있다.

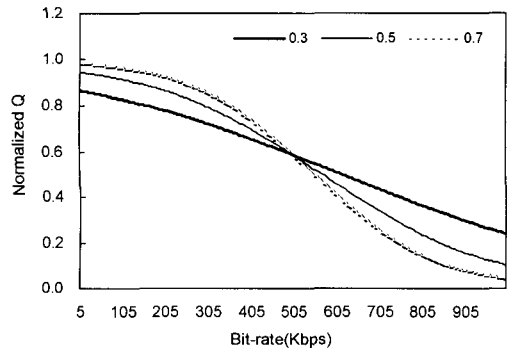


그림 5. 파라미터 β 에 따른 그래프의 변화(단, $\alpha=0.7$)

3.2 복잡도와 양자화 레벨간의 관계

영상의 복잡도는 영상을 주파수 영역으로 변환하였을 때 고주파 부분의 분포가 얼마나 넓고 큰 값을 가지는가 하는 것으로 나타낼 수 있다. 영상을 부호화하기 위하여 같은 비트가 주어졌을 때 영상내의 복잡도가 높은 경우는 비트 율을 맞추기 위하여 양자화 과정에서 고주파 영역의 계수 값을 잘라내게 되어 결과적으로는 PSNR의 감소를 가져온다. 반대로 복잡도가 낮은 경우는 고주파 영역에 계수 값이 적어서 부호화 시에 많은 비트를 필요로 하지 않으므로 원 영상에 가깝게 부호화되고 이로 인해 PSNR은 증가하게 된다. 그러므로 복잡도에 관계없이 일정한 PSNR을 유지하기 위해서는 복잡도가 높은 경우에는 많은 비트를, 복잡도가 낮은 경우에는 적은 비트를 할당하여야 한다. 이를 위해 본 논문에서는 복잡도를 이용한 양자화 레벨의 선정을 위하여 다음의 수식을 제안한다.

$$Q_I \propto \frac{1}{e^{\gamma \cdot MAV_s}} \tag{9}$$

단, γ 는 수식을 표현한 그래프의 형태를 보정하기 위한 파라미터이며 MAV_h 는 다음과 같이 정의된다.

$$MAV_h = \frac{1}{N \times M} \left(\sum_{u=0}^N \sum_{v=0}^M |F(u,v)| - \sum_{u=0}^N \sum_{v=0}^M |F(u,v)| \right) \quad (10)$$

$F(u,v)$ 는 DCT 변환 후의 계수 값이며, N 은 수평 픽셀 수, M 은 수직 픽셀 수를 의미한다. 수식 (10)에 나타낸 MAV_h 는 전체 DCT 계수 값 중 저주파에 해당하는 일부분과 DC를 제거한 값이다.

그림 6에는 수식 (9)의 형태를 나타내었다.

MAV_h 의 값은 복잡도가 높아서 고주파의 값이 클수록 커지며 양자화 레벨은 낮아진다. 이는 할당되는 비트 수가 많아짐을 의미한다. 전체적인 그래프의 곡선은 고주파가 많아지면 급격하게 양자화 레벨이 줄어들다가 어느 정도 이상인 경우에 변화가 완만하다. 파라미터 γ 의 역할은 곡선의 기울기와 영향을 주는 MAV_h 의 범위를 결정한다.

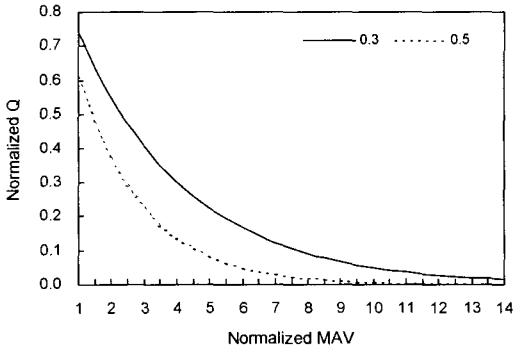


그림 6. 파라미터 γ 에 의한 그래프의 변화

3.3 I 프레임의 거리와 양자화 레벨의 관계

GOP의 구조는 한 개의 I 프레임과 다수의 B, P 프레임으로 구성될 수 있다. 본 논문에서는 심플 프로파일(Simple profile)을 목표로 하고 있으며 이러한 경우의 GOP 구조는 한 개의 I 프레임과 나머지 수 개의 P 프레임(Nr)으로 구성된다. 표준안에서는 GOP의 각 프레임별 구조를 사용자가 정의할 수 있도록 융통성을 제공하고 있으며 비트율 제어 알고리즘들도 초기의 비트율 할당 시에 남은 P 프레임의 수에 따라 반비례하여 비트를 할당한다. 그러나 이들 알고리즘들은 묵시적으로 P 프레임의 수가 일정한 범위 내에 있을 것을 가정하고 있으며 할당되는 비트 또한 선형성을 유지하므로 실제적인 구조의 다양한

변화에 대처하지 못하고 있다. 특히, Q2 알고리즘에서는 P 프레임의 수에 무관하게 I 프레임의 비트 할당이 이루어지고 있어서 GOP 내에 P 프레임의 수가 많은 경우에는 빈번한 프레임 스킵이 발생하게 된다. 그러므로 본 논문에서는 GOP 내의 P 프레임 수를 이용하여 I 프레임의 비트 할당을 조절할 수 있도록 식 (11)을 제안하여 사용한다.

$$Q_I \propto \tanh(\delta \times Nr) \quad (11)$$

그림 7에는 파라미터와 P 프레임 구조의 변화에 따른 I 프레임의 양자화 레벨 변화를 나타내고 있다. X 축을 따라 GOP 내의 P 프레임 수가 많아지면 I 프레임에 할당되는 비트 수를 줄이고 남은 비트 수를 P 프레임에 할당하여 P 프레임의 화질 개선이 이루어지도록 양자화 레벨을 높인다. 그러나 P 프레임의 수가 계속해서 늘어나는 경우에도 I 프레임에 할당되는 비트의 수를 무한정 줄이게 되면 P 프레임의 참조 프레임으로 사용되는 I 프레임의 화질의 열화가 발생되어 전체적인 PSNR이 오히려 줄어들 수 있다. 그러므로 일정한 수준에서 포화 상태를 만들어서 P 프레임의 증가에 따른 I 프레임의 양자화 레벨이 변화가 없도록 구성해야 한다. 본 논문에서 제안한 식을 이용하는 경우 그림 7에서 보는 바와 같이 P 프레임의 증가에 따라 적절하게 양자화 레벨을 높일 수 있으며 일정한 수 이상이 될 때 양자화 레벨은 포화 상태에 이르게 된다. 파라미터 δ 의 역할은 포화 상태에 이르게 되는 속도와 양자화 레벨을 결정하는 기울기를 변화시켜 P 프레임의 수와 I 프레임의 양자화 레벨이 적합한 관계를 가질 수 있도록 조절한다.

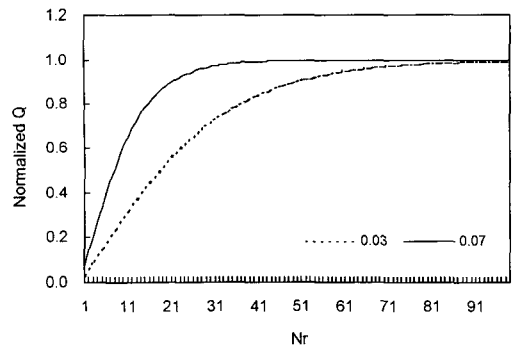


그림 7. 파라미터 δ 에 의한 그래프

3.4 수식 병합 및 가중치 조절

앞 절에서 논의된 여러 가지 인자들을 이용하여

실제 양자화 레벨을 계산하기 위해서는 각 인자들이 양자화 레벨을 결정하는 비율을 조절하는 것이 필요하다. 본 논문에서는 4절의 실험에서 사용된 영상을 제외한 각 Class 당 2개씩의 영상을 이용하여 위에서 설명한 4개의 파라미터들과 3개의 가중치에 대하여 일정한 범위의 값을 주고 큰 스텝사이즈로 PSNR을 조사하였다. 가장 높은 PSNR을 보인 파라미터와 가중치를 중심으로 다시 작은 스텝사이즈를 사용하여 최적의 파라미터와 가중치 값을 선정하였으며 그 결과 비트율, MAV_n, 그리고 Nr의 비율을 1.5 : 0.55 : 0.3으로 설정하였다. 이를 이용하여 전체 수식을 병합하면,

$$Q_I = \left[\frac{1.5}{1 + \alpha e^{(0.01\beta - 3)\beta}} + \frac{0.55}{e^{\gamma \cdot MAV}} + 0.3 \cdot \tanh(\delta \cdot Nr) \right] \times Q_{\max} \quad (12)$$

여기서, Q_{max}는 각 항의 최대 값을 더한 후 32에서 나는 값을 의미하며 양자화 레벨이 일정한 범위 내로 수렴하도록 만들기 위해 사용되었다. 위에서 결정된 Q_I 값은 다시 [1,31]의 정수 값을 가지도록 (13)식을 적용한다.

$$Q_I = \max[1, \min(Q_I, 31)] \quad (13)$$

4. 실험 결과 및 고찰

본 논문에서는 제안한 방법의 효율성을 증명하기 위하여 MPEG-4 Simple profile 부호화기를 이용하여 실험하였으며 Q2 알고리즘은 MPEG-4 Video 검증모델19.0에 소개된 방법을 사용하였다. 사용된 영상은 Akiyo(Class A), Foreman(Class B), Mobile & Calendar(Class C) 각각 150프레임을 이용하였다. 검증모델에 의하면 Class A는 낮은 복잡도와 적은 움직임, Class B는 중간 정도의 복잡도와 적은 움직임, Class C는 높은 복잡도와 중간 정도의 움직임을 가지는 표준 영상으로 정의되어 있으므로 본 논문의 제안 점을 증명하기 위한 실험 데이터로 적합하리라 생각된다. 프레임율은 15fps, 비트율은 40Kbps에서 1Mbps까지 가변시키며 실험하였고 이 때 사용된 파라미터들은 3.4절에서 설명한 실험 방법에 의하여 α = 20, β = 0.3(256Kbps 이하) / 1.0(256Kbps 이후), γ = 0.5, δ = 0.03으로 고정하였다. 실험의 결과는 검증모델19.0을 구현한 마이크로소프트 MPEG4 비주얼 참조 소프트웨어(Visual Reference Software) 2.5[13]

와 비교하였으며, 앞서 소개되었던 Pan의 결과와도 비교하여 나타내었다. 이 중 마이크로소프트의 부호화기는 비트율 제어 부분의 버그가 있어 약간의 수정 후 사용하였다.

표 1에는 각 알고리즘의 실험 결과를 나타내었다.

표 1. PSNR 비교(Nr=14, Unit : dB)

| Sequence | 비트율 (bps) | 각 방법의 PSNR | | | |
|-------------------|-------------------|------------|-------|-------|------|
| | | Q2 | Pan | 제안 | 이득 |
| Akiyo | 40K | 35.47 | 38.27 | 35.52 | 0.05 |
| | 64K | 35.47 | 38.27 | 35.55 | 0.08 |
| | 128K | 35.47 | 38.27 | 35.64 | 0.17 |
| | 256K | 35.47 | 38.27 | 35.79 | 0.32 |
| | 512K | 35.47 | 38.27 | 38.27 | 2.80 |
| | 1M | 35.47 | 38.27 | 38.96 | 3.49 |
| Foreman | 40K | 32.11 | 32.54 | 32.42 | 0.31 |
| | 64K | 33.76 | 34.72 | 34.08 | 0.32 |
| | 128K | 33.86 | 35.07 | 34.30 | 0.44 |
| | 256K | 33.86 | 35.07 | 34.43 | 0.57 |
| | 512K | 33.86 | 35.07 | 35.40 | 1.54 |
| | 1M | 33.86 | 35.07 | 35.61 | 1.75 |
| Mobile & Calendar | 40K ¹⁾ | 23.98 | 24.15 | 26.07 | 2.09 |
| | 64K | 25.56 | 25.40 | 25.57 | 0.01 |
| | 128K | 28.21 | 28.29 | 28.40 | 0.19 |
| | 256K | 30.45 | 31.44 | 31.29 | 0.84 |
| | 512K | 30.45 | 31.44 | 32.05 | 1.60 |
| | 1M | 30.45 | 31.44 | 32.36 | 1.91 |

¹⁾ Mobile & Calendar 영상의 복잡도가 높아서 주어진 40Kbps에 수렴하지 못하였으므로 PSNR의 비교는 큰 의미가 없음.

표를 살펴보면 제안한 방법이 Q2의 방법보다 전반적으로 우수함을 알 수 있다. 낮은 비트율에서는 높은 양자화 레벨이 정의되어야 하지만 Q2의 경우는 고정되어 있는 낮은 값이 사용되어 PSNR이 높아지게 되므로 제안한 방법과 많은 차이가 발생하지 않았다. 높은 비트율에서는 Q2의 방법은 상대적으로 높은 양자화 레벨을 할당하는 본 논문의 방법이 더욱 효과적임을 알 수 있다. 또한, Class A에 속하는 Akiyo 영상의 경우 Q2의 방법은 비트 할당이 적절치 못하여 주어진 비트를 다 소진하지 못하고 남음에 따라

비트 율에 따른 PSNR의 변화가 없지만 제안한 방법은 비트 율에 따라 선형적인 분포를 보이고 있다. Pan의 방법이 Akiyo 영상에서 비트 율에 상관없이 항상 같은 PSNR 분포를 보이는 것은 제시한 양자화 레벨의 산출 방법이 Akiyo 영상에 부적합하기 때문인 것으로 생각된다. Pan의 식에 의해 계산된 양자화 레벨의 값은 40Kbps의 경우에 2.99, 1Mbps에서는 0.06이 산출되어 클립 함수에 의하여 항상 양자화 레벨 5가 할당되었다. 그러므로 이 방법은 Class B 이상의 복잡도를 가지는 영상에서 낮은 비트 율을 가지는 경우에만 적용되어야 할 것으로 생각된다. 이에 비하여 본 논문의 제안 방법은 복잡도와 비트 율에 관계없이 전체적으로 안정된 결과를 보이고 있다.

다음의 표 2는 Nr의 값에 의한 제안 방법의 결과를 보여주기 위하여 실험하였다.

표 2에서는 Foreman과 Mobile & Calendar 영상 300프레임을 64Kbps와 128Kbps의 비트 율에 대하여 Nr 값을 10에서부터 50까지 변화시킨 후 PSNR을 구하였다. 대부분 Q2의 방법에 비하여 나은 결과를 보여준다. 이득은 Nr 값이 커질수록 Q2와 차이가 줄어들다가 50일 때는 오히려 Q2의 PSNR이 높아진다. 이는 제안한 방법의 효율성이 떨어지는 것이 아니라 제안한 방법이 적용되는 I 프레임의 수가 줄어들어서 발생하는 현상으로 분석할 수 있다. P 프레임의 부호화에 소요되는 비트 수가 I 프레임에 비해 아주 적으므로 비록 Q2 알고리즘이 I 프레임의 부호화에 의한 PSNR이 감소하더라도 연속된 몇 개의 P 프레임만 영향을 받으며 이후의 GOP 내에 남아 있는 P 프레

임 부호화를 통하여 충분히 보상받을 수 있다. 그러므로 표에서 보는바와 같이 같은 조건에서 Nr이 클수록 이득이 감소하긴 하지만 복호화기에서는 오류 전파를 방지하거나 임의의 접근을 가능토록 하기 위해서 크지 않은 Nr 값을 요구하므로 이러한 관점에서 Nr이 낮을수록 효율이 좋은 본 논문의 제안 방법의 의미가 있다고 할 수 있다.

그림 8에서는 Mobile & Calendar 300프레임을 이용하여 512Kbps, 15fps에서의 제안한 방법과 Q2, Pan의 방법에 대한 비트할당 분포를 나타내었다. 높은 비트 율에도 불구하고 Q2 알고리즘은 고정된 I 프레임 양자화 레벨에 의하여 충분한 비트를 소비하지 못함을 알 수 있다. 이후의 P 프레임도 이전 비트의 소비량을 참조하게 되며 양자화 레벨 또한 급격하게 변화하지 못하도록 제한되므로 비트 소비량은 천천히 증가하게 되어 결국 주어진 비트를 모두 소비하지 못하는 결과를 보이게 된다. 반면, 제안한 방법과

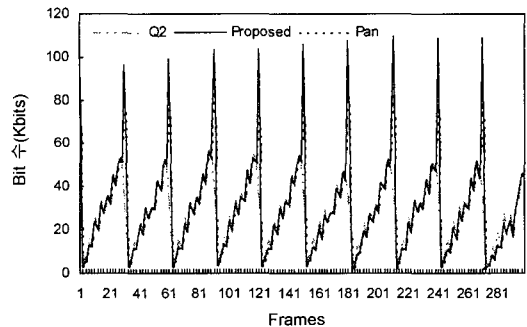


그림 8. 프레임 수의 변화에 의한 비트할당 분포

표 2. Nr의 변화에 따른 PSNR 분포(Unit: dB)

| Sequence | Nr | 64Kbps | | | 128Kbps | | |
|-------------------|----|--------|-------|------|---------|-------|-------|
| | | Q2 | 제안 | 이득 | Q2 | 제안 | 이득 |
| Foreman | 10 | 32.02 | 32.70 | 0.68 | 32.14 | 32.98 | 0.84 |
| | 20 | 33.73 | 33.95 | 0.22 | 35.12 | 35.39 | 0.27 |
| | 30 | 34.18 | 34.25 | 0.07 | 36.92 | 37.02 | 0.10 |
| | 40 | 34.37 | 34.38 | 0.01 | 37.38 | 37.41 | 0.03 |
| | 50 | 34.44 | 34.42 | -0.2 | 37.77 | 37.71 | -0.06 |
| Mobile & Calendar | 10 | 25.77 | 25.87 | 0.10 | 28.19 | 28.74 | 0.55 |
| | 20 | 25.87 | 26.02 | 0.14 | 28.57 | 28.81 | 0.24 |
| | 30 | 25.95 | 25.98 | 0.03 | 28.85 | 28.95 | 0.10 |
| | 40 | 26.05 | 26.07 | 0.03 | 28.94 | 29.02 | 0.08 |
| | 50 | 26.06 | 26.06 | 0.00 | 29.00 | 29.05 | 0.05 |

Pan의 방법은 비트 율에 따라 I 프레임의 양자화 레벨이 변하게 되므로 그래프에서 보이는 바와 같이 높은 비트 율에서 많은 비트를 소비한다. 제안한 방법은 Pan의 방법에 비하여 I 프레임에 더 많은 비트가 할당되었으며 이에 가까운 P 프레임에도 영향을 주어 많은 비트가 소비된다. 그 결과 PSNR의 측면에서도 GOP 내의 초기 부분에 높은 결과치를 보인다. 다만, 한정된 비트 율에서 초반의 많은 비트 소비로 인하여 GOP 내의 후반부에서 다소 낮은 비트 소비량을 보인다. 하지만, GOP 전체에 걸쳐 Pan의 방법보다 높은 PSNR을 가지는 프레임 수가 많으므로 제안한 방법이 영상의 복호화 시 주관적으로 나온 결과를 보여줄 수 있다. 또한 I 프레임의 PSNR이 높으므로 PVR(Personal Video Recorder) 등의 응용에서 빨리 찾기 기능과 같이 I 프레임만을 보여주는 경우에 Pan의 방법보다 향상된 화질을 보여줄 것으로 생각된다. 그림 9에는 그림 8의 일부분을 확대하여 나타내었다.

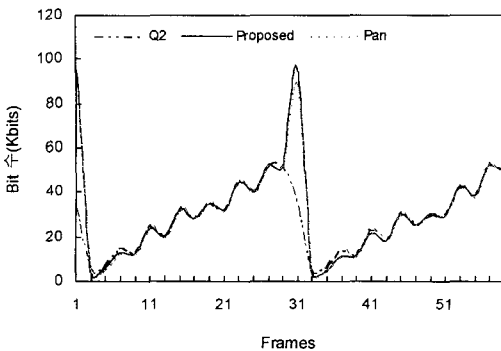


그림 9. 프레임 수의 변화에 의한 비트할당 분포(확대 화면)

5. 결 론

본 논문에서는 I 프레임의 부호화에 필요한 양자화 레벨을 할당하는 방법을 제시하였다. 항상 고정된 값을 이용하여 부호화하는 Q2 알고리즘에 비하여 현재 입력되는 영상의 특성과 부호화 조건에 따라 양자화 레벨을 적절하게 변경하여 적용함으로써 효율성이 증대됨을 증명하였다. 이를 위하여 비트 율을 반영할 수 있는 방법, 영상의 복잡도를 이용하는 방법, 그리고 GOP의 구조를 이용하는 방법을 제시하였고 다양한 영상에 적용하여 실험한 결과 Q2 알고리즘에 비하여 최대 3.49dB까지 개선을 할 수 있었다. 낮은 비트 율에서 Pan의 방식보다 PSNR 측면에서 일부

뒤지는 영상도 있었으나 Pan의 방식이 오직 낮은 비트 율에 대해서만 제어를 하고 있다는 점을 고려할 때 모든 비트 율에서 안정된 결과를 나타내고 있는 본 논문의 방법이 더욱 효과적이고 실용적임을 알 수 있다.

하지만, 복잡도 계산을 위하여 많은 계산량이 필요하다는 점과 비록 제안된 식의 파라미터 값을 많은 실험에 의해 구하였지만 좀 더 효율적인 부호화를 위하여 세밀한 조절이 필요하다는 것이 앞으로 해결해야 할 문제점이라 하겠다.

참 고 문 헌

- [1] 후지와라, *최신 MPEG 교과서*. 아스키 출판국, Aug. 1994.
- [2] ITU-T Rec. H.261, "Video Codec for Audio-visual Services at $p \times 64$ kbits," International Standard, Mar. 1993.
- [3] ITU-T Rec. H.263, "Video Coding for Low Bit Rate Communication," International Standard, Feb. 1998.
- [4] ISO/IEC 11172-2:1993, "Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbits/s-Part 2: Video," International Standard, Dec. 1999.
- [5] ISO/IEC 13818-2:2000, "Information Technology-Generic Coding of Moving Pictures and Associated Audio Information: Video," International Standard, Mar. 2002.
- [6] ISO/IEC 14496-2:2001(E), "Information Technology-Coding of Audio-Visual Objects-Part 2: Visual," International Standard, Dec. 2001.
- [7] ISO/IEC JTC1/SC29/WG11 N1796, "MPEG-4 Video Verification Model Version 8.0," Committee Draft, Jul. 1997.
- [8] A. Vetro, H. Sun, and Y. Wang, "MPEG-4 Rate Control for Multiple Video Objects," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol.9, No.1, pp.186-199, Feb. 1999.
- [9] H. J. Lee, T. Chiang, and Y. Q. Zhang, "Scalable Rate Control for MPEG-4 Video," *IEEE*

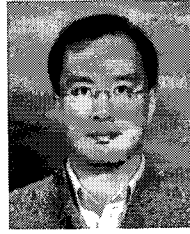
Trans. on Circuits and Systems for Video Technology, Vol.10, No.6, pp.878-894, Sep. 2000.

[10] ISO/IEC JTC1/SC29/WG11 N6184, "MPEG-4 Video Verification Model Version 19.0," Committee Draft, Dec. 2003.

[11] F. Pan, Z. Li, K. Lim, and G. Feng, "A Study of MPEG-4 Rate Control Scheme and Its Improvements," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol.13, No.5, pp.440-446, May 2003.

[12] Edmund Y. Lam, "A Mathematical Analysis of the DCT Coefficient Distribution for Images," *IEEE Trans. on Image Processing*, Vol.9, No. 10, pp.1661-1666, Oct. 2000.

[13] Simon A. J. Winder, "MPEG-4 Visual Reference Software FDAM1-version 2.5," Microsoft Corporation, Feb. 2004.



김 정 우

1991년 2월 성균관대학교 전자공학과 졸업(학사)
 1993년 2월 성균관대학교 대학원 전자공학과 졸업(공학석사)
 1998년 8월 성균관대학교 대학원 전자공학과 졸업(공학박사)

2000년 2월 성균관대학교 부설 과학기술 연구소 연구원
 2000년 3월~현재 LG전자 DTV연구소 책임연구원
 관심분야: 영상처리, 영상압축, HDTV, 멀티미디어통신



조 성 환

1980년 성균관대학교 전자공학과 졸업(학사)
 1982년 성균관대학교 대학원 전자공학과 졸업(공학석사)
 1991년 성균관대학교 대학원 전자공학과 졸업(공학박사)
 1982년~1985년 해군사관학교

전기 및 전자공학과 전임강사
 1997년 미국 Columbia 대학 CATT Visiting Scholar
 1985년~2002년 동서울대학 컴퓨터공학과 부교수
 2002년~현재 금강대학교 컴퓨터전공 부교수
 관심분야: 영상처리, 영상통신, 패턴인식, DRM