

논문 2005-42SP-6-8

선형 모델 기반 매크로블록 레이어 비트율 제어 기법

(Macroblock Layer Bit-rates Control Algorithm based on the Linear Source Model)

서 동 완*, 최 윤 식**

(Dong-Wan Seo and Yoonsik Choe)

요 약

본 논문에서는 H.263, H.263+, MPEG-4와 같은 블록 기반 비디오 압축을 위한 비트-율 제어 기법을 제안한다. 제안 알고리즘은 선형 모델에 기반한 라그랑지안 최적화 기법을 통해서 매크로블록 당 할당되는 양자화 파라미터의 조합을 산출한다. 제안 알고리즘에 사용하는 라그랑지안 비용함수는 선형 모델에 의해서 산출되는 비트-율과 왜곡을 통해 구한다. H.263과 MPEG-4의 양자화 파라미터를 부호화하는 Dquant 방법에 맞도록 하기 위해 Dquant 방법을 고려하면서 동시에 라그랑지안 최적화 기법에 의한 해를 구하는 비터비 알고리즘에 의해 양자화 파라미터 조합을 구한다. 제안 알고리즘은 모델 기반 비트-율 제어 기법 중 하나인 TMN8 기법에 비해 1.5 dB 정도의 비디오 화질 개선 효과를 가지며, 선형 모델 기반 비트-율 제어 기법인 일정화질 접근 방법에 비해 화질은 거의 동일하게 유지하면서 비트-율의 변동폭이 줄어드는 효과가 있다.

Abstract

In this paper, we propose the bit-rate control algorithm for the block based image compression like H.263, H.263+ or MPEG-4. The proposed algorithm is designed to identify the quantization parameter set through the Lagrangian optimization technique based on the well-known linear source model. We set the Lagrangian cost function with the rates and distortion calculated from the linear source model. We calculate the quantization parameter set using the Viterbi algorithm to solve the Lagrangian optimization problem considering the Dquant method of H.263 and MPEG-4. The proposed algorithm improves the video quality by up to 1.5 dB compared with the TMN8 scheme, and is more effective in the video sources with dynamic activities than the consistent quality approaches.

Keywords : Lagrangian Optimization, MPEG-4, Rate Control, Viterbi

I. 서 론

H.26x 나 MPEG 등의 압축 표준에서는 DCT(Discrete Cosine Transform)에 이어 양자화, VLC (Variable Length Coding)의 과정을 사용한다. 양자화의 과정에서 비트 양과 왜곡의 양이 결정되므로, 비트-율 제어 기법은 목표 비트-율 내에서 왜곡의 양이 최소가

되도록 양자화 파라미터를 선택해야 한다. 즉, 비트-율 조정 기법은 목표 비트-율을 맞추면서 평균 왜곡의 양을 최소화하는데 그 목표가 있다. 채널용량과 부/복호기의 버퍼 컨트롤을 고려하여 프레임 별 출력 비트-율의 변동 폭을 줄여야 한다.

일반적으로 비트-율 제어 기법은 전체 영상의 왜곡을 최소화 하도록 영상의 프레임이나 매크로블록에 소요될 비트와 왜곡의 양을 고려하여 매크로블록마다 양자화 파라미터를 설정하여 부호화한다. 과거의 비트-율 제어 기법은 두 가지 방법에 근간하여 양자화 파라미터를 구하는 연구가 진행 되어 왔다. 하나는 통계적 특성에 따라 비트-율 모델과 왜곡 모델에 근간하여 각 매크로블록의 양자화 파라미터를 산출하는 모델 기반 비트-율

* 학생회원, ** 정회원, 연세대학교 전기전자공학과
(Department of Electrical & Electronics
Engineering, Yonsei University)

※ 본 논문은 차세대방송기술연구센터의 지원을 받았
습니다.

접수일자: 2005년3월16일, 수정완료일: 2005년9월22일

제어 기법^[1-7]이고, 다른 하나는 실제 비트-율과 왜곡에 근간하여 최적 양자화 파라미터를 산출하는 방법^[8-11]이다. 모델 기반의 비트-율 할당 기법은 양자화 파라미터와 비트-율, 왜곡의 양 사이의 관계를 수식적인 모델링을 통하여 비트-율과 왜곡을 모델링하고, 주어진 비트-율 내에서 왜곡을 최소화하는 양자화 파라미터 조합을 찾아 부호화하는 방식이다. 모델 기반의 접근 방법은 딜레이를 최소화하는 실시간 부호화 환경에서 많이 사용하고 있지만 통계적 특성에 따른 모델을 사용하기 때문에 장면이 급격히 변화하는 프레임에서는 비트-왜곡 특성의 예측이 어려워 최적의 양자화 파라미터 조합을 구할 수 없다. 즉, 비트-율은 어느 정도 맞출 수 있으나 화질 면에서 성능이 떨어지는 현상이 일어난다. 기존의 모델 기반의 이러한 단점을 줄이기 위해서 Non-zero DCT(Discrete Cosine Transform) 계수와 비트-율, 왜곡과의 관계를 선형 모델로 정의한 접근 방식이 제안되었다^[5-7]. 하지만 선형 모델 기반의 접근 방식은 프레임 당 하나의 양자화 파라미터를 설정하므로 화질의 일정한 유지에는 용이하나 비트-율의 심한 변동을 막을 수 없는 한계가 있었다. 실제 비트-율과 왜곡에 기반한 최적의 양자화 파라미터를 찾는 기법은 다이나믹 프로그래밍 기법(Dynamic Programming)^[8]과 비터비(Vitervi) 알고리즘을 통한 라그랑지안(Lagrangian) 최적화 기법^[9-11]을 이용하여 가능한 양자화 파라미터 조합 중에서 주어진 비트-율 내에서 최소의 왜곡 양을 가지는 양자화 파라미터 조합을 찾는다. 하지만 비트-왜곡 면에서 최적의 방법이나 실시간 적용은 불가능하다.

본 논문에서는 비트-율 변동을 줄이면서 화질을 유지하는 선형 모델에 기반한 매크로블록 레이어 비트-율 제어 기법을 제안한다. 선형모델은 실제 비트-왜곡 특성과 거의 일치하는 것으로 알려져 있으므로 제안 알고리즘의 비트 모델과 왜곡 모델은 선형모델을 통해서 얻는다^[5-7]. 제안 알고리즘에서 비트 모델과 왜곡 모델을 통해서 라그랑지안 비용함수를 생성하게 된다. 제안 비트-율 조정 기법에서는 H.263^[12]과 MPEG-4^[13]의 Dquant 방식의 제약을 뛰어 넘기 위해서 비터비 알고리즘에 Dquant의 제약을 두어 비용함수가 최소가 되는 양자화 파라미터를 얻는다. 본 논문의 II장에서는 비트-율 할당 문제를 정의하고 최적 양자화 파라미터 조합을 구하는 방법을 소개하였다. III장에서는 기존의 수학적 모델에 기반한 비트-율 제어 기법과 그의 문제점을 설명하였다. IV장에서는 본 논문에서 제안하는 알고리즘이 양자화 파라미터 조합을 구하는 방법을 구체적으로

로 제시하였다. V장과 VI장에서 각각 실험결과와 결론을 보이도록 하였다.

II. 비트-율 할당 문제

본 장에서는 비트-율 할당 문제를 정의하고, 이를 바탕으로 각 매크로블록의 최적 양자화 파라미터라는 해를 구하는 방법을 설명한다. 비트-율 할당 문제는 주어진 비트-율(R_T) 내에서 전체 왜곡(D)을 최소화하는 각 매크로블록의 양자화 파라미터를 구하는 문제로 정의할 수 있다. 비트-율 할당 문제는 라그랑지 멀티플라이어(Lagrange multiplier)에 의해 비트-율과 왜곡을 연계한 문제로 식 (1)에 의해 정의될 수 있다^[9].

$$Q_1^*, Q_2^*, \dots, Q_N^* = \underset{Q_1, Q_2, \dots, Q_N, \lambda}{\operatorname{argmin}} \sum_{i=1}^N (D(Q_i) + \lambda R(Q_i)) \quad (1)$$

여기서, Q_i^* 는 최종적으로 구해지는 i 번째 매크로블록의 양자화 파라미터이고, Q_i 는 i 번째 매크로블록의 선택 가능한 양자화 파라미터이다. λ 는 양의 실수 값을 가지는 라그랑지 멀티플라이어이고, $D(Q_i)$ 와 $R(Q_i)$ 는 양자화 파라미터 Q_i 로 i 번째 매크로블록을 부호화할 때 발생하는 왜곡과 비트-율이다.

최적 양자화 파라미터를 구하는 기법에서는 실제 부호화를 통해 발생하는 왜곡과 비트-율을 사용하며, 모델 기반 비트-율 제어 기법에서는 수학적으로 모델링된 왜곡과 비트-율을 사용한다.

최적 양자화 파라미터를 구하는 기법에서는 식 (1)에서와 같이 왜곡과 비트-율로 이루어진 비용함수를 놓고 비터비 알고리즘이나 다이나믹 프로그래밍 방법으로 누적 비용함수가 최소가 되는 최적 양자화 파라미터를 구한다^[8-10].

III. 모델 기반 비트-율 제어 기법

본 장에서는 모델 기반 비트-율 제어 기법을 통계적 특성에 따라 TMN8 기법^[1]과 선형모델 기반의 기법^[5-7]을 통해 설명하도록 하였다. TMN8은 선택된 모델과 실제 비트-율/왜곡의 양과의 차이로 인해 그 효율이 떨어진다. 선형모델 기반의 기법은 모델은 실제 비트-율과 왜곡의 양을 잘 반영하나 프레임 당 하나의 양자화 파라미터를 모든 매크로블록에 할당하므로 비트의 변동

이 심한 단점이 있다.

1. TMN8 비트-올 제어 기법

TMN8 비트-올 제어 기법은 DCT 계수의 통계적 특성을 라플라시안 분포로 가정하여 식 (2)에 의해 표현하고, 왜곡모델은 균등 양자화기를 통한 유니폼(uniform) 분포를 가지는 DCT 계수를 가정하여 식 (3)와 같이 표현한다.

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

$$D(Q_i) = \frac{1}{N} \sum_{i=1}^N \alpha_i^2 \frac{Q_i^2}{12} \tag{3}$$

여기서, Q_i , σ_i , α_i 는 각각 i -번째 매크로블록의 양자화 파라미터, 분산, 가중치를 나타내고, C 는 헤더 비트 수, K 는 모델 파라미터, A 는 매크로블록 내의 픽셀 수, N 는 프레임 내의 매크로블록 수이다.

위의 비트-올 모델과 왜곡 모델을 식 (1)에 적용하여 미분방정식에 의해 양자화 파라미터를 구하면 식 (4)로 나타낼 수 있다.

$$Q_i^* = \sqrt{\frac{AK}{(B-ANC)} \frac{\sigma_i}{\alpha_i} \sum_{k=1}^N \alpha_k \sigma_k}, 1 \leq i \leq N \tag{4}$$

TMN8에서는 매크로블록 별로 큰 계산량 없이 양자화 파라미터를 식 (4)를 통해 계산할 수 있다. H.263이나 MPEG-4에서는 현재 매크로블록의 양자화 파라미터가 이전 매크로블록의 양자화 파라미터에서 $\{-2, -1, 0, 1, 2\}$ 의 값 내의 변화를 가져야 한다. 하지만 TMN8 비트-올 제어 기법은 이러한 변화를 따라가지 못하여 강제적으로 양자화 파라미터 값을 변화시켜야 한다. 비트 모델과 왜곡모델이 실제 비트와 왜곡을 반영하지 못하므로 최적 양자화 파라미터 조합에 접근을 하지 못한다.

2. 선형 모델에 기반한 비트-올 제어 기법

선형 모델을 기반한 고정 양자화 파라미터 비트-올 제어 기법은 "0"이 아닌 DCT 계수의 수와 출력 비트 수가 선형 관계임을 이용한다^[5-7]. "0"이 아닌 DCT 계수의 수를 NZC(Non Zero Coefficients)로 정의하여 프레임에 고정 양자화 파라미터가 적용될 때 NZC에 따른 DCT 계수의 비트 올 모델을 식 (5)와 같이 정의한

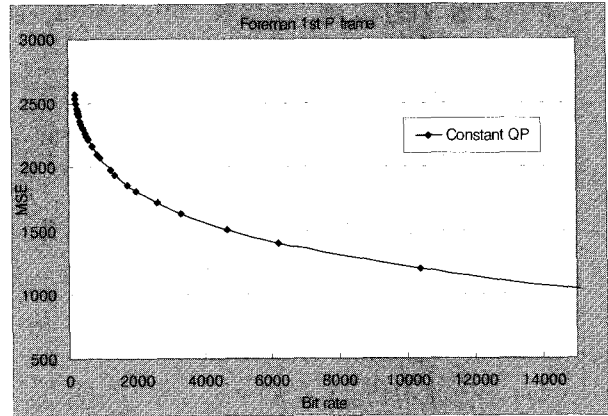


그림 1. 고정 양자화 파라미터의 비트-왜곡 특성
Fig. 1. R-D characteristics of the constant quantization parameter.

다^[7].

$$R(Q) = \sum_{qt \in I, N} (\alpha_{qt} \times NZC_{qt}(Q)) \tag{5}$$

여기서, α_{qt} 는 모델 파라미터, NZC_{qt} 는 "0"이 아닌 DCT 계수, Q 는 양자화 파라미터, qt 는 매크로블록 모드($qt \in \{Intra, Non-Intra\}$)를 나타낸다. 양자화 파라미터는 프레임 당 하나로 설정되어 해당 프레임 전체 매크로블록에 같은 양자화 파라미터로 설정된다.

선형 모델에 기반한 고정 양자화 파라미터 비트-올 제어 기법에서 사용하는 왜곡은 모델을 통해서 구하지 않고 실제로 일어난 DCT 계수에 해당하는 값이 양자화 되었을 때 일어나는 왜곡을 통해 구한다. 따라서 기존의 모델 기반 알고리즘보다 정확한 왜곡 값을 사용한다^[7].

비트-올 모델과 왜곡 모델이 실제 비트-올과 왜곡과 거의 일치하므로 효과적이다. 하지만 프레임 당 하나의 양자화 파라미터에 의해 제어하므로 프레임 별로 목표 비트에 근접하기는 힘든 특성이 나타난다. 이러한 특성은 프레임 별로 하나의 양자화 파라미터를 사용하는 모든 고정 양자화 파라미터 비트-올 제어 기법에서 나타난다. 이러한 이유로 인해 CBR (Constant Bit Rates) 환경에서는 적합하지 못하다. 비트-올의 변동 특성은 그림 1에서 나타내고 있다. 그림 1과 같이 목표 비트가 8000 비트 14000 비트 등으로 설정되는 경우 그 프레임 내에서는 실제 비트에 다다를 수 없어 비트 변동이 심하게 나타나게 된다.

IV. 선형모델 기반의 매크로블록 레이어 비트-율 조정 기법

본 장에서는 TMN8과 같은 통계적 특성을 이용한 모델 접근의 비트-왜곡과 실제 비트-왜곡 특성간의 차이로 인한 문제와 기존의 선형 모델 기반 고정 양자화 파라미터 제어 기법의 비트율 변동 문제를 해결하기 위해 선형 모델 기반 매크로블록 레이어 비트율 제어 기법을 제안한다.

기존의 라그랑지안 최적화 기법은 최적 양자화 파라미터를 산출하기 위해 모든 매크로블록을 모든 양자화 파라미터로 부호화 한 후, 실제 비트 양과 왜곡의 양을 통해 식 (1)과 같은 라그랑지안 문제로 만들고 이를 비터비 알고리즘이나 다이나믹 프로그래밍을 통해 해결하였다. 우리의 이전 연구^[11]에서는 이를 H.263이나 MPEG-4에 적용할 수 있도록 Dquant 제한 조건을 해결하는 비터비 알고리즘을 제안하였다. 우리의 이전 연구^[11]에서는 모든 양자화 파라미터에 의해 모든 매크로블록을 부호화하여 실제 비트와 실제 왜곡을 구하여 Dquant 제한 조건이 있을 당시 최적 양자화 파라미터 조합을 찾았다. 본 논문은 이와 같은 방법을 통해서 Dquant의 문제를 해결한다. 또한 실시간 적용을 위하여 실제 비트-율과 왜곡을 사용하지 않고 선형 모델에 기반하여 비트-율과 왜곡을 산출하여 사용한다.

1. Dquant 제약을 해결하는 비터비 알고리즘

H.263이나 MPEG-4와 같은 비디오 압축 표준은 Mquant로 부호화하는 MPEG-2와는 달리 현재 매크로블록의 양자화 파라미터를 이전 매크로블록의 양자화 파라미터와의 차이 값으로 부호화한다. 차이 값에 의해 부호화하는 방식을 Dquant라 부르며, Dquant의 값은 $\{-2, -1, 0, 1, 2\}$ 의 범위를 가진다. 식 (1)과 같은 문제를 비터비 알고리즘을 이용하여 라그랑지안 최적화 기법을 통해 양자화 파라미터를 구하면 Dquant의 조건을 만족하지 못한다. 식 (1)에 의해 주어진 비트-율 할당 문제를 Dquant 조건을 반영하여 새로이 정의하면 식 (6)과 같다.

$$Q_1^*, DQ_2^*, \dots, DQ_N^* \\ = \underset{Q_i, DQ_2, \dots, DQ_N, \lambda}{\operatorname{argmin}} \sum_{i=1}^N (D(Q_i) + \lambda R(Q_i)) \quad (6)$$

식 (6)에서 Dquant를 만족하는 양자화 파라미터 사이의 관계는 식 (7)을 통해 표현할 수 있다.

$$Q_i = Q_{i-1} + DQ_i \quad (7)$$

여기서, DQ_i 는 i 번째 매크로블록의 Dquant 값으로 $DQ_i \in \{-2, -1, 0, 1, 2\}$ 의 조건을 만족해야 한다.

제안 비트-율 제어 기법에서는 라그랑지안 비용함수를 최소화하는 각각의 양자화 파라미터(Mquant)를 찾는 대신에 첫 번째 매크로블록의 양자화 파라미터와 나머지 매크로블록의 Dquant 값을 찾는다. 사실, Mquant를 찾는 방법이 매크로블록의 최적 양자화 파라미터를 찾는 방법이다. Mquant 방법에서는 매크로블록의 양자화 파라미터에 5 비트를 할당하지만 Dquant 방법에서는 Dquant 값이 "0"인 경우에는 제로 비트를 할당하고 이외의 경우에는 2 비트를 할당하게 된다. 저 비트-율 환경에서는 양자화 파라미터를 부호화하는데 드는 비트 차이로 인해 부호화 효율이 더 좋아진다.^[11] 식 (6)과 같이 Dquant라는 제약이 있는 라그랑지안 최적화 방법을 해결하기 위해서 비터비 알고리즘에 Dquant라는 제약을 추가하여 양자화 파라미터를 구한다.

본 논문에서 사용하는 라그랑지안 비용함수를 비터비 알고리즘의 비용함수로 사용한다. 우선, 모든 양자화 파라미터(1,2,...,31)에 따른 매크로블록 별 비트양과 왜곡의 양을 구하여야 한다. 이 과정은 선형 모델 기반의 고정 양자화 파라미터 비트율 제어 기법^[6]과 같은 방법으로 행한다. 총 라그랑지안 비용함수를 최소화 시키는 라그랑지안 멀티플라이어, λ 를 구한다.^[7,8,9] 다음으로 매크로블록 별로 양자화 파라미터를 구한다. 비터비 알고리즘을 통한 매크로블록 별로 양자화 파라미터를 구하는 과정은 아래와 같다.

목표 비트율과 유한개의 양자화 파라미터가 주어질 때 매크로블록에 최적의 양자화 파라미터를 설정하는 문제는 라그랑지 비용함수를 통하여 재 정리된다. 즉, M 개의 양자화 변수 $\{Q_1, Q_2, \dots, Q_M\}$, 프레임 내의 N 개의 매크로블록, 파라미터 λ 가 주어질 때, 각각의 매크로블록의 비용함수는 식 (8)과 같이 주어진다. 본 논문에서 제안하는 비트율 제어 기법에서는 식 (8)의 비용함수가 최소가 되도록 Dquant 조건을 만족하는 양자화 파라미터 조합 $Q^* = \{Q_1^*, Q_2^*, \dots, Q_N^*\}$ 를 구하는데 그 목표가 있다.

$$J_i(j) = D_i(j) + \lambda R_i(j) \quad (8)$$

여기서, $J_i(j)$, $D_i(j)$, $R_i(j)$ 는 각각 $Q_i = j$

($\in \{1, 2, \dots, 31\}$)로 부호화할 때 발생하는 i 번째 매크로블록의 라그랑지안 비용함수, 왜곡, 비트가 된다.

식 (8)의 비용함수를 모든 매크로블록의 모든 양자화 파라미터에 대하여 구한 후, 누적 비용함수를 그림 2에서와 같이 구한다. 이전 매크로블록이 존재하지 않으므로 첫 번째 매크로블록의 누적 비용함수는 첫 번째 매크로블록의 라그랑지안 비용함수와 같다. 그림 2는 i 번째 매크로블록의 양자화 파라미터가 j 와 같을 경우 누적 비용함수를 구하는 계산하는 과정을 나타내고 있다. i 번째 매크로블록의 양자화 파라미터 j 로 오는 최적의 경로는 ($i-1$) 번째 매크로블록의 양자화 파라미터 $j-2, j-1, j, j+1, j+2$ 중에 최소의 누적 비용함수를 가지는 양자화 파라미터가 된다. $\hat{J}_{i-1}(j-1)$ 가 최소라면, i 번째 매크로블록의 양자화 파라미터, j 의 누적 비용함수는 식 (9)와 같이 되며, i 번째 매크로블록의 양자화 파라미터, j 의 이전 노드는 ($i-1$) 번째 매크로블록의 양자화 파라미터, $j-1$ 이 된다.

$$\hat{J}_i(j) = \hat{J}_{i-1}(j-1) + J_i(j) \quad (9)$$

위의 과정을 반복 수행하면, N 번째 매크로블록의 모든 양자화 파라미터(1, 2, ..., 31)의 누적 비용함수를 산출할 수 있다. Dquant 조건을 만족하는 최적 양자화 파라미터 조합은 N 번째 매크로블록에서 최소의 누적 비용함수 값을 가지는 경로를 형성하는 양자화 파라미터 조합과 같다.

그림 3은 비터비 알고리즘을 이용하여 최종적인 양

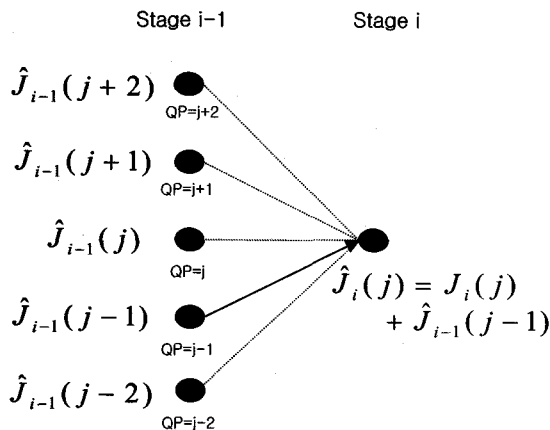


그림 2. Dquant 조건을 만족하는 누적 비용함수
Fig. 2. The cumulative cost function with the condition of the Dquant.

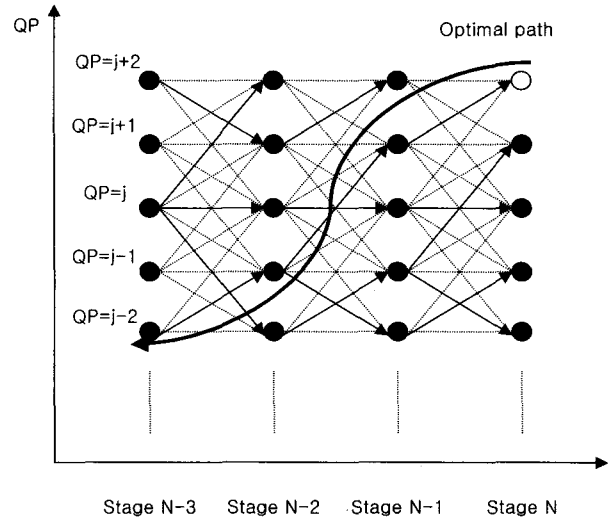


그림 3. 누적 비용함수를 통한 비터비 알고리즘
Fig. 3. Viterbi algorithm with the cumulative cost function.

자화 파라미터 조합을 구하는 과정을 나타낸다. N 번째 매크로블록에서 최소의 누적 비용함수를 가지는 양자화 파라미터가 $j+2$ 와 같다면, N 번째 매크로블록의 노드 ($j+2$)는 최소의 ($N-1$) 번째 누적 비용함수를 가지는 양자화 파라미터를 가진다. 그러므로 그림 3과 같이 ($N-1$) 번째 매크로블록의 양자화 파라미터는 $j+1$ 이 된다. 즉, $Q_N = j+2, Q_{N-1} = j+1, Q_{N-2} = j-1, Q_{N-3} = j-2$ 가 되며, $DQ_N = 1, DQ_{N-1} = 2, DQ_{N-2} = 1$ 가 된다. N 번째 매크로블록에서 첫 번째 매크로블록까지 위의 과정을 반복하여 Dquant 조건을 만족하는 최종 양자화 파라미터 조합 $\{Q_1, Q_2, \dots, Q_N\}$ 를 선택할 수 있다. 위의 양자화 파라미터 조합은 $\{Q_1, DQ_2, DQ_3, \dots, DQ_N\}$ 를 찾는 것과 같다. 위의 결과를 통해 양자화 파라미터를 Mquant(5 비트)로 부호화하지 않고 Dquant(2 비트)로 부호화할 수 있다.

2. 선형 모델 기반의 양자화 파라미터 조합

본 절에서는 선형 모델에 기반한 비트-왜곡 값 예측 기법과 라그랑지 비용함수에 의한 비터비 알고리즘을 적용시켜 양자화 파라미터 조합을 구하는 알고리즘을 구체적으로 설명한다. 전체 부호화 과정을 설명하고 구체적으로 라그랑지 멀티플라이어와 양자화 파라미터 조합을 구하는 과정을 설명하도록 한다.

우선 전체 부호화 과정을 설명하면, 과정 1에 의해 최종적으로 선형 모델 기반의 매크로블록 레이어 비트를

제어가 이루어진다. 첫 번째 단계에서는 파라미터 초기화로 오버헤드 비트와 모델 파라미터의 초기 값을 구한다. 두 번째 단계에서는 양자화 파라미터 조합을 계산한다. 다음으로 모든 매크로블록을 부호화한 후, 파라미터를 갱신하고 다음 프레임으로 진행한다.

과정 1

Step 1 : 파라미터 초기화

a. "0"이 아닌 DCT 계수 당 비트 수를 초기화한다. 가변 길이 부호화하므로 실제 크기는 다양한 분포를 보이지만 다양한 영상에서 구한 평균값으로 파라미터를 추정하여 시퀀스마다 고정 값을 사용한다.

$$\alpha_{intra} = \alpha_{inter} = 7 \quad (10)$$

b. DCT에 사용할 수 있는 목표비트를 정하기 위해 DCT 이외에 사용될 오버헤드 비트(헤더 비트)를 추정하여 사용한다. Inter 매크로블록에서 움직임벡터를 4개 가지는 INTER4V 모드의 오버헤드 비트 파라미터는 따로 설정하여 예측 정확성을 높인다.

$$\begin{aligned} Obit_{intra} &= Obit_{inter} = 10 \\ Obit_{inter4v} &= 40 \\ Obit_{frameheader} &= 54 \end{aligned} \quad (11)$$

Step 2 : 양자화 파라미터 조합 계산

움직임 추정에 의한 차이 픽셀 값의 DCT 계수 정보와 매크로블록의 모드 정보를 이용하여 양자화 파라미터 조합을 찾는다.

a. 비트 수 추정 : 매크로블록마다 주어진 DCT 계수들은 그 크기에 따라 다음 변수의 크기를 하나씩 증가시키는 단순한 연산으로 식 (12)과 같이 변수를 구한다. 모델 파라미터들을 사용하여 매크로블록마다 비트수를 추정한다.

$$NZC_i [Q=1], \dots, NZC_i [Q=31] \quad (12)$$

$$\begin{aligned} \text{Intra 모드} : R_i [Q] &= \alpha_{intra} \times NZC_i [Q] \\ \text{Inter 모드} : R_i [Q] &= \alpha_{inter} \times NZC_i [Q] \end{aligned} \quad (13)$$

b. 왜곡 값 추정 : N 개의 매크로블록마다 주어진 양자화 파라미터에 의한 왜곡의 양을 [7]과 같이 구한다. 이는 식 (14)로 표현된다.

$$D_i [QP=1], \dots, D_i [QP=31] \quad (14)$$

c. DCT 계수에 사용할 수 있는 총 비트 수 계산

$$\begin{aligned} B_{DCT} &= B_{frame} - Obit_{intra} \times N_{intra} \\ &\quad - Obit_{inter} \times N_{inter} \\ &\quad - Obits_{inter4v} \times N_{inter4v} \\ &\quad - Obit_{frameheader} \end{aligned} \quad (15)$$

d. 비터비 알고리즘에 의한 양자화 파라미터 조합 계산 : 예측한 비트 값과 왜곡 값에 따라 IV.1에서와 같이 Dquant를 고려한 비터비 알고리즘에 적용하여 전체 비용 함수를 최소화하는 양자화 파라미터 조합을 찾는다. $(Q_1^*, Q_2^*, \dots, Q_N^*)$

Step 3 : 모든 매크로블록 부호화

앞서 구한 양자화 파라미터 조합 $(Q_1^*, Q_2^*, \dots, Q_N^*)$ 으로 프레임 내의 모든 매크로블록을 부호화 한다.

Step 4 : 파라미터 업데이트

식 (8)과 (9)의 파라미터를 이 전 프레임까지의 평균으로 계산한다.

Step 5 : 프레임 반복

다음 프레임으로 Step 2의 과정부터 반복한다.

과정 1의 비터비 알고리즘에 의한 양자화 파라미터 조합 계산의 단계를 구체적으로 살펴보면, 다음과 같다. 비용함수를 구하기 위해서는 왜곡, 비트-율과 함께 최적의 라그랑지 멀티플라이어(λ^*)가 필요하다. 라그랑지 멀티플라이어 하나가 하나로 선택되었을 때 본 장의 1절에서 설명한 비터비 알고리즘에 의한 양자화 파라미터와 현재 프레임의 왜곡과 비트 양을 구하는 과정을 설명한 후, 최적의 라그랑지 멀티플라이어를 구하는 과정을 설명하도록 하겠다.

주어진 λ 에 대해 라그랑지안 비용 함수를 모든 매크로블록의 모든 양자화 파라미터에 대해 구한다. i 번째 매크로블록의 양자화 파라미터가 j 일 경우, 이전 매크로블록의 노드($j-2, j-1, j, j+1, j+2$) 중에서 최소의 누적 비용함수를 가지는 노드를 찾는다. 찾은 노드를 저장하고 식 (9)와 같은 방법으로 i 번째 매크로블록의 노드 j 의 누적 비용함수를 계산한다. 위와 같은 방법을 반복하여 N 번째 매크로블록의 모든 양자화 파라미터에 대한 누적 비용함수를 계산하여 최소의 N 번째 누적 비용함수를 가지는 노드를 찾는다. 이전 경로를 찾아서 최종적인 양자화 파라미터 조합을 찾는다. 최종적인 양자화 파라미터 조합이 형성하는 왜곡과 비

트의 양이 현재 프레임의 왜곡과 비트 양이 된다.

최적 라그랑지 멀티플라이어를 구하는 과정은 다음과 같다. 우선 λ_s 를 매우 작은 값(0.01)으로 설정하여 현재 프레임의 왜곡(D_s)과 비트(R_s)를 위에서 설명한 양자화 파라미터를 구하는 방법을 통해서 구한다. λ_l 를 매우 큰 값(1000)으로 설정하여 현재 프레임의 왜곡(D_l)과 비트(R_l)를 위의 양자화 파라미터를 구하는 방법을 통해서 구한다. 구해진 왜곡과 비트에 의해 새로운 λ 값을 식 (16)을 통하여 구한다.

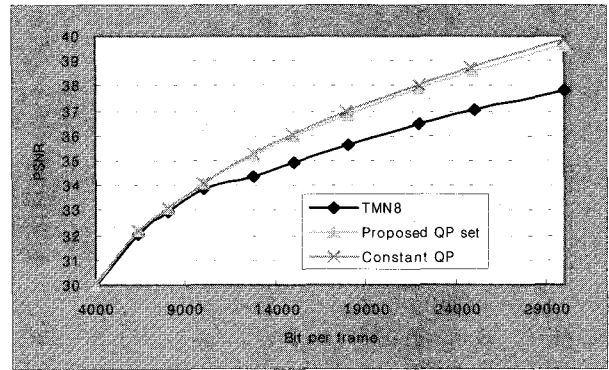
$$\lambda = \frac{D_l - D_s}{R_s - R_l} \quad (16)$$

$D(\lambda)$ 와 $R(\lambda)$ 를 양자화 파라미터를 구하는 방법을 통해서 새로이 구한 후, $R(\lambda) > R_T$ 일 경우, $D_l = D(\lambda)$, $R_l = R(\lambda)$ 로 대체하고, 그렇지 않은 경우, D_s 와 R_s 값을 대체하여 식 (16)에 의해 새로운 λ 를 $R(\lambda) = R_T$ 가 될 때까지 반복하여 수행한다. $R(\lambda) = R_T$ 를 만족하게 되는 λ 가 최적의 라그랑지 멀티플라이어, λ^* 가 되며, 이 때의 양자화 파라미터 조합이 실제 현재 프레임을 부호화하는 양자화 파라미터가 된다.

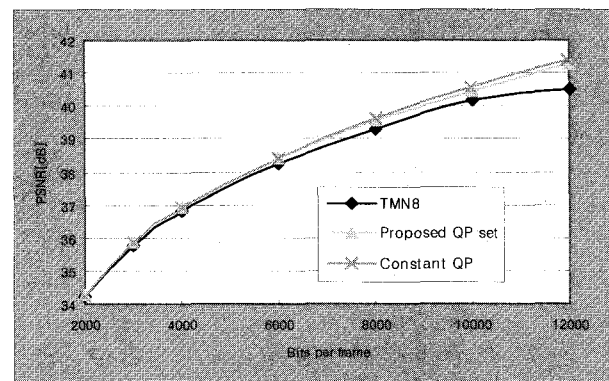
V. 실험

본 논문에서 제안한 선형 모델에 기반한 매크로블록 레이어 비트-율 제어 기법의 성능을 검증하기 위해서 모델 기반 비트-율 제어 기법 중 하나인 TMN8과 선형 모델에 기반한 고정 양자화 기법과 비교하였다. 실험은 MPEG-4 Simple profile을 기준으로 하여 I 프레임에 이어 P 프레임만이 연속되는 구조로, QCIF 영상을 대상으로 10f/s의 프레임-율로 낮은 비트-율에서 실험하였다. B 프레임의 경우, P 프레임과 마찬가지로 IV장에서의 서술한 방식에 의해 적용할 수 있으므로 특별한 실험 결과를 포함하지 않았다.

TMN8과 비교하면 같은 비트를 사용하여 전체적인 PSNR이 평균 0.7 dB에서 최고 1.5 dB까지 개선됨을 알 수 있으며, Foreman 영상의 갑작스런 변화가 일어날 때 비트의 변동을 막음을 알 수 있다. 선형 모델에 기반한 고정 양자화 파라미터 방식과 비교하면 평균 PSNR은 거의 차이가 없으면서 비트율의 변동 폭을 상당히 줄이는 것을 볼 수 있다. 입력 영상의 움직임이나 복잡도의



(a)



(b)

그림 4. R-D 그래프. (a) Foreman 시퀀스, (b) Susie 시퀀스
Fig. 4. The cumulative cost function with the condition of the Dquant.

변화가 심한 Foreman 영상에서 Susie 영상에 비해 효과적이다.

그림 4는 Foreman 영상과 Susie 영상의 목표 비트-율을 변화시키면서 R-D 특성을 나타내는 그래프이다. Foreman 영상의 경우 비트-율을 40kbps에서 300kbps까지 변화시켰으며, 상대적으로 복잡도가 적은 Susie 영상의 경우 20kbps에서 120kbps까지 변화시키며 실험하였다. 그림 4에서 알 수 있듯이 제안 알고리즘은 선형 모델에 기반한 비트-율 제어 기법에 비해 PSNR 차이가 거의 없으며, TMN8 비트-율 제어 기법에 비해서 평균 0.7dB에서 최고 1.5 dB까지 화질을 개선함을 알 수 있다. 그림 4에서와 같이 제안 알고리즘이 선형 모델에 기반한 고정 양자화 기법에 비해 PSNR이 떨어지는 경향이 나타난다. 고정 양자화 기법에서는 모델 파라미터의 조정이 프레임 단위로 일어나며, 제안 알고리즘에서는 매크로블록 마다 모델 파라미터의 변동이 일어난다. 매크로블록 마다의 특성 변화가 프레임간의 특성 변화보다 심해서 정확한 특성 변화를 따라가지 못하기 때문이

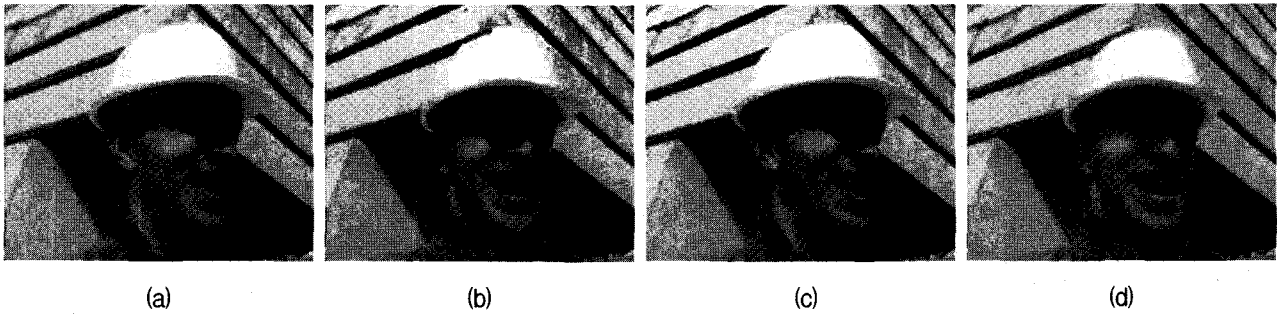


그림 6. Foreman 결과 영상 비교. (a) TMN8 38th 프레임, (b) TMN8 39th 프레임, (c) 제안 알고리즘 38th 프레임, (d) 제안 알고리즘 39th 프레임

Fig. 6. The result image of Foreman sequence. (a) TMN 38th frame, (b) TMN 39th frame, (c) The proposed algorithm 38th frame, (d) The proposed algorithm 39th frame.



그림 7. Foreman 결과 영상 비교. (a) TMN8 62th 프레임, (b) TMN8 63th 프레임, (c) 제안 알고리즘 62th 프레임, (d) 제안 알고리즘 63th 프레임

Fig. 7. The result image of Foreman sequence. (a) TMN 62th frame, (b) TMN 63th frame, (c) The proposed algorithm 62th frame, (d) The proposed algorithm 63th frame.

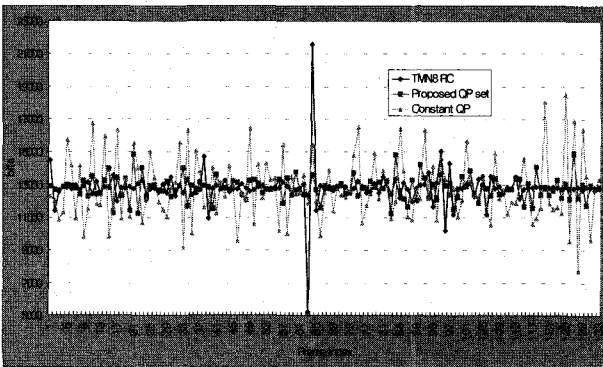


그림 5. 프레임 별 비트 발생 양(Foreman, 128kbps)

Fig. 5. Generated bits for each frame. (Foreman, 128kbps)

다. 그림 5는 128kbps에서의 Foreman 영상을 10f/s로 부호화한 경우 부호화되는 프레임 별 비트 양을 나타내는 그래프이다. Foreman 영상의 경우 장면의 변화가 심하고 움직임이 많이 존재하므로 프레임 별 비트의 변화가 심하다. 제안 알고리즘이 선형 모델에 기반한 고정 양자화 파라미터 비트-율 제어 기법에 비해 프레임 별 비트 변형이 훨씬 덜하고, 63, 64 번째 프레임에서 TMN8 비트-율 제어 기법에서는 비트 변동이 심하나

제안 알고리즘에서는 비트 변동을 줄이는 것을 볼 수 있다.

그림 6과 7에서는 주관적으로 화질을 비교하기 위해서 이전 프레임의 화질이 거의 동일한 특성을 보이는 반면 현재 프레임의 화질이 특히 차이는 39와 63번째 프레임을 예로 들어 보여주고 있다. 그림 6에서 38번째 프레임의 화질은 TMN8과 제안 알고리즘에서 거의 동일하나 39번째 프레임의 화질 면에서 얼굴, 특히 눈 주변에서 제안 알고리즘의 성능이 우월함을 보여주고 있다. 그림 7에서도 마찬가지로 63번째 프레임의 화질 면에서 얼굴의 윤곽선과 눈 주변에서 제안 알고리즘의 성능이 우월함을 보여주고 있다. 그림 6과 7에서 알 수 있듯이 제안 알고리즘에 의해 화질적인 성능이 TMN8에 비해 증가하는 것을 알 수 있다.

VI. 결 론

본 논문에서는 매크로블록 기반의 영상 압축 표준에서 양자화 파라미터 추정을 통한 비트-율 제어 알고리즘을 제시하였다. 기존 모델 기반 알고리즘에서는 영상

의 통계적 특성을 이용한 정해진 모델을 사용하나 이는 실제 비트-왜곡과는 상당한 차이를 보인다. 이런 통계적 특성의 한계를 극복하기 위해 선형 모델 기반의 비트율 제어 기법이 제안 되었다. 선형 모델 기반의 고정 양자화 파라미터 기법은 비트율의 변동이 심한 단점이 있다. 본 논문에서 제안하는 선형 모델 기반 비트율 제어 기법은 라그랑지안 최적화 기법을 Dquant 라는 제약을 극복하는 새로운 비터비 알고리즘을 제안하여 선형 모델을 사용하는 매크로블록 별로 양자화 파라미터를 설정할 수 있도록 하였다. 제안 알고리즘은 TMN8 기법에 비해서 비트당 화질을 개선하는 한편 선형 모델 기반의 고정 양자화 파라미터에 비해 비트율의 변동을 줄이는 효과가 있다.

참 고 문 헌

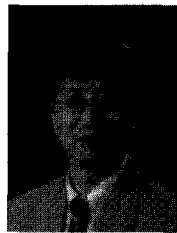
- [1] J. Ribas-Cobera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. Circuits and Systems for Video Tech.*, Vol. 9, no. 1, pp. 172-185, Feb. 1999.
- [2] T. Chiang and Y. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits and Systems for Video Tech.*, Vol. 7, no. 1, pp. 246-250, Feb. 1997.
- [3] J. Ribas-Cobera and S. Lei, "A frame-layer bit allocation for H.263+," *IEEE Trans. Circuits and Systems for Video Tech.*, Vol. 10, no. 7, pp. 1154-1158, Oct. 2000.
- [4] L. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *IEEE Trans. Circuits and Systems for Video Tech.*, Vol. 8, no. 4, pp. 446-459, Aug. 1998.
- [5] Z. He and S. K. Mitra "A linear source model and a unified rate control algorithm for DCT video coding," *IEEE Trans. Circuit and Systems for Video Tech.*, Vol. 12, no. 11, pp. 970-982, Nov. 2002.
- [6] Z. He and S. K. Mitra "Optimum bit allocation and accurate rate control for video coding via ρ -domain source modeling," *IEEE Trans. Circuit and Systems for Video Tech.*, Vol. 12, no. 10, pp. 840-849, Oct. 2002.
- [7] S. H. Hong and S. J. Yoo, "Rate control of MPEG Video for consistent picture quality," *IEEE Trans. Broadcasting*, Vol. 49, no. 1, pp. 1-13, Mar. 2003.
- [8] A. Ortega and K. Ramchandran, "Optimal trellis-based buffered compression and fast approximations," *IEEE Trans. Image Processing*, Vol. 3, no. 1, pp. 26-40, Jan. 1994.
- [9] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acousitics, Speech and Signal Processing*, Vol. 36, no. 9, pp. 1445-1453, Sept. 1988.
- [10] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. Image Processing*, Vol. 2, pp. 160-175, Apr. 1993.
- [11] D. Seo and S. Han, "Optimal quantization parameter set for MPEG-4 bit-rate control," *IEICE Trans. Communications*, Vol. E87-B, no. 11, pp. 3338-3342, Nov. 2004.
- [12] ITU-T, Video coding for low bit rate communication, Draft, ITU-T Recommendation H.263, Sept 1997.
- [13] ISO/IEC, Information technology - Coding of audio-visual objects - Part 2 : Visual, ISO/IEC 14496-2, Dec. 1999.

저 자 소 개



서 동 완(학생회원)
 1999년 연세대학교대 전기공학과
 학사 졸업.
 2001년 연세대학교 전기전자공학
 과 석사 졸업.
 2001년~현재 연세대학교 전기
 전자공학과 박사과정
 재학.

<주관심분야 : 영상신호처리, 동영상 압축, 비트
 율 조정 기법, 비트율-왜곡(R-D) 최적화>



최 윤 식(정회원)
 1979년 연세대학교 전기공학과
 학사 졸업.
 1984년 Case Western Reserve
 Univ. 시스템공학과 졸업.
 1987년 Pennsylvania State Univ.
 전기공학과 석사 졸업.

1990년 Purdue Univ. 전기공학부 박사 졸업
 1990년~1993년 (주)현대전자 산업전자 연구소
 책임 연구원

1993년~현재 연세대학교 전기전자공학부 정교수
 <주관심분야: 영상신호처리, HDTV, 동영상 압
 축>