

2단계 최적화 방법을 이용한 비트할당 기법

정회원 김 옥 중*, 김 성 대*

A Novel Bit Allocation Method Using Two-phase Optimization Technique

Wook Joong Kim*, Seong Dae Kim* *Regular Members*

요 약

비트할당 문제는 영상을 손실 부호화 하는데 있어서 가장 기본적인 것으로, 최대 허용 비트량을 초과하지 않으면서도 복원 영상의 발생 왜곡정도를 최소화 하는 것을 목적으로 하는 것이다. 이러한 비트할당 문제의 최적해를 구하는데 있어서 기존의 최적 비트할당 방법들은 동적 프로그래밍(dynamic programming) 방식에 근간하기 때문에 영상의 수가 증가하거나 움직임 보상을 이용하는 동영상 부호화의 경우에는 해를 구하기 위해서는, 그 방식이 요구하는 계산량과 메모리량이 너무 많아 최적해를 구하는데 있어서 그 한계가 있다. 본 논문에서는 기존의 최적 비트할당 방법들이 지니는 문제점들을 해결하기 위하여 '비트할당 문제의 분할'과 '2단계 최적화 방법'을 이용하여 기존의 최적 비트할당 방법들에 비해 적은 계산량과 메모리를 요구하면서도 최적 비트할당 결과를 얻을 수 있는 방법을 제안한다. 또한 제안한 방식을 효과적으로 근사화 할 수 있는 방법에 대해서도 제안한다. 최적 비트할당 결과는 영상 부호화 시스템의 분석, 복원화질의 최고치 파악, 효과적인 근사화 방법의 개발등에 효과적으로 이용될 수 있다. 따라서 본 논문에서 제안하는 방법은 비트할당에 관련된 분야에 효과적으로 이용될 수 있으리라 기대된다.

ABSTRACT

In this work, we propose a novel bit allocation method that is to minimize overall distortions subject to the bit rate constraint. We partition the original bit allocation problem into 'macroblock level bit allocation' problems that can be solved by conventional Lagrangian multiplier methods and a 'frame level bit allocation' problem. To tackle the frame level problem, 'two-phase optimization' algorithm is used with inter-frame dependency model. While the existing approaches are almost impossible to find the macroblock-unit result for the moving picture coding system due to high computational complexity, the proposed algorithm can drastically reduce the computational loads by the problem partitioning and can obtain the result close to the optimal solution.

Because the optimally allocated results can be used as a benchmark for bit allocation methods, the upper performance limit, or a basis for approximation method development, we expect that the proposed algorithm can be very useful for the bit allocation related works.

* 한국과학기술원 전기 및 전자공학과
 論文番號 : 98016-0113
 接受日字 : 1998年 1月 13日

I. 서 론

영상을 손실 부호화 하는데 있어서 가장 큰 관심사는 적은 비트량으로도 좋은 복원 화질을 만들 수 있도록 하는데 있다. 이러한 개념은 널리 알려진 비트할당 문제(bit allocation problem)로 대표되는데, 이것은 손실 부호화의 가장 기본적인 문제로 주어진 사용 가능한 총 비트량을 주어진 신호원들에 효과적으로 분배하여 전체의 왜곡을 최소화 하는 것을 의미한다. 예를 들어 DCT 부호화 된 정지 영상에서의 블록별 비트할당 문제라던가, 비디오 시퀀스(sequence)에서 프레임 당 비트할당, 대역별 부호화(subband coding)에서의 각 대역별 비트할당의 경우를 들 수 있다.

동영상 부호화에서의 비트할당 관련 연구들은 실제 시스템에의 적용을 위해 적은 계산량과 빠른 처리속도를 지니면서도 복원된 영상의 화질을 향상시키는 것을 목적으로 하는 방법들[1-6]과 비트할당 문제를 조건이 있는 최적화 문제(constrained optimization problem)로 고려하여 주어진 입력 영상에 대하여 최적의 비트할당 결과를 구해내는 방식들[7-12]로 구분될 수 있다. 첫번째 종류의 방식들은 실시간 구현에 적합하지만 그 제안 방식들의 근거 제시가 미흡하고, 그 방법의 타당성에 대해 불확실한 측면이 있다. 이에 비해 최적화에 기반을 둔 접근들은 해를 구하기 위한 계산량이 비교적 많아 부담이 되지만, 주어진 영상 부호화 시스템에 대한 복원 화질의 최고치(upper limit)를 파악한다는 점과, 이 결과를 바탕으로 부호화 시스템의 분석, 효과적인 근사화(approximation) 방식의 개발에서의 성능 평가 기준, 그리고 여러 비트할당 방법들의 벤치마크(benchmark)로의 역할을 할 수 있고 실제적인 적용에 있어서 영상정보의 저장 등 실시간 처리가 필요 없는 경우에 대하여 적용이 가능하기 때문에 그 의미를 지니게 된다. 본 논문에서는 효과적인 최적 비트할당에 관한 내용이다.

기존의 최적 비트할당 방법으로 널리 알려진 방법으로는 라그랑지안 멀티플라이어(Lagrangian multiplier)를 이용하여 비트할당 문제를 '제한이 없는 문제(unconstrained problem)'으로 변환 시키고 임의의 라그랑지안 멀티플라이어(Lagrangian multiplier)에 대해 동적 프로그래밍(dynamic programming)을 이용하여 최적 결과를 구하는 방법이 대표적이다[7-9]. 그런데 이러한 최적화 방법은 가능한 해의 경우의 수가 증가함

에 따라 동적 프로그래밍(dynamic programming)을 수행하기 위한 복잡도가 기하급수적으로 증가하기 때문에 그 한계가 존재한다. 그 일 예로 표준화된 동영상 부호화 시스템(MPEG-x[13][14], H.26x[15][16])의 최종 비트할당은 모든 매크로블록(macroblock, MB)에 허용 가능한 양자화기들 중 하나를 선택하는 것으로 구성이 되는데, 이 경우 입력 MB들의 상호 상관성과 시간축 중복성(temporal redundancy) 제거를 위한 움직임 보상(motion compensation)에 의하여 발생하게 되는 MB별 상호 의존성(dependency) 등으로 인하여 기존의 최적 비트할당 방법들이 요구하는 계산량을 감당하기가 곤란하다. 따라서, 기존의 방법으로는 움직임 보상이 없는 경우에 대한 최적 비트할당이라던가, 움직임 보상을 하는 경우라도 각각의 영상에 하나의 동일한 양자화기만을 할당하는 경우 등 비교적 단순화된 경우들에만 적용을 하고 있다. 결국 기존의 최적 비트할당 방법들은 앞에서 제시한 최적 비트할당 결과의 의미를 이용하기 위한 목적에서 접근하였으나, 과도한 계산량의 요구로 인하여 그 의미를 만족시킬 수 있는 결과를 구하는 데는 성공하지 못하고, 단지 최적해를 구할 수 있는 방법의 제안이라는 수준에 머물러 있다.

본 논문에서는 이러한 문제점들을 해결하기 위하여 비트할당 문제의 분할과 2단계 최적화 방법(two-phase optimization technique)[17]을 이용하여 기존의 방법들에 비해 적은 계산량과 메모리를 요구하면서도 최적 비트할당 결과를 얻을 수 있는 방법을 제안한다. 또한 제안한 방식을 효과적으로 근사화 할 수 있는 방법에 대해서도 제안한다. 제안 방식은 기존의 방식들에 비해서는 많은 계산량을 줄일 수 있지만, 제안 방식이 요구하는 계산량도 그 양을 실제로 수용하기에는 무리가 있을 수 있기 때문에, 근사화하지 않은 경우의 해에 근접하면서도 계산량을 효과적으로 감소시킬 수 있는 근사화 방안이 요구되기 때문이다. 본 논문의 구성은 다음과 같다. 일단 제 2 장에서는 비트할당문제의 수학적 표현에 대하여 서술하고, 문제의 세분화를 통한 새로운 표현방법에 대하여 설명한다. 그리고 본 논문에서 이용하고자 하는 2 단계 최적화 방법에 대하여 설명한다. 제 3 장에서는 최적 비트할당을 위하여 가정하는 모델들을 설명하고 비트할당 알고리즘에 대하여 설명한다. 제 4 장에서는 제안하는 방법을 적은 계산량 만으로도 효과적으

로 근사화 하는 방법에 대하여 제시한다. 제 5 장에서 본 논문에서 제안하는 방법들의 비트할당 결과를 나타내고, 제 6 장에서 결론을 맺는다.

II. 비트할당 문제의 분할과 2 단계 최적화 방법

영상 부호화 시스템에서 주어진 최대 허용 비트량을 초과하지 않으면서 복원영상의 왜곡을 최소화 하는 것을 목적으로 하는, 일반적인 비트할당 문제는 식(1)와 같이 조건이 있는 최적화 문제로 표현될 수 있다.

$$\begin{aligned} & \text{minimize } \sum_{i=1}^N \sum_{j=1}^M d(i, j) \\ & \text{subject to } \sum_{i=1}^N \sum_{j=1}^M r(i, j) \leq R_{budget} \end{aligned} \quad (1)$$

단, $N =$ 입력된 영상(frame)의 수,
 $M =$ 하나의 영상에 있는 MB의 수,
 $r(i, j) =$ i 번째 영상의 j 번째 MB에 발생하는 비트량,
 $d(i, j) =$ i 번째 영상의 j 번째 MB에 발생하는 왜곡량,
 $R_{budget} =$ 최대 허용 가능한 비트량

이렇게 표현된 비트할당 문제는 $M \times N$ 개의 변수에 대한 특정 값을 구하는 문제이고, 입력 영상의 수가 증가함에 따라 구해야 하는 변수의 수가 M 배씩 증가하게 된다. 바로 이러한 증가가 기존의 최적화 방법의 적용을 용이치 않게 하는 가장 큰 이유이다.

식(1)의 조건식을 만족시키는 $r(i, j)$ 에 대하여 다음의 조건을 만족시키는 R_i 를 다음과 같이 정의할 수 있다.

$$\sum_{j=1}^M r(i, j) \leq R_i \quad \text{and} \quad \sum_{i=1}^N R_i \leq R_{budget} \quad (2)$$

그런데, $\sum_{j=1}^M d(i, j)$ 의 값이 $r(i, 1), r(i, 2), \dots, r(i, M)$ 의 값에 의하여 결정되므로 새로운 값 D_i 를 식(3)과 같이 정의할 수 있다.

$$D_i = \min \left\{ \sum_{j=1}^M d(i, j) \right\} \quad \text{subject to } \sum_{j=0}^M r(i, j) \leq R_i \quad (3)$$

따라서, 식(1)과 같이 표현된 비트할당 문제는 식(2)와 식(3)을 이용하면 식(4)와 같이 새로이 표현될 수 있다.

$$\text{minimize } \sum_{i=1}^N D_i \quad \text{subject to } \sum_{i=1}^N R_i \leq R_{budget} \quad (4.a)$$

$$\text{단, } D_i = \min \left\{ \sum_{j=1}^M d(i, j) \right\} \quad \text{subject to } \sum_{j=1}^M r(i, j) \leq R_i \quad (4.b)$$

즉, 식(1)은 전체 입력 영상을 MB 단위로 분리하여 비트할당 문제를 표현한 형태인 반면, 식(4)의 경우는 프레임 단위에서의 비트할당 문제를 표현하는 식(4.a)와 하나의 프레임에 포함되는 MB들에 대한 비트할당을 표현하는 식(4.b)로 분리한 것이 된다. 따라서, 비트할당 문제의 해결은 식(4)의 해를 구하는 것과 동일한 것이다. 그런데, 임의의 R_i 값이 주어진 경우 식(4.b)의 해는 기존의 최적화 방식들[7-9]에 의하여 구할 수 있기 때문에, 결국 N 개만의 인수를 가지는 프레임별 비트할당 문제인 식(4.a)만을 해결하면 전체 비트할당 문제를 해결할 수 있다.

본 논문에서는 식(4.a)의 제한이 있는 최적화 문제를 해결하기 위하여 이단계 최적화 방법을 이용하였다. 일반적으로 제한이 있는 최적화 문제(constrained optimization problem)는 다음과 같이 표현된다. 벡터 \vec{x} 를 인수로 지니는 $f(\vec{x})$ 에 대해서,

Minimize $f(\vec{x})$ subject to constraints

$$\begin{aligned} g_1(\vec{x}) &\leq 0, \dots, g_r(\vec{x}) \leq 0 \\ h_1(\vec{x}) &= 0, \dots, h_m(\vec{x}) = 0 \end{aligned} \quad (5)$$

단, f, g_i 그리고 h_j 는 R^n (n 차원 실수 공간)에서 정의되는 함수이고, $m \leq n$ 이다.

이러한 제한이 있는 문제의 해결방법으로 Maa와 Shanblatt[17]은 식(6)과 (7) 같은 2단계 탐색(two phase search) 방식을 제시하였다. 이 방법은 반복 수행시 i 번째에 구한 값 \vec{x}_i 에 대하여 다음 반복 수행이 결정되는 \vec{x}_{i+1} 값이 $\vec{x}_{i+1} - \vec{x}_i = \lim_{\substack{\vec{x} \rightarrow \vec{x}_i \\ dt \rightarrow 0}} \dot{\vec{x}} \cdot dt$ 의 관계로 구

성될 때, $\dot{\vec{x}}$ 값을 미리 결정된 변환시간(switching time)

t_1 에 의하여 다음과 같이 결정하는 것이다.

(1) 제 1 단계 (for $0 \leq t \leq t_1$):

$$\begin{aligned} \dot{\vec{x}} = & -\nabla f(\vec{x}) - s \left[\sum_{i=1}^N g_i^+(\vec{x}) \nabla g_i(\vec{x}) \right. \\ & \left. + \sum_{j=1}^M h_j(\vec{x}) \nabla h_j(\vec{x}) \right] \end{aligned} \quad (6)$$

단, s 는 충분히 큰 양의 실수, $g_i^+ = \max(0, g_i)$

(2) 제 2 단계 (for $t \geq t_1$):

$$\begin{aligned} \dot{\vec{x}} = & -\nabla f(\vec{x}) - \left[\sum_{i=1}^N g_i(\vec{x})(s g_i^+(\vec{x}) + \lambda_i) \right. \\ & \left. + \sum_{j=1}^M \nabla h_j(\vec{x})(s h_j(\vec{x}) + \mu_j) \right] \end{aligned} \quad (7)$$

단, $\dot{\lambda} = \epsilon \cdot s g_i^+$, $\dot{\mu} = \epsilon \cdot s h_j$, ϵ 은 작은 양의 수.

이 방법은 일반적인 gradient 방식들을 지나는 infeasibility 문제¹⁾를 해결할 수 있는 특징을 지니는 것으로, 제 1 단계에서 구한 해는 'Penalty function theorem' [18]에 근거하여 s 가 무한대이기 전에는 최적해일 수 없기 때문에 반복이 진행됨에 따라 최적 해에 근접하는 infeasible 영역의 한 점으로 수렴되는 결과를 나타내고, 제 2 단계에서는 이전 단계에서 구한 점을 초기값으로 보다 미세한 탐색을 통하여 $t \rightarrow \infty$ 함에 따라 feasible 한 최적해로 수렴하게 된다.

III. 2단계 최적화 방식을 이용한 비트할당 문제의 해결

식(4.a)는 앞 절에서 제시한 2단계 최적화 방식에 대하여 식(8)인 경우로 고려될 수 있다.

$$\vec{x} = [R_1, R_2, \dots, R_N]^T \quad (8.a)$$

$$f(\vec{x}) = \sum_{i=1}^N D_i \quad (8.b)$$

$$g(\vec{x}) = \sum_{i=1}^N R_i - R_{budget} \quad (\because r = 1) \quad (8.c)$$

1) 해가 존재할 수 있는 영역 밖의 값을 최적해로 판단하는 것. 일반적인 미분을 이용하는 최적화 방법들은 항상 존재한다는 것을 보장하지 못한다[18][19].

$$h(\vec{x}) = 0 \quad (8.d)$$

따라서, $\nabla g(\vec{x})$ 는 $g(\vec{x})$ 의 정의에 의하여 식(9)와 같이 구해지고,

$$\nabla g(\vec{x}) = [1 \dots 1]^T \quad (9)$$

D_i 가 R_1, \dots, R_i 를 인수로 가지는 함수 이므로 $\nabla f(\vec{x})$ 는 식 (10)과 같이 정의 된다.

$$\nabla f(\vec{x}) = \begin{bmatrix} \frac{\partial D_1}{\partial R_1} + \frac{\partial D_2}{\partial R_1} + \dots + \frac{\partial D_{N-1}}{\partial R_1} + \frac{\partial D_N}{\partial R_1} \\ \frac{\partial D_2}{\partial R_2} + \frac{\partial D_3}{\partial R_2} + \dots + \frac{\partial D_N}{\partial R_2} \\ \frac{\partial D_3}{\partial R_3} + \dots + \frac{\partial D_N}{\partial R_3} \\ \vdots \\ \frac{\partial D_N}{\partial R_N} \end{bmatrix} \quad (10)$$

실제적으로는 R_i 의 값이 비연속적(discrete)인 값이기 때문에 $\frac{\partial D_i}{\partial R_j}$ ($N \geq i \geq j \geq 1$)를 정의할 수 없지만, R_i 값이 일정한 범위 안에서 충분히 많은 경우의 값을 지니므로 정보이론[19]에 근거하여 식(11) 연속함수의 형태로 왜곡률(distortion-rate) 관계를 가정할 수 있다.

$$D_i = \hat{\sigma}_i^2 \cdot H_i(R_i) \quad (11)$$

where, $H_i(R_i) = a_i \exp[-b_i R_i]$, a_i, b_i 는 임의의 상수

단, $\hat{\sigma}_i^2$ 는 i 번째 영상과 $i-1$ 번째 복원 영상간의 움직임 보상 오차.

그런데, Cheng, Li, Kuo[20]은 $\hat{\sigma}_i^2$ 이 식(12)과 같이 이전 원 영상과 현재 영상의 움직임 오차와 이전 영상의 복원 오류의 선형 결합의 형태로 표현될 수 있다고 제안하였다.

$$\hat{\sigma}_i^2 = \sigma_i^2 + a_i \cdot D_{i-1} \quad (12)$$

단, σ_i^2 는 현재 프레임과 이전 원 영상과의 움직임 보상 오차.

이러한 모델링을 기반으로 하면 $\frac{\partial D_i}{\partial R_i}$ ($N \geq j \geq i \geq 1$)는 식(13)과 같이 정의할 수 있다.

$$\begin{aligned} \frac{\partial D_i}{\partial R_i} &= \hat{\sigma}_i^2 \cdot \frac{\partial H_i(R_i)}{\partial R_i} = -\hat{\sigma}_i^2 \cdot a_i \cdot b_i \cdot \exp[-b_i R_i] \\ \frac{\partial D_{i+1}}{\partial R_i} &= \frac{\partial}{\partial R_i} [\hat{\sigma}_{i+2}^2 \cdot H_{i+1}(R_{i+1})] \\ &= \frac{\partial}{\partial R_i} [(\sigma_{i+1}^2 + a_{i+1} \cdot D_i) \cdot H_{i+1}(R_{i+1})] \\ &\quad (\because \hat{\sigma}_{i+1}^2 = \sigma_{i+2}^2 + a_{i+1} \cdot D_i) \\ &= \frac{\partial}{\partial R_i} [\sigma_{i+1}^2 \cdot H_{i+1}(R_{i+1}) + a_{i+1} \cdot H_{i+1}(R_{i+1}) \cdot D_i] \\ &= a_{i+1} \cdot H_{i+1}(R_{i+1}) \cdot \frac{\partial D_i}{\partial R_i} \quad (\because \frac{\partial H_{i+1}(R_{i+1})}{\partial R_i} = 0) \\ &\vdots \\ \frac{\partial D_j}{\partial R_i} &= a_j \cdot H_j(R_j) \cdot \frac{\partial D_{j-1}}{\partial R_i} \quad \text{where, } j > i \end{aligned} \quad (13)$$

따라서, 임의의 영상에서 R_i 와 D_i 의 관계식인 식(11)의 a_i, b_i 를 결정하고, 식(12)을 결정지어 주는 α_i 값을 구하면 식(13)을 이용하여 $\nabla f(\vec{x})$ 를 결정할 수 있다.

이상의 내용을 바탕으로 다음과 같은 비트할당 알고리즘을 구성할 수 있다.

- 단계 1. $R_i(1 \leq i \leq N)$ 에 임의의 초기값을 할당한다.
- 단계 2. 할당된 R_i 값에 대하여 각 프레임 내에서 식(4.b)를 만족시키는 값을 구한다. 그리고 적당한 수의 데이터 들을 이용하여 식(11)의 a_i, b_i 값을 결정하고, 움직임 보상 관계를 이용하여 현재 프레임에 적합한 a_i 값을 결정한다.
- 단계 3. 앞의 단계에서 구한 parameter 들과 식(13)를 이용하여 $\nabla f(\vec{x})$ 를 구하고, 식(6)의 관계를 이용하여 새로운 $R_i(1 \leq i \leq N)$ 값을 결정한다.
- 단계 4. $f(\vec{x})$ 값의 변화량이 임의의 문턱치 보다 적은 경우가 정해진 수만큼 반복되는 경우 다음 단계로 진행하고, 그렇지 않은 경우는 단계 2로 간다.

단계 5. 앞 단계에서 결정된 $R_i(1 \leq i \leq N)$ 에 대하여 단계 2의 동작을 수행하고 식(7)을 이용하여 제 2 단계 최적화 과정을 진행한다.

단계 6. 제 1 단계에서와 마찬가지로 $f(\vec{x})$ 값의 변화량이 임의의 문턱치 보다 적은 경우가 임의의 경우만큼 반복되는 경우 수행을 중지하고, 그렇지 않은 경우 단계 4로 진행한다.

이상의 과정을 통하여 반복을 계속해 나가면 결국 최적 비트할당에 근접하는 결과를 얻을 수 있다.

IV. 제안 최적 비트할당 방법의 근사화 방안

앞 절에서 제시한 비트할당 방법은 기존의 방법들 [7-9]이 요구하는 계산량을 상당히 감소시킬 수 있지만, 메모리량과 허용된 계산 시간의 한계가 존재하는 실제적인 적용에 있어서 제안 방법이 요구하는 계산량도 큰 부담이 될 수 있다. 따라서 적은 계산량을 지니면서도 최적해에 근접한 해를 구할 수 있는 효과적인 근사화 방안이 요구된다. 본 논문에서는 A.Ortega [8]에 의하여 제시된 것과 유사한 방법을 통하여 제안 방법을 근사화 하였다.

4.1 Sliding Windowing을 이용한 프레임별 비트할당 근사화

입력 프레임의 수가 증가함에 따라 요구되는 많은 계산량 부담을 해결하는 방안으로는 sliding window 형태의 분할 비트할당 방법을 이용할 수 있다. 즉, 식(4.a)의 해결을 위하여 전체 입력 프레임들을 중복이 존재하는 window 크기 만큼의 세그먼트(segment)들로 분리하여 각각에 대하여 비트할당 방법을 적용하고, 최종적으로는 각 결과들을 결합시키는 방법이다. 이런 식으로 분리하여 비트할당 하는 것은 프레임간의 거리가 멀어질수록 상호 연관성이 떨어지는 성질을 이용하는 것이다. 이러한 성질은 A.Ortega[8] 등에 의하여 실험을 통하여 확인 되었다.

그림 1은 본 논문에서 이용한 방법을 간략히 표현한 것이다. 일단 전체 입력 영상들을 몇 개의 중복이 존재하는 부분들로 분리한 후, 각각에 제안하는 비트할당 방법을 적용하여 그 결과를 구하게 된다. 그림에서는 각각의 입력 영상을 비어있는 사각형으로 표시하였고 전체 영상을 5 개의 세그먼트(segment)들로

분리한다고 가정 하였다. 각 세그먼트의 비트할당 결과를 오른쪽사선, 격자선, 왼쪽사선, 얇은 회색, 짙은 회색으로 표시하였고, 전체 프레임의 근사화 된 최종 비트할당 결과는 각 세그먼트의 처음 프레임 결과 부분과 마지막 프레임 결과 부분을 제외한 부분들이 되도록 하여 결정하였다.

실제 적용에 있어서도 세그먼트들이 중복되는 부분의 비트할당 결과에 대해서는 중복되는 영역의 반은 이전 세그먼트의 값으로 채우고, 그 나머지 영역은 다음 세그먼트의 값으로 채워 근사화 하였다.

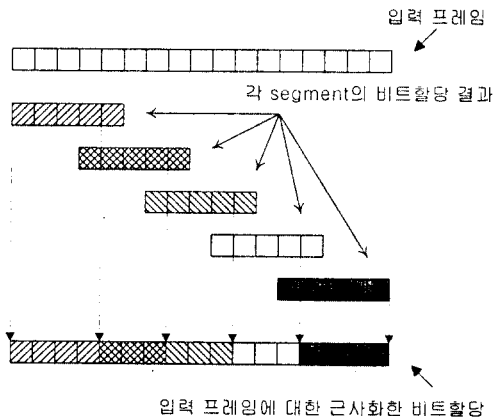


그림 1. Windowing을 통한 프레임간 비트할당의 근사화
Fig. 1 Approximation of frame level bit allocation using windowing technique.

4.2 Knapsack 문제 해결 알고리즘의 유사 노드 합병(similar node merging)을 통한 프레임 내에서의 비트할당 근사화

Knapsack 문제[22]는 각기 다른 크기와 가치를 가지는 T 가지 종류의 물건을 S 만큼의 크기를 가지는 가방에 담는 경우 가방에 담긴 물건의 가치를 최대화 하는 문제로 많은 경우의 정수 프로그래밍(integer programming)들은 Knapsack 문제로 치환되어 해결되어 진다. 이 문제의 대표적인 해결 방법은 동적 프로그래밍을 이용하는 것으로, 해를 구하기 위하여 필요한 시간은 $T \times S$ 에 비례한다.

그런데, 식(4.b)의 문제는 T 의 값이 MB에 대하여 허용된 양자화기의 수이고, S 가 R 인 경우의 Knapsack 문제로 고려할 수 있다. 따라서 Knapsack 문제 해결

방법의 적용이 가능한데, 일반적으로 비트할당의 경우 R 의 값이 비교적 큰 값이기 때문에 동적 프로그래밍을 위한 트리(tree)를 구성하면서 발생하는 노드들의 수가 증가하여 많은 계산시간과 메모리 양을 요구한다. 따라서, 본 논문에서는 임의의 단계에서 발생하는 노드의 수가 미리 결정한 최대값보다 큰 경우, 인접하는 노드들은 비슷한 이전 경로를 지닌다는 가정하에, 유사한 노드들을 하나로 합병시켜 발생 노드수를 제한하는 방식으로 근사화 하였다.

그림 2는 가방의 크기 $S=5$ 이고 가방에 담을 물건의 종류의 수 $T=3$ 인 경우에 해를 구하는 과정을 나타낸다. 각 물건은 t_1, t_2, t_3 이라 하였고, 물건들의 (크기, 가치) 값은 (1, 1), (2, 5), (3, 7)으로 설정하였다. 초기 상태에서는 가방이 비어 있기 때문에 가방에 담긴 물건의 크기 합이 0 인 상태에서 시작하여 임의의 다른 2가지 이상의 경로가 하나의 노드에서 만나는 경우에 대해서는 많은 가치를 가지는 경로를 선택하는 식으로 진행한다. 이렇게 진행하면 최종적으로 5 가지 경우에 대한 값을 얻을 수 있고, 이중 최적의 경우는 가방의 크기를 넘지 않으면서 가장 높은 가치를 지니는 경우인 (5, 12)를 선택하면 된다.

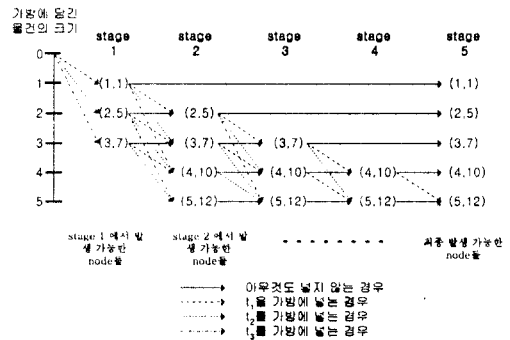


그림 2. Knapsack 문제 해결의 예.
Fig. 2 Example of Knapsack programming.

V. 시뮬레이션 결과

제안된 비트할당 방법을 H.263로 부호화 하는 경우에 적용해 보았다. R 의 초기값으로는 각 프레임별 평균 비트량으로 하였고, 2 단계 최적화를 위한 파라미

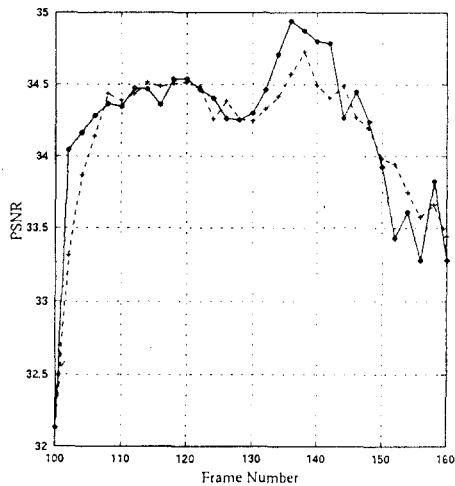
터들은 $s = 200$, $\epsilon = 0.5$, $dt = 0.001$ 로 설정하였다²⁾.

우선 제안 방식을 구현하는 경우 근사화 하는 경우와 그렇지 않은 경우에 대하여 비교해 보았다. 그림 3은 근사화 시 window의 크기를 10 frame으로 하고, window를 이동시킬 때 각 window 별 중복되는 프레임의 크기를 4로 하여, 표 1과 같은 경우의 영상에 대하여 제안 방식으로 비트할당 하는 경우의 복원 영상의 화질을 비교한 결과이다.

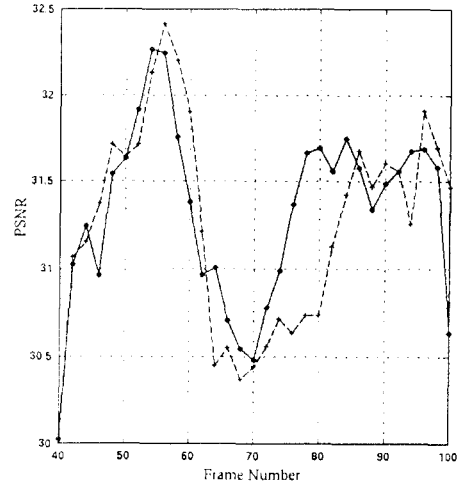
그림 3에서 보이듯이 제안 방식을 근사화 한 경우와 그렇지 않은 경우 복원 영상의 화질에는 큰 차이가 없음을 확인 할 수 있었다. 표 2는 복원 영상의 평균 화질을 비교한 결과이다.

표 1. 근사화 방안의 성능 비교를 위한 입력 영상
Table 1. Input images for the performance test of approximation method

입력 영상	Mother & Daughter	Foreman
영상 크기	QCIF	
전송 속도	30 Kbps	50 Kbps
프레임 전송률	15 frames/sec	
시작프레임 / 끝 프레임	100 / 160	40 / 100
부호화된 총 영상의 수	31	



(a) Mother & Daughter 영상의 경우



(b) Foreman 영상의 경우

그림 3. 제안방식을 근사화 하는 경우와 그렇지 않은 경우의 복원영상의 화질 비교.

(실선: 근사화 하지 않은 경우, 점선: 근사화 한 경우)

Fig. 3 The reconstructed image quality test for the approximation method.

(solid lines: not-approximated case, dashed lines: approximated case)

표 2. 비트할당 방식에 대한 복원 영상의 평균 PSNR 비교
Table 2. The comparison of average PSNR results between approximated case and not-approximated case.

영상종류 \ 방식	근사화 하지 않은 경우	근사화 한 경우
Mother & Daughter	34.272 dB	34.215 dB
Foreman	31.368 dB	31.216 dB

다음으로는 제안 방법의 성능 비교를 위하여 표 3의 영상들에 대하여 TMN6[21]에 의하여 비트할당 하는 경우와 제안 근사화 방법에 의하여 비트할당 하는 경우에 대하여 비교해 보았다³⁾. 그림 4는 각기 다른 입력 영상에 대한 복원 영상의 화질을 나타낸다. 결과에서도 알 수 있듯이, 최적해는 최소 TMN6에 비해서

3) 제안 방식은 최적화에 근거한 방법이기에 때문에 TMN6와의 비교가 적절치 않은 측면이 있으나, TMN6가 H.263 비트할당에 대한 가장 널리 알려진 방법이고, 또한 현재 적절한 비교 판단 기준이 부재하여 기타 다른 문헌들에서도 이 방법을 채택하는 등의 이유로 이 방식을 성능의 비교 기준으로 삼음.

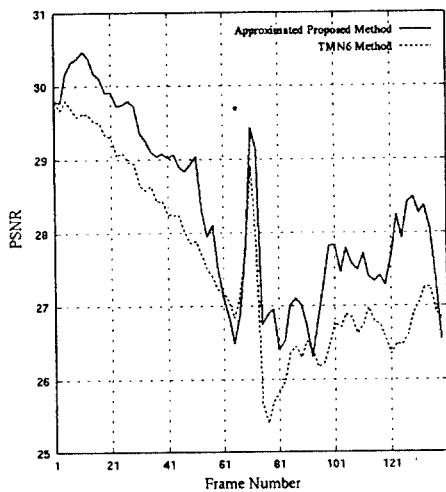
2) 이 값들은 일반적인 2단계 최적화 방법에서 초기값으로 일반적으로 사용되는 값들로 설정한 등임.

약 1dB 정도의 평균 향상을 얻을 수 있을 것으로 예측할 수 있다. 그리고 결과 그래프를 살펴보면, TMN6의 비트할당 방법은 이전 프레임에 할당된 비트량을 현재 프레임의 비트할당에 이용하기 때문에 프레임간 화질 변화가 크지 않지만, 제안 방식은 입력영상의 특성에 맞추어 비트할당을 하기 때문에 프레임간 화질 변화가 많은 편이다. 표 4는 각 비트할당 방법을 적용한 경우 복원 영상의 평균 화질을 나타낸다.

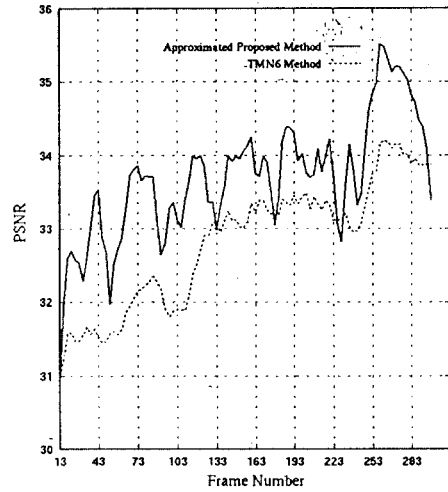
표 3. 제안 근사화 방법과 TMN6의 성능 비교를 위한 입력 영상.

Table 3. Input images for the performance comparison between the approximated proposed method and TMN6 method.

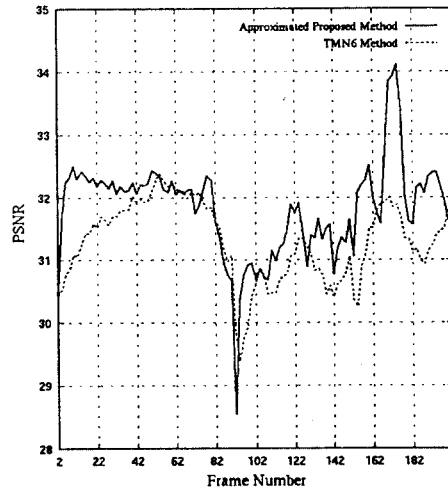
입력 영상	Coast Guard	Hall Monitor	News
영상 크기	QCIF	QCIF	QCIF
전송 속도	40 Kbps	25 Kbps	35 Kbps
프레임 전송률	15 frames/sec	10 frames/sec	15 frames/sec
시작프레임 / 끝 프레임	1/140	13/298	2/200
부호화된 총 영상의 수	70	96	100



(a) Coast Guard 영상의 경우



(b) Hall Monitor 영상의 경우



(c) News 영상의 경우

그림 4. 제안 근사화 방법과 TMN6에 의한 비트할당의 경우 복원 영상의 화질 (실선: 제안 근사화 방법, 점선: TMN6 방법)

Fig. 4 The reconstructed image quality comparison between the proposed approximation method and TMN6 method.

(solid lines: proposed approximation method, dashed lines: TMN6 method)

표 4. 제한 근사화 방법과 TMN6에 의한 비트할당의 경우 복원 영상의 평균 PSNR 비교

Table 4. The comparison of average PSNR results between the proposed approximation method and TMN6 method.

비트할당방식 영상종류	TMN6에 의한 비트할당	제한 근사화 방식에 의한 비트할당
Coast Guard	27.28 dB	28.28 dB
Hall Monitor	32.83 dB	33.72 dB
News	31.30 dB	31.83 dB

VI. 결 론

본 논문에서는 2 단계 최적화 방법을 이용하는 새로운 최적 비트할당 방법을 제안하였다. 기존의 최적 방법들은 구조적인 한계로 인하여 동영상 부호화와 같이 영상의 상호 상관성이 높은 경우를 완전히 해결하지 못하고 대단히 단순화된 환경에 대한 결과만을 제시한 반면, 제안 방식은 비트할당 문제의 분할과 2 단계 최적화 방법을 이용하여 그러한 문제점들을 해결하였다. 그리고, 제안 방식의 효과적인 적용을 위한 근사화 방법을 제시하였고, 이 방법의 성능 평가를 위하여 H.263 부호화 환경에 적용시켜 TMN6[21]의 경우와 비교해 보았다. 또한, 최적 비트할당 하는 경우의 성능 정도를 여러 실험 영상들에 대하여 분석해 보았다. TMN6에 비해서 최적 비트할당을 하는 경우는 평균적으로 1dB 정도를 향상 시킬 수 있는 것으로 분석 되었다.

최적 비트할당 방법은 많은 계산량이 부담이 되긴 하나, 시스템의 분석, 성능평가, 여러 비트할당 방법들 간의 benchmark로의 역할, 효과적인 근사화 방법의 제시 등에 이용될 수 있다. 따라서 기존의 방법들이 단순화된 환경에 대해서만 결과를 나타낸 것에 비해 본 논문에서 제안된 최적화 방법은 영상 부호화의 최소 단위인 MB에 대한 최적 양자화기 할당 결과를 구할 수 있으므로 기존의 방법들에 비해 보다 효과적으로 이용될 수 있다.

참 고 문 헌

1. ISO-IEC JTC1/SC29/WG11, Coding representation

of pictures and audio information, MPEG-2 test model 5, Apr. 1993.

2. 이팔진, 강인근, 외, 신경회로망을 이용한 비디오의 고정 비트율 제어에 관한 연구, 한국통신학회 논문지, vol. 20, no. 4, pp.877-883, Apr. 1995.

3. G. Keesman, I. Shah, R. Klein-Gunnewiek, Bit-rate Control for MPEG encoders, Signal Processing : Image Communication, vol. 6, issue 6, pp.545-560, Feb. 1995.

4. S.C. Liew, C. Tse, A Control-Theoretic Approach to Rate-Controlled Video Compression, ICIP vol II, pp.65-68, 1996.

5. T. Chiang, Y.Q. Zhang, A New Rate Control Scheme Using Quadratic Rate-Distortion Model, IEEE trans. on CSVT, vol. 7, no. 1, pp.246-250, Feb. 1997.

6. Jennifer J.H. Webb, K. Oehler, A Simple Rate-Distortion Model, Parameter Estimation, and Application to Real-time Rate Control for DCT-based Coders, ICIP vol. II, pp.13-16, 1997.

7. Y. Shoham, A. Gersho, Efficient Bit Allocation for an Arbitrary Set of Quantizers, IEEE trans. on ASSP, vol. 36, no. 9, pp.1445-1453, Sep. 1988.

8. A. Ortega, K. Ramchandran, M. Vetterli, Optimal Trellis-Based Buffered Compression and Fast Approximations, IEEE trans. on Image Processing, vol. 3, no. 1, pp.26-40, Jan. 1994.

9. K. Ramchandran, A. Ortega, M. Vetterli, Bit Allocation for Dependent Quantization with Applications to Multiresolution and MPEG Video Coders, IEEE trans. on Image Processing, vol. 3, no. 5, pp. 533-545, Sep. 1994.

10. K. M. Uz, M. Shapiro, M. Czigler, Optimal Bit Allocation in the presence of Quantizer Feedback, ICASSP, vol. V, pp.385-388, 1993.

11. J.J. Chen, D.W. Lin, Optimal Bit Allocation for Video Coding under Multiple Constraints, ICIP, vol. III, pp.403-406, 1996.

12. G.M. Schuster, A.K. Katsaggelos, Optimal Bit Allocation among Dependent Quantizers for the Minimum Maximum Distortion Criterion, ICASSP CD-ROM, 1997.

13. MPEG-1, ISO-IEC 11172, Coding of moving picture and associated audio - for storage at up to about 1.5 Mbits/s, Nov. 1992.
14. Recommendation ITU-T H.262, ISO/IEC 13818, Generic Coding of Moving Pictures and Associated Audio, Mar. 1995.
15. ITU-T Recommendation H.261, Video codec for audiovisual services at 64 - 1920 Kbit/s, 1993.
16. ITU-T SG15 Experts Group for Very Low Bitrate Visual Telephony, Draft Recommendation H.263, Mar. 1997.
17. C.Y. Maa, M.A. Shanblatt, A two-phase optimization neural network, IEEE trans. Neural Networks, vol. 3, no. 6, pp.1003-1009, Nov. 1992.
18. G.N. Vanderplaats, Numerical Optimization Techniques for Engineering Design with Application, McGraw-Hill, pp121-150, 1993.
19. T. Berger, Rate Distortion Theory, Prentice-Hall, 1971.
20. P. Cheng, J. Li, C. Kuo, Rate Control for an Embedded Wavelet Video Coder, IEEE trans. on CSVT, vol. 7, no. 4, pp.696-702, Aug. 1997.
21. ITU-T SG15 Experts Group for Very Low Bitrate Visual Telephony, "Video Codec test model, TMN6", Apr. 1996.
22. Robert Sedgewick, Algorithms 2nd Edition, Addison Wesley, 1988.



金 煜 中(Wook Joong Kim) 정회원
1989년 3월~1992년 2월: 한국과학기술원 전기 및 전자공학과 졸업(공학사)
1993년 3월~1995년 2월: 한국과학기술원 전기 및 전자공학과 졸업(공학석사)

1995년 3월~현재: 한국과학기술원 전기 및 전자공학과 졸업(공학사)

※주관심분야: 초저속 영상 압축 부호화, 영상 통신 시스템 등

e-mail : wjk@sdvision.kaist.ac.kr



金 聖 大(Seong Dae Kim) 정회원
1997년: 서울대학교 전자공학과 졸업(공학사)
1979년: 한국과학기술원 전기 및 전자공학과 졸업(공학석사)
1983년: 프랑스 INPT ENSEEITH 졸업(공학박사)

1994년~현재: 한국과학기술원 전기 및 전자공학과 교수

※주관심분야: 영상처리, 영상통신, 컴퓨터 비전, VLSI, 구현 등

e-mail : sdkim@ee.kaist.ac.kr