

비디오 코딩시 Lagrange 승수를 조정하여 적응 모드 선택에 따른 비트율의 제어

정회원 류철*, 김승필**

Efficient Rate Control by Lagrange Multiplier Using Adaptive Mode Selection in Video Coding

Chul Ryu*, Seung P. Kim** *Regular Members*

요 약

본 논문은 블록 단위의 비디오 코더 (video coder)에서 최적 모드들을 적절하게 선택할 수 있는 방법을 설명한다. 비트율의 제어에 관한 문제는 버퍼가 차 있는 양에 따라 양자화 레벨이 변하는 방법이 여러 학자들로부터 연구되어 왔다. 제시된 방법은 각각의 매크로블록(macroblock)에 관한 최적 모드만을 검색하지 않고 최적 결정 곡선을 검색한다는 점에서 이전의 것, 즉 [4]와는 다르다. 제시된 방법은 비트율의 제어를 위하여 운동 보상 또는 비 운동 보상 및 intra/inter 코딩에 관한 결정 방법까지 포함한다. 이러한 결정 방법은 주어진, 즉 고정된 결정 곡선을 이용하는 것 보다 주어진 프레임에서 설정된 비트율(bit-rate)에 따른 총체적인 왜곡을 최소화시키는 유동적인 결정 곡선을 이용하는 것이 더 바람직하다. 가열냉각 최적화 기술 (annealing optimization technique)의 모의실험을 통해 곡선을 도출할 수 있다. 알고리즘은 H.261 비디오 코덱(video codec) 범위에서 수행되고 모의실험 또한 그 범위에서 결과들을 보여주고 이들을 여타의 접근방법들과 비교한다.

ABSTRACT

This paper presents an approach for rate control by adaptively selecting macroblock modes in video coding. The problem of rate control has been investigated by many authors where quantizer level is adjusted based on the buffer fullness. The proposed approach is different from the previous ones [4] in that it finds the optimal decision curve rather than finding a set of the modes. Proposed algorithm extends the coding decision options for rate control to motion/no-motion compensation as well as inter/intra decisions. Instead of having a fixed motion/no-motion compensation or inter/intra decision curve, one can utilize an adaptive decision curve based on the characteristics of input frames so that the PSNR at a given bit rate is maximized.

Therefore, the proposed approach provides better rate control than simple quantizer feedback approach in terms of visual quality. The curve is obtained by utilizing simulated annealing optimization technique. The algorithm is implemented and simulations are compared with other approaches within H.261 video codec.

1. 서 론

공중 전화 교환망(PSTN)을 통한 전송과 같은 비디오 코딩 응용 프로그램들에게는 통신망의 간단한

* 동국대학교 정보통신공학과 (cryu@dgu.ac.kr)

** InterDigital Com. (skim@interdigital.com)

논문번호: 98213-0512, 접수일자: 1998년 5월 12일

※ 본 논문은 1999년도 동국대학교 신입교원연구비 지원으로 이루어 졌음.

구성 혹은 고정된 대역 폭을 위해서 일정한 비트율이 바람직하다. 비디오 신호들을 일정한 비트율로 전송하기 위해서는 코더 출력과 채널 사이의 버퍼를 사용하여 비트율의 변화 폭을 줄여야 한다. 이에 반해서, 코더 내에 해당 비율을 제어하게 되면 코더의 출력 비트율은 가변적인 코딩 품질에 상응하는 비율을 일정하게 유지할 수 있다. 최선의 품질/비용의 비율(quality/cost ratio)을 얻기 위해서는 이러한 두 가지 방법을 동시에 채택하여야 한다. 이 논문은 주어진 버퍼의 크기와 전송 비율에 따라 결정되는 용 비트율의 범위 내에서 코더의 비트율을 제어하는 접근방법을 제시한다.

표준 디코더 [10], [13], [16]과 호환되며, 설정된 채널 비트율(channel bit-rate)의 가시적인 품질(visual quality)을 극대화하기 위해 많은 방법이 소개되었다. [1], [2]에서는 버퍼에 차 있는 양에 따라 양자화 파라미터(quantization parameter)만 다루었다. 제시된 솔루션에서는 현재의 매크로블록(macro-block) 또는 블록의 서브그룹(sub-group of-block)^[8]에 관한 비트 카운트의 예상 값으로서 이전의 비트 카운트 값을 사용하며 이에 따라 각각의 양자화 레벨을 조절한다. 또 다른 접근방법은 일정한(stationary) 가정^[11] 하에서 현재의 매크로블록에 관한 비트의 갯수를 훈련열(training sequences)로부터 예상하기 위해서 어림 값을 사용한다. [4]에서는 제품 코드에 한정되는 비율(rate-constrained product code)의 공식을 만들어서 양자화 선택항목들의 결합을 최적화시킨다. 일반적으로 양자화가 자주 조정될수록 출력 비트율의 변화는 줄어들고 필요한 버퍼의 크기는 작아진다. 반면에, 양자화의 조정이 버퍼 예상 값의 도출 시간에 의존하게 되면 더 큰 버퍼 사이즈가 필요하게 된다. 버퍼의 점유 또는 제한 량에 따라 양자화 조절을 결정할 때는 모델링 기술이 자주 사용되는데 예를 들면, [3], [16]에서 주어진 버퍼의 점유 량에 관해서 양자화 단계의 크기를 결정할 때는 버퍼의 점유와 양자화 관계의 분석 모델들이 사용된다. 본 주제에 관한 이전의 연구는 [5]에 수록되었다.

본 논문에서는 기존의 양자화 장치(quantizer)의 제어방법과 아울러 비트율(rate)을 제어할 수 있는 방법 즉, 모드 선택에 관해서 설명한다. 이 방법을 이용하면 간단한 양자화 장치의 제어방법에 의한 것보다 가시적인 품질의 변화가 적은 코딩 구성이 가능하다. 즉, 블록 단위의 비디오 코덱에서는 매크로블록들에 관한 최적 모드들을 선택할 수 있는 효

과적인 알고리즘을 제시한다. 이러한 모드들은 목표 비트율과 왜곡 관계에 의해서 순서대로 결정되는 최적 결정 곡선들에 의해서 선택된다. 각각의 프레임에 관한 고유 결정 곡선들은 모든 매크로블록들에 적용되어 사용된다. 이러한 결정 곡선들은 코딩 왜곡을 최소화시키는 목표 비트율과 출력 비트율을 반복해서 비교하여 결정된다. 채널 버퍼의 점유와 관계없는 이러한 반복 절차는 가열냉각 최적화기술(simulated annealing optimization technique)을 이용하여 최적 결정 곡선들이 최적임을 보증한다. 최적 결정 곡선들이 결정되면 주어진 비트율에 대한 총체적인 왜곡을 최소화시키는 결정 곡선들에 따라 매크로블록들의 최적 모드가 선택된다. 본 논문에서는 두 가지의 상이한 결정 곡선 즉, 운동/비 운동 보상 곡선 및 intra/inter 코딩 결정 곡선이 최적화된다. 본 논문에서 제시된 방법은 양자화 장치를 조작하여 목표 비트율을 만족시키는 것이 아니라 일관성을 가지는 가시적 품질을 도출하는 최적 매크로블록 모드들을 찾아서 목표 비트율을 만족시킨다는 점에서 기존의 방법들과는 다르다. 제시한 알고리즘은 H.261 비디오 코덱의 범위에서 수행된다.

본 논문은 다음과 같이 구성되어 있다. 문제 설정의 배경과 최적 모드의 선택 방법이 2절에서 설명된다. 3절에서는 최적 모드의 선택방법을 이용하여 원하는 비트율을 성취하는 방법이 설명된다. 이 방법은 기존의 방법보다 향상된 시각적 품질을 보장한다. 설명된 방법의 성능을 입증하기 위해 4절에서 모의 실험을 수행하고 마지막으로 5절에서는 결론과 차후의 연구방향을 제시한다.

II. 문제의 공식화

1. 배경

비디오 코딩 표준 모델 [1], [2]에서 각각의 매크로블록에 관한 모드를 결정하기 위해서 고정된 함수들이 사용된다. 이러한 함수들을 사용하여 인코더는 어떤 모드로 코드화 할 것인지를 각각의 매크로블록에 관해서 결정하여야 한다. 이러한 결정 방법에는 가장 좋은 운동 벡터들을 어떻게 결정하는지, 매크로블록을 intra 모드 또는 inter 모드로 결정할 것인지, 그리고 어떻게 양자화 장치의 규모를 설정할 것인지 등의 방법들이 있다. 이처럼 인코더는 다른 형태의 매크로블록들 중에서 선택을 해야 하는 어려운 문제를 갖고 있다. 철저한 검색방법(exhaustive search method)은 각각의 형태를 사용하는 갑

은 레벨의 정확성으로 매크로블록을 코딩하고 가장 적은 갯수의 코딩 비트들을 사용하는 모드를 선택하는 것이다. 명백하게 철저한 검색방법은 실제 코딩 구현에는 직결하지 않다.

컴퓨터를 사용하여 더 쉽게 계산할 수 있는 간단한 하위최적화 방법(suboptimal method)은 일련의 결정 방법을 작성하는 것이다. 이러한 결정 방법의 한 예로 다음과 같은 설정순서로 나타낼 수 있다.

(단계 1) 운동 보상(motion compensation) 또는 비운동 보상(no motion compensation); 운동 벡터가 전송되는지 아니면 '0'으로 처리가 되는지를 결정한다.

(단계 2) intra 코딩 또는 inter 코딩; 매크로블록의 모드가 intra인지 또는 단계 (1)에서 찾아진 운동 벡터를 사용한 inter 모드인지를 결정한다.

(단계 3) inter 코딩의 경우, 코딩을 할 것인지 안 할 것인지; 잔여 오류(residual error)가 DCT 변형을 사용하여 코드 화될 수 있을 정도로 큰 지의 여부를 결정한다.

(단계 4) 양자화 장치의 규모가 만족스러우지 또는 변경되어야 하는지 결정한다.

각각의 단계에서 함수 또는 규칙이 효과적인 의사결정을 위해서 사용된다. 예를 들면, [2]에서 운동/비운동 보상 및 intra/inter 코딩에 관한 결정 방법

은 일정한 함수들 (그림 1에 나타낸 것과 같은 함수들)을 사용하는 반면, 코드/비코드의 결정은 에러 신호들의 크기의 차이에 따라 결정된다. 양자화 파라미터들은 버퍼가 차 있는 양에 따라 결정된다. 일련의 결정 과정을 통해 각각의 매크로블록과 양자화 장치들에 관한 모드가 결정되고 매크로블록이 코드 화된다.

비트율 제어에 관해 발표된 대부분의 방법들은 단계 (4)에 집중된다. 예를 들면, 양자화 장치의 레벨은 버퍼의 점유 량에 따라 조절된다. 현재 제시된 알고리즘은 단계 (1)의 운동/비운동 보상뿐만 아니라 단계 (2)의 intra/inter 코딩에 대한 비트율 제어를 위한 옵션들로도 확장된다. 코드 화된 출력 비트율은 결정 곡선들의 모양에 대해서 매우 민감하다는 점에 유의하여야 한다.

일정한 운동/비운동 보상 곡선 및 intra/inter 코딩 곡선을 이용하는 것 보다 주어진 비트율에서 PSNR이 극대화될 수 있도록 입력 프레임들의 특성에 따라 적절한 결정 곡선들을 찾는 것이 더 바람직하다. 양자화 에러는 가시적인 품질에 직접적인 영향을 미치지 때문에 양자화 장치를 사용해서 비트율을 제어하는 것은 그다지 바람직한 것이 아니다. 이 논문에서 제시된 방법은 양자화 장치의 간단한 피드백 접근방법 보다 정교한 비트율의 제어를 제공한다. 철저한 검색 [4]를 사용하는 것보다 최적 결정 곡선들을 사용하면 매크로블록 모드를 간단하게 선택할 수 있다. 그리고 결정 곡선들은 다음에 설명되는 비율-왜곡(rate-distortion) 곡선들에 의하여 결정된다.

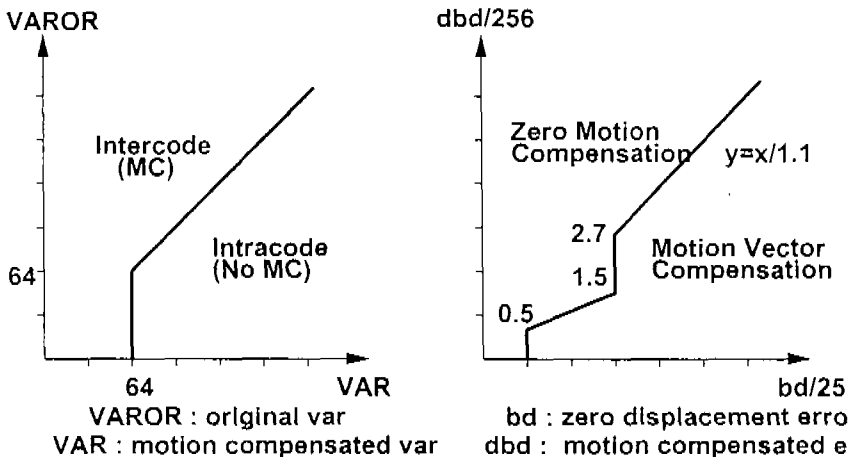


그림 1. MPEG 제공: P 프레임에서의 운동 보상/비운동 보상 그리고 intra/inter 코딩에 쓰이는 결정 곡선

2. 비율-왜곡(Rate-Distortion) 이론

설정된 비트율의 소스 왜곡을 최소화시키는 방법을 다루는 비율-왜곡의 고전적인 구조는 임의의 양자화 장치들에 관한 비트 할당 문제 [6]라고 간주할 수 있다. 비율 및 왜곡 측정방법의 부가성으로 인해서 목표 평균 비트율 R에 관한 평균 왜곡값 D를 최소화시키는 비트의 할당 문제에 한정되는 문제 (constrained problem)는 라그랑제 승수(Lagrange multiplier) λ 를 이용하여 비율과 왜곡을 합병함으로써 비한정 문제 (unconstrained problem)로 변형될 수 있다. 따라서, 비한정 문제는 다음과 같이 정의된 라그랑제 계열(Lagrangian)의 비용함수를 최소화하는 데 있다^[7].

$$J(\lambda) = D + \lambda R \tag{1}$$

원하는 최적 상수값 λ^* 는 우선값으로서 인식되지 않으며, 이는 원하는 목표 비트 예산값 (target bit budget) R 또는 품질 강제값 (quality constraint) 및 데이터 소스 특성에 따라 다르다. 라그랑제 계열의 비한정 비용 함수에 대해서 발표된 대다수의 해결 방법 [4], [12], [15]는 양자화 파라미터들에 관한 선택 방법을 제시한다. 본 논문에서는 비율 (R) 과 왜곡 (D) 사이의 관계가 양자화 장치에 대한 것보다 더욱 복잡하여도 식 (1)과 동일한 비용 함수 (cost function)를 사용한다. 그러나 이들의 사용은 양자화 장치들에 관한 것이 아니라 모드 선택 결정 곡선들에 관한 파라미터들을 결정하기 위해서 활용된다.

3. 모드의 선택

F 와 \mathcal{F} 를 각각 비디오 코덱의 입력 프레임 및 재구성된 프레임이라고 하자. j 번째 프레임 F_j 와 \mathcal{F}_j 는 매크로블록들로 구성된 그룹으로 분할될 수 있다.

$$F_j = (X_{0,j}, X_{1,j}, X_{2,j}, \dots, X_{L-1,j}) \tag{2}$$

$$\mathcal{F}_j = (\hat{X}_{0,j}, \hat{X}_{1,j}, \hat{X}_{2,j}, \dots, \hat{X}_{L-1,j}) \tag{3}$$

프레임 F_j 에서 각각의 매크로블록 $X_{i,j}$ 는 집합 S 에 의해서 주어진 가능한 N 개의 모드들 중 1개를 사용하여 코드화될 수 있다.

$$S = \{M_0, M_1, M_2, \dots, M_{N-1}\} \tag{4}$$

$M_k \in S, k=0, \dots, N-1$ 을 매크로블록 $X_{i,j}$ 를 코드화하기 위해서 선택된 모드라고 하자. 숫자 i 와 j 는 각각 매크로블록과 프레임의 인덱스를 의미한다. $y_{i,j}$ 를 선택된 모드 M_k 의 처리에 의해서 발생하는 코드화된 매크로블록의 데이터라고 하자. $\hat{X}_{i,j}$ 를 대응 블록의 디코딩한 후의 출력이라고 하자 (그림 2).

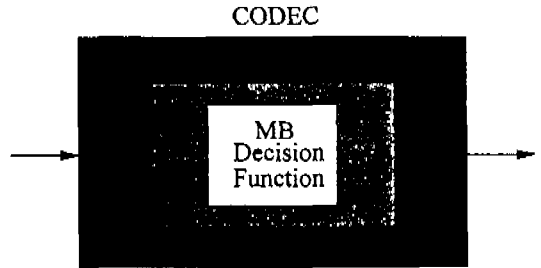


그림 2. 간단히 표현된 코덱의 구조

일반적으로 $y_{i,j}$ 는 할당되는 모드에 따라 다음과 같이 표현될 수 있다.

$$y_{i,j} = \begin{cases} 0, & \text{if skipped} \\ X_{i,j}, & \text{if intra coded} \\ X_{i,j} - X_{i,\Delta,j-1}, & \text{if inter coded} \end{cases} \tag{5}$$

$X_{i,\Delta,j-1}$ 은 운동 벡터 Δ 을 갖는 $(j-1)$ 번째 프레임의 i 번째 매크로블록을 의미하고 운동 벡터 Δ 는 다음과 같이 표현된다.

$$\Delta = \begin{cases} 0, & \text{if no MC} \\ \delta_{xy}, & \text{if MC} \end{cases} \tag{6}$$

δ_{xy} 는 운동 보상 알고리즘의 결과이다. 운동/비운동 보상에 관한 결정 곡선은 차수가 $P-1$ 인 다항식으로 표현된다고 가정하자.

$$g(x) = \sum_{k=0}^{P-1} a_k x^k \tag{7}$$

식 (7)은 매크로블록이 운동 보상을 사용하여 코딩이 될 것인지 또는 비운동 보상을 사용하여 코딩이 될 것인지를 첫 번째 4 분면을 2개의 부분으로 나누는 곡선을 통하여 결정한다. 그림 3은 운동 보상 및 비운동 보상에 관해서 영역을 A와 B로 분할하는 결정 곡선의 예를 나타낸 것이다.

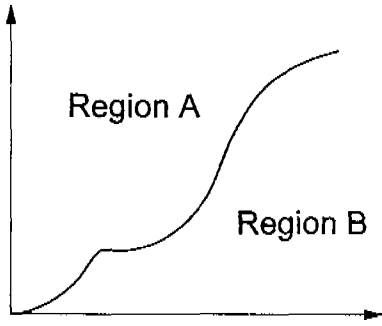


그림 3. 운동 보상/비운동 보상구간을 나누는 결정 곡선의 예

그림 3에서 x 축은 주어진 i 와 j 에 관한 $X_{i,j}$ 및 $X_{i,j-1}$ 블록들에서 픽셀 값들 사이의 절대적인 차의 합으로 결정된다.

$$\begin{aligned} x(X) &= |X_{i,j} - X_{i,j-1}| \quad \forall i, j \\ &= \frac{1}{251} \sum_m \sum_n |X_{i,j}(m, n) \\ &\quad - X_{i,j-1}(m, n)| \end{aligned} \quad (8)$$

m 과 n 은 매크로블록내의 인덱스들이다. 변수 y 는 $X_{i,j}$ 와 $X_{i,j-1}$ 의 픽셀 값들 사이의 절대 차의 합으로 정의된다.

$$\begin{aligned} x(X) &= |X_{i,j} - X_{i,j-1}| \quad \forall i, j \\ &= \frac{1}{251} \sum_m \sum_n |X_{i,j}(m, n) \\ &\quad - X_{i,j-1}(m, n)| \end{aligned} \quad (9)$$

운동 벡터 Δ 는 다음과 같이 곡선 $g(x)$ 를 통해서 결정된다.

$$\Delta = \begin{cases} 0, & \text{if } y \geq g(x) \\ \delta_{xy}, & \text{if } y < g(x) \end{cases} \quad (10)$$

프레임 $X_{i,j}$ 와 $X_{i,j-1}$ 그리고 프레임 $X_{i,j}$ 와 $X_{i,j-1}$ 사이의 절대 차의 합을 사용하면 운동 보상과 비운동 보상에 관한 모드가 결정된다. 즉, 주어진 절대 값으로부터 좌표 (x, y) 가 결정되고 그 좌표가 $g(x)$ 와 비교하여 어디에 위치하는가에 따라서 모드가 결정된다.

그렇다면 해결해야 할 문제는 주어진 비트율 R 에 관해서 j 번째 프레임의 범용 코딩 왜곡값(global

coding distortion) $D = D_{F_j}(a_k)$ 를 최소화시키는 다항식 $g(x)$ 의 계수 a_k 를 찾는 것이다.

$$\arg \{ \min_{a_k} D_{F_j}(a_k) \} \quad (11)$$

$$\text{with } R_{F_j}(a_k) \leq R \quad (12)$$

앞에서 논한 것처럼 식 (11)의 한정된 최적화 문제를 해결하는 대신에 다음과 같은 비한정 최적화 문제에 관한 해를 구할 수 있다.

$$\arg \{ \min_{a_k} J_{F_j}(a_k) \} = D_{F_j}(a_k) + \lambda \cdot R_{F_j}(a_k) \quad (13)$$

그러나, 이 경우에는 양자화 장치의 파라미터 λ 가 아닌 파라미터 a_k 를 찾으려고 한다는 것을 유의하여야 한다. 식 (13)을 풀기 위해서 λ 를 찾아야 하지만 λ 가 한정조건(the constraint)과 불충족한 R-D 곡선의 특성을 만족시키는 선택된 최적 점에서 R-D 곡선의 기울기라는 사실을 사용하면 문제를 간단하게 만들 수 있다. R-D 곡선을 이용할 수 있으면 x_1 은 원하는 비트율이고 y_1 은 미지의 최소 왜곡점인 R-D 곡선 상의 (x_1, y_1) 점에서 기울기를 찾으면 λ 를 구할 수 있다. 그러므로, 비한정 문제는 일정한 λ 를 사용해서 다음과 같이 설명된다. 따라서, 주어진 $\lambda > 0$ 에 관해서 라그랑제 계열의 비용 함수 식 (14)를 최소화시키는 a_k^* 에 관해서 해를 구할 필요가 있다.

$$J_{F_j}(a_k^*) = \min_{a_k} \{ D_{F_j}(a_k) + \lambda R_{F_j}(a_k) \} \quad (14)$$

식 (11)에 대응하는 한정문제는 다음과 같다.

$$\begin{aligned} D_{F_j}(a_k^*) &= \min_{a_k} D_{F_j}(a_k) \\ \text{such that } R_{F_j}(a_k^*) &\leq R \end{aligned} \quad (15)$$

식 (15)의 한정 문제는 식 (11)의 비한정 문제와 동일한 해를 갖게 되는 것이 분명하고 이는 λ 를 선택하여 한정된 최적 값을 만족시키기 때문에 $R = R_{F_j}(a_k^*)$ 일 때 가능하다. 사실, 이것은 [7]에서 주어진 이론을 일반화한 것이다.

최적 해는 R-D 특성에 따라 $\lambda > 0$ 일 때 존재하지 않을 수 있다는 것을 유의해야 한다. 옴의 값이 아닌 모든 λ 에 대하여, 한정된 문제의 해가 상응하는 비한정 문제의 해와 동일하다고 할 수 있다. 그러나, 중요한 것은 $R_{F_j}(a_k^*)$ 가 주어진 λ 값에 관해서

R과 같은 경우가 발생하게 되면 a_k^* 는 한정 문제에 대한 원하는 해법이 된다. 다시 말해서, R이 주어지면 λ 를 구할 수 있고 이에 따라서 비한정 문제에 관해서 해를 구할 수 있으며 $R_F \leq R$ 을 만족시키는 원하는 해법 a_k^* 를 얻을 수 있다. 단, 기울기가 λ 인 R-D 곡선상의 점을 구하면 R을 구할 수 있다. 그러나, 실제로 있어서 여러 개의 결정 곡선 파라미터들 (양자화 장치는 상수를 취한 경우)에 관해 R-D 곡선을 구하는 것은 매우 어렵다 (많은 시간이 소요된다). 2개의 결정 곡선에 관한 계수들(예를 들면, 운동/비운동 보상 곡선 및 intra/non-intra 코딩 결정 곡선)의 개수가 N과 M인 곳의 R-D 곡선의 각각의 데이터 점에 관해서 $N \times M$ 차원의 검색이 필요하다. 계산을 극대화시키기 위해서 0 (최저 왜곡, 최상 비율)에서 무한대 (최상 왜곡, 최저 비율) 사이의 λ 의 값들을 검색하고 최적 계수 a_k^* 를 구한다. 이 후에는 원하는 비트율에 충분히 근접하는지 확인하기 위해서 비트율 $R_F(a_k^*)$ 를 계산한다. 그 결과 얻어진 비트율이 충분한 근접 값이면 계산할 때 사용된 파라미터 a_k 들에 따라 최적 결정 곡선들이 정의된다.

4. λ 를 구할 수 있는 빠른 해법

식 (14)를 구하려면 a_k 뿐만 아니라 λ 를 검색하여 한정된 최소 비용함수 값을 얻어야 한다. 프레임들 사이에서 장면의 변화가 일관성을 유지하면 프레임들 사이에서 λ 의 값은 변화가 심하지 않다고 생각할 수 있다. 이것은 각각의 프레임에서 코딩된 비트율의 변화가 프레임들 사이에서 작다는 것을 암시한다. 그러므로, 각각의 프레임에 관해서 λ 를 구하는 대신, 이어서 발생하는 프레임들에 관해서 이전의 λ 를 사용하면 컴퓨터로 처리하는 데 있어서의 복잡성을 감소시킬 수 있다. λ 는 코딩된 프레임들의 품질 지수로서 해석될 수 있기 때문에 첫 번째 프레임의 코딩을 시작 시점으로 화면 변화가 발생 (예를 들면, 화면의 컷트(cut) 등의 경우로 인해서)될 때까지 최적 λ 를 구하는 것이 중요하다. 매크로블록들에 관해서 j번째 프레임의 왜곡은 다음과 같이 확장될 수 있다.

$$TD_{F_j}(a_k) = E \{ | F_j - \hat{F}_j |^2 \}$$

$$= \frac{1}{L} \sum_{i=0}^{L-1} | X_{i,j} - (X_{i,j-1} + \hat{X}_{i,j}) |^2 \quad (16)$$

따라서, Lagrangian 비용함수는 다음과 같이 변경된다.

$$J_{F_j}(a_k^*) = \min_{a_k} \{ D_{F_j}(a_k) + \lambda R_{F_j}(a_k) \}$$

$$= \min_{a_k} \left\{ \frac{1}{L} \sum_{i=0}^{L-1} | X_{i,j} - (X_{i,j-1} + \hat{X}_{i,j}) |^2 + \lambda R_{F_j}(a_k) \right\} \quad (17)$$

단,

$$\hat{X}_{i,j} = D^{-1} \cdot Q^{-1} \cdot Q \cdot DX_{i,j} - X_{i,j-1}$$

$$F_j = \text{Frame } j \quad (= \sum_i X_{i,j})$$

$$\hat{F}_j = \text{Reconstructed frame } j \quad (= \sum_i \hat{X}_{i,j})$$

D와 D^{-1} 은 DCT와 IDCT를 의미하며 Q와 Q^{-1} 은 각각 양자화 및 역 양자화를 의미한다. intra 및 inter 프레임 코딩의 경우, 식 (8)과 식 (9)는 다음과 같은 수식으로 대체된다.

$$x(X) = \text{var}(X_{i,j} - X_{i,j-1}) \quad \forall i, j$$

$$= \frac{1}{251} \sum_m \sum_n (X_{i,j}(m, n) - X_{i,j-1}(m, n))^2 \quad (18)$$

이것은 우리 문제에 대해서 비한정 최적 공식이며, 확률론 적인 가열냉각 최적 알고리즘은 식 (17)에서 왜곡을 최소화시키는 식 (7)에서의 계수 a_k 를 구할 때 채택된다.

5. 고속 가열냉각 모의실험(Very Fast Simulated Reannealing)

가열냉각의 모의실험은 무한 반복 후에만 전체적인 최소 값에 접근할 수 있지만 알고리즘의 漸近線 역함(asymptotic behavior)로 종종 적절한 시간이 경과하면 근사값으로 접근될 수 있다. 복잡성의 부담을 줄이기 위해서 VFSA (Very Fast Simulated Annealing) 알고리즘 [17]을 적용하기로 한다. Cauchy 분포도를 사용하면 더 빠른 가열냉각 스케줄을 얻을 수 있다. 또한, Cauchy 분포의 특성상 두터운 꼬리 (fatter tail) 때문에 검색 과정 동안에 현재의 영역별 모방 상태로부터 더 많은 상태를 실험할 수 있다. T_j 에 관해서 관련된 가열냉각 스케줄은 다음과 같다.

$$T_i(k) = T_i(0) \exp(-c - ik^{1/b}) \quad (20)$$

$T_i(0)$ 는 초기 비용온도와 최고 및 최근의 비용 함수의 절대값 차와 그들의 차이의 최대 값의 최소 값으로 적용된다. T_i 는 비용함수의 가장 최근에 저장된 모방 값들과 이전 값과의 차액의 절대 값과 현재 값의 최대 값으로서 계산되며 초기의 비용 온도를 초과하지 않도록 억제된다. 이 과정은 현재 가장 최근에 저장된 최소 값의 크기 내에서 비용 온도의 가열냉각의 크기를 필연적으로 리셋시키며 통계적인 전체 최소 값이 얻어질 수 있도록 만든다.

III. 적응 모드 선택을 통한 비트율 제어

H.261 코딩 방법에 대해서 응용할 때 현재 프레임의 비트 카운트 값을 만족시키기 위해 적절한 모드를 생성하기 위해서는 결정 픽셀들을 발생시키도록 비한정 함수의 라그랑제 승수가 제어되어야 한다. 버퍼의 내용은 프레임 비율, 채널 비율 및 버퍼의 크기와 같은 입력 코딩 파라미터들에 의하여 이전에 코드화된 프레임들의 비트 카운트를 사용하여 추측된다. 이 절에서는 적응 모드 선택방법을 사용하는 비트율 제어 방법에 관련된 내용을 자세히 살펴본다.

비트율을 원하는 값으로 제어하려면 버퍼를 제어하는 것이 필수적이다. 일반적으로 버퍼의 오버플로우(overflow) 또는 언더플로우(underflow)를 피할 수 있도록 양자화 장치들을 선택하는 코더에 버퍼의 상태가 다시 전달된다^[12]. 그러나, 여러 모드의 프레임 (I, P 및 B)으로부터 발생된 코드화된 비트들의 개수는 서로 넓게 변하기 때문에 버퍼에 기록되는 것과 양자화 장치들 사이의 매핑 방법을 이용하는 전통적인 비트율 제어방법으로 부터는 좋은 결과를 얻을 수 없다. 이 절에서는 Adaptive Mode Selection (AMS) 알고리즘에 따른 간단한 버퍼 제어의 방법을 설명한다.

1. 오버플로우 및 언더플로우 경고 상태의 설정 방법

R_f 와 R_c 는 프레임 속도 (frame rate, fps) 및 채널 속도(channel rate, bps)를 의미하고, B 가 채널 속도의 k msec (kR_c)인 버퍼의 크기를 의미한다고 하자. 그리고 버퍼에 대한 입력 속도가 다음과 같이 표현된다고 하자.

$$\begin{aligned} R_m &= R_1 + R_2 + R_3 + \dots + R_n \\ &= R_{f_1 r_1} + R_{f_2 r_2} + R_{f_3 r_3} \dots + R_{f_n r_n} \end{aligned}$$

코더로 입력되는 프레임의 개수가 n 이고 $R_i (i=1, \dots, n)$ 의 단위를 bps라고 하면 위의 표기법을 사용하여 다음과 같이 버퍼 내용을 공식으로 만들 수 있다.

$$B_L(t=n^-) = \sum_{i=0}^{n-1} R_{f_i r_i} - nR_c \quad (21)$$

$$B_L(t=n^+) = \sum_{i=0}^{n-1} R_{f_i r_i} - (n+1)R_c \quad (22)$$

식 (21)은 n 번째 프레임이 버퍼에 임시로 기록되고 코드화 된 직후에 버퍼의 내용을 의미한다. 반면 식 (22)는 $(n+1)$ 번째 프레임이 버퍼에 임시 기록되기 바로 직전을 의미한다. 버퍼가 오버플로우 또는 언더플로우 되는 것을 피하기 위해서 버퍼 레벨 (B_L)은 다음과 같이 제어된다.

$$aB \leq B_L < \beta B \quad (23)$$

단, $a + \beta = 1$, $a \geq 0$, $\beta \geq 0$, $a \geq \beta$ 식 (21)과 식 (22)가 식 (23)을 만족하면 버퍼가 오버플로우 또는 언더플로우 되는 것을 피할 수 있다. 그러므로,

$$aB \leq \sum_{i=0}^{n-1} R_{f_i r_i} - (n+1)R_c \quad (24)$$

$$\sum_{i=0}^{n-1} R_{f_i r_i} - nR_c < \beta B \quad (25)$$

B 는 채널 속도로 표현되기 때문에 위의 식들은 다음과 같이 확장될 수 있다.

$$aKr_c \leq \sum_{i=0}^{n-1} \left(\frac{R_i}{R_c}\right) r_i - (n+1) \quad (26)$$

$$\sum_{i=0}^{n-1} \left(\frac{R_i}{R_c}\right) r_i - n < \beta k \quad (27)$$

a 와 β 를 변경하면 버퍼가 언더플로우 또는 오버플로우될 때 경고(alert)를 해줄 범위들을 미리 설정할 수 있고 그 범위의 샘플로 상단 및 하단 경계를 그림 4에 나타내었다.

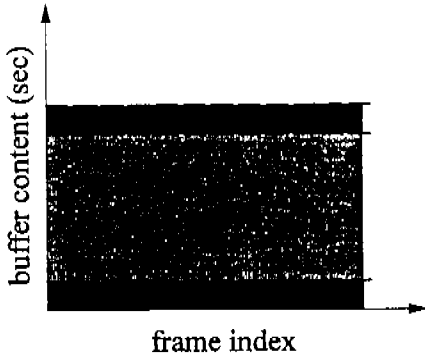


그림 4. 버퍼의 오버플로우 또는 언더플로우를 경보해줄 미리 설정된 경계

2. 품질 계수 λ에 따른 비트율 제어

이와 같이 제시된 접근방법에서의 비트율은 일반적인 양자화 장치 제어와 차이가 있게 만드는 품질 계수 λ에 의해서 제어되며, 라그랑제 승수는 버퍼의 상태에 따라 조절된다. 버퍼 내용이 언더플로우 정보 상태에 있을 때 (예를 들면, $0 < B_L < aB$ 식 (17)의 라그랑제 승수 λ는 비트율의 억제 보다 왜곡의 억제를 더 제공하도록 감소된다. 버퍼의 내용이 상단 경계 aB 에 도달할 때 λ는 매크로블록들의 모드가 스킵(skipping) 모드를 사용할 수 있도록 증가된다. 또한 λ를 증가시키면 알고리즘은 왜곡의 억제보다 비트율의 억제에 더 충실하게 된다.

각각의 프레임에서 오버플로우 또는 언더플로우가 발생할 때 최적의 λ를 구하는 것은 시간을 소비하는 것이다. 프레임 단위에 관해서 파라미터 λ의 값을 사용하면 과정을 빨리 진행시킬 수 있다. 버퍼의 오버플로우를 피하기 위해서 시뮬레이션에서 사용된 방법은 다음과 같다.

$$\lambda_{(j+1)} = \lambda_j + \lambda_j \left(\frac{R_{Fj}}{R - R_{Fj}} \right) - \lambda_j \left(\frac{R}{R - R_{Fj}} \right) \quad (28)$$

λ의 증가분은 이전 프레임의 코드화된 비트율에 의존된다. 식 (28)이 의미하는 것은 원하는 비트율과 j번째 프레임의 코드화된 비트 속도 사이의 차이가 클 때까지 λ에서 작은 증가분을 만들고 차이가 작을 때 큰 증가분을 만든다는 것이다. 마찬가지로 버퍼의 내용이 언더플로우 상태에 도달할 때 다음과 같이 그 값을 감소시키면 λ 값을 변경할 수 있다.

$$\lambda_{(j+1)} = \lambda_j - \lambda_j \left(\frac{R_{Fj}}{R - R_{Fj}} \right) - \lambda_j \left(\frac{R - 2R_{Fj}}{R - R_{Fj}} \right) \quad (29)$$

IV. 실험 결과 및 분석

제시된 접근방법을 실험하기 위해 식 (7)에서의 차수 $P=2$ 에서 9까지 여러 등급에 관한 CIF 포맷 (352 픽셀 × 24 라인)으로 여러 가지 sequence를 사용하여 모의실험을 수행하였다. 이 실험에 프레임의 속도는 초당 30장의 속도로 일정하게 유지되고 프레임 속도는 평균 비트 속도 128 Kbps에서 352 Kbps의 범위에서 변한다. 인코딩 과정의 일부분으로서 모드들은 위에서 설명된 과정 즉, 먼저 최적인 운동 보상/비운동 보상 코딩 픽션을 구하고 최적의

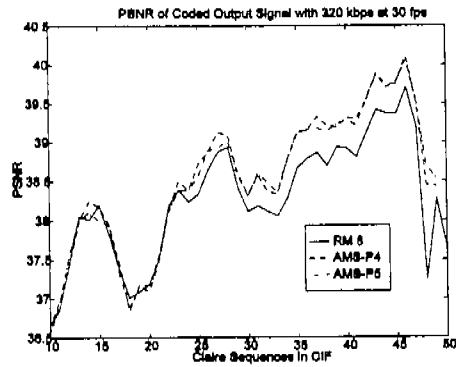
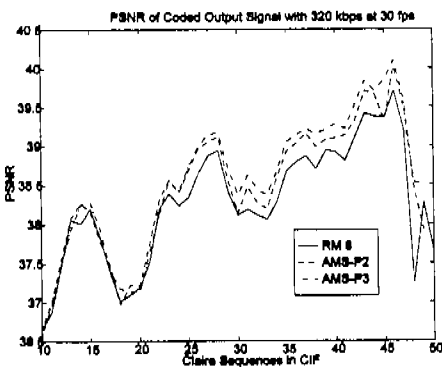


그림 5. 운동 보상/비운동 보상의 결정 픽션 차수(2에서 5)와 PSNR의 관계

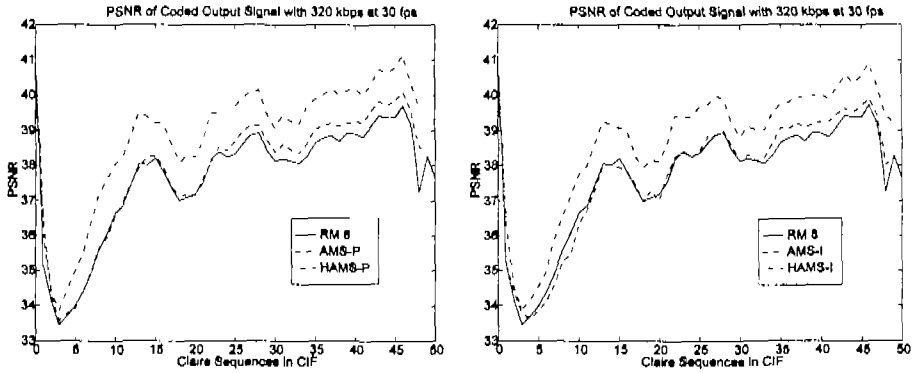


그림 6. Claire에서 HAMS와 AMS로 코딩시 PSNR과 프레임 코딩 속도의 비교

intra/inter 코딩 픽션을 구하는 방법을 사용하여 선택된다. 코딩 결과는 비디오 코드 실험 모델인 RM 8에 의해서 코드화된 프레임들과 비교되었다.

실험하는 동안 운동/비운동 보상 결정 픽션에 관해서 $P=3$ 일 때 다항식의 차수가 최적상태라는 것을 알았다. 3 이하의 차수를 적용하면 좋은 보조 최적 계수들이 구해지지 않았고 3 이상의 차수가 사용되면 계산이 커질 때 PSNR이 매우 작은 양으로 증가하였다. Intra/inter 코딩 결정 픽션에 관한 다항식의 최적 차수는 $P=2$ 이라는 것이 실험에 의하여 구해졌다. 두 가지 경우에서 다항식들의 최적 차수들을 그림 5에 나타내었다. 계층적인 검색 운동 값 (Hierarchical Search Motion Estimation) [14]은 운동 벡터들을 구하는 복잡성을 줄이기 위해서 그리고 코딩 성능을 증가시키기 위해서 채택되었다. 계층적인 검색 운동 값을 사용하여 매크로블록의 모드들은 운동 벡터들로부터 영향을 받는다. 계층적인 검색 방법에 의한 적응 모드 선택 (HAMS, Hierarchical Search Based Adaptive Mode Selection)을 사용하여 구해진 모드들은 적응 모드 선택

에 따라 선택된 모드들과 다르다. 이것은 서로 다른 운동 벡터들이 이전 프레임으로부터의 운동 때문에 대응하는 매크로블록과 현재의 매크로블록 사이에서 다른 절대적인 차 또는 편차가 다르기 때문이다. AMS 접근방법 보다 HAMS 접근방법이 더 잘 수행된다는 것을 그림 6에 나타내었다. HAMS를 사용하여 최적 결정 픽션을 구하는 두 가지 모델은 예를 들면, 운동/비운동 (HAMS-P) 및 inter/ intra (HAMS-I)에 관한 최적 결정 픽션을 구하는 것은 RM 8과 비교된다. 제시된 모든 모델은 그림 7에 나타낸 것과 같은 프레임들 사이 및 그 범위에서 일관된 가시적 품질을 제공한다.

다른 비트율과 여러 가지 순서들 사용하는 많은 모의실험 결과들 그림 8에서 나타내었다. 일반적인 (일정한) 코딩에 관해서 인지할 수 있는 품질의 차이는 프레임의 반복 때문에 비트 속도가 특히 낮을 때 버퍼 오버플로우가 발생하면 매우 좋지 않다는 것을 유의하여야 한다. RM 8을 사용해서 실험한 결과 버퍼는 프레임 13 뒤에 상당히 채워지기 시작하였고 그로 인하여 그 다음 프레임부터는 이전의 프레임율 반복

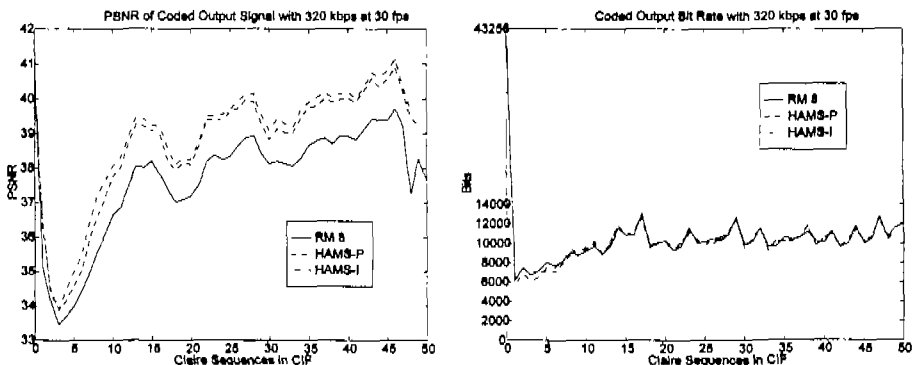


그림 7. Claire에서 HAMS로 코딩시 PSNR과 프레임 코딩 속도의 관계

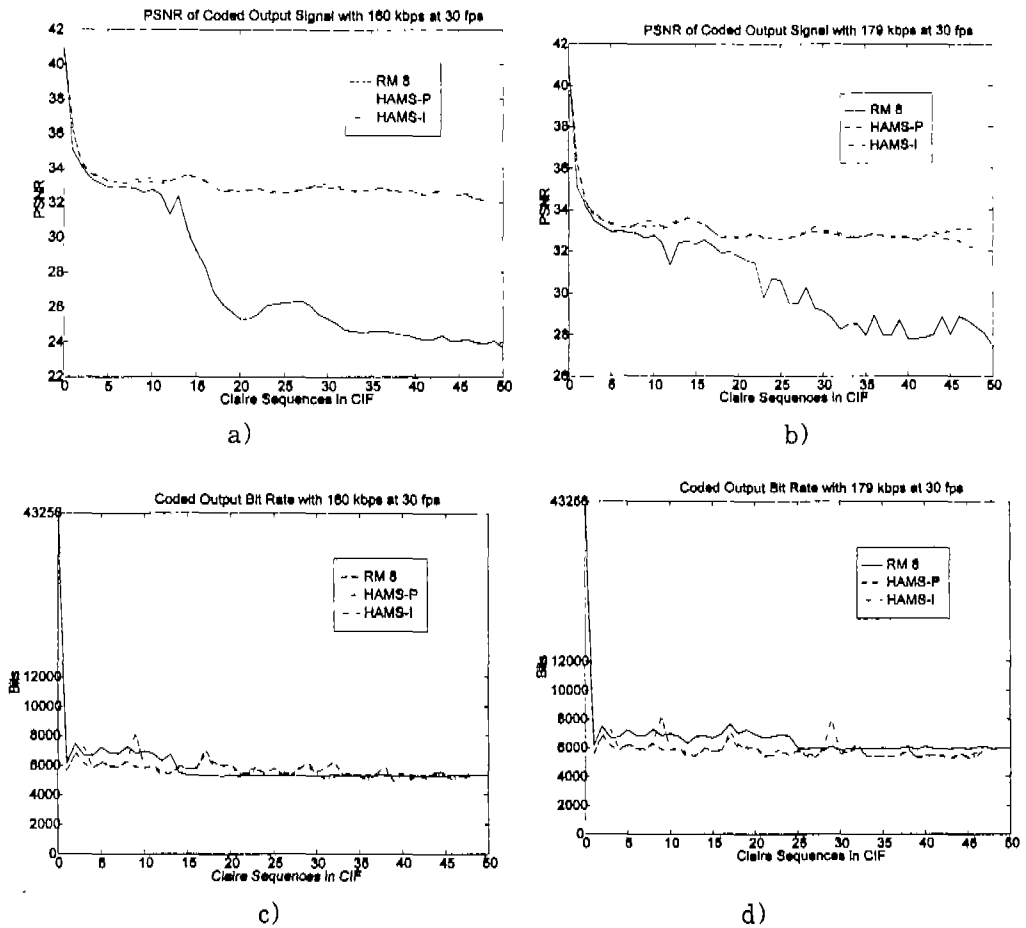


그림 8. Claire에서 HAMS를 이용하여 160kbps와 179kbps로 각각 코딩시 PSNR과 코딩된 프레임 속도와의 관계

하도록 강요되었다 8-a). 그러나, 제시된 방법을 사용하면 2개의 모델 모두에서 최적 결정 곡선에 따라 매크로블록 모드들이 선택될 때 버퍼는 오버플로우 상태에 도달하지 않는다. 또한, 버퍼가 오버플로우가 되기 전 프레임 13까지의 제시된 방법의 가시적인 품질도 고정된 모델보다 더 좋다. 버퍼의 언더플로우의 경우에는 제시된 접근방법 및 RM 8의 가시적인 품질은 가장 작은 양자화 파라미터들을 사용하기 때문에 거의 같았다. 제시된 접근방법이 사용될 때 가시적 품질은 낮은 비트 속도에서 10 dB만큼 향상된다. 또한, 제시된 접근방법은 프레임들을 통해서 일정한 가시적 품질을 제공한다.

V. 결 론

이 논문에서 우리는 블록 단위의 비디오 코딩 프

레이م 작업에 관하여 최적 운영 모드를 적절하게 선택하여 비디오 코딩시의 비트율 제어 알고리즘에 관한 새로운 방법을 제시하였다. 알고리즘은 주어진 비트율에서 최적 비트 할당을 위해서 라그랑제 방법을 기초로 한다. 우리의 접근방법은 H.261 비디오 codec 표준에 따라 수행되었다. 모의실험 결과 표준 양자화 장치 피드백(standard quantizer feedback)에 따른 비트율 제어 접근방법과 비교할 때 더 높은 PSNR과 더 좋은 가시적 품질을 나타내고 있다. 이 방법의 결점은 가열냉각의 모의실험 때문에 좀 더 복잡하다는 것이다. 그러나, 복잡성은 CD-ROM에 저장되어 있는 비디오와 같은 실시간으로 처리되지 않는 중요한 문제가 아니라면 특정 응용 프로그램에서 유용할 수 있다. 이 논문에서는 기본적인 알고리즘 중 세 가지 변형 방법이 설명되었다. 앞으로의 연구를 통해 동일한 기본적인 개념에 따

라 알고리즘을 개선할 수 있고 해당 복잡성을 줄일 수 있다. 본 논문에서 제시된 연구 결과는 알고리즘이 효율적이라는 것과 여러 가지 응용프로그램에서 매우 유용함을 보여주고 있다.

참고 문헌

- [1] ISO/IEC 11172-2, "Information technology- Generic coding of moving picture and associated audio for digital storage media at up to about 1.5 Mb/s -Part 2: Video", International Standard, Nov. 1993.
- [2] CCITT SC XV, "Description of Reference Model 8 (RM 8)", Document 525, June 1989.
- [3] ISO-IEC/JTC1/SC29/WG11, "Test Model 5", Draft, April 1993.
- [4] T. Wiegand, M. Lightstone, D. Mukherjee, T.G. Campbell and S. Mitra, "Rate-Distortion Optimized Mode Selection for Very Low Bit Rate Video Coding and the Emerging H.263 Standard", *IEEE Trans. on Circuit Syst. Video Technol.*, vol. 6, no. 2, pp. 182-190, April 1996.
- [5] C. Ryu and S. P. Kim, "Rate Control in Video Coding by Adaptive Mode Selection", *IEEE Int. Symposium on Circuit and Systems*, Monterey, CA, May 1998.
- [6] Claude E. Shannon, "Coding Theorems for Discrete Source with a Fidelity Criterion", *IRE Nat. Conv. Rec.* pt. 4, pp. 142-163, March 1959
- [7] Y. Shoham and A. Gersho, "Efficient Bit Allocation for an Arbitrary Set of Quantizers", *IEEE Trans. on Acoust. Speech, Signal Processing*, vol. 36, pp. 1445-1453, Sep. 1988.
- [8] C. T. Chen and A. Wong, "A Self-Governing Rate Buffer Control Strategy for Pseudoconstant Bit Rate Video Coding", *IEEE Trans. on Image Processing*, vol. 2, No. 1, pp. 50-59, Jan. 1993. *전에 관한 연구*, 한국출판사, pp.11-21, 1999
- [9] W. Ding and B. Liu, "Rate Control of MPEG Video Coding and Recording by Rate-Quantization Modeling", *IEEE Trans. on Circuits Syst. Video Technol.*, vol. 6, no. 1, pp. 12-20, Dec. 1991.
- [10] C. A. Gonzales and e. Viscito, "Motion Video Adaptive Quantization in the Transform Domain", *IEEE Trans. on Circuit Syst. Video Technol.*, vol. 1, no. 4, pp. 374-378, Dec. 1991.
- [11] A. Puri and R Aravind, "Motion Compensated Video Coding with Adaptive Perceptual Quantization", *IEEE Trans. on Circuits Syst. Video Technol.*, vol. 1, no. 4, pp. 351-361, Dec. 1991.
- [12] K. Ramchandran and M. Vetterli, "Best Wavelet Packet Bases in a Rate-Distortion Sense", *IEEE Trans. on Image Processing*, vol. 2, no. 2, pp. 160-175, April 1993.
- [13] K. Ramchandran, A. Ortega, and M. Viteerli, "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.
- [14] C. Ryu and S. P. Kim, "Deinterlacing Using Motion Compensated Local Spectra", *29th Asilomar Conference*, Pacific Grove, CA, Oct. 1995.
- [15] H. Sun, W Kwork, M. Chien, and C. H. Ju, "MPEG Coding Performance Improvement by Jointly Optimizing Coding Mode Decisions and Rate Control", *IEEE Trans. on Circuit Syst. Video Technol.*, vol. 7, no. 3, pp. 449-458, June 1997.
- [16] E. Viscito and C. Gonzales, "A Video Compression Algorithm with Adaptive Bit Allocation and Quantization", *SPIE Visual Comm. Image Processing*, vol. 1605-1, pp. 58-72, Nov. 1991.
- [17] L. Ingber and B. Rosen, "Generic Algorithm and Very Fast Simulated Reannealing: A Comparison", *Mathematical and Computer Modeling*, vol. 16, no. 11, pp. 87-100, 1992.

류 철(Chul Ryu)

1989년 3월 : 미국 Florida Institute of Technology
컴퓨터 공학과졸업

1991년 3월 : 미국 Polytechnic University 전기/전자
공학과 석사

1997년 10월 : 미국 Polytechnic University 전기/전
자공학과 박사

1998년 4월~1999년 3월 : LG 정보통신 연구소 선
임연구원

1999년 4월~현재 : 동국대학교 정보통신공학과
<주관심 분야> Multimedia and visual data com-
munications, compression, encryption,
watermarking

김 승 필(Seung P. Kim)

정회원

현재 : InterDigital Com. 연구원